REC2012

Proceedings of the 5th International Conference on

Reliable Engineering Computing

Practical Applications and Practical Challenges

June 13–15, 2012
Brno, Czech Republic

Editors:

Miroslav Vořechovský
Václav Sadílek
Stanislav Seitl
Václav Veselý
Rafi L. Muhanna
Robert L. Mullen

The organization of this conference is supported by the Czech Scientific Foundation, under project No. P105/10/J028 – Development and implementation of a simulation framework for quasi-ductile composites combining short fibers and aligned continuous reinforcement.

Published by

Ing. Vladislav Pokorný – LITERA
Tábor 43a, 612 00 Brno, Czech Republic
e-mail: vopelkova@literabrno.cz

5th INTERNATIONAL CONFERENCE ON RELIABLE ENGINEERING COMPUTING (REC 2012)

Editors: Miroslav Vořechovský, Václav Sadílek, Stanislav Seitl, Václav Veselý, Rafi L. Muhanna and Robert L. Mullen

Liability
The Conference Organizer, which term includes the Organizing Committee or the Publisher, is not liable for personal accidents or loss/damage of private property during the Conference. Delegates are encouraged to make their own arrangements for personal insurance.

Disclaimer
Whilst every attempt will be made to ensure that all aspects of the Conference mentioned will be as scheduled, the Conference Organizer reserves the right to make any changes should the need arise. The Conference Organizer is not responsible for any personal expenses incurred or any loss suffered by any Conference Delegates or his/her guests during or in connection with the event. All statements, opinions (whether of a technical nature or not), evidence, conclusions, exhibits, demonstrations, expressions, issues debated, which are said, made or expressed either orally or in writing and in any form or manner whatsoever including papers, paper prints, discussion material, articles, materials, notes, presentational slides, discussions or speeches are solely the views and opinions of the individual and not that of the Conference Organizer or the Publisher. The Conference Organizer or the Publisher will not be responsible whatsoever for the accuracy, correctness or representations of any statements or documents presented or discussed in this Conference.

ISBN 978-80-214-4507-9

Preface

Reliable Engineering Computing has emerged as a multi-disciplinary quality brand with this conference series, which has been hosted at Georgia Tech Savannah in 2004, 2006, 2008 and at the National University of Singapore in 2010. The REC-meetings provide a unique symbiosis of various engineering and associated disciplines with the kernel areas of Civil/Mechanical Engineering, Computer Science, and Mathematics. Central issue of the discussions is the reliability of engineering computations. Cross-disciplinary advisements generate synergy and impulses of a new quality for research and development, as well as for innovative applications.

REC2012 is focusing on providing solutions for Practical Applications and dealing with Practical Challenges in incorporating Reliable Computing in engineering practice. While there is an underlying theoretical framework for Reliable Computing, translation from theory to practice in engineering is still needed. As, in the previous conferences, papers in this proceeding address different fields of engineering, sciences, and mathematics within the context of risk and uncertainty. While providing solutions for real life problems represents the ultimate goal of the engineering profession, the papers endeavor to retain the rigor of mathematical formulations and their computational implementations to ensure safety, reliability, and more accurate predictions. The keystone in this aspect is the development of appropriate models of the practical applications and the use of realistic information in parameters' evaluation.

REC2012 continues the tradition of the conference series with a unique multi-disciplinary character to achieve advancements in the field of Reliability Engineering Computing.

This conference is supported by:

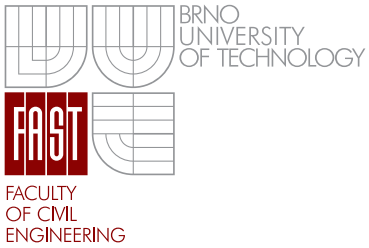
- Cervenka Consulting Ltd.
- Dynardo GmbH
- The South Moravian Region

The organizers appreciate the support of these organizations.

Rafi L. Muhanna, Robert L. Mullen and Miroslav Vořechovský
Chairmen of REC 2012

Organization Committee

Organized by



In Collaboration with

Honorary Chair
Gerhart I. Schuëller, *University of Innsbruck, Austria*

Founder and General Chair
Rafi L. Muhanna, *Georgia Institute of Technology, USA*
Robert L. Mullen, *University of South Carolina, USA*

Chair
Miroslav Vořechovský, *Brno University of Technology (BUT), Czech Republic*

Local Organizing Committee
Václav Sadílek, *Institute of Structural Mechanics, BUT*
Stanislav Seitl, *Institute of Structural Mechanics, BUT*
Václav Veselý, *Institute of Structural Mechanics, BUT*

Technical Support Team
Dana Halíková, *Institute of Structural Mechanics, BUT*
Martina Kolaciová, *Institute of Structural Mechanics, BUT*
Jana Kaděrová, *Institute of Structural Mechanics, BUT*

REC 2012 Secretariat
Dita Vořechovská, *Institute of Structural Mechanics, BUT*

International Scientific Committee

Ivan Au S.K.,
City University of Hong Kong, Hong Kong

Michael Beer,
University of Liverpool, United Kingdom

Daniel Berleant,
University of Arkansas at Little Rock, USA

Christian Bucher,
Vienna University of Technology, Austria

Jianye Ching,
National Taiwan University, Taiwan

George Deodatis,
Columbia University, USA

Scott Ferson,
Applied Biomathematics, USA

Wolfgang Graf,
Dresden University of Technology, Germany

Vladik Kreinovich,
University of Texas at El Paso, USA

Baoding Liu,
Tsinghua University, China

Dagang Lu,
Harbin Institute of Technology, China

Kurt Marti,
Federal Armed Forces University Munich, Germany

Mehdi Modares,
Illinois Institute of Technology, USA

Zissimos Mourelatos,
Oakland University, USA

Drahomír Novák,
Brno University of Technology, Czech Republic

Arnold Neumaier,
University of Vienna, Austria

Michael Oberguggenberger,
University of Innsbruck, Austria

Andrzej Pownuk,
University of Texas at El Paso, USA

M.V.Rama Rao,
Vasavi College of Engineering, Hyderabad, India

Siegfried M. Rump,
Hamburg University of Technology, Germany

Gerhart I. Schuëller,
University of Innsbruck, Austria

Pol D. Spanos,
Rice University, USA

Mark Stadtherr,
University of Notre Dame, USA

Lev Utkin,
St.Petersburg State Forest Technical Academy, Russia

Dirk Vandepitte,
Catholic University of Leuven, Belgium

Reinhard Viertl,
Vienna University of Technology, Austria

Hao Zhang,
The University of Sydney, Australia

Enrico Zio,
Polytechnic of Milan, Italy

Abstracts of Plenary Lectures

Reliability-based Optimization: An Overview and Recent Advances

G.I. Schuëller[1)], **M.A. Valdebenito**[2)]

[1)]Professor, Institute of Engineering Mechanics, University of Innsbruck, Technikerstraße 13, A-6020, Innsbruck, Austria, EU, e-mail: Mechanik@uibk.ac.at (presenter)

[2)]Assistant Professor, Departamento de Obras Civiles, Universidad Tecnica Federico Santa Maria, Av. Espana 1680, Valparaiso, Chile

Abstract

Engineering aims at designing systems that can fulfill prescribed performance objectives within a certain life time. For example, consider the design of a metallic structure subject to cyclic loading. In this particular case, a relevant design objective would be ensuring that fatigue damage leads neither to loss of serviceability nor to collapse. This design objective may be achieved, e.g. by sizing structural members such that stresses are below a certain threshold or by scheduling appropriate maintenance activities. As resources for constructing and maintaining any structure are always scarce, the final design of the target system should not only comply with prescribed performance objectives but also its life time cost should be as economic as possible. In view of this last statement, it is clear that the design task may be interpreted as the solution of an optimization problem, i.e. the objective is to minimize overall costs, while ensuring that the structural performance is within acceptable limits [1].

Although the formulation of the design problem within the context of optimization is certainly advantageous, there is a major issue present in almost all practical design situations: several parameters which are relevant for design cannot be quantified by precise, deterministic values, as they are inherently uncertain. Typical examples of these parameters include loadings, member sizes, material properties, etc. The uncertainties in these parameters affect the structural response. In consequence, the behavior of the structure will be uncertain as well. Therefore, the presence of these uncertainties should be reflected in the design process of a system. A possible means to quantify the effects of uncertainty in the system's response is resorting to probability concepts, as they allow calculating reliability, i.e. the probability that the performance objectives will be fulfilled [2]. The consideration of the effects of uncertainty by means of reliability within the design process (formulated as an optimization problem) is known in the literature as reliability-based optimization (RBO) [3].

Although RBO constitutes a most powerful tool for design in engineering, its application to problems of engineering interest has remained limited in the past due to high numerical costs involved in its solution. These high costs are due to repeated evaluation of structural response (by means of numerical methods such as finite elements) required for solving problems of optimization and structural reliability. Nonetheless, in recent years several new methodologies have been developed which render involved RBO problems tractable. Within this context, the objective of this lecture is presenting some of the most recently developed tools for RBO. The focus is on two aspects. The first one comprises the application of advanced simulation techniques, which have opened the possibility for assessing structural reliability for large structures [4], particularly for problems involving a large number of uncertain parameters (in the order of thousands) as well as a large number of failure criteria. The second aspect is the efficient assessment of reliability sensitivity [5], i.e. how much does the structural reliability vary due to a perturbation in variables that are controlled by the designer. Theoretical as well as practical aspects on the application of tools for solving RBO problems are discussed. Case studies are also analyzed in order to show the applicability and efficiency of the tools introduced. Special emphasis is given to applications involving optimal structural design for stochastic linear and nonlinear dynamics [5,6] as well as optimal maintenance scheduling for fatigue-prone structures [7].

Keywords: *Uncertainties; structural reliability; advanced simulation techniques; reliability-based optimization; reliability sensitivity; optimal maintenance scheduling.*

References

[1] R.T. Haftka and Z. Gürdal. *Elements of Structural Optimization.* Kluwer, 3rd edition, 1992.

[2] A.M. Freudenthal. Safety and the probability of structural failure. *ASCE Transactions*, 121:1337–1397, 1956.

[3] I. Enevoldsen and J.D. Sorensen. Reliability-based optimization of series systems of parallel systems. *Journal of Structural Engineering*, 119(4):1069–1084, 1993.

[4] G.I. Schuëller, H.J. Pradlwarter, J.L. Beck, S.K. Au, L.S. Katafygiotis and R. Ghanem. Benchmark study on reliability estimation in higher dimensions of structural systems – An overview. In C. Soize and G.I. Schuëller (eds.), *Structural Dynamics EURODYN 2005 – Proc. of the 6^{th} Int. Conf. on Structural Dynamics*, pages 717–722, Paris, France, September 4–7, 2005. Millpress, Rotterdam.

[5] M.A. Valdebenito and G.I. Schuëller. Efficient strategies for reliability-based optimization involving non linear, dynamical structures. *Computers & Structures*, 89(19–20):1797–1811, 2011.

[6] M.A. Valdebenito, H.A. Jensen, G.I. Schuëller, and F.E. Caro. Reliability sensitivity estimation of linear systems under stochastic excitation. *Computers & Structures*, 92–93:257–268, 2012.

[7] P. Beaurepaire, M.A. Valdebenito, H.A. Jensen, and G.I. Schuëller. Reliability-based optimization of maintenance scheduling of mechanical components under fatigue. *Computer Methods in Applied Mechanics and Engineering*, 221–222:24–40, 2012.

Terror, Security and Money: Balancing the Risks, Benefits, and Costs of Critical Infrastructure Protection

M.G. Stewart[1)], **J. Mueller**[2)]

[1)]Australian Research Council Professorial Fellow Director,
Centre for Infrastructure Performance and Reliability, The University of Newcastle
New South Wales, 2308, Australia
Corresponding Author – phone: +61 2 49216027, email: mark.stewart@newcastle.edu.au

[2)]Professor and Woody Hayes Chair of National Security Studies
Mershon Center for International Security Studies and Department of Political Science
Ohio State University, Columbus, Ohio 43201, United States

Abstract

The loading and response of structures to explosive blast loading is subject to uncertainty and variability. This uncertainty can be caused by variability of dimensions and material properties, model errors, environment, etc. Limit state and LRFD design codes for reinforced concrete and steel have been derived from probabilistic and structural reliability methods to ensure that new and existing structures satisfy an acceptable level of risk. These techniques can be applied to the area of structural response of structures subject to explosive blast loading. The use of decision theory to determine acceptability of risk is crucial to prioritise protective measures for built infrastructure. Government spending on homeland security will reach $141.6 billion worldwide in 2009 and is projected to reach $300 billion by 2016. The question is, is this or other expenditure necessary? Clearly, scientific rigour is needed when assessing the effectiveness and the need for protective measures to ensure that their benefits exceed the cost. The paper will assess terrorist threats to buildings, bridges and transportation infrastructure and the cost-effectiveness of protective and counter-terrorism measures. Structural reliability and probabilistic methods are used to assess risk reduction due to protective measures. The key innovation is incorporating uncertainty modelling in the decision analysis, which in this case will maximise net benefit. This analysis will then consider threat likelihood, cost of security measures, risk reduction and expected losses to compare the costs and benefits of security measures to decide which security measures are cost-effective, and those which are not.

For additional and wider-ranging assessments of the issues raised and the approaches used, see the reference.

Reference

J. Mueller and M.G. Stewart. *Terror, Security, and Money: Balancing the Risks, Benefits, and Costs of Homeland Security.* Oxford University Press, forthcoming September 2011.

Reliable Computing and Fuzzy Information

R. Viertl

Professor, Department of Statistics and Probability Theory,
Vienna University of Technology, Wiedner Hauptstr. 8, 1040 Wien, Austria
phone: +43-1-58801-10720, email: R.Viertl@tuwien.ac.at

Abstract

In engineering computations there exist different kinds of uncertainty. The most important are variability, imprecision of data, model uncertainty, and uncertainty of a-priori information. Whereas variability is modeled since a long time by probability models, the quantitative mathematical description of imprecision by so-called fuzzy models was done more recently. Especially in reliability calculations in the Bayesian context also a-priori information is best modeled by so-called fuzzy probability distributions. Examples of non-precise data and related fuzzy models and relationships to stochastic models will be given in the contribution.

Reference

R. Viertl. *Statistical Methods for Fuzzy Data*. Wiley, Chichester, 2011.

Contents

Power Transformer Protective System Reliability Assessment

Abdelkader ABEDLMOUMANE*, Hamid BENTARZI* and Dris OUAHDI**
**Signals and Systems Laboratory, IGEE, UMBB University, Boumerdes, ALGERIA*
E-mail: sisylab@yahoo.com
*** Department of Electrification and Automation, FHC,*
UMBB University, Boumerdes, ALGERIA

Abstract: The reliability and the safety of the power systems are mainly related to the power transformers. Many precautions have been taken for protecting well these transformers from defects because when they break down causing large fires and explosions, and hence the consequences can be very heavy in terms of damage as well as in terms of economic loss.

Published statistics on failures, fires and explosions of these large power transformers show that these types of incidents frequently appear, so that power companies pay a special attention to the effectiveness and the reliability of the protective systems. In this work, a new approach allowing an evaluation of the reliability of the protective systems of power transformer with and without redundancy has been presented. Models of protective system (without redundancy and with redundancy) have been developed using the fault tree. Then, a comparative study has been carried out using our software program ERPT[1] which has shown that the failure rate significantly decreases in the case of protective system with redundancy.

Keywords: Reliability Evaluation, Redundancy, Protection system, Power transformer, Fault tree.

1. Introduction

The safety of the power systems is mainly related to the power transformers. Many precautions have been taken for protecting these transformers against faults because when they break down they often cause large fires and explosions, and hence the consequences can be very heavy in terms of damage as well as in terms of economic loss.

A great number of experts have noticed that an important increase in transformers failures during the last years. In the majority of the countries, the privatization of companies of production and distribution of electricity gives place to a reduction of investments. Moreover, the consumption of electricity increases regularly by 2% per year around the world. The old transformers are, therefore, often overloaded. Moreover, published statistics on failures, fires and explosions of these large power transformers show that these types of incidents frequently appear, so that power companies pay a special attention to the effectiveness and the reliability of the protective systems.

In this work, a new approach allowing an evaluation of the reliability of the protective systems of power transformer with and without redundancy has been presented. The first part of the paper is devoted with the presentation of the various defects undergone by the transformers and their adequate protection

[1]ERPT (version 1.1 2011) software of calculation and simulation developed by the authors in the university of Boumerdes, Algeria.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

techniques, (Abdelmoumene, 2009). Then, models using the fault tree, (Chafai, Refoufi, and Bentarzi, 2009), have been developed for both protective systems: without redundancy and with redundancy. A comparative study shows that the reliability of the protective system with redundancy is increased several times as compared to its reliability without redundancy using our software program ERPT.

2. Faults and their Origins in Power Transformer

Transformers are subjected to many external electrical stresses upstream and downstream. The consequences of any failure can be very great in terms of damage as well as in terms of operating losses. Therefore, knowledge of faults and their origins is very important to choose the adequate interruption devices, protective relays and their adjustments. The determination and the evaluation of these faults become difficult in some cases. The use of specialized software program helps us to carry out these calculations and simulations more quickly and more precisely during steady state as well as transient state. Among these electrical stresses such as overvoltage, overload and short circuit may cause faults as illustrated in Fig.1.

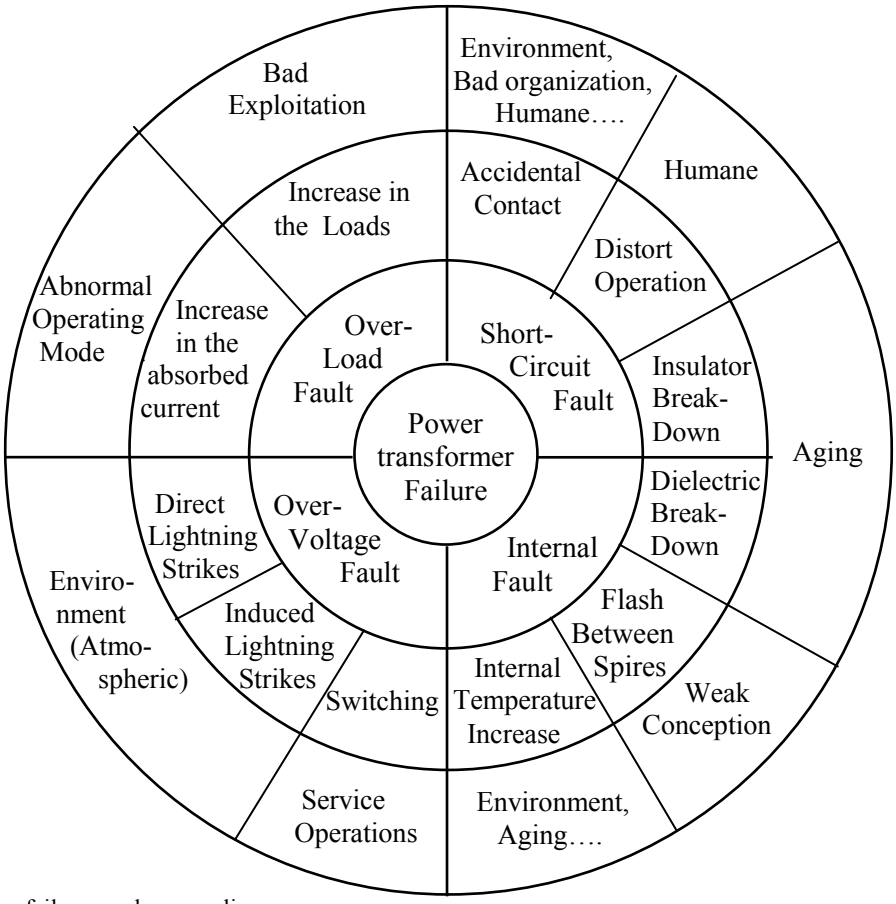


Figure 1: Transformer failures and causes diagram.

2.1. OVERVOLTAGE

The transformers are subjected to transient over-voltages coming from the networks to which they are connected. These over-voltages have produced either by direct or induced lightning which strokes on HV or LV networks, or from switching actions of elements in the upstream network. Besides, they can be produced from ferro-resonance that is a non-linear resonance phenomenon. This phenomenon typically involves the saturated magnetizing inductance and a capacitive of any elements of the network. Its occurrence is more likely in the absence of adequate damping (Boutora, Bentarzi and Ouadi, 2010).

2.1.1 *Lightning*

Studies showed that incidents which have occurred in transformers are due to the internal over-voltages created by lightning shocks, whose propagation on line excited the specific resonance of the transformers.It was shown that for a thunderbolt striking a high voltage line to about 12 km up to the transformer, an important overvoltage of resonance may be generated and hence the dielectric breakdown noted in the tap changer (Ahmad, 1992). At the time of a direct thunderbolt fall on lines or transformers (this is rarely happened, once every 100 years in a determined place) these last don't dissolve, they vaporize!! (Gérard, 2008), (Tessier, 2006). In France, about 5000 power transformers per year destroyed by the lightning. (France Transfo, 2007).

2.1.2 *Stakes under and over voltage*

During the energizing or de-energizing by switchgear situated immediately upstream, over-voltages can be generated by the combined transformer- supply circuit, switchgear set leads to a dielectric stress in the transformer. This stress causes premature ageing, or even an insulation fault between turns or to ground.

2.2. OVERLOADS

The overload can be due to the increase in the number of loads fed simultaneously or to the increase in the power absorbed by one or more loads. It results in an over-current which causes a rise in temperature prejudicial to the characteristics of insulators and the longevity of the transformer.

2.3. SHORT-CIRCUIT

A short-circuit is an accidental connection between conductors with small or null impedance such as solid short-circuit. It can be internal within the transformer or external.

2.4. INTERNAL FAULT

The main causes of internal faults are: dielectric breakdown, flash between spires or windings and the inner temperature increasing.

3. Protection of Transformer

3.1. OVERVOLTAGE PROTECTION

Techniques that may be used to protect the power transformers against over-voltage are described below:

Earth wires. The earth wires placed above the power lines attract the thunderbolt, and hence avoid the thunder-striking on the phase conductors.

Spark-gaps with horns. The spark-gap is a simple device consists of two electrodes, the first connected to the conductor to be protected, the second connected to the ground. At the place where it is installed in the network, it presents a weak point for discharging the over-voltages to the ground and thus protects the equipments. The arcing-voltage setting of the spark-gap is adjusted by acting on the distance in the air between electrodes.

Lightning arresters. The lightning arrestors are apparatuses intended to limit the over-voltages imposed on electric transformers, instruments and machines by the lightning and by maneuvers of commutation. The upper part of the lightning arrestor is connected to one of wire of the power line to be protected and the lower part is connected to the ground by a low resistance earthing, generally with less than one ohm. Their principle of operation is based on strongly non-linear resistances which present an important reduction in their inner resistance above of a certain terminal voltage value (Boutora, Bentarzi, Ouadi, 2011).

3.2. OVERLOADS PROTECTION

An overload protection must act with a threshold values ranging between 110 and 150 % of the rated current and preferably operate in a time dependant manner. The protective devices that have above mentioned characteristics are fuses and thermal relays. The fuses are widely used in the distribution transformers, primarily because of simplicity and of the limited cost. However, the technological limits of their realization involve a certain number of disadvantages.

The protection system can be placed on either primary or secondary side of power transformer. For low transformer power, the position of the protection is suitable on the low voltage side. While, for high power, the more chosen place of protection is on the HV side (Fulchiron, 1992).

3.3. PROTECTION AGAINST THE SHORT-CIRCUITS

The protective devices that may be used for protecting the power transformer against the short-circuits are: electromagnetic relays and fuses: Due to disadvantages of the fuses such as they can be used once time and they cannot be adjusted but they are still used as back up protection as well as they are used for protecting the circuit breaker.

3.4. PROTECTION AGAINST THE INTERNAL FAULTS

The protection of the transformers against the internal defects (internal Breakdowns HT/Ground or between turns) is ensured by the Buchholz and differential protection. The latter (differential protection) is sensitive and a fast clearing technique. This technique of protection detects nonzero differential current, and then activates a circuit breaker that disconnects the transformer (Bouderbala, Bentarzi 2011).

The choice of power transformer protections is relatively complex because that requires to take into account a great number of parameters and several technical specifications can be retained to ensure the same type of protection.

Even if protections are well chosen and dimensioned, the degradation of their characteristics and the consequences of a possible failure leads to search for the technical solutions to overcome these drawbacks in order to reach a high level of reliability.

4. Failure and Reliability Evaluation

Reliability is the aptitude of an entity to achieve a required function, under certain conditions and for a given time t (TEBBI, 2005).It is given by

$$R(t) = P_r(T > t) \tag{1}$$

The function of failure indicates the probability that a failure occurs before a given time, it concerns the distribution function:

$$F(t) = P_r(T \leq t) \tag{2}$$

4.1. USUAL MATHEMATICAL MODELS OF RELIABILITY

Reliability is described by decreasing mathematical laws during time. The density of failure is defined by the following relation:

$$f(t) = \frac{\mathrm{d}F(t)}{\mathrm{d}t} \tag{3}$$

The failure rate is the probability that a system is failing between t and $t + \mathrm{d}t$ knowing that it functioned at t:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P[\mathrm{Efailureover}[t, t+\Delta t] \& \mathrm{Enotfailureover}[0, t]]}{\Delta t \cdot P[\mathrm{EEnotfailureover}[0, t]]}$$

Then,

$$\lambda(t)\mathrm{d}t = \frac{f(t)\mathrm{d}t}{R(t)} \tag{4}$$

From the expression (4), the general expression of the law of reliability as a function of the failure rate can be deduced by (Fulchiron, 1998):

$$R(t) = \mathrm{e}^{-\int_0^t \lambda(u)\mathrm{d}u} \tag{5}$$

If the failure rate is constant $\lambda(t) = \lambda$, the expression of reliability is:

$$R(t) = \mathrm{e}^{-\lambda t} \tag{6}$$

Equation (6) is the exponential law. If the failure rate depends on time, the used function is generally the function of Weibull which can be written:

$$R(t) = \mathrm{e}^{-\left(\frac{t}{\eta}\right)^{\beta}} \tag{7}$$

Where β is the parameter of the form and η is the parameter of the scale.

Reliability of a series system. A series system functions if and only if all the components function. Its reliability is calculated by the following relation:

$$\begin{aligned} R(t) &= P_r(T > t) \\ &= P_r[(T_1 > t) \cap (T_2 > t) \cap \dots] \\ &= \prod_i P_r\,(T_i > t) \\ &= \prod_i R_i\,(t) \end{aligned} \tag{8}$$

Reliability of a parallel system. A parallel system functions so at least one of its components functions. Reliability for this system is given by:

$$\begin{aligned} R(t) &= 1 - F(t) \\ &= 1 - P_r[(T_1 \le t) \cap (T_2 \le t) \cap \dots.] \\ &= 1 - \prod_i P_r\,(T_i \le t) \end{aligned} \tag{9}$$

4.2. SAFETY

Safety is the probability of the system does not have any catastrophic failure between the initial time and the time t. The concept of safety is directly related to the criticality of the failures. The failures can present several classes of external consequences on the application of the system: benign, severe, and critical catastrophic. Safety is the privileged criterion of the applications of high criterion for which the consequences of certain failures are catastrophic. This criterion measures confidence ascribable to the product not to present a failure whose external consequences are catastrophic.

5. Redundancy

Redundancy is the existence in an entity of more than one means for accomplishing a required function. It can be noted that there exists several techniques of redundancies such as active redundancy, passive redundancy, the redundancy m among N, the differential redundancy and the redundancy with voting system etc. (Ward, 2004).

5.1 ACTIVE REDUNDANCY

We speak about active redundancy, when all the elements permanently function. We distinguish the total and partial active redundancy. In *total active redundancy,* the system becomes failing only with the failure of the last surviving element. In *partial active redundancy*, when a system comprises *N* elements, where *m (m < N)* components are strictly necessary for its functioning. The system can thus accept *(N − m)* failures.

If the components are independent and are identically distributed ($R = R_i$, $i = 1,2,\ldots,N$), then the reliability of the system is given by:

$$R(t) = \sum_{j=m}^{n} \binom{n}{j} R^j (1-R)^{n-1} \tag{10}$$

with

$$\binom{n}{j} = \frac{n!}{j!\,(n-j)}$$

5.2. PASSIVE REDUNDANCY

In a passive redundancy, the redundant components are brought into service only when the components of the system are failing. If the components have the same law of reliability R(t), the reliability of this system is given by:

$$R_{sys}(t) = R(t) + \int_0^t f(\tau) R(t-\tau) \mathrm{d}\tau \tag{11}$$

5.3. SEQUENTIAL REDUNDANCY

The redundancy is qualified as sequential when the superabundant elements are brought into service only at time of need; that means that among *N* elements only *K* is in service. This implies that certain elements will be in reserve or stock.

5.4. DIFFERENTIAL REDUNDANCY

In the differential redundancy, different tools are used to assure the same function.

6. Fault Tree Method

The Method of fault tree is widely used in the field of the Reliability. It offers a framework privileged to the deductive and inductive analysis by means of a tree structure of logical gates.

The principal treatments carried out on the fault tree are the research of the minimal cuts and the quantitative evaluation.

The minimal cuts represent the smallest combinations of events whose simultaneous realization involves that of the undesirable event. They have as an order the number of events which constitutes them.

7. Protective System Without Redundancy

Generally; the power transformer's protection system is constituted of several elementary protections, each of them assure one or more function.

In our study we will consider that the protection system is composed of:

- Overvoltage protection: lightning arrester.
- Overload protection: fuse.
- Short-circuits protection: circuit-breaker.
- Internal faults protection: Buchholz relay.

7.1 MODELING

Figure 1 represents the fault tree model of this system. Where,

E1: is the event of overvoltage protection failure.
E2: event of overvoltage appearance.
E3: event of overload protection failure.
E4: event of an overload appearance.
E5: event of short-circuit protection failure.
E6: event of appearance of short-circuit.
E7: event of failure of protection against internal fault.
E8: event of appearance of an internal fault.
Eu: undesirable event→ transformer in defect

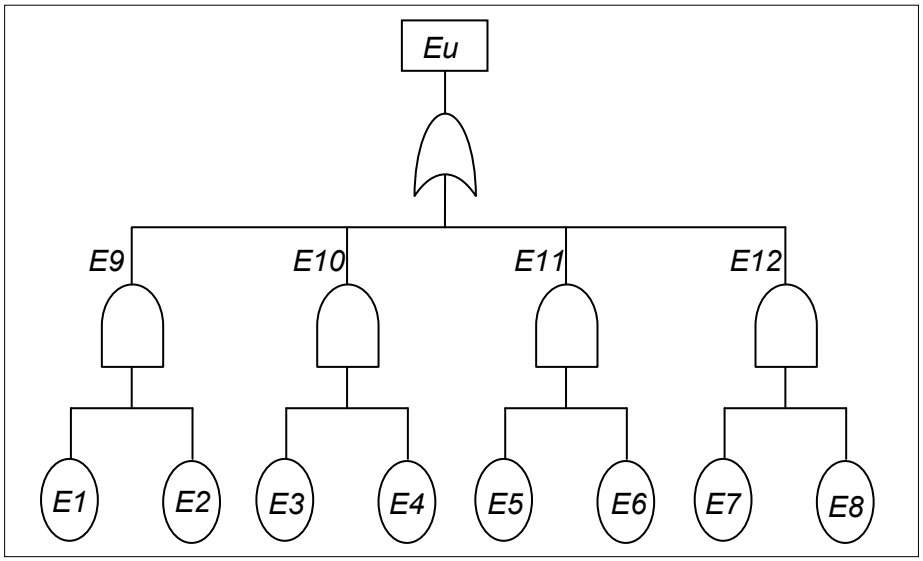


Figure 2: **F**ault **T**ree model of the **P**rotection system without **R**edundancy (FTPOR).

7.2 QUALITATIVE ANALYSIS

According to the model shown in Fig.2, there are four immediate causes of the undesirable event; which are the intermediate events: E9, E10, E11 and E12.

E9: represents an overvoltage,
E10: overload fault,
E11: short-circuit fault,

E12: internal fault.

The decomposition of these events leads to the initial events*E1, E2,… E8*.

$$\begin{aligned} E9 &= E1 \wedge E2 \\ E10 &= E3 \wedge E4 \\ E11 &= E5 \wedge E6 \\ E12 &= E7 \wedge E8 \\ Eu &= E9 \vee E10 \vee E11 \vee E12 \\ &= (E1 \wedge E2) \vee (E3 \wedge E4) \vee (E5 \wedge E6) \vee (E7 \wedge E8) \end{aligned} \quad (12)$$

Equation (12) is the logical equation of the FTPOR model. The minimal cuts are as follow: *E1 E2, E3E4, E5E6& E7E8.*

7.3. QUANTITATIVE ANALYSIS

A step of such analysis exploits information of quantitative nature or quantified: it is, for example, the failure rates of components, the conditions of working and of environment, probabilities of events, etc. Failure rate of over-voltages protection is 0.0570. Failure rate of overloads protection is 0.0690. Failure rate of short-circuits protection is 0.0750. A failure rate of internal protection is 0.0953. Probability of an overvoltage appearance is 0.1600. Probability of an overload appearance is 0.1900. Probability of a short-circuit appearance is 0.0800. Probability of an internal defect appearance is 0.0329.

The calculation carried out by our software program ERPT gives the following results (see Fig.3). According to this model, it can be noted that:

- The probability of failure of the protective system is about: 3%.
- The highest failure rate is that of protection against the internal faults.
- The most critical protection is the overload protection.

By using ERPT software, we can plot reliability or failure curves of the different elements of the FTPOR model as shown in Fig.4. These curves make it possible to determine the reliability of the elements in a given instant. They also allow the determination of the most critical elements, from reliability point of view, in the protective system. More the curve is concave, more the represented element is critical.

The analysis of criticality is the object of another analysis method; it is the FMECA method: Failure Mode Effect and Criticality Analysis.

We can deduce easily that:

$$\lambda_{IFP} > \lambda_{SCP} > \lambda_{OLP} > \lambda_{OVP} \quad (13)$$

8. Reliability Modeling of Protective System with Redundancy

8.1. MODELING

We will consider that the protective system is composed of:

- Overvoltage protection: the earth-wire and the lightning arrestor.
- Overload protection: circuit breaker and fuse.

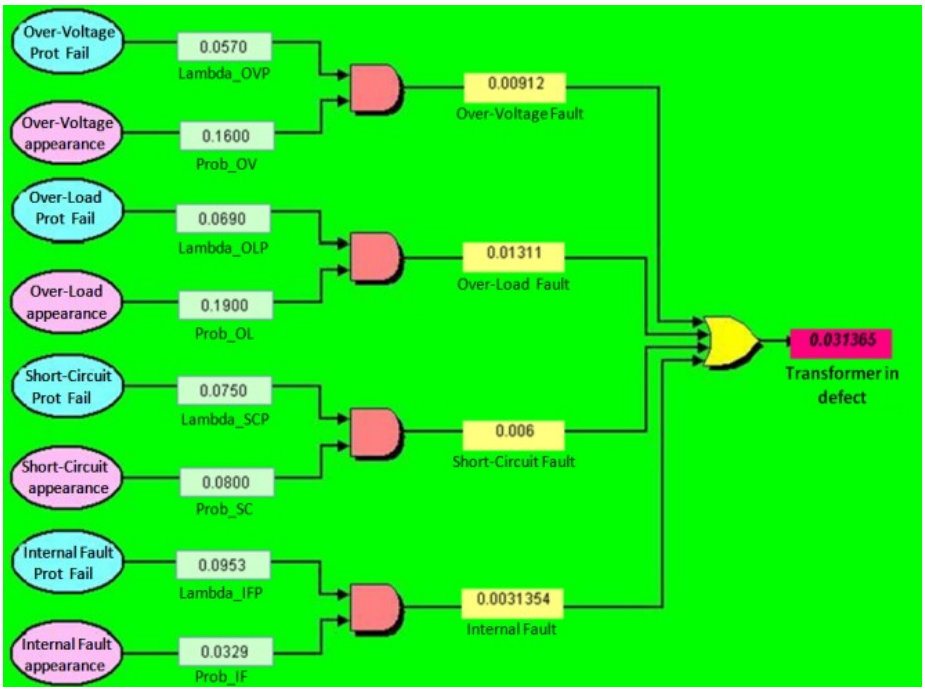


Figure 3: Results of calculation of the FTPOR model.

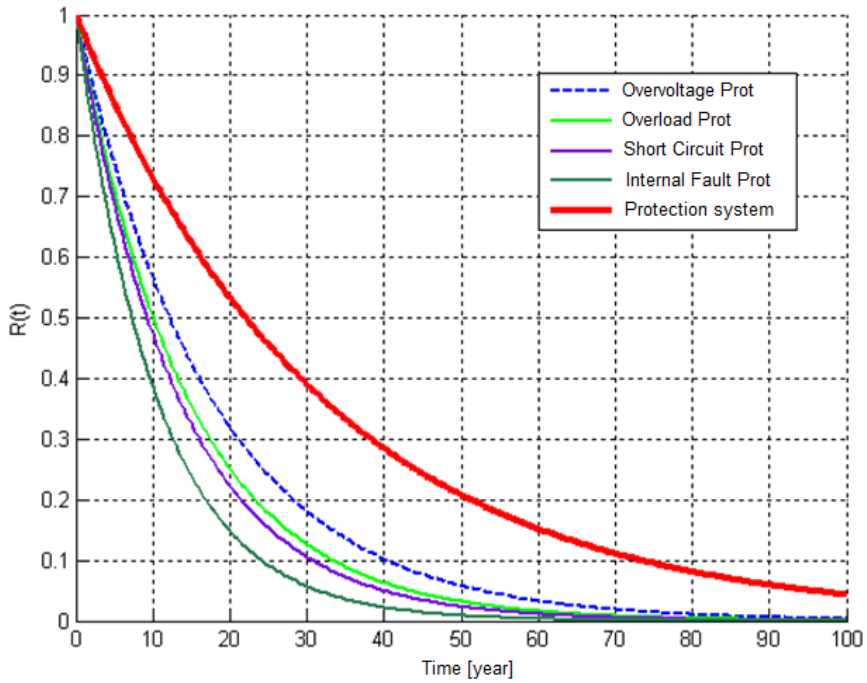


Figure 4: Reliability of protections (without Redundancy).

- Protection against the short-circuits: circuit breaker and fuse.
- Protection against the internal defects: Buchholz and temperature sensors.

The fault tree model of this system is represented on the figure5. Where,

E1: is the event of Fusible failure.
E2: event of circuit breaker failure.
E3: event of failure of Lightning arrestor.
E4: event of Earth-wire failure.
E5: event of fuse failure.
E6: event of circuit breaker failure.
E7: event of failure of Buchholz relay.
E8: event of failure of thermal protection.
E10: event of appearance of a short-circuit.
E12: event of appearance of an overvoltage.
E14: event of overload appearance.
E16: event of appearance of an internal defect.
Eu: undesirable event→ transformer in defect.

8.2. QUALITATIVE ANALYSIS

As in the previous model, we have four immediate causes which lead to the undesirable event.

E17: overvoltage fault.
E18: overload fault.
E19: short-circuit fault.
E20: internal fault.

Decomposition of these events:

$$E17 = E9 \wedge E10 = (E1 \wedge E1) \wedge E10$$
$$E18 = E1 \wedge E12 = (E3 \wedge E4) \wedge E12$$
$$E19 = E13 \wedge E14 = (E5 \wedge E6) \wedge E14$$
$$E20 = E15 \wedge E16 = (E7 \wedge E8) \wedge E16$$

Finally, we find the logical equation of the model:

$$Ei = [(E1 \wedge E2) \wedge E10] \vee [(E3 \wedge E4) \wedge E12] \vee [(E5 \wedge E6) \wedge E14] \vee [(E7 \wedge E8) \wedge E16]$$

The minimal cuts are the following ones: E1E2E10, E3E4E12, E5E6E14 & E7E8E16.

8.2. QUANTITATIVE ANALYSIS

The computation results of the model which is obtained by software ERPT are shown in Figs.6 and 7.
It may be noted that:

- The probability of failure of the protective system is approximately 3‰;
- Overvoltage protection has the highest probability of failure.

- The most critical protection is the protection against overvoltage faults.
- Within the operational limits, it is more practical to model the function R(t) of this protective system by the following equation: : $R(t) = -0.003\,t + 1$

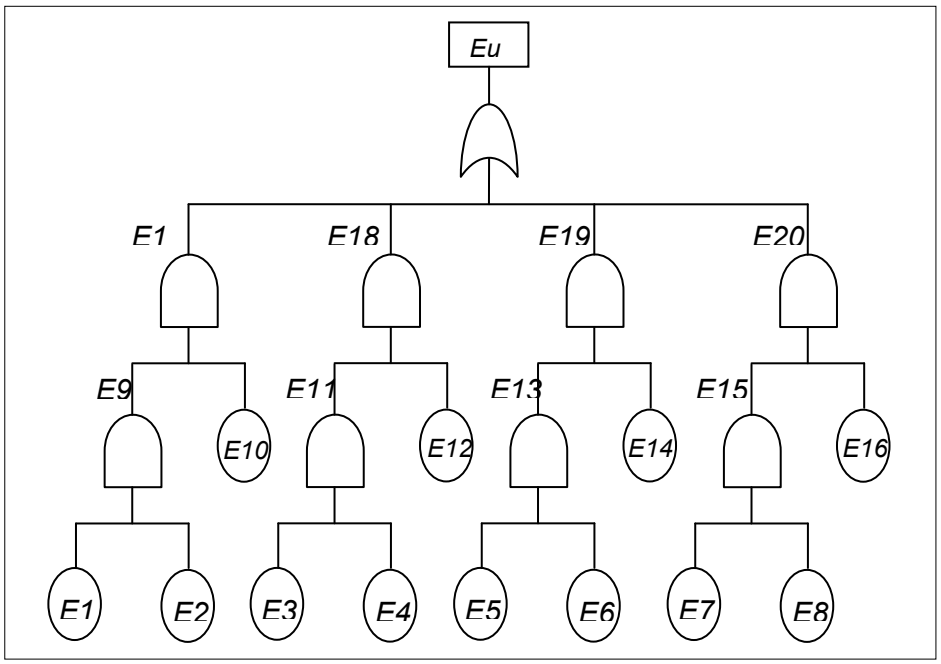


Figure 5: **F**ault **T**ree model of the **P**rotection system with **R**edundancy (FTPR).

The decreasing order of failure rates is:

$$\lambda_{\mathrm{OVP}} > \lambda_{\mathrm{IFP}} > \lambda_{\mathrm{SCP}} > \lambda_{\mathrm{OLP}} \tag{14}$$

We can also model the failure function of the protective system in the following manner:

$$F(t) + R(t) = 1$$
$$F(t) = 1 - R(t) = -0.003\,t$$

9. Comparison Between Two Models

In order to judge the effectiveness of the redundancy technique; we will introduce the concept of the improvement factor.

The improvement factor $\Gamma = \frac{\lambda_{FTPOR}}{\lambda_{FTPR}}$, then $\Gamma \approx 10$. The failure rate is approximately ten times smaller in the protective system with redundancy relatively to the protective system without redundancy. The reliability of the protective system with redundancy is increased several times the reliability of protective

system without redundancy. According to Figure 8, it clearly that the factor of improvement of reliability is not constant; it is a function of time $\Gamma(t)$ such as:

$$\Gamma(t) = e^{\lambda_a \cdot t}$$

with

$$\lambda_a = \lambda_{FTPOR} - \lambda_{FTPR}$$

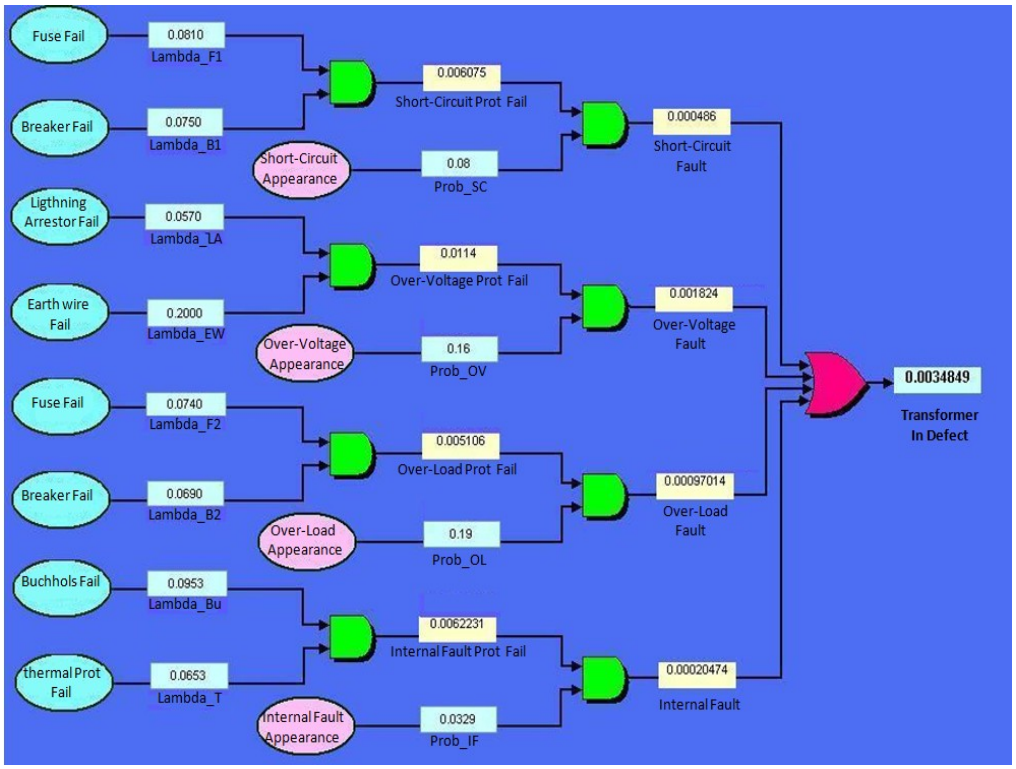


Figure 6: Results of calculation of the FTPR model.

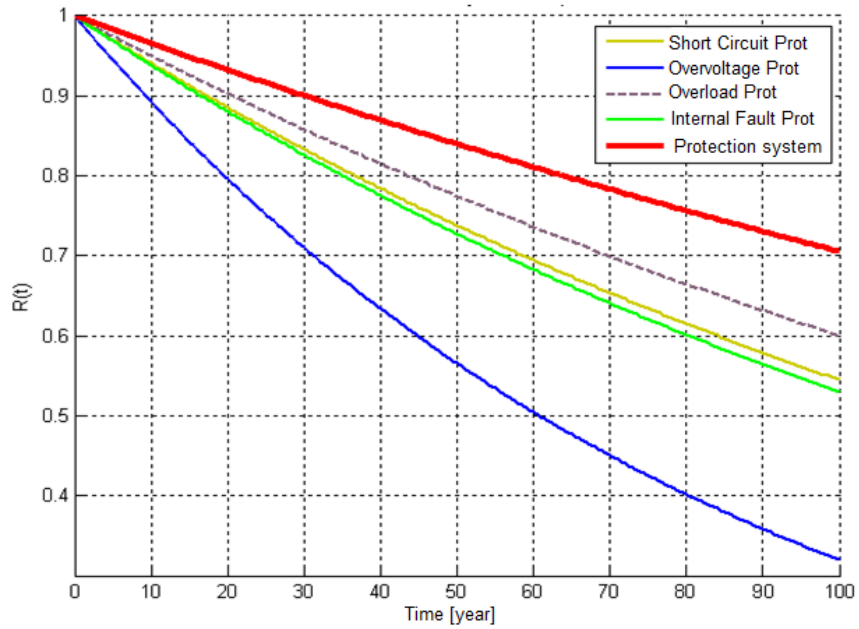


Figure 7: Reliability of protections (with Redundancy)

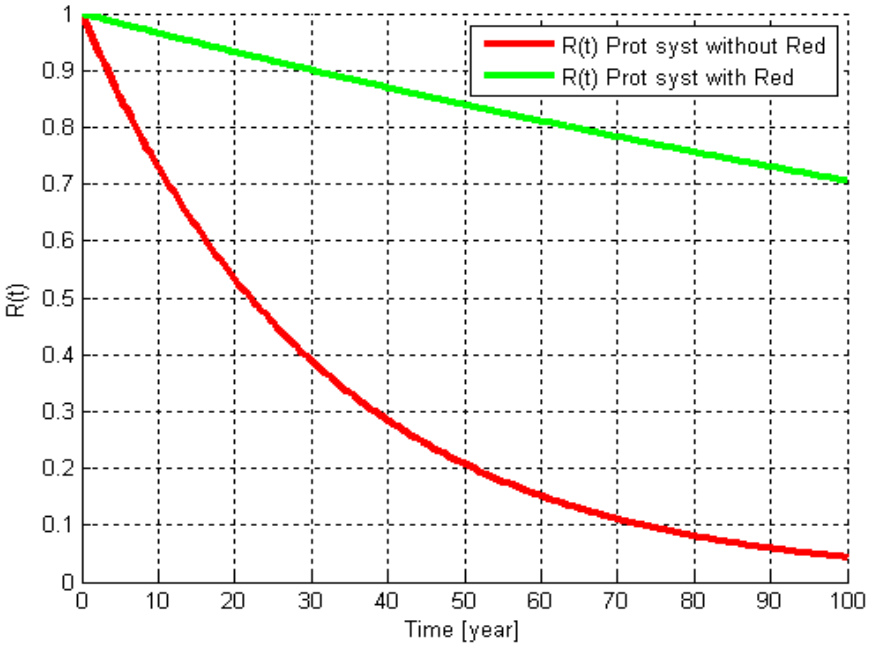


Figure 8: Reliability of protection systems: without & with redundancy.

10. Conclusion

The evaluation and the optimization of the reliability of protection system are essential to safeguard any element in large electrical systems.

In this paper, we proposed a comprehensive approach leading to reduce the failure rate to about one tenth in the power transformer protected by protective system with redundancy and hence its reliability has been enhanced.

The use of the redundancy technique in the power transformer protection systems leads:

- to reduce the failure rate;
- to increase the reliability and the lifespan of the transformers;
- to ensure a continuity of service more reliable and more secure;
- to improve the technical and economic indices of the exploitation.

References

Abdelmoumene A., Etude de la redondance sur le système de protection du transformateur de puissance, *Magistère thesis*, BoumerdesUniversity, Algeria, 2009.

Chafai, M., Refoufi, L. and Bentarzi, H. Reliability Assessment and Improvement of Medium Power Induction Motor Winding Insulation Protection System Using Predictive Analysis, in *Advances in Numerical Methods*, Book-Chapter Verlag-Springer, Berlin Heidelberg,367-378, 2009.

BOUTORA S., H. BENTARZI and A. OUADI, "Ferroresonance Avoidance by Distributed Resources Reconfiguration in Islanded Power Grid, *International Conference (CSECS '10*), pp. 117-121, Athena, Greece, December 29-31. 2010.

Boutora S., Bentarzi H., Ouadi A., Analysis of the Disturbances in Distribution Networks using Matlab and ATP, *INTERNATIONAL JOURNAL of ENERGY*, 2011

Gerard F., Généralités sur la foudre et les surtensions, *ADEE France*, Octobre 2008.

FRANCE TRANSFO, *Présentation de la société France Transfo,* Septembre 2007.

Bouderbala R., Bentarzi, H. Ouadi A., Digital Differential Relay Reliability Enhancement of Power Transformer, *INTERNATIONAL JOURNAL of CIRCUITS, SYSTEMS and SIGNAL PROCESSING*, 2011.

Fulchiron D., protection des transformateurs de postes MT/BT, *, cahier technique,* Schniederelectric, 1998.

TEBBI O., Estimation des lois de fiabilité en mécanique par les essais accélérés, *Thèse de doctorat*, Université d'Angers, Mars 2005.

Ward S., Dahlin T., Higinbotham W., Improving Reliability for Power System Protection,*58th Annual Protective Relay Conference,,*Atlanta, GA, 2004.

Efficient Sparse Polynomial Chaos Expansion Methodology for Computationally-expensive Deterministic Models

Tamara Al-Bittar[1)], and Abdul-Hamid Soubra[2)]
[1)] *PhD student, University of Nantes, Bd. de l'université, BP 152, 44603 Saint-Nazaire cedex, France; E-mail: Tamara.Al-Bittar@univ-nantes.fr*
[2)] *Professor, University of Nantes, Bd. de l'université, BP 152, 44603 Saint-Nazaire cedex, France; E-mail: Abed.Soubra@univ-nantes.fr*

Abstract: The sparse polynomial chaos expansion (SPCE) methodology is an efficient approach that deals with uncertainties propagation in case of high-dimensional problems (i.e. when a large number of random variables is involved). This methodology significantly reduces the computational cost with respect to the classical full polynomial chaos expansion (PCE) methodology. Notice however that when dealing with computationally-expensive deterministic models, the time cost remains important even with the use of the SPCE. In this paper, an efficient combined use of the SPCE methodology and the global sensitivity analysis (GSA) is proposed to solve such a problem. The proposed methodology is validated using a relatively non-expensive deterministic model.

Keywords: Sobol indices; spatial variability; sparse polynomial chaos expansion, global sensitivity analysis.

1. Introduction

An efficient approach to deal with uncertainties propagation in case of high-dimensional problems (i.e. when a large number of random variables is involved) was recently presented by Blatman and Sudret (2010). This approach is based on a Sparse Polynomial Chaos Expansion (SPCE) for the system response and leads to a reduced computational cost as compared to the classical Polynomial Chaos Expansion (PCE) methodology. Notice that both, the PCE and the SPCE methodologies, aim at replacing the original expensive deterministic model which may be an analytical model or a finite element/finite difference model by a meta-model. This allows one to calculate the system response using a simple analytical equation (e.g. Isukapalli et al., 1998; Huang et al., 2009; Mollon et al., 2011; Mao et al., 2012). Notice however that when dealing with computationally-expensive deterministic models with a large number of random variables, the time cost remains important even with the use of the SPCE. Consequently, a method that can reduce once again the cost of the probabilistic analysis is needed. In this paper, an efficient combination between the SPCE methodology and the Global Sensitivity Analysis (GSA) is proposed to solve such a problem. In this method, a small SPCE order is firstly selected to approximate the system response by a meta-model. A GSA based on Sobol indices is then performed on this small SPCE order to determine the weight of each random variable in the variability of the system response. As a result, the variables with very small values of their Sobol indices (i.e. those that have a small weight in the variability of the system response) can be discarded.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

Consequently, a response which only depends on a smaller number of random variables is obtained. In other words, one obtains a response with an 'effective dimension'. This dimension is smaller than the initial dimension where the total number of random variables was considered. As it will be shown later, the use of a small SPCE to perform the GSA is not a concern since higher SPCE orders lead to the same influential random variables. Once the 'effective dimension' was determined, a higher SPCE order that makes use of only the most influential random variables can be used. This significantly reduces the computation time. The use of a higher SPCE order is necessary in order to lead to an improved fit of the SPCE.

The proposed methodology is validated using a relatively non-expensive model which was extensively investigated by Al-Bittar and Soubra (2011, 2012). This model involves the computation of the ultimate bearing capacity of a strip footing resting on a weightless spatially varying (c, φ) soil where c is the soil cohesion and φ is the soil angle of internal friction. It should be noticed here that the random fields of c and φ are discretized into a finite number of random variables. This number is small for very large autocorrelation distances and significantly increases for small values of the autocorrelation distances.

The paper is organized as follows: The next two sections aim at briefly presenting both the sparse polynomial chaos expansion (SPCE) and the global sensitivity analysis (GSA). Then, the proposed efficient combination between the SPCE methodology and the GSA is presented. It is followed by the numerical results. The paper ends with a conclusion.

2. Sparse polynomial chaos expansion (SPCE) methodology

In this section, one first presents the polynomial chaos expansion (PCE) and then its extension, the sparse polynomial chaos expansion (SPCE). The Polynomial Chaos Expansion (PCE) methodology allows one to replace an expensive deterministic model which may be an analytical model or a finite element/finite difference numerical model by a meta-model. Thus, the system response may be calculated using a simple analytical equation. This equation is obtained by expanding the system response on a suitable basis which is a series of multivariate polynomials that are orthogonal with respect to the joint probability density function of the random variables.

The PCE theory was originally formulated with standard Gaussian random variables and Hermite polynomials (Ghanem and Spanos, 1989). It was later extended to other types of random variables that use other types of polynomials (Xiu and Karniadakis, 2002). In this paper, standard normal random variables in conjunction with Hermite polynomials are used. The coefficients of the PCE may be efficiently computed using a non-intrusive technique where the deterministic calculations are done using for example an analytical model or a finite element/finite difference software treated as a black box. The most used non-intrusive method is the regression approach (e.g. Isukapalli et al., 1998; Huang et al., 2009; Blatman and Sudret, 2010; Mollon et al., 2011; Mao et al., 2012). This method is used in the present work. The PCE methodology can be briefly described as follows:

For a deterministic model Γ with M random variables, the system response can be expressed by a PCE of order p fixed by the user as follows:

$$Y = \Gamma_{PCE}(\xi) = \sum_{\beta=0}^{\infty} a_\beta \Psi_\beta(\xi) \cong \sum_{\beta=0}^{P-1} a_\beta \Psi_\beta(\xi) \tag{1}$$

where P is the number of terms retained in the truncation scheme, $\xi = \{\xi_i\}_{i=1,\ldots,M}$ is a vector of M independent standard random variables that represent the M random variables, a_β are unknown coefficients to be computed and Ψ_β are multivariate Hermite polynomials which are orthogonal with respect to the joint probability density function (PDF) of the standard normal random vector ξ. These multivariate Hermite polynomials can be obtained from the product of one-dimensional Hermite polynomials as follows:

$$\Psi_\beta = \prod_{i=1}^{M} H_{\alpha_i}(\xi_i) \tag{2}$$

where α_i (i = 1, …, M) are a sequence of M non-negative integers and $H_{\alpha_i}(.)$ is the α_i^{th} one-dimensional Hermite polynomial. The expressions of the one-dimensional Hermite polynomials are given in Ghanem and Spanos (1989) among others.

In practice, the PCE with an infinite number of terms should be truncated by retaining only the multivariate polynomials Ψ_β of degree less than or equal to p. For this purpose, the classical truncation scheme based on the determination of the first order norm $\|\alpha\|_1 = \sum_{i=1}^{M} \alpha_i$ is used. This first order norm should be less than or equal to the order p of the PCE. This leads to a number P of the unknown PCE coefficients equal to $\frac{(M+p)!}{M!\,p!}$. This number is significant in the present case of random fields (especially when considering small values of the autocorrelation distances) and thus, one needs a great number of calls of the deterministic model (see Al-Bittar and Soubra 2011, 2012). The SPCE methodology presented by Blatman and Sudret (2010) is an efficient alternative that can significantly reduce the number of calls of the deterministic model. In this methodology, Blatman and Sudret (2010) have shown that the number of significant terms in a PCE is relatively small since the multivariate polynomials Ψ_β corresponding to high-order interaction (i.e. those resulting from the multiplication of the H_{α_i} with increasing α_i values) are associated with very small values for the coefficients a_β. Thus, a truncation strategy (called the hyperbolic truncation scheme) based on this observation was suggested by these authors. Within this strategy, the multivariate polynomials Ψ_β corresponding to high-order interaction were penalized. This was performed by considering the hyperbolic truncation scheme which suggests that the q-norm should be less than or equal to the order p of the PCE. The q-norm is given by:

$$\|\alpha\|_q = \left(\sum_{i=1}^{M} \alpha_i^q \right)^{1/q} \tag{3}$$

where q is a coefficient ($0 < q < 1$). In this formula, q can be chosen arbitrarily. Blatman and Sudret (2010) have shown that sufficient accuracy is obtained for $q \geq 0.5$.

The proposed SPCE methodology leads to a sparse polynomial chaos expansion that contains a small number of unknown coefficients which can be calculated from a reduced number of calls of the deterministic model. This strategy was used in Al-Bittar and Soubra (2011, 2012) and will also be used in this paper to build up a SPCE of the system response. The iterative procedure suggested by Blatman and Sudret (2010) for building up a SPCE is detailed in Blatman and Sudret (2010) and Al-Bittar and Soubra (2011, 2012) and is not repeated herein. Once the coefficients a_β have been computed, the statistical moments (mean, standard deviation, skewness, and kurtosis) can be calculated with no additional cost. The next subsection is devoted to the method used for the computation of the coefficients a_β of the SPCE using the regression approach.

2.1. COMPUTATION OF THE SPCE COEFFICIENTS BY THE REGRESSION APPROACH

Consider a set of K realizations $\{\xi^{(1)}=(\xi_1,...,\xi_M),...,\xi^{(K)}=(\xi_1,...,\xi_M)\}$ of the standard normal random vector ξ. These realizations are called experimental design (ED) and can be obtained from Monte Carlo (MC) simulations or any other sampling scheme (e.g. Latin Hypercube (LH) sampling or Sobol set). We note $\Gamma=\left\{\Gamma\left(\xi^{(1)}\right),...,\Gamma\left(\xi^{(K)}\right)\right\}$, the corresponding values of the response determined by deterministic calculations.

The computation of the SPCE coefficients using the regression approach is performed using the following equation:

$$\hat{a}=(\eta^T\eta)^{-1}\eta^T\Gamma \tag{4}$$

where the data matrix η is defined by:

$$\eta_{i\beta}=\Psi_\beta(\xi^{(i)}), \qquad i=1,...,K, \qquad \beta=0,...,P-1 \tag{5}$$

In order to ensure the numerical stability of the treated problem in Eq.(4), the size K of the ED must be selected in such a way that the matrix $(\eta^T\eta)^{-1}$ is well-conditioned. This implies that the rank of this matrix should be larger than or equal to the number of unknown coefficients. This test was systematically performed while solving the system of equations of the regression approach.

The quality of the output approximation via a SPCE closely depends on the SPCE order p. To ensure a good fit between the meta-model and the true deterministic model (i.e. to obtain the optimal SPCE order), the simplest error estimate is the well-known coefficient of determination R^2 given by:

$$R^2=1-\frac{\frac{1}{K}\sum_{i=1}^{K}\left[\Gamma\left(\xi^{(i)}\right)-\Gamma_{SPCE}\left(\xi^{(i)}\right)\right]^2}{\frac{1}{K-1}\sum_{i=1}^{K}\left[\Gamma\left(\xi^{(i)}\right)-\bar{\Gamma}\right]^2}, \tag{6}$$

where

$$\bar{\Gamma}=\frac{1}{K}\sum_{i=1}^{K}\Gamma\left(\xi^{(i)}\right) \tag{7}$$

The value $R^2 = 1$ indicates a perfect fit of the true model response Γ, whereas $R^2 = 0$ indicates a nonlinear relationship between the true model Γ and the SPCE model Γ_{SPCE}. The coefficient R^2 may be a biased estimate since it does not take into account the robustness of the meta-model (i.e. its capability of correctly predicting the model response at any point which does not belong to the experimental design). As a consequence, one makes use of a more reliable and rigorous error estimate, namely the leave-one-out error estimate (Blatman and Sudret, 2010). This error estimate consists in sequentially removing a point from the experiment design composed of K points. Let $\Gamma_{\xi \backslash i}$ be the meta-model that has been built from the experiment design after removing the i^{th} observation and let $\Delta^i = \Gamma(\xi^{(i)}) - \Gamma_{\xi \backslash i}(\xi^{(i)})$ be the predicted residual between the model evaluation at point $\xi^{(i)}$ and its prediction based on $\Gamma_{\xi \backslash i}$. The corresponding coefficient of determination is often denoted by Q^2:

$$Q^2 = 1 - \frac{\frac{1}{K}\sum_{i=1}^{K}\left(\Delta^i\right)^2}{\frac{1}{K-1}\sum_{i=1}^{K}\left[\Gamma\left(\xi^{(i)}\right) - \overline{\Gamma}\right]^2} \tag{8}$$

This coefficient will be used in the present paper to check the accuracy of the fit.

3. Global sensitivity analysis (GSA)

Once the SPCE coefficients are determined, a global sensitivity analysis (GSA) based on Sobol indices can be easily performed. Notice that the first order Sobol index of a given random variable ξ_i (i = 1,…, M) gives the contribution of this variable in the variability of the system response. The first order Sobol index is given by Saltelli (2000) and Sobol (2001) as follows:

$$S(\xi_i) = \frac{Var\left[E\left(Y \mid \xi_i\right)\right]}{Var\left(Y\right)} \tag{9}$$

where Y is the system response, $E\left(Y \mid \xi_i\right)$ is the expectation of Y conditional on a fixed value of ξ_i, and *Var* denotes the variance.

In the present paper, the system response Y is represented by a SPCE. Thus, by replacing Y in Eq. (9) with the SPCE expression, one obtains the Sobol index formula as a function of the different terms of the SPCE (Sudret, 2008). This formula is given by:

$$S(\xi_i) = \frac{\sum_{\beta \in I_i}\left(a_\beta\right)^2 E\left[\left(\Psi_\beta\right)^2\right]}{D_{PC}} \tag{10}$$

where a_β are the obtained SPCE coefficients, Ψ_β are the multivariate Hermite polynomials, $E[.]$ is the expectation operator, and D_{PC} is the variance of the response approximated by the SPCE. The response variance D_{PC} is given by Sudret (2008) as follows:

$$D_{PC} = \sum_{\beta \in [I_1, I_2, ..., I_M]} \left(a_\beta\right)^2 E\left[\left(\Psi_\beta\right)^2\right] \tag{11}$$

Notice that the term $E\left[\left(\Psi_\beta\right)^2\right]$ that appears in both Eq. (10) and Eq. (11) is given by Sudret (2008) as follows:

$$E\left(\Psi_\beta^2\right) = \prod_{i=1}^{M} \alpha_i ! \tag{12}$$

where the α_i are the same sequence of M non-negative integers $\{\alpha_1, ..., \alpha_M\}$ used in Eq. (2). Notice finally that I_i in Eq. (10) denotes the set of indices β for which the corresponding Ψ_β is only a function of the random variable ξ_i (i.e. it only contains the variable ξ_i), and I_i (i = 1, …, M) regroup all the indices β for which the corresponding Ψ_β is only a function of the random variable ξ_i (i = 1, …, M).

In order to illustrate the PCE theory and the global sensitivity analysis based on Sobol indices in a simple manner, an illustrative example of a PCE of order p = 3 using only M = 2 random variables (ξ_1 and ξ_2) is presented in Appendix A.

4. Efficient combination between the SPCE methodology and the global sensitivity analysis

As mentioned previously, the time cost of the probabilistic analysis remains important even with the use of the SPCE when dealing with computationally-expensive deterministic models. Consequently, a procedure that can reduce once again this time cost is needed. An efficient combination between the SPCE methodology and the GSA is proposed in this section. The basic idea of this combination is that, for a given discretized random field, the obtained random variables do not have the same weight in the variability of the system response. The variables with a very small contribution in the variability of the system response can be discarded which significantly reduces the dimensionality of the treated problem. This allows one to perform a probabilistic analysis using a reduced Experiment Design (ED) and thus a smaller number of calls of the computationally-expensive deterministic model. The main challenge remains in detecting the most influential random variables in order to reduce the dimensionality of the problem. For this purpose, a procedure that makes use of both the SPCE and the GSA (denoted hereafter by SPCE/GSA) is proposed in this regard. The SPCE/GSA procedure can be summarized by the following steps:

a) Discretize the random field(s): This step was made in this paper using EOLE method and its extensions by Vořechovsky (2008). Let us consider N_{RF} anisotropic non-Gaussian cross-correlated random fields $Z_i^{NG}(x, y)$ ($i = 1, ..., N_{RF}$) described by: (i) constant means and standard deviations (μ_i, σ_i; $i = 1, ..., N_{RF}$), (ii) non-Gaussian marginal cumulative distribution functions CDFs named G_i ($i = 1, ..., N_{RF}$), (iii) a target cross-correlation matrix C^{NG} and (iv) a common square exponential

autocorrelation function $\rho_Z^{NG}[(x,\ y),\ (x',\ y')]$ which gives the values of the correlation function between two arbitrary points (x, y) and (x', y'). This autocorrelation function is given as follows:

$$\rho_z^{NG}[(x,\ y),\ (x',\ y')] = \exp\left(-\left(\frac{x-x'}{a_x} \right)^2 - \left(\frac{y-y'}{a_y} \right)^2 \right) \tag{13}$$

where a_x and a_y are the autocorrelation distances along x and y respectively. The Expansion Optimal Linear Estimation method (EOLE) and its extension by Vořechovsky (2008) to cover the case of correlated non-Gaussian random fields are used herein to generate the N_{RF} random fields. Notice that EOLE was first proposed by Li and Der Kiureghian (1993) for the case of uncorrelated Gaussian fields, and then extended by Vořechovsky (2008) to cover the case of correlated non-Gaussian fields. In this method, one should first define a stochastic grid composed of q grid points (or nodes) $\left\{ (x_1,\ y_1),\ ...,\ (x_q,\ y_q) \right\}$ for which the values of the field are assembled in a vector $\chi = \left\{ Z(x_1,\ y_1),\ ...,\ Z(x_q,\ y_q) \right\}$. Then, one should determine the common correlation matrix for which each element $\left(\Sigma_{\chi;\chi} \right)_{i,\ j}^{NG}$ is calculated as follows:

$$\left(\Sigma_{\chi;\chi} \right)_{i,\ j}^{NG} = \rho_Z^{NG} \left[(x_i,\ y_i),\ (x_j,\ y_j) \right] \tag{14}$$

The common non-Gaussian autocorrelation matrix $\Sigma_{\chi;\chi}^{NG}$ and the target non-Gaussian cross-correlation matrix C^{NG} should be transformed into the Gaussian space using Nataf model (Nataf, 1962) since the discretization of the random fields using EOLE is done in the Gaussian space. As a result, one obtains N_{RF} Gaussian autocorrelation matrices $\Sigma_{\chi;\chi}^{i}\ (i = 1, ..., N_{RF})$, and a Gaussian cross-correlation matrix C that can be used to discretize the two random fields. The value $\tilde{Z}_i$ of a random field obtained using this method is given by the following equation (cf. Al-Bittar and Soubra, 2011, 2012):

$$\tilde{Z}_i(x,\ y) \cong \mu_i + \sigma_i \sum_{j=1}^{N} \frac{\kappa_{i,\ j}^{D}}{\sqrt{\lambda_j^i}} \cdot \left(\phi_j^i \right)^T \cdot \Sigma_{Z(x,\ y);\chi}^{i} \qquad i = 1,\ ...,\ N_{RF} \tag{15}$$

where N_{RF} is the number of random fields, N is the number of terms retained in the series expansion, $\kappa_{i,\ j}^{D}$ are N_{RF} cross-correlated blocks of independent standard normal random variables obtained using the Gaussian cross-correlation matrix C between the N_{RF} fields, $(\lambda_j^i, \phi_j^i; i = 1, ..., N_{RF})$ are the eigenvalues and eigenvectors of the N_{RF} Gaussian autocorrelation matrices $\Sigma_{\chi;\chi}^{i}$ evaluated at the different points of the stochastic mesh, and $\Sigma_{Z(x,y);\chi}$ is the correlation vector between the value of the

field $\tilde{Z}_i$ at an arbitrary point (x, y) and its values at the different points of the stochastic mesh. Notice finally that $\kappa_{i,j}^{D}$, ϕ_j^i, and $\Sigma_{Z(x,y);\chi}$ in Eq.(15) are vectors whose size is equal to N.

Once the two Gaussian random fields are obtained, they should be transformed into the non-Gaussian space (in case of non-Gaussian random fields) by applying the following formula:

$$\tilde{Z}_i^{NG}(x,\, y) = G_i^{-1}\left\{\Phi\left[\tilde{Z}_i(x,\, y)\right]\right\} \qquad i = 1, ..., N_{RF} \tag{16}$$

where $\Phi(.)$ is the standard normal cumulative density function (CDF). For more details about the EOLE method and its extensions to cover the case of cross-correlated non-Gaussian random fields, the reader may refer to Vořechovsky (2008) and Al-Bittar and Soubra (2011, 2012).

After the discretization procedure, a random field is represented by N independent standard normal random variables. For the N_{RF} random fields that have the same autocorrelation function, the total number of random variables is $N_T = N_{RF} \cdot N$ which can be relatively large especially for small values of the autocorrelation distances.

b) Use a preliminary small order of the sparse polynomial chaos expansion (e.g. $p = 2$) to approximate the system response by a meta-model. The main reason for selecting a small order is the exploration of the most influential random variables (i.e. those that have a significant weight in the variability of the system response) using a small Experiment Design (ED). It should be emphasized here that the reduced number of the unknown SPCE coefficients related to the small value of the SPCE order leads to a significant decrease in the size of the experiment design, i.e. in the number of calls of the deterministic model.

c) Perform a GSA based on Sobol indices (using the obtained second order SPCE) to determine the weight of each random variable in the variability of the system response. The variables with very small values of their Sobol indices have no significant impact in the variability of the system response and can thus be discarded. Consequently, a response that only depends on a smaller number of random variables is obtained. In other words, one obtains a response with an 'effective dimension' N_e that is smaller than the initial dimension where the total number N_T of random variables was considered. It should be mentioned here that the small SPCE order (i.e. $p = 2$) used firstly to perform the GSA is sufficient to provide the weight of each random variable in the variability of the system response since higher SPCE orders lead to the same influential random variables as will be seen later in the numerical results.

d) Use the same Experiment Design (ED) which was employed in step (b) but this time by only keeping the most influential random variables. By reducing the number of random variables from N_T to N_e ($N_e < N_T$), one has the possibility to use a higher SPCE order (i.e. $p > 2$). The use of a higher SPCE order is necessary to lead to an improved fit of the SPCE since the leave-one-out error estimate Q^2 given in Eq. (8) increases when the SPCE order increases as it will be shown in the numerical results.

As a conclusion, the use of the SPCE/GSA procedure has the advantage of performing a good fit of the deterministic model with a reduced number of model evaluations as compared to the classical SPCE approach.

5. Numerical results

The aim of this section is to validate the present SPCE/GSA procedure. For this purpose, a comparison between the results obtained by using the classical SPCE methodology and those given by the proposed SPCE/GSA procedure is presented. A computationally non expensive deterministic model was used for the validation. The problem used for the validation was presented in Al-Bittar and Soubra (2011, 2012). It aims at computing the ultimate bearing capacity of a strip footing resting on a $c-\varphi$ spatially varying soil. The input parameters are similar to those considered in Al-Bittar and Soubra (2011, 2012). They are briefly presented in Table I. For a more detailed description on these data, the reader may refer to Al-Bittar and Soubra (2011, 2012). The deterministic model is based on numerical simulations using FLAC3D and it involves the case of a weightless soil. Thus, one obtains the soil bearing pressure due to only the soil cohesion; the contribution of the soil friction angle being neglected in the present paper. It should be mentioned here that when neglecting the soil weight γ, the computation time decreases from 10 to 5 min per simulation. This significantly reduces the computation time for the validation of the present SPCE/GSA procedure.

As shown in Figure 1, the adopted soil domain considered in the analysis is 15 m wide by 6 m deep. For the boundary conditions, the horizontal movement on the vertical boundaries of the grid is restrained, while the base of the grid is not allowed to move in both the horizontal and the vertical directions.

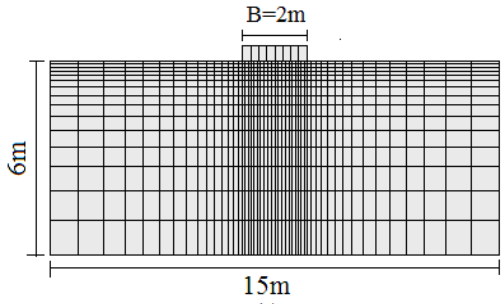


Figure 1. The adopted soil domain

The validation of the SPCE/GSA procedure is done for the illustrative case [a_x = 10 m, a_y = 1 m, $r(c, \varphi) = -0.5$] referred to hereafter as the reference case. For this configuration, the discretization of the two random fields c and φ has led to a total number of random variables N_T equal to 24 (12 random variables for each random field as was shown in Al-Bittar and Soubra (2011, 2012)). By using the total number of random variables N_T, Al-Bittar and Soubra (2011, 2012) have shown that a third order SPCE was sufficient to reach a target accuracy of 0.999. An ED involving 800 points was needed to solve the regression problem given in Eq. (4) (i.e. to obtain a well-conditioned regression problem for which the rank of the matrix $(\eta^T\eta)^{-1}$ is larger than or equal to the number of unknown coefficients). By using the present SPCE/GSA procedure, a GSA was performed to detect the most influential random variables. Different SPCE orders (i.e. orders 2, 3, and 4) were considered in order to check if the SPCE order has an impact on the most influential random variables.

Table I. Deterministic and probabilistic input parameters

Model	Parameters	Parameters type	Deterministic parameters values	Stochastic parameters		
				Statistical characteristics	Coefficient of correlation r	Autocorrelation function $\rho_z^{NG}[(x, y), (x', y')]$
Soil: Elastic-perfectly plastic (Mohr-Coulomb)	Young modulus E	deterministic	$E = 60\text{MPa}$	-	-	-
	Poisson ratio ν	deterministic	$\nu = 0.3$	-	-	-
	Cohesion c	Random field	-	Lognormal $\mu_c = 20\text{kPa}$, $Cov_c = 25\%$	Reference case: $C^{NG} = \begin{bmatrix} r(c,c)=0 & r(c,\varphi)=-0.5 \\ r(\varphi,c)=-0.5 & r(\varphi,\varphi)=0 \end{bmatrix}$ Parametric study: $-0.5 \le r(c,\varphi) \le 0$	Square exponential $\rho_Z^{NG} = \exp\left(-\left(\frac{x-x'}{a_x}\right)^2 - \left(\frac{y-y'}{a_y}\right)^2\right)$ a_x and a_y are the autocorrelation distances along x and y respectively. Reference case: a_x = 10m and a_y = 1m. Parametric study: $2\text{m} \le a_x \le 50\text{m}$; $0.5\text{m} \le a_y \le 8\text{m}$
	Friction angle φ	Random field	-	Beta $\mu_\varphi = 30^0$, $Cov_\varphi = 10\%$		
	Dilation ψ	Random field	-	Beta $\mu_\psi = 20^0$, $Cov_\psi = 10\%$		
Foundation: Elastic	Young modulus E	deterministic	$E = 25\text{GPa}$	-	-	-
	Poisson ratio ν	deterministic	$\nu = 0.4$	-	-	-
Interface Elastic-perfectly plastic (Mohr-Coulomb)	Shear stiffness K_s	deterministic	$K_s = 1\text{GPa}$	-	-	-
	Normal stiffness K_n	deterministic	$K_n = 1\text{GPa}$	-	-	-
	Cohesion c_{int}	deterministic	$c_{\text{int}} = 20\text{kPa}$	-	-	-
	Friction angle φ_{int}	deterministic	$\varphi_{\text{int}} = 30^0$	-	-	-
	Dilation ψ_{int}	deterministic	$\psi_{\text{int}} = 20^0$	-	-	-

Figure 2 depicts the values of Sobol indices for the 24 random variables, as given by SPCEs of orders 2, 3 and 4. The first 12 random variables [i.e. ξ_i for i = 1, …, 12] correspond to the cohesion random field and the last 12 random variables [i.e. ξ_i for i = 13, …, 24] are those corresponding to the friction angle random field. Figure 2 shows that whatever the SPCE order is, the two first random variables of both fields, (i.e. ξ_1, ξ_2, ξ_{13}, ξ_{14}) are the most influential. For the two random fields, a very fast decay in the weight of the random variables is noticed with quasi negligible values beyond the first two random variables. In fact, the first two random variables of the two random fields, which correspond to the first two eigenmodes of both fields involve 95% of the response variability as may be seen from Table II. This is logical since the system response (i.e. the ultimate bearing capacity) is an averaged quantity over the soil domain which is therefore quite insensitive to small-scale fluctuations of the spatially varying shear strength parameters c and φ.

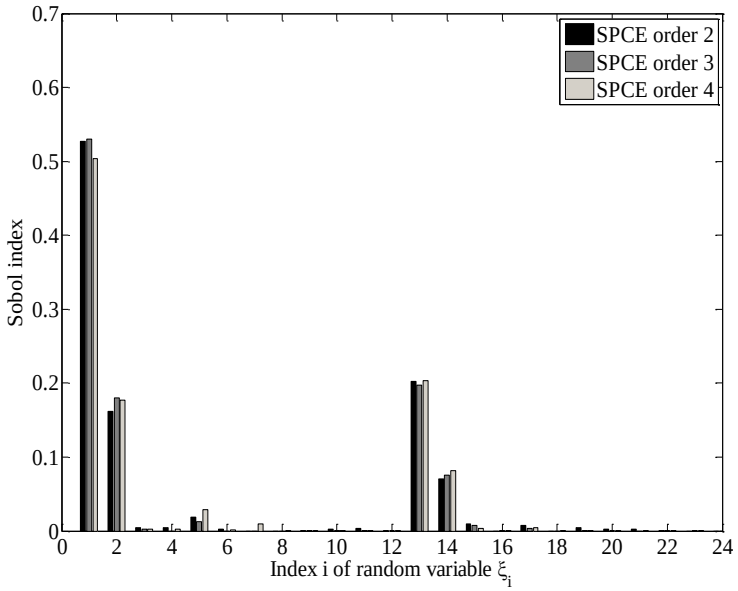


Figure 2. Sobol indices for SPCEs of orders 2, 3 and 4 using the total number of eigenmodes ξ_i (i = 1, …, 24)

Table II. Sobol indices for the reference case where a_x = 10 m, a_y = 1 m, and $r(c,\varphi) = -0.5$

	ξ_i (i = 1, ..., 12) for the cohesion random field											
	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8	ξ_9	ξ_{10}	ξ_{11}	ξ_{12}
Sobol index	0.50	0.17	0.002	0.002	0.03	0.002	0.009	0.0002	0.0002	9 x10^{-05}	0.0002	7 x10^{-05}
	ξ_i (i = 13, ..., 24) for the friction angle random field											
	ξ_{13}	ξ_{14}	ξ_{15}	ξ_{16}	ξ_{17}	ξ_{18}	ξ_{19}	ξ_{20}	ξ_{21}	ξ_{22}	ξ_{23}	ξ_{24}
Sobol index	0.2	0.08	0.001	0.0008	0.002	0.0005	0.0006	0.0003	0.0001	4 x10^{-05}	4 x10^{-05}	5 x10^{-05}

Figure 2 clearly shows that the Sobol indices of the different random variables do not significantly change with the SPCE order. Thus, a second order SPCE is sufficient to identify the influential random variables (i.e. those that have a significant weight in the variability of the ultimate bearing capacity). Increasing the SPCE order has led to the same influential random variables which justify the small SPCE order chosen to perform the preliminary investigations. The main advantage of a small SPCE order is that a small ED is sufficient to solve the regression problem. As shown in Table III, 150 calls of the deterministic model are needed to solve the regression problem for a second order SPCE. This number attains 800 for a fourth order SPCE. This significant increase is because the number of unknown coefficients significantly increases from 29 to 144 when one chooses a fourth SPCE order instead of a second SPCE order.

Table III. Number of unknown coefficients and model evaluations for different SPCE order

SPCE order	2	3	4
Number of unknown coefficients P	29	35	144
Number of model evaluations	150	350	800

To choose the number of random variables which will be retained hereafter, the different random variables of the two random fields are firstly sorted in a descending order according to the values of their Sobol indices (cf. first and second columns in Table IV). A threshold of acceptance t_a is then fixed as a percentage of the most influential (weighted) random variable. In the present paper, the most influential random variable is ξ_1 and it has a Sobol index equal to 0.5. Different values of the threshold were tested (cf. first line in Table IV). The random variables having a Sobol index smaller than the prescribed threshold t_a are discarded. In this paper, a threshold of 2% of the Sobol index of the most weighed random variable is considered as sufficient; the corresponding retained random variables provide 98% of the total variance of the system response as may be seen from column 6 of Table IV. For this threshold, an 'effective dimension' $N_e = 5$ is obtained (i.e. 5 random variables are considered to be the most weighed). The 5 retained random variables will now be used with the already existing 150 model evaluations which were firstly employed to approximate the second order SPCE with the total number of random variables $N_T = 24$.

The reduction in the number of random variables from $N_T = 24$ to $N_e = 5$ provides the possibility to use higher SPCE orders (i.e. $p > 2$) with the same ED (i.e. the 150 model evaluations). The use of a higher SPCE order is necessary to lead to an improved fit of the SPCE since the leave-one-out error estimate Q^2 given in Eq. (8) increases when the SPCE order increases as shown in Table V for both the classical SPCE approach (using the total number of random variables $N_T = 24$) and the present SPCE/GSA procedure (where the effective dimension is equal to 5 (i.e. $N_e = 5$)). Using the SPCE/GSA procedure, an SPCE up to $p = 8$ was reached using only 150 model evaluations. From Table V, one can notice that with the use of the SPCE/GSA procedure, the Q^2 increases with the increase of the SPCE order and stabilizes beyond the order 5. This means that no improvement in the fit is obtained beyond this order. On the other hand, the value of Q^2 given by the present approach is smaller than the classical SPCE approach with a fourth order. This is because 19 random variables were discarded which slightly affect the goodness of the fit.

Table IV. Sobol indices of the different random variables and the retained random variables for the different values of the threshold of acceptance

Random variable	Sobol index	t_a = 0.5% x ξ_1 = 0.0025	t_a = 1% x ξ_1 = 0.005	t_a = 1.5% x ξ_1 = 0.0075	t_a = 2% x ξ_1 = 0.01	t_a = 2.5% x ξ_1 = 0.0125	t_a = 3% x ξ_1 = 0.015	t_a = 4% x ξ_1 = 0.02	t_a = 5% x ξ_1 = 0.025
ξ_1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
ξ_{13}	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
ξ_2	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
ξ_{14}	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
ξ_5	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
ξ_7	0.009	0.009	0.009	0.009					
ξ_6	0.002	0.002							
ξ_{17}	0.002	0.002							
ξ_3	0.002	0.002							
ξ_4	0.002	0.002							
ξ_{15}	0.001								
ξ_{16}	0.0008								
ξ_{19}	0.0006								
ξ_{18}	0.0005								
ξ_{20}	0.0003								
ξ_8	0.0002								
ξ_9	0.0002								
ξ_{11}	0.0002								
ξ_{21}	0.0001								
ξ_{10}	9.0 x10^{-05}								
ξ_{12}	7.0 x10^{-05}								
ξ_{24}	5.0 x10^{-05}								
ξ_{22}	4.0 x10^{-05}								
ξ_{23}	4.0 x10^{-05}								
Sum of Sobol indices	1.001	0.997	0.989	0.989	0.98	0.98	0.98	0.98	0.98

Table V. SPCE using the total and the reduced number of random variables

	SPCE order	2	3	4	5	6	7	8
Total number of random variables N_T	Coefficient of determination R^2	0.998	0.999	0.999	-	-	-	-
	Leave-one-out cross-validation Q^2	0.824	0.932	0.9943	-	-	-	-
Reduced number of random variables N_e	Coefficient of determination R^2	0.961	0.963	0.968	0.970	0.972	0.972	0.972
	Leave-one-out cross-validation Q^2	0.791	0.883	0.957	0.961	0.963	0.963	0.963

Figure 3 shows the PDF of the ultimate bearing capacity as obtained by both the classical SPCE approach (with the total number of random variables $N_T = 24$) and the proposed SPCE/GSA procedure (using only five random variables). Table VI provides the corresponding statistical moments and error estimates. Notice that the results of the present SPCE/GSA approach are given in Table VI for different values of the model evaluations (from 150 to 800). From this table, one can see that the error estimate of the SPCE/GSA procedure is quasi constant with the increase in the number of model evaluations. This means that 150 model evaluations are sufficient and there is no need for more model evaluations to improve the accuracy of the fit. On the other hand, one can observe (see Figure 3 and Table VI) that the first two statistical moments are well estimated with the present SPCE/GSA approach using the 150 model evaluations. However, the third and fourth statistical moments need more model evaluations (800 model evaluations) in order to converge to their reference values given by the SPCE approach (cf. Table VI). This demonstrates the efficiency of the present SPCE/GSA procedure to compute the first two statistical moments with a much reduced number of the model evaluations with respect to the classical SPCE approach.

As for the Sobol indices of the two random fields c and φ, Table VII shows that the SPCE/GSA procedure with only 150 model evaluations gives the same results obtained by the classical SPCE approach using 800 model evaluations which demonstrates once again the efficiency of the present SPCE/GSA procedure.

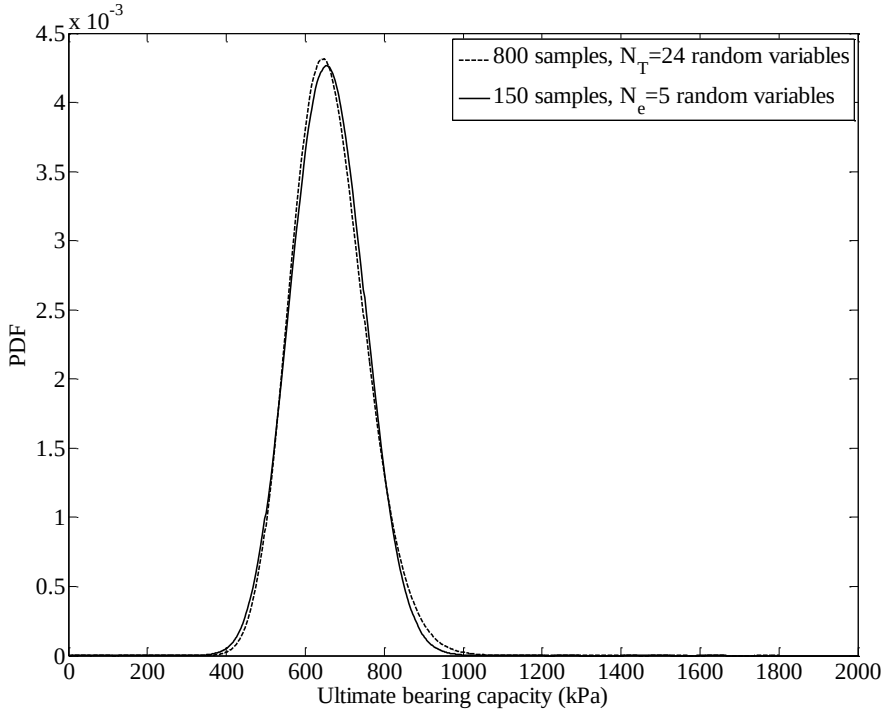


Figure 3. PDF of the ultimate bearing capacity for both the classical SPCE with the total number of random variables $N_T = 24$ and the proposed SPCE/GSA procedure with only five random variables $N_e = 5$.

Table VI. Error estimates of the SPCE and statistical moments of the ultimate bearing capacity as given by the classical SPCE approach and by the present SPCE/GSA procedure

	Number of model evaluations	Mean μ_{qult} (kPa)	Standard deviation σ_{qult} (kPa)	Skewness δ_u (-)	Kurtosis κ_u (-)	R^2	Q^2
With the total number of random variables $N_T = 24$	800	658.2	93.57	0.287	0.163	0.999	0.995
With the reduced number of random variables $N_e = 5$	150	657.84	90.80	0.105	0.0129	0.972	0.957
	200	658.98	91.53	0.168	0.0563	0.972	0.951
	250	659.90	92.10	0.188	0.0630	0.964	0.956
	300	659.73	92.15	0.202	0.0600	0.962	0.963
	400	660.05	90.95	0.291	0.0500	0.969	0.960
	500	659.50	90.81	0.296	0.0430	0.970	0.963
	600	659.75	90.99	0.272	0.116	0.968	0.963
	700	659.50	90.85	0.280	0.1637	0.968	0.963
	800	659.85	91.20	0.30	0.160	0.970	0.967

Table VII. Sobol indices as computed from the classical SPCE approach and the present SPCE/GSA procedure.

	Number of model evaluations	i	S_i (i = 1, ..., 12)	i	S_i (i = 13, ..., 24)	$S(c) = \sum_{i=1}^{12} S_i$	$S(\varphi) = \sum_{i=13}^{24} S_i$
With the total number of random variables $N_T = 24$	800	1	0.5	13	0.2	0.715	0.285
		2	0.17	14	0.08		
		3	0.002	15	0.001		
		4	0.002	16	0.0008		
		5	0.03	17	0.002		
		6	0.002	18	0.0005		
		7	0.009	19	0.0006		
		8	0.0002	20	0.0003		
		9	0.0002	21	0.0001		
		10	9.0×10^{-05}	22	4.0×10^{-05}		
		11	0.0002	23	4.0×10^{-05}		
		12	7.0×10^{-05}	24	5.0×10^{-05}		
	Number of model evaluations	i	S_i (i = 1, 2, 3)	i	S_i (i = 4, 5)	$S(c) = \sum_{i=1}^{3} S_i$	$S(\varphi) = \sum_{i=4}^{5} S_i$
With the reduced number of random variables $N_e = 5$	150	1	0.510	4	0.076	0.721	0.279
		2	0.200	5	0.190		
		3	0.010				

6. Conclusions

An efficient combined use of the SPCE methodology and the global sensitivity analysis (GSA) has been proposed. The aim is to reduce the cost of the probabilistic analysis of computationally-expensive deterministic models. This methodology was validated in this paper using a relatively non-expensive deterministic model. The validation consists in comparing the results of both the classical SPCE methodology with the total number of random variables and the proposed combination between the SPCE and the GSA. Satisfactory results were obtained using a much smaller number of model evaluations with the proposed methodology. The first two statistical moments and the Sobol indices have been well estimated with the very small number of model evaluations. On the other hand, the third and fourth statistical moments need more model evaluations in order to converge to their reference values obtained using the classical SPCE. Since the present SPCE/GSA procedure was shown to be efficient for the probabilistic computation with a reduced calculation cost with respect to the classical SPCE approach, this approach may now be applied with confidence to costly deterministic models.

Appendix A

Illustrative Example

In order to illustrate the PCE theory in a simple manner, a PCE of order $p = 3$ using only $M = 2$ random variables (ξ_1 and ξ_2) will be considered in this illustrative example. Using the classical truncation scheme, Table A.1 presents the retained PCE terms which are those having a first order norm $\|\alpha\|_1$ smaller than or equal to p (i.e. $p = 3$). These terms are presented in Table A.1 in bold characters. As may be easily seen from Table A.1, the PCE basis contains $P = 10$ terms whose expressions are computed using Eq.(2).

Table A.1. Terms retained using the classical truncation scheme for $M = 2$ and $p = 3$

α_1	0	1	0	1	2	0	2	1	2	3	0	3	1	3	3	3
α_2	0	0	1	1	0	2	1	2	2	0	3	1	3	2	3	3
$\|\alpha\|_1$	**0**	**1**	**1**	**2**	**2**	**2**	**3**	**3**	4	**3**	**3**	4	4	5	6	6

Table A.2 presents the expressions of the PCE basis Ψ_β. Using Table A.2, on can write the PCE expression as function of the input random variables (ξ_1 and ξ_2) as follows:

$$Y = \Gamma_{PCE}(\xi) = a_0\Psi_0 + a_1\Psi_1 + ... + a_9\Psi_9 = \\ a_0 + a_1\xi_1 + a_2\xi_2 + a_3\xi_1\xi_2 + a_4(\xi_1^2 - 1) + a_5(\xi_2^2 - 1) + a_6\left(\xi_1^2 - 1\right)\xi_2 + a_7\xi_1\left(\xi_2^2 - 1\right) + a_8(\xi_1^3 - 3\xi_1) + a_9(\xi_2^3 - 3\xi_2) \quad (A.1)$$

In this expression, the unknown coefficients can be computed using Eq.(4) by simulating an ED which contains K initial realizations of the two random variables (ξ_1, ξ_2) and computing the corresponding responses from deterministic calculations. It should be mentioned here that the size K of the ED should ensure the numerical stability of the regression problem and thus it can be enriched each time the matrix $(\eta^T\eta)^{-1}$ is badly-conditioned.

The first order Sobol indices for the two random variable (ξ_1 and ξ_2) can be easily obtained once the coefficients $a_0,\ldots, a_9$ are computed using Eq. (10). The only additional step is to compute $E\left(\Psi_\beta^2\right)$ corresponding to these two random variables. Table A.2 shows the values of $E\left(\Psi_\beta^2\right)$ computed using Eq. (12) for the different Ψ_β terms. The expression of the first order Sobol indices of the two random variables ξ_1 and ξ_2 are written as follows:

$$S(\xi_1)=\frac{a_1^2+2a_4^2+6a_8^2}{a_1^2+2a_4^2+6a_8^2+a_2^2+2a_5^2+6a_9^2}; \qquad S(\xi_2)=\frac{a_2^2+2a_5^2+6a_9^2}{a_1^2+2a_4^2+6a_8^2+a_2^2+2a_5^2+6a_9^2} \tag{A.2}$$

with

$$I_1=(1,4,8); \qquad I_2=(2,5,9). \tag{A.3}$$

Table A.2. Basis of the PCE with the classical truncation scheme for M = 2 and p = 3

β	PCE order p	$\Psi_\beta=\prod_{i=1}^{M}H_{\alpha_i}(\xi_i)$	$E\left(\Psi_\beta^2\right)=\prod_{i=1}^{M}\alpha_i!$
0	P = 0	$H_0(\xi_1)\times H_0(\xi_2)=1$	$\alpha_1!\times\alpha_2!=0!\times0!=1$
1	P = 1	$H_1(\xi_1)\times H_0(\xi_2)=\xi_1$	$\alpha_1!\times\alpha_2!=1!\times0!=1$
2		$H_0(\xi_1)\times H_1(\xi_2)=\xi_2$	$\alpha_1!\times\alpha_2!=0!\times1!=1$
3		$H_1(\xi_1)\times H_1(\xi_2)=\xi_1\,\xi_2$	$\alpha_1!\times\alpha_2!=1!\times1!=1$
4	P = 2	$H_2(\xi_1)\times H_0(\xi_2)=\xi_1^2-1$	$\alpha_1!\times\alpha_2!=2!\times0!=2$
5		$H_0(\xi_1)\times H_2(\xi_2)=\xi_2^2-1$	$\alpha_1!\times\alpha_2!=0!\times2!=2$
6		$H_2(\xi_1)\times H_1(\xi_2)=\left(\xi_1^2-1\right)\xi_2$	$\alpha_1!\times\alpha_2!=2!\times1!=2$
7	P = 3	$H_1(\xi_1)\times H_2(\xi_2)=\xi_1\left(\xi_2^2-1\right)$	$\alpha_1!\times\alpha_2!=1!\times2!=2$
8		$H_3(\xi_1)\times H_0(\xi_2)=\xi_1^3-3\xi_1$	$\alpha_1!\times\alpha_2!=3!\times0!=6$
9		$H_0(\xi_1)\times H_3(\xi_2)=\xi_2^3-3\xi_2$	$\alpha_1!\times\alpha_2!=0!\times3!=6$

References

Al-Bittar, T. and Soubra, A. H. Bearing capacity of strip footing on spatially random soils using sparse polynomial chaos expansion. *International Journal for Numerical and Analytical Methods in Geomechanics*, accepted paper, 2012.

Al-Bittar, T. and Soubra, A. H. Bearing capacity of strip footing on spatially random soils using sparse polynomial chaos expansion. *GeoRisk* (GSP 224), ASCE, Atlanta, USA, 26-28 June 2011.

Blatman, G. and Sudret, B. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics*, 25: 183-197, 2010.

Huang, S. P., Liang, B. and Phoon K. K. Geotechnical probabilistic analysis by collocation-based stochastic response surface method: An Excel Add-in implementation. *Georisk*, 3: 75-86, 2009.

Isukapalli, S. S., Roy, A. and Georgopoulos, PG. Stochastic response surface methods (SRSMs) for uncertainty propagation: Application to environmental and biological systems. *Risk Analysis*, 18: 357-363, 1998.

Mollon, G., Dias, D. and Soubra, A. H. Probabilistic analysis of pressurized tunnels against face stability using collocation-based stochastic response surface method. *Journal of Geotechnical and Geoenvironmental Engineering ASCE*, 137: 385-397, 2011.

Mao, N., Al-Bittar, T. and Soubra, A. H. Probabilistic analysis and design of strip foundations resting on rocks obeying Hoek-Brown failure criterion. *International Journal of Rock Mechanics and Mining Sciences*, 49: 45-58, 2012.

Nataf, A. Détermination des distributions de probabilités dont les marges sont données. *CR Acad Sci*, 225: 42-3, 1962.

Saltelli, A., Chan, K. and Scott E. M. *Sensitivity Analysis*. John Wiley and sons, 2000.

Spanos, P. D. and Ghanem, R. Stochastic finite element expansion for random media. *Journal of Engineering Mechanics*, 115: 1035-1053, 1989.

Sobol, I. M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mechanics and Computation in Simulation*, 55: 271-280, 2001.

Sudret, B. Global sensitivity analysis using polynomial chaos expansion. *Reliability Engineering and System Safety*, 93: 964-979, 2008.

Vořechovský, M. Simulation of simply cross-correlated random fields by series expansion methods. *Structural Safety*, 30: 337-363, 2008.

Xiu, D. and Karniadakis, G. E. The Wiener-Askey polynomial chaos for stochastic differential equations. *Journal of Scientific Computing*, 24: 619-44, 2002.

Evaluation of Recycling Concrete

Naser Alenezi
Research Associate, Kuwait Institute for Scientific Research; P.O.Box 24885 Safat, 13109; Kuwait
PH (965) 24989282; FAX (965) 24989099; email: naserq8@hotmail.com

Abstract: Recycling is becoming essential to reduce the amount of solid waste, which is becoming a global environmental crisis, in addition to conserving the world recourses, which are becoming scarce, to the future generations. Demolish concrete may be utilized for making a new concrete. The durability of concrete depends on heavily of aggregates. The properties of recycled concrete are different than those of new aggregates because of its exposed to many environments. The concrete mix design will be affected to its strength because of the properties of aggregates. Properties of recycled aggregates were comparing to the new aggregates.

The consistency was evaluated for three batches of recycled aggregates which are collected from different structure with different types of concrete. Their properties expected to differ from each other. The aim of this study is to investigate the possibility of producing good concrete from recycled aggregate. Concrete samples were produced from mixed that used recycled aggregates. The properties of concrete were evaluated and compare to the control mix which used a new virgin aggregates.

Keywords: aggregates, recycling, concrete, demolish, durability.

1. Introduction

Waste concrete may be collected and utilized in the production of new concrete. Properties of aggregates have a great impact on the durability and strength of the resulting concrete. As recycled aggregates contain concrete rebels and have been exposed to different environments for different durations, their properties are expected to differ compared to those of virgin new aggregates. Properties of aggregates affect the mix design and the proportions of the constituent materials. The suitability of the recycled aggregates for concrete production was assessed through standard aggregates tests. Properties were compared to of new aggregates.

As recycled aggregates are collected from different structures with different types of concrete with different types and durations of exposure, their properties are expected to differ between batches. Most of the properties were determined for three different batches of recycled concrete and a preliminary statistical evaluation was performed to investigate consistency.

2. Methodology

A plane of experiments and tests laboratories were carried out to complete the aims of this limited study, this included choosing recycled aggregate sizes to be evaluated and set for laboratory tests.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

3. Laboratory Evaluation

Laboratory estimation was exploited to judge the features of the aggregates used for concrete. The selections of recycled aggregate were taken for test such as: water absorption and density (ASTM C-127, C-128), abrasion and degradation resistance test. Fine and coarse aggregate samples were analyzed to determine the chloride and sulfate contents in the lab. The prepared concrete specimen which used a recycled aggregates were cured in the lab and tested for compressive strength (ASTM C39)

3.1. Water Absorption and Density Test

Water absorption and density were assessed according to ASTM C-127 and C-128. This test is of importance since it identifies the density of the essentially solid portion of a large number of aggregates particles and provides an average value of representing samples.

3.2. Resistance to Degradation by Abrasion and Impact Test

Aggregates were also tested for abrasion and degradation resistance according to ASTM C131 using the Los Angeles machine. This test is to identify the relative quality or competence of various sources of aggregate having similar mineral compositions.

3.3. Chloride and Sulphate Contents Tests

Chloride and sulphate contents play an important role in the rate and type of deterioration in concrete. They also affect the way concrete mix should be designed. Chloride and sulphate contents were evaluated for both the fresh and three batches of recycled crushed aggregates collected at different times from the crushing factory
This test is to identify percentage of water absorbed which indicates the quantity of voids in aggregate, and this would give the idea of aggregate capacity to keep hold of water and its potential to deterioration.

4. Tests Results

4.1. Laboratory Evaluation

4.1.1. *Water absorption and density test*

Water absorption and density were assessed according to ASTM C-127 and C-128 the results for the virgin and three sizes of recycled aggregates are shown in Table I. The average values, standard deviations and coefficients of variations for both the density and water absorption are given for the three sizes along with the overall aggregates in Table II. As can be seen the values of the standard deviations and coefficient of variations are very small and insignificant which means that small variations between samples were detected. The coefficient of variations between the three average values of densities and water absorption were calculated and found to be 1.33 and 15.06%, in order. The values indicate that no significant changes in densities were detected and the value of 2.48 may be considered constant for recycled aggregates. For water

absorption, however, although no changes for each size were detected, the values have some dependence on aggregates size. These results, however, should be verified by continuous testing or long term monitoring of recycled aggregates, as only three samples were considered in this work.

Table I. Density and water absorption of aggregates

Aggregate Type	Density (g/cm^3)	Water Absorption (%)
New Aggregate 3/4"	2.81	0.73
New Aggregate 1/2"	2.65	0.98
New Aggregate 3/8"	2.65	1.30
Crushed Aggregate 3/4", batch 1	2.64	4.97
Crushed Aggregate 3/4", batch 2	2.444	4.706
Crushed Aggregate 3/4", batch 3	2.468	4.69
Crushed Aggregate 1/2", batch 1	2.55	5.18
Crushed Aggregate 1/2", batch 2	2.437	5.334
Crushed Aggregate 1/2", batch 3	2.461	5.20
Crushed Aggregate 3/8", batch 1	2.52	7.18
Crushed Aggregate 3/8", batch 2	2.383	6.735
Crushed Aggregate 3/8", batch 3	2.408	6.71

4.1.2. *Abrasion and degradation resistance test*

Aggregates were also tested for abrasion and degradation resistance according to ASTM C131 using the Los Angeles machine. The Los Angeles values are shown for new and three recycled batches of aggregates, using grade B, in Table III. The resistance to crushing and abrasion using steel balls was also assessed using the 10% fines method according to BS 882, wet method. The values are also shown in Table III.

Table II. Statistical values for density and water absorption of aggregates

	Density			Water absorption		
	Average	S.D.	V	Average	S.D	V
¾ in aggregates	2.517	0.087	3.46	4.79	0.128	2.67
½ in aggregates	2.483	0.049	1.97	5.24	0.068	1.30
3/8 in aggregates	2.437	0.032	1.31	6.76	0.245	3.62
Overall	2.48	0.056	2.25	5.6	0.147	2.53

Table III. LA and 10% fines values of aggregates

	LA value	10% fines value
New aggregates	18	373.3
Recycled batch 1	33.92	150.49
Recycled batch 2	30.41	190.8
Recycled batch 3	33.3	150.64
Average for recycled	32.54	163.98
S.D.	1.40	18.97
V.	4.30	11.57

4.1.3. *Chloride and sulphate contents tests*

Chloride and sulphate contents play an important role in the rate and type of deterioration in concrete. They also affect the way concrete mix should be designed. Chloride and sulphate contents were evaluated for both the fresh and three batches of recycled crushed aggregates collected at different times from the crushing factory. The results along with the average values and standard deviations for the recycled aggregates are shown in Table IV.

In general, the changes in aggregates properties between different batches are tolerable. The properties of recycled aggregates vary considerably compared to those of virgin new aggregates. Summary of those properties and the difference percentages are shown in Table V. Changes of more than 100% may be seen in water absorption, impact, abrasion resistance, chloride content and sulphate content.

Table IV. Chloride and sulphate contents

Aggregate Type	Chloride Content (%)	Sulphate Content (%)
New Sand	0.003	0.17
Recycle Crushed Concrete batch 1	0.052	0.75
Recycle Crushed Concrete batch 2	0.054	0.52
Recycle Crushed Concrete batch 3	0.060	0.82
Average for recycled concrete	0.055	0.70
Standard deviation for recycled concrete	.0034	0.128
Coeff. of Variation	6.18	18.28

5. Conclusions

The investigation done under this study have reached the following results:

Table V. Comparison between recycled and new aggregates

Property	New	Recycled	Difference %
Water absorption 3/4	0.73	4.79	556.2
Water absorption 1/2	0.98	5.24	434.7
Water absorption 3/8	1.30	6.76	420
Density	2.70	2.48	8.1
Impact Value	6.3	13	106.7
LA abrasion	15.01	32.54	116.8
10% fines	373.3	163.98	56.1
Sulphate content	0.17	0.7	311.8
Chloride content	0.003	0.055	1733

1. The values indicate that no significant changes in densities were detected and the value of 2.48 may be considered constant for recycled aggregates.
2. For water absorption, however, although no changes for each size were detected, the values have some dependence on aggregates size.
3. Lack of covering and storing concrete a materials and mixing machinery on site from direct sun light and open air have increased temperature of concrete and exposed its materials to all kind of pollution.
4. The changes in aggregates properties between different batches are tolerable.
5. The properties of recycled aggregates vary considerably compared to those of virgin new aggregates.

References

ASTM C1202. 1990. Electrical indication of concrete's ability to resist ion penetration. American Society for Testing and Materials, Philadelphia, Pennsylvania.

BSI 1881: Part 119. 1983. Method for determination of compressive strength of concrete cubes.

ASTM C39. 1990. Standard test method for compressive strength of cylindrical concrete specimens. American Society for Testing and Materials, Philadelphia, Pennsylvania.

ASTM C231. 1990. Standard test method for air contest of freshly mixed concrete by the pressure method. American Society for Testing and Materials, Philadelphia, Pennsylvania.

ASTM C642. 1990. Standard test method for specific gravity, absorption and voids in hardened concrete. American Society for Testing and Materials, Philadelphia, Pennsylvania.

Verified Parameter Identification for Solid Oxide Fuel Cells

Ekaterina Auer[1], Stefan Kiel[1] and Andreas Rauh[2]
[1]*University of Duisburg-Essen, INKO, {auer,kiel}@inf.uni-due.de,*
[2]*University of Rostock, Chair of Mechatronics, andreas.rauh@uni-rostock.de*

Abstract.
In the last decades, a lot of research in the area of decentralized energy supply systems has been focused on design and development of solid oxide fuel cells (SOFC). These devices convert chemical energy directly into electricity and represent an environmentally friendly alternative for the use as auxiliary power supply units. The advantages of SOFCs include high efficiency and flexibility with respect to the kind of fuel, whereas the main disadvantages are the complicated production process and the necessity for advanced control procedures to deal with instationary operating points. Contrary to this demand, most state-of-the-art control strategies for fuel cells cover stationary operating points only. Another difficulty is that many system parameters are influenced by significant uncertainty.

An important goal of a current joint project between the Universities of Rostock and Duisburg-Essen is to develop dynamic system models which accurately describe the instationary behavior of SOFCs. Here, one possibility to deal with parameter and model uncertainty is the use of interval analysis. Aside from providing a natural representation of bounded uncertainties, interval and similar methods guarantee the correctness of simulation results. We apply a verified global optimization algorithm based on that from (Hansen and Walster(2004)) to identify uncertain parameters of a dynamic SOFC model by (Rauh et al.(2011)). The model covers the effects of preheated air and fuel gas supply along with the corresponding reaction enthalpies on the thermal behavior. The parameters of interest describe the thermal resistances of the stack materials, the dependency of heat capacities on temperature, and the heat produced during the electrochemical reactions on the surface of each individual fuel cell. Because of the complex structure of the goal function, the optimization software has to be adjusted to the problem, which was one of the reasons we chose the solver UNIVERMEC by (Dyllong and Kiel(2010)) allowing for additional flexibility.

Keywords: Interval analysis, verified optimization, SOFC systems, software design

1. Introduction

Solid oxide fuel cells (SOFCs) convert chemical energy directly into electricity and represent an environmentally friendly alternative for the use as auxiliary power supply units in, for example, different types of vehicles or stationary industrial or domestic systems. These devices are currently in the focus of research on decentralized energy supply systems due to their high efficiency and flexibility with respect to the kind of fuel. However, SOFCs are difficult to produce and in need of procedures for dealing with instationary operating points. Control strategies for SOFCs are mostly designed for constant operating conditions and are based on simplifying assumptions which are not valid for wide operation ranges (Bove and Ubertini(2008); Pukrushpan et al.(2005)). This makes development of approaches taking into account

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

also instationary points necessary, which is the major goal of a current joint project between the Universities of Rostock and Duisburg-Essen (Rauh et al.(2011); Rauh et al.(2012a); Dötschel et al.(2012); Rauh et al.(2012b)).

Control-oriented mathematical models for SOFCs should be both accurate and applicable in real time for engineers to be able to develop robust controllers and state estimators. The thermal behavior of SOFC systems is usually described by partial differential equations. To obtain a model of the fuel cell stack temperature suitable for control design, the stack is semi-discretized into $L \times M \times N$ finite volume elements (cf. Figure 1). This leads to a set of $L \times M \times N$ nonlinear ordinary differential equations (ODEs). They are generated by the use of the first law of thermodynamics for each finite volume element to express the spatial temperature distribution in the interior of the SOFC stack. The ODEs are derived in such a way as to be valid in a wide operating range including not only the neighborhood of the desired operating point, but also the SOFC system's heating and cooling phases. The influence of varying electrical load conditions can be included in the thermal system model by means of a disturbance input. At the moment, the models are developed under the assumption that the temperature is homogenous in each volume element. However, using finite element approaches which approximate the actual situation of temperature inhomogeneities is also possible and a topic for our future work.

In this paper, we discuss possibilities to parameterize such control-oriented mathematical models in an accurate and robust way. The parameters of interest describe the thermal resistances of the stack materials, the dependency of heat capacities on the temperature, and the heat produced during the electrochemical reactions on the surface of each individual fuel cell. The parametrization is performed on the basis of measured data for the SOFC test rig available at the Chair of Mechatronics at the University of Rostock. The basics for parameter identification in our case are shown in (Rauh et al.(2011)). Note that parameters of the models developed for SOFCs are influenced by a considerable uncertainty. Aside from the inevitable model simplification, its sources are temporal and spatial discretizations as well as measurement and rounding errors.

One possibility to deal with the uncertainty (which can be accepted as bounded for the purpose of this first study) is the use of verified methods and, in particular, interval analysis (Moore et al.(2009)). They provide

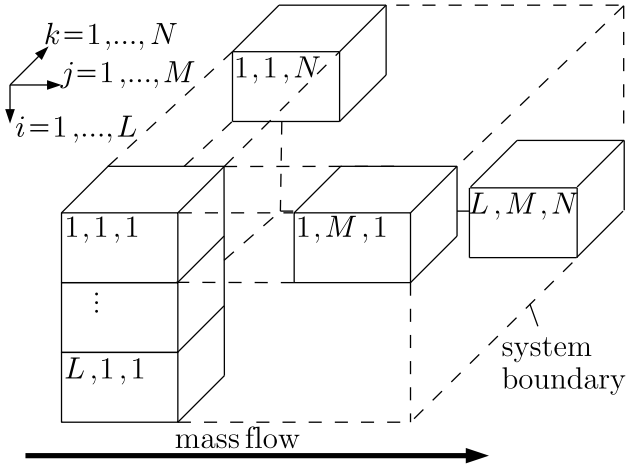


Figure 1. Semi-discretization of the fuel cell stack module into finite volume elements.

a natural representation of such kind of incertitude and guarantee the correctness of results. A successful attempt at using verified techniques for parameter identification has been made in (Rauh et al.(2012a)). By employing a basic interval optimization routine for three possible system orders (corresponding to spatial resolutions $1 \times 1 \times 1$, $1 \times 3 \times 1$, and $3 \times 3 \times 1$, cf. Figure 1), the authors managed to reduce the estimation error in comparison to the non-verified procedure which used Nelder-Mead simplex algorithm provided by MATLAB (`fminsearch`).

Optimization tasks for practical applications often have to deal with a complex goal function structure. It might include, for example, solutions to initial value problems and many summands, as is the case for the identification problem considered in this paper. Because of the complex structure of the goal function, the optimization software has to be adjusted to the problem at hand. These two reasons (complexity and the need for adjustments) usually render the use of such well known verified optimizers as GLOBSOL (Kearfott(1996)) difficult. In this paper, we consider the simplest situation of the $1 \times 1 \times 1$ spatial discretization in detail and show how the results obtained using the basic procedure from (Rauh et al.(2012a)) can be improved by exploiting the flexibility of the newly developed solver UNIVERMEC (Dyllong and Kiel(2010)). Besides implementing the usual global optimization algorithm from (Hansen and Walster(2004)) in C++, this solver allows users to choose the underlying data type for the evaluation of the goal function (e.g. naive interval or Taylor model (Berz(1995))) or the optimization strategy (e.g. with/ without differentiation, cf. Section 4.1) freely. Additionally, it is implemented in such a way as to allow for parallelization. Note that although the issue of real-time applicability is not important for the considered case of the offline parametrization, the complexity of the problem makes parallelization necessary if we want to obtain results in an acceptable time, especially if models for fluidic and electrochemical SOFC subsystems are to be included in the simulation process along with the thermal one (Rauh et al.(2011)).

The use of UNIVERMEC has one more reason. In the context of the already mentioned joint project, we plan to develop a framework for modeling, simulation, and control of SOFC systems based on the strategies developed in (Rauh et al.(2011); Dötschel et al.(2012)). This framework, supplemented by a graphical interface, should allow users to perform both verified and usual floating point computations with SOFC models of their choice easily, making flexibility of the underlying routines with respect to basic data types and algorithms unavoidable. Obviously, optimization software should also comply with this requirement, which UNIVERMEC does.

The paper is structured as follows. In Section 2, the considered problem is described in detail. A basic identification procedure for this task is summarized in Section 3. The new optimizer, parameter identification results and a comparison between the basic procedure and UNIVERMEC are reported on in Section 4. Finally, conclusions are in Section 5.

2. Problem Formulation

To describe the dynamics of fuel cell systems such as SOFCs, it is necessary to subdivide the overall model into three parts characterizing its fluidic, electrochemical and thermal behavior. Each of these subsystems is modeled by nonlinear ODEs expressing the corresponding dynamics in terms of the dominant physical phenomena such as temperature dependent heat capacities, current density dependent partial pressures of the gas mixtures, and internal Ohmic losses combined with electrical storage effects. For the modeling of

the thermal subsystem, the electrochemical reactions

$$2\,H_2 + 2O^{2-} \rightarrow 2\,H_2O + 4e^- \quad \text{and} \quad O_2 + 4e^- \rightarrow 2O^{2-} \tag{1}$$

at the anode and the cathode of the fuel cell are considered. A finite dimensional model is generated by using integral balances of the internal energy to characterize the temperature distribution. The modeling procedure described in detail in (Rauh et al.(2011)) leads to the following ODE for the discretization with $L = M = N = 1$:

$$\begin{aligned}
\dot{\theta}_{FC} \;=\; & \dot{m}_{H_2} \cdot \Big(p_{\Delta H,2} \cdot \theta_{FC}^2 + p_{\Delta H,1} \cdot \theta_{FC} + p_{\Delta H,0}\Big) + 6 \cdot p_A \cdot (\theta_A - \theta_{FC}) + (\theta_{AG} - \theta_{FC}) \\
& \cdot \; (\dot{m}_{H_2} \cdot (p_{H_2,2} \cdot \theta_{FC}^2 + p_{H_2,1} \cdot \theta_{FC} + p_{H_2,0}) + \dot{m}_{H_2O} \cdot (p_{H_2O,2} \cdot \theta_{FC}^2 + p_{H_2O,1} \cdot \theta_{FC} + p_{H_2O,0}) \\
& + \; \dot{m}_{N_2} \cdot (p_{N_2,A,2} \cdot \theta_{FC}^2 + p_{N_2,A,1} \cdot \theta_{FC} + p_{N_2,A,0})) + I_{FC} \cdot p_{el} - \dot{m}_A \cdot (\theta_{FC} - \theta_{CG}) \\
& \cdot \; (77 \cdot p_{N_2,C,0}/100 + 11 \cdot p_{O_2,0}/50 + 77 \cdot p_{N_2,C,1} \cdot \theta_{FC}/100 \\
& + \; 11 \cdot p_{O_2,1} \cdot \theta_{FC}/50 + 77 \cdot p_{N_2,C,2} \cdot \theta_{FC}^2/100 + 11 \cdot p_{O_2,2} \cdot \theta_{FC}^2/50)
\end{aligned} \tag{2}$$

with the initial condition $\theta_{FC}(0) = 299.7053$ K. Initial guesses for the time invariant parameters along with their meanings are shown in Table I. Here, the temperature dependent heat capacities c_{H_2} of hydrogen, c_{H_2O} of water vapor, $c_{N_2,A}$ of nitrogen at the anode, $c_{N_2,C}$ of nitrogen at the cathode, and c_{O_2} of air as well as the reaction enthalpy $\Delta_r H$ are replaced by their second-order polynomial approximations with certain coefficients $\alpha_{g,i}$ for $g \in \{H_2, H_2O, N_2, O_2\}$ and $i = 0, 1, 2$. After some expression manipulations, the corresponding coefficients $p_{g,i}$ from Table I appear, which are proportional to $\alpha_{g,i}$ (Rauh et al.(2011); Rauh et al.(2012a)).

The identification is performed with respect to parameters for which interval bounds are given in the table, namely, the zero-order terms of the polynomial approximations of the temperature-dependent specific heat capacities and the reaction enthalpy. Time variant inputs are shown in Table II. Their values along with the temperature θ_{FC} are measured each second ($h = 1$) for the time period of $T = 19963$ seconds. The measurement device is known to cause the uncertainty of $\Delta y_m \in [\Delta y_m]$. Aside from including the information about the measurement error of the device, the interval $[\Delta y_m]$ should enclose the effects of not being able to measure the temperature at exactly the same point as defined by the output variable of a certain finite volume element in the model.

Since we plan to use verified techniques, the task is to minimize the upper bound $\overline{J}$ of the goal function

$$J = \sum_{k=1}^{T} \left(y(t_k, p) - y_m(t_k)\right)^2 \tag{3}$$

with respect to the six parameters $p = [p_{H_2,0}\;\; p_{H_2O,0}\;\; p_{N_2,A,0}\;\; p_{N_2,C,0}\;\; p_{O_2,0}\;\; p_{\Delta H,0}]$, where $y(t_k, p) = \theta_{FC}(t_k, p)$ is the simulated temperature of the fuel cell at the time $t_k = 1, \ldots, T$ obtained from Eq. (2) and $y_m(t_k)$ the measured temperature at the same point. Note that here and in the following, the goal function is written down for our situation of $h = 1$, $t_k = h \cdot k$, $k = 1, \ldots, T$. The function J quantifies deviations between the measured output vector and the simulated temperature vector acquired with T samples and a

Table I. Meanings and values of the constant parameters in Eq. (2)

Name	Value	Physical meaning (proportional to)
p_A	$-2.111896 \cdot 10^{-5}$	inverse of the thermal insulation resistance
p_{el}	$1.645381 \cdot 10^{-3}$	resistance of SOFC materials
$p_{H_2,0}$	-7.614159 ± 1	heat capacity of hydrogen (order 0)
$p_{H_2,1}$	$-6.023259 \cdot 10^{-5}$	heat capacity of hydrogen (order 1)
$p_{H_2,2}$	$9.513841 \cdot 10^{-8}$	heat capacity of hydrogen (order 2)
$p_{H_2O,0}$	0.6273529 ± 1	heat capacity of water vapor (order 0)
$p_{H_2O,1}$	$-6.479947 \cdot 10^{-4}$	heat capacity of water vapor (order 1)
$p_{H_2O,2}$	$-1.583060 \cdot 10^{-7}$	heat capacity of water vapor (order 2)
$p_{N_2,A,0}$	-0.8367768 ± 1	heat capacity of nitrogen at the anode (0)
$p_{N_2,A,1}$	$6.250080 \cdot 10^{-4}$	heat capacity of nitrogen at the anode (1)
$p_{N_2,A,2}$	$-6.366022 \cdot 10^{-9}$	heat capacity of nitrogen at the anode (2)
$p_{N_2,C,0}$	0.4525605 ± 1	heat capacity of nitrogen at the cathode (0)
$p_{N_2,C,1}$	$-1.453636 \cdot 10^{-4}$	heat capacity of nitrogen at the cathode (1)
$p_{N_2,C,2}$	$1.974873 \cdot 10^{-8}$	heat capacity of nitrogen at the cathode (2)
$p_{O_2,0}$	-0.5931353 ± 1	heat capacity of air (0)
$p_{O_2,1}$	$-8.060370 \cdot 10^{-6}$	heat capacity of air (1)
$p_{O_2,2}$	$-2.049129 \cdot 10^{-9}$	heat capacity of air (2)
$p_{\Delta H,0}$	-217.3967 ± 1	reaction enthalpy (0)
$p_{\Delta H,1}$	$-5.236888 \cdot 10^{-2}$	reaction enthalpy (1)
$p_{\Delta H,2}$	$-5.014673 \cdot 10^{-6}$	reaction enthalpy (2)

constant sampling time h. An additional condition, which is induced by the accuracy of the measurements $[\Delta y_m]$, is

$$y(t_k) \subseteq [y_m(t_k) - 15, y_m(t_k) + 15] =: [\Delta y_m(t_k)] \quad \text{for} \quad t_k = 1, \ldots, T \ . \tag{4}$$

Note that since we have only one temperature to model the thermal behavior of the whole stack in the $1 \times 1 \times 1$ case, the uncertainty of ± 15 also includes the incertitude arising from spacial discretization.

The global optimization problem (3) can be solved in the following two general ways, if there is no possibility to find an analytical solution to Eq. (2):

1. Verified approximation. The true solution $\theta_{FC}(t) = y(t)$ of the Eq. (2) is approximated by the explicit Euler method as

$$[y_k] := [y_{k-1}] + h \cdot f([y_{k-1}], [p]) \ , \tag{5}$$

where f denotes the right side of the Eq. (2). The approximation $[y_k]$ is substituted for the exact solution $y(t_k)$ in the goal function (3) and the discretization error ignored. In our setup, the sampling time h of 1 second is by at least two orders of magnitude smaller than the dominant time constants of the thermal process (2). For this reason, $[y_k]$ is an acceptable approximation of the true solution $y(t_k)$. Although we cannot verify the whole process by applying interval procedures in this case, the results of optimization of

Table II. Time-dependent inputs in Eq. (2)

Name	Meaning
$\dot{m}_{H_2}$	mass flow of hydrogen
$\dot{m}_{N_2}$	mass flow of nitrogen at the anode
$\dot{m}_{H_2O}$	mass flow of water vapor
$\dot{m}_A$	mass flow of air at the catode
θ_{AG}	initial temperature of the anode gas in K
θ_{CG}	initial temperature of the cathode gas in K
θ_A	temperature of the environment in K
I_{FC}	electric current in A

the approximated goal function

$$J_{app} = \sum_{k=1}^{T} \left(y_{k-1} - y_m(t_k) + h \cdot f(y_{k-1}, p)\right)^2 \tag{6}$$

are verified. It is easy (although costly) to find the first (and second) derivatives of J_{app} with the help of algorithmic differentiation using, for example, FADBAD++ (Stauning and Bendtsen(2006)), should the chosen optimization algorithm (Kearfott(1996); Hansen and Walster(2004); Rauh et al.(2012b)) require that.

2. Verified solution. The true solution $\theta_{FC}(t) = y(t)$ of the Eq. (2) is enclosed numerically by a verified IVP solver such as VNODE-LP (Nedialkov(2011)). It is more difficult to compute the derivatives $\partial J/\partial p$ of the cost function J, because it requires solving an additional IVP of the form

$$\dot{s}_i = \frac{\partial f}{\partial \theta_{FC}} \cdot s_i + \frac{\partial f}{\partial p_i} \quad \text{with} \quad s_i = \frac{\partial \theta_{FC}}{\partial p_i} \tag{7}$$

in our one-dimensional case for each parameter p_i in each iteration step since

$$\frac{\partial J}{\partial p_i} = 2 \cdot \sum_{k=1}^{T} (y(t_k, p_i) - y_m(t_k)) \cdot s_i \ . \tag{8}$$

This procedure can be very expensive, especially if the sensitivity equations cannot be derived analytically and must be obtained by algorithmic differentiation. Therefore, it is advisable to prefer derivative-free optimization techniques in this case, although the additional information provided by derivatives usually improves the performance of a method.

For the purpose of an initial verified study, we consider only the first situation in this paper. Note that even the approximated situation (2), (6), (4) is not a simple one. Although the summands from (6) are easier to compute, the approximated problem is affected by the same difficulties as (3) from the point of view of verification in addition to suffering from discretization errors. For example, it involves at least T occurrences of the same (interval) parameters leading to the dependency problem and overestimation (Kieffer et al.(2011)). Since T is equal to 19963 in our case, overestimation reduction is not trivial for the considered optimization task.

3. Basic Approach to Verified Optimization

In this section, we describe the basic procedure shown in Figure 2, which was applied to the problem (2), (6), (4) in (Rauh et al.(2012a)). It is implemented as a C++ routine using PROFIL/BIAS (Keil(2008)) for basic interval evaluations and FADBAD++ for algorithmic differentiation. The first step is to define the interval vector bounding the search space. To find the globally optimal solution, this vector is subdivided further in the direction μ, chosen in each iteration step in such a way as to maximize the sensitivity measure

$$\Delta J^{<l>} = \sum_{\nu=1}^{T} \left[\frac{\partial [y^{<l>}(t_\nu)]}{\partial [p^{<l>}]} \cdot \frac{\mathrm{w}([p^{<l>}])^2}{\mathrm{w}([p^{<0>}])} \right] \tag{9}$$

for each parameter p, where L is the current list length, $l = 0, \ldots, L$, and $\mathrm{w}(\cdot)$ the width of the box. Intervals from the list which produce a small upper bound for $\overline{J}^{<l>}$ are preferred for subdivision. For the resulting subintervals, a validity criterion based on the condition in (4) is applied additionally. The following three cases are distinguished:

Consistent parameter vectors are no longer subdivided. They are characterized by

$$[y(t,p)] \subseteq y_m(t) + [\Delta y_m] \tag{10}$$

for each $t \in [0,T]$ with the worst-case measurement error $[\Delta y_m] = [-15, 15]$.

Inconsistent parameter vectors are excluded. They are identified by

$$[y(t,p)] \cap (y_m(t) + [\Delta y_m]) = \emptyset \tag{11}$$

for at least one point of time $t \in [0,T]$.

Undecided parameter vectors are subdivided further either until they fall below a minimum diameter or until a maximum number of subdivisions is reached. They are characterized by

$$[y(t,p)] \cap (y_m(t) + [\Delta y_m]) \neq \emptyset \tag{12}$$

for each $t \in [0,T]$ and

$$[y(t,p)] \not\subseteq y_m(t) + [\Delta y_m] \tag{13}$$

for at least one point of time.

Note that this procedure is not an all-encompassing one. The basic branch-and-bound algorithm is tailored for the problem at hand by exploiting the condition in Eq. (4) as its constituent part (cf. the validity test in Figure 2). In return, it can handle higher system orders ($1 \times 3 \times 1$ and $3 \times 3 \times 1$ finite volume elements along with the $1 \times 1 \times 1$ problem considered here) and more parameters (e.g. all from Table I), a difficult task for a general-purpose solver. The results produced by this procedure for the problem described in Section 2 in comparison to those obtained with UNIVERMEC (which is a highly adjustable general-purpose solver) will be shown in Section 4.2.

Initialization of the parameter range $[p^{<0>}] = [p^{<L>}]$, $L = 1$

while (stopping criterion is not reached)

- **Determine the parameter vector** $[p^{<l>}]$ to be subdivided from the complete list $[p^{<1>}], \ldots, [p^{<L>}]$
- **Determine the component** μ of the parameter vector $[p^{<l>}]$ to be subdivided
- **Splitting procedure:** $[p^{<L+1>}] := [p^{<l>}]$

$$[p_\mu^{<l>}] := \left[\inf\left([p_\mu^{<l>}]\right), \text{mid}\left([p_\mu^{<l>}]\right)\right]$$
$$[p_\mu^{<L+1>}] := \left[\text{mid}\left([p_\mu^{<L+1>}]\right), \sup\left([p_\mu^{<L+1>}]\right)\right]$$

- **Evaluate state equations** for $[p^{<l>}]$
- **Validity test:** Delete $[p^{<l>}]$ from the parameter list if guaranteed to be inconsistent
- **Evaluate state equations** for $[p^{<L+1>}]$
- **Validity test:** Delete $[p^{<L+1>}]$ from the parameter list if guaranteed to be inconsistent
- **Recount** the length L of the list

Figure 2. The basic routine for verified parameter identification.

4. Application of a General-Purpose Optimizer and Comparison of Results

In this section, we show how to apply the general-purpose optimizer UNIVERMEC to the problem (2), (6), (4). The major difficulty here is to overcome overestimation. Besides, we need to incorporate the condition in Eq. (4) in an appropriate form. First, we describe the software-oriented foundation making our optimizer more flexible than other general-purpose ones such as GLOBSOL. After that, we show the results of its application to the problem of SOFC parameter identification. Finally, we compare the outcome to that of the basic routine from Section 3.

4.1. UNIVERMEC: DESIGN, ALGORITHMS, FEATURES

The optimization algorithm (Dyllong and Kiel(2010)) is implemented in the uniform framework UNIVERMEC (**Uni**fied Framework for **Ver**ified Geo**Me**tric **C**omputations), originally developed for geometric computations (Dyllong and Kiel(2011)). UNIVERMEC provides a conceptual and a software basis for

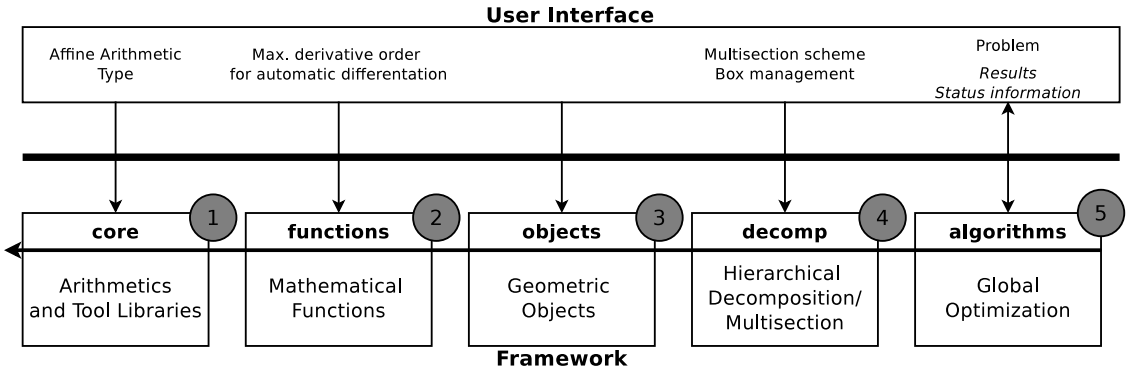


Figure 3. The basic structure of UNIVERMEC. Every layer depends only on the ones left of it. The parameters of each layer can be changed through the user interface. Words in italics denote information from the framework.

handling different kinds of (verified) arithmetics (e.g interval, affine (de Figueiredo and Stolfi(1997)), or Taylor model) and algorithms uniformly. Owning to its design, the framework allows users to evaluate their models (e.g. for geometric objects) with the arithmetic suitable for their task. If necessary, various decomposition or branching schemes can be applied to user-defined models regardless of their actual mathematical representation. Higher level algorithms, such as parameter identification or distance computation, can access these different methods easily, allowing for ready reuse and exchange. Additionally, UNIVERMEC provides a fair comparison between specific techniques inside high level algorithms because the overhead and implementations of the algorithms themselves are identical.

An overview of the framework components is given in Figure 3. The code is organized into five conceptual layers. The first layer (`Core`) implements arithmetic concepts in form of an abstract algebra and wrappers for actual libraries (e.g. C-XSC (Hofschuster and Krämer(2004)) for intervals or YALAA (Kiel(2012)) for affine arithmetic). The second layer (`Functions`) defines an interface for scalar functions formally and independently of the chosen kind of arithmetic. The layer `Objects` specifies a uniform formal representation for different models of objects, which is important for geometric computations and can be skipped in our context of parameter identification. The forth layer `Decomp` encapsulates various hierarchical decomposition and multisection strategies. Finally, the layer `Algorithms` provides a basis for flexible implementations of different kinds of high level algorithms, in particular, the optimization algorithm we use in this paper to solve the problem (2), (6), (4). Although the algorithm can be implemented without the framework, UNIVERMEC ensures the maximum flexibility necessary to solve the task.

The optimizer is developed to solve a general problem with inequality constraints according to the approach described in (Hansen and Walster(2004)):

$$\begin{aligned} &\min \phi(x) \quad \text{with } x \in \mathbb{R}^{\mathrm{d}} \\ &\text{subject to } \; g_i(x) \leq 0 \;\; \text{for } i = 1, ..., m, \end{aligned} \tag{14}$$

where the objective function $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ is scalar. The algorithm calculates a verified interval enclosure of the minimum and maintains three lists: the working list $\mathcal{L}$, the temporary list $\mathcal{L}_{\mathrm{tmp}}$, and the result list $\mathcal{L}_{\mathrm{final}}$. All lists contain parts of the search space in the form of boxes. $\mathcal{L}$ accommodates the boxes to be processed further, while $\mathcal{L}_{\mathrm{final}}$ is filled with the final results. The basic structure of the algorithm is outlined in Figure 4.

To allow users to adapt the algorithm to their problems, it is subdivided into several stages (PHASE_ in the figure), the behavior of which can be changed individually. Following (Hansen and Walster(2004)), we choose contractors and strategies for box reduction in dependence on the feasibility of the current box. That is, it is necessary to distinguish the stages PHASE_POS_INFEAS, PHASE_FEAS, and PHASE_STRICT_FEAS for boxes with unknown, certain and strictly certain feasibility, respectively (cf. Figure 4). The contractors in the stages PHASE_A to PHASE_D are called independently of the feasibility of the current box in each iteration, while PHASE_SPLIT is called on boxes directly after the multisection step.

Inside an `apply_contractor` call, a box may be pruned, completely discarded or moved to $\mathcal{L}_{\mathrm{final}}$ if it satisfies the termination criterion. In the latter two cases, the main loop is restarted. Unlike other global optimization algorithms, UNIVERMEC maintains not only the lists $\mathcal{L}$ and $\mathcal{L}_{\mathrm{final}}$, but also $\mathcal{L}_{\mathrm{tmp}}$ similarly to the distance computation algorithm in (Dyllong and Kiel(2011)). If a box is subdivided below a certain minimum width ϵ_t, it is temporarily deleted from $\mathcal{L}$ and moved into $\mathcal{L}_{\mathrm{tmp}}$. This strategy ensures that the problem domain is subdivided more uniformly and prevents heuristics such as best-first from causing a deep subdivision in the wrong region. When $\mathcal{L}$ becomes empty, all boxes from $\mathcal{L}_{\mathrm{tmp}}$ are moved back into $\mathcal{L}$, and the user can alter the algorithm stages again. In this way, accelerating devices such as interval Newton can be configured dynamically.

The algorithm can be parallelized efficiently, since it can work on parts of the search region independently of other parts. The search region cannot be subdivided a priori, because it is unknown where a deep search will take place. For workload sharing, a permanent synchronization among all threads is important. In our current implementation, $\mathcal{L}$ is shared among all threads, which is suitable only for shared-memory architectures. Besides, there is a possible bottleneck in the algorithm, because every access to $\mathcal{L}$ represents a critical section. However, the time spent there can be reduced by an intelligent initialization approach for the verified upper bound of the minimum $\tilde{\phi}$ (e.g. using IPOPT (Wächter and Biegler(2006))).

In practice, evaluating the goal function is expensive. While derivatives are theoretically available through FADBAD++ in UNIVERMEC, we do not use them, because their recursive evaluation at every time step slows down the computations considerably. Instead, we employ the following derivative-free strategies:

PHASE_SPLIT: Bounds the goal function with interval arithmetic and test condition (4).

PHASE_PA: Tests the feasibility.

PHASE_POS_FEAS: Tests the box consistency on constraints.

PHASE_TMP: Tries to find a verified upper bound on the minimum using the midpoint test.

PHASE_FINAL: Bounds the goal function using affine arithmetic and pruning by (15)–(16).

We use the implicit linear interval estimation (ILIE) technique (Bühler(2002)) to take the condition in Eq. (4), which cannot be written down in terms of expressions valid globally for all $t = 1, \ldots, T$, into account in our general purpose solver. Originally, this technique was developed and used for computer graphics applications. However, it also helps to incorporate such heuristic conditions as Eq. (4) into the overall optimization procedure.

Input: Search region X_0
Output: Enclosure of minimum ϕ^*

$\mathcal{L} := \{X_0\}$; $S := RUNNING$;
while $S == RUNNING$ **do**
  if $\mathcal{L} = \emptyset$ **then**
    Get configuration for next phase;
    forall $X' \in \mathcal{L}_{\text{tmp}}$ **do**
      `apply_contractors(`*PHASE_TMP,* X'`)`;
    end
    $\mathcal{L} := \mathcal{L}_{\text{tmp}}$; $\mathcal{L}_{\text{tmp}} := \emptyset$;
    if $\mathcal{L} == \emptyset$ **then** $S := FINISHED$; **continue;**
  end
  $X := head(\mathcal{L})$; $\mathcal{L} := tail(\mathcal{L})$;
  `apply_contractors(`*PHASE_A,* X`)`;
  if $\neg$`feasible(`X`)` **then**
    `apply_contractors(`*PHASE_POS_INFEAS,* X`)`;
  end
  `apply_contractors(`*PHASE_B,* X`)`;
  if `feasible(`X`)` **then** `apply_contractors(`*PHASE_FEAS,* X`)`;
  `apply_contractors(`*PHASE_C,* X`)`;
  if `strictly_feasible(`X`)` **then**
    `apply_contractors(`*PHASE_STRICT_FEAS,* X`)`;
  end
  `apply_contractors(`*PHASE_D,* X`)`;
  if $\mathrm{w}(X) < \epsilon_t$ **then** $\mathcal{L}_{\text{tmp}} \leftarrow X$;
  else
    $\mathcal{N}$:= `split(`X`)`;
    forall $X' \in \mathcal{N}$ **do**
      `apply_contractors(`*PHASE_SPLIT,* X'`)`;
    end
    $\mathcal{L} \leftarrow \mathcal{N}$;
  end
end
forall $X' \in \mathcal{L}_{\text{final}}$ **do**
  `apply_contractors(`*PHASE_FINAL,* X'`)`;
end
$\underline{\phi^*} := \min_{X \in \mathcal{L}_{\text{final}}} (\underline{\phi(X)})$; $\overline{\phi^*} := \min(\max_{X \in \mathcal{L}_{\text{final}}} (\overline{\phi(X)}), \phi^*)$;

Figure 4. Interval optimization algorithm UNIVERMEC: $\tilde{\phi}$ is a verified upper bound on the minimum.

Let $y(t_k, p)$ be the solution of the IVP (2) at the point t_k, p the parameter vector ($p \in \mathbb{R}^6$ in our case), and $L(t_k, p)$ its linearized enclosure

$$L(t_k, p) := \sum_{i=1}^{6} a_i p_i + [l] \quad \text{with} \quad [l] \in \mathbb{I}, p_i \in \mathbb{R} \ . \tag{15}$$

Further, if $[\Delta y_m(t_k)]$ is the measurement with the corresponding uncertainty at the point t_k and $[p]$ the current interval box for the (unknown) parameters p, then the interval $L(t_k, p)$ for each $p \in [p]$ should not exceed $[\Delta y_m(t_k)]$ for the condition in Eq. (4) to be valid. After factoring out a parameter p_i, $i = 1 \ldots 6$,

from this relation, we obtain the following transformed condition:

$$p_i \in \frac{[\Delta y_m(t_k)] - [l] - \sum_{j=1, j\neq i}^{6} a_j p_j}{a_i} \ , \tag{16}$$

which can be incorporated as a solution strategy into the overall algorithm. Here, the coefficients a_i and the enclosure of the nonlinearity $[l]$ are determined by using affine arithmetics and, in particular, the library YALAA, which UNIVERMEC allows us to access.

4.2. RESULTS AND COMPARISON

We solved the optimization problem with UNIVERMEC on an Intel Xeon CPU x5570, 2.9 GHz, 12Gb RAM virtualized computer with four cores. The algorithm was parallelized with the help of OpenMP (`http://openmp.org`). To obtain results comparable to those from the basic routine in Section 3, we limited the number of iterations to 54000 as had been done there. Besides, we used only the midpoint test as a discarding strategy. Note that the search space is bisected inside the basic routine, whereas UNIVERMEC uses the multisection scheme according to (Ratz(1992)). To reduce the computational effort, we updated the upper bound only between stage switches (cf. Figure 4). The value for ϵ_t was initially set to $\mathrm{w}(X_0)/5$ and then reduced by half during every change of the stage.

After 54000 iterations, we obtained 36942 candidate boxes for optimal parameters p. No verified solution strictly complying with (4) could be found, as was also the case for the basic procedure from Section 3. The results are shown in Figure 5, right, in comparison to those from the basic procedure (Figure 5, left). In the figure, the difference $\delta(t_k)$ between the simulated and the measured temperature value is shown for the midpoints of the set of candidate optimal parameters $[\hat{p}]$, chosen in such a way as for the absolute error value $|\delta(t_k)|$ to be the smallest. Note that by selecting such $[\hat{p}]$, we no longer work with entirely verified results, but only with candidates. The results obtained by UNIVERMEC have less deviation from the measured values than those from the basic routine.

In Table III, we compare the outcome for the problem (2), (6), (4) obtained by a floating-point optimization in MATLAB, the basic interval routine from Section 3, and UNIVERMEC. The non-verified results are described in more detail in (Rauh et al.(2012a)). As the comparison measure e given in the table, we use the following practice-motivated expression:

$$e = \sqrt{\frac{\sum_{k=1}^{T} \left(y_{k-1} - y_m(t_k) + f(y_{k-1}, \mathrm{mid}([\hat{p}]))\right)^2}{T}} \ . \tag{17}$$

This choice leads to values of e similar to the usual root mean square error measure in the floating point case. The table shows that the interval based strategies are less accurate than the floating point one for this rudimentary model of order $1 \times 1 \times 1$, although the results obtained in UNIVERMEC are better than those from the basic routine. More accurate interval solutions can be produced by using higher order temperature models as shown in (Rauh et al.(2012a)). We plan to test these enhanced models with UNIVERMEC in the future.

In the third line of Table III, we provide CPU times for all three kinds of computations. Note that they cannot be considered absolutely, because they were obtained on different computers (Intel dual core

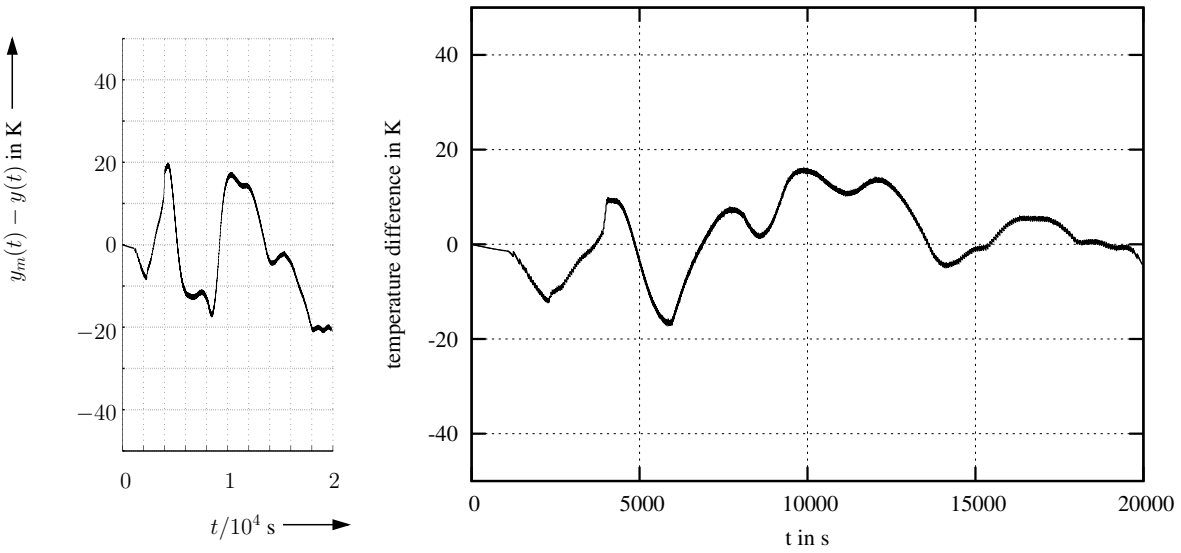


Figure 5. Difference between the simulated and measured temperatures for the model of order $1 \times 1 \times 1$ for the basic optimization routine from Section 3, left, and UNIVERMEC, right.

CPU, 2.8 GHz, 2 GB RAM for the basic interval and floating point routines) and in different environments (MATLAB for floating point vs. C++ for the others). Besides, the floating-point optimization was carried for all 20 parameters from Table I. However, the CPU times can give an idea of what the advantage of parallelization can be. For UNIVERMEC, we provide both the CPU time and the real time to show how long the computations took in reality. The results demonstrate that the usage of derivative free techniques in UNIVERMEC speeds up the computations (3.2 vs. 10 hours of CPU time) and that parallelization is effective in this case (1.33 hours of real time). Note that the CPU times include not only the ones needed to obtain candidate solutions, but also those for constructing the solutions in Figure 5 in all cases.

Table III. Empirical comparison between floating point routine, the basic interval routine, and UNIVERMEC for the SOFC model of order $1 \times 1 \times 1$

	floating point	basic interval	UNIVERMEC
error e in K	5.17	11.84	7.68
CPU time in hours	≈ 6	≈ 10	3.2 (1.33 real time)

Other experiments we performed showed that the settings we chose represented a good compromise between performance and accuracy for this rudimentary model. Increasing the number of iterations up to 270000 lead to a slightly smaller value of 7.5 for the root mean square error measure while increasing the CPU time considerably. The linearization strategy (15)–(16) does not offer much improvement in such situations because the resulting boxes are still too large. Setting the termination criterion to the box width of

10^{-1} (for which the linearization might work) is accompanied by computing times of over a week, which is not justified for this simple model. That is, if better accuracy is requested, better models should be used.

5. Conclusions

In this paper, we considered the simplest model for the temperature of a fuel cell stack in detail from the point of view of obtaining accurate and robust results for the problem of identification of its parameters. We showed how to apply a range of interval based algorithms to it. Although the results obtained using interval procedures are slightly less accurate than the floating point ones in this simplest case, the research in this direction seems promising because the accuracy of interval techniques improves if better models are used (Rauh et al.(2012a)). The presented verified optimizer UNIVERMEC helped to increase both the accuracy and the performance of interval techniques. The reasons were its flexible implementation, the use of derivative-free techniques and parallelization.

Our future work includes three general directions. The first one concerns SOFC modeling and control. The developed models will be used for a nonlinear state and disturbance observer design. Sensitivity based parameter identification routines will be implemented for a further improvement of the model quality. The second direction is to incorporate all the developed models and algorithms into a unified tool for modeling, simulation, and control of SOFCs with the help of both verified and floating point methods.

The last future work direction concerns the improvement of the quality of the obtained parameter set. For that purpose, we plan to implement the fully verified identification technique mentioned in Section 2 in UNIVERMEC. Moreover, the use of better models of orders $1 \times 3 \times 1$ and $3 \times 3 \times 1$ might help to actually verify the optimum of the problem (2), (6), (4). Besides, more thought should be given to the empirically motivated condition (4) itself. Although it helps to reduce the overestimation in the goal function (6), our tests have shown that the midpoints of the obtained parameter sets are infeasible with respect to it in spite of a high subdivision depth. Therefore, we should either prove the validity of the condition mathematically (at least, under certain assumptions), or try to reduce the overestimation in the goal function in other ways.

Acknowledgements

This work is founded by the German Research Council.

References

M. Berz. Modern Map Methods for Charged Particle Optics. *Nuclear Instruments and Methods A363*, pages 100–104, 1995.

R. Bove and S. Ubertini, editors. *Modeling Solid Oxide Fuel Cells*. Springer, Berlin, 2008.

K. Bühler. Implicit Linear Interval Estimations. In *Proceedings of the 18th spring conference on Computer graphics*, page 132. ACM, 2002.

L. de Figueiredo and J. Stolfi. *Self-Validated Numerical Methods and Applications*. IMPA, Rio de Janeiro, 1997.

T. Dötschel, A. Rauh, and H. Aschemann. Reliable Control and Disturbance Rejection for the Thermal Behavior of Solid Oxide Fuel Cell Systems. In *Proc. of Vienna Conference on Mathematical Modelling MATHMOD 2012*, Vienna, Austria, 2012. Accepted.

E. Dyllong and S. Kiel. A Comparison of Verified Distance Computation between Implicit Objects Using Different Arithmetics for Range Enclosure. *Computing*, 2011.

E. Dyllong and S. Kiel. Verified Distance Computation Between Convex Hulls of Octrees Using Interval Optimization Techniques. *PAMM Special Issue: 81st Annual Meeting of the International Association of Applied Mathematics and Mechanics (GAMM)*, 10(1):651–652, 2010.

E. Hansen and G. W. Walster. *Global Optimization Using Interval Analysis*. Marcel Dekker, New York, 2004.

W. Hofschuster and W. Krämer. C-XSC 2.0 — A C++ Library for Extended Scientific Computing. *LNCS 2991: Numerical Software with Result Verification*. Springer, 2004

R. Kearfott. *Rigorous Global Search: Continuous Problems*. Nonconvex optimization and its applications. Kluwer Academic Publishers, 1996.

C. Keil. PROFIL/BIAS, Version 2.0.8, 2008. `www.ti3.tu-harburg.de/keil/profil/`.

M. Kieffer, M. Csaba, H. Schichl, and E. Walter. Verified Global Optimization for Estimating the Parameters of Nonlinear Models. In *Modeling, Design and Simulation of Systems with Uncertainties*, Mathematical Engineering. Springer, 2011.

S. Kiel. YALAA – Yet Another Library for Affine Arithmetic. *Reliable Computing*. Submitted.

R. Moore, B. Kearfott, and M. Cloud. *Introduction to Interval Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, 2009.

N. S. Nedialkov. Implementing a Rigorous ODE Solver through Literate Programming. In A. Rauh and E. Auer, editors, *Modeling, Design, and Simulation of Systems with Uncertainties*, Mathematical Engineering. Springer, 2011.

J. Pukrushpan, A. Stefanopoulou, and H. Peng. *Control of Fuel Cell Power Systems: Principles, Modeling, Analysis and Feedback Design*. Springer, Berlin, 2nd edition, 2005.

D. Ratz. *Automatische Ergebnisverifikation bei Globalen Optimierungsproblemen*. Universität Karlsruhe, 1992.

A. Rauh, T. Dötschel, and H. Aschemann. Experimental Parameter Identification for a Control-Oriented Model of the Thermal Behavior of High-Temperature Fuel Cells. In *CD-Proc. of IEEE Intl. Conference MMAR 2011*, Miedzyzdroje, Poland, 2011.

A. Rauh, T. Dötschel, E. Auer, and H. Aschemann. Interval Methods for Control-Oriented Modeling of the Thermal Behavior of High-Temperature Fuel Cell Stacks. In *Proc. of SysID 2012*, 2012a. Accepted.

A. Rauh, L. Senkel, and H. Aschemann. Sensitivity-Based State and Parameter Estimation for Fuel Cell Systems. In *Proc. of 7th IFAC Symposium on Robust Control Design ROCOND 2012*, Aalborg, Denmark, 2012b. Submitted.

O. Stauning and C. Bendtsen. FADBAD++ web page. `http://www.fadbad.com/` .

A. Wächter and L. T. Biegler. On the Implementation of an Interior-Point Filter Line-Search Algorithm for Large-Scale Nonlinear Programming. *Math. Program.*, 106(1):25–57, 2006. ISSN 0025-5610.

Random Sets-based estimation of soundings density for geotechnical site investigation

D. Boumezerane[1)], S. Belkacemi[2)]
[1)] *Civil Engineering Department, University of Bejaïa, 06000 Bejaïa, Algeria, dboumezerane@gmail.com*
[2)] *Civil Engineering Department, Ecole Nationale Polytechnique, Algiers 16000, smainbelkacemi@yahoo.com*

Abstract: Geotechnical site investigation is a process conducted generally in two steps. One preliminary stage consisting in collecting available information and executing a limited number of soundings on site, and a second step of investigation based on the first stage using more soundings for soil testing. The optimal number of soundings is not known, it depends on number of factors such as geology of the site, soil variability and the type of project to build. Intervals of values are proposed by number of experts (engineers) concerning the soundings density based on preliminary information from site (soil variability, geology, type of project). Each engineer can give an interval of values based on his degree of belief. He will support his judgment by available information. As a first step we used only one parameter (soil variability) to construct the random sets. For certain soil variability degree between [1,10] the expert will give the corresponding number of soundings (with a degree of belief m_{ij}). Using Eurocode7 recommendations for site soundings, we constructed an "objective function" $f(X)$ to rely "soil variability" to the number of soundings. This function permits constructing the random set and obtain the number of soundings by unit area for each expert (engineer). The next step consists in aggregating information from other parameters (Geology, Project type…) and computing the random sets. The construction of upper and lower probabilities permits us optimizing the number of soundings to carry out on site.

Keywords: Geotechnical Investigation; Soundings density; Random Sets; Upper and Lower probabilities.

1. Introduction

Geotechnical site investigation is a process conducted generally in two steps. One preliminary stage consisting in collecting available information and executing a limited number of soundings on site, and a second step of investigation based on the first stage using more soundings for soil testing. The optimal number of soundings is not known, it depends on number of factors such as geology of the site, soil variability and the type of project to build.

The geotechnical engineer in charge of investigation should take into account all those factors and preliminary information to decide on the number of soundings to conduct in the second stage of investigation.

According to Cambefort (1980) there is no specific rule on the number of soundings to execute for geotechnical investigations. If an arbitrary loose mesh of soundings used in the preliminary study shows that the project area is relatively homogeneous then this quantity is satisfactory. However, if the results of the preliminary study show erratic conditions, the site characterization requires more soundings.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

Previous information on the site is generally given in form of geological and topographic maps and eventual results from adjoining sites. The engineer's judgement is important on the way investigation should be conducted, and especially on the number of soundings to make. The engineer will have a degree of belief about the density of boreholes depending on the preliminary information (soil variability, geologic profiles and type of project). The more important the preliminary information the more significant will be the degree of belief on the density of soundings to carry out. The degree of belief could be expressed as a subjective probability by experts. For a geotechnical investigation to be done, and given preliminary information, experts can formulate a subjective probability (degree of belief) related to the density of soundings to carry out. When variability of soil properties is very important for example experts will propose a "high" degree of belief that the density of boreholes should be "important" to complete an efficient investigation, on the other hand if the given preliminary information indicates a "low" variability of soil this will suggest an "important" degree of belief that the number of soundings to execute out will be "low". The approach will need certain number of experts to be reliable. The more experts the more reliable will be the result.

In a previous study we proposed to use fuzzy sets for this purpose (Boumezerane et al., 2011). In the present work we will use random sets to estimate intervals of soundings to carry out on site.

2. Parameters Affecting Geotechnical Investigations

The distribution of soundings to be made in a project area does not follow particular rules (Magnan, 2000), it depends on preliminary information among which;

- The geologic context of the project area,
- The preliminary results of investigation,
- The project type, and
- The knowledge of the neighbouring areas.

2.1. GEOLOGY

The available information about site's geology helps engineers to plan efficient geotechnical investigations. Information is obtained using geological and topographical maps. A visit on site is necessary, it permits having a reliable idea about the visible ground and formations constituting the soil.

The degree of information (knowledge) depends mainly on the scale of geological maps used, on the quality of information available (rough or precise) and on the on-site engineer's judgement. Geological published maps are fundamental tools for any of the analysis; however details have to be revealed by more specific studies. The use of maps is essential to have a first idea on the geologic formations constituting the site, their properties, as well as the possibilities of inadequate or adverse geologic details. Clayton et al (2005) recommend for geotechnical studies to use geologic maps in the scale 1/2500.

The spacing of borings depends on the geology of the area and may vary from a site to another. Boring spacing should be selected to intersect distinct geological characteristics of the project. Borings should be situated to confirm the location of significant changes in subsurface conditions as well as to confirm the continuity of apparently consistent subsurface conditions (US Corps of Engineers, 1994).

2.2. SOIL VARIABILITY

The preliminary step of geotechnical investigation consists in few soundings that permit to have a rough idea about the variability of soil properties. The parameter "Soil Variability" is important for the engineer to decide on how many soundings will be necessary in the second stage of site investigation. The soil variability is related to the number of different soil layers, their orientations and thickness. Average values of soil parameters can be obtained from different points of the site. For important variability of soil properties the density of soundings should be significant.

3. Random Sets Concept and Uncertainty

According to Helton (1997) uncertainty is classified into two distinct groups: aleatory and epistemic. *Aleatory uncertainty* is related to the natural variability of the variables involved. *Epistemic uncertainty* is related to lack of knowledge or data, therefore it can be reduced when new information is available. Possibility, evidence, interval analysis and random set (RS) theories have shown to be appropriate to deal with this type of subjective uncertainty, and here the information is expressed by means of intervals and linguistic terms (Alvarez, 2008).

3.1. RANDOM SETS

A *random set*, sometimes also referred to as a *Dempster-Shafer structure*, is given by finitely any subsets A_i, $i = 1, \ldots, n$ of a given set X, called the *focal sets*, each of which comes with a *probability weight* $m_i = m(A_i)$, $\sum m(A_i) = 1$. An example of a random set is shown in Fig. (Oberguggenberger, 2005).

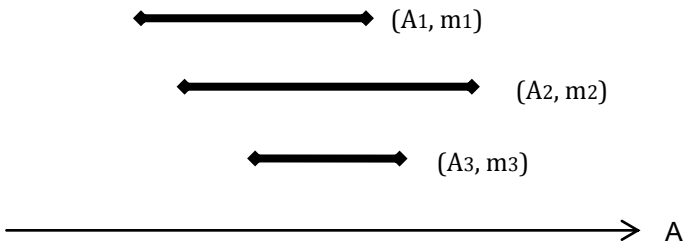


Figure 1. A random set

In the Dempster-Shafer approach (Alvarez, 2008), the random set allows to define a degree of belief $\gamma(S)$ and a degree of plausibility $\eta(S)$, respectively, that the realizations of the parameter A lie in S by;

$$\gamma(S) = \sum_{A_i \subset S} m(A_i)$$

$$\eta(S) = \sum_{A_i i \cap S \neq \emptyset} m(A_i)$$

The belief function $\gamma(S)$ or *Bel*, of a subset *S* is a set-valued function obtained through summation of basic probability assignments of subsets *Ai* included in *S* and the plausibility function $\eta(S)$, or *Pl*, of subset *S* is a set-valued function obtained through summation of basic probability assignments of subsets A_i having a non-zero intersection with *S.* They are envelopes of all possible cumulative distribution functions compatible with the data.

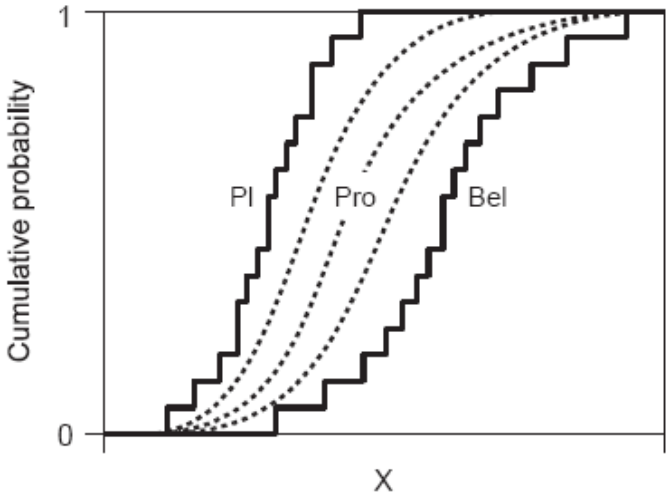


Figure 2. Upper bound (*Pl*) and lower bound (*Bel*) on precise probability (*Pro*)

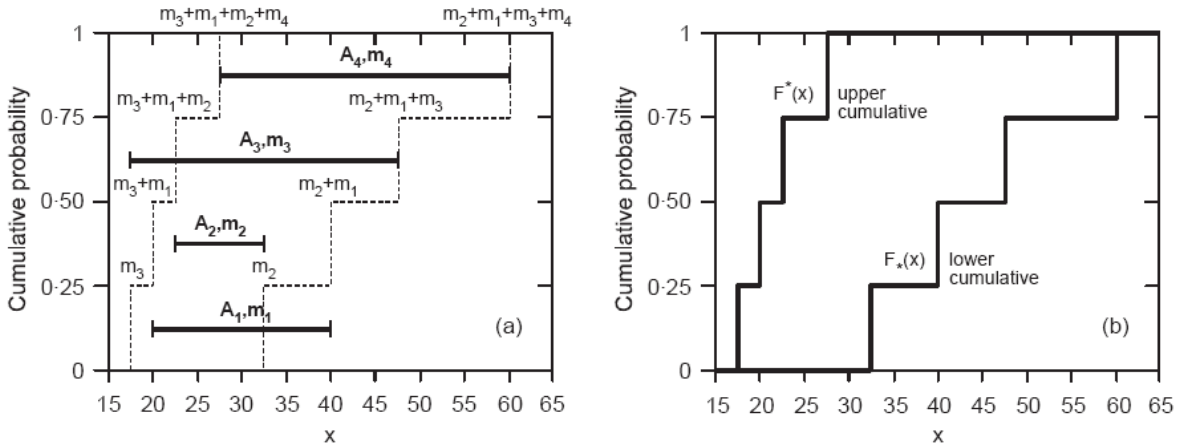


Figure 3. Random set ; (a) construction, (b) upper and lower discrete cumulative distribution function

Let's consider for instance a Dempster Shafer (D.S) structure which is formed by gathering the information provided by four different sources (e.g. books, experts, previous analysis, etc.) on the friction angle of some soil; each of those opinions will form one element A_i of the focal set A. Suppose that $A = \{A_1 = [20°,22°], A_2 = [21°,23°], A_3 = [18°,20°], A_4 = [20°,25°]\}$. The basic mass assignment given to each of those focal elements will represent the importance of each of those opinions in our assessments. Suppose for example that ($m(A_1) = 0.4$, $m(A_2) = 0.2$, $m(A_3) = 0.1$, $m(A_4) = 0.3$; this means that we are giving to our first source of information the largest relevance (Alvarez, 2008).

4 Situation of the Problem

The idea underlying the use of random sets as a tool to estimate soundings density is supported by their ability to handle vague and uncertain information. The degree of belief an engineer could have given preliminary information is used to construct the upper and lower probabilities to estimate the number of soundings for geotechnical investigation. The calibration is done upon minimal number of soundings per surface recommended by Eurocode7.

Let's have the opinions of different engineers concerning the soundings density based on preliminary information from site (soil variability, geology, type of project). Each engineer can give an interval of values based on his degree of belief. He will support his judgment by available information. If "Soil Variability" obtained from preliminary soundings is "Very Important" for instance then he will propose an important number of soundings with a strong degree of belief.

How Soil Variability is quantified by experts? A scale between 1 and 10 is proposed representing intervals of "Very Low", "Low", "Medium" and "High" Variability. Eurocode 7 recommends 1sounding per 40 × 40 m² as a minimum for an investigation. The degree 1 of soil variability corresponds to a "very low" variability. We consider this degree of variability necessitating the minimal number of soundings recommended by Eurocode7. The maximum number of soundings recommended by codes and some authors (Hunt, 2007) is given by 1 sounding / area of 15 × 15 m². Globally the number of soundings per unit area of 40 × 40 m² varies between 1 as a minimum and 6 as maximum, but it is possible to have more soundings if information is not enough.

Let's consider the opinions of experts about the density of soundings on site, given soil variability. For certain soil variability degree between [1,10] the expert will give the corresponding number of soundings (with a degree of belief m_{ij}).

5. Point to point approach

We try first with only one expert 1. According to eurocode7 the minimum number of soundings is 1 for an area of 40x40m². This minimum number as explained before could be used for a "Very low" soil variability (which is comprised in the interval [0,2] on the scale). If the maximum number of soundings (6 to 7 / unit area 40x40m²) corresponds to a "High variability" (10 on the scale) we could argue a linear variation and construct an "objective function" to rely "soil variability" to the number of soundings. This function will permit us to construct Upper and Lower probability boxes as a decision aid tool.

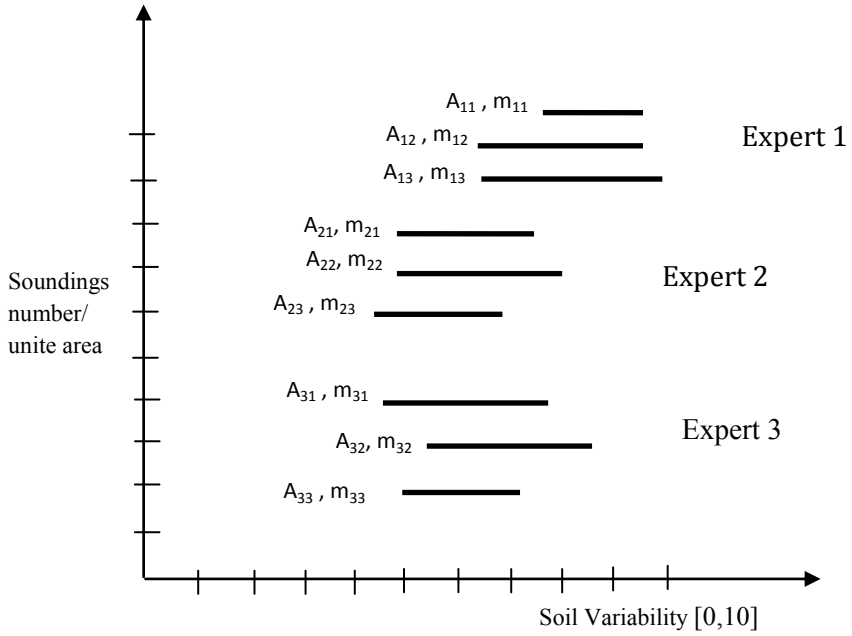


Figure 4. Scheme of random set "Soil Variability" given different experts

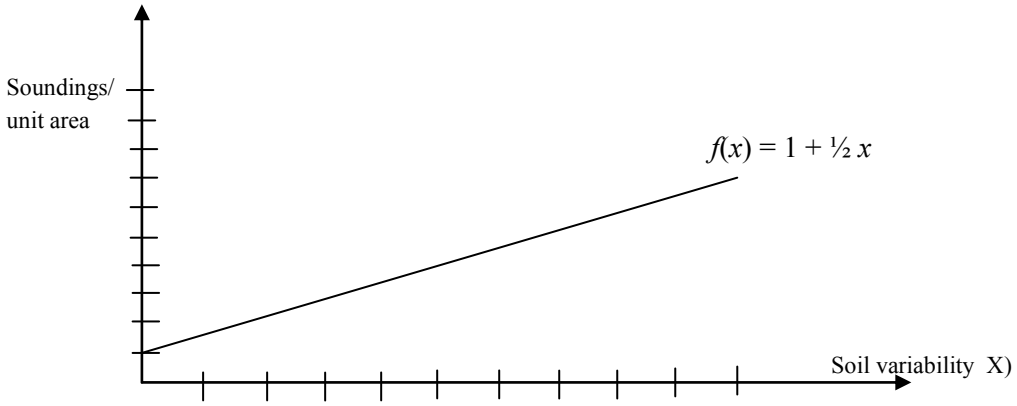


Figure 5. Objective function f(x), point to point approach

The obtained "objective function" is given by $f(x) = (1 + \frac{1}{2}x)$, this function permits to calculate the number of soundings based on the degree of variability.

Example: Expert 1 gave the following number of soundings upon a "low" soil variability, with a degree of belief m_i for each case.

Soil variability = [0,3]; $n_1 = 2$; $m_1 = ¼$

Soil variability = [1,5]; $n_2 = 3$; $m_2 = ¼$

Soil variability = [2,4]; $n_3 = 3$; $m_3 = ½$

Using $f(x)$ we obtain for the number of soundings / unit area (40 × 40m²) as shown in fig.6 :

Number of soundings [1, 2.5], $m_1 = 1/4$;
[1.5, 3.5], $m_2 = 1/4$,
[2, 3], $m_3 = 1/2$

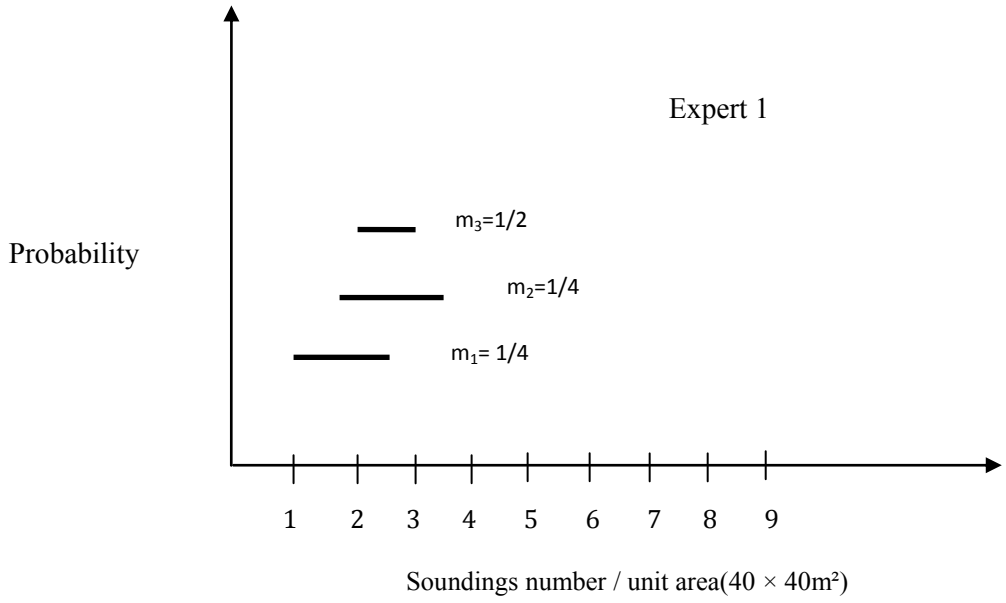


Figure 6. Construction of a random set (Number of soundings) upon $f(x) = (1 + \frac{1}{2}x)$, according to expert 1

The next step is the construction of the upper and lower probabilities (cumulative distribution function) which will give an estimation of the soundings number on site according to one expert.

From the constructed random set we assemble upper and lower probabilities.

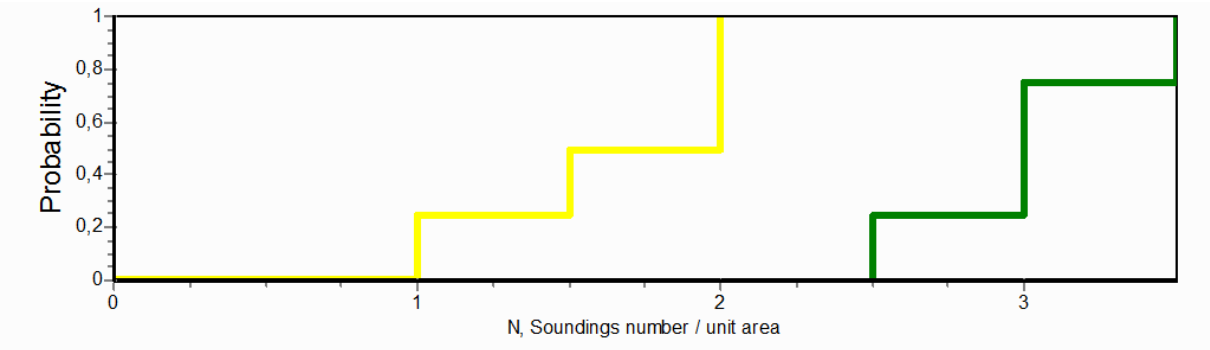


Figure 7. The constructed Upper and Lower probabilities according to one expert.

The suitable number of soundings is around 2 in this case. This example was run to illustrate the way we use the random set-based approach.

When considering more than one expert an aggregation is necessary. According to Hall et al. (2004) when there are "n" alternative random sets describing some variable *x*, each one corresponding to an independent source of information (expert in this case) for each focal element *A*,

$$m(A) = \frac{1}{n}\sum m_i(A)$$

In the case when random sets (A_i, m_i) : $i = 1,\ldots, n$ from different sources do not contain the same focal elements a merged random set is obtained using union and $m(A)$ is obtained from the previous equation.

There are other combination rules such as "Dempster rule", Yager's modified Dempster's rule, Inkagi's unified rule of combination, Dubois and Prade's rule and others (Sentz, 2002).

The Dempster's rule combines multiple belief functions through their basic probability assignments (m). The combination (joint) m_{12} is calculated from the aggregation of two pba (probability basic assignment) m_1 and m_2 as it follows:

$$m_{12}(A) = \frac{\sum_{B\cap C=A} m_1(B)m_2(C)}{1-K} \quad \text{when} \quad A \neq \emptyset$$

$$m_{12}(\emptyset) = 0$$

where

$$K = \sum_{B\cap C=\emptyset} m_1(B).m_2(C).$$

The result of aggregation is still a cumulative function of distribution which could be used as a tool for decision making.

Conclusion

A random set-based approach is introduced to estimate the number of soundings for geotechnical investigations. As a first step we used only one parameter (soil variability) to construct the random sets. For certain soil variability degree between [1,10] the expert will give the corresponding number of soundings (with a degree of belief m_{ij}). Using Eurocode7 recommendations for site soundings, we constructed an "objective function" f(x) to rely "soil variability" to the number of soundings. This function permits constructing the random set and obtain the number of soundings by unit area for each expert (engineer). The construction of upper and lower probabilities permits us optimizing the number of soundings to carry out on site. The proposed system needs to be run on real sites, with different experts opinions and then aggregate them together to obtain a suitable upper and lower probabilities for estimating the number of soundings on site.

References

Alvarez Marín D. A., 2008. *Infinite Random Sets and Applications in Uncertainty Analysis.* Phd Thesis. Unit of Engineering Mathematics. University of Innsbruck, Austria.

Boumezerane D., Belkacemi S. Zlender B. Fuzzy-sets decision-support system for geotechnical site soundings. *Acta Geotechnica Slovenica*, 2011/1: 51-63

Cambefort H. *Géotechnique de l'Ingénieur, reconnaissance des sols*. Eyrolles Editeur - Paris, 1980.

Clayton CRI, Matthews MC, Simons NE. HYPERLINK "http://www.geotechnique.info/" \t "_blank" Site Investigation. Department of Civil Engineering, University of Surrey, 2005.

FHWA (2002a). Sabatini PJ, Bachus RC, Mayne PW, Schneider JA, Zettler TE. *Geotechnical Engineering Circular 5 (GEC5) - Evaluation of Soil and Rock Properties*, Report No FHWA-IF-02-034. Federal Highway Administration, U.S. Department of Transportation, 2002.

Hunt RE. *Geotechnical investigation methods: a field guide for geotechnical engineers*. Taylor & Francis Group LLC, 2007.

Magnan JP. Quelques spécificités du problème des incertitudes en géotechnique. *Revue Française de Géotechnique*, 2000; 93: 3-9.

Oberguggenberger M. Mathematics of uncertainty; Models, Methods and Interpretations; in Felin W., Oberguggenberger M., Vieider R. *Analyzing Uncertainty in Civil Engineering*, Springer Verlag 2005

Sentz K. & S. Ferson, 2002. *Combination of Evidence in Dempster- Shafer Theory*, SANDIA Report, SAND2002-0835, April 2002.

U.S. Corps of Engineers. *Geotechnical Investigation*, Report N° EM 1110-2-2504, 1994.

Markovian Analysis of Aging Distributed Systems Remaining Life

A.V. Bushinskaya[1)], S.A. Timashev[2)],
[1)] *Science and Engineering Center "Reliability and Safety of Large Systems and Machines" Ural Branch, Russian Academy of Sciences, Ekaterinburg 620049, Russia, timashevs@cox.net*
[2)] *Science and Engineering Center "Reliability and Safety of Large Systems and Machines" Ural Branch, Russian Academy of Sciences, Ekaterinburg 620049, Russia, bushinskaya@gmail.com*

Abstract: Reliable assessment of the remaining life of distributed systems such as pipeline systems (PS) with defects plays a crucial role in solving the problems of their integrity.

Authors propose a methodology which allows estimating the *random* residual time (remaining life) of transition of a PS from its current state to a critical or limiting state, based on available information on the sizes of a set of growing defects found during an in line inspection (ILI), followed by verification or direct assessment (DA).

PS with many actively growing defects is a physical *distributed* system, which transits from one state to another. This transition finally leads to (conditional) failure of its components, each component physically being a defect. Such process can be described by a Markov process.

The degradation of the PS (measured as monotonous deterioration of its failure pressure) is considered as a non-homogeneous pure death Markov process (NPDMP) of the continuous time and discrete states type. Failure pressure is calculated using one of the internationally recognized pipeline design codes: B13G, B31Gmod, DNV, Battelle and Shell-92.

The NPDMP is described by a system of non-homogeneous differential equations, which allows calculating the probability of defects failure pressure being in each of its possible states. On the basis of these probabilities the conditional remaining life of defects is calculated, i.e. the time from the last ILI or DA to the moment of conditional failure. In other words, until the moment of time $t = t_f$, which is a *random variable*, when the failure pressure of pipeline defect $P_f(t_f) \leq P_{op}$, where P_{op} is the operating pressure. The developed methodology was successfully applied to real life case, which is presented and discussed.

Keywords: remaining life, residual strength, Markov processes, pipeline systems.

1. Pure Death Markov Process of Pipeline Strength Degradation

A PS with many actively growing defects is a physical *distributed* system, which transits from one state to another. This transition finally leads to (conditional) failure of its components, each component physically being a defect. In this paper such process is described by a Markov process.

Markov processes are described by systems of differential equations and do not depend on the nature of objects and their physical properties. In this sense they are universal and are widely used in various fields of science and technology: nuclear physics, biology, astronomy, queuing theory, reliability theory, etc. (Ventzel, 1969; Gnedenko, 1988; Timashev, 1982; Feller, 1967; Bolotin, 1988; Gnedenko, Belyaev and Solovyev, 1965; Bogdanoff and Kozin, 1989).

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

Apply the empirical Markov model to describing the degradation (reduction of the residual strength) of the pipeline, which is assessed by the international design codes described below.

1.1 Degradation Model of Residual Strength of Pipeline Cross Section with a Defect

Consider the cross section of pipeline with a defect. The burst pressure of a performing pipeline defective cross section at some fixed time t is a random variable (RV) $P_f(t) \geq P_{op}$, where Pop is the operating pressure in the pipeline. The burst pressure $P_f(t)$ can be assessed using one of the five internationally recognized pipeline design codes: B13G (ANSI/ASME B31G, 1991), B31Gmod (Kiefner and Vieth, 1989), DNV (DNV-RP-F101, 2004), Battelle (Stephens and Leis, 2000) and Shell-92 (Ritchie and Last, 1995).

Divide the possible range of change of the burst pressure of a pipeline defective cross section $\left(P_{op}; P_f(0)\right]$ into *M*-1 non-overlapping intervals $I_i\ (i=1,..,M)$. Here $P_f(0)$ is the burst pressure of defect at the initial moment of time $t = 0$.

Take the last interval (conditional failure state), which includes the lowest values of burst pressure, as equal to $(0; P_{op}]$.

The burst pressure of the defective cross section can only monotonically decrease over time, i.e., at random moments of time transit from *i*-th state only to the (*i* + 1)-th state, where the state is one of the intervals $I_i\ (i=1,..,M)$. Consider this process in detail.

Assume that at some moment of time *t* the value of burst pressure $P_f(t)$ is in the *i*-th state (interval), and assume that the probability that during a small time span Δt the RV $P_f(t)$ will move to the next (*i* + 1)-th interval is approximately equal to $\mu_i(t)\Delta t + o(\Delta t)$. Here quantity $\mu_i(t) \geq 0$ [1/time] is called *intensity of transition* from the *i*-th state to the (*i* + 1)-th state, and determines the rate of decrease of RV $P_f(t)$. At the same time the quantity $\mu_i(t)$ does not depend on how the RV $P_f(t)$ arrived at its current state. This means that the process in consideration is a Markov process, i.e. when the future depends on the past only via the present. The probability that during the small time span of Δt more than one transition will occur is of higher order of magnitude smaller than Δt.

The system of differential equations (SDE) that describes this process has the form:

$$\begin{cases} \dfrac{\mathrm{d}P_1(t)}{\mathrm{d}t} = -\mu_1(t)P_1(t) \\ \dfrac{\mathrm{d}P_i(t)}{\mathrm{d}t} = \mu_{i-1}(t)P_{i-1}(t) - \mu_i(t)P_i(t), \ (i=2,..,M-1) \\ \dfrac{\mathrm{d}P_M(t)}{\mathrm{d}t} = \mu_{M-1}(t)P_{M-1}(t) \end{cases} \tag{1}$$

where $P_i(t)$ is the probability that the RV $P_f(t)$ is in the *i*-th state at the moment of time *t*, $\mu_i(t)$ is intensity of transition from the *i*-th state to the (*i* + 1)-th state.

System (1) describes the non-homogeneous Markov process of pure death which is characterized by discrete number of states and continuous time. For such process the flow of events, which transit the RV $P_f(t)$ from one state to another, is a non-stationary Poisson process (Ventzel, 1991). Then, according to the definition of a non-stationary Poisson flow of events, the expected number of events, which transit the value of burst pressure from one state to another within a time interval (0, *t*], is calculated by formula:

$$\rho(t) = \int_0^{\Delta t} \mu(\tau)\,\mathrm{d}\tau \tag{2}$$

The physical meaning of the intensity (density) $\mu(t)$ of the non-stationary Poisson flow of events (transitions from one state to another) is the the average number of events per unit time for an elementary time interval (0, Δ*t*]. The quantity $\rho(t)$ is the average number of states through which the random variable $P_f(t)$ passes within small time interval (0, Δ*t*]. The intensity $\mu(t)$ can be expressed by any nonnegative function and has the dimension [1/time] (Ventzel, 1991).

Consequently, the quantity $\mu(t)$ may be associated with the rate of change of RV $P_f(t)$ as follows:

$$\mu(t) = -\frac{P_f'(t)}{\Delta I} \tag{3}$$

Where ΔI is the length of intervals, $P_f'(t)$ is the derivative of the function $P_f(t)$ with respect to time at time *t*. The minus in this formula is due to the fact that the derivative of monotonically decreasing function has negative values in the whole domain of its definition.

Hence, the system (1) can be rewritten as:

$$\begin{cases} \dfrac{\mathrm{d}P_1(t)}{\mathrm{d}t} = -\mu(t)P_1(t) \\ \dfrac{\mathrm{d}P_i(t)}{\mathrm{d}t} = \mu(t)P_{i-1}(t) - \mu(t)P_i(t), \ (i = 2,..,M-1) \\ \dfrac{\mathrm{d}P_M(t)}{\mathrm{d}t} = \mu(t)P_{M-1}(t) \end{cases} \tag{4}$$

It is obvious that at the initial moment of time *t* = 0 RV $P_f(0) \in I_1$. Hence the initial conditions for the SDE (4) have the form:

$$P_1(0) = 1,\ P_i(0) = 0,\ (i = 2,..,M) \tag{5}$$

Solving the SDE (4) by the method of variation of the constant, with initial conditions (5), obtain:

$$\begin{cases} P_1(t) = \exp\{-\rho(t)\} \\ P_i(t) = \dfrac{\rho^{i-1}(t)}{(i-1)!} \cdot \exp\{-\rho(t)\},\ i = 2,..,M-1 \\ P_M(t) = 1 - \left[\exp\{-\rho(t)\} + \sum\limits_{i=2}^{M-1} \dfrac{\rho^{i-1}(t)}{(i-1)!} \cdot \exp\{-\rho(t)\} \right] \end{cases} \tag{6}$$

Where $P_i(t)$ is the probability that the value of RV $P_f(t)$ is in the *i*-th state at time *t*, and $\rho(t)$ is calculated by formula:

$$\rho(t) = \int_0^t \mu(\tau)\,\mathrm{d}\tau = -\int_0^t \frac{P_f'(\tau)}{\Delta I}\,\mathrm{d}\tau = \frac{P_f(t) - P_f(0)}{\Delta I} \tag{7}$$

1.2 Degradation Model of Distributed Pipeline System Residual Strength with a Finite Set of Defects

The Markov model for a distributed PS with a finite set of defects is constructed on the same reasoning as for single defective cross section of a pipeline.

Assume that at the initial moment of time $t = 0$, using some diagnostic tools, $N(0)$ defects were found, their geometric parameters identified, and for each defect an estimate of burst pressure was obtained. Now calculate the frequency of occurrence of the burst pressure in each of the intervals (states) at time $t = 0$:

$$P_i(0) = p_i^* = \frac{n_i(0)}{N(0)} \quad (i = 1,2,..,M) \tag{8}$$

Where $n_i(0)$ is the number of defects, which burst pressure at the moment of time $t = 0$ is in the *i*-th interval, $N(0)$ is the overall number of defects.

Expressions (8) are the initial conditions for the SDE (4). Solution of SDE (4) with initial conditions (8) has the following generalized form:

$$\begin{cases} P_1(t) = p_1^* \exp\{-\rho(t)\} \\ P_k(t) = \sum\limits_{i=1}^{k} \left[p_i^* \dfrac{\rho^{k-i}(t)}{(k-i)!} \right] \cdot \exp\{-\rho(t)\},\ k = 2,..,M-1, \\ P_M(t) = 1 - \left[p_1^* + \sum\limits_{k=2}^{M-1} \left(\sum\limits_{i=1}^{k} \left[p_i^* \dfrac{\rho^{k-i}(t)}{(k-i)!} \right] \right) \right] \exp\{-\rho(t)\} \end{cases} \tag{9}$$

Where $P_i(t)$ is the probability that the value of burst pressure $P_f(t)$ of the PS with a set of defects is in the *i*-th state at time *t*.

To determine the function $\rho(t)$ use the quantile approach, which holds a special place in the arsenal of probabilistic tools for analysis of statistical data. By controlling the quantile, it is possible to assess the probability of failure (POF) of the entire PS with required reliability (confidence level).

For a PS with defects the value of $\rho(t)$ at a given moment of time *t* is defined as the sample quantile of the order α of the set of values $\rho_i(t)$ $(i=1,..,N)$ comprised of all defects found by the ILI.

Here $\rho_j(t)$ is the value of $\rho(t)$ of the *j*-th defect at considered time *t*, calculated using formula (7). In other words, $\rho(t)$ = $\rho^{(k)}(t)$, where $\rho^{(k)}(t)$ is the *k*-th order statistic of ordered series $\rho^{(1)}(t) \le \rho^{(2)}(t) \le,..,\le \rho^{(N)}(t)$ of the sample values of $\rho_1(t)\rho_2(t),..,\rho_N(t)$ at time *t*, and $k=[\alpha N+1]$, where […] is the integer part. The order α is chosen out of safety reasons.

The probability of finding the burst pressure in the last state (in interval I_M) is the *conditional probability of failure* of the defect or the pipeline system under the "rupture" type failure.

The conditionality of probability of failure is due to possible reserves of the residual strength, "hidden" in the existing regulations, which have a variety of embedded safety factors.

Markov models describing the growth rate of defects (or degradation of the pipe wall thickness), which fail by "leaking" are given in papers (Timashev, Malyukova, Poluian and Bushinskaya, 2008; Poluyan, Bushinskaya, Malyukova and Timashev, 2009; Valor, Caleyo, Alfonso and others, 2007).

2. Assessment of system residual life

Assessment of residual life is based on the above-constructed Markov pure death process.

Assume that the Markov process of residual strength degradation of pipeline system or a defective cross section of pipeline is successfully constructed, along the lines described above. Now denote as T_i the time the burst pressure is in the subset of states $S_i=[I_1,..,I_i]$, $i=1,..,M-1$. According to (Ventzel, 1991) the probability distribution function PDD of RV T_i is equal to:

$$F_{T_i}(t)=P(T_i<t)=\tilde{P}_i(t) \tag{10}$$

Where $\tilde{P}_i(t)$ is the probability that at time t the burst pressure will transit from set S_i to the subset of states $S_{i+1}=\{I_{i+1},..,I_M\}$, where *M* is the number of the subset states. The set S_{i+1} is considered to be absorbing (i.e., that burst pressure, once entering the set, cannot leave it).

Consequently, the PDD of RV T_i is equal to:

$$F_{T_i}(t)=P_{i+1}(t) \tag{11}$$

Where $P_{i+1}(t)$ is the probability that the value of burst pressure is in the (i + 1)-th state at time t, calculated using formulas (6) or (9).

Considering that the set S_{i+1} is absorbing, the PDD of RV T_i will be:

$$f_{T_i}(t) = F'_{T_i}(t) = \mu(t) P_i(t) \tag{12}$$

Where $\mu_i(t)$ is the intensity of transition from the i-th state to the (i + 1)-th state, calculated by formula (3).

Knowledge of the PDD of RV T_i allows finding its numerical characteristics:

$$\begin{aligned} M[T_i] &= \int_0^\infty t \cdot f_{T_i}(t)\,\mathrm{d}t \\ D[T_i] &= \int_0^\infty t^2 \cdot f_{T_i}(t)\,\mathrm{d}t - M[T_i] \end{aligned} \tag{13}$$

Where $M[T_i]$ is the mean (average time of the burst pressure being in the subset of states S_i); $D[T_i]$ is its variance.

Define the average time of burst pressure being in each subset of states. The average time of burst pressure being in the first state:

$$M[T_1] = \int_0^\infty t \cdot \mu(t) P_1(t)\,\mathrm{d}t = \int_0^\infty \exp\{-\rho(t)\}\,\mathrm{d}t - t \cdot \exp\{-\rho(t)\} \tag{14}$$

The average time of burst pressure being in the first and second state, i.e., the average time at which the burst pressure will move into the third state:

$$M[T_2] = \int_0^\infty t \cdot \mu(t) P_2(t)\,\mathrm{d}t = \frac{1}{2}\int_0^\infty t \cdot \rho(t)\mu(t)\exp\{-\rho(t)\}\,\mathrm{d}t \tag{15}$$

In generalized form, the average time of burst pressure being in a subset of states $S_i = [I_1,..,I_i]$, $i = 2,..,M-2$, i.e., the time at which the burst pressure will move into the (i + 1)-th state:

$$M[T_i] = \frac{1}{(i-1)!}\int_0^\infty t \cdot \mu(t)\rho^{i-1}(t)\exp\{-\rho(t)\}\,\mathrm{d}t,\ (i = 2,..,M-2) \tag{16}$$

Then, the time the burst pressure is dwelling in the first (M –1) states, i.e. the *average* time after which the burst pressure will move into the last, *conditional* failure state, is determined by the formula:

$$M[T_{M-1}] = \frac{1}{(M-2)!}\int_0^\infty t \cdot \mu(t)\rho^{M-2}(t)\exp\{-\rho(t)\}\,\mathrm{d}t \tag{17}$$

Thus, formula (17) allows assessing the conditional residual life of a single defective cross section of pipeline, as well as of a distributed PS with a finite set of defects, i.e., the time from the last inspection to

the time of conditional failure $t = t_f$, at which the burst pressure $P_f\left(t_f\right) \le P_{op}$, where P_{op} is the pipeline operating pressure.

3. Example of application of the proposed methodology

Consider the example of two defective cross sections of the pipeline. Calculate the residual life by the "gap" of defective sections of the pipeline, which fail in the rupture mode and whose parameters are defined in Table I.

Table I. Initial data. Pipeline parameters

Parameters	Symbol	Value
External pipeline diameter, mm	*D*	325.0
Pipeline wall thickness, mm	*wt*	9.0
Specified minimum yield strength of pipe material , MPa	*SMYS*	245.0
Minimum tensile strength of the pipe material, MPa	*UTS*	410.0
Design operating pressure, MPa	P_{op}	6.4

Consider two defect of the “metal loss” type with parameters as defined in Table II.

Table II. Initial data. Defects parameters

Parameters	Symbol	Value, mm
Parameters of defect #1		
Depth	d_{01}	2.25
Length	l_{01}	246.00
Parameters of defect #2		
Depth	d_{02}	5.62
Length	l_{02}	70.00

To estimate the parameters of the defects growth rates use the method of estimating the defects corrosion rate according to their conditional maximal growth, as described in (Enterprise Standard 0-03-22-2008, 2008).

According to this method the conditional maximal size of the defects depth and length were calculated, with probability γ = 0.95, as well as the growth rates, which turned out to be equal to $a_d = 0.20$ mm/years for defects depth, and $a_l = 2.34$ mm/year for defects length.

Assume that the growth rates of the corrosion defects are close to linear. Then the sizes of defect parameters at time *t* are ruled by formulas:

$$d(t) = d_0 + a_d \cdot t$$
$$l(t) = l_0 + a_l \cdot t \tag{18}$$

Where d_0, l_0 are the defect depth and length at the time of inline inspection of the pipeline; a_d, a_l are the radial and longitudinal-axis corrosion rates, respectively.

Consider the process of residual strength degradation of defective pipeline cross sections using the B31Gmod code, according to which (Kiefner and Vieth, 1985) the burst pressure is calculated by formula:

$$P_f = \frac{2wt\left(SMYS + 68.95МПа\right)}{D} \cdot \frac{\left(1 - 0.85\frac{d(t)}{wt}\right)}{\left(1 - 0.85\frac{d(t)}{wt \cdot M_{B31G\,\mathrm{mod}}}\right)} \tag{18}$$

Where $M_{B31G\,\mathrm{mod}}$ is the Folias factor:

$$M_{B31G\,\mathrm{mod}}(t) = \begin{cases} \sqrt{1 + \frac{0.6275 \cdot l^2(t)}{D \cdot wt} - \frac{0.003375 \cdot l^4(t)}{\left(D \cdot wt\right)^2}}, & l(t) \le 7.07\sqrt{D \cdot wt} \\ \frac{0.032 \cdot l^2(t)}{D \cdot wt} + 3.3, & l(t) > 7.07\sqrt{D \cdot wt} \end{cases} \tag{19}$$

Calculate the ranges of the burst pressure of defects under consideration. Calculate the right boundary for the burst pressure of each defect. To do so, calculate the burst pressure at time $t = 0$, using eqn. (18). For each defect under consideration:

$$P_{f_1}(0) = 15.49\,\mathrm{MPa},\ P_{f_2}(0) = 13.73\,\mathrm{MPa}$$

Where $P_{f_1}(0)$ is the burst pressure of defect #1 at time $t = 0$; $P_{f_2}(0)$ is the burst pressure of defect #2 at moment of time $t = 0$.

Divide the possible range of the burst pressure into $M = 10$ non-overlapping intervals. The first interval (conditional failure state) is equal to $\left(0; P_{op}\right] = \left(0; 6.40\right]$. The length ΔI of other nine intervals is determined by dividing the value of the expression $P_f(0) - P_{op}$ by the number of states (intervals):

$$\Delta I_1 = 0.91MPa,\ \Delta I_2 = 0.73MPa,$$

Where ΔI_1 is the length of is first nine states of burst pressure of defect #1, is the length of first nine states of burst pressure of defect #2.

The obtained intervals (states) for the considered defects are presented in Table III.

Table III. States(intervals) of burst pressures of the defects

Interval #	Boundaries of intervals, MPa	
	defect #1	defect #2
10	[0.00; 6.40)	[0.00; 6.40)
9	[6.40; 7.31)	[6.40; 7.13)
8	[7.31; 8.21)	[7.13; 7.86)
7	[8.21; 9.12)	[7.86; 8.59)
6	[9.12; 10.03)	[8.59; 9.32)
5	[10.03; 10.94)	[9.32; 10.06)
4	[10.94; 11.84)	[10.06; 10.79)
3	[11.84; 12.75)	[10.79; 11.52)
2	[12.75; 13.66)	[11.52; 12.25)
1	[13.66; 14.57)	[12.25; 12.98)

Estimate the probability of failure (of the rupture type) and the residual life of the defects. Using formulas (6) find the probability of burst pressure being in each state for three future moments of time t = 1,2, and 3 years. The results are given in Table IV.

In Table IV the probabilities of finding the burst pressure in the last (tenth) failure state are the *probabilities of defects failing in the rupture mode.*

Table IV. Predicted probabilities of defects burst pressures being in each state

Interval (state) #	The probability of finding the burst pressure in each state for					
	defect #1			defect #2		
	t = 1 year	t = 2 years	t = 3 years	t = 1 year	t = 2 years	t = 3 years
1	0.74	0.54	0.40	0.59	0.34	0.19
2	0.22	0.33	0.37	0.31	0.37	0.32
3	0.03	0.10	0.17	0.08	0.20	0.26
4	$3.45\cdot10^{-3}$	0.02	0.05	0.02	0.07	0.14
5	$2.62\cdot10^{-4}$	$3.17\cdot10^{-3}$	0.01	$1.92\cdot10^{-3}$	0.02	0.06
6	$1.59\cdot10^{-5}$	$3.89\cdot10^{-4}$	$2.24\cdot10^{-3}$	$2.03\cdot10^{-4}$	$4.17\cdot10^{-3}$	0.02
7	$8.08\cdot10^{-7}$	$3.97\cdot10^{-5}$	$3.45\cdot10^{-4}$	$1.79\cdot10^{-5}$	$7.42\cdot10^{-4}$	$5.33\cdot10^{-3}$
8	$3.51\cdot10^{-8}$	$3.47\cdot10^{-6}$	$4.56\cdot10^{-5}$	$1.35\cdot10^{-6}$	$1.14\cdot10^{-4}$	$1.26\cdot10^{-3}$
9	$1.33\cdot10^{-9}$	$2.66\cdot10^{-7}$	$5.27\cdot10^{-6}$	$8.95\cdot10^{-8}$	$1.54\cdot10^{-5}$	$2.58\cdot10^{-4}$
10	$\mathbf{4.65\cdot10^{-11}}$	$\mathbf{1.92\cdot10^{-8}}$	$\mathbf{5.96\cdot10^{-7}}$	$\mathbf{5.55\cdot10^{-9}}$	$\mathbf{2.06\cdot10^{-6}}$	$\mathbf{5.64\cdot10^{-5}}$

To estimate the remaining life of the defects estimate the average time of burst pressure being in each subset of states. Consider the following subsets of states:

$$S_1 = \{I_{i1}\}$$
$$S_2 = \{I_{i2}, I_{i1}\}$$
$$S_3 = \{I_{13}, I_{i2}, I_{i1}\}$$
$$.....$$
$$S_{M-1} = \{I_{1M-1}, I_{iM-2}, .., I_{i1}\}$$

where I_{ij} is the *j*-th state of the burst pressure of *i*-th defect, $i = 1,2$, $j = 1,..,9$.

Calculate using formula (16) the average time of burst pressures being in these subsets of states. The obtained values are given in Table V and in Figure, which is a graphical interpretation of this table.

The last row of Table V gives the values of the residual life of the pipeline defective cross sections, i.e., the time which will elapse from the time of the last inspection of the pipeline to the time of conditional failure of pipeline defect ($P_f\left(t_f\right) \le P_{op}$).

Table V. Average time of burst pressures being in each subset of states for

subsets of states #	defect #1	defect #2
1	3.2	1.8
2	6.2	3.5
3	9.2	5.2
4	12.0	6.7
5	14.8	8.2
6	17.5	9.7
7	20.1	11.1
8	22.7	12.5
9	**25.2**	**13.9**

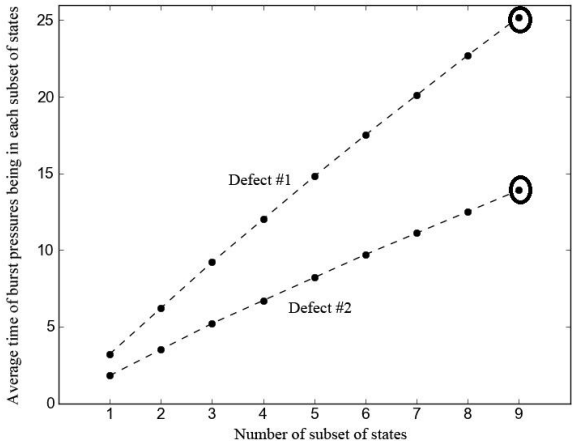


Figure . Average time of burst pressures being in each subset of states S_i, $i = 1,..,9$

Conclusion

The developed model of residual strength degradation for a single defective cross section of pipeline and a pipeline system with a finite set of discrete defects in the form of a non-homogeneous Markov pure death process, allows estimating the probability of failure/reliability and residual life of a single defect and of a distributed pipeline system with multiple defects.

References

ANSI/ASME B31G-1991. Manual for Determining the Remaining Strength of Corroded Pipelines. National Standard of USA. – New York: ASME. – 1991. – 140 p.

Bogdanoff J. and Kozin F. Probabilistic models of cumulative damage. John Wiley & Sons NY, NY 1985 344 p.

Bolotin V.V. Predicting life of machines and structures. – Moscow: Mashinostroenie, 1988. – 240 p. In Russian.

DNV-RP-F101. Corroded pipelines. Recommended Practice. – Norway: Det Norske Veritas, 2004. – 42 p.

Feller V. Feller W. An Introduction to Probability Theory and Its Applications. Vol.1, 2. – Moscow: Mir, 1984. – 1266 p. In Russian.

Gnedenko B.V. Course on Probability Theory Manual.-6-th ed. Moscow: Nauka, 1988 – 448 p. In Russian.

Gnedenko B.V., Belyaev J.K. and Soloviev A.D. Mathematical methods in reliability theory. The main characteristics of reliability and statistical analysis. – Moscow: Nauka, 1965. – 524 p. In Russian.

Kiefner J.F. and Vieth P.H. A Modified Criterion for Evaluating the Remaining Strength of Corroded Pipe // AGA Pipeline Research Committee. Report PR 3-805. – 1989. – 78 p.

Poluyan L.V., Bushinskaya A.V., Malyukova M.G. and Timashev S.A. Reliability Based Inspection and Maintenance of Pipelines with Markov Type Corrosion Defects Growth // ICOSSAR. – Japan. – 2009. Paper #ICOSSAR09-0156.

Ritchie D. and Last S. Burst Criteria of Corroded Pipelines - Defect Acceptance Criteria // Proceedings of the EPRG/PRC 10th Biennial Joint Technical Meeting on Line Pipe Research. – Cambridge, UK. – 1995. – Paper

Stephens D.R. and Leis B.N. Development of an Alternative Criterion for Residual Strength of Corrosion Defects in Moderate to High-Toughness Pipe // Proceedings of the Third International Pipeline Conference. – Calgary, Canada, American Society of Mechanical Engineers. – 2000. – P. 781-792.

Enterprise Standard #0-03-22-2008. Technological safety of operation of pipelines pumping gas containing sulphur hydrogen, and pipelines carrying unstable condensate. «OrenburgGasprom», VNIIGAZ, ANO NTP «Technological Park, Orenburg State University», 2008, In Russian

Timashev S.A. Reliability of Large Mechanical Systems. – M.: Nauka, 1982 . – 183 p. In Russian

Timashev S.A., Malyukova M.G., Poluyan L.V. and Bushinskaya A.V. Markov Description Of Corrosion Defects Growth and Its Application To Reliability – Based Inspection and Maintenance of Pipelines // Proceedings of IPC Conference. – 2008. – Paper № IPC2008-64545

Valor A., Caleyo F., Alfonso L. and others. Stochastic modeling of pitting corrosion: a new model for initiation and growth of multiple corrosion pits // Corrosion Science. – 2007. – № 49(2) . – P. 559-579.

Ventzel E.S. Theory of Probability. – M.: Nauka, 1969. – 576 p. In Russian

Ventzel E.S. Theory of Random Processes and its Engineering Applications. – M.: Nauka, 1991. – 384 p. In Russian

Simulation of high variable random processes through the spectral-representation-based approach

Pierfrancesco Cacciola
School of Environment and Technology, University of Brighton, Brighton, BN24GJ,
p.cacciola@brighton.ac.uk

Abstract: In this paper the traditional spectral-representation-method for simulating stochastic processes is revisited. A modification aimed to control the variability of the simulated samples of the random process is proposed. Specifically, in order to avoid that the simulated samples possess similar Fourier spectrum, it is proposed to randomize the power spectral density function through a pass band filter with random parameters. The filters is selected in such a way the ensemble averaged power spectral density determined by the samples will match the original power spectrum, but each individual sample will possess different frequency distributions. Comparison between the traditional simulation technique and the new one proposed in this paper will be also discussed. Results show that despite the ensemble averaged power spectral density is the same, related quantities, such as the distribution of peaks, will be significantly different highlighting the needing to consider the variability of frequency distributions when stochastic models are calibrated from experimental data.

Keywords: Monte Carlo simulation, variability, spectral-based representation

1. Introduction

Several problems in science and engineering involve uncertain quantities generally modeled as stochastic processes or stochastic fields. The evaluation of the response of structures to seismic or wind actions, the reliability of composite materials, the action potential generation in neurons, the drifting of particles suspended in fluids, the distribution of nanoparticles, the mutation of populations subjected to genetic drift are just few examples of the plethora of problems involving stochastic modeling. The solution of the stochastic problem is generally pursued through *ad hoc* strategies including simulation techniques, stochastic calculus and perturbation methods. The Monte Carlo simulation seems to be, up to now, the only universal method able to cope with different problems using a systematic approach: i) simulation of samples of the stochastic process (field), ii) solution of the deterministic problem for each individual sample, iii) statistical evaluation of the results. The simulation of the samples represents the first step and requires the definition of a proper stochastic model representing reliably the uncertainty embedded in the problem. Gaussian models are certainly the most used models used to represent and simulate stochastic processes especially for engineering applications. Accordingly, the Gaussian random process is fully defined by the mean and by power spectral density function (or alternatively by the mean and the autocorrelation function). Therefore, after defining a suitable model of the power spectral density the samples can be simulated through various techniques such as the spectral representation method (SRM), Karhunen-Loeve (KL) decomposition, ARMA models. In the framework of simulation of Gaussian processes (fields), the spectral representation method based on the superposition of harmonics with random

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

phases (see e.g. Shinozuka and Deodatis, 1988) is certainly one of the most diffused approaches. This method has been applied in the last three decades to solve efficiently several problems in the fields of civil, mechanical and aeronautical engineering.

The increasing number of data collected from laboratory experiments or recorded through measuring stations and the sustained progress in experimental measurement techniques is compelling the research in stochastic modeling to validate (and eventually improve) models and the methodologies traditionally adopted by the comparison of observed and simulated random data. In the field of earthquake engineering, as an example, the increasing number of strong-motion networks installed worldwide revealed that the current methodologies for simulating artificial earthquakes possess the drawback that the simulated time-histories do not manifest large variability of the seismological parameters observed for natural accelerograms. As a consequence, the dispersion of the consequent structural response analysis can be underestimated. This issue has been recently addressed by calibrating some of parameters embedded in traditional stochastic models by using experimental data. Specifically, Pousse et al. (2006) proposed a methodology for simulating accelerograms through a stochastic approach by using the K-Net Japanese database. The basic idea of this approach is to define an evolutionary power spectral density function possessing random variables determined through empirical attenuation equations. Recently, Rezaeian and Der Kiureghian (2010) proposed a method for simulating synthetic ground motion time histories through a parameterized stochastic model based on a modulated filtered white-noise process. The parameters of the model are random variables calibrated on a set of recorded earthquakes. Cacciola and Zentner (2012) introduced the natural variability of ground motion accelerograms in the model through a pertinent evolutionary power spectrum with random coefficients. The distribution of the random coefficients has been determined by the further matching with a given mean ± standard deviation response spectra evaluated using an European earthquake database.

The above recent contributions addressed the issue of the discrepancy between the experimental and simulated samples including random parameter in the model. This approach leads to more realistic simulation that in general manifest greater dispersion in peak values and in general in the energy distribution. In this paper, based on the same motivations, the traditional spectral-representation-method for simulating stochastic processes is revisited. A modification aimed to control the variability of the simulated samples of the random process is proposed. Specifically, in order to avoid that the simulated samples possess the same Fourier spectrum, it is proposed to filter the power spectral density through a pass band filter with random parameters. The filters is selected in such a way the ensemble averaged power spectral density will match the original power spectrum, but each individual samples will possess different frequency distributions. Comparison between the traditional simulation technique and the new one proposed in this paper will be also discussed.

2. Simulation of random processes via the spectral representation method

In this section the spectral-representation-method (see e.g. Shinozuka and Deodatis, 1988, Deodatis 1996) is briefly described. The formulation will be based on stationary and non-stationary random process, but can be easily extended to homogeneous or non-homogeneous random field. Consider the one-dimensional and uni-variate Gaussian non-stationary stochastic process, $f(t)$, having zero mean:

$$E\left[f(t)\right]=0, \tag{1}$$

and the correlation function $R(t,t+\tau)$ given by the following equation

$$R(t,t+\tau)=E\left[f(t)f(t+\tau)\right]. \tag{2}$$

The process $f(t)$ can be defined alternatively by evolutionary power spectral density function defined (Priestley, 1965):

$$S(\omega,t)=\left|A(\omega,t)\right|^2 S(\omega)=\frac{1}{T}E\left[\left|\int_t^{t+T} f(\vartheta)e^{-i\omega\vartheta}\mathrm{d}\vartheta\right|^2\right] \tag{3}$$

being $(t,t+T)$ with $T<<t$ a *small* interval; also $A(\omega,t)$ and $S(\omega)$ are the (non-separable) modulating function and the (stationary) power spectral density function of the stochastic process, $f(t)$. It is noted that the ensemble average in equation (3) is not commonly used to define the evolutionary spectrum due the difficulties in its numerical evaluation related to the Uncertainty Principle and alternative representations of joint time-frequency representation are also proposed (see e.g. Spanos and Failla, 2004). Therefore, the evolutionary power spectral density is usually defined indirectly from the knowledge of the autocorrelation function, that is

$$S(\omega,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty} R(t,t+\tau)e^{-i\omega\tau}\mathrm{d}\tau. \tag{4}$$

The correlation function therefore is related to the power spectral density function through the following transformation:

$$R(t,t+\tau)=\int_{-\infty}^{\infty} A(\omega,t)A(\omega,t+\tau)S(\omega)e^{i\omega\tau}\mathrm{d}\omega. \tag{5}$$

Due to the non-stationarity of the vector process, the autocorrelation function is function of both time t and time lag τ , while the power spectral density function is a function of both frequency ω and time t . It has to be emphasized that under the hypothesis of fully non-stationary processes (non-stationary processes with amplitude and frequency modulation), the power spectral density function is a non-separable function of frequency ω and time t . For the special case of uniformly modulated non-stationary stochastic process (or generally known also as quasi-stationary process), the modulating function $A(\omega,t)$ is independent of the frequency ω , that is:

$$A(\omega,t)=A(t). \tag{6}$$

In this special case, equation (5) reduces to

$$R(t,t+\tau)=A(t)A(t+\tau)\int_{-\infty}^{\infty} S(\omega)e^{i\omega\tau}\mathrm{d}\omega. \tag{7}$$

Finally in the case in which $A(t)=1$, i.e. stationary case, the following equations hold

$$R(\tau) = \int_{-\infty}^{\infty} S(\omega) e^{i\omega\tau} \mathrm{d}\omega; \tag{8}$$

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau) e^{-i\omega\tau} \mathrm{d}\tau. \tag{9}$$

Equations (8) and (9) are generally known as the Wiener-Khintchin relationships. For the stationary case the power spectral density function can be also determined directly from experimental data

$$S(\omega) = \lim_{T\to\infty} \frac{1}{2\pi T} E\left[\left| \int_{-T/2}^{T/2} f(\vartheta) e^{-i\omega\vartheta} \mathrm{d}\vartheta \right|^2 \right]. \tag{10}$$

Once defined the power spectral density function either through experimental or physical/theoretical approaches the simulation of the sample of the non-stationary random process through the spectral representation method is performed using the following equation

$$f(t) = \sqrt{2} \sum_{j=1}^{N} \sqrt{2S(\omega_j, t)\Delta\omega} \cos\left(\omega_j t + \phi_j\right), \tag{11}$$

where

$$\omega_j = j\Delta\omega \qquad j = 1,2,...,N \tag{12}$$

with upper bound (circular) frequency $\omega_N = N\Delta\omega$ generally known also as the cut-off frequency beyond which the power spectral density function $S(\omega,t)$ is considered negligible. Also ϕ_j $(j = 1,2,...,N)$ are independent random variable distributed over the range $(0, 2\pi)$. The simulated process is asymptotically Gaussian as N tend to infinity due to the Central Limit Theorem. Also it can be shown (Shinozuka and Deodatis 1988, Deodatis 1996) that the ensemble averaged $E[f(t)]$ and $E[f(t)f(t+\tau)]$ tend to the corresponding target. Finally, samples of stationary processes can be simulated through equation (11) by substituting $S(\omega,t)$ with $S(\omega)$.

3. Enhancing the variability of the spectral representation method

It has emphasized that the simulation of random samples through the spectral representation method defined in equation (11) is based on the definition of the power spectral density function $S(\omega,t)$. Consider the power spectrum determined through the ensemble averages of experimental data (see e.g. Shillinger and Papadopoulos, 2011). Depending from the degree of randomness, the experimental data might exhibit strongly different distributions of the Fourier spectrum. In such a case, to better represent the underlying uncertainty , the simulated samples are expected not only converge to the target mean value and correlation function as equation (11) guarantees, but also it is expected that they manifest the variability of the spectrum around is expected value. In other words it is expected that the simulated data also capture the variability of the Fourier distribution of the experimental data. To enhance the variability of the simulated

random samples in this paper it is proposed to introduce a random filter acting in series with the expected value of the power spectrum, that is

$$S(\omega,t,\boldsymbol{\alpha}) = H(\omega,\boldsymbol{\alpha})S(\omega,t) \quad , \tag{13}$$

where $\boldsymbol{\alpha}$ is the vector collecting the random parameter of the filter. $H(\omega,\boldsymbol{\alpha})$ is real positive function that satisfies the following equation

$$\int_A S(\omega,t,\boldsymbol{\alpha})p_A(\boldsymbol{\alpha})\mathrm{d}\boldsymbol{\alpha} = \int_A H(\omega,\boldsymbol{\alpha})S(\omega,t)p_A(\boldsymbol{\alpha})\mathrm{d}\boldsymbol{\alpha} = S(\omega,t) \quad . \tag{14}$$

Or alternatively

$$\int_A H(\omega,\boldsymbol{\alpha})p_A(\boldsymbol{\alpha})\mathrm{d}\boldsymbol{\alpha} = 1 \quad \forall\,\omega\in\Omega \tag{15}$$

In equations (14) and (15) $p_A(\boldsymbol{\alpha})$ is the joint probability density function of random parameter of the filter. It is noted the proposed definition of the resultant random power spectrum is similar to the physical stochastic model recently defined by Li et al. (2012), but with different meaning. Embedding the proposed random spectrum in the traditional spectral representation method (SRM) the following simulation formula is herein derived

$$f_H(t) = \sqrt{2}\sum_{j=1}^{N}\sqrt{2H(\omega_j,\boldsymbol{\alpha})S(\omega_j,t)\Delta\omega}\cos\left(\omega_j t + \phi_j\right) \tag{16}$$

The samples generated by equation (16) are Gaussian as N tends to infinity due to the Central Limit Theorem and converge to the target mean $E\left[f_H(t)\right] = 0,$ and correlation function $R(t,t+\tau) = E\left[f_H(t)f_H(t+\tau)\right]$ as the traditional SRM.

Proof : the proof of convergence is based on the corresponding one proposed by Shinozuka and Deodatis (1988) and Deodatis (1996) to prove the convergence of equation (11). Due to the statistical independence of the random phase ϕ_j ($j = 1,2,...,N$) and assuming $\boldsymbol{\alpha}$ independent from ϕ_j , the expected value $E\left[f_H(t)\right]$ using equation (16) becomes

$$E\left[f_H(t)\right] = \int_A \int_0^{2\pi} \overset{N-\text{fold}}{\cdots} \int_0^{2\pi} \sqrt{2}\sum_{j=1}^{N}\sqrt{2H(\omega_j,\boldsymbol{\alpha})S(\omega_j,t)\Delta\omega}\cos\left(\omega_j t + \phi_j\right)p_A(\boldsymbol{\alpha})\left(\prod_{i=1}^{N} p_{\Phi_i}(\phi_i)\mathrm{d}\phi_i\right)\mathrm{d}\boldsymbol{\alpha}, \tag{17}$$

Where $p_{\Phi_i}(\phi_i)$ $(=1/2\pi,\ 0\le\phi_i\le 2\pi\ ;\ =0,\ \textit{otherwise})$ is the probability density function of ϕ_i . Rearranging terms in equation (17)

$$E\left[f_H(t)\right]=\sqrt{2}\sum_{j=1}^{N}\int_{A}\sqrt{H(\omega_j,\boldsymbol{\alpha})}p_A(\boldsymbol{\alpha})\mathrm{d}\boldsymbol{\alpha}\left[\int_0^{2\pi}\overset{N-\text{fold}}{\ldots}\int_0^{2\pi}\sqrt{2S(\omega_j,t)\Delta\omega}\cos\left(\omega_j t+\phi_j\right)\left(\prod_{i=1}^{N}p_{\Phi_i}(\phi_i)\mathrm{d}\phi_i\right)\right]=$$
$$=\sqrt{2}\sum_{j=1}^{N}\int_{A}\sqrt{H(\omega_j,\boldsymbol{\alpha})}p_A(\boldsymbol{\alpha})\mathrm{d}\boldsymbol{\alpha}\sqrt{2S(\omega_j,t)\Delta\omega}\left[\int_0^{2\pi}\frac{1}{2\pi}\cos\left(\omega_j t+\phi_j\right)\mathrm{d}\phi_i\right]=0 \tag{18}$$

The autocorrelation function of the simulated process

$$R(t,t+\tau)=E\left[f_H(t)f_H(t+\tau)\right]=$$
$$=\int_A\int_0^{2\pi}\overset{N-\text{fold}}{\ldots}\int_0^{2\pi}\sqrt{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sqrt{2H(\omega_i,\boldsymbol{\alpha})S(\omega_i,t)\Delta\omega}\sqrt{2H(\omega_j,\boldsymbol{\alpha})S(\omega_j,t)\Delta\omega} \tag{19}$$
$$\times\cos\left(\omega_i t+\phi_i\right)\cos\left(\omega_j(t+\tau)+\phi_j\right)p_A(\boldsymbol{\alpha})\left(\prod_{k=1}^{N}p_{\Phi_k}(\phi_k)\mathrm{d}\phi_k\right)\mathrm{d}\boldsymbol{\alpha}$$

After simple algebra

$$R(t,t+\tau)=2\sum_{i=1}^{N}\sum_{j=1}^{N}\int_A\sqrt{2H(\omega_i,\boldsymbol{\alpha})S(\omega_i,t)\Delta\omega}\sqrt{2H(\omega_j,\boldsymbol{\alpha})S(\omega_j,t+\tau)\Delta\omega}p_A(\boldsymbol{\alpha})\mathrm{d}\boldsymbol{\alpha}$$
$$\times\int_0^{2\pi}\overset{N-\text{fold}}{\ldots}\int_0^{2\pi}\cos\left(\omega_i t+\phi_i\right)\cos\left(\omega_j(t+\tau)+\phi_j\right)\left(\prod_{k=1}^{N}p_{\Phi_k}(\phi_k)\mathrm{d}\phi_k\right)= \tag{20}$$
$$=\sum_{j=1}^{N}\int_A H(\omega_j,\boldsymbol{\alpha})p_A(\boldsymbol{\alpha})\mathrm{d}\boldsymbol{\alpha}\,2\sqrt{S(\omega_j,t)}\sqrt{S(\omega_j,t+\tau)}\Delta\omega\cos\left(\omega_j\tau\right)$$

Taking into account equation (3), the condition (15) , the limit of equation (20) as $\Delta\omega\to 0$ and $N\to\infty$ will lead to equation (5).

Interestingly, in principle any real positive function satisfying equation (15) can be used for controlling the variability of the simulated process.

4. Butterworth filter

Experimental data of random processes (e.g. earthquake, wind or ocean waves) manifest quite different (joint) frequency distributions, therefore $H(\omega,\boldsymbol{\alpha})$ can be determined considering the distribution of the energy around the expected vale for each frequency. To follow this approach is necessary a large number of data that make this approach practically unfeasible. An alternative strategy is to consider synthetic parameters defining the variability of the energy distribution such as the bandwidths and central frequency. To this aim the following pass-band Butterworth filter will be adopted:

$$H_B(\omega,\alpha_1,\alpha_2)=\frac{1}{2\alpha_1\left[1+\left(\frac{\alpha_2-\omega}{\alpha_1}\right)^{2j}\right]} \tag{21}$$

Where $j=1,2,...,n$ is a positive integer number defining the order of the filter, α_1 and α_2 are the random filter parameters defining respectively the bandwidth and the central frequency. Figure 1a shows for selected $\alpha_1=20\,\text{rad/s}$ and $\alpha_2=40\,\text{rad/s}$ the influence of the order of the filter j while Figure 1b shows the influence of the bandwidth parameter on the shape of the filter.

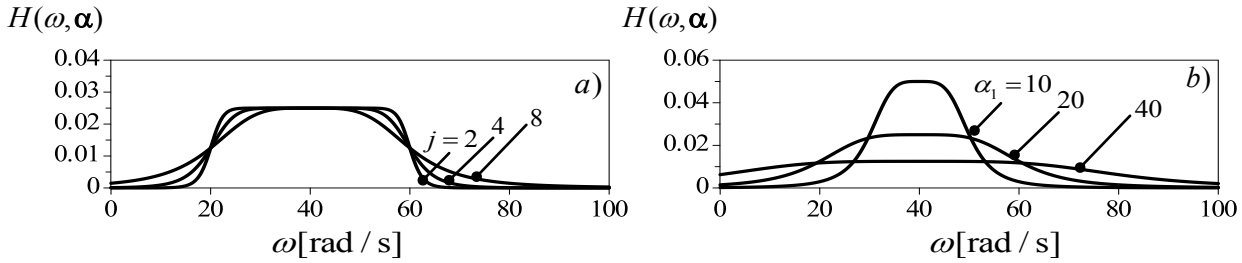


Figure 1. Butterworth filter for (a) different orders and (b) bandwidth ($\alpha_2=40\,\text{rad/s};\ j=2$)

The distribution of the filter parameters α_1 and α_2 can be defined through experimental data measuring the central frequency and bandwidth of the squared Fourier spectrum of the recorded samples. In the case in which the experimental data are not enough to define a proper distribution of the data α_1 and α_2 will be assumed as statistical independent and uniformly distributed. Therefore,

$$p_A(\boldsymbol{\alpha})=p_{A_1}(\alpha_1)p_{A_2}(\alpha_2)=\frac{1}{\alpha_{u1}-\alpha_{l1}}\frac{1}{\alpha_{u2}-\alpha_{l2}}, \tag{22}$$

where α_{ui} and α_{li} ($i=1,2$) represents the upper and lower bounds of the random variables α_1 and α_2. The simulation formula will be then modified as follows

$$f_H(t)=\sqrt{2}\sum_{j=1}^{N}\sqrt{2C(\omega)H_B(\omega_j,\alpha_1,\alpha_2)S(\omega_j,t)\Delta\omega}\cos\left(\omega_j t+\phi_j\right), \tag{23}$$

where the function $C(\omega)$ is introduced to satisfy equation (15) , i.e.

$$C(\omega)=\left(\frac{1}{\alpha_{u1}-\alpha_{l1}}\frac{1}{\alpha_{u2}-\alpha_{l2}}\int_{\alpha_{l1}}^{\alpha_{u1}}\int_{\alpha_{l2}}^{\alpha_{u2}}H_B(\omega,\alpha_1,\alpha_2)\mathrm{d}\alpha_1\mathrm{d}\alpha_2\right)^{-1} \tag{24}$$

It is noted that being $C(\omega)$ function of the circular frequency ω the pass-band filter will be distorted, however being $C(\omega)$ a smooth function the main features of the filters will be preserved.

5. Numerical results

In this section stationary and non-stationary samples are generated following the traditional spectral representation method and the proposed one enhancing the variability. Consider the stationary case first. It is assumed that the process is zero-mean, Gaussian with power spectral density function defined by the Clough-Penzien model, that is

$$S(\omega) = S_0 \frac{\left(1+4\zeta_g^2\left(\frac{\omega}{\omega_g}\right)^2\right)}{\left(1-\left(\frac{\omega}{\omega_g}\right)^2\right)^2+4\zeta_g^2\left(\frac{\omega}{\omega_g}\right)^2} \frac{\left(\frac{\omega}{\omega_f}\right)^4}{\left(1-\left(\frac{\omega}{\omega_f}\right)^2\right)^2+4\zeta_f^2\left(\frac{\omega}{\omega_f}\right)^2} \tag{25}$$

with parameters

$$\omega_g = 20 \text{ rad/s}; \quad \zeta_g = 0.6; \quad \omega_f = 0.1\omega_g; \quad \zeta_f = \zeta_g \tag{26}$$

and

$$S_0 = \frac{\sigma^2}{2\pi\omega_g\left(2\zeta_g+\frac{1}{2\zeta_g}\right)}; \qquad \sigma = 100 \text{ cm/s}^2. \tag{27}$$

Samples of the stationary processes have been simulated using equation (11) by substituting $S(\omega,t)$ with $S(\omega)$ and depicted in Figure 2a. The power spectral density are then randomized using the Butterworth pass-band filter defined in equation (21), in which the random variables α_1 and α_2 are assumed statistical independent and uniformly distributed. Pertinent selected parameters of the distributions are

$$\alpha_{l1} = 10 \text{ rad/s}; \quad \alpha_{u1} = 80 \text{ rad/s}; \quad \alpha_{l2} = 18 \text{ rad/s}; \quad \alpha_{u1} = 40 \text{ rad/s}. \tag{28}$$

Samples of the random processes with enhanced variability are then simulated using equation (23). It is noted that each individual sample requires *N+2* realizations of independent uniformly distributed random variables (*N* phase angles ϕ_j and 2 filter parameter α_1 and α_2). Figure 2b shows three simulated trajectories. The effect of the Butterworth filter in the simulation of the process is more appreciated in Figures 3 where the squared modulus of the Fourier transform of the simulated samples is depicted. It is noted that the samples simulated from the traditional spectral-representation method exhibit Fourier spectra very similar to each other, while the proposed modification lead to spectra possessing clearly visible different distributions.

Comparison with the target power spectral density is shown in Figure 4. It is noted that the ensemble average power spectra determined using 500 samples oscillates around the target spectrum.

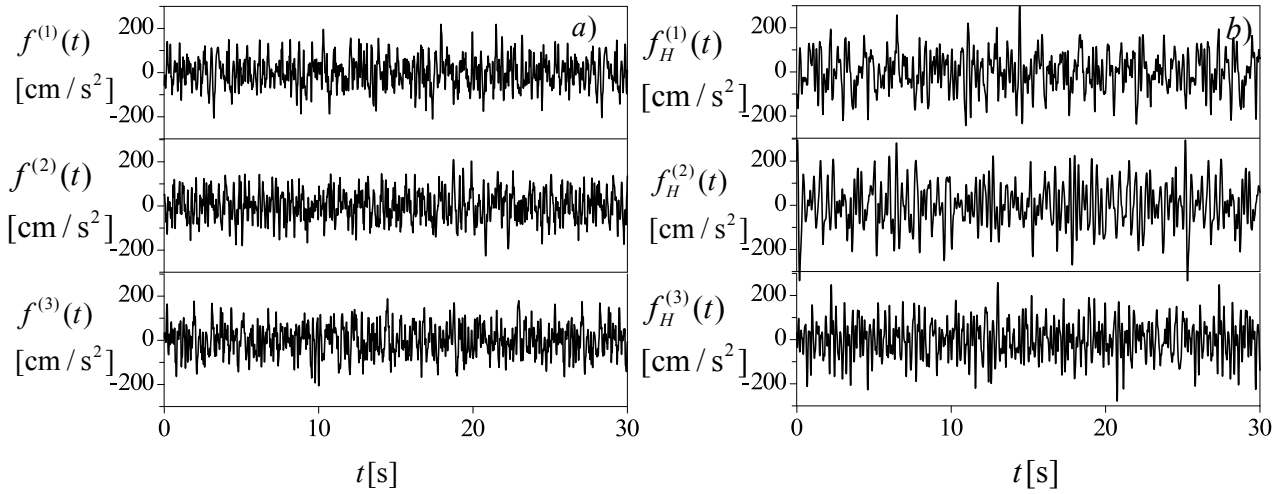


Figure 2. Simulated trajectories of the zero-mean Gaussian, stationary process through a) spectral representation method b) proposed with enhanced variability.

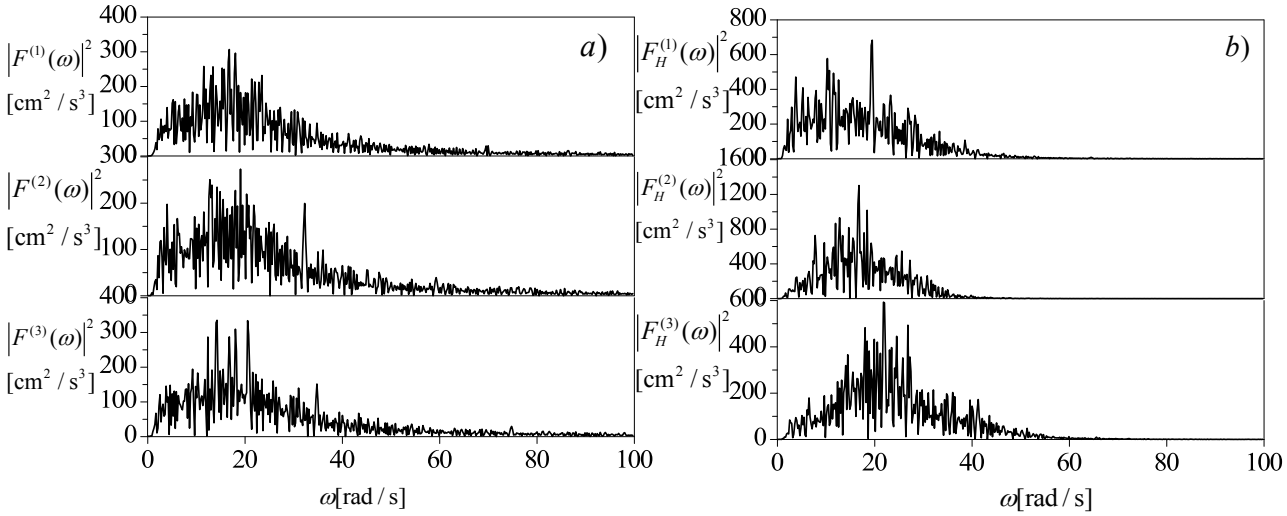


Figure 3. Squared Fourier transform of the simulated samples a) spectral representation method b) proposed with enhanced variability.

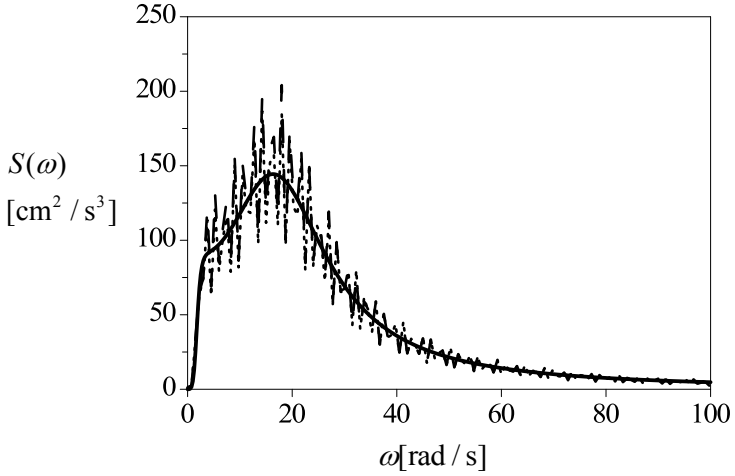


Figure 4. Comparison between the target power spectral density function (solid line), the ensemble averaged determined by using the spectral representation method (dotted line) and the SRM with enhanced variability (dash-dotted line).

The non stationary case is then investigated. The evolutionary power spectral density defined in equation (29) is selected for this purpose

$$S(\omega,t)=A(t)^2 S_0(t)\frac{\left(1+4\zeta_g^2(t)\left(\frac{\omega}{\omega_g(t)}\right)^2\right)}{\left(1-\left(\frac{\omega}{\omega_g(t)}\right)^2\right)^2+4\zeta_g^2(t)\left(\frac{\omega}{\omega_g(t)}\right)^2}\frac{\left(\frac{\omega}{\omega_f(t)}\right)^4}{\left(1-\left(\frac{\omega}{\omega_f(t)}\right)^2\right)^2+4\zeta_f^2(t)\left(\frac{\omega}{\omega_f(t)}\right)^2} \tag{29}$$

pertinent parameters used in this numerical example include:

$$\omega_g(t)=20-7\frac{t}{30};\quad \zeta_g(t)=0.6-0.2\frac{t}{30};\quad \omega_f(t)=0.1\omega_g(t);\quad \zeta_f(t)=\zeta_g(t); \tag{30}$$

and

$$A(t)=a_1 t\exp\left(-a_2 t\right);\quad a_1=0.68\,\mathrm{s}^{-1};\quad a_2=1/4\,\mathrm{s}^{-1}. \tag{31}$$

Furthermore

$$S_0(t)=\frac{\sigma^2}{2\pi\omega_g(t)\left(2\zeta_g(t)+\frac{1}{2\zeta_g(t)}\right)};\quad \sigma=100\,\mathrm{cm}/\mathrm{s}^2. \tag{32}$$

Samples of the non-stationary processes simulated using equations (11) and (23) are depicted in Figures 5 and their related squared Fourier transform in Figure 6. Also in this case it can be appreciated as the random filter modifies the distribution of each sample.

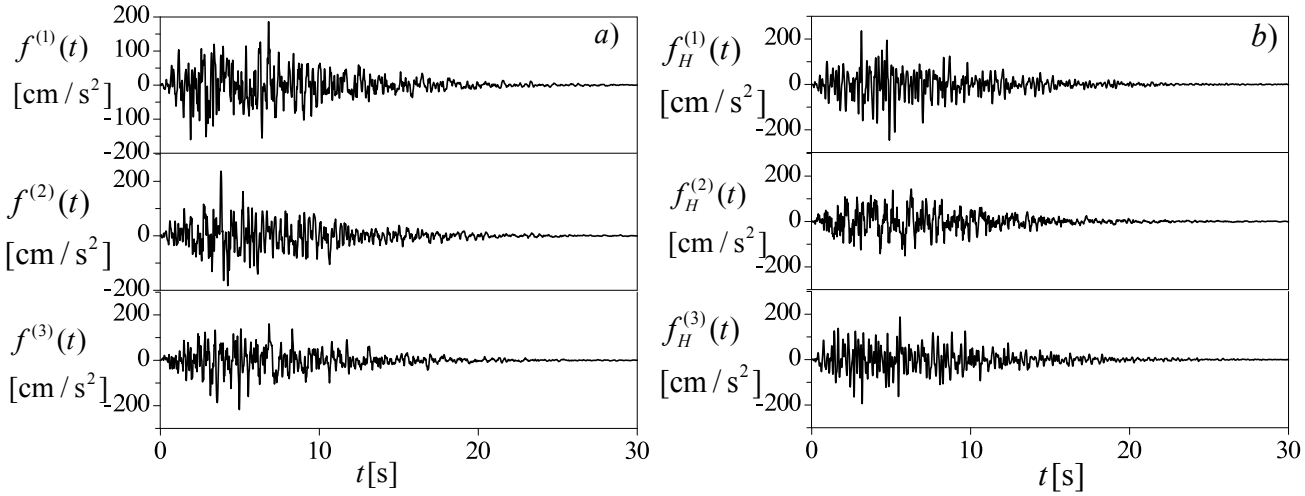


Figure 5. Simulated trajectories of the zero-mean Gaussian, non-stationary process through a) spectral representation method b) proposed with enhanced variability.

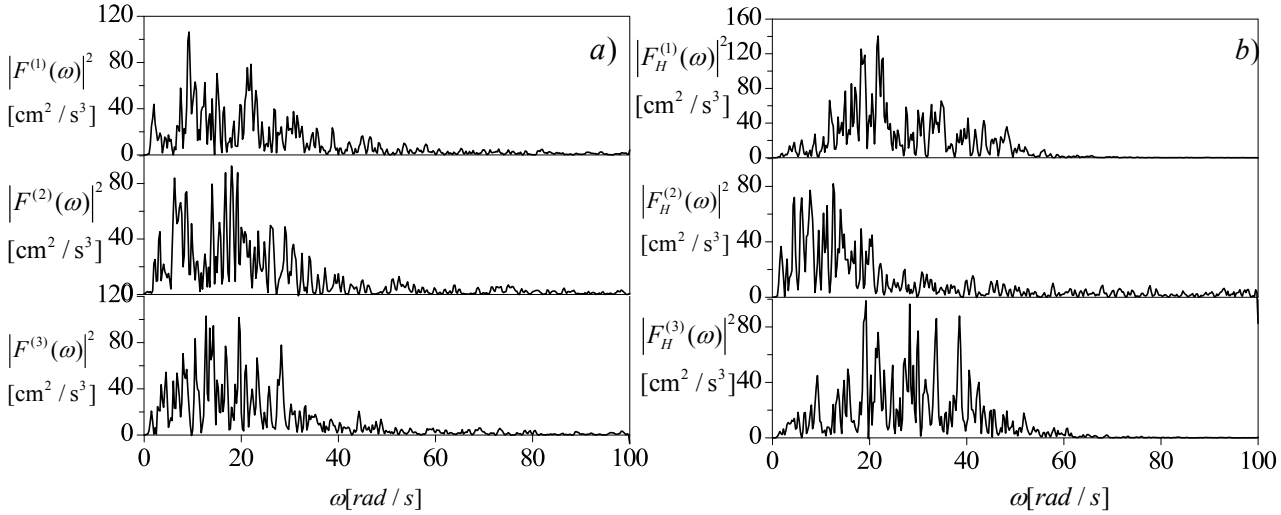


Figure 6. Squared Fourier transform of the simulated samples a) spectral representation method b) proposed with enhanced variability.

Comparisons with the target spectrum is shown in Figures 7 where the marginal distributions (Cohen, 1989)

$$\lambda(t) = \int_0^{\omega_N} 2S(\omega,t)\mathrm{d}\omega\,; \qquad S(\omega) = \frac{1}{t_f}\int_0^{t_f} 2S(\omega,t)\mathrm{d}t \tag{33}$$

are compared respectively with the ensemble average of the squared samples $E[f(t)^2]$, $E[f_H(t)^2]$ and with ensemble average of the squared Fourier transform $E[|F(\omega)|^2]$, $E[|F_H(\omega)|^2]$. The comparison shows the excellent agreement between the simulated and the target ones

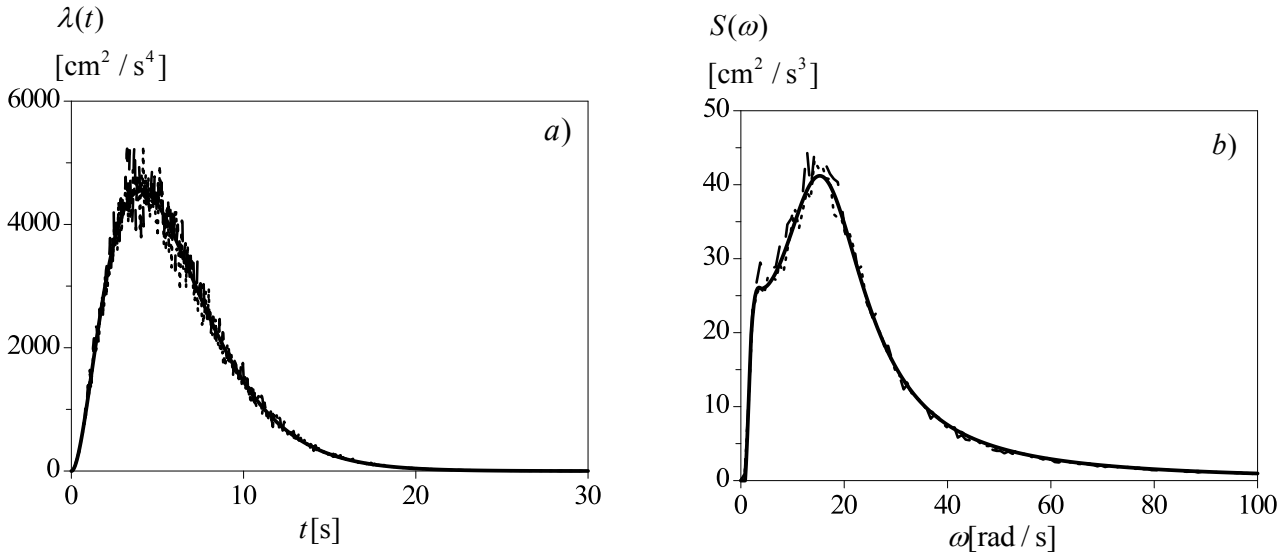


Figure 7. Comparison between the marginal distribution of the evolutionary power spectral density function a) mean instantaneous energy and b) energy density spectrum; target (solid line), the ensemble averaged determined by using the spectral representation method (dotted line) and SRM with enhanced variability (dash-dotted line).

The influence of the enhanced spectrum variability is then investigated through the Monte Carlo study of the distribution of peaks value. Figure 8 shows the comparison between the mean value of the peak of the stationary (Figure 8a) and non-stationary process (Figure 8b). Interestingly both the simulation tends to very similar value and the rate of convergence is comparable. Figures 9 shows the variance of the distribution of maxima. It is noted that the SRM with enhanced variability leads to larger values of the variance. Finally the distributions of maxima along its cumulative distribution are depicted in Figures 10 and 11. It is noted that in both stationary and non stationary case the SRM with enhanced variability leads to wider distribution and as a consequence sensitive different values of the tails of the distribution.

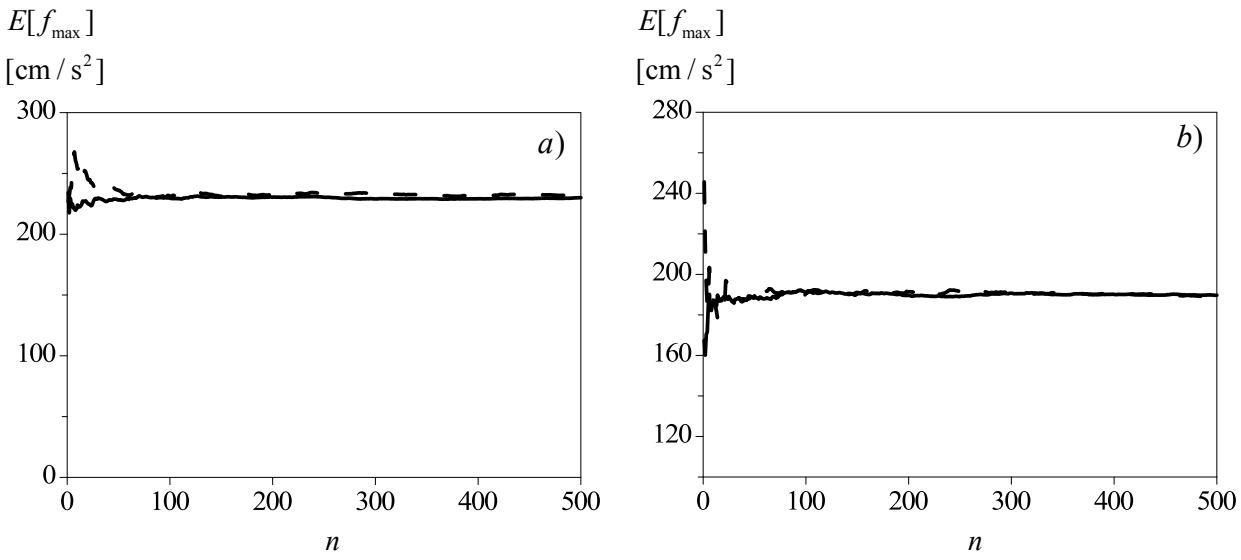


Figure 8. Convergence of the mean value of the peak versus the number of samples *n* for the a) stationary and b) non-stationary process: traditional SRM (solid line), SRM with enhanced variability (dash-dotted line).

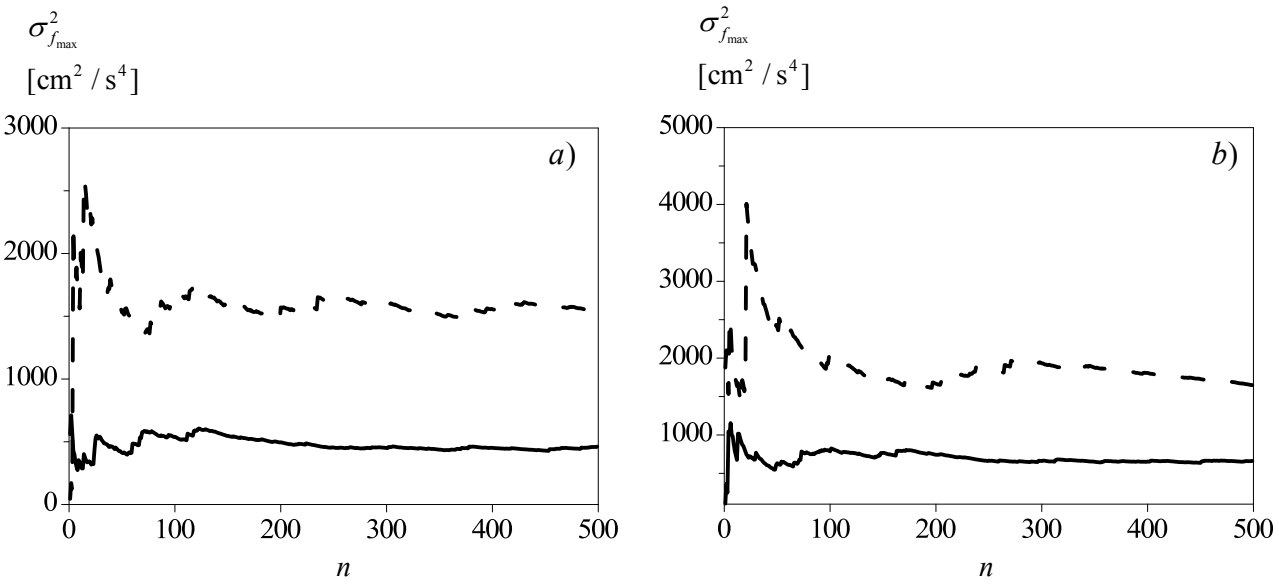


Figure 9. Convergence of the variance of the peak value versus the number of samples *n* for the a) stationary and b) non-stationary process: traditional SRM (solid line), SRM with enhanced variability (dash-dotted line).

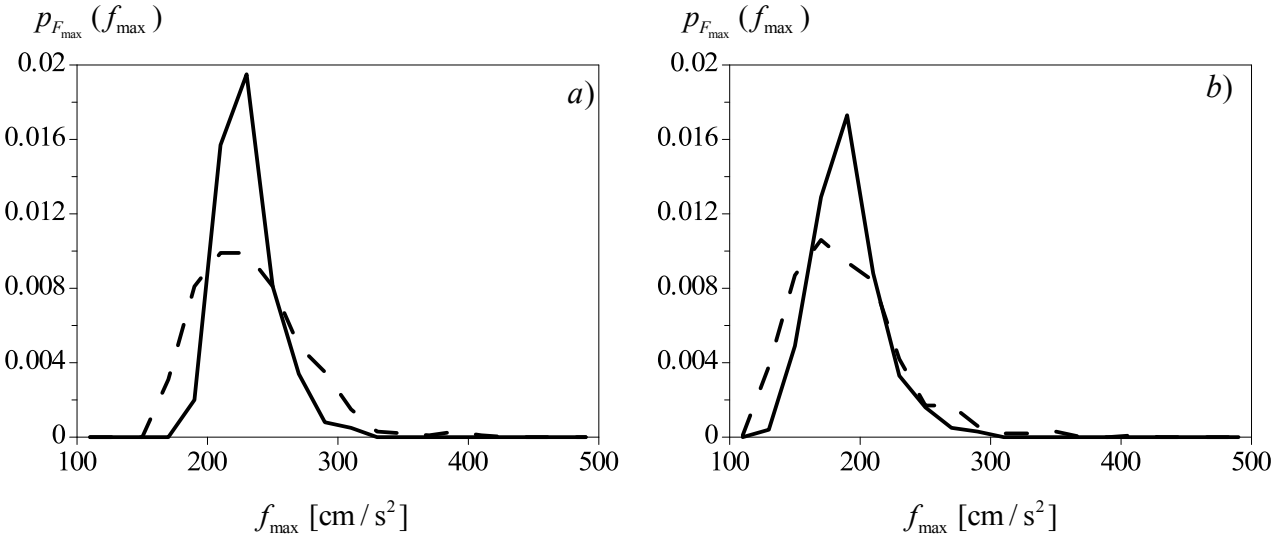


Figure 10. Comparison between the distribution of peaks for the a) stationary and b) non-stationary process: traditional SRM (solid line), SRM with enhanced variability (dash-dotted line).

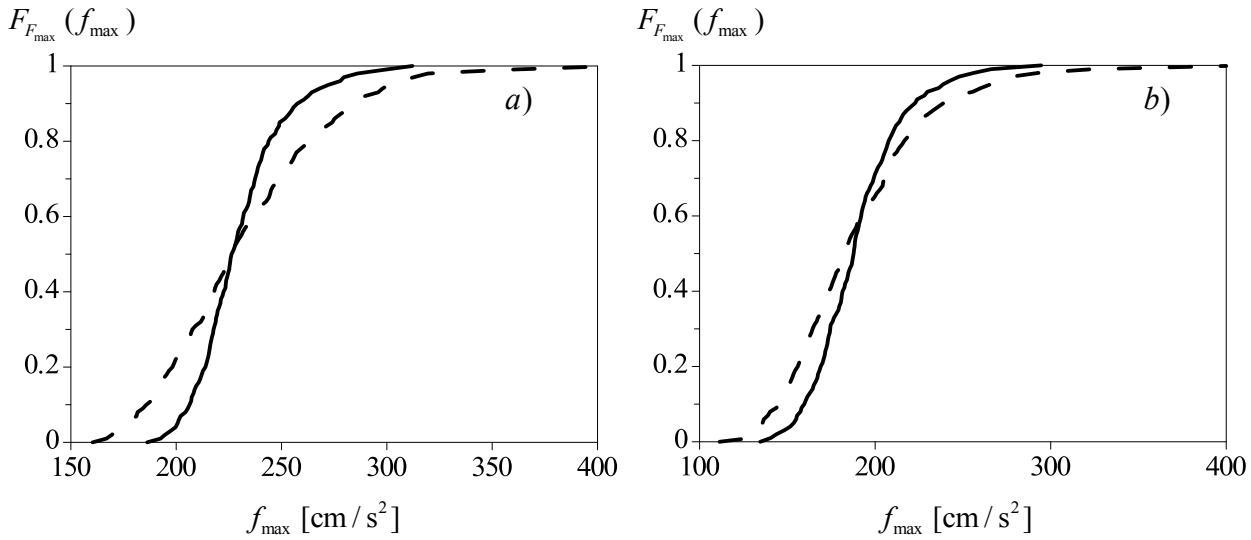


Figure 11. Comparison between the cumulative distribution of peaks for the a) stationary and b) non-stationary process: traditional SRM (solid line), SRM with enhanced variability (dash-dotted line).

6. Concluding remarks

In this paper the traditional spectral-representation-method for simulating stochastic processes is revisited. A modification aimed to control the variability of the simulated samples of the random process is proposed. Specifically a Butterworth pass-band filter with random parameters has been included in the simulation formula to generate samples with different Fourier spectra. It has been shown that the use of this filter do not alter the convergence of the traditional spectral representation method, offering a possible simulation strategy whereas it is required a variability larger than those currently offered by the traditional simulation technique. Comparison between the traditional spectral representation method and the enhanced variability introduced in this paper showed how the randomized power spectral density influence the ensemble simulated samples. Remarkably the peak distribution is significantly sensible to the spectrum variability and should be carefully considered when reliability analyses are performed. It is also expected in general that the spectrum variability influence the features of the process non-linearly related to the power spectrum.

7. References

Cacciola P and Zentner I, Generation of response-spectrum-compatible artificial earthquake accelerograms with random joint time–frequency distributions Probabilistic Engineering Mechanics (28), 52-58, 2012

Cohen L. Time-Frequency Distributions – A Review. Proceedings of the IEEE; 77(7): 941-981,1989.

Deodatis, G. Non-stationary stochastic vector processes: seismic ground motion applications. Probabilistic Engineering Mechanics 11: 149-168, 1996.

Li J, Yan Q ,Chen JB, Stochastic modeling of engineering dynamic excitation for stochastic dynamics of structures, Probabilistic Engineering Mechanics, 27, 19-28, 2012.

Pousse, G., Bonilla, L.F., Cotton, F., Margerin, L., Nonstationary Stochastic Simulation of String Ground Motion Time Histories Including Natural Variability: Application to the K-Net Japanes Database, Bullettin of the Seismological Society of America, 96(6), 2103-2117, 2006.

Priestley, M.B. Evolutionary spectra and non-stationary processes, Journal of the Royal Statistical Society, Series B, Methodological, 27(2), 204-237, 1965.

Rezaeian S and Der Kiureghian A, Simulation of synthetic ground motions for specified earthquake and site characteristics, Earthquake Engineering and Structural Dynamics, 39, 1155-1180, 2010

Schillinger D and Papadopoulos V, The method of separation: a novel approach for accurate estimation of evolutionary power spectra, Computational methods in Stochastic Dynamics, Papadrakakis M, Stefanou G, Papadopolous V eds. Springer, 209-228, 2011.

Shinozuka, M. and Deodatis, G. "Stochastic Process Models for Earthquake Ground Motion," Probabilistic Engineering Mechanics, 3, (3), 114-123, 1988

Spanos PD, Failla G, Evolutionary spectra estimation using wavelets, Journal of Engineering Mechanics (ASCE), 130(8), 952-960,2004.

Viallet E., Humbert N. Considerations on the use of natural and artificial time histories for seismic transient non-linear analyses of structures and variability assessment. Proceedings of SMiRT 19, Toronto, Canada, 2007.

Analysis of Tri-axial Stress-strain Conditions of Pre-stressed Masonry Corner

Radim Cajka, Pavlina Mateckova, Lucie Mynarzova, Marie Stara
Department of Building Structures, VSB-Technical University of Ostrava, Czech Republic,
pavlina.mateckova@vsb.cz

Abstract: During the revitalization of masonry structures the post-tensioning of masonry is used frequently for improving the structural mechanical properties. Masonry is heterogeneous material with varied material properties, in older buildings affected with different type of damage and that is why in engineering practice the intensity of pre-stressing is often designed according to engineering judgement of the designer. Mathematical modelling of post-tensioned masonry structures is very valuable. Authors are interested particularly in so-called micromodels and macromodels of masonry structures. Structural parts are analyzed using ANsys computer program. Mathematical model has been verified with testing. At VSB – Technical University of Ostrava unique equipment was designed for experimental testing of tri-axial state of stress and strain of pre-stressed masonry corner. Plan dimensions of the tested corner are 900 × 900 mm, the thickness of the wall is 450 mm and the height is 900 mm. Experiments started in 2011 with masonry corner made of clay bricks and general purpose mortar. Ongoing experiments and appropriate mathematical modelling should contribute to higher reliability of engineering computations of masonry structures.

Keywords: micromodel; macromodel; FEM, masonry structures, post-tensioning

1. Testing the Pre-stressed Masonry Corner

1.1. MOTIVATION

Masonry structures are often historical buildings exposed at present to higher action than they were designed to, as a result of changing the use of the building or e.g. undermining and related change in soil-structure interaction (Cajka, Manasek, 2005). In engineering practice post-tensioning of masonry structure is used frequently because it is an effective method to enhance the static action and improve the rigidity of the building. Other advantages are that the existing cracks could be eliminated and the outward of the building does not change. Masonry is heterogeneous material with varied material properties, in older buildings affected with different type of damage and that is why the intensity of pre-stressing is often designed according to engineering judgement of the designer. From experience of practical design and also from available literature (Bazant, Klusacek, 2004) could be concluded that during the post-tensioning the failure occurs most often in the zone of pre-stressing force anchoring. Authors aim is to contribute to better understanding of post-tensioned masonry and therefore it was decided to perform the testing of tri-axial stress/strain conditions in post-tensioned masonry corner.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

1.2. DESCRIPTION

At VSB-Technical University of Ostrava unique equipment for was designed for testing of tri-axial stress/strain conditions, Fig. 1. Plan dimensions of the tested corner are 900 × 900 mm, the thickness of the wall is 450 mm and the height is 900 mm. It is possible to accomplish number of tests for different input parameters, e.g. using various materials (brick, mortar), laying out and number of pre-stressing bars, value of pre-stressing force, size and shape of anchoring plate, value of vertical load, number of brick layers or the thickness of bed joint, brick bond and the supporting of the masonry corner (simple, with slide joint and others). Deformation in network of points during the pre-stressing is measured. Possible tension-metrical measurements have not been fully accomplished yet.

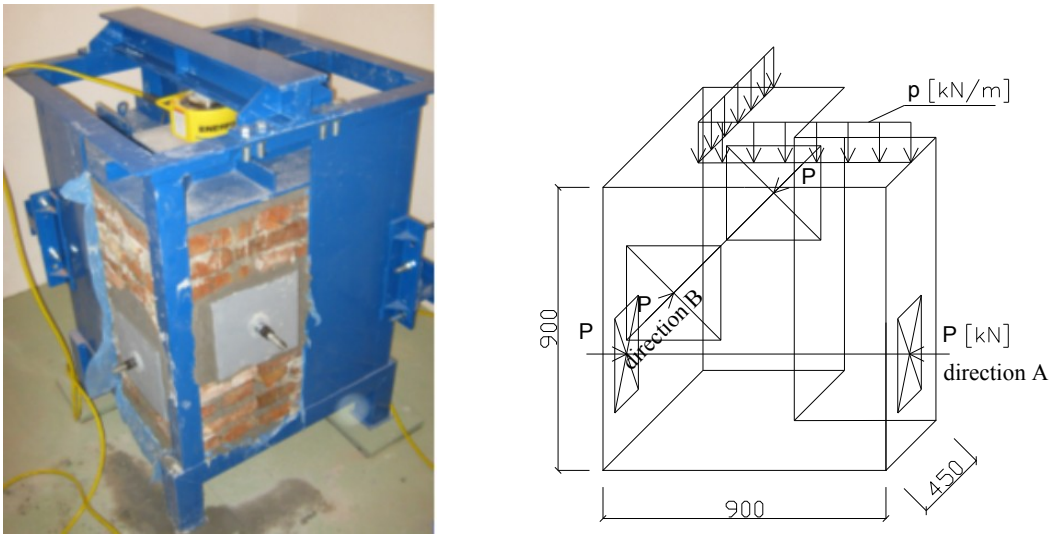


Figure 1. Testing equipment

Experiments started in 2011 with masonry corner made of clay bricks and general purpose mortar. Bricks were obtained from demolished building and limecement mortar was prepared from designed dry mixture. Pre-stressing was installed with two pre-stressing steel bars and square anchoring plates with dimensions 300 mm, 200 mm and 100 mm, the thickness is 10 mm. The strength of bricks and the mortar was checked in Faculty laboratory according to valid codes. Normalised mean strength of bricks was f_b = 16.08 MPa and mean strength of mortar f_m = 9.8 MPa. Characteristic strength of tested masonry is settled in (1) according to (EN 1996-1-1, 2005), where K is constant for different groups of bricks.

$$f_k = K.f_b^{0.7}.f_m^{0.3} = 0.55 \times 16.08^{0.7} \times 9.8^{0.3} = 7.6 \text{ MPa} \quad (1)$$

Masonry corner was exposed to vertical load 0.125 MPa and 0.250 MPa corresponding to vertical load in common building and arbitrary pre-stressing force 50 kN and 100 kN. Pre-stressing was installed in one direction and released, than in second direction and released and in the end in both directions. Only short-time deformations were measured. Primary pre-stressing force was settled according to recommendation that the pre-stressing force should cause the stress value of 10% vertical masonry strength.

1.3. TEST RESULTS

The deformations are measured in regular network of measuring points in two directions, network in the *direction A is* in Fig. 2.

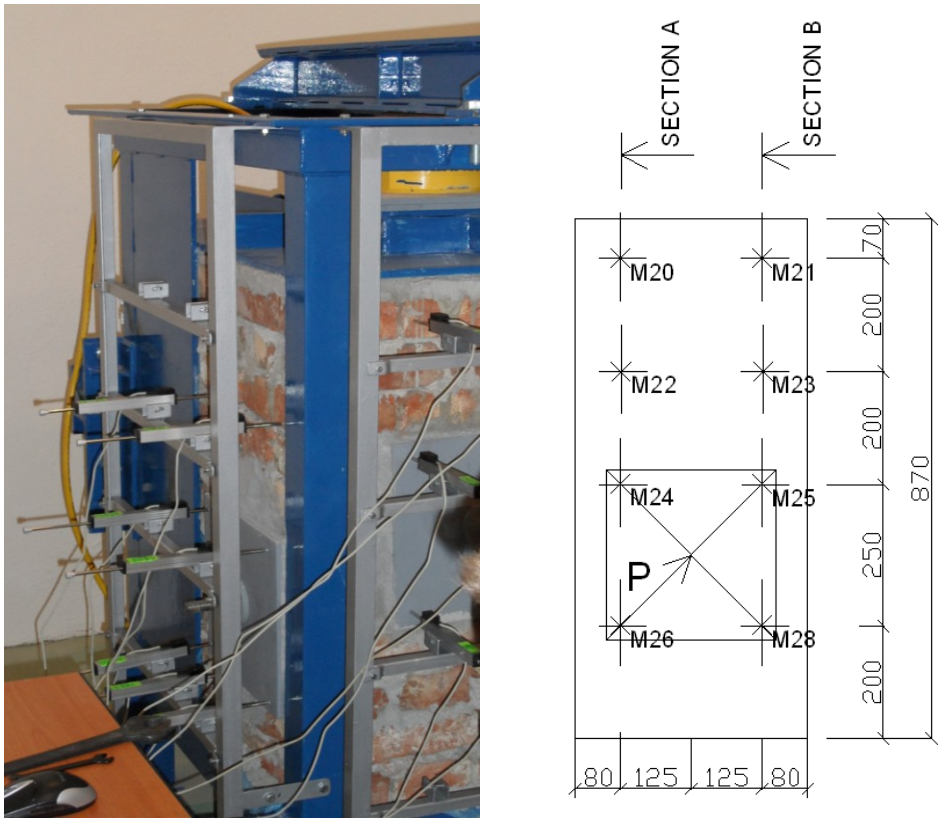


Figure 2. Network of deformation measuring points in the direction *A*

In the charts, Fig. 3 there are measured deformations in the direction A, section A and B (according to Fig. 2) for vertical load 0.125 MPa and pre-stressing force 100 kN in A direction anchored with square plate 0.3 × 0.3 m. Horizontal line indicates the location of pre-stressing force. Deformations in particular vertical section are unexpectedly higher in measuring points farther from the pre-stressing force. However the course of measured deformations in two symmetrical sections is similar. The same course of deformations is also in the Fig. 4 where there are deformations in the direction A for different dimensions of anchoring plates. Values in section A and section B are averaged. When anchoring the pre-stressing force with plate 0.150 × 0.150 m the stress is probably transferred more locally and the course of deformations is different. Possible inaccuracies could be caused also with uneven mortar joint under the anchoring plate. Analogous measurements were done also for pre-stressing in B direction and pre-stressing in both directions and the course of the deformations is similar. The deformations could grasp only the surface stress/strain

conditions. Critical is distribution of pre-stressing in the whole cross-section. Tension-metrical measurements are planned in the newly bricked masonry corner.

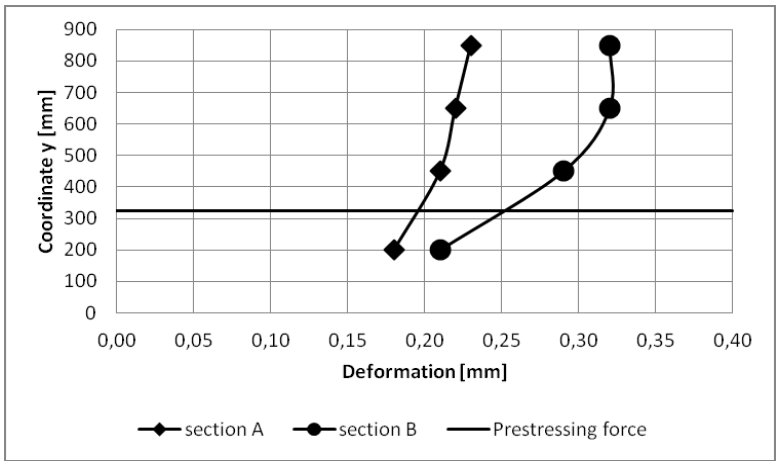


Figure 3. Deformations for pre-stressing 100 kN in the direction *A*, vertical load 0.125 MPa, anchoring plates 0.3 × 0.3 m

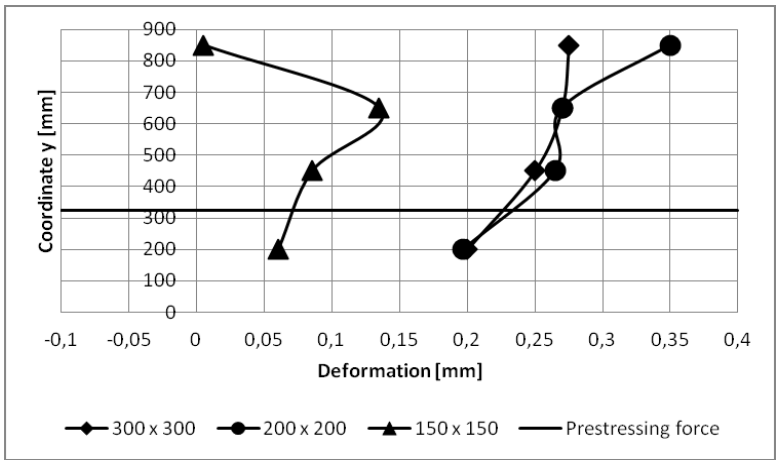


Figure 4. Average deformations for pre-stressing 100 kN in the direction *A*, vertical load 0.125 MPa

2. Masonry numerical modeling

2.1. BASIC WAYS OF MODELING

Complicated numerical modeling of masonry as anisotropic and heterogeneous material is connected with the main following factors:

- Different material properties of basic components (brick/mortar)

- Different dimensions of basic components (dimension of brick/dimension of mortar joint)
- Narrow dimension of mortar joint
- Geometrical arrangement of bricks
- Different structural response for different load action direction
- Interaction between components
- Quality of manufacturing
- Environmental influence

Investigation into masonry structures and into its mathematical modeling has brought number of different approaches, (Materna, Brozovsky, 2007). According to (Lourenco, 1996) there are three basic strategies, Fig. 5:

- *Detailed micromodel* – bricks and mortar are assumed as two different materials, with their real dimensions and real geometrical arrangement in the structure.
- *Simplified micromodel* – brick and surrounding mortar joints is assumed as one block. Material properties of bricks and mortar have to be simplified.
- *Macromodel* – masonry is assumed as homogenous material. It is necessary to determine the most fitting anisotropic material properties.

2.2. Determination Of Heterogeneous Material Properties

Detailed micromodel, with its modeling of particular masonry components and geometrical arrangement, could be precise and accurate, but also very laboring and demanding the powerful computer. Modeling the whole structure in this way is inconceivable. Micromodel is useful for modeling of structural details or local action. Micromodel could be used for determination of heterogeneous material characteristics for macromodel, (Cajka, Kalocová, 2007). Part of the masonry wall micromodel is exposed to deformation load in different direction, Fig. 6, and from the result structural response F_x modulus of elasticity and shear modulus are determined (2). Input parameters for illustrative example and settled material heterogeneous parameters are listed in the Table 1.

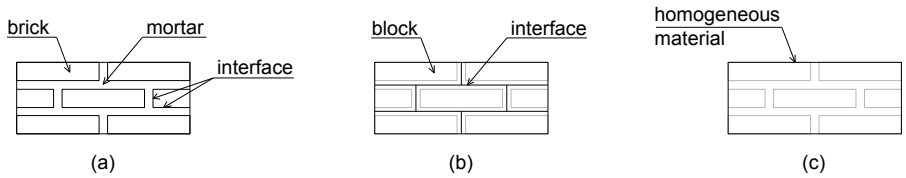


Figure 5. Basic way of masonry modeling

$$E_x = \frac{F_x \cdot L_x}{u_x \cdot A_x}, \qquad G_{xy} = \frac{F_{xy}}{\gamma \cdot A} \tag{2}$$

Table I. Input and output parameters for illustrative example

	Material	*Modulus of Elasticity* [GPa]	*Shear modulus* [GPa]	*Poissons' ratio* [-]
Input	Bricks 290/140/65	15.00		0.15
	Mortar 10MPa	10.00		0.2
Output	Masonry *x* direction	13.96	3.65	0.202
	Masonry *y* direction	14.00	5.56	0.198
	Masonry *z* direction	13.84	4.30	0.197

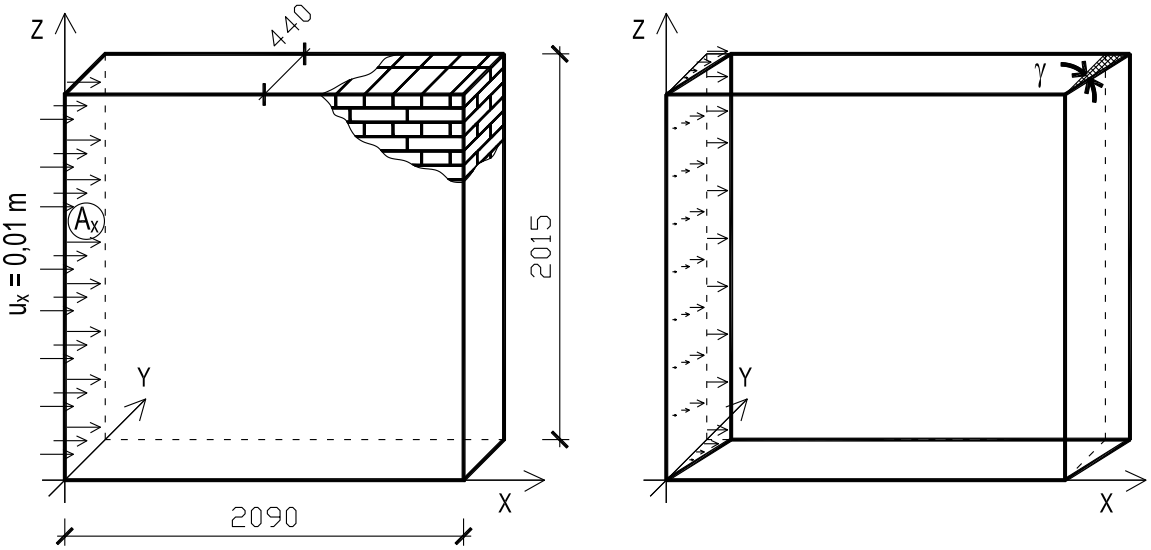


Figure 6. Masonry wall micromodel exposed to deformation load in x direction

2.3. MODELING OF TESTED MASONRY CORNER

Before the testing started the micromodel and macromodel of post tensioned masonry corner was prepared in ANsys computer program, Fig. 7. Elements Solid 45 are used for bricks, mortar, anchoring plates and also in macromodel. Pre-stressing is incorporated with the element Link8.

Micromodel and macromodel of masonry corner is made assuming the same conditions as in experiment, i.e. bricks 209/140/65 and lime cement mortar 10 MPa, vertical load 0.250 MPa and pre-stressing force 100 kN. In the Fig. 8 there are deformations in the section in location of upper pre-stressing bar. In macromodel the deformations (and consequently stress) are spread to larger area. In micromodel the deformation is concentrated in anchoring area and especially bricks and mortar adjacent to anchoring plate are affected. Deformation (and consequently stress) in micromodel corresponds with practical experiences with masonry post-tensioning. In case of masonry resistance is exhausted usually only bricks and mortar in anchoring area are affected while the farther components are not concerned.

Currently the mathematical model of masonry corner is improved so that it better fits the real structure. Presently it is possible to state that the settled deformations correspond to measured values approximately.

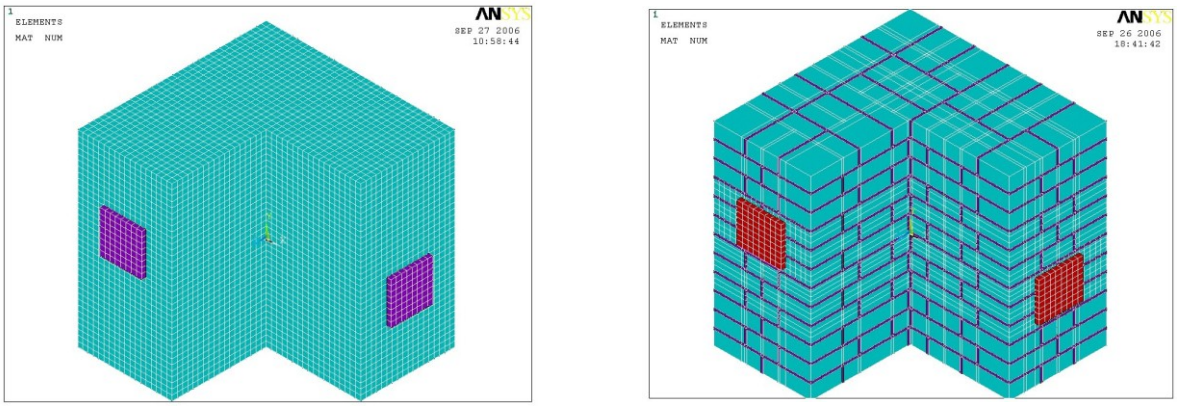


Figure 7. Macromodel and micromodel of post-tensioned masonry corner

Together with improving the mathematical model walling of new masonry corner is prepared again with clay bricks and general purpose mortar with lower strength.

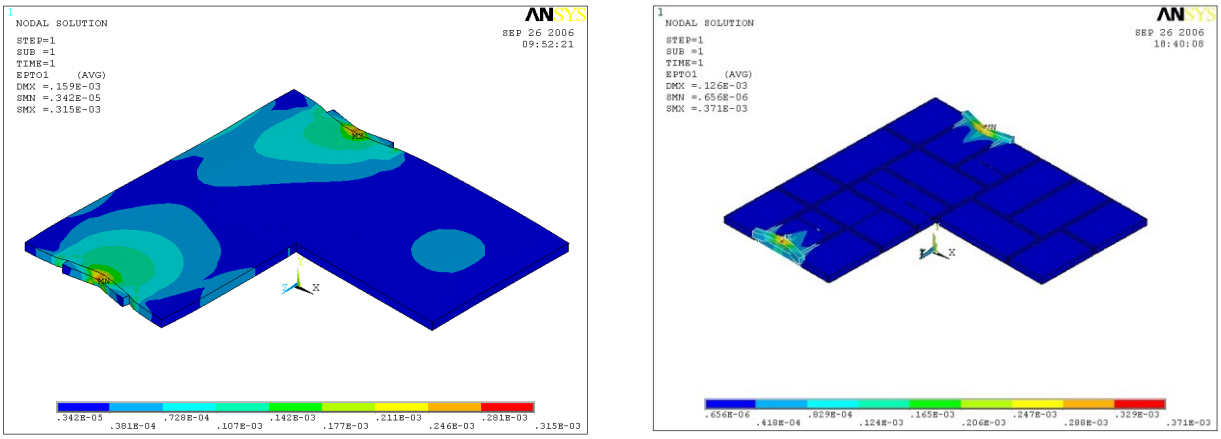


Figure 8. Macromodel and micromodel – deformation in section in place of upper pre-stressing bar

3. Conclusion

In the paper authors introduce the experimental testing of masonry corner exposed to tri-axial load and partial measured deformations. Masonry as heterogeneous and anisotropic material requires specific way of mathematical FEM modeling. Micromodel and macromodel of tested masonry corner is introduced. Ongoing experiments and following FEM modeling should contribute to better understanding of masonry strength characteristics and thus improve the structural reliability of masonry in case of post-tensioning.

Acknowledgements

This outcome has been achieved with the financial support of the Ministry of Industry and Trade of the Czech Republic, program TIP, project No. FR-TI2/746 Rheological sliding joint with thermo-controlled viscoelastic properties.

References

Bazant, Z.; Klusacek, L. Statics at building reconstruction (in Czech). VUT Brno, 2004

Cajka, R.; Kalocova, L**.** Modelling and Analysis of Post-Tensioned Masonry. *In Proceedings of the Eleventh International Conference on Civil, Structural and Envirinmental Engineering Computing 2007*, Malta 2007

Cajka R., Manasek P. Building structures in danger of flooding. *In proceedings IABSE Conference on Role of Structural Engineers Towards Reduction of Powerty*, New Delhi, India, 2005

EN 1996-1-1 Eurocode 6 – Design of masonry structures – Part 1-1: General rules and rules for reinforced and unreinforced masonry structures. CEN Brussels 2005

Lourenco, P. B. Computational strategies for masonry structures. Delft University Press, 1996

Materna, A., Brozovsky, J.: Constitutive model for two-dimensional modeling of masonry. *In Proceedings of the Eleventh International Conference on Civil, Structural and Envirinmental Engineering Computing 2007*, Malta 2007

Different models of soil-structure interaction and consequent reliability of foundation structure

Radim Čajka
Department of Building Structures, VSB – Technical University, 708 33 Ostrava, Czech Republic,
radim.cajka@vsb.cz

Abstract: In the calculation of structural reliability often variation of material characteristics and action effect is considered. The accuracy of reliability assessment depends on how precisely it is possible to grasp statistical concepts of material characteristics and action effect. In this paper author would like to discuss the fact that concerning foundation structures the highest variation in reliability assessment is not caused with material characteristics or load effect but in the model of soil-structure interaction itself. Above mentioned problem is demonstrated in the example of strip foundation / foundation slab.

Keywords: foundation structures, interaction models, FEM, elastic half space

1. Introduction

In buildings, the foundation structures are required to transfer all load components from the upper construction onto subsoil. Typically, attention is paid to the transfer of the vertical load components which is applied in the direction of Earth's attraction. The interaction between various types of environment has been discussed for several years. In order to define the state of stress more precisely, in particular that of foundation structures, it is essential to define, on one hand, how rigidity of the foundation structure influences the settling process and, on the other hand, how rigidity or elasticity of subsoil influences internal forces within constructions. First works about this topic include those written by Gorbunov-Posadov, Winkler and Pasternak (Cajka, 2008).

Application of numerical methods in practice started upon launch of computers. A general variational method for analysis of building constructions – Finite Element Method (FEM) – has been developed in detail by now. Several scientists were dealing with a surface model, the best known being a multi-parameter model of subsoil processed by (Kolář and Němec, 1989). Authors dealing with the state of stress in subsoil caused by vertical and horizontal forces include (Poulos and Davis, 1974). The other theory of soil–structure interaction and subsoil–foundation contact tasks were investigated (Abdel Rahman and Edil, 1991; Qian and Zhang, 1993; Reitinger and Svejda, 1998; Provenzano, 2003; Katzenbach, Schmitt and Turek, 2005; Cajka, 2003, 2005; Cajka and Manasek, 2005; Souli and Shahrour, 2012).

2. Foundation slab with stiff walls

Pregnancy of various models and stiffness of foundation in the foundation–subsoil interaction system was solved by the authors software in the example below taken from (Reitinger and Svejda, 1998). The software

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

MKPINTER (Cajka, 2010) is based on FEM with thick slab theory (Mindlin, 1951), numerical integration (Davis and Rabinowitz, 1956) and nonlinear elastic half-space modified by means of the structural strength of the soil (Cajka, 2003, 2005, 2008).

Let us assume a foundation slab on subsoil. The slab is reinforced longitudinally with stiff walls. The subsoil is modelled by means of 3D FEM as a linearly elastic half-space. But non-linearity is not taken into account and the structural strength is not modified.

Dimensions and loading data are evident from Fig. 1 which was taken from (Reitinger and Svejda, 1998). But there is a correction in the Poisson's ratio for concrete and clayey subsoil which were evidently confused with each other. Results of the published solution are in Fig. 2.

The published example (Reitinger and Svejda, 1998) deals, for purposes of comparison, with an interaction task where a slab is located on a half-space and on Winkler's subsoil. The modulus of subsoil, $k = 1250\ \mathrm{kN \cdot m^{-3}}$, was chosen in such a way so that subsidence in the defined A point could be same for the both models.

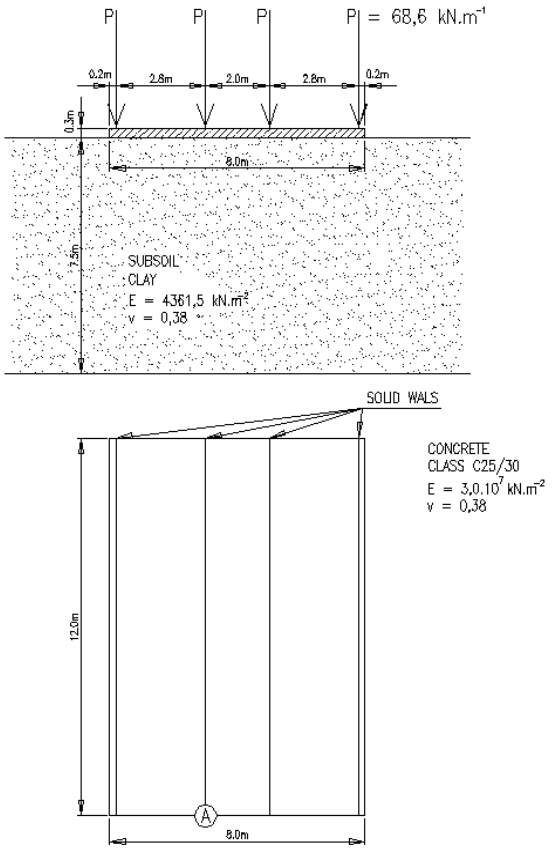


Figure 1. Foundation slab with longitudinal walls

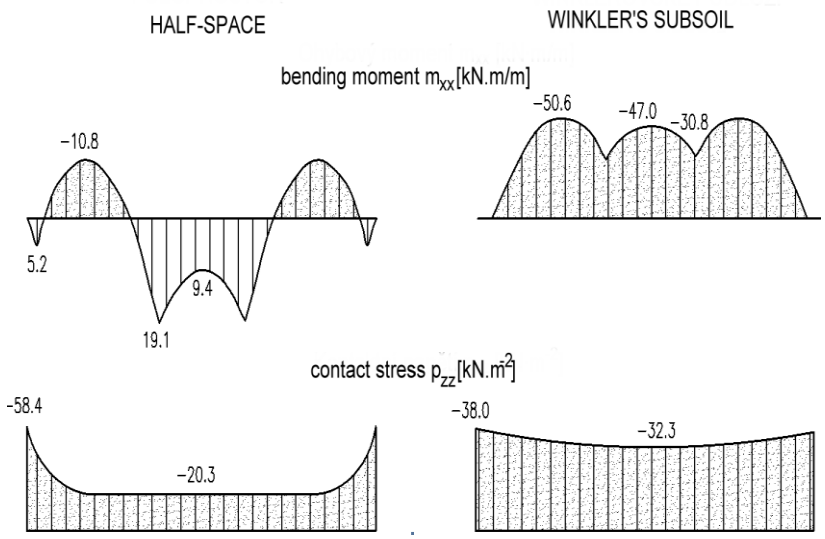


Figure 2. Internal forces in the transversal direction obtained by solution of a slab in a subsoil model with 3D elements and in a Winkler's subsoil model according to 1

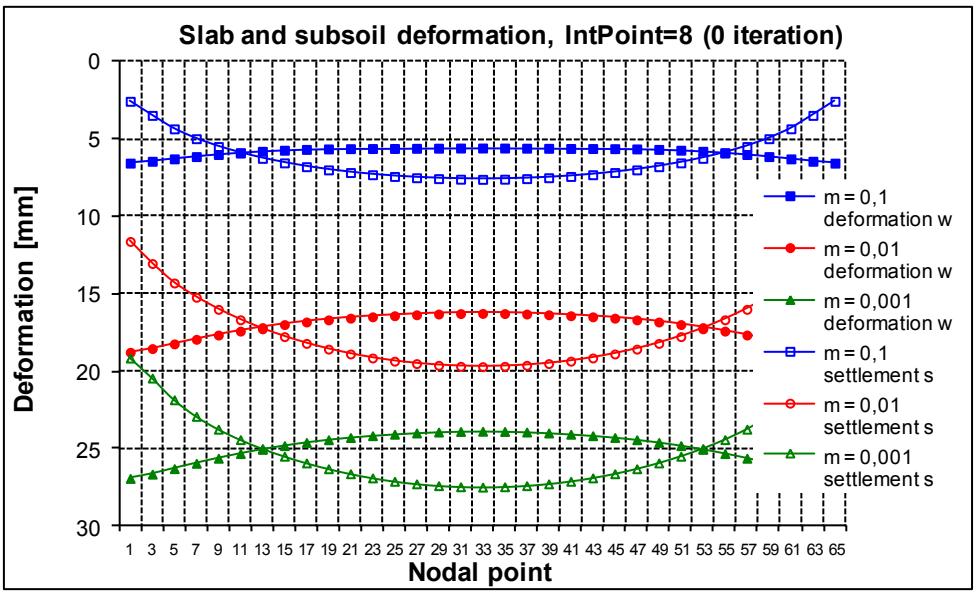


Figure 3. Deformation of the slab and settlement of subsoil vs. structural strength of the soil in subsoil of a contact element – without iterations

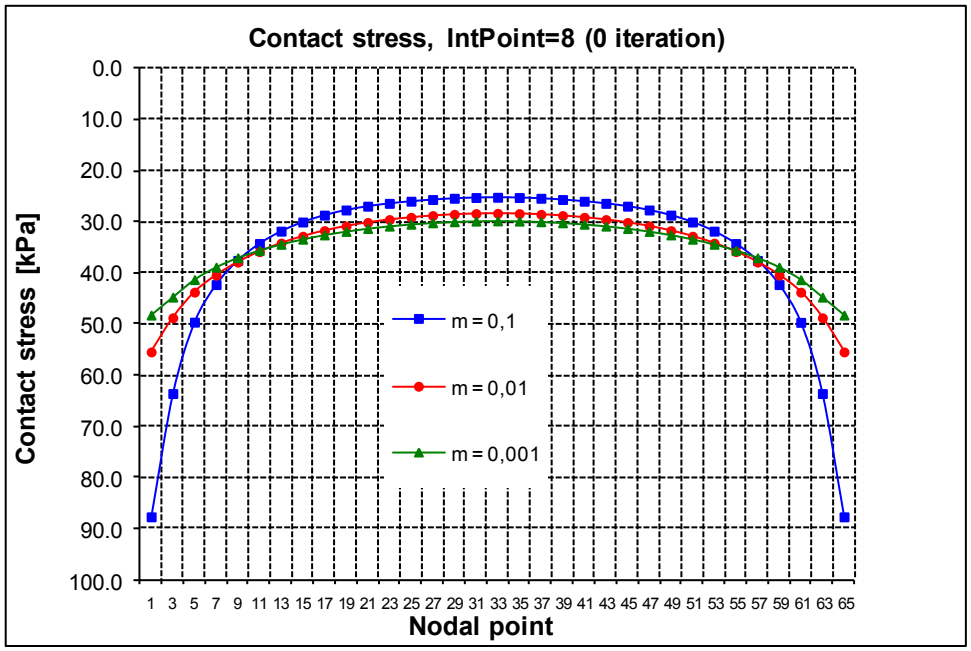


Figure 4. Contact stress vs. structural strength of the soil in subsoil and depth of the deformation zone if a contact element is used – without iterations

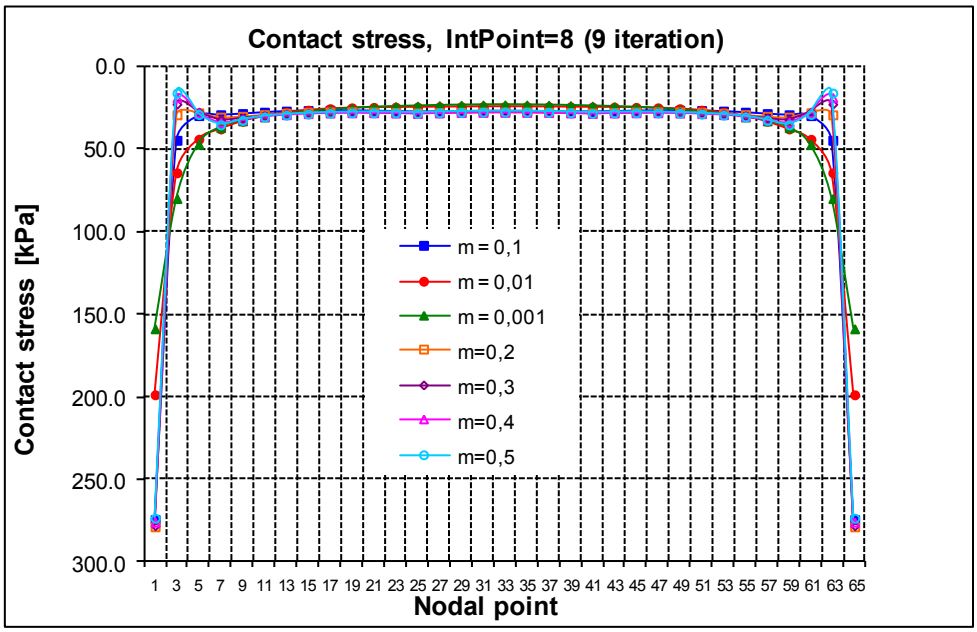


Figure 5. Contact stress vs. structural strength of the soil in subsoil and depth of the deformation zone if a contact element is used – 9th iteration

The same example can be used for an iteration task consisting in a slab on an elastic half-space which is modified by means of the structural strength of the soil pursuant to ČSN 73 1001, CSN EN 1997-1 and CSN EN 1992-1-1.

A typical representative of the mentioned parameters of the subsoil is clayey soil, F4 class, with solid consistence. The reference value of the modulus of plasticity is $E_{def} = 4 – 6$ MPa. Poisson's ratio is $\nu = 0.35$, volumetric weight is $\nu = 18{,}5\ \mathrm{kN \cdot m^{-3}}$ and the coefficient of structural strength of the soil in subsoil is $m = 0.2$. The calculation was also carried out for other coefficients − $m = 0.1$; 0.3; 0.4 and 0.5 − which model various rigidities of the subsoil. The coefficient which approaches zero for $m = 0.01$ and 0.001 model *the subsoil of a standard linear elastic half-space.*

If the deformation and state of stress in soil environment are modelled by means of 3D finite elements and if a sufficiently big domain is chosen, the results should be same as those calculated from explicitly derived relations.

The solution to a 3D task of a linear elastic half-space is among few tasks which have been derived from general equations of the theory of elasticity and fulfil all conditions applicable to solutions in a closed shape.

Thus, the Finite Element Methods as well as the approximate numerical method should have, or at least should converge to, same results for the task if the 3D element should be regarded as a correctly derived element. If the domain of the 3D subsoil (or 3D subsoil, in case of a planar task) is made smaller the results are different for 3D FEM elements because the domain and, in particular, the depth of the domain are chosen by estimates. This situation indicates well presence of non-compressible subsoil (such as rock) which corresponds to the specified zero deformations on the lower edge of the area. In other cases, the scope of the domain should be determined by calculations.

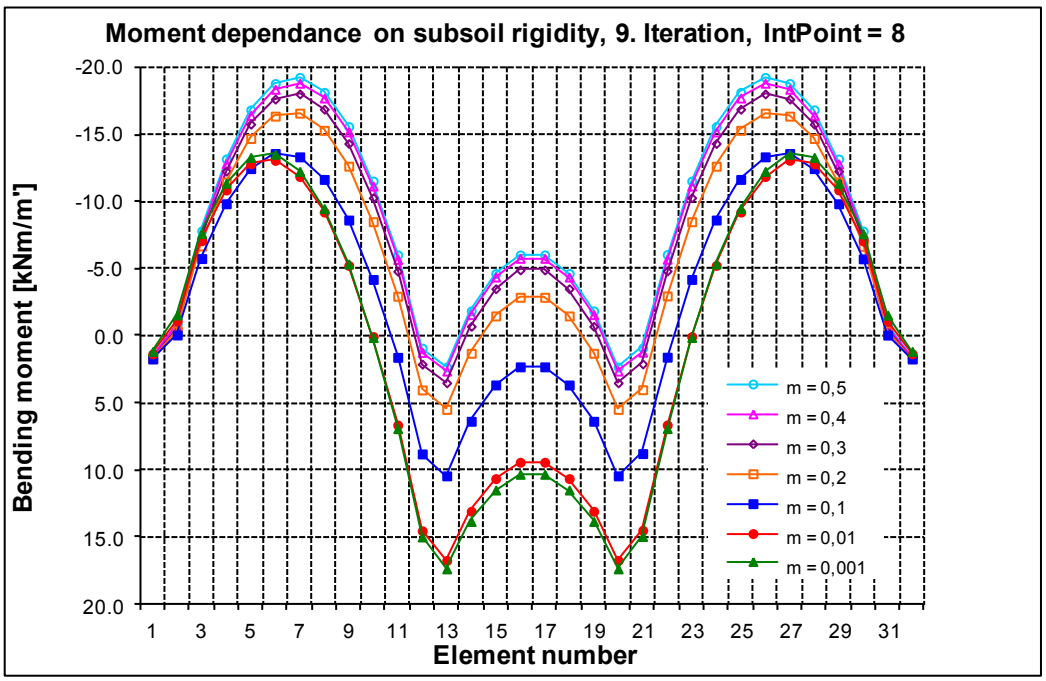


Figure 6. Transversal moment vs. structural strength of the soil in subsoil and depth of the deformation zone if a contact element is used

Results of the FEM interaction method are clear from Fig. 3 to Fig. 6 where development of the subsoil settlement and deflection of the slab is plotted for calculations without iteration (0^{th} iteration step). The figures also show development of the contact stress for the initial (0^{th} iteration) iteration and for the last iteration step (9^{th} iteration) and development of bending moment in a slab for various rigidities of the subsoil. *The used contact element (Cajka, 2003, 2005, 2008) satisfies non-linear deformation properties of subsoil pursuant to ČSN 73 1001 and European Standards CSN EN 1997-1 and CSN EN 1992-1-1. Solution results achieved with the structural strength coefficient being close to zero (m = 0.01 through m = 0.001) correspond to a big deformation zone. The solution with non-real settlement and moments converges towards results of iteration of a slab on a linear elastic half-space (without influence of structural strength of the soil) and, in turn, towards the solution achieved if FEM 3D elements are used in line with (Reitinger and Svejda, 1998).*

3. Convergence towards the exact solution

As it follows from general formulation of FEM, theory of integral computations and accuracy of numerical integrations (Davis and Rabinowitz, 1956), *two key factors affect the convergences towards the theoretically exact solution of the stress and deformation in the foundation- subsoil model (Cajka, 2008):*

- division of the construction into finite elements, the applied slab theory and the degree of the approximation polynomial of the element (the convergence of the slab),
- approximation accuracy of development of the subsoil settlement and stress which is influenced by the number of Gaussian integration points. The number of the Gaussian integration points determines the degree of a polynomial which approximates development of stress in a linearly elastic half-space (the convergence of subsoil).

The convergence of a slab element towards the exact solution has been verified for a freely supported slab without any subsoil. An even continuous load and a single load in the centre of the slab were considered. It follows from the comparison of the results with the exact solution from the literature that the solution converges in accordance with the FEM theory.

Accuracy of the numerical integration in calculations of the stress and settlement of the half-space was tested in reference examples which were confronted to data available in the literature. Comparison calculations indicate that an acceptable technical accuracy is reached when 6 integration points are used.

Because development of the contact stress influences deformation of both the slab and subsoil, it is clear that the division of the construction into finite elements affects directly description of the contact stress development in the subsoil. The more finite elements are used, the more accurate is the contact stress.

In each iteration step, it is possible to check vertical balance as a difference in the sum of the load and the resulting force (the integral) of the contact stress. The more iteration steps are used, the lesser is the difference.

The fineness of the FEM network division influences also calculations of the stress and settlement of the half-space because the network divides the domain of the loaded half-space which is being integrated into partial sub-domains where the individual increases in stress caused by the elements should be added up. *The more elements are available, the less integration points are needed for the same accuracy (Cajka, 2008). From the mathematical point of view, this finding results also from the characteristics of the composite integration formulae in integration of partial intervals and FEM convergence.* New possibilities

of FEM solution and decreasing the time for solvers and integration procedures offer the methods of parallel programming (Konečný, Brožovský and Křivý, 2010).

Acknowledgements

This outcome has been achieved with the financial support of the Ministry of Industry and Trade of the Czech Republic, program TIP, project No. FR-TI2/746 Rheological sliding joint with thermo-controlled viscoelastic properties.

References

ČSN 73 1001 Zakládání staveb. Základová půda pod plošnými základy. Foundation of structures. Subsoil under shallow foundations. 8/1988, Z1 9/2009, zrušena 04/2010

ČSN EN 1997-1 (73 1000) Eurokód 7: Navrhování geotechnických konstrukcí - Část 1: Obecná pravidla. Eurocode 7: Geotechnical design - Part 1: General rules, ČNI 9/2006

ČSN EN 1992-1-1 (73 1201) Eurokód 2: Navrhování betonových konstrukcí - Část 1-1: Obecná pravidla a pravidla pro pozemní stavby. Eurocode 2: Design of concrete structures - Part 1-1: General rules and rules for buildings ČNI 11/2006

Abdel-Rahman, M. A. and Edil, T. B. (1991), An Integration Algorithm for Constitutive Models with Two Intersecting Yield Surfaces. Computer-Aided Civil and Infrastructure Engineering, 6: 99–108. doi: 10.1111/j.1467-8667.1991.tb00180.x

Qian, L. and Zhang, H. (1993), An improved Non-reflecting Boundary Method for Soil-Structure Interaction Analysis. Computer-Aided Civil and Infrastructure Engineering, 8: 291–297. doi: 10.1111/j.1467-8667.1993.tb00214.x

Provenzano, P. (2003), A Fuzzy-Neural Network Method for Modeling Uncertainties in Soil-Structure Interaction Problems. Computer-Aided Civil and Infrastructure Engineering, 18: 391–411. doi: 10.1111/1467-8667.00326

Katzenbach, R., Schmitt, A. and Turek, J. (2005), Assessing Settlement of High-Rise Structures by 3D Simulations. Computer-Aided Civil and Infrastructure Engineering, 20: 221–229. doi: 10.1111/j.1467-8667.2005.00389.x

Souli, M. and Shahrour, I. (2012), A coupling method for soil structure interaction problems. Int. J. Numer. Anal. Meth. Geomech.. doi: 10.1002/nag.2080

Čajka, R. (2003) Numerical Analysis of Contact Pressure under Shallow Foundation. International Symposium on Shallow Foundations FONDSUP 2003, 5.-7. November 2003, Paris, France, ISBN 2-7208-0355-3.

Čajka, R. (2005) Soil – structure interaction in case of exceptional mining and flood actions, COST 12 – Final Conference Proceedings, 20th – 22nd January 2005, University of Innsbruck, Austria, ISBN 04 1536 609 7.

Čajka, R., Manasek, P. (2005) Building structures in danger of flooding. IABSE Conference on Role of Structural Engineers Towards Reduction of Powerty, February 19-22, 2005, New Delhi, India, ISBN 3-85748-111-0

Čajka, R. (2008) Interakce základových konstrukcí s podložím (in Czech). Monografie, Ediční středisko VŠB – TU Ostrava, leden 2008, 177 s., ISBN 978-80-248-1754-5

Čajka, R. (2010) Program MKPINTER pro řešení interakce mezi základovou konstrukcí a podložím pomocí izoparametrických prvků s využitím Jacobiánu transformace a numerické integrace (in Czech). Software evidenční AR-SW-2010-04, ARMING spol. s r.o. 2010

Davis, P., Rabinowitz, P. (1956) Abscissas and Weights for Gaussian Quadratures of High Order. Journal of Research of the National Bureau of Standards, Vol. 56, No.1, January 1956, pp.35-37.

Kolář, V., Nemec, I. (1989) Modelling of Soil Structures Interactions. Praha, Academia 1989. 336 s,

Konečný, P., Brožovský, J., Křivý, V. (2010) Towards Parallel Computing using the Simulation-based Reliability Assessment Method. In Proceedings of the First International Conference on Parallel, Distributed and Grid Computing for Engineering. Pécs : Civil Comp Press, Hungary, 2010, ISBN 978-1-905088-29-4

Mindlin, R.D. (1951) Influence of Rotatory Inertia and Shear on Flexural Motions of Isotropic, Elastic Plates. Journal Appl. Mechanic 18, 1951, č.1, str. 31-38

Poulos H.G., Davis, E.H. (1974) Elastic Solutions for Soil and Rock Mechanics. John Wiley & Sons, Inc., New York 1974

Reitinger, R., Švejda, L. (1998) Výpočet vrstevnatého podloží metodou konečných prvků (in Czech). Stavitel, 5/98, pp. 32 –33.

Large Power Transformer Reliability Modelling

Mahfoud CHAFAI, Hamid BENTARZI and Abderrahmane OUADI
Signals and Systems Laboratory, IGEE, UMBB University, Boumerdes, ALGERIA
E-mail: sisylab@yahoo.com

Abstract: In this paper a simplified reliability model is developed on the basis of knowledge, from field data, of the dominating failure modes and mechanisms of high voltage transformer. Field data failures distribution indicates a dominance of failure modes pertaining to winding-tank oil insulation and terminal connections.

The transformer system can be regarded as a combination of three fundamental parts: winding inter-turn insulation, tank oil insulation and the terminals, which are respectively physical and electrical in nature. On this basis, the transformer system reliability block diagram is modelled in a series configuration comprising the above mentioned parts. The individual reliability functions developed for each part will yield together to the overall transformer system reliability.

Keywords: Transformer, failure rate, mechanism, reliability.

1. Introduction

Power transformers are critical elements in power systems. The reliability of power grids depends on the condition of high-voltage transformers that are expensive and time-consuming to be repaired or replaced by new one. Transformer failures can cause costly disruptions and sometimes involve oil spills, fires, and equipment damage. Therefore, there is a great need to maximize the availability and safety of this machine and hence its reliability.

Reliability is a very important performance parameter of transformer but its modeling remains a very complex problem. However, in the light of field data, dominant failure modes of some transformer system parts can be used to develop a simple but credible reliability model.

2. Transformer Failure Mechanisms

Better understanding of transformer failure mechanisms and risk factors has helped development efforts to improve condition assessment and life extension. The causes of transformer failure can generally be classified accordingto the failure distribution statistics given in (Hattangadi, 1999) as follows:

- Failure of inter-turns insulation in the main windings,
- Oil insulation failure between the winding and the tank,
- Defect in internal terminals (Bushing).

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

2.1. WINDINGS INSULATION FAILURE MECHANISM

The winding insulation is always subjected to the combined thermal, electrical, mechanical and environmental stresses during the transformer operation.

- *Thermal stress.* After a certain time, the insulation will start to lose its characteristic due to the normal thermal aging process. But, the occurrence of premature failures, which are predominant, are a direct result of an over-current caused generally by an overload and over-voltage.
- *Electrical stresses.* Most of electrical failures are caused by a combination of over-voltage spikes and normal deterioration. This fast over-voltage can be caused by switching, lightning, and surges to propagate through the machine which leads to a fast breakdown.
- *Electromechanical stresses and contamination.* The level of electromechanical factors (vibrations, expansion.) and environmental contamination are other stresses that can have effect on the dielectric strength of the winding wire insulation.
  The failure mechanisms sequence of the winding insulation is summarized in Fig. 1.

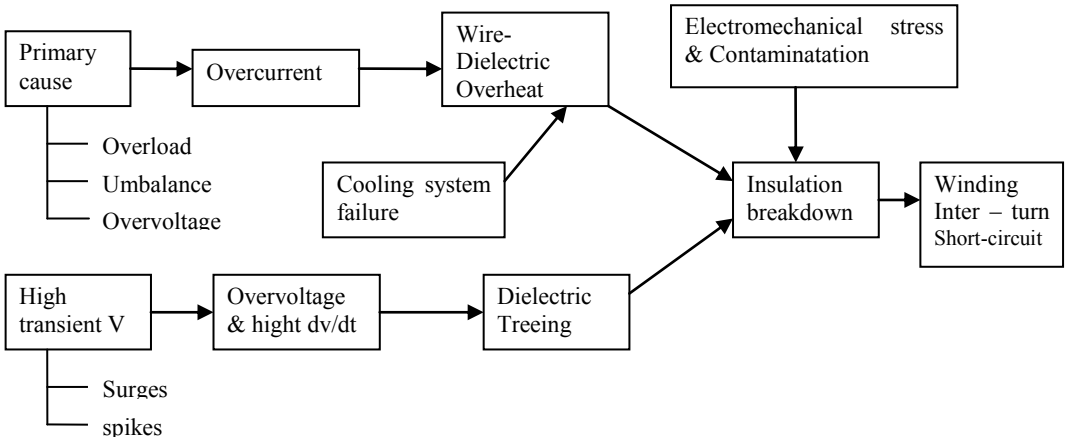


Figure 1. Failure mechanism sequence of the electrical windings insulation.

2.2. OIL INSULATION FAILURE MECHANISM

Oil in power transformers serves as a cooling medium as well as an insulation. Like the winding insulation, the oil insulation is always subjected to the thermal, electrical and environmental stresses. The factors that has greatly influence on the aging process of oil insulation are: moisture, acid formation and contaminant, over-temperature, and the addition of the air (oxygen) (Aquilino, 1983) and (Minhas, 1999) as shown in Fig. 2. The transformer oil is oxidized under the influence of excessive oxygen and temperature, particularly in the presence of small metallic particles which act as catalysts, resulting in an increase in acid amount.The failure mechanisms sequence of oil is illustrated in the Fig. 2.

2.3. TERMINALS FAILURE MECHANISM

Degradation of connector quality occurs at the separate or permanent interfaces due to loss in contact area by several mechanisms including fretting corrosion, wear and loss in contact force.

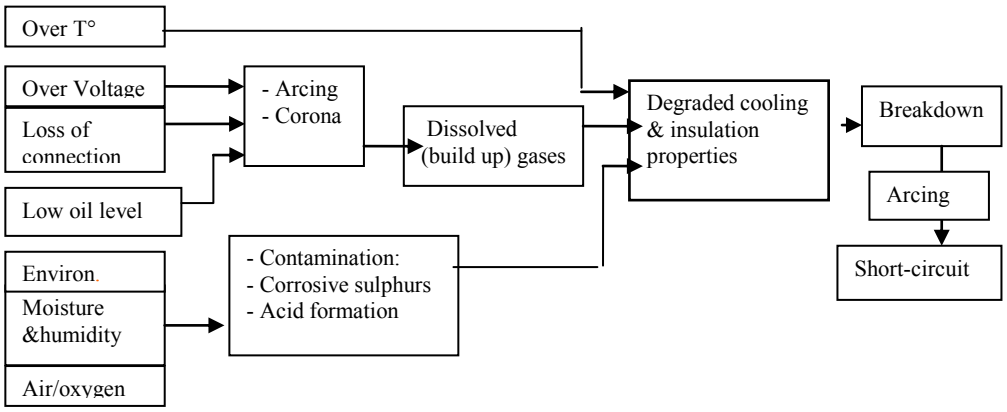


Figure 2. Failure mechanisms sequence of the oil insulation.

The field data experience indicates that common problem of terminal connection follows failure mechanisms with cumulative effects such fretting corrosion (Bloch and Geitner, 1999) and (Braunovic, Konchits, and Myshkin 2006). In this latter, motion-induced corrosion, driving forces include vibrations, mechanical and thermal shock, and thermal expansion mismatch due to temperature cycling and the number of fretting cycles. Loss of asperity contact surface, due to the generated corrosion film or contamination, can result in contact interface (constriction) resistance increases that are sufficient to lead to connector failures. This gradual degradation is accelerated under the effects of environmental influences such as humidity, temperature, bias voltage leading after time duration to much consequential damages such as, overheat and possible fracture and arcing and contamination as illustrated in Fig. 3.

3. Power Transformer Reliability Model

The instantaneous failure rate experienced by power transformers is not constant but increases with time. The model is dictated by three failure modes: winding insulation, oil insulation and terminals.

The transformer system can be considered as combined electrical and mechanical parts where it may be assumed that the transformer fails when any of its parts fails such that its reliability bloc-diagram is modeled in series configuration, as shown in Fig. 4, which can be expressed mathematically as follows:

$$R_T = R_t . R_w . R_o . \quad (1)$$

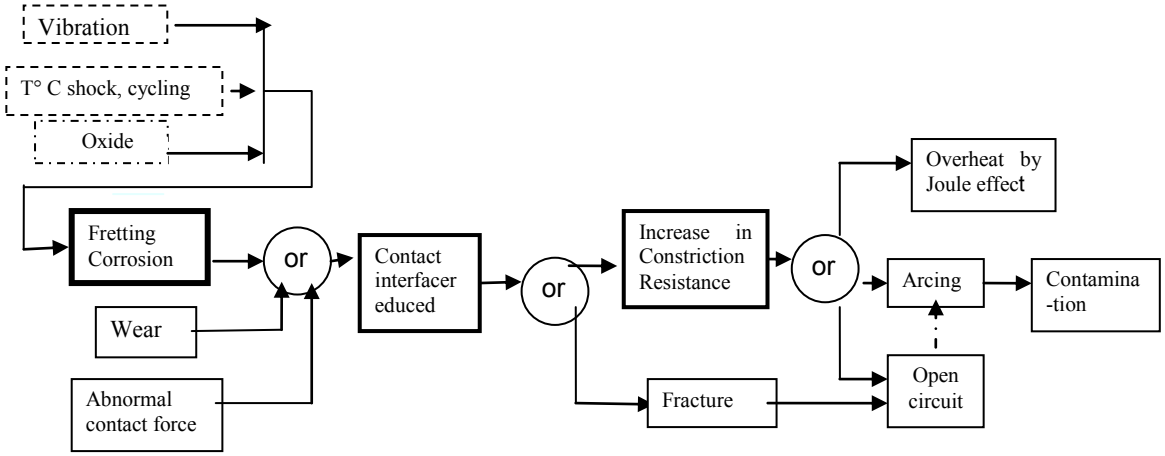


Figure 3. Failure mechanism of terminal connection.

3.1. WINDING INSULATION RELIABILITY MODEL

Since the insulation element can be represented by capacitance of high dielectric strength in parallel with a high value resistance its reliability is modeled in exponential form where the failure rate is constant. This failure rate is expressed as base failure rate times series of multiplicative stress factors related failure mechanisms discussed in section 2 with the following general form (Standard, MIL-HDBK-217 F):

$$\lambda p = \lambda b \,.\, \pi_Q \,.\, \pi_E \,.\, \pi_T \,.\, \pi_S \,, \tag{2}$$

where λb = base failure rate, π_Q = quality adjustment factor, π_E = environment adjustment factor, π_T = temperature adjustment factor, π_S = electrical stress adjustment factor. Assuming that the MTBF of class A insulation is 105h then the base failure rate is given by:

$$\lambda b = 1/\text{MTBF} = 1/100000 = 10 \text{ F}/106 \text{ h} \,. \tag{3}$$

3.1.1. *Thermalacceleration stress factor*

The temperature acceleration factor is given by the Arrhenius model (MIL-HDBK-217 F):

$$\pi_T = \frac{\lambda_2}{\lambda_1} = \frac{\text{MTBF}_1}{\text{MTBF}_2} = e^{B_T\left(\frac{1}{T_1}-\frac{1}{T_2}\right)}, \tag{4}$$

where constant B_T is determined using the half life rule for each 10°C rise of the winding insulation temperature (Standard, MIL-HDBK-217 F) as:

$$B_T = \frac{\ln\left(\frac{1}{2}\right)}{\left(\frac{1}{T_1} - \frac{1}{T_2}\right)} = -7.007 \cdot 10^3. \quad (5)$$

The activation energy of the insulation is deduced as follows:

$$E_a = B_T \,.\, K_B \text{ (eV)} , \quad (6)$$

where K_B = Boltzman constant (eV/K).

3.1.2. *Electrical acceleration stress factor* π_s

When a transformer is new the dielectric strength of the winding insulation system is very high. For a typical 33 kV transformer, it is assumed that the maximum initial dielectric strength is over 600 kV to ground while the dielectric strength breakdown is greater than or equal to 120 kV. The fast over-voltage stress will accelerate the degradation process of the insulation strength and hence this will lead rapidly to a breakdown. According to the degradation mode of the dielectric strength whose voltage dependence is modelled with the following hyperbolic equation (Lanham, 2002):

$$V_s = 600 \cdot \sqrt{1 - \left(t / 20\right)^2} \quad (7)$$

Whose graphical representation is shown in Fig. 5.

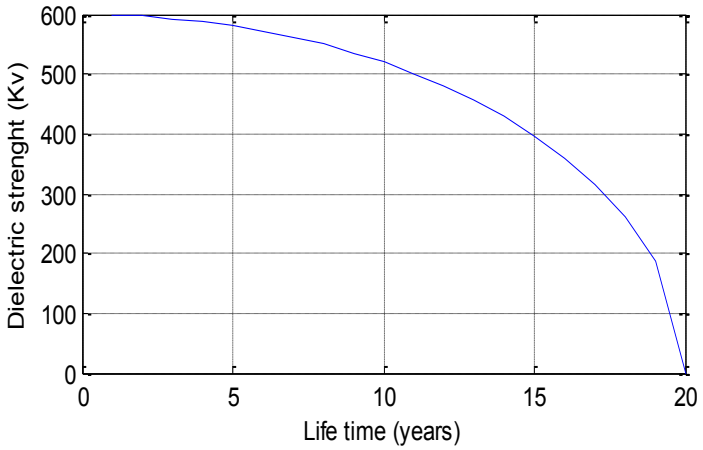


Figure 5. Dielectric strength degradation of the winding insulation.

While, the time for given voltage may be determined from:

$$t = 20 \cdot \sqrt{1 - \left(V_s / 600\right)^2} \quad (8)$$

The time to failure (breakdown) is evaluated as follows:

$$t_f = 19.34 - t \tag{9}$$

Hence, the degradation acceleration factor of the dielectric strength is evaluated as follows:

$$\pi_S = t_{f1}/t_{f2} \tag{10}$$

Table I. Acceleration voltage stress factor

Parameters	V_s (kV)	t (s)	t_f (s)	$\pi_S = t_{f1}/t_{f2}$
Values	600	0	19.34	1
	500	11.05	9.29	2.4

3.1.3. *The quality and environment stress factors*

Assuming that class F insulation of this transformer has a best quality factor ($\pi_Q = 1$) and that the environment includes dust, dirt, corrosion and humidity is classed as GF according to (Standard, MILHDBK), whose accelerating stress factor is corresponding to $\pi_E = 2$.

3.1.4. *The overall winding insulation failure rate*

In the case of the following conditions such as 5°C temperature rise and the dielectric strength of the insulation reaching 500 kV, the π–correction factors are given in the table II.

Table II. π-correction factors

Parameters	λb	π_Q	π_E	π_T	π_s	λp
Values	10	1	2	1.4	2.4	67.2

The overall failure rate is calculated as:

$$\lambda p = \lambda b \,.\, \pi_Q \,.\, \pi_E \,.\, \pi_T \,.\, \pi_S = 59.40 \text{ F}/10^6 \text{ h} \,. \tag{11}$$

Hence, the reliability expression of the windings insulation is given as follows:

$$R_w(t) = e^{-67.2t} \,. \tag{12}$$

This is represented graphically in Fig. 6.

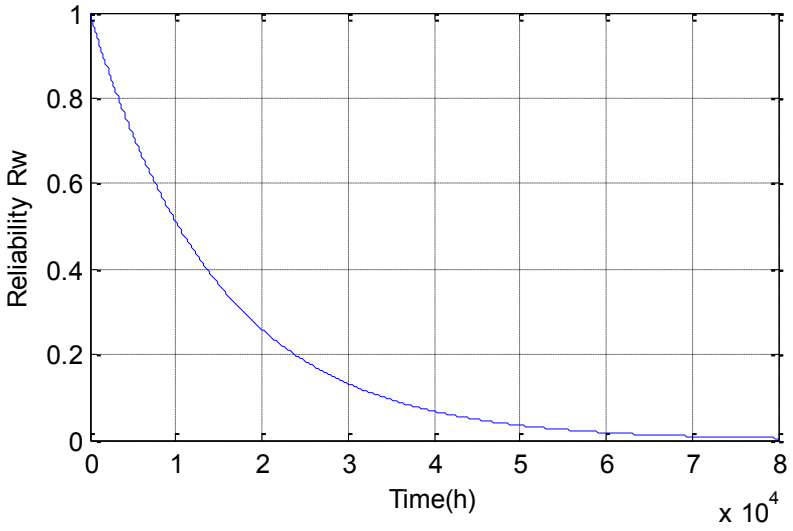


Figure 6. Reliability curve of the winding insulation.

3.2. OIL INSULATION RELIABILITY MODEL

As in the case of winding insulation, the part failure rate of the oil insulation is expressed as base failure rate times series of multiplicative stress factors related failure mechanisms given with the following general form:

$$\lambda p = \lambda b \,.\, \pi_Q \,.\, \pi_E \,.\, \pi_T \,.\, \pi_S \,. \tag{13}$$

If the life duration of the used oil specimen at 60°C is equal to 249,526 h, then the base failure rate is λb = 1/MTBF = 1/249,526 h = 4.007 F/10^6 h.

3.2.1. *Temperature acceleration factor*

The critical factor which determines the aging is the temperature of the transformer oil. The durability of the transformer is reduced by a period proportional to the duration of the overload and exponentially proportional to the excess of oil temperature above 90° C (Hattangadi, 1999). For every 10°C rise in the oil temperature above 90°, the age of the transformer is reduced by a half. Using equation (3) and (4), π_T is calculated for 5° C rise and is given further in table III.

3.2.2. *Dielectric strength*

According to the degradation mode of the dielectric strength similar to that of the winding insulation with the following assumed data:

- Maximum (initial) dielectric strength: 400 kV;
- Dielectric strength breakdown: 100kV,
- Life duration: 30 years.

Its degradation mode is modeled with the following hyperbolic equation (Lanham, 2002):

$$V_s = 400 \cdot \sqrt{1-(t/30)}^2 \tag{14}$$

Whose graphical representation is shown in Fig. 7.

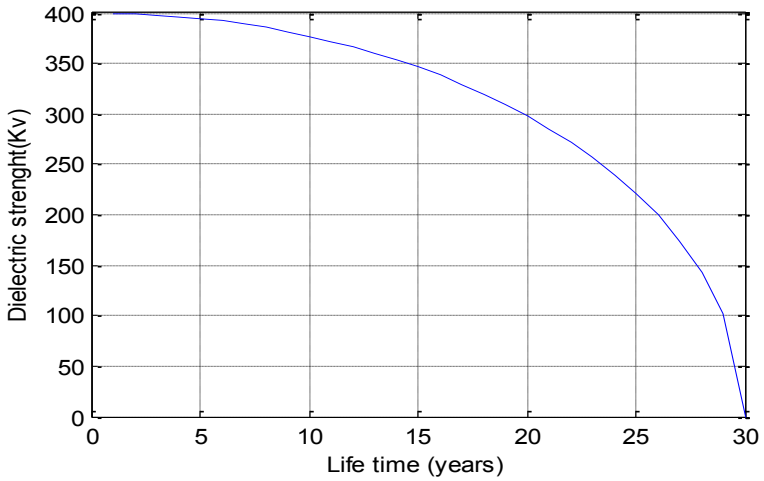


Figure 7. Dielectric strength degradation of the oil insulation.

Hence, the degradation acceleration factor of the dielectric strength is evaluated as $\pi_S = 2.51$.

3.2.3. *Quality and environment stress factors*

Assuming that class of the oil insulation of this transformer has a best quality factor ($\pi_Q = 1$) and that the environment is similar to that of the winding ($\pi_E = 2$).

3.2.4. *Overall oil insulation failure rate*

In the case of the following conditions such as 5°C temperature rise and the dielectric strength of the insulation has reached 300kV. The π-correction factors are given in the table III.

Table III π-correction factors

Parameters	λb	π_Q	π_E	π_T	π_s	λp
Values	4.07	1	3	1.4	2.51	42.16

The overall failure rate is calculated as:

$$\lambda p = \lambda b \,.\, \pi_Q \,.\, \pi_E \,.\, \pi_T \,.\, \pi_S = 42.16 \text{ F}/10^6 \text{ h} \,. \tag{15}$$

Hence, the reliability expression of the oil insulation is given as follows:

$$R_o(t) = e^{-42.16t} \tag{16}$$

This is represented graphically in Fig. 8.

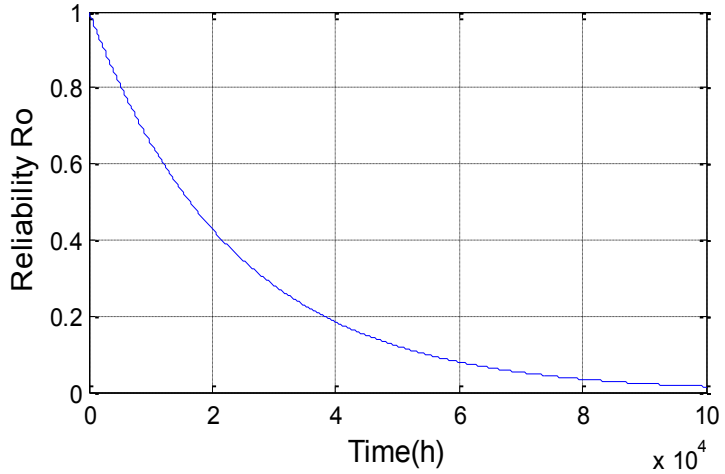


Figure 8. Reliability curve of the oil insulation.

3.3. TERMINALS RELIABILITY MODEL

Reliability modeling of connectors requires knowledge of the relationship between degradation mechanisms and connector performance parameters such as contact resistance. Fretting is known to be a major cause of contact deterioration and failure, particularly in nickel plated contacts (Braunovic, Konchits, and Myshkin, 2006).

During fretting corrosion, the contact resistance generally increases slowly as function of time.

Assuming that the contact resistance *Re* is simulated to n equal resistances, of value *R*, in parallel configuration then the equivalent resistance will be:

$$Re = R/n \tag{18}$$

The fretting corrosion expansion is equivalent to loss of one or several paralleled resistance *R* and hence leading to an increase in the equivalent contact resistance *Re*.

The direct measurement of dynamic constriction resistance of terminal remains difficult and unsafe in HV substation. For this, Non-contact and distant temperature measurement based on infrared technique is used for ensuring the safety.

According to heat transfer theory, the total thermal power losses in a resistive connector are mainly due to conduction and radiation processes in a given environment. It is expressed as follows:

$$P_{los} = P_{cond} + P_{rad} = (K \,.\, A_s/L)\,(T - T_a) + \sigma \,.\, \varepsilon \,.\, A \,.(T4 - T_a4) \tag{19}$$

where K = thermal conductivity of copper (= 401)
L = length (thickness) of the portion of wire
A_s = area of section of the wire
T = temperature of the connector

T_a = ambient temperature
σ = Stephan–Boltzmann constant (= 5.67. 10^{-8} $W.m^{-2}.K^{-4}$)
ε = emission coefficient or emissivity
A_c = area of the connector or exchange
In the other side, according to Joule effect the rate of heatdissipation in a resistiveconductor (such as the connector) is proportional to the square of the current through it and to its resistance as expressed below:

$$P_j = R \,.\, I^2$$

Assuming that the rate of dissipated electrical energy is equal to the sum of the rates of the conductive and radiated thermal energies, then, the ohmic resistance of the terminal connector, through which flows an ac current of 100A rms value, can be deduced as follows:

$$R = P_j / I^2 = P_{Loss} / I^2 \tag{20}$$

The handheld pyrometers enables temperature precision point measurement once the hot spot or area of interest is identified by means of a two dimensional thermal imager.

Connector temperature measurements were performed on nickel plated copper alloy contacts sockets of 100A rated transformer of cement industry manufactured by ENEL at a given fretting cycle and load. For each rise in temperature T, due to an increase in contact resistance R, time durations are censored and ranked in table IV.

Table IV. Life duration of terminal connectors resistance (Swingler, 2002)

Measured T variation (°C)	Contact resistance (μΩ)	Time duration (h)	*Fi*
5	39.5	4530	10
5	80.9	6080	20
10	124.3	7800	30
10	217.3	9020	40
15	373.2	11070	50
15	550.1	12500	60
25	851	14100	70
20	1200	15120	80
>20	10000	17060	90

Where, Fi = Repartition function.

The value of resistance change is varying from tens of micro-ohms to ohms, and even open circuit (Halliday & Resnick, 2011). However, when the contact resistance increases sufficiently to raise the local temperature, a self-accelerating deterioration resulting from the interaction of thermal, chemical, mechanical and electrical processes will be triggered, and the contact resistance will rise abruptly as shown in Fig. 8.

The cumulative effect of fretting corrosion is modeled according the Weibull distribution (Swingler and McBride 2002) where the cumulative probability function of failure is given as follows:

$$F(t) = 1 - e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}} \,. \tag{19}$$

Where β is a shape parameter of the distribution and η is a scale parameter (life characteristic of contact interface).

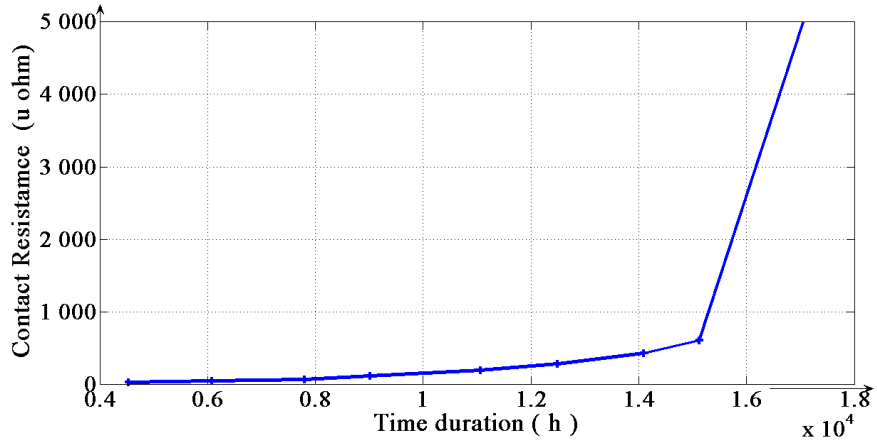


*Figure 8.*Constriction resistance change as function of time.

Assuming that starting point $\gamma = 0$, the Weibull scale and shape factor were evaluated using Allen Plait Chart as $\eta = 120000$; $\beta = 2.3$ that is characterizing the cumulative effect such as corrosion. And hence the reliability will be expressed as follows:

$$R_t(t) = e^{-\left(\frac{t}{12000}\right)^{2.3}} \tag{20}$$

This is represented graphically as shown in Fig. 9.

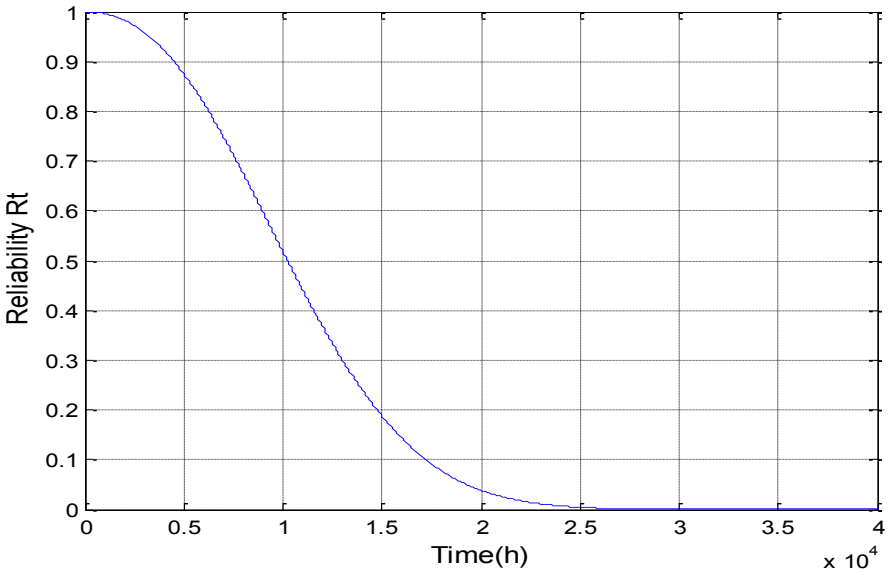


Figure 9. Reliability curve of the oil insulation.

3.4. Overall Reliability Of The Transformer

According to equation (1) the overall reliability of the large power transformer is obtained:

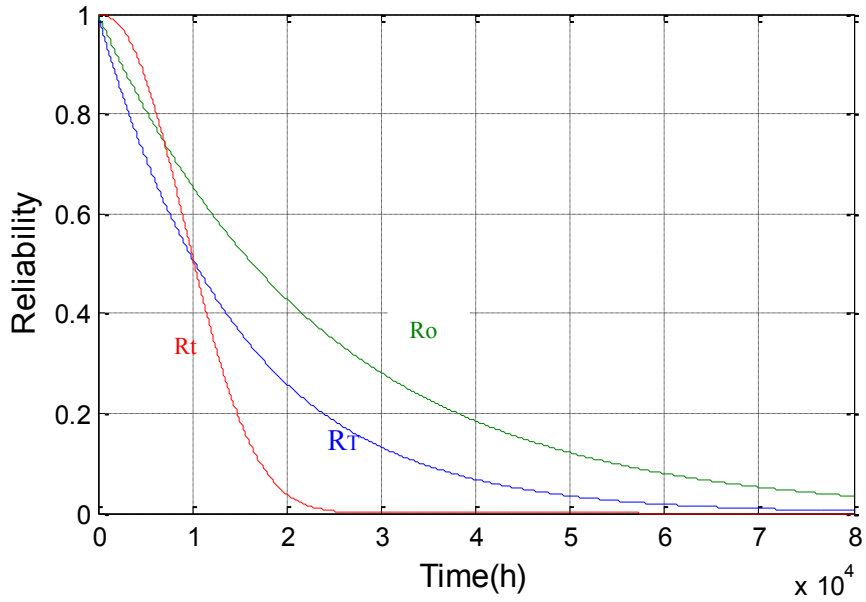
$$R_T(t) = e^{-67.2t} \cdot e^{-42.16t} \cdot e^{-\left(\frac{t}{12000}\right)^{2.3}} \tag{21}$$

The corresponding graphical representation is given in Figs. 10 and 11.

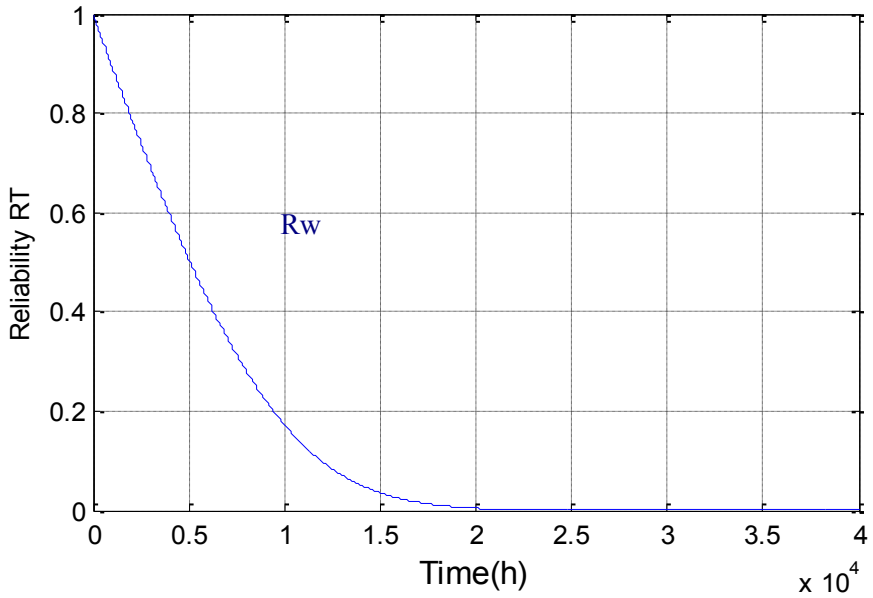
4. Conclusion

In this study, it is shown that both winding and oil insulation reliabilities, because of their electro-chemical nature, are adequately represented by an exponential form function, while the Weibull law is well suited to the cumulative effect of fretting corrosion at the terminals.

The obtained overall reliability of the transformer essentially follows an exponential form function for a limited time interval, but with a relatively much higher constant failure rate, than that of the winding and oil insulations as shown in Fig. 11, because of the 'precipitating' effect of corrosion in terminal reliability slope. Hence, it may be concluded that the remaining life of large transformer is decided by insulation to a great extent and the life of the terminal connection system.



a - Separated reliabilities



b - Combined reliability

Figure 10. The overall reliability of the Power transformer: a - Separated reliabilities, b -Combined reliability

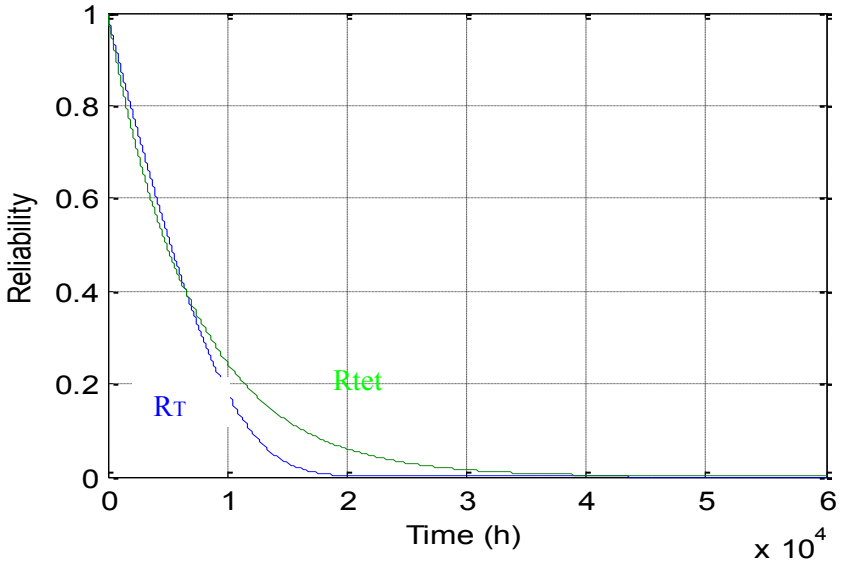


Figure 11. Graphical comparison of the overall reliability of the transformer with an exponential form reliability of constant failure rate (≈ 170 F/10^6 h)

References

Hattangadi A., *Electrical Fires and failures*, Mc.Graw–Hill, 1999.

Aquilino J. W., IEEE Report of Transformer Reliability Survey-Industrial Plants and Commercial Buildings, *IEEE Transactions on Industry Applications*, vol. IA 19, no. 5, September /October ,1983.

Minhas M., Reynders PJ de Klerk, Failures in power system transformers and appropriate monitoring technique; High Voltage Engineering Symposium, 22-*27 August 1999 Conference Publication,* No. 467, 0 IEE, 1999.

Bloch H.P., and Fred K. Geitner, *Machinery failure analysis and troubleshooting*, Vol 2, 3th ed, Elsevier , 1999.

Braunovic M., Konchits V., and Myshkin NK, *Fundamentals of Electrical Contacts*, CRC Press, Taylor and Francis group 2006.

Standard MIL-HDBK-217 F

Lanham C., *Understanding the Tests that are Recommended for Electric Motor Predictive Maintenance*, Baker Instrument Company, Energy publication, 2002.

SwinglerJ.and McBride J.W., Fretting Corrosion and the Reliability of Multicontact Connector Terminals, *IEEE Transactions on components and packaging*, Vol .25 ,No 4,December 2002.

Halliday & Resnick, *Principles of Physiscs*, John Wiley & Sons, Inc , 9th Edition, 2011

ENEL Company, Rapports internes de suivi et d'inspection des transformateurs HTinstallés dans les postes de distribution Cimenteries de Chlef Algeria.

Mirzai M., Gholami A., Aminifar F., Failures Analysis and Reliability Calculation for Power Transformers, *J. Electrical Systems* 2-1: 1-12, 2006.

On the effect of material spatial randomness in lattice simulation of concrete fracture

Jan Eliáš and Miroslav Vořechovský
Institute of Structural Mechanics, Faculty of Civil Engineering, Brno University of Technology,
elias.j@fce.vutbr.cz, vorechovsky.m@fce.vutbr.cz

Abstract. The paper presents stochastic discrete simulations of concrete fracturing. The spatial material randomness of local material properties is introduced into a discrete lattice-particle model via an autocorrelated random field generated by the Karhunen–Loève expansion method. The stochastic discrete model is employed to simulate failure of three-point-bent beams with and without a central notch notch. The effect of spatial randomness on the peak load and energy dissipation is studied.

Keywords: lattice model, concrete, fracture, stochastic simulations, material randomness, fracture energy, flexural failure

1. Introduction

It has been widely recognized that mechanical properties of materials exhibit a spatial variability. The seminal theory of (Weibull, 1939) offered simple and powerful tool to determine the probabilistic distribution of structural strength. However, applicability of the Weibull theory is limited to brittle structures with no redistribution prior to the peak load. The Weibull theory lacks any length scale and rupture of infinitely small volume directly causes failure of the whole structure. The absence of any characteristic length scale also results in spurious infinite strength of infinitely small structures (Vořechovský, 2010). Moreover, the Weibull theory assumes that strength of every material point is independent of its surroundings. However, many structures are made of quasibrittle materials like concrete, ceramics, rocks or ice. These structures have the ability to partially redistribute released stresses and thus their failure is triggered by rupture of some representative volume of finite size. Also the Weibull assumption of independence stands out against the natural expectation that the local strength fluctuate rather continuously inside a structure.

The advantage of Weibull theory comes from the fact that the mechanics of failure does not interact with the stochastic model – only elastic stress field is needed. Extension of the Weibull theory for finite internal material length scale requires knowledge of changes in the stress field during the redistribution prior to the peak load. The redistribution can be mimicked by the nonlocal Weibull theory of (Bažant and Xi, 1991) and (Bažant and Novák, 2000), where probability of failure of material point depends not only on its local stress but also on stress in its surroundings. Therefore, local stress is replaced by nonlocal stress obtained by nonlocal averaging of the (local) elastic stress field (Bažant and Jirásek, 2002). The nonlocal Weibull theory agrees for the large sizes with the local one. For intermediate structural sizes, it predicts higher strengths than the local Weibull theory thanks to possible stress redistribution. Unfortunately, in the in the case of very small structures, the theory is not applicable because the approximation or stress redistribution by nonlocal

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

averaging is too simplistic. Though the nonlocal averaging helps to introduce the material internal length, it is not able to correctly reflect possible spatial correlations of local material properties.

A laborious option of structural strength estimation is represented by stochastic failure simulations that include proper mechanics of stress redistribution. Such a stochastic analysis can be performed using the finite element method with a sophisticated material constitutive law (Vořechovský, 2007; Vořechovský and Sadílek, 2008). Failure of highly heterogeneous materials can also be advantageously modeled via discrete models. These models can be *deterministic*: (Grassl and Rempling, 2008; Van Mier and Van Vliet, 2003; Bolander and Saito, 1998) or *stochastic*: (Grassl and Bažant, 2009; Alava et al., 2006). In this study, we adopt the lattice particle-model developed by G. Cusatis (Cusatis and Cedolin, 2007) for modeling of concrete fracturing. Spatial material fluctuations are introduced by modeling the material properties as realizations of a random field.

The following Section 2 briefly describes the deterministic mechanical (lattice) model and Section 3 elucidates how the spatial randomness is incorporated into the model. The model is then used for numerical simulations of failure of notched and unnotched three-point bent beams. The results are presented in Sections 4 (notched beams) and 5 (unnotched beams).

2. Deterministic model

Modeling of the initiation and propagation of cracks in quasibrittle materials exhibiting strain softening has been studied for several decades. Although this is a difficult task complicated by the distributed damage dissipating energy within a fracture process zone (FPZ) of non-negligible size, realistic results have been achieved by several different approaches; see e.g. (Bažant and Planas, 1998). The present study is based on the cohesive crack model (Barenblatt, 1962; Hillerborg et al., 1976; Bažant and Planas, 1998) called sometimes the fictitious crack model. It relies on an assumption that the cohesive stress transmitted across the crack is released gradually as a decreasing function of the crack opening, called the cohesive softening curve. Its main characteristic is the total fracture energy, G_F – a material constant representing the area under the softening curve.

In heterogeneous materials, the dissipation of energy takes place within numerous meso-level cracks inside the FPZ. Direct modeling of such distributed cracking calls for representation of the material meso-level structure. Models capable to efficiently incorporate the concrete meso-structure should be used. For this purpose, the present analysis will be based on the discrete lattice-particle developed by (Cusatis and Cedolin, 2007), which is an extension of (Cusatis et al., 2003; Cusatis et al., 2006).

The material is represented by a discrete three-dimensional assembly of rigid cells. The cells are created by tessellation according to pseudo-random locations and radii of computer generated aggregates/particles. Every cell contains one aggregate (Fig. 1a,b). The cells are interconnected by set of three nonlinear springs (normal - n and two tangential - t_1, t_2) placed at the interfaces between the cells, representing the mineral aggregates in concrete and its surroundings. On the level of rigid cell connection, the cohesive crack model is used to represent cracking in the matrix between the adjacent grains. The inter-particle fracturing is assumed to be of damage-mechanics type and is modeled using a single damage variable ω applied to all three directions $i = n, t_1$ and t_2. Forces F_i in the springs can thus be evaluated from their extensions Δu_i by

$$F_i = (1 - \omega) k_i \Delta u_i \tag{1}$$

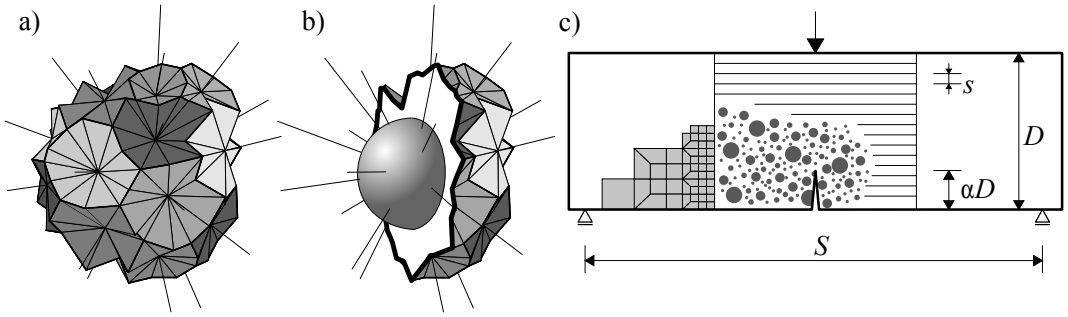


Figure 1. a) One cell of the lattice-particle model and b) its section revealing the aggregate. c) Geometry of the beams simulated in three-point-bending.

where k_i is elastic spring stiffness. The damage parameter ω depends on Δu_i and on the previous loading history of each connection. For a detailed description of the connection constitutive law or other model features, see (Cusatis and Cedolin, 2007). The confinement effect (present in the full version of the model) is neglected here as it was estimated that confinement does not play any important role in the studied type of experiment.

Beams of depths $D = 300$ mm, span-depth ratio $S/D = 2.4$ and thickness $t = 0.04$ m, were modeled. The maximal aggregate diameter was 9.5 mm. The minimal grain diameter was selected as 3 mm. Aggregates' diameters within the chosen range were generated using the Fuller curve. The parameters of the connection constitutive law, which were mostly taken similar to those in (Cusatis and Cedolin, 2007), were: matrix elastic modulus $E_c = 30$ GPa; aggregate elastic modulus $E_a = 90$ GPa; meso-level tensile strength $\sigma_t = 2.7$ MPa; meso-level tensile fracture energy $G_t = 30$ N/m; meso-level shear strength $\sigma_s = 3\sigma_t = 8.1$ MPa; meso-level shear fracture energy $G_s = 480$ N/m; meso-level compressive strength $\sigma_c = 42.3$ MPa; $K_c = 7.8$ GPa; $\alpha = 0.25$; $\beta = 1$; $\mu = 0.2$; $n_c = 2$.

To save computer time, the lattice-particle model covers only the region in which cracking was expected. Surrounding regions of the beams were assumed to remain linear elastic and were therefore modeled by standard 8-node isoparametric finite elements. The elastic constants for these elements were identified by fitting a displacement field with homogeneous strain to displacements of particle system subjected to low-level uniaxial compression. The macroscopic Young's modulus and Poisson ratio were found to equal $\bar{E} = 34.7$ GPa and $\bar{\nu} = 0.19$. The finite element mesh was connected to the system of particles by introducing interface nodes treated as auxiliary zero-diameter particles (Eliáš and Bažant, 2011). These auxiliary particles have their translational degrees of freedom prescribed by shape (or interpolation) functions of the nearest finite element. The rotations of the auxiliary particles were unconstrained.

3. Stochastic model

In the described discrete model, we assign material properties of each inter-particle connection according to a stationary autocorrelated random field. The value of the c-th realization of the discretized field at spatial coordinate $\boldsymbol{x}$ will be denoted $\boldsymbol{H}^c(\boldsymbol{x})$. For a given coordinate $\boldsymbol{x}_0$, $\boldsymbol{H}(\boldsymbol{x}_0)$ is a random variable H of cumulative distribution function (cdf) $F_H(h)$. Since we work with stationary random fields, the cdf $F_H(h)$ is identical for any position $\boldsymbol{x}_0$. Recent studies by Bažant and co-workers (Bažant and Pang, 2007; Bažant et al.,

2009) showed that, when H represents strength of a quassibrittle material, $F_H(h)$ can be approximated by a Gaussian distribution onto which a power-law tail is grafted from the left at a probability about 10^{-4}–10^{-3}.

$$F_H(h) = \begin{cases} r_f \left(1 - e^{-\langle h/s_1 \rangle^m}\right) & 0 \le h \le h_{gr} \quad (2a) \\ F_H(h_{gr}) + \dfrac{r_f}{\delta_G \sqrt{2\pi}} \displaystyle\int_{h_{gr}}^{h} e^{-(h-\mu_G)^2/2\delta_G^2} \mathrm{d}h & h > h_{gr} \quad (2b) \end{cases}$$

where $\langle x \rangle = \max(x, 0)$, $s_1 = s_0 r_f^{1/m}$, m is the Weibull modulus (shape parameter) and s_0 is scale parameter of the Weibull tail, μ_G and δ_G are the mean value and the standard deviation of the Gaussian distribution that provides the Gaussian core. The Weibull-Gauss juncture at point at h_{gr} requires that that $(\mathrm{d}F_H/\mathrm{d}h)|_{h_{gr}^+} = (\mathrm{d}F_H/\mathrm{d}h)|_{h_{gr}^-}$. r_f is a scaling parameter normalizing the distribution to satisfy $F_H(\infty) = 1$. The distribution has in total 4 independent parameters.

The spatial fluctuation of the field is characterized through an autocorrelation function. It determines the spatial dependence pattern between the random variables at any pair of nodes. The correlation coefficient ρ_{ij} between two field variables at coordinates $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ can be assumed to obey the squared exponential function:

$$\rho_{ij} = \exp\left[-\left(\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{d}\right)^2\right] \quad (3)$$

It brings a new parameter d called the autocorrelation length.

To digitally simulate the stationary random field described by the random variable cdf F_H and correlation length d in the discrete model, we need to generate N realizations of the discretized random field $\boldsymbol{H}^0(\boldsymbol{x})$, $\boldsymbol{H}^1(\boldsymbol{x}), \ldots,$ $\boldsymbol{H}^{N-1}(\boldsymbol{x})$ at the facet centers of the model. This is achieved using the the Karhunen–Loève expansion based on the spectral decomposition of covariance matrix $\boldsymbol{C}$, where $C_{ij} = \rho_{ij}$. This procedure decompose the correlated Gaussian variables $\widehat{\boldsymbol{H}}(\boldsymbol{x}_i)$ into independent standard Gaussian variables ξ_k that are easy to generate. $c-$th realization of the Gaussian random field $\widehat{\boldsymbol{H}}^c(\boldsymbol{x})$ is then obtained using K standard Gaussian random variables by the following expression

$$\widehat{\boldsymbol{H}}^c(\boldsymbol{x}) = \sum_{k=1}^{K} \sqrt{\lambda_k} \xi_k^c \boldsymbol{\psi}_k(\boldsymbol{x}) \quad (4)$$

where λ and $\boldsymbol{\psi}$ are the eigenvalues and eigenvectors of the covariance matrix $\boldsymbol{C}$. The value K is the number of eigenmodes/variables considered. In practice, it suffices to employ only a reduced number of eigenmodes $K \ll$ order of $\boldsymbol{C}$. In particular, K can be selected such that $\sum_{k=1}^{K} \lambda_k$ corresponds to about 99% of the trace of the covariance matrix $\boldsymbol{C}$ (Vořechovský, 2008). The vectors of independent standard Gaussian variables $\boldsymbol{\xi}$ are generated by Latin Hypercube Sampling using the mean value of each subinterval. The spurious correlation of the variables is then minimized by reordering their K realizations (Vořechovský and Novák, 2009).

A non-Gaussian random field can be generated by isoprobabilistic transformation of the underlying Gaussian field as

$$\boldsymbol{H}^c(\boldsymbol{x}) = F_H^{-1}(\Phi(\widehat{\boldsymbol{H}}^c(\boldsymbol{x}))) \quad (5)$$

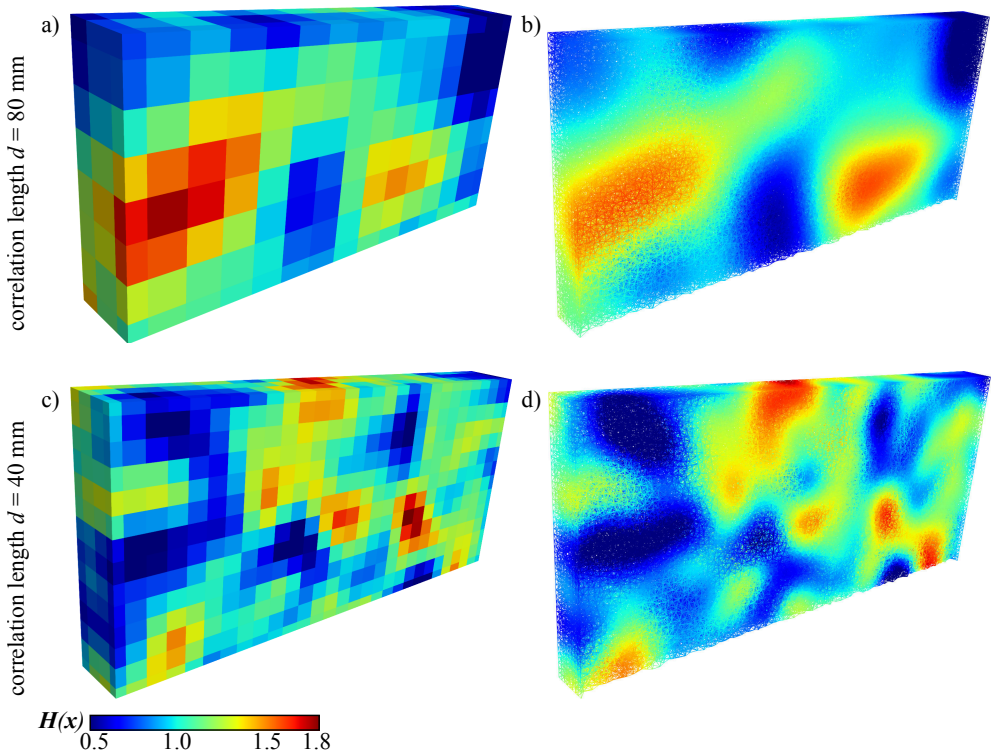


Figure 2. Left: one realization of the autocorrelated random field $\boldsymbol{H}$ on a grid of spacing $d/3$ for $d = 80$ mm (top) and $d = 40$ mm (bottom). Right: realization of the field $\boldsymbol{H}$ at the element centers of the lattice-particle model.

Such a transformation, however, distorts the correlation structure of the field. Thus, when generating Gaussian field $\widehat{\boldsymbol{H}}$, the correlation coefficients must be modified (Vořechovský, 2008). This is here performed using the approximate method of (HongShuang et al., 2008).

The realizations of the random field need to be evaluated for every shared facet (inter-particle bond) of the discrete mechanical model (at its center). This can be computationally extremely demanding for a large number of facets (large covariance matrix) and a short correlation length d (many eigenvalues needed, large K). We therefore adopted the expansion optimal linear estimation method - EOLE (Li and Kiureghian, 1993), which can significantly reduce the time of random field generation. Instead of the facet centers, the random field is initially generated on a regular grid of nodes with spacing $d/3$ (see Fig. 2). The values of the random field at the facets are then obtain from expression

$$\widehat{\boldsymbol{H}}^c(\boldsymbol{x}) = \sum_{k=1}^{K} \frac{\xi_k^c}{\sqrt{\lambda_k}} \boldsymbol{\psi}_k^T \boldsymbol{C}_{xg} \tag{6}$$

where λ and $\boldsymbol{\psi}$ are now eigenvalues and eigenvectors of the covariance matrix of the grid nodes, and $\boldsymbol{C}_{xg}$ is a covariance matrix between facet center at coordinates $\boldsymbol{x}$ and the grid nodes. After the Gaussian random field values at facet centers are obtained by EOLE (Eq. 6), they need to be transformed to the non-Gaussian space by Eq. 5.

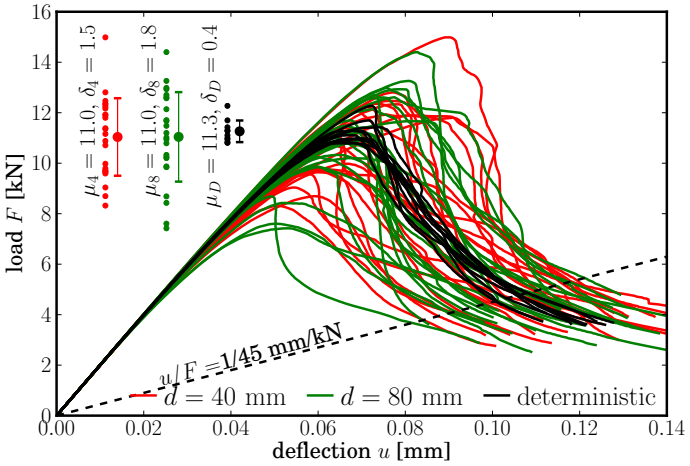


Figure 3. Load-deflection curves for simulations of TPB beams with notch.

Besides the significant time savings, another advantage of using EOLE is that one can simply use the same field realization for several different granular positions. By keeping the c-th realization of decomposed variables $\boldsymbol{\xi}^c$ unchanged, the field realization can be adapted for any configuration of the facets in the discrete model.

Structural strength of a quasibrittle material is typically governed by two important material properties, namely the material strength and fracture energy. Realistic fracture models should therefore incorporate random spatial variability of at least these two variables. It is reasonable to consider the material strength fully correlated with the fracture energy (Grassl and Bažant, 2009). Furthermore, in the proposed lattice model, we also include the shear strength f_s and mode-II fracture energy G_s, which are again assumed to fully be correlated to the tensile strength f_t and mode-I fracture energy G_t, respectively. Assuming identical coefficient of variation (cov), we can use the same realizations of the random field to generate values of material strengths and fracture energies. For X substituted by any of the four mentioned mechanical properties, we can write

$$X(\boldsymbol{x}) = \bar{X}\boldsymbol{H}(\boldsymbol{x}) \tag{7}$$

where $\bar{X}$ stands for mean value of the particular property. The mean value of the (field) random variable H has to equal 1.

In this study, the following parameters of the Weibull-Gauss grafted distribution (Eq. 2a) were used: $m = 24$; $s_1 = 0.486$ MPa; $h_{gr} = 0.364$ MPa; $\delta_G = 0.25$. These values provide overall mean value μ_H=1; standard deviation δ_H ≈0.25 and grafting probability $F_H(h_{gr}) \approx 10^{-4}$. Two correlation lengths d were considered: a shorter length $d_4 = 40$ mm (according to (Grassl and Bažant, 2009)) and a longer length $d_8 = 80$ mm (according to (Vořechovský, 2007)). The computation is performed with $N = 24$ realizations of the random field for each correlation length.

4. Simulations of bending of notched beams

The first set of beams (depth $D = 300$ mm, span $S = 2.4D$, thickness $t = 40$ mm) loaded in three-point-bending were modeled with a central notch up to 1/6 of its depth. Ten deterministic simulations were computed. These simulations exhibit a certain scatter because of the pseudo-random granular positions differing for each realization. For both correlation lengths 40 and 80 mm, 24 simulations with spatial material randomness were performed. All the simulations were terminated when the magnitude of the loading force dropped to 1/3 of the maximal reached load $F_{\max}$. To ensure numerical stability in the presence of softening, the simulations were controlled by prescribing an increase of the crack mouth opening displacement (CMOD) in every step.

The notch present in the model induces a stress concentration at the notch tip. Therefore, high stresses occur only in a small area above the notch tip. Therefore, a crack initiates and propagates always from the notch tip. In stochastic calculations with rather large correlation length, local strength fluctuations within the region of high-stresses diminishes because of the imposed spatial correlation. Thus, the peak load $F_{\max}$ depends mostly on a single value of the random field realization at the notch tip location. In other words, a random field with correlation length greater than the length/width of FPZ can be, in the vicinity of the crack tip, viewed as a random constant – random field becomes a random variable at that region.

The obtained load-deflection curves are shown in Fig. 3. The figure also shows the maximal loads $F_{\max}$ in its upper left corner. The effect of the spatial strength fluctuations on the mean value of maximum load is negligible. The mean value of $F_{\max}$ is, for the deterministic calculation, $\mu_d = 11.3$ kN and, for stochastic simulations with $d = 40$ and 80 mm $\mu_4 = \mu_8 = 11.0$ kN. However, the standard deviations of the peak load are significantly influenced by the material randomness. The standard deviation of deterministic calculations (given solely by random aggregate position) is $\delta_d = 0.4$ kN. Significant increase in the standard deviation is observed for both correlation lengths: $\delta_4 = 1.5$ kN ($d = 40$ mm) and $\delta_8 = 1.8$ kN ($d = 80$ mm). Since the maximal load of the beam is given by local meso-level strength of a small area above the notch tip, we believe that the fluctuation rate does not influence the standard deviation (unless it is so small that material parameters vary significantly inside the FPZ).

For several selected realizations, the computed damage patterns (damage parameter ω from Eq. 1) at the peak load and at the termination of the simulations are showed in Fig. 4 together with the corresponding random field realization. Even though one can notice that the crack is slightly attracted (repelled) by areas of low (high) strength, the macrocrack trajectory is similar to the deterministic case (dictated by the singular stress field).

In order to compare energy dissipation in the beams, we need to determine simulation stages where the same portion of the ligament has already been damaged. Therefore, we select a stage when equivalent crack lengths (according to LEFM) are equal. Thus, all the models should have at that (reference) stage the same (reference) compliance, chosen as 1/45 mm/kN (Fig. 3). The depth of specimen was divided into horizontal stripes of depth s (Fig. 1c). All the energy dissipated at inter-particle contacts within a specific stripe was summed into variable G_d. One can normalize that energy by ligament area as $g_d = G_d/st$. The mean values and standard deviations of g_d are plotted in Fig. 5 for every stripe at the peak load and at the reference compliance stages. The figure confirms that the mean energy dissipation in notched tests does not change when the spatial material randomness is applied. Similarly to the peak force behavior, standard deviations of dissipated energy increase when randomness is present.

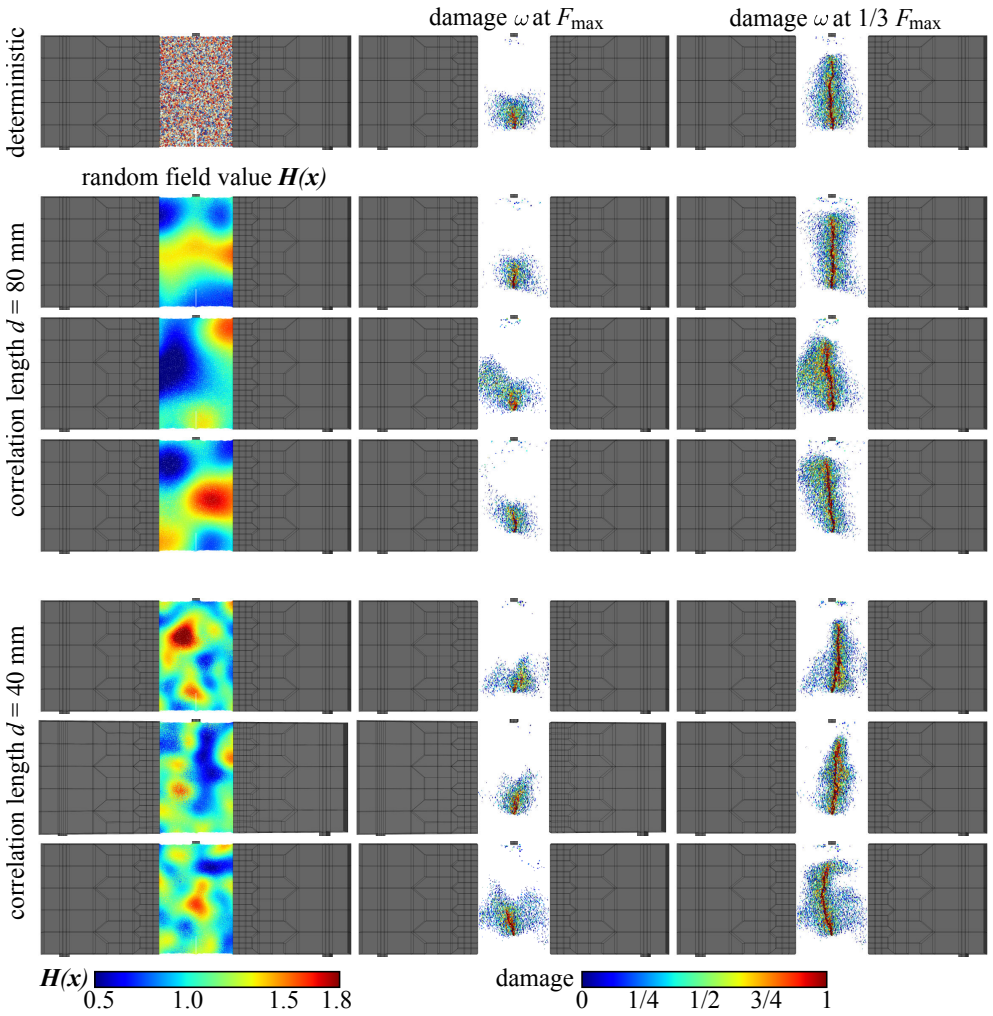


Figure 4. Realizations of random field $\boldsymbol{H}$ (left) and corresponding damage patterns developed in bent notched beams at the peak force (middle) and after the load dropped to $1/3$ of its maximum (right).

5. Simulations of bending of unnotched beams

The second simulation set focused on bending of unnotched beams where cracks initiate from a smooth bottom surface. Ten deterministic simulations and $N = 24$ simulations with random field for each correlation length were performed. To control the simulation, one needs to find some monotonically increasing variable, here again the CMOD was used. For unnotched beams with spatially fluctuating meso-level strength, the location of the macrocrack and thus the position of the crack mouth is not known in advance. Therefore, several short overlapping intervals were monitored simultaneously and the controlling CMOD was chosen to be the maximum one over them. Note, that other possibility of controlling variable might be the total energy dissipation in the specimen (Gutiérrez, 2004).

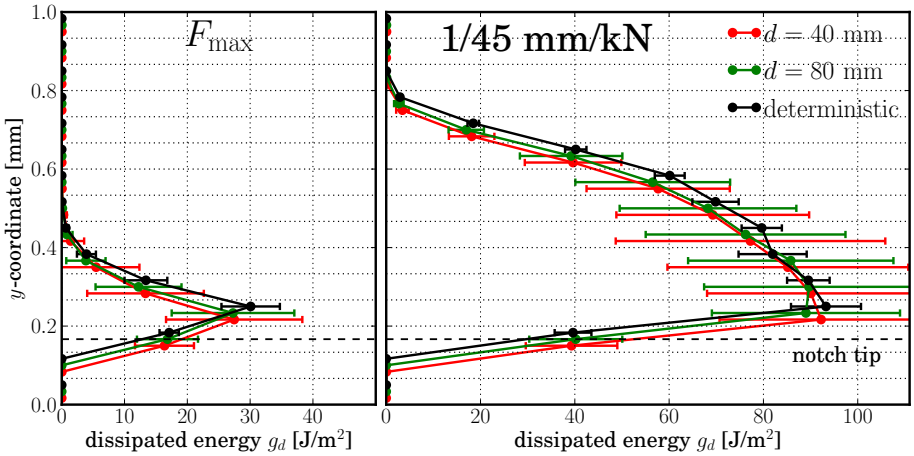


Figure 5. Energy per unit ligament area g_d dissipated in notched beams up to a) maximal load and b) reference beam compliance $1/45$ mm/kN in dependence on the vertical position in the beam.

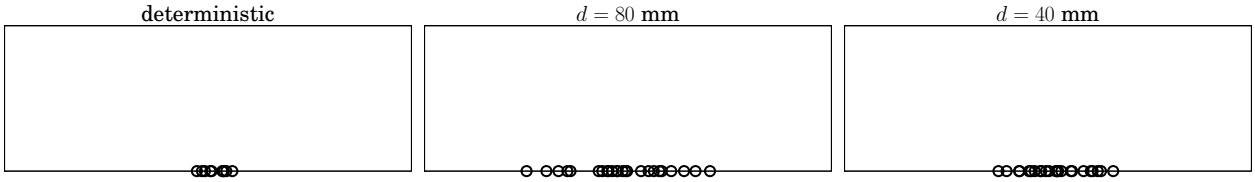


Figure 6. Points of crack initiation of unnotched beams for various degrees of randomness.

The variations in position of the crack mouth of the macrocrack are documented in Fig. 6. The Figure demonstrates the fundamental difference between notched and unnotched simulations. When no notch is present, the high-level stress region is much larger, located along the bottom central part of the specimen. Material strength and fracture energy fluctuate within the region and allow the macrocrack to "choose a weak spot" to initiate from. The higher is the distance form the midspan, the lower tensile stress appears. In the process of crack(s) formation, the stress field with a certain ability of redistribution increases towards the barrier (randomly varying strength and energy). The crack would start far from the midspan only when the material strength (and energy) of all points closer to the midspan is higher than in the surrounding. It is thus expectable (and confirmed by Fig. 6) that short correlation length, resulting in fluctuations that generate the weak spots more frequently, shrinks the zone where the macrocrack initiates. Indeed, the initiation zone for correlation length $d = 80$ mm is wider than for $d = 40$ mm.

Load deflection curves obtained from all the performed simulations are plotted in Fig. 7. The upper left corner shows the mean values and standard deviations of the peak load $F_{\max}$. The more fluctuating is the local strength, the weaker spot is statistically present and thus the lower is the mean value: $\mu_d = 22.4$ kN (deterministic), $\mu_8 = 17.0$ kN ($d = 80$ mm), $\mu_4 = 16.2$ kN ($d = 40$ mm). The standard deviation of the maximal force is low for the deterministic set, where $\delta_d = 0.6$ kN (cov$_d$=2.7%). For the correlation length 80 mm, it increases rapidly to $\delta_8 = 3.5$ kN (cov$_8$=21%). When the fluctuation rate increases more ($d = 40$ mm), the standard deviation of $F_{\max}$ decreases back to $\delta_4 = 2.1$ kN (cov$_4$=13%). This trend simply comes from the fact that the standard deviation of the local strength in the weakest spot inside some fixed

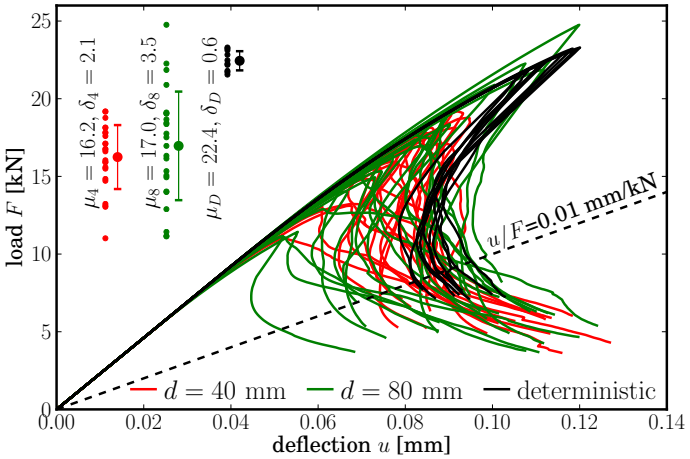


Figure 7. Load-deflection curves for simulations of TPB beams without notch.

region decreases with decreasing correlation length. Theoretically, the maximal standard deviation of $F_{\max}$ should be obtained for $d \approx \infty$ (a situation when the random field can be represented by a random variable – a random constant over the specimen volume).

Fig. 8 presents several selected realizations of the random field $\boldsymbol{H}$ and the computed damage patterns. One can see that the damage patterns differ for different levels of randomness. In the deterministic case, the damaged region at the peak load stage spans continuously the whole bottom area and the damage intensity directly depends on the distance from the midspan. For a random local strength and local fracture energy, the damage regions are more localized around low random field values. There is usually one such region for correlation length $d = 40$ mm and several low strength regions for $d = 80$ mm.

To compare the energy dissipation, we again choose some reference compliance that marks stages with the same LEFM crack length. The reference compliance now equals to 1/100 mm/kN (Fig. 7). Contrary to results from notched simulations, summation of total energy dissipated in stripes (per unit ligament area) is dependent on material randomness. In Fig. 9, deterministic calculations show higher values of dissipated energy g_d both for the peak force stage and for the stage at the reference compliance. This is caused by two factors: i) the localized macrocrack propagates in stochastic simulations through areas of lower meso-level strength and meso-level fracture energy, thus less energy is dissipated in total; ii) Distributed pre-peak cracking outside the macrocrack occurs mostly for deterministic simulation and thus it increases its total energy dissipation. Note that from about the middle of the specimens depth upwards, the energy dissipation of deterministic and stochastic simulations again match each other. This is because the crack at that depth cannot choose the weak region as it has already localized and the stress field forces the crack to grow from the current crack tip; and no pre-peak distributed cracking takes place there.

Finally, we focus on a deeper analysis of the energy dissipation along the bottom surface. In the bottom boundary stripe of width $2d_{\max} = 19$ mm, the dissipated energies (per unit ligament area) inside and outside the macrocrack were evaluated for stages at the peak load and at the reference compliance. These values are plotted in Fig. 10 separately for each simulation. The results document that distributed cracking outside macrocrack in the most bottom layer after the peak is reached is close to zero. The amount of energy dissipated *outside* a macrocrack is much higher for the deterministic simulations than for those with random

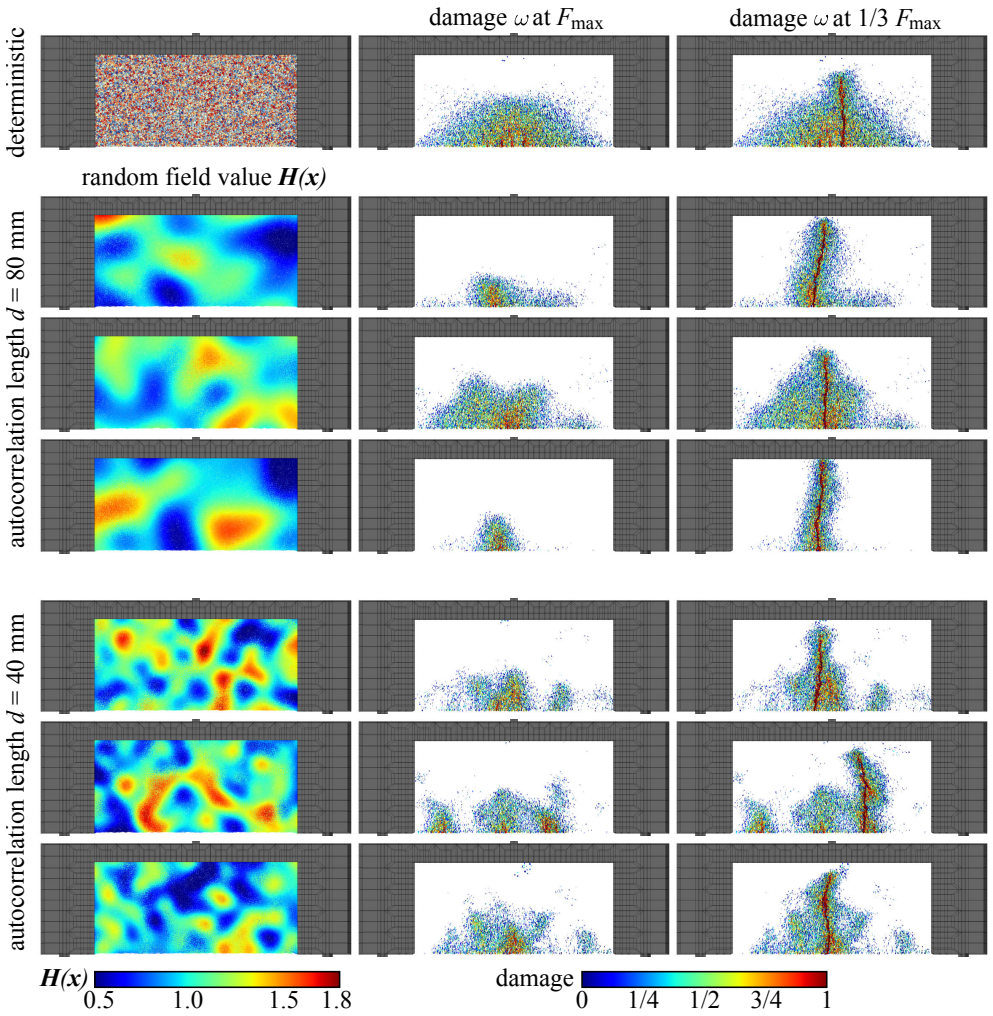


Figure 8. Realizations of random field $\boldsymbol{H}$ (left) and corresponding damage patterns developed in bent beams without notch at the peak force (middle) and after the load dropped to $1/3$ of its maximum (right).

fields. Some of the simulations for $d = 80$ mm reached the value of the deterministic model, which can be explained by an absence of a locally weak spot and subsequent extensive pre-peak distributed cracking (see Fig. 8, third row). The energy dissipated *inside* the macrocrack at the reference compliance is clearly higher in the deterministic case than in the stochastic one. This is due to the positive correlation of local meso-level energy and meso-level strength at the inter-particle bonds. Since the macrocrack propagates through locally weaker areas, it also dissipates less energy there. Aspects related to correlation between the local tensile strength and fracture energy have been discussed by (Vořechovský and Novák, 2004).

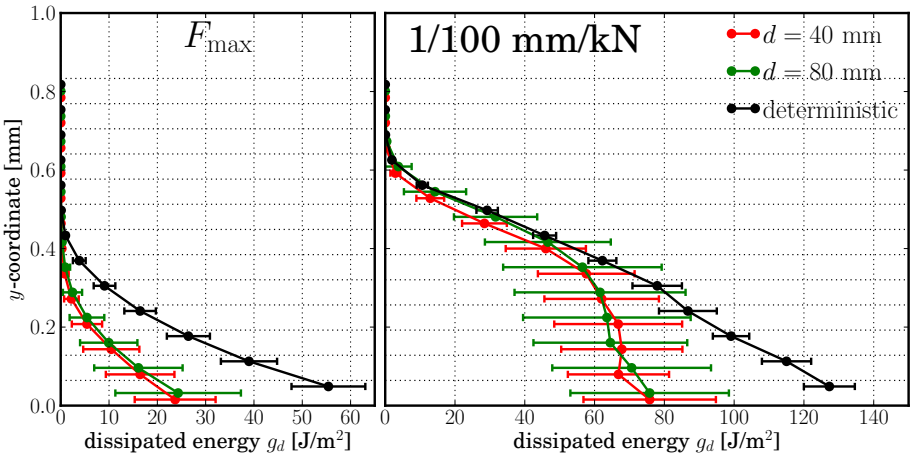


Figure 9. Energy per unit ligament area dissipated in unnotched beams up to a) maximal load and b) reference beam compliance $1/100$ mm/kN in dependence on vertical position in the beam.

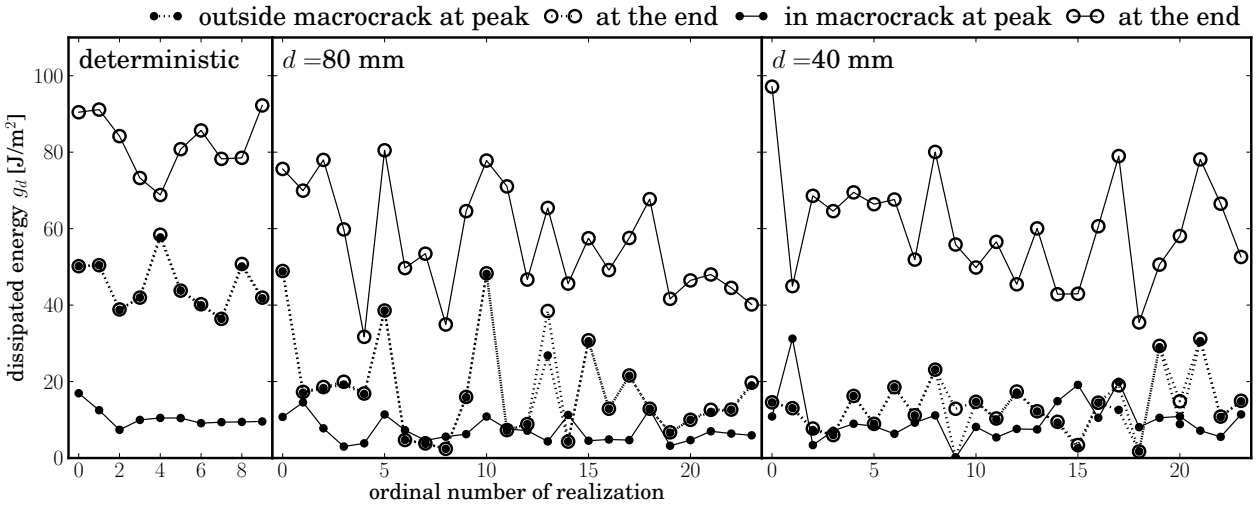


Figure 10. Energy dissipation inside and outside the macro-crak at the peak load and at the reference compliance stages for every simulation.

6. Conclusions

We analyzed the influence of material spatial randomness on the peak load and the energy dissipation using a discrete lattice-particle model that reflects the concrete meso-scopic structure, i.e. the aggregate composition. The spatial material randomness was introduced by simultaneous scaling of the local meso-level strength and fracture energy of inter-particle bonds by realizations of autocorrelated random field. Two basic cases of three-point-bent beams were investigated: i) beams with a notch and ii) beams without a notch (the modulus of rupture test). Numerical results generally confirm theoretical expectations.

It has been found that:

– for the simulation with a sufficiently deep notch, the crack is forced to start at the notch tip. Therefore, the mean value of the maximal load for notched beam simulations does not change when material spatial randomness applies. However, the standard deviation of the maximal load increases when strength randomness is introduced. Also, the energy dissipation in deterministic and random media exhibit the same mean but an increasing standard deviation for the random cases.

– In the case of unnotched beams, the macrocrack initiates in a locally weaker spot.When a shorter correlation length of material properties is applied, the weaker is statistically the initiation spot and therefore the mean of the maximal load is lower. Standard deviations of the maximal load increase when randomness applied, however the shorter correlation lengths lead to a decrease of the standard deviation.

– Energy dissipated in unnotched beams is dependent on the randomness of the material. Two effects responsible for the dependency were identified. i) Change of the dissipated energy due to correlation of the local meso-level fracture energy and low meso-level strength of inter-particle bonds through which the macrocrack propagates. Depending on the sign of the energy-strength cross-correlation, this effect may increase or decrease the dissipated energy. For the current settings of the model, the lower is the local meso-level strength, the lower is also the local fracture energy and the lower is the energy dissipated inside the macrocrack. ii) The pre-peak distributed cracking has a tendency to localize only in weaker areas and thus the material dissipated less energy outside the macrocrack when random field is applied.

Acknowledgments

The financial support received from the Czech Science Foundation under Project No. P105/11/P055 and the Ministry of Education, Youth and Sports of the Czech Republic under Project No. LH12062 is gratefully acknowledged.

References

Alava, M. J., Nukala, P. K. V. V., and Zapperi, S. Statistical models of fracture. *Advances in Physics*, 55(3-4):349–476, 2006.

Barenblatt, G. I. Mathematical theory of equilibrium cracks in brittle fracture. *Advances in Applied Mechanics*, 7:55–129, 1962.

Bažant, Z. P., and Jirásek, M. Nonlocal integral formulations of plasticity and damage: Survey of progress. *ASCE's Journal of Engineering Mechanics*, 128(11):1119–1149, 2002.

Bažant, Z. P., Le, J.-L., and Bazant, M. Z. Scaling of strength and lifetime distributions of quasibrittle structures based on atomistic fracture mechanics. *Proceeding of the National Academy of Sciences, USA*, 106(28):11484–11489, 2009.

Bažant, Z. P., and Novák, D. Probabilistic nonlocal theory for quasibrittle fracture initiation and size effect. I. Theory. *ASCE's Journal of Engineering Mechanics*, 126(2):166–174, 2000.

Bažant, Z. P., and Pang, S.-D. Activation energy based extreme value statistics and size effect in brittle and quasibrittle fracture. *Journal of the Mechanics and Physics of Solids*, 55:91–131, 2007.

Bažant, Z. P., and Planas, J. *Fracture and size effect in concrete and other quasibrittle materials.* CRC Press, 1998.

Bažant, Z. P., and Xi, Y. Statistical size effect in quasi-brittle structures: II. Nonlocal theory. *ASCE's Journal of Engineering Mechanics*, 117(7):2623–2640, 1991.

Bolander, J. E., and Saito, S. Fracture analyses using spring networks with random geometry. *Engineering Fracture Mechanics*, 61:569–591, 1998.

Cusatis, G., Bažant, Z., and Cedolin, L. Confinement-shear lattice model for concrete damage in tension and compression: I. theory. *ASCE's Journal of Engineering Mechanics*, 129:1439–1448, 2003.

Cusatis, G., Bažant, Z., and Cedolin, L. Confinement-shear lattice csl model for fracture propagation in concrete. *Computer Methods in Applied Mechanics and Engineering*, 195:7154–7171, 2006.

Cusatis, G., and Cedolin, L. Two-scale study of concrete fracturing behaviour. *Engineering Fracture Mechanics*, 6:3–17, 2007.

Eliáš, J., and Bažant, Z. Fracturing in concrete via lattice-particle model. In Onate, E., and Owen, D., editors, *2nd international conference on particle-based methods - fundamentals and applications*, pages 12, Barcelona, Spain. CD ROM, 2011.

Grassl, P., and Bažant, Z. P. Random lattice-particle simulation of statistical size effect in quasi-brittle structures failing at crack initiation. *ASCE's Journal of Engineering Mechanics*, 135:85–92, 2009.

Grassl, P., and Rempling, R. A damage-plasticity interface approach to the meso-scale modelling of concrete subjected to cyclic compressive loading. *Engineering Fracture Mechanics*, 75:4804–4818, 2008.

Gutiérrez, M. A. Energy release control for numerical simulations of failure in quasi-brittle solids. *Communications in Numerical Methods in Engineering*, 20(1):19–29, 2004.

Hillerborg, A., Modéer, M., and Petersson, P.-E. Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements. *Cement and Concrete Research*, 6:773–782, 1976.

HongShuang, L., ZhenZhou, L., and XiuKai, Y. Nataf transformation based point estimate method. *Chinese Science Bulletin*, 53(17):2586–2592, 2008.

Li, C.-C., and Kiureghian, A. D. Optimal discretization of random fields. *ASCE's Journal of Engineering Mechanics*, 119(6):1136–1154, 1993.

Van Mier, J. G. M., and Van Vliet, M. R. A. Influence of microstructure of concrete on size/scale effects in tensile fracture. *Engineering Fracture Mechanics*, 70:2281–2306, 2003.

Vořechovský, M., and Novák, D. Modeling statistical size effect in concrete by the extreme value theory. In Walraven, J., Blaauwendaad, J., Scarpas, T., and Snijder, B., editors, *5^{th} International Ph.D. Symposium in Civil Engineering, held in Delft, The Netherlands*, Vol. 2, pages 867–875, London, UK. A.A. Balkema Publishers. ISBN 90 5809 676 9, 2004.

Vořechovský, M. Interplay of size effects in concrete specimens under tension studied via computational stochastic fracture mechanics. *International Journal of Solids and Structures*, 44:2715–2731, 2007.

Vořechovský, M. Simulation of simply cross correlated random fields by series expansion methods. *Structural Safety*, 30:337–363, 2008.

Vořechovský, M. Incorporation of statistical length scale into weibull strength theory for composites. *Composite Structures*, 92(9):2027–2034, 2010.

Vořechovský, M., and Novák, D. Correlation control in small-sample monte carlo type simulations I: A simulated annealing approach. *Probabilistic Engineering Mechanics*, 24:452–462, 2009.

Vořechovský, M., and Sadílek, V. Computational modeling of size effects in concrete specimens under uniaxial tension. *International Journal of Fracture*, 154:27–49, 2008.

Weibull, W. The phenomenon of rupture in solids. In *Royal Swedish Institute of Engineering Research*, Vol. 153, pages 1–55, Stockholm, 1939.

Interval Solution for Inverse Problems under uncertainty

Francesco Fedele[1] and Rafi L. Muhanna[2]
[1,2]*School of Civil and Environmental Engineering, Georgia Institute of Technology,*
[1]*School of Electrical and Computer Engineering, Georgia Institute of Technology,*
Atlanta, GA 30332-0355, USA
fedele@gatech.edu, rafi.muhanna@gtsav.gatech.edu

Abstract: Inverse problems in science and engineering aim at estimating model parameters of a physical system using observations of the model's response. Variational least square type approaches are typically adopted, solving the forward model, and then comparing the resulting modeled data with the actual measured data. The data mismatch is minimized and the process is iterated until the best match is achieved. However, data measurements are associated with uncertainties, and deterministic inverse algorithms hardly provide the associated error estimates for the model parameters. In this work, an interval-based iterative solution is presented to predict bounds on such errors, using optimization and the containment-stopping criterion.

Keywords: Inverse, Interval, FEM

1. Introduction

Inverse problems in science and engineering aim at estimating model parameters of a physical system from available observations (data) of his response or output (see Tarantola 1987). A classical example is that of wave tomography in geophysics for a full seismic waveform inversion (see, for example, Fichtner 2010), or the optical tomography for the recognition of cancer in breast tissue via fluorescence (see, for example, Eppstein et al. 2003). In both cases, a forward model is given to predict the (seismic or light) wave propagation through a heterogeneous medium (soil subsurface or human tissue). The forward model is solved only if the (elastic or optical) properties of the medium are known in advance. These, however, are exactly what are not known and what one wants to predict. This leads to a formulation of an inverse problem if measurements of wave amplitudes and phases at given points on the accessible boundaries of medium are available. Using these data, an appropriate 'inverse' algorithm can be formulated to estimate maps of the properties of the medium, from which regions of high/low stiffness can be localized, or malign tissue detected. Variational least square type approaches are typically adopted by making an initial guess (either random or educated) for the unknown variables, solving the forward model, and then comparing the resulting modeled data with the actual measured data. The initial guess is then corrected by minimizing the data mismatch to yield a better match. The process is iterated until the best match is achieved.

A deep look into the mathematics used to model the wave propagation through a medium and the associated physical phenomena, such as scattering and absorption, will reveal an underlying mathematical structure, characterized by Helmholtz wave equations. These, as other partial differential equations encountered in engineering and sciences are typically solved by finite-element methods (FEM) on an unstructured mesh to adequately model the geometry of the medium, and to increase discretization density where appropriate. The inverse algorithm highly depends upon the forward model. If FEM is used the domain is discretized into elements and the number of unknowns depends on the mesh size and on the element type. Typically, the number of unknowns to be estimated exceeds the number of boundary

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

measurements available and that will result in an ill-posed problem. Ill-posedness is treated by regularization procedures (Tikhonov & Arsenin 1977), by adding appropriate additional constraints that yield well-posed inverse algorithms. Robustness is typically achieved by a course-to-fine regularization that exploits arc-length or surface-area minimizers.

Clearly, data measurements are affected by errors, whose nature depends upon both controllable and uncontrollable factors, such as, for example, the precision of the adopted instrumentation or the environmental conditions during the measurement campaign, respectively. Deterministic inverse algorithms hardly provide error bounds on the parameter estimates given uncertainties in the data. Indeed, this would require a combinatory approach that explores all the possible combinations of data within the given bounds, and this is computationally unfeasible even for small-to-medium scale problems. On the other hand, a probabilistic approach to solve the inverse problem, as in Kalman Filter estimation (Kalman 1960, see also Brown and Hwang 1992), allows identifying the propagation of uncertainties and it also provides errors on the parameter estimates. However, such approaches have their own limitations since they require a prior assumption on the nature of uncertainties, i.e. data errors are usually assumed as Gaussian. It is desirable to have inverse algorithms that do not rely on the type of uncertainties.

This work addresses this issue, by proposing an interval-based iterative solution for inverse problems that not only minimize the overestimation in the target quantities, but also exploits the same overestimation to track propagation of uncertainties of the target estimates. The paper is structured as follows. First, to illustrate the proposed theoretical approach, we present a one-dimensional (1-D) inverse problem that is estimating the Young's modulus of an elastic bar from known measurements of displacements due to traction/compression. The inverse algorithm is then introduced by combining an 'optimize-then-discretize' strategy with interval FEM in order to minimize the mismatch functional between modeled and actual data. Examples are finally presented and discussed.

2. Formulation of inverse problem in elastostatics

2.1. Deterministic Formulation

Consider an elastic bar of length L subject to distributed tensional forces $f(x)$. The differential equation

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\alpha\frac{\mathrm{d}u}{\mathrm{d}x}\right)+f=0, \qquad 0<x<L, \tag{1a}$$

with prescribed boundary conditions

$$u(0)=u_0, \qquad \alpha\frac{\mathrm{d}u}{\mathrm{d}x}=Q_0, \tag{1b}$$

define the 'continuous' forward model that allows to predict the displacements $u(x)$ given the parameter $\alpha(x)=E(x)A(x)$, where $E(x)$ is the Young's Modulus and $A(x)$ is the cross-sectional area, both assumed as spatially varying. When α is unknown, it can be estimated if measurements $\tilde{u}_j$ of u are available at N points $x=x_j$, j = 1, ...N on the bar surface. To solve for α we consider the following functional

$$F(u,\alpha,\tilde{u},w)=\frac{1}{2}\sum_{j=1}^{N}(u(x_j)-\tilde{u}_j)^2+\int_0^L w\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(\alpha\frac{\mathrm{d}u}{\mathrm{d}x}\right)+f\right]\mathrm{d}x+\gamma\int_0^L\left(\frac{\mathrm{d}\alpha}{\mathrm{d}x}\right)^2\mathrm{d}x, \tag{2}$$

here, the first term in the right-hand side is the square of the mismatch between data and the unknown theoretical displacements u (modelled data) at the locations x_j in accord to the forward model (1). The second term introduces the Lagrange multiplier $w(x)$ to enforce the 'strong' constraint (1), and the third integral is a standard course-to-fine regularization term to control the smoothness/roughness of α, and to guarantee the well-posedness of the inverse problem (γ is the regularization parameter).

To find the optimal α that minimizes (2), we introduce an imaginary time that rules the evolution/convergence of an initial guess for α toward the minimal solution of (2). Thus, u, w and α also depend upon the fictitious t, and we wish to find the rate $\dot{\alpha} = d\alpha/dt$ at which α should change in time so that F always decreases, i.e. $\dot{F} < 0$. The time derivative of F follows after several integrations by parts and some algebra as (see appendix A)

$$\begin{aligned}\dot{F} = &-\int_0^L \dot{\alpha}\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}w}{\mathrm{d}x}\mathrm{d}x - 2\gamma\int_0^L \dot{\alpha}\frac{\mathrm{d}^2\alpha}{\mathrm{d}x^2}\mathrm{d}x + w(0)\left(\dot{\alpha}\frac{\mathrm{d}u}{\mathrm{d}x} + \alpha\frac{\mathrm{d}\dot{u}}{\mathrm{d}x}\right)\Bigg|_{x=0} - \dot{u}(L)\left(\alpha\frac{\mathrm{d}w}{\mathrm{d}x}\right)\Bigg|_{x=L} + \\ &+\int_0^L \dot{u}\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(\alpha\frac{\mathrm{d}w}{\mathrm{d}x}\right) + \sum_{j=1}^{N}(u_j - \tilde{u}_j)\delta(x - x_j)\right]\mathrm{d}x,\end{aligned} \tag{3}$$

where we have set $u_j = u(x_j)$. Since the multiplier w is arbitrary, it can be properly chosen to further simplify (3). Indeed, if we impose the following boundary value problem

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\alpha\frac{\mathrm{d}w}{\mathrm{d}x}\right) + \sum_{j=1}^{N}(u_j - \tilde{u}_j)\delta(x - x_j) = 0, \qquad 0 < x < L, \tag{4a}$$

with boundary conditions as

$$w(0) = 0, \qquad \alpha\frac{\mathrm{d}w}{\mathrm{d}x}\Bigg|_{x=L} = 0, \tag{4b}$$

then (3) reduces to the minimal form

$$\dot{F} = -\int_0^L \dot{\alpha}\left(\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}w}{\mathrm{d}x} + 2\gamma\frac{\mathrm{d}^2\alpha}{\mathrm{d}x^2}\right)\mathrm{d}x \tag{5}$$

We are still free to choose the time rate of α so that F is always decreasing. To do so, $\dot{F} < 0$ is always satisfied at any t if we choose

$$\dot{\alpha} = \frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}w}{\mathrm{d}x} + 2\gamma\frac{\mathrm{d}^2\alpha}{\mathrm{d}x^2} \tag{6}$$

This yields the evolution equation of the unknown parameter α so that at steady state, i.e. $t \to \infty$, F is minimized. Observe that (6) depends upon the field u, which satisfies (1), and the associated adjoint or multiplier w, given by the boundary value problem (4). If we approximate the time derivative of α as

$$\dot{\alpha} \cong \frac{\alpha_{i+1} - \alpha_i}{\Delta t}$$

and follow the FEM-based 'discrete' version of the 'continuous' equations (1-4-6), a deterministic inverse algorithm can be formulated as

$$\begin{cases} K(\alpha_i)u_i = P \\ K(\alpha_i)w_i = (u_i - \tilde{u}) \\ \alpha_{i+1} = \alpha_i + Du_i \circ Dw_i \Delta t + 2\gamma \Delta t D_2 \alpha_i. \end{cases} \tag{7}$$

here, α_i is the ($m \times 1$) vector at iteration i that lists the individual values of α parameters within each of the m elements, u_i is a ($n \times 1$) vector of the nodal displacements, and $K(\alpha_i)$ is the assembled FEM stiffness matrix, which depends upon α_i. Further, P ($n \times 1$) is the vector of nodal forces, $\tilde{u}$ ($n \times 1$) is the data vector of measured displacements interpolated at the nodes, Du_i ($m \times 1$) and Dw_i ($m \times 1$) are the vector of element strains, and the vector of first derivative of w_i, respectively. $D_2\alpha_i$ ($m \times 1$) is an approximation of the second space derivative of $\alpha(x)$. The Hadamard product $a \circ b = (a_j b_j)$ is the element-by-element product. We point out that (7) can also be obtained via a 'discretize-then-optimize' strategy. To do so, one first discretizes the forward model (1) and then optimizes the 'discrete' version of the functional (2) with respect to the vector α.

The free parameter Δt can be chosen to control the smallness of the correction $\|\Delta\alpha\| = \|\alpha_{i+1} - \alpha_i\| << \|\alpha_i\|$ during the iterations, where $\|b\|$ is the norm of a vector b. Typically, one starts with an initial guess for α, say α_0, and iterates Eq. (7) until convergence is achieved, viz. when the relative error $\|\Delta\alpha\|/\|\alpha_i\|$ is smaller than a prescribed threshold ε.

In the following, we present an interval formulation of Eq. (7) that will provide bounds on the uncertainties of the estimates for α.

3. Interval FEM Formulation

One of the main features of interval arithmetic is its capability of providing guaranteed results. However, it has *the disadvantage of overestimation if variables have multiple occurrences in the same expression. For example, if x is an interval, the function f(x) = x – x is not equal to zero but to an interval that contains zero. Such dependencies lead to meaningless results, and have discouraged some researchers* of pursuing further developments of FEM techniques using interval representations.

Only recently, Interval Finite Element Methods (IFEM) have been developed to handle the analysis of systems for uncertain parameters described as intervals. Since the early development of IFEM during the mid-1990s of the last century (Koyluoglu et al., 1995; Muhanna and Mullen, 1995; Nakagiri and Yoshikawa, 1996; Rao and Sawyer, 1995; Rao and Berke, 1997; Rao and Chen 1998) researchers have focused, among other issues, on two major problems: the first is how to obtain solutions with reasonable bounds on the system response that make sense from a practical point of view, or in other words, with the least possible overestimation of their bounding intervals; the second is how to obtain reasonable bounds on the derived quantities that are functions of the system response.

The most successful approaches for overestimation reduction are those that relate the dependency of interval quantities to the physics of the problem being considered (for details see Muhanna and Mullen, 1995; Muhanna and Mullen, 2001; Zhang, 2005). A brief description of IFEM formulation is presented below, but a detailed explanation of the method can be found in Rama Rao et al., 2011. The two major issues resolved by this formulation are:

1. Reducing of overestimation in the bounds on the system response due to the coupling and transformation in the conventional FEM formulation as well as due to the nature of used interval linear solvers (Muhanna and Mullen, 2001).
2. Obtaining the secondary variables (derived) such as forces, stresses, and strains of the conventional displacement FEM along with the primary variables (displacements) and with the same accuracy of the primary ones.

3.1. DISCRETE STRUCTURAL MODELS

The FEM variational formulation for a static discrete structural model is given by minimizing the total potential energy functional

$$\Pi = \frac{1}{2} U^T K_c U - U^T P, \tag{8}$$

which yields

$$\frac{\partial \Pi}{\partial U} = K_c U - P = 0 \ ,$$

where Π, K_c, U, and P are total potential energy, stiffness matrix, displacement vector, and load vector respectively. For structural problems this formulation includes both direct and indirect approaches. For the direct approach, the strain ε is selected as a secondary variable of interest, where a constraint can be introduced as $C_2\, U = \varepsilon$. For the indirect approach, constraints are introduced on displacements of the form $C_1 U = V$ in such a way that Lagrange multipliers will be equal to the internal forces. C_1 and C_2 are matrices of orders $m \times n$ and $k \times n$, respectively, and m is the number of displacements' constraints, k is the number of strains, and n is the number of displacements' degrees of freedom. We note that V is a constant and ε is a function of U. We amend the right-hand side of Eq. (8) to obtain

$$\Pi^* = \frac{1}{2} U^T K_c U - U^T P + \lambda_1^T (C_1 U - V) + \lambda_2^T (C_2 U - \varepsilon), \tag{9}$$

where λ_1 and λ_2 are vectors of Lagrange multipliers with the dimensions m and k, respectively. Invoking the stationarity of Π^*, that is $\delta\Pi^* = 0$, we obtain

$$\begin{pmatrix} K_c & C_1^T & C_2^T & 0 \\ C_1 & 0 & 0 & 0 \\ C_2 & 0 & 0 & -I \\ 0 & 0 & -I & 0 \end{pmatrix} \begin{pmatrix} U \\ \lambda_1 \\ \lambda_2 \\ \varepsilon \end{pmatrix} = \begin{pmatrix} P \\ V \\ 0 \\ 0 \end{pmatrix} \tag{10}$$

The solution of Eq. (10) will provide the values of dependent variable U and the derived ones λ_1, λ_2, and ε at the same time and with the same accuracy. The present interval formulation is an extension of the Element-By-Element (EBE) finite element technique developed by Muhanna and Mullen (2001).

The main sources of overestimation in IFEM are the multiple occurrences of the same interval variable (*dependency problem*), the width of interval quantities, the problem size, and the problem complexity, in addition to the nature of the used interval solver of the interval linear system of equations.

The current formulation is modifying the displacements' constraints used in the previous EBE formulation to yield the element forces as Lagrange Multipliers directly and the system strains. *All interval*

quantities will be introduced in non-italic boldface font. Following the procedures given in Rama Rao et al. (2011) we obtain the interval linear system $\mathbf{KU} = \mathbf{P}$, or explicitly,

$$\begin{pmatrix} \mathbf{K}_c & C_1^T & B^T & 0 \\ C_1 & 0 & 0 & 0 \\ B & 0 & 0 & -I \\ 0 & 0 & -I & 0 \end{pmatrix} \begin{pmatrix} \mathbf{U}_c \\ \boldsymbol{\lambda}_1 \\ \boldsymbol{\lambda}_2 \\ \boldsymbol{\varepsilon} \end{pmatrix} = \begin{pmatrix} \mathbf{P}_c \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{11}$$

here, $\mathbf{K}_c$ is a ($k \times k$) interval matrix, which contains the individual elements' local stiffness and zeros corresponding to the free nodes' degrees of freedom, where k is the sum of number of elements and free nodes.

The accuracy of the system solution depends mainly on the structure of Eq. (11) and on the nature of the used solver. The associated solution provides the enclosures of the values of dependent variables which are the interval displacements $\mathbf{U}$, interval element forces $\boldsymbol{\lambda}_1$, the multiplier $\boldsymbol{\lambda}_2$, and the elements' interval strains. An iterative solver is discussed in the next section.

3.2. INTERVAL SOLVERS AND ITERATIVE ENCLOSURE

Any solver for interval linear system of equations can be used to solve for u_i and w_i in Eq. (7), however, the best known method for obtaining very sharp enclosures of interval linear system of equations that have the structure introduced in Eq. (11) and with large uncertainty is the iterative method developed in the work of Neumaier and Pownuk (2007). The current formulation results in the interval linear system of equations given in (11) which can be transformed to have the general form:

$$(K + B\,\mathbf{D}\,A)\mathbf{u} = a + F\,\mathbf{b} \tag{12}$$

where $\mathbf{D}$ is diagonal. Furthermore, defining

$$C := (K + BD_0A)^{-1} \tag{13}$$

where D_0 is chosen to ensure invertability (often D_0 is selected as the midpoint of $\mathbf{D}$), the solution $\mathbf{u}$ can be written as:

$$\mathbf{u} = (Ca) + (CF)\mathbf{b} + (CB)\mathbf{d} \tag{14}$$

To obtain a solution with tight interval enclosure we define two auxiliary interval quantities,

$$\begin{aligned} \mathbf{v} &= A\mathbf{u} \\ \mathbf{d} &= (D_0 - \mathbf{D})\mathbf{v} \end{aligned} \tag{15}$$

which, given an initial estimate for $\mathbf{u}$, we iterate as follows:

$$\mathbf{v}^{k+1} = \{ACa) + (ACF)\mathbf{b} + (ACB)\mathbf{d}^k\} \cap \mathbf{v}^k, \quad \mathbf{d}^{k+1} = \{(D_{c0} - \mathbf{D}_c)\mathbf{v}^{k+1} \cap \mathbf{d}^k \tag{16}$$

until the enclosures converge, from which the desired solution $\mathbf{u}$ can be straightforwardly obtained.

Observe that not only are the interval displacements $\mathbf{U}$ obtained but also the derived quantities $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$, and $\boldsymbol{\varepsilon}$ with the same accuracy. The next section will discuss the use of this formulation in the solution of the inverse problem Eq. (2) under interval uncertainties.

4. Interval Inverse Problem

The interval solution of the inverse problem of Eq. (2) is based on the notion that the measurements (data) are given as intervals. In this work we are introducing an initial attempt to provide such a solution using interval finite element. From a closer look at the deterministic solution presented in Eq. (7), it can be seen that the iterative update of the sought parameters is given by: $\alpha_{i+1} = \alpha_i + Du_i \circ Dw_i \Delta t + 2\gamma \Delta t D_2 \alpha_i$, where the terms Du_i and Dw_i are the first derivative of u_i and w_i respectively. A naive interval FEM formulation will result in an enormous overestimation of the solution enclosure and with additional excessive overestimations in derived quantities such as stresses and strains. In our case, the solution is the displacement and the derived quantity is the strain. The IFEM formulation described in the previous section provides an exact solution for the interval loads and the tightest possible enclosure when both load and stiffness being intervals. Moreover, the formulation provides the stresses ($\boldsymbol{\lambda}_1$) and strains ($\boldsymbol{\varepsilon}$) as part of Eq. (11) solution and of course with the same accuracy as that of the displacements. Furthermore, we speculate that the course-to-fine regularizer can be neglected, viz. set $\gamma_1 = 0$, because one can exploit the natural relaxation induced by intervals, which allows to seek for a 'thick vector' $\boldsymbol{\alpha} = \mathbf{E}A$, a vector with thickened (relaxed) values that can span the range naturally imposed by the uncertainty of the data. This is similar to the Tikhonov regularization that imposes vector solutions with small norm.

In summary, the solution of the inverse problem of Eq. (2) as interval is accomplished by implementing the following steps:

1. Solve for $\mathbf{u}_i$ and $D\mathbf{u}_i$ using Eq. (11) as $\mathbf{U}_i = \mathbf{K}^{-1}\mathbf{P}_i$, where the interval vector $\mathbf{U}_i$ contains $\mathbf{u}_i$ and $D\mathbf{u}_i$.
2. Solve for $D\mathbf{w}_i$ using Eq. (11) in the form $\mathbf{W}_i = \mathbf{K}_i^{-1}(\mathbf{u}_i - \tilde{\mathbf{u}})$, where the interval vector $\mathbf{W}_i$ contains $\mathbf{w}_i$ and $D\mathbf{w}_i$. Instead of computing $(\mathbf{u}_i - \tilde{\mathbf{u}})$ as a conventional interval operation, the subtraction is done on bounds due to inherited dependency of $\mathbf{u}_i$ upon $\tilde{\mathbf{u}}$, since $\mathbf{u}_i \to \tilde{\mathbf{u}}$ when convergence is attained (see Eq. 7). In particular, $\mathbf{u}_i - \tilde{\mathbf{u}} = [\underline{\mathbf{u}}_i - \underline{\tilde{\mathbf{u}}}, \overline{\mathbf{u}}_i - \overline{\tilde{\mathbf{u}}}]$, where $\underline{\mathbf{u}}$ and $\overline{\mathbf{u}}$ are the lower and upper bounds of $\mathbf{u}$, respectively.
3. Compute the updated interval value of $\boldsymbol{\alpha}_{i+1}$ as
$$\boldsymbol{\alpha}_{i+1} = \boldsymbol{\alpha}_i + D\mathbf{u}_i \circ D\mathbf{w}_i \Delta t \tag{17}$$
with $\Delta t = c \times \min(\underline{\boldsymbol{\alpha}}_i / \overline{D\mathbf{u}_i \circ D\mathbf{w}_i})$, and $\gamma = 0$. The optimal choice of the constant c is problem dependent, and in our case we set c = 0.005.
4. The iterations are stopped when the estimated displacements $\mathbf{u}_i$ contain the data $\tilde{\mathbf{u}}$ (containment-stopping criterion), or in other words when $\mathbf{u}_i \supseteq \tilde{\mathbf{u}}$.

4.1. Example

For an illustrative example, we are using a 5 m long bar, pinned at one end and simply supported at the other as shown in Fig. (1). The bar has a constant cross sectional area A = 0.005 m^2 and is subjected to an axial force of 1000 kN applied at C. The bar is modelled using 25 finite elements each has a different modulus of elasticity. The values: 100, 105,110,115,120, 120, 115, 110, 105, 100, 105,110, 115, 120, 130, 140, 150, 140, 130, 125, 120, 115, 105, 100, and 90 GPa are the assumed moduli of elasticity of elements 1 through 25, respectively.

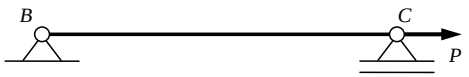


Figure 1. Truss bar.

The problem to be solved is that to predict the values of elasticity moduli for each element given that the displacements at the 26 nodes are known intervals (measurements with interval uncertainty). As an initial guess for the Young's modulus we set $E(x) = 60$ GPa.

4.2. DETERMINISTIC SOLUTION

First, the algorithm in Eq. (7) has been tested for the case where the measured data are assumed to be deterministic. The solution converged to the measured data and the moduli were predicted correctly (results are not reported). Hereafter, we will apply the interval-based formulation of the algorithm.

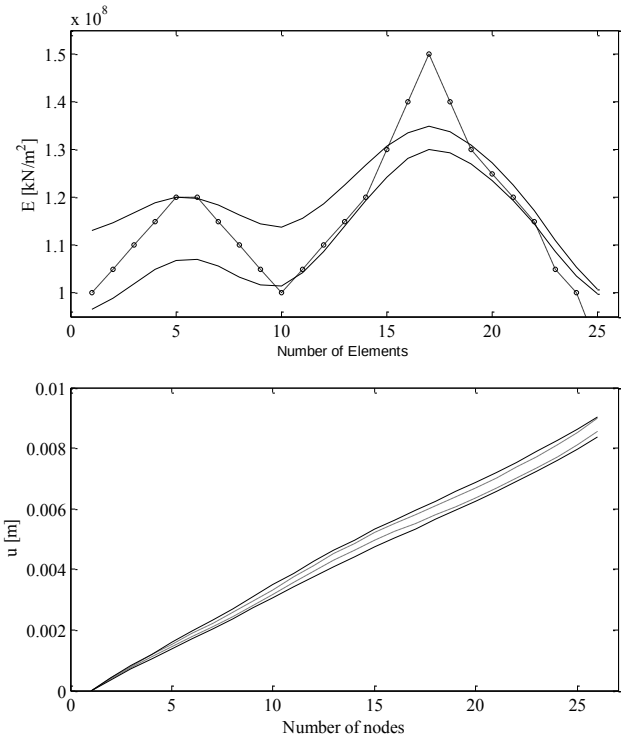


Figure 2. Premature solution of the Interval Inverse Problem. (top) Exact Young's modulus E (dash) and upper and lower bounds (solid) of the interval estimate $\mathbf{E} = \boldsymbol{\alpha}/A$, where A is the cross-sectional area; (bottom) Given interval data $\tilde{\mathbf{u}}$ (dash) and associated $\mathbf{u}$ displacements. Note that estimates contain data.

4.3. SOLUTION FOR UNCERTAIN MEASUREMENTS

For the uncertain case, a 5% interval uncertainty is considered in the measurements. Fig. 2 shows the obtained interval solution and the associated containment of the measurements, i.e. the estimated $\mathbf{u}$

displacements contain the measured $\tilde{\mathbf{u}}$ data. However, a pre-mature prediction of the elasticity moduli occurred. This phenomenon is due to the overestimation in the solution (the measurements are contained before the final solution is attained). Work is in progress to improve overestimation reduction of Eq. (17) by simultaneously solving for **u** and **w** (see Eq. 7) in an interval block-matrix form similar to that of Eq. (11).

A simpler alternative strategy has been adopted to avoid a significant overestimation and to obtain the correct solution. We first proceed with the solution in a deterministic form until the $\Delta\boldsymbol{\alpha} = \boldsymbol{\alpha}_{i+1} - \boldsymbol{\alpha}_i$ update becomes insignificant after several iterations (usually of the order of hundreds). At this stage the update for $\boldsymbol{\alpha}$ is switched to a full interval form using the interval algorithm based on Eq. (17). Fig. 3 shows the resulting mature solution, where both the measurements and the estimated unknown Young's moduli are contained.

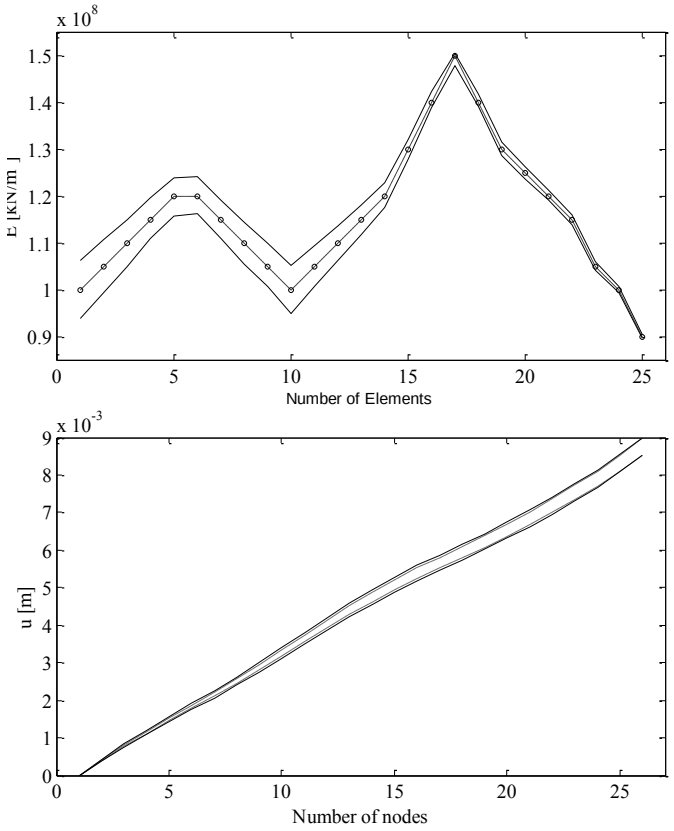


Figure. 3. Mature solution of the Interval Inverse Problem. (top) Exact Young's modulus *E* (dash) and upper and lower bounds (solid) of the interval estimate **E** = **α**/*A*, where *A* is the cross-sectional area; (bottom) Given interval data $\tilde{\mathbf{u}}$ (dash) and associated **u** displacements. Note that estimates contain data.

5. Conclusion

An initial formulation for interval inverse problems is introduced. Uncertainty in the measurements is considered in an interval form. The containment stopping criterion is used which is intrinsic for interval arithmetic. Overestimation control and reduction play crucial role in achieving correct solutions. Results show a great potential for further developments.

References

Brown, R. G. and P. Y. C. Hwang. (1992). *Introduction to Random Signals and Applied Kalman Filtering*, Second Edition, John Wiley & Sons, Inc.

Eppstein M. , Fedele F., Laible J. P., Zhang C. , Godavarty A. & Sevick-Muraca E. M. (2003). A comparison of exact and approximate adjoint sensitivities in fluorescence tomography. IEEE Transactions on Medical Imaging, vol. **22**, No. 10, pp. 1215-1222.

Fichtner A. (2010). *Full seismic waveform modelling and inversion*. Spinger-Verlag, Heidelberg.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems, *Transaction of the ASME—Journal of Basic Engineering*, pp. 35-45.

Koyluoglu, U., Cakmak, S., Ahmet, N. and Soren, R.K. (1995). Interval algebra to deal with pattern loading and structural uncertainty, *Journal of Engineering Mechanics*, Vol. **121**, No. 11, pp.1149–1157.

Muhanna, R.L. and Mullen, R.L. (1995). Development of interval based methods for fuzziness in continuum mechanics', IEEE Proceedings of ISUMA - NAFIPS '95, 17-19, pp. 145-150.

Muhanna, R.L. and Mullen, R.L. (2001). Uncertainty in mechanics problems—interval-based approach, *Journal of Engineering Mechanics*, Vol. **127**, No. 6, pp.557–566.

Nakagiri, S. and Yoshikawa, N. (1996). Finite element interval estimation by convex model, Proceedings of 7th ASCE EMD/STD *Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability*, WPI, Worcester, MA, USA, 7–9 August, pp.278–281.

Neumaier, A. and Pownuk, A. (2007). Linear systems with large uncertainties, with applications to truss structures, *Reliable Computing*, Vol. **13**, No. 2, pp.149–172.

Rao, S.S. and Berke, L. (1997). Analysis of uncertain structural systems using interval analysis, *American Institute of Aeronautics and Astronautics Journal*, Vol. **35**, No. 4, pp.727–735.

Rao, S. S. and Chen, L. (1998). Numerical solution of fuzzy linear equations in engineering analysis. International *Journal of Numerical Methods in Engineering*, Vol. **43**, pp.391–408.

Rao, S.S. and Sawyer, P. (1995). Fuzzy finite element approach for analysis of imprecisely defined systems, *American Institute of Aeronautics and Astronautics Journal*, Vol. **33**, No. 12, pp.2364–2370.

Rama Rao, M.v., Mullen, R.L. and Muhanna R. L. (2011). A new interval finite element formulation with the same accuracy in primary and derived variables, *Int. J. Reliability and Safety*, Vol. **5**, Nos. ¾.

Tarantola, A. (1987). *Inverse problem theory: Methods for data fitting and model parameter estimation*. Elsevier, ISBN 0444427651.

Tikhonov, A. & Arsenin, V. (1977). *Solution of ill-posed problems*. Wiley, New York.

Zhang, H. , (2005). Nondeterministic Linear Static Finite Element Analysis: An Interval Approach, PhD Dissertation, *School of Civil and Environmental Engineering, Georgia Institute of Technology*, Atlanta, GA.

Appendix A

In Eq. (2), we set $u_j = u(x_j)$ and integrate by parts once to obtain

$$F(u,\alpha,\tilde{u},w) = \frac{1}{2}\sum_{j=1}^{N}(u_j-\tilde{u}_j)^2 - \int_0^L \alpha \frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}w}{\mathrm{d}x}\mathrm{d}x + w\alpha\frac{\mathrm{d}u}{\mathrm{d}x}\Big|_0^L + \int_0^L wf\,\mathrm{d}x + \gamma\int_0^L\left(\frac{\mathrm{d}\alpha}{\mathrm{d}x}\right)^2\mathrm{d}x \tag{A1}$$

Since *u*, *w* and α are assumed time dependent, the time derivative of *F* follows from (A1) as

$$\dot{F} = \sum_{j=1}^{N}(u_j-\tilde{u}_j)\dot{u}_j + \dot{w}\alpha\frac{\mathrm{d}u}{\mathrm{d}x}\Big|_0^L + \left(w\dot{\alpha}\frac{\mathrm{d}u}{\mathrm{d}x} + w\alpha\frac{\mathrm{d}\dot{u}}{\mathrm{d}x}\right)\Big|_0^L - \int_0^L \dot{\alpha}\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}w}{\mathrm{d}x}\mathrm{d}x - \underbrace{\int_0^L \alpha\frac{\mathrm{d}\dot{u}}{\mathrm{d}x}\frac{\mathrm{d}w}{\mathrm{d}x}\mathrm{d}x}_{A} - \underbrace{\int_0^L \alpha\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}\dot{w}}{\mathrm{d}x}\mathrm{d}x}_{B} + \tag{A2}$$

$$\int_0^L \dot{w}f\,\mathrm{d}x + \underbrace{2\gamma\int_0^L \frac{\mathrm{d}\alpha}{\mathrm{d}x}\frac{\mathrm{d}\dot{\alpha}}{\mathrm{d}x}\mathrm{d}x}_{C}$$

Here, applying integration by parts once to terms *A*, *B* and *C* yield (for simplicity, we set $\gamma = 0$ at the boundaries)

$$\dot{F} = \sum_{j=1}^{N}(u_j-\tilde{u}_j)\dot{u}_j + w\left(\dot{\alpha}\frac{\mathrm{d}u}{\mathrm{d}x} + \alpha\frac{\mathrm{d}\dot{u}}{\mathrm{d}x}\right)\Big|_0^L - \int_0^L \dot{\alpha}\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}w}{\mathrm{d}x}\mathrm{d}x - \dot{u}\alpha\frac{\mathrm{d}w}{\mathrm{d}x}\Big|_0^L - 2\gamma\int_0^L \dot{\alpha}\frac{\mathrm{d}^2\alpha}{\mathrm{d}x^2}\mathrm{d}x \tag{A3}$$

$$\int_0^L \dot{u}\frac{\mathrm{d}}{\mathrm{d}x}\left(\alpha\frac{\mathrm{d}w}{\mathrm{d}x}\right)\mathrm{d}x + \underline{\int_0^L \dot{w}\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(\alpha\frac{\mathrm{d}u}{\mathrm{d}x}\right) + f\right]\mathrm{d}x}$$

Note that the underlined term vanishes because of (1*a*). Further, taking the time derivative of the boundary conditions (1*b*) for *u* yields

$$\dot{u}\big|_{x=0} = 0, \qquad \left(\dot{\alpha}\frac{\mathrm{d}u}{\mathrm{d}x} + \alpha\frac{\mathrm{d}\dot{u}}{\mathrm{d}x}\right)\Big|_{x=L} = 0 \tag{A4}$$

and (A3) simplifies to

$$\dot{F} = \sum_{j=1}^{N}(u_j-\tilde{u}_j)\dot{u}_j + w(0)\left(\dot{\alpha}\frac{\mathrm{d}u}{\mathrm{d}x} + \alpha\frac{\mathrm{d}\dot{u}}{\mathrm{d}x}\right)\Big|_{x=0} - \int_0^L \dot{\alpha}\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}w}{\mathrm{d}x}\mathrm{d}x - \dot{u}(L)\left(\alpha\frac{\mathrm{d}w}{\mathrm{d}x}\right)\Big|_{x=L} + \int_0^L \dot{u}\frac{\mathrm{d}}{\mathrm{d}x}\left(\alpha\frac{\mathrm{d}w}{\mathrm{d}x}\right)\mathrm{d}x \tag{A5}$$

$$-2\gamma\int_0^L \dot{\alpha}\frac{\mathrm{d}^2\alpha}{\mathrm{d}x^2}\,dx.$$

Eq. (3) follows from (A5) after re-writing the mismatch term as

$$\sum_{j=1}^{N}(u_j-\tilde{u}_j)\dot{u}_j = \sum_{j=1}^{N}\int_0^L (u_j-\tilde{u}_j)\dot{u}(x)\delta(x-x_j)\mathrm{d}x \tag{A6}$$

where $\delta(x-x_j)$ is the Dirac function centered at x_j.

A Particle Swarm Optimization Approach for Training Artificial Neural Networks with Uncertain Data

Steffen Freitag[1,2], Rafi L. Muhanna[1] and Wolfgang Graf [2]
[1]*School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA, steffen.freitag@gtsav.gatech.edu, rafi.muhanna@gtsav.gatech.edu*
[2]*Institute for Structural Analysis, Technische Universität Dresden, 01062 Dresden, Germany, steffen.freitag@tu-dresden.de, wolfgang.graf@tu-dresden.de*

Abstract. Artificial neural networks are powerful tools to learn functional relationships between data. They are widely used in engineering applications. Recurrent neural networks for fuzzy data have been introduced to map uncertain structural processes with deterministic or uncertain network parameters. Based on swarm intelligence, a new training strategy for neural networks is presented in this paper. Accounting for uncertainty in measurements, particle swarm optimization (PSO) approaches using interval and fuzzy numbers are developed. Applications are focused on the description of time-dependent material behavior with recurrent neural networks for uncertain data within interval and fuzzy finite element analyses. Network training with PSO allows to create special network structures with dependent parameters in order to consider physical boundary conditions of investigated materials.

Keywords: particle swarm optimization; neural network; uncertainty; interval numbers; fuzzy numbers, constitutive material description, finite element method

1. Introduction

Reliability assessment of structures requires knowledge of its behavior under environmental influences. Information on the structural behavior may be obtained by structural monitoring. Existing structures can be investigated by in situ monitoring whereas material tests can be performed to investigate new materials. As a result of experimental investigations, data series for measured structural actions and responses are available. Measured results are more or less characterized by data uncertainty due to varying boundary conditions, inaccuracies in measurements, and / or incomplete sets of observations. Interval or fuzzy numbers can be used to represent imprecise parameters, see e.g. (Möller and Beer, 2008). Time-dependent structural parameters are quantified as interval or fuzzy processes.

Functional relationships between uncertain data are required to describe the observed physical phenomena. Commonly, constitutive models are used for stress-strain relationships. Their parameters must be identified by an inverse analysis. If no closed-form expression can be obtained, optimization approaches can be applied to determine the unknown parameters of a predefined model.

An alternative approach to get functional relationships between uncertain data is the application of artificial intelligence. Artificial neural networks are widely used in engineering. Fields of applications in civil engineering are presented, e.g. in (Adeli, 2001). Often, multilayer perceptrons with feed forward architecture are utilized to learn functional relationships in deterministic data. For this purpose, several training strategies

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

are available (Haykin, 1999). In (Graf et al., 2011), neural network approaches for structural analysis with uncertain data are discussed. For time-dependent phenomena, recurrent neural networks can be applied. These advanced network architectures enable the consideration of the whole history for the computation of current states, see e.g. (Oeser and Freitag, 2009). Recurrent neural networks for fuzzy data (Freitag et al., 2011a) have been developed to identify deterministic dependencies (Graf et al., 2010) or uncertain dependencies (Freitag et al., 2011c) in fuzzy processes. In (Freitag et al., 2010a), a backpropagation training algorithm for recurrent neural networks with trainable fuzzy network parameters has been introduced. It is a gradient based approach using the derivatives of fuzzy activation functions. In general, interval arithmetic (Moore, 1979) or α-level optimization (Möller et al., 2000) can be used to compute the signals of recurrent neural networks for fuzzy data, see (Freitag, 2010). In this paper, a new training strategy for recurrent neural networks is introduced considering both ways of computation. It is based on swarm intelligence (Kennedy et al., 2001).

Particle swarm optimization (PSO) (Kennedy and Eberhart, 1995) is an optimization concept motivated by social behavior of group individuals. It is a random search strategy and requires multiple evaluations of an objective function. After random initialization, each individual (denoted as particle) share its information with other particles in the swarm in order to define its new position in the space of search variables. In (Eberhart and Shi, 2001), developments and applications of PSO are discussed. Applications in civil engineering are presented e.g. in (Perez and Behdinan, 2007) and (Li et al., 2007). The approaches in these works can be used for optimization tasks with constraints.

One of the first applications of PSO was the training of artificial neural networks, see e.g. (Kennedy and Eberhart, 1995). Algorithms for feed forward neural networks are presented e.g. in (Mendes et al., 2002) and (Kuok et al., 2010). A hybrid training strategy, combining backpropagation and PSO for training of feed forward neural networks, is shown in (Zhang et al., 2007). Accounting for uncertain training data, PSO approaches using interval and fuzzy numbers are developed in this paper. They can be applied to feed forward and recurrent neural networks. The advantage of PSO for recurrent neural networks is that all network parameters can be modified during training. Additionally, special network structures with dependent parameters can be created. This is helpful to consider physical boundary conditions, if neural networks are used as constitutive models.

Recurrent neural networks for interval or fuzzy data are used to describe uncertain stress-strain-time dependencies. A finite element formulation for neural network based material descriptions is shown. Neural networks can be applied as constitutive models within interval finite element methods (Muhanna et al., 2007), (Rao et al., 2011), fuzzy finite element methods (Möller et al., 2000), (Moens and Vandepitte, 2005) or fuzzy stochastic finite element methods (Graf et al., 2011), (Sickert et al., 2011). Examples are presented to show the applicability of the new approach.

2. Uncertain data

2.1. Intervals and fuzzy numbers

Uncertain data can be represented as intervals or fuzzy numbers. An interval

$$\bar{x} = [{}_l x,\ {}_r x] \tag{1}$$

is defined by its left ${}_l x$ and right ${}_r x$ bounds. A bar $\bar{}$ is used to indicate intervals. It is also common to define an interval by its midpoint

$$ {}_m x = \frac{{}_l x + {}_r x}{2} \tag{2} $$

and its width

$$ {}_w x = {}_r x - {}_l x \ . \tag{3} $$

If midpoints and widths are used, the left and right interval bounds are obtained by

$$ {}_l x = {}_m x - \frac{{}_w x}{2} \tag{4} $$

and

$$ {}_r x = {}_m x + \frac{{}_w x}{2} \ , \tag{5} $$

respectively.

Fuzzy numbers $\tilde{x}$ are uncertain sets gradually assessed by membership functions $\mu(x)$. The tilde $\tilde{}$ is used to indicate fuzziness. The functional values of $\mu(x)$ are defined in $[0,\ 1]$. For each realisation x, its level of membership to the set $\tilde{x}$ is between 0 and 1. Considering convex fuzzy numbers, an interval

$$ {}_s\bar{x} = [{}_{sl} x,\ {}_{sr} x] \tag{6} $$

is obtained for each level s of membership $\alpha_s = \mu\left({}_{sl} x\right) = \mu\left({}_{sr} x\right)$. A set of $s = 1, \ldots, S$ cuts (α-cuts) can be used to approximate the membership function of a fuzzy number $\tilde{x}$ by piecewise linear functions, see Figure 1. The interval bounds of each α-cut are given by

$$ {}_{sl} x = \min \left[x \in \mathbb{R} \mid \mu(x) \geq \alpha_s \right] \tag{7} $$

and

$$ {}_{sr} x = \max \left[x \in \mathbb{R} \mid \mu(x) \geq \alpha_s \right] \ , \tag{8} $$

respectively. A fuzzy number can be represented by its α-cuts as a discrete set of the corresponding interval bounds. The fuzzy number

$$ \tilde{x} = \langle {}_{1l} x, \ldots, {}_{Sl} x, {}_{Sr} x, \ldots, {}_{1r} x \rangle \tag{9} $$

contains all left and right interval bounds as a sorted sequence. In general, the α-cut S (with $\mu(x) = 1$) can be an interval or a deterministic number. If it is a deterministic number, i.e. ${}_{Sl} x = {}_{Sr} x = {}_S x$, the number of elements in Eq. (9) is an odd number. At least three elements are required to define a fuzzy number. In this case, the fuzzy number $\tilde{x} = \langle {}_{1l} x,\ {}_2 x,\ {}_{1r} x \rangle$ has a membership function with triangular shape. With four elements $\tilde{x} = \langle {}_{1l} x,\ {}_{2l} x,\ {}_{2r} x,\ {}_{1r} x \rangle$, a membership function with trapezoidal shape is created. The α-cut representation of fuzzy numbers is common in engineering. It allows to handle fuzzy numbers similar to intervals in numerical simulations, i.e. interval operations can be performed for each α-cut.

2.2. Interval and fuzzy processes

Interval and fuzzy processes can be represented by series of interval or fuzzy numbers. The interval process

$$ \bar{x}(\tau) = \left\{ {}^{[1]}\bar{x}, \ldots, {}^{[n]}\bar{x}, \ldots, {}^{[N]}\bar{x} \right\} \tag{10} $$

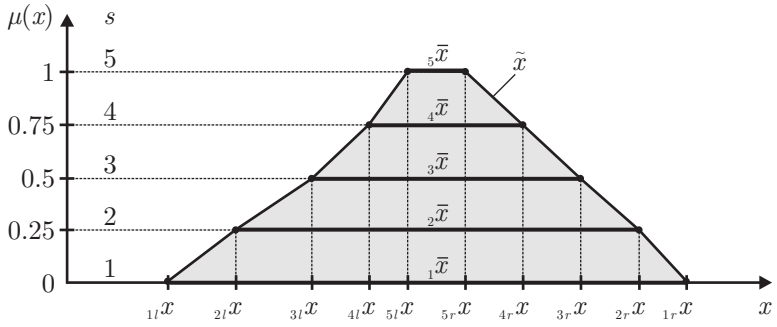


Figure 1. Fuzzy number represented by its α-cuts.

has discrete functional values (intervals ${}^{[n]}\bar{x}$) for each time point ${}^{[n]}\tau$. The time steps are equidistant, i.e. $\Delta\tau = {}^{[n]}\tau - {}^{[n-1]}\tau \quad \forall \; n = 2, \ldots, N$. Eq. (10) can also be formulated for fuzzy processes

$$\tilde{x}(\tau) = \left\{ {}^{[1]}\tilde{x}, \ldots, {}^{[n]}\tilde{x}, \ldots, {}^{[N]}\tilde{x} \right\} \; . \tag{11}$$

2.3. Functional relationships between interval or fuzzy processes

Mappings can be created to describe functional relationships between interval or fuzzy processes. Here, three types of mapping are regarded (exemplified for fuzzy processes – a formulation for interval processes is straightforward):

– **Type 1 mapping**

$$\tilde{\mathbf{x}}(\tau) \; \mapsto \; \tilde{\mathbf{z}}(\tau) \tag{12}$$

The vector of fuzzy processes $\tilde{\mathbf{x}}(\tau)$ is mapped onto the vector of fuzzy processes $\tilde{\mathbf{z}}(\tau)$ with deterministic mapping parameters.

– **Type 2 mapping**

$$\mathbf{x}(\tau) \; \tilde{\mapsto} \; \tilde{\mathbf{z}}(\tau) \tag{13}$$

The vector of deterministic processes $\mathbf{x}(\tau)$ is mapped onto the vector of fuzzy processes $\tilde{\mathbf{z}}(\tau)$ with fuzzy mapping parameters.

– **Type 3 mapping**

$$\tilde{\mathbf{x}}(\tau) \; \tilde{\mapsto} \; \tilde{\mathbf{z}}(\tau) \tag{14}$$

The vector of fuzzy processes $\tilde{\mathbf{x}}(\tau)$ is mapped onto the vector of fuzzy processes $\tilde{\mathbf{z}}(\tau)$ with fuzzy mapping parameters.

The Type 3 mapping is the general case. Type 1 and Type 2 mappings can be treated as special cases of the Type 3 mapping. The vector $\tilde{\mathbf{x}}(\tau)$ contain $j = 1, \ldots, J$ fuzzy components $\tilde{\mathbf{x}}_j(\tau)$, which are related to the $k = 1, \ldots, K$ fuzzy components $\tilde{\mathbf{z}}_k(\tau)$ of vector $\tilde{\mathbf{z}}(\tau)$. With respect to the representation of fuzzy processes in Eq. (11), the fuzzy number ${}^{[n]}\tilde{z}_k$ of time step $[n]$ can depend on all $j = 1, \ldots, J$ fuzzy numbers ${}^{[r]}\tilde{x}_j$ of prior and current time steps $[r] = [1], \ldots, [n]$.

Mathematical formulations are required to describe the mappings. These formulations contain unknown parameters which have to be identified by an inverse analysis. For Type 2 and Type 3 mappings, these parameters are fuzzy numbers or intervals. An optimization task can be formulated to identify deterministic, interval or fuzzy parameters.

2.4. Optimization task

The objective of an inverse analysis is to identify unknown deterministic, interval or fuzzy parameters. Forward analyses with deterministic processes $\mathbf{x}(\tau)$, interval processes $\bar{\mathbf{x}}(\tau)$ or fuzzy processes $\tilde{\mathbf{x}}(\tau)$ and predefined sets of parameters lead to interval processes $\bar{\mathbf{z}}^*(\tau)$ or fuzzy processes $\tilde{\mathbf{z}}^*(\tau)$. An optimization task can be performed to minimize the difference between computed results ($\bar{\mathbf{z}}^*(\tau)$ or $\tilde{\mathbf{z}}^*(\tau)$) and available data ($\bar{\mathbf{z}}(\tau)$ or $\tilde{\mathbf{z}}(\tau)$). The difference between computed and collected interval data is obtained by

$$E^h = \frac{1}{N}\frac{1}{K}\sum_{n=1}^{N}\left[\sum_{k=1}^{K}\left\{\left({}^{[n]}_{l}z_k - {}^{[n]}_{l}z^*_k\right)^2 + \left({}^{[n]}_{r}z_k - {}^{[n]}_{r}z^*_k\right)^2\right\}\right] , \tag{15}$$

whereas

$$E^h = \frac{1}{N}\frac{1}{K}\frac{1}{S}\sum_{n=1}^{N}\left[\sum_{k=1}^{K}\left\{\sum_{s=1}^{S}\left[\left({}^{[n]}_{sl}z_k - {}^{[n]}_{sl}z^*_k\right)^2 + \left({}^{[n]}_{sr}z_k - {}^{[n]}_{sr}z^*_k\right)^2\right]\right\}\right] \tag{16}$$

is used to evaluate the distance between computed and collected fuzzy data. If different patterns $h = 1, \ldots, H$ are available for parameter identification, the averaged error can be computed by

$$E^{av} = \frac{1}{H}\sum_{h=1}^{H}\left[E^h\right] . \tag{17}$$

The scaling with the number of patterns H, the number of time steps N, the number of components K, and the number of α-cuts S in the above Eqs. (15) to (17) is done due to practical reasons. It is easier to compare and evaluate errors with different selected and available numbers of H, N, K, and S.

Deterministic, interval or fuzzy parameters can be identified using Eq. (17) as objective function to be minimized. The optimization task can be solved by application of swarm intelligence.

3. Particle swarm optimization

Particle swarm optimization is a random search strategy motivated by social behavior of group individuals. Individuals of the group (swarm) are denoted as particles. Each particle is represented by a vector including all unknown parameters of the objective function – the space of search variables. First, the parameters of all particles $i = 1, \ldots, I$ of the swarm are randomly initialized. Then, the objective function is evaluated for each particle. The new position of each particle, i.e. a new set of parameters, is defined by its own search history, information of other particles, and random influences. Different ways of sharing information between particles in the swarm can be chosen, see e.g. (Fontan et al., 2011). Here, a fully connected topology is selected, i.e. each particle shares its information with all other particles in the swarm. This procedure is

applied to multiple runs $(r) = (1), \ldots, (R)$ until a predefined number of runs R is reached or the functional value of the objective function (Eq. (17)) is less than a predefined error value.

3.1. DETERMINISTIC PARAMETERS

Particle swarm optimization with deterministic particles can be used for parameter identification in case of Type 1 mapping (Eq. (12)). Each particle i is represented by a vector $\mathbf{a}^i$, which has $q = 1, \ldots, Q$ components a_q^i. The number of components Q is equal to the number of search variables, i.e. the dimension of the search space.

In each run (r), the objective function (Eq. (17)) is evaluated for each particle. The position of the best particle in the swarm, i.e. the set of parameters with the least value of the objective function in all runs, is stored as vector $\mathbf{g}$ (global best). Additionally, the best positions of each particle i are stored as vectors $\mathbf{p}^i$ (individual best).

Each component q of particle i is updated by

$$ {}^{(r+1)}a_q^i = {}^{(r)}a_q^i + {}^{(r)}\Delta a_q^i \ , \tag{18} $$

with

$$ {}^{(r)}\Delta a_q^i = c_3 \cdot {}^{(r-1)}\Delta a_q^i + c_1 \cdot d \cdot \left(p_q^i - {}^{(r)}a_q^i \right) + c_2 \cdot e \cdot \left(g_q - {}^{(r)}a_q^i \right) \tag{19} $$

for the next run $(r+1)$. In Eq. (19), d and e are realizations of independent uniformly distributed random variables in $[0,\ 1]$. For each particle i and each component q, different samples of d and e are chosen. The constants c_1, c_2, and c_3 are introduced to control the search behavior of the swarm. They are used to allow the selection of different weights for historical, individual best, and global best influences. It is common to restrict ${}^{(r)}\Delta a_q^i$ to

$$ \Delta_{\min} a_q \leq {}^{(r)}\Delta a_q^i \leq \Delta_{\max} a_q \ , \tag{20} $$

where $\Delta_{\min} a_q$ and $\Delta_{\max} a_q$ can be defined with respect to the assumed width of the q-th component, i.e. the q-th dimension of the search space, see e.g. (Eberhart and Shi, 2001).

The following conditions are defined for the first run $(r) = (1)$:

- random initialization of all particles ${}^{(1)}\mathbf{a}^i$ in the search space
- initial position is equal to individual best
- after evaluation of the objective function for all particles, best initial position is equal to global best
- prior incremental update is zero (${}^{(0)}\Delta a_q^i = 0$)

3.2. INTERVAL PARAMETERS

If interval parameters are required to map deterministic processes $\mathbf{x}(\tau)$ or interval processes $\bar{\mathbf{x}}(\tau)$ onto interval processes $\bar{\mathbf{z}}(\tau)$ (Type 2 and Type 3 mappings), an extension of the presented well known PSO algorithm is necessary. Particles $\bar{\mathbf{a}}^i$, global best $\bar{\mathbf{g}}$, and individual best $\bar{\mathbf{p}}^i$ of each particle are defined as interval numbers according to Section 2.

The update

$$ {}^{(r+1)}\bar{a}_q^i = {}^{(r)}\bar{a}_q^i + {}^{(r)}\Delta\bar{a}_q^i \tag{21} $$

is done by interval arithmetic, see e.g. (Moore, 1979). For the left bound

$$ {}^{(r)}_{l}\Delta a_q^i = {}^{(r)}_{m}\Delta a_q^i - \frac{{}^{(r)}_{w}\Delta a_q^i}{2} \tag{22} $$

and the right bound

$$ {}^{(r)}_{r}\Delta a_q^i = {}^{(r)}_{m}\Delta a_q^i + \frac{{}^{(r)}_{w}\Delta a_q^i}{2} \tag{23} $$

of ${}^{(r)}\Delta\bar{a}_q^i$, the midpoint and width representation of intervals is used, compare Eqs. (2) to (5). The incremental update of the midpoint is computed by

$$ {}^{(r)}_{m}\Delta a_q^i = c_3 \cdot {}^{(r-1)}_{m}\Delta a_q^i + c_1 \cdot d \cdot \left({}_{m}p_q^i - {}^{(r)}_{m}a_q^i\right) + c_2 \cdot e \cdot \left({}_{m}g_q - {}^{(r)}_{m}a_q^i\right) \ . \tag{24} $$

The width of interval ${}^{(r)}\Delta\bar{a}_q^i$ is obtained by

$$ {}^{(r)}_{w}\Delta a_q^i = \begin{cases} {}^{(r)}_{w}\Delta\hat{a}_q^i, & \text{if } {}^{(r)}_{w}\Delta\hat{a}_q^i \geq 0 \\ 0, & \text{if } {}^{(r)}_{w}\Delta\hat{a}_q^i < 0 \ , \end{cases} \tag{25} $$

with

$$ {}^{(r)}_{w}\Delta\hat{a}_q^i = c_3 \cdot {}^{(r-1)}_{w}\Delta a_q^i + c_1 \cdot d \cdot \left({}_{w}p_q^i - {}^{(r)}_{w}a_q^i\right) + c_2 \cdot e \cdot \left({}_{w}g_q - {}^{(r)}_{w}a_q^i\right) \ . \tag{26} $$

It should be noted, that different realizations (d and e) of independent uniformly distributed random variables are used in Eqs. (24) and (26).

3.3. Fuzzy parameters

For Type 2 and Type 3 mappings of deterministic processes $\mathbf{x}(\tau)$ or fuzzy processes $\tilde{\mathbf{x}}(\tau)$ onto fuzzy processes $\tilde{\mathbf{z}}(\tau)$, fuzzy parameters are required. PSO can be extended to fuzzy particles, i.e. particles $\tilde{\mathbf{a}}^i$, global best $\tilde{\mathbf{g}}$, and individual best $\tilde{\mathbf{p}}^i$ of each particle are fuzzy numbers, see Section 2.

Fuzzy arithmetic operations are performed (interval arithmetic for each α-cut) to get the updated particle position

$$ {}^{(r+1)}\tilde{a}_q^i = {}^{(r)}\tilde{a}_q^i + {}^{(r)}\Delta\tilde{a}_q^i \ . \tag{27} $$

For ${}^{(r)}\Delta\tilde{a}_q^i$, the left and right bounds of each α-cut s are given by

$$ {}^{(r)}_{sl}\Delta a_q^i = {}^{(r)}_{sm}\Delta a_q^i - \frac{{}^{(r)}_{sw}\Delta a_q^i}{2} \tag{28} $$

and

$$ {}^{(r)}_{sr}\Delta a_q^i = {}^{(r)}_{sm}\Delta a_q^i + \frac{{}^{(r)}_{sw}\Delta a_q^i}{2} \ , \tag{29} $$

respectively. For α-cut $s = 1$, the incremental update of the midpoint

$$ {}^{(r)}_{1m}\Delta a^i_q = c_3 \cdot {}^{(r-1)}_{1m}\Delta a^i_q + c_1 \cdot d \cdot \left({}_{1m}p^i_q - {}^{(r)}_{1m}a^i_q \right) + c_2 \cdot e \cdot \left({}_{m}g_q - {}^{(r)}_{1m}a^i_q \right) \tag{30} $$

is computed similar to the interval approach in Section 3.2, compare Eq. (24). The same holds for the width of the interval ${}^{(r)}_{1}\Delta \bar{a}^i_q$, i.e.

$$ {}^{(r)}_{1w}\Delta a^i_q = \begin{cases} {}^{(r)}_{1w}\Delta \hat{a}^i_q, & \text{if } {}^{(r)}_{1w}\Delta \hat{a}^i_q \geq 0 \\ 0, & \text{if } {}^{(r)}_{1w}\Delta \hat{a}^i_q < 0 \,, \end{cases} \tag{31} $$

with

$$ {}^{(r)}_{1w}\Delta \hat{a}^i_q = c_3 \cdot {}^{(r-1)}_{1w}\Delta a^i_q + c_1 \cdot d \cdot \left({}_{1w}p^i_q - {}^{(r)}_{1w}a^i_q \right) + c_2 \cdot e \cdot \left({}_{1w}g_q - {}^{(r)}_{1w}a^i_q \right) \,. \tag{32} $$

For all other α-cuts ($s > 1$), three cases are distinguished for the incremental update of the midpoint

$$ {}^{(r)}_{sm}\Delta a^i_q = \begin{cases} {}^{(r)}_{sm}\Delta \hat{a}^i_q, & \text{if } {}^{(r)}_{s-1l}\Delta a^i_q \leq {}^{(r)}_{sm}\Delta \hat{a}^i_q \leq {}^{(r)}_{s-1r}\Delta a^i_q \\ {}^{(r)}_{s-1l}\Delta a^i_q, & \text{if } {}^{(r)}_{s-1l}\Delta a^i_q > {}^{(r)}_{sm}\Delta \hat{a}^i_q \\ {}^{(r)}_{s-1r}\Delta a^i_q, & \text{if } {}^{(r)}_{s-1r}\Delta a^i_q < {}^{(r)}_{sm}\Delta \hat{a}^i_q \,, \end{cases} \tag{33} $$

with

$$ {}^{(r)}_{sm}\Delta \hat{a}^i_q = c_3 \cdot {}^{(r-1)}_{sm}\Delta a^i_q + c_1 \cdot d \cdot \left({}_{sm}p^i_q - {}^{(r)}_{sm}a^i_q \right) + c_2 \cdot e \cdot \left({}_{sm}g_q - {}^{(r)}_{sm}a^i_q \right) \tag{34} $$

and for the incremental update of the width

$$ {}^{(r)}_{sw}\Delta a^i_q = \begin{cases} {}^{(r)}_{sw}\Delta \hat{a}^i_q, & \text{if } 0 \leq {}^{(r)}_{sw}\Delta \hat{a}^i_q \leq {}^{(r)}_{sw}\Delta_{\max} \hat{a}^i_q \\ {}^{(r)}_{sw}\Delta_{\max} \hat{a}^i_q, & \text{if } {}^{(r)}_{sw}\Delta \hat{a}^i_q > {}^{(r)}_{sw}\Delta_{\max} \hat{a}^i_q \\ 0, & \text{if } {}^{(r)}_{sw}\Delta \hat{a}^i_q < 0 \,, \end{cases} \tag{35} $$

with

$$ {}^{(r)}_{sw}\Delta \hat{a}^i_q = c_3 \cdot {}^{(r-1)}_{sw}\Delta a^i_q + c_1 \cdot d \cdot \left({}_{sw}p^i_q - {}^{(r)}_{sw}a^i_q \right) + c_2 \cdot e \cdot \left({}_{sw}g_q - {}^{(r)}_{sw}a^i_q \right) \tag{36} $$

and

$$ {}^{(r)}_{sw}\Delta_{\max} \hat{a}^i_q = 2 \cdot \min \left[\left({}^{(r)}_{sm}\Delta a^i_q - {}^{(r)}_{s-1l}\Delta a^i_q \right), \left({}^{(r)}_{s-1r}\Delta a^i_q - {}^{(r)}_{sm}\Delta a^i_q \right) \right] \,. \tag{37} $$

Different realizations d and e of independent uniformly distributed random variables are used for each α-cut. If α-cut $s = S$ is restricted to give deterministic numbers, only midpoints are updated for α-cut S.

4. Artificial neural networks for interval and fuzzy data

Artificial neural network concepts can be applied to map deterministic processes $\mathbf{x}(\tau)$, interval processes $\bar{\mathbf{x}}(\tau)$ or fuzzy processes $\tilde{\mathbf{x}}(\tau)$ onto interval processes $\bar{\mathbf{z}}(\tau)$ or fuzzy processes $\tilde{\mathbf{z}}(\tau)$. Two ways of computation are possible to process interval or fuzzy data with neural networks:

1. interval arithmetic (for each α-cut)

2. optimization (α-level optimization)

Interval arithmetic approaches for deterministic network parameters are presented in (Graf et al., 2010). Extensions for a priori defined and trainable interval or fuzzy network parameters are published in (Freitag et al., 2011c) and (Freitag et al., 2011a), respectively. Algorithms for signal computation with α-level optimization can be found in (Freitag, 2010) and (Freitag et al., 2011b). In the following, the neural network approaches are formulated for fuzzy data and Type 3 mapping, see Eq. (14). However, they can also be applied to interval data or Type 1 and Type 2 mappings, which is straightforward.

4.1. FEED FORWARD NEURAL NETWORKS

If the fuzzy number ${}^{[n]}\tilde{z}_k$ of time step $[n]$ depends on the $j = 1, \ldots, J$ current fuzzy numbers ${}^{[n]}\tilde{x}_j$ only, feed forward networks can be used as mathematical formulation of the mappings introduced in Section 2.3. Neural networks with feed forward architecture consist of (M) layers, i.e. an input layer, $(M-2)$ hidden layers and an output layer. The number of input and output neurons is given by the number of components J and K, respectively. The number of hidden layers and neurons has to be defined with respect to the complexity of the formulation. In general, fully connected networks are considered, i.e. each neuron in layer (m) has synaptic connections to all neurons in the following layer $(m+1)$, see Figure 2. For specific applications, special network structures may be created, see Section 5.

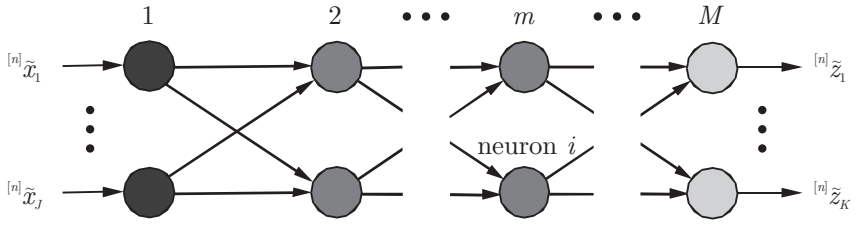


Figure 2. Feed forward neural network.

In each time step $[n]$, the fuzzy components ${}^{[n]}\tilde{x}_j$ (e.g. structural actions) may be transformed to dimensionless fuzzy network input signals, e.g.

$$ {}^{[n]}\tilde{x}_j^{(1)} = \frac{{}^{[n]}\tilde{x}_j}{x_j^{sc}} \, . \tag{38} $$

The dimensionless network output signals ${}^{[n]}\tilde{x}_k^{(M)}$ may be scaled to fuzzy components (e.g. structural responses)

$$ {}^{[n]}\tilde{z}_k = {}^{[n]}\tilde{x}_k^{(M)} \cdot z_k^{sc} \, . \tag{39} $$

For each component j and k, the scaling parameters x_j^{sc} and z_k^{sc} can be defined as the maximum absolute value of its possible positive and / or negative values.

The signals of feed forward neural networks are computed layer by layer. In the hidden and output neurons, fuzzy output signals

$$ {}^{[n]}\tilde{x}_i^{(m)} = \tilde{\varphi}_i^{(m)} \left(\sum_{h=1}^{H} \left[{}^{[n]}\tilde{x}_h^{(m-1)} \cdot \tilde{w}_{ih}^{(m)} \right] + \tilde{b}_i^{(m)} \right) \tag{40} $$

are computed by means of a fuzzy activation function $\tilde{\varphi}_i^{(m)}(.)$. These fuzzy output signals are transfered by synaptic connections to the neurons of the next layer. The argument of the fuzzy activation function of neuron i in layer (m) contains all fuzzy output signals ${}^{[n]}\tilde{x}_h^{(m-1)}$ of the previous layer $(m-1)$ multiplied by the fuzzy weights $\tilde{w}_{ih}^{(m)}$ and a fuzzy bias value $\tilde{b}_i^{(m)}$. Various types of monotonic and differentiable fuzzy activation functions can be used, see e.g. (Freitag, 2010).

The fuzzy weights, fuzzy bias values and perhaps parameters of the fuzzy activation function are unknown fuzzy network parameters. The PSO approaches presented in Section 3 can be used for parameter identification. It is proposed to initialize the particle components representing fuzzy weights and fuzzy bias values randomly, e.g. in $[-1,\ 1]$. In general, the values of these fuzzy parameters are not restricted in the search space. The search space can be restricted with respect to selected particle components representing the fuzzy factors of the fuzzy activation functions.

4.2. Recurrent neural networks

More general is the assumption, that all $j = 1, \ldots, J$ fuzzy numbers ${}^{[r]}\tilde{x}_j$ of prior and current time steps $[r] = [1], \ldots, [n]$ have influences to the current fuzzy number ${}^{[n]}\tilde{z}_k$ of time step $[n]$. In this case, recurrent neural networks are suitable to formulate the mappings according to Section 2.3.

In addition to feed forward networks, context neurons are used to consider the whole history for the computation of the current fuzzy number ${}^{[n]}\tilde{z}_k$ of time step $[n]$. All hidden and output neurons are connected to their context neurons, see Figure 3.

In each context neuron, the fuzzy output signal is transfered to the fuzzy context signal

$$ {}^{[n]}\tilde{y}_i^{(m)} = {}^{[n]}\tilde{x}_i^{(m)} + {}^{[n-1]}\tilde{y}_i^{(m)} \cdot \tilde{\lambda}_i^{(m)} \ . \tag{41} $$

The influence of the previous fuzzy context signal ${}^{[n-1]}\tilde{y}_i^{(m)}$ is considered by the fuzzy feedback factor $\tilde{\lambda}_i^{(m)}$. Fuzzy feedback factors $\tilde{\lambda}_i^{(m)}$ are additional fuzzy network parameters defined in the interval $[0,\ 1]$.

Each context neuron sends weighted fuzzy signals with a time delay of one time step to all hidden or output neurons in its layer. Hence, Eq. (40) must be extended to

$$ {}^{[n]}\tilde{x}_i^{(m)} = \tilde{\varphi}_i^{(m)} \left(\sum_{h=1}^{H} \left[{}^{[n]}\tilde{x}_h^{(m-1)} \cdot \tilde{w}_{ih}^{(m)} \right] + \sum_{q=1}^{I} \left[{}^{[n-1]}\tilde{y}_q^{(m)} \cdot \tilde{c}_{iq}^{(m)} \right] + \tilde{b}_i^{(m)} \right) \tag{42} $$

in order to consider the fuzzy context signals ${}^{[n-1]}\tilde{y}_q^{(m)}$ multiplied by the fuzzy context weights $\tilde{c}_{iq}^{(m)}$. It should be noted, that a feed forward neural network is obtained as a special case of the discussed recurrent neural network, if all fuzzy context weights are set to zero.

The fuzzy context weights $\tilde{c}_{iq}^{(m)}$ and fuzzy feedback factors $\tilde{\lambda}_i^{(m)}$ are additional unknown fuzzy network parameters, which can be identified by the introduced PSO approaches, see Section 3. The search space

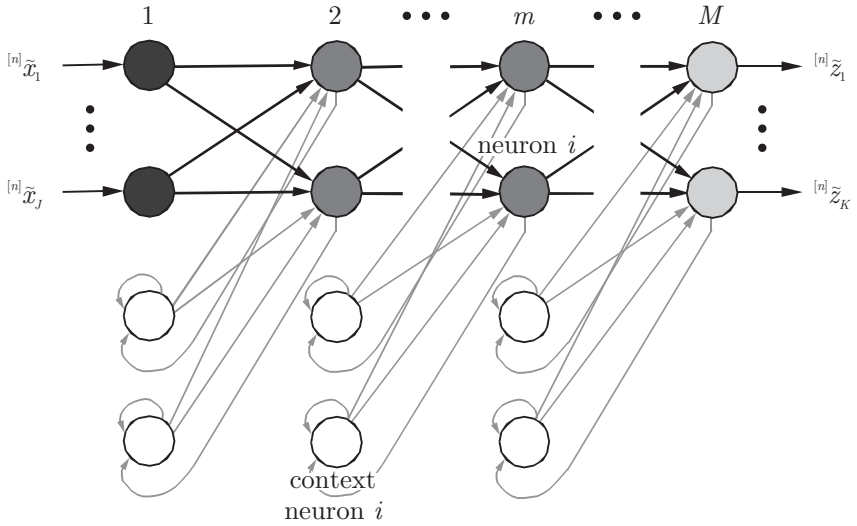


Figure 3. Recurrent neural network.

must be restricted to $[0,\ 1]$ for the particle components, which represent fuzzy feedback factors $\tilde{\lambda}_i^{(m)}$. If an updated interval bound is less than zero or greater than one, it is set to zero or to one, respectively. The search space is not restricted for fuzzy context weights $\tilde{c}_{iq}^{(m)}$. They can be initialized randomly in $[-1,\ 1]$.

5. Application for time-dependent material behavior

The presented neural network approaches can be applied to describe uncertain material behavior. Uncertain nonlinear stress-strain dependencies can be identified with feed forward neural networks for elastic material behavior. For nonlinear stress-strain-time dependencies (viscous material behavior), recurrent neural networks can be utilized. An α-level optimization is applied to compute the network outputs.

Fuzzy strain processes can be mapped onto fuzzy stress processes or vice versa. Here, an approach for strain to stress mapping is presented. In this case, the fuzzy processes $\tilde{\mathbf{x}}(\tau)$ represent fuzzy strain processes $\tilde{\varepsilon}(\tau)$ and the fuzzy processes $\tilde{\mathbf{z}}(\tau)$ correspond to fuzzy stress processes $\tilde{\sigma}(\tau)$. The strain and stress vectors include all components, which are required for strain and stress tensors ($J = K = 6$ for 3D, $J = K = 3$ for 2D and $J = K = 1$ for 1D material models). The fuzzy network parameters can be identified by results of experimental investigations.

5.1. Tangential stiffness

Applications of neural network based constitutive models within the finite element method require the tangential stiffness matrix of the material description ${}^{[n]}\tilde{\mathbf{C}}$ in order to get the tangential system stiffness matrix. The components of the uncertain tangential stiffness matrix

$$ {}^{[n]}\tilde{C}_{kj} = \frac{\partial {}^{[n]}\Delta\tilde{\sigma}_k}{\partial {}^{[n]}\Delta\tilde{\varepsilon}_j} \tag{43} $$

are determined in linearized form by the partial derivatives of the incremental fuzzy stress components

$$ {}^{[n]}\Delta\tilde{\sigma}_k = {}^{[n]}\tilde{\sigma}_k - {}^{[n-1]}\tilde{\sigma}_k \tag{44} $$

with respect to the incremental fuzzy strain components

$$ {}^{[n]}\Delta\tilde{\varepsilon}_j = {}^{[n]}\tilde{\varepsilon}_j - {}^{[n-1]}\tilde{\varepsilon}_j \ . \tag{45} $$

The incremental fuzzy stresses

$$ {}^{[n]}\Delta\tilde{\sigma}_k = \left({}^{[n]}\tilde{x}_k^{(M)} - {}^{[n-1]}\tilde{x}_k^{(M)} \right) \cdot z_k^{sc} \tag{46} $$

contain the fuzzy output signals of the neural network ${}^{[n]}\tilde{x}_k^{(M)}$ (time step $[n]$) and ${}^{[n-1]}\tilde{x}_k^{(M)}$ (time step $[n-1]$). The chain rule is applied two times in Eq. (43), which leads to

$$ \frac{\partial^{[n]}\Delta\tilde{\sigma}_k}{\partial^{[n]}\Delta\tilde{\varepsilon}_j} = \frac{\partial\left({}^{[n]}\tilde{\sigma}_k - {}^{[n-1]}\tilde{\sigma}_k \right)}{\partial^{[n]}\Delta\tilde{\varepsilon}_j} = \frac{\partial^{[n]}\tilde{\sigma}_k}{\partial^{[n]}\Delta\tilde{\varepsilon}_j} = \frac{\partial^{[n]}\tilde{\sigma}_k}{\partial^{[n]}\tilde{x}_k^{(M)}} \cdot \frac{\partial^{[n]}\tilde{x}_k^{(M)}}{\partial^{[n]}\tilde{x}_j^{(1)}} \cdot \frac{\partial^{[n]}\tilde{x}_j^{(1)}}{\partial^{[n]}\Delta\tilde{\varepsilon}_j} \ . \tag{47} $$

The partial derivatives of the fuzzy stress components with respect to the fuzzy output signals in Eq. (47) are obtained by

$$ \frac{\partial^{[n]}\tilde{\sigma}_k}{\partial^{[n]}\tilde{x}_k^{(M)}} = \frac{\partial\left({}^{[n]}\tilde{x}_k^{(M)} \cdot z_k^{sc} \right)}{\partial^{[n]}\tilde{x}_k^{(M)}} = z_k^{sc} \ . \tag{48} $$

The partial derivatives of the fuzzy input signals with respect to the incremental fuzzy strain components

$$ \frac{\partial^{[n]}\tilde{x}_j^{(1)}}{\partial^{[n]}\Delta\tilde{\varepsilon}_j} = \frac{\partial\left(\frac{{}^{[n]}\tilde{\varepsilon}_j}{x_j^{sc}} \right)}{\partial^{[n]}\Delta\tilde{\varepsilon}_j} = \frac{1}{x_j^{sc}} \cdot \frac{\partial\left({}^{[n]}\Delta\tilde{\varepsilon}_j + {}^{[n-1]}\tilde{\varepsilon}_j \right)}{\partial^{[n]}\Delta\tilde{\varepsilon}_j} = \frac{1}{x_j^{sc}} \tag{49} $$

in Eq. (47) are evaluated using Eq. (38) (with ${}^{[n]}\tilde{x}_j = {}^{[n]}\tilde{\varepsilon}_j$) and (45). Eqs. (48) and (49) are substituted in Eq. (47) and hence, the components of the tangential stiffness matrix are obtained by

$$ {}^{[n]}\tilde{C}_{kj} = \frac{z_k^{sc}}{x_j^{sc}} \cdot \frac{\partial^{[n]}\tilde{x}_k^{(M)}}{\partial^{[n]}\tilde{x}_j^{(1)}} \ . \tag{50} $$

The partial derivatives of the network output signals ${}^{[n]}\tilde{x}_k^{(M)}$ with respect to the network input signals ${}^{[n]}\tilde{x}_j^{(1)}$ are evaluated using multiple applications of the chain rule. An efficient algorithm to compute these partial derivatives is presented in (Freitag et al., 2011b).

5.2. SPECIAL NETWORK STRUCTURES

Physical boundary conditions of investigated materials can be considered by creating special network structures. The tangential stiffness matrix should be symmetric for materials with isotropic properties. This condition (${}^{[n]}\tilde{C}_{kj} = {}^{[n]}\tilde{C}_{jk}$) can be fulfilled by

$$\frac{\partial {}^{[n]}\tilde{x}_k^{(M)}}{\partial {}^{[n]}\tilde{x}_j^{(1)}} = \frac{\partial {}^{[n]}\tilde{x}_j^{(M)}}{\partial {}^{[n]}\tilde{x}_k^{(1)}} \ , \tag{51}$$

if $x_j^{sc} = x_k^{sc}$ and $z_k^{sc} = z_j^{sc}$, see Eq. (50). Symmetric partial derivatives of the network output signals ${}^{[n]}\tilde{x}_k^{(M)}$ with respect to the network input signals ${}^{[n]}\tilde{x}_j^{(1)}$ can be guaranteed for networks with three layers and linear activation functions (with the same slope parameter) in the output layer. In Figure 4, the symmetry of the synaptic connections is exemplified. The symmetry condition for deterministic, interval or fuzzy weights

$$\tilde{w}_{ij}^{(2)} = \tilde{w}_{ji}^{(3)} \tag{52}$$

is also valid for recurrent neural networks and arbitrary numbers of neurons in the three layers.

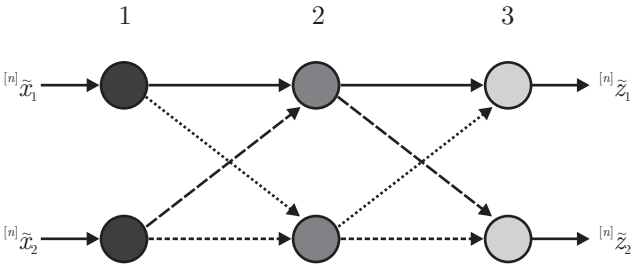


Figure 4. Neural network with symmetric derivatives.

Some stress and strain components are decoupled for isotropic or orthotropic behavior, i.e. the stress component ${}^{[n]}\tilde{\sigma}_k$ only depends on its corresponding strain component ${}^{[n]}\tilde{\varepsilon}_k$. This can be achieved by partially connected neural networks, see Figure 5. Selected weights and context weights are set to zero, which is equivalent to cut synaptic connections in a fully connected feed forward or recurrent neural network.

A simple feed forward network with two layers (no hidden layers) and linear activation functions (identity function) is equivalent to Hooke's law for linear elastic material. In Figure 6, the neural network representation of linear elastic material is demonstrated for the 3D case. The deterministic, interval or fuzzy weights are $\tilde{w}_{11} = \tilde{w}_{22} = \tilde{w}_{33} = \tilde{c}_1$, $\tilde{w}_{44} = \tilde{w}_{55} = \tilde{w}_{66} = \tilde{c}_2$, $\tilde{w}_{12} = \tilde{w}_{13} = \tilde{w}_{23} = \tilde{w}_{21} = \tilde{w}_{31} = \tilde{w}_{32} = \tilde{c}_3$ for isotropic material behavior, see Eq. (53).

$$\tilde{\mathbf{C}} = \begin{bmatrix} \tilde{c}_1 & \tilde{c}_3 & \tilde{c}_3 & 0 & 0 & 0 \\ \tilde{c}_3 & \tilde{c}_1 & \tilde{c}_3 & 0 & 0 & 0 \\ \tilde{c}_3 & \tilde{c}_3 & \tilde{c}_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{c}_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{c}_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \tilde{c}_2 \end{bmatrix} \tag{53}$$

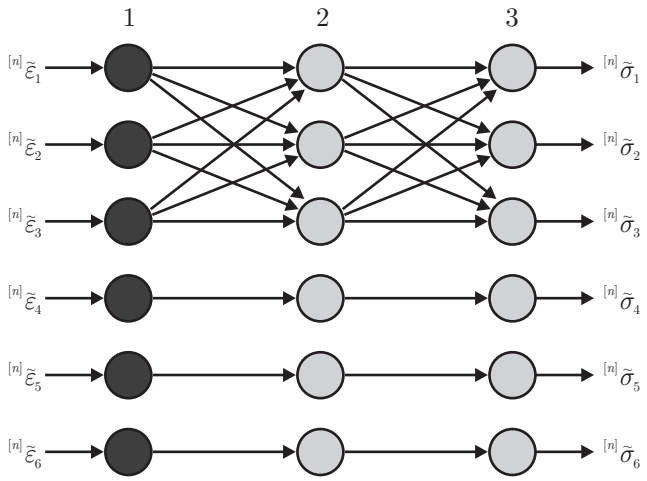


Figure 5. Partially connected neural network with symmetric derivatives.

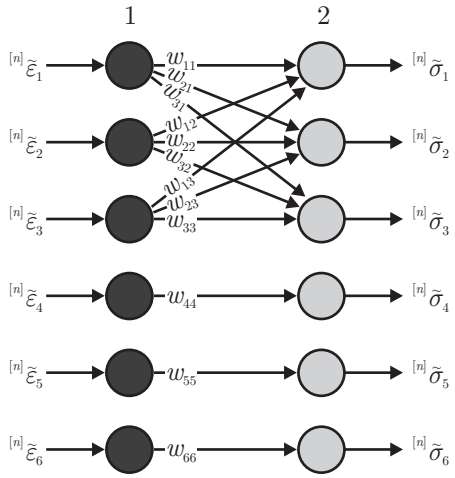


Figure 6. Feed forward neural network for linear elastic material behavior.

6. Examples

6.1. Verification with 1D fractional rheological model

The presented recurrent neural network approach is applied to identify and to predict uncertain stress-strain-time dependencies of the fuzzy fractional Newton element. The differential equation

$$\tilde{\sigma}(\tau) = \tilde{p}\,\frac{\mathrm{d}^{\tilde{r}}}{\mathrm{d}\tau^{\tilde{r}}}\,\tilde{\varepsilon}(\tau) \tag{54}$$

of this rheological element, see e.g. (Oeser and Freitag, 2009), contains a fractional derivative of strain $\tilde{\varepsilon}(\tau)$ with respect to time τ. In this example, $\tilde{p}$ is defined as deterministic parameter $\tilde{p} = p = 101\,000\ \ (\mathrm{MPa\,s}^r)$.

The operator $\tilde{r}$ represents the order of the derivative. It is a fuzzy number between zero (linear elastic spring) and one (dashpot). Here, it is defined as a fuzzy number with triangular shape $\tilde{r} = \langle 0.13,\ 0.14,\ 0.15 \rangle$).

The fractional differential equation (54) is solved by the Laplace transform. The strain boundary condition $\tilde{\varepsilon}(\tau) = \tilde{\varepsilon}^*$ is used to obtain the relaxation function of the fuzzy fractional Newton element. Convolution of the relaxation function and time step discretization of the fuzzy strain process (equidistant time steps $\Delta\tau$) lead to

$$ {}^{[n]}\tilde{\sigma} = \sum_{i=1}^{n} \left\{ \frac{p \cdot {}^{[i]}\Delta\tilde{\varepsilon}}{\Gamma(2-\tilde{r}) \cdot \Delta\tau^{\tilde{r}}} \left[(n+1-i)^{(1-\tilde{r})} - (n-i)^{(1-\tilde{r})} \right] \right\} . \tag{55} $$

It can be seen, that the stress in time step $[n]$ dependents on the current strain and the whole strain history. Eq. (55) is utilized to verify the presented recurrent neural network approach. Training and validation patterns are computed by solving Eq. (55) within a fuzzy analysis (FA) (α-level optimization (Möller et al., 2000)). The time step length $\Delta\tau = 100$ s is chosen. Three α-cuts ($\alpha_1 = 0$, $\alpha_2 = 0.5$ and $\alpha_3 = 1$) are evaluated. The same fuzzy stress and fuzzy strain processes as presented in (Freitag et al., 2010b) are utilized, see Figures 7 to 10.

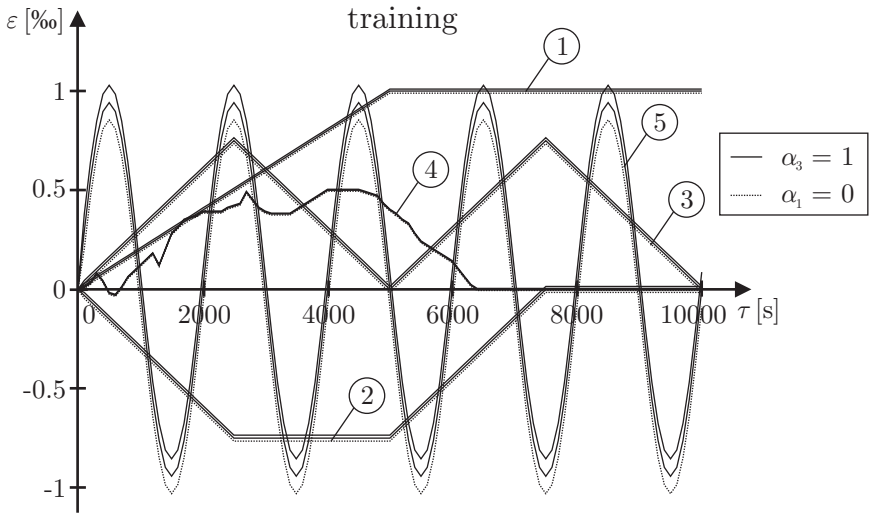


Figure 7. Fuzzy strain processes for network training.

The five fuzzy strain processes and the corresponding five fuzzy stress processes plotted in Figures 7 and 8 are used to train a recurrent neural network for Type 3 mapping ($\tilde{\varepsilon}(\tau) \tilde{\mapsto} \tilde{\sigma}(\tau)$). Nonlinear activation functions in the form of the area hyperbolic sine (arsinh) are used in the hidden neurons and a linear activation function is used in the output neuron. The signals of the recurrent neural network are computed by interval arithmetic operations for each α-cut. The developed PSO approach for fuzzy numbers is applied to identify the fuzzy network parameters. The number of particles is selected as $I = 20$. The control parameters ($c_1 = c_2 = 1.494$ and $c_3 = 0.729$) are defined according to (Eberhart and Shi, 2001). The training results of the recurrent neural network (RNN) are shown in Figure 8.

The five additional fuzzy strain processes in Figure 9 are used to validate the identified uncertain stress-strain-time dependency. The recurrent neural network predictions show a very good agreement with the desired responses obtained by a fuzzy analysis using Eq. (55), see Figure 10.

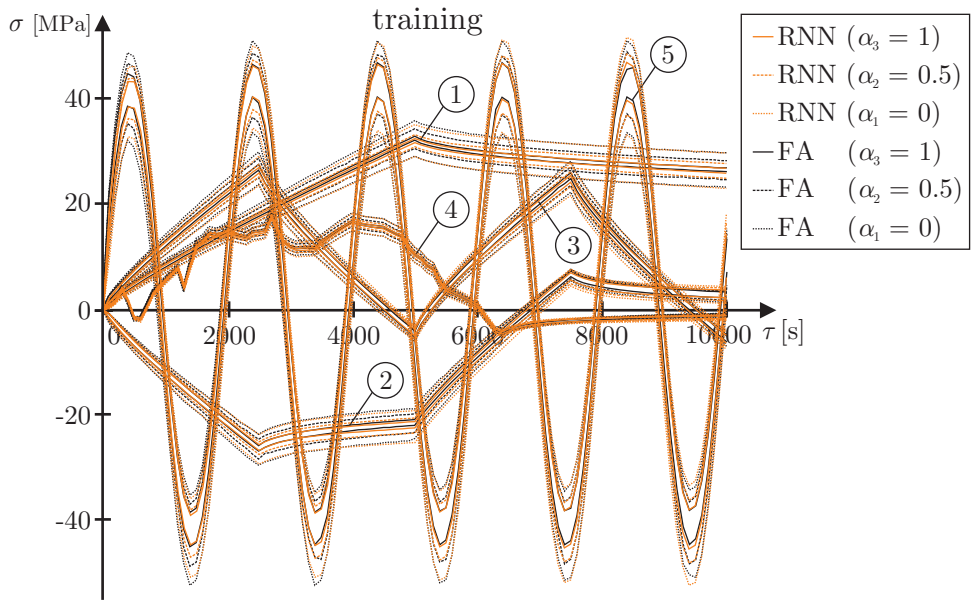


Figure 8. Fuzzy stress processes for network training.

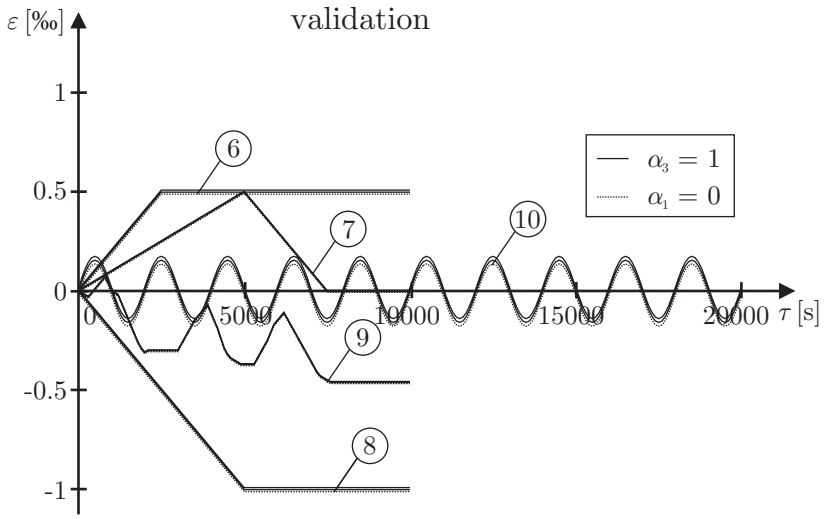


Figure 9. Fuzzy strain processes for network validation.

The same quality is achieved in comparison with the results in (Freitag et al., 2010b), where a backpropagation training algorithm has been applied. But here, a recurrent neural network with three hidden neurons and four context neurons ($1-3-1$ architecture) was sufficient for PSO training, whereas three hidden layers with 13 hidden and 14 context neurons in total ($1-5-5-3-1$ architecture) were required for backpropagation training.

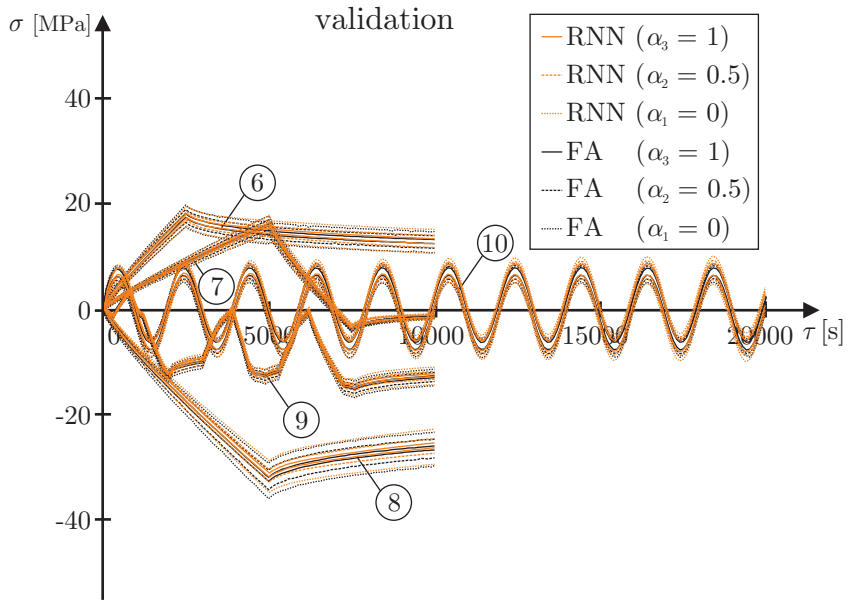


Figure 10. Fuzzy stress processes for network validation.

6.2. VERIFICATION WITH 3D MATERIAL MODEL

The proposed strategy for symmetric and decoupled stiffness is verified by a 3D linear elastic material model. Here, results for the mapping of deterministic strain processes onto deterministic stress processes are presented (special case of Type 1 mapping). The modulus of elasticity $E = 210\,000\,\mathrm{MPa}$ and Poisson's ratio $\nu = 0.2$ lead to the deterministic tangential stiffness matrix

$$\mathbf{C} = \begin{bmatrix} 233333 & 58333 & 58333 & 0 & 0 & 0 \\ 58333 & 233333 & 58333 & 0 & 0 & 0 \\ 58333 & 58333 & 233333 & 0 & 0 & 0 \\ 0 & 0 & 0 & 87500 & 0 & 0 \\ 0 & 0 & 0 & 0 & 87500 & 0 \\ 0 & 0 & 0 & 0 & 0 & 87500 \end{bmatrix} \mathrm{MPa}\ . \tag{56}$$

Deterministic stress and strain processes (two patterns with $N = 1000$ time steps each) are used to train and validate a recurrent neural network with one hidden layer comprising 6 neurons ($6 - 6 - 6$ architecture). Linear activation functions are used in the output neurons and nonlinear activation functions in the form of the area hyperbolic sine are used in the hidden neurons. The network has 12 context neurons to consider possible history dependencies in the data series. However, the time-independent mapping of the strain vector ${}^{[n]}\varepsilon$ onto the stress vector ${}^{[n]}\sigma$ should be learned by the recurrent neural network.

The discussed PSO approach is applied to identify the deterministic network parameters. The same number of particles ($I = 20$) and values for the constants $c_1 = c_2 = 1.494$ and $c_3 = 0.729$ are used as in the previous example. The symmetry condition of Eq. (52) is used to get a symmetric tangential stiffness matrix, which is not possible by applying backpropagation training algorithms, see (Freitag et al., 2011b).

The training pattern has been presented 10^5 times to the network to identify the linearity and the time-independence between the strain and stress processes. As a result, the tangential stiffness matrix of the training pattern (Tr)

$$\mathbf{C}^{Tr} = \begin{bmatrix} 232826 & 58225 & 58191 & 0 & 0 & 0 \\ 58225 & 232769 & 58211 & 0 & 0 & 0 \\ 58191 & 58211 & 232743 & 0 & 0 & 0 \\ 0 & 0 & 0 & 87373 & 0 & 0 \\ 0 & 0 & 0 & 0 & 87372 & 0 \\ 0 & 0 & 0 & 0 & 0 & 87369 \end{bmatrix} \text{MPa} \tag{57}$$

is obtained, which contains the mean values of the partial derivatives of the stress components with respect to the strain components considering all 1000 time steps.

The network prediction has been verified with a second pattern comprising $N = 1000$ time steps, too. The mean values of the partial derivatives of the stress components with respect to the strain components of the validation pattern (V) are summarized as

$$\mathbf{C}^{V} = \begin{bmatrix} 232911 & 58247 & 58210 & 0 & 0 & 0 \\ 58247 & 232856 & 58232 & 0 & 0 & 0 \\ 58210 & 58232 & 232824 & 0 & 0 & 0 \\ 0 & 0 & 0 & 87385 & 0 & 0 \\ 0 & 0 & 0 & 0 & 87381 & 0 \\ 0 & 0 & 0 & 0 & 0 & 87364 \end{bmatrix} \text{MPa} \, . \tag{58}$$

The error is less than 0.25% for all components of the tangential stiffness matrix computed with the training and the validation patterns. In future works, symmetric network structures will also be applied to describe uncertain stress-strain-time dependencies.

7. Conclusion

In this paper, a new training strategy for artificial neural networks is presented. It is based on swarm intelligence. PSO approaches for interval and fuzzy numbers are developed accounting for uncertainty in measurements. These approaches have the flexibility of modifying all parameters during training of recurrent neural networks. Additionally, special network structures can be created, which is important for using neural networks as constitutive models. An application for time-dependent material behavior is presented. Results of verifications with a fuzzy fractional Newton element and a 3D linear elastic material model show high approximation quality of the developed neural network approaches. The new approaches can be applied to measured interval and fuzzy data. Recurrent neural networks for uncertain data can be utilized as constitutive models within interval, fuzzy, and fuzzy stochastic finite element analyses.

References

Adeli, H. Neural Networks in Civil Engineering: 1989-2000. *Computer-Aided Civil and Infrastructure Engineering*, 16:126–142, 2001.

Eberhart, R. C. and Y. Shi. Particle Swarm Optimization: Developments, Applications and Resources. In *Proceedings of the 2001 IEEE Congress on Evolutionary Computation*, pp. 81–86, Seoul, 2001.

Fontan, M., A. Ndiaye, D. Breysse, F. Bos, and C. Fernandez. Soil–structure interaction: Parameters identification using particle swarm optimization. *Computers and Structures*, 89(17-18):1602-1614, 2011.

Freitag, S. *Modellfreie numerische Prognosemethoden zur Tragwerksanalyse*. Veröffentlichungen – Institut für Statik und Dynamik der Tragwerke, Heft 19, Technische Universität Dresden, 2010.

Freitag, S., W. Graf, and M. Kaliske. Identification and prediction of time-dependent structural behavior with recurrent neural networks for uncertain data. In M. Beer, R. L. Muhanna, and R. L. Mullen, editors, *Proceedings of the 4th International Workshop on Reliable Engineering Computing*, pp. 577–596, Singapore, 2010. Research Publishing Services, Singapore.

Freitag, S., W. Graf, and M. Kaliske. Prediction of Time-Dependent Structural Responses with Recurrent Neural Networks. *Proceedings in Applied Mathematics and Mechanics*, 10:155–156, 2010.

Freitag, S., W. Graf, and M. Kaliske. Recurrent Neural Networks for Fuzzy Data. *Integrated Computer-Aided Engineering*, 18(3):265–280, 2011.

Freitag, S., W. Graf, and M. Kaliske. Recurrent Neural Networks for Fuzzy Data as a Material Description within the Finite Element Method. In Y. Tsompanakis and B. H. V. Topping, editors, *Proceedings of the Second International Conference on Soft Computing Technology in Civil, Structural and Environmental Engineering*, paper 28, Chania, 2011. Civil-Comp Press, Stirlingshire.

Freitag, S., W. Graf, M. Kaliske, and J.-U. Sickert. Prediction of time-dependent structural behaviour with recurrent neural networks for fuzzy data. *Computers and Structures*, 89(21-22):1971–1981, 2011.

Graf, W., S. Freitag, M. Kaliske, and J.-U. Sickert. Recurrent neural networks for uncertain time-dependent structural behavior. *Computer-Aided Civil and Infrastructure Engineering*, 25(5):322–333, 2010.

Graf, W., J.-U. Sickert, S. Freitag, S. Pannier, and M. Kaliske. Neural Network Approaches in Structural Analysis under Consideration of Imprecision and Variability. In Y. Tsompanakis and B. H. V. Topping, editors, *Soft Computing Methods for Civil and Structural Engineering*, Chapter 4, pp. 59–85. Saxe-Coburg Publications, Stirlingshire, 2011.

Haykin, S. *Neural Networks – A Comprehensive Foundation*. Prentice-Hall, Upper Saddle River, 1999.

Kennedy, J. and R. C. Eberhart. Particle Swarm Optimization. In *Proceedings of the IEEE International Conference on Neural Networks IV*, pp. 1942–1948. IEEE, Piscataway, 1995.

Kennedy, J., R. C. Eberhart, and S. Yuhui. *Swarm Intelligence*. Morgan Kaufmann, San Francisco, 2001.

Kuok, K. K., S. Harun, and S. M. Shamsuddin. Particle swarm optimization feedforward neural network for modeling runoff. *International Journal of Environmental Science and Technology*, 7(1):67–78, 2010.

Li, L. J., Z. B. Huang, F. Liu, and Q. H. Wu. A heuristic particle swarm optimizer for optimization of pin connected structures. *Computers and Structures*, 85:340–349, 2007.

Mendes, R., P. Cortez, M. Rocha, and J. Neves. Particle Swarms for Feedforward Neural Network Training. In *Proceedings of the 2002 International Joint Conference on Neural Networks*, pp. 1895–1899. IEEE, Honolulu, 2002.

Möller, B. and M. Beer. Engineering computation under uncertainty – Capabilities of non-traditional models. *Computers and Structures*, 86(10):1024–1041, 2008.

Möller, B., W. Graf, and M. Beer. Fuzzy structural analysis using α-level optimization. *Computational Mechanics*, 26(6):547–565, 2000.

Moens, D. and D. Vandepitte. A survey of non-probabilistic uncertainty treatment in finite element analysis. *Computer Methods in Applied Mechanics and Engineering*, 194:1527–1555, 2005.

Moore, R. E. *Methods and Applications of Interval Analysis*. SIAM, Studies in Applied Mathematics, 2, Philadelphia, 1979.

Muhanna, R. L., H. Zhang, and R. L. Mullen. Interval Finite Elements as a Basis for Generalized Models of Uncertainty in Engineering. *Reliable Computing*, 13(2):173–194, 2007.

Oeser, M. and S. Freitag. Modeling of materials with fading memory using neural networks. *International Journal for Numerical Methods in Engineering*, 78(7):843–862, 2009.

Perez, R. E. and K. Behdinan. Particle swarm approach for structural design optimization. *Computers and Structures*, 85:1579–1588, 2007.

Rao, M. V. R., R. L. Mullen, and R. L. Muhanna. A new interval finite element formulation with the same accuracy in primary and derived variables. *International Journal for Reliability and Safety*, 5(3/4):336–357, 2011.

Sickert, J.-U., S. Freitag, and W. Graf. Prediction of uncertain structural behaviour and robust design. *International Journal for Reliability and Safety*, 5(3/4):358–377, 2011.

Zhang, J.-R., J. Zhang, T.-M. Lok, and M. R. Lyu. A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training. *Applied Mathematics and Computation*, 185:1026–1037, 2007.

Yield Maximization Technique for Microwave Circuits using NEWUOA and Space Mapping Algorithm

A. S. O. Hassan, A. S. A. Mohamed, and A. Y. El-Sharabasy
Department of Engineering Mathematics and Physics, Faculty of Engineering, Cairo University, Giza, 12211, Egypt; E-mail: a.y.elsharabasy@gmail.com

Abstract: An efficient yield optimization technique for the microwave circuits is introduced. Generally, in the classical design and optimization problem, we look for one single point, in the designable parameter space which satisfies the design specifications. This solution is impractical from the manufacturing point of view since there are manufacturing tolerances and uncertainties (i.e., statistical fluctuations of the circuit designable parameters about their nominal values).These fluctuations may reduce parametric yield which is the percentage of the circuits in the manufacturing processes satisfying the design specifications. In this paper, reducing the undesirable effects of these statistical fluctuations is achieved.

Yield optimization of microwave circuits is obstructed by the high expense of electromagnetic simulations required in the yield estimation process in addition to the absence of any gradient information. In this paper, Space Mapping (SM) surrogates using the generalized space mapping (GSM) technology is integrated with a derivative-free trust region optimization method (NEWUOA: new unconstrained optimization algorithm). Moreover, a variance reduction technique is used in the sampling process. Implementing VRT reduces the number of samples required to estimate acceptable yield values. In this paper the Latin Hypercube sampling (LHS) is employed in the sampling process. The technique exploits a SM-developed surrogate using minimax optimizer in the yield maximization process. Our novel approach is illustrated by practical examples showing its efficiency.

***Keywords*:** CAD; yield optimization; EM simulation; NEWUOA; space mapping; statistical analysis.

1. Introduction

Yield optimization is of a great interest for the microwave CAD society. Yield optimization is a design problem which looks for nominal values of circuit parameters that minimizes the undesirable effects on circuit performance. Statistical yield optimization depends on performing multitude of circuit simulations. Yet using the full–wave EM simulator in the design centering process, in the traditional way, incorporates high computational effort which may be prohibitive.

In general, circuit parameters are subject to known but unavoidable statistical fluctuations inherent to the manufacturing process used, due to environmental effects during operations, or due to manufacturing tolerances and uncertainties. Generally, the traditional microwave circuit optimization trend is to find a minimax design of the circuit. However, in the design space the location of the minimax point may be closer to the boundaries of the feasible region (a region in the design space where the design specifications are satisfied). This may cause the circuit performance to violate the design specifications. Yield optimization problem seeks for nominal values of circuit parameters which minimizes the undesirable

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

effects of these statistical fluctuations. Namely, it seeks for the design center which maximizes the probability of satisfying the design specification.

In manufacturing, the probability of satisfying the design specification by an outcome is called the parametric yield which can be approximated by the percentage of the manufactured circuits satisfying the design specifications (Bandler *et. al.,* 1993 and 2002). The process seeking the nominal values of designable parameters that maximize the parametric yield is usually known as yield maximization or design centering.

Yield optimization approaches can be classified as statistical and geometrical. Geometrical approaches approximate the feasible region using a convex body, e.g., a hyperellipsoid. Then the center of this body is considered as the design center. One of the popular methods used for yield optimization is the ellipsoidal technique (Abdel-Malek *et. al.,* 2006). These approaches have fast convergence for convex and small dimensional problems. Statistical approaches, on the other hand, optimize the yield function in a straightforward way, regardless the size of the problem or its convexity (Hassan *et. al.,* 2006 and 2011). Hybrid methods, combining both approaches, may also be used for solving such problems.

The statistical yield optimization process has some permanent special difficulties. One of these difficulties is the cost of finding a yield value for a given nominal design parameters. Yield values are estimated through a statistical analysis algorithm, e.g., Monte Carlo analysis. EM simulations required for the generated samples represent an obstacle in applying a traditional optimization process. Another difficulty in statistical yield optimization is the need for a derivative-free optimizer due to the absence of any exact or approximate gradient information about the yield. Any method can be used to approximate the gradient of the yield highly increases the computational overhead.

The new proposed approach integrates three strategies to overcome these difficulties. First, a variance reduction technique (VRT) is used in the sampling process (McKay *et. al.,* 1979). Implementing VRT reduces the number of samples required to estimate acceptable yield values. In this paper the Latin Hypercube sampling (LHS) (Stein, 1987) is employed in the sampling process. Second, a derivative free trust region method is utilized in the optimization process. The general framework of these methods is to iteratively build and update a smooth model to locally approximate the expensive yield function in a trust region around the current solution (Hassan *et. al.,* 2006). Then the model is optimized over the trust region. In this paper, NEWUOA algorithm developed by Powell (Powell, 2006 and 2007) which employs quadratic models to approximate objective function is used. Third, the space mapping (SM) technique (Bandler *et al.*, 2004; Koziel *et. al.*, 2006; Cheng *et. al.*, 2010) is used to reduce the simulation computational effort. SM utilizes fast physically-based coarse models to replace the time-consuming, computationally-expensive full-wave fine models so that a great reduction in computations and simulations can be achieved. In our proposed approach, we employ the generalized space mapping (GSM) (Koziel *et. al.*, 2006) to construct a space-mapped surrogate model, based on the fast coarse model, matching the fine model response with high accuracy. Therefore, in our proposed approach presented in this paper, we integrate NEWUOA algorithm (Powell, 2006) with the GSM technique (Koziel *et. al.*, 2006) in addition to LHS technique (McKay *et. al.,* 1979) to get our yield maximization technique for microwave circuits. The technique makes use of a surrogate, a SM-based, developed by any other circuit optimizer, e.g., minimax optimizer, in the yield maximization process. This technique may be of great benefits to whom interested in modeling and constructing surrogates. Practical examples to demonstrate the new approach are included.

2. The New Approach for Yield Optimization

2.1. INTRODUCTION

The desired performance of a microwave circuit is described by some performance specifications set by the designer. These specifications or constraints define a region in the designable parameter space known as the feasible region, which can be defined as

$$F = \{ \boldsymbol{x} \in \mathbb{R}^n \mid f_i(\boldsymbol{R}_f(\boldsymbol{x})) \leq b_i \quad i = 1,2,\dots,m \} \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^{\mathrm{n}}$ is the vector of the design parameters, $\boldsymbol{R}_f$: $\mathbb{R}^{\mathrm{n}} \rightarrow \mathbb{R}^{\mathrm{m}}$ is the fine model response vector, n is the number of design parameters, m is the number of constraints, f_{i} is the i-th performance function, and b_{i} is corresponding specification bound. Every $\boldsymbol{x} \in F$ is considered as an acceptable microwave circuit. In general, circuit parameters are subject to known but unavoidable statistical fluctuations. These arise, for example, during manufacturing or due models uncertainties. To simulate these statistical fluctuations circuit parameters are assumed to be random variables with a joint probability density function $p(\boldsymbol{x}, \boldsymbol{x}_0)$, where $\boldsymbol{x}_0 \in \mathbb{R}^{\mathrm{n}}$ is the vector of nominal parameter values. Therefore, the yield Y can be defined as the probability of satisfying the design specifications (Hassan *et. al.,* 2006)

$$Y(\boldsymbol{x}_0) = \int_{\boldsymbol{R}_f} p(\boldsymbol{x}, \boldsymbol{x}_0) \; d\boldsymbol{x} \tag{2}$$

Hence, the yield optimization problem is formulated as:

$$\max_{\boldsymbol{x}_0} Y(\boldsymbol{x}_0) \tag{3}$$

The yield integral in Eq. (2) cannot be exactly evaluated as the feasible region F is not explicitly defined. Instead, the yield values at a nominal vector $\boldsymbol{x}_0$ can be estimated by generating a set of samples in the parameter space using the pdf of design parameters. Let $\boldsymbol{x}^j, j$ =1, 2,…, K be the generated samples around the nominal parameter vector $\boldsymbol{x_o}$. For each sample, an acceptance index $I_a : \mathbb{R}^{\boldsymbol{n}} \rightarrow \mathbb{R}$, defined by

$$I_a(\boldsymbol{x}^j) = \begin{cases} 1 & if \; \boldsymbol{x}^j \in F \\ 0 & if \; \boldsymbol{x}^j \notin F \end{cases} \tag{4}$$

where F , defined by Eq. (1), is evaluated. If K is sufficiently large, the yield function at the nominal parameter value $\boldsymbol{x_o}$ can be estimated as

$$Y(x_o) \approx \frac{1}{K} \sum_{j=1}^{K} I_a\left(x^j\right) \tag{5}$$

Hence, the yield value at a nominal vector $\boldsymbol{x}_0$ can be estimated by generating a set of samples in the parameter space using the pdf of design parameters. The circuit is simulated for each sampling point. The percentage of acceptable circuits gives an estimate of the yield value at $\boldsymbol{x}_0$ (Bandler *et. al.*, 2002).

In our new approach, the yield function in Eq. (2) will be optimized using the NEWUOA algorithm. Also, a lot of the computational efforts with the high-fidelity "fine" model EM simulations can be alleviated by creating an accurate, efficient and simpler SM surrogate model instead of the fine one. Here, the GSM (Koziel *et. al.*, 2006) is employed where the feasible region can be approximated by

$$F = \{ \boldsymbol{x} \in \mathbb{R}^n \mid f_i(\boldsymbol{R}_s(\boldsymbol{x})) \leq b_i \quad i = 1,2,\dots,m\} \tag{6}$$

where $\boldsymbol{R}_s : \mathbb{R}^{\boldsymbol{n}} \rightarrow \mathbb{R}^{\boldsymbol{m}}$ is the GSM surrogate model response vector. In the proposed technique, a well-constructed SM surrogate will be used by the NEWUOA algorithm to optimize the yield.

2.2. GENERALIZED SPACE MAPPING (GSM) SURROGATE (Koziel *et. al.*, 2006)

The GSM-based surrogate model is constructed, using a computationally fast coarse mode with input and output mappings, in the form

$$\boldsymbol{R}_s^i(\boldsymbol{x}) = \boldsymbol{A}^i . \boldsymbol{R}_c \left(\boldsymbol{B}^i . \boldsymbol{x} + \boldsymbol{c}^i\right) + \boldsymbol{d}^i + \boldsymbol{E}^i . \left(\boldsymbol{x} - \boldsymbol{x}_0^i\right) \tag{7}$$

where $\boldsymbol{x}_0^i$ is the current optimal parameter vector, $\boldsymbol{A}^{\mathrm{i}} \in \mathrm{M_{m\times m}}$ is a diagonal matrix, $\boldsymbol{R}_{\mathrm{c}} : \mathrm{X_c} \rightarrow \mathbb{R}^{\mathrm{m}}$ is the coarse model response vector, $\boldsymbol{B}^{\mathrm{i}} \in \mathrm{M_{n\times n}}$, $\boldsymbol{c}^{\mathrm{i}} \in \mathrm{M_{n\times 1}}$, and $\boldsymbol{d}^{\mathrm{i}} \in \mathrm{M_{m\times 1}}$ is given by

$$\boldsymbol{d}^i = \boldsymbol{R}_f \left(\boldsymbol{x}_0^i\right) - \boldsymbol{A}^i . \boldsymbol{R}_c \left(\boldsymbol{B}^i . \boldsymbol{x}_0^i + \boldsymbol{c}^i\right) \tag{8}$$

where $\boldsymbol{R}_{\mathrm{f}} : \mathrm{X_f} \rightarrow \mathbb{R}^{\mathrm{m}}$ is the fine model response vector, and $\boldsymbol{E}^{\mathrm{i}} \in \mathrm{M_{m\times n}}$ is given by

$$\boldsymbol{E}^i = \boldsymbol{J}_f \left(\boldsymbol{x}_0^i\right) - \boldsymbol{A}^i . \boldsymbol{J}_c \left(\boldsymbol{B}^i . \boldsymbol{x}_0^i + \boldsymbol{c}^i\right) . \boldsymbol{B}^i \tag{9}$$

where $\boldsymbol{J}_f : \mathrm{X_f} \rightarrow \mathbb{R}^{\mathrm{m\times n}}$ and $\boldsymbol{J}_c : \mathrm{X_c} \rightarrow \mathbb{R}^{\mathrm{m\times n}}$ are the Jacobian matrices of the fine and coarse model responses w.r.t. the corresponding points, respectively. The mapping parameters $\boldsymbol{A}^i, \boldsymbol{B}^i, \boldsymbol{c}^i$ are obtained by the parameter extraction (PE) optimization process given by

$$\left(\boldsymbol{A}^i, \boldsymbol{B}^i, \boldsymbol{c}^i\right) = \arg\min_{\boldsymbol{A},\boldsymbol{B},\boldsymbol{c}} \quad \boldsymbol{e}^i \left(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{c}\right) \tag{10}$$

where $\boldsymbol{e}^i$ represents the response deviation residual of the surrogate from the fine model and is given by

$$\begin{aligned} \boldsymbol{e}^{(k)}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{c}) = & \sum_{i=0}^{k} w_i \left\| \boldsymbol{R}_f\left(\boldsymbol{x}^{(i)}\right) - \boldsymbol{A}. \boldsymbol{R}_c\left(\boldsymbol{B}. \boldsymbol{x}^{(i)} + \boldsymbol{c}\right)\right\| \\ & + \sum_{i=0}^{k} v_i \left\| \boldsymbol{J}_f\left(\boldsymbol{x}^{(i)}\right) - \boldsymbol{A}. \boldsymbol{J}_c\left(\boldsymbol{B}. \boldsymbol{x}^{(i)} + \boldsymbol{c}\right). \boldsymbol{B}\right\| \end{aligned} \tag{11}$$

where the coefficients w_i and v_i are chosen according to the nature of the design problem. Hence, the surrogate model is updated by the new mapping parameters obtained from Eq. (11), then the new surrogate model obtained is optimized according to the design specification required to Eq. (7), and so, the process will be terminated when the final design reached the user-predefined accuracy.

2.3. NEWUOA ALGORITHM (Powell, 2006)

Assume that $\boldsymbol{R}_{\mathrm{S}}^{\mathrm{i}}(\boldsymbol{x})$ is the surrogate model obtained in the $(i-1)$th SM iteration with a design center (nominal parameter vector) $\boldsymbol{x}_0^{\mathrm{i}}$. The estimated yield value, *Y*, is submitted to the optimizer by a subroutine that employs LHS technique. The computationally expensive yield function is locally approximated around a current iterate $\boldsymbol{x}_0^{\mathrm{i}}$ by a much cheaper quadratic model $M(\boldsymbol{x})$ in the form

$$M(\boldsymbol{x}) = a + \boldsymbol{g}^T\left(\boldsymbol{x} - \boldsymbol{x}_0^i\right) + \frac{1}{2}\left(\boldsymbol{x} - \boldsymbol{x}_0^i\right)^T \boldsymbol{H}\left(\boldsymbol{x} - \boldsymbol{x}_0^i\right) \tag{12}$$

where $a \in \mathbb{R}$, $\boldsymbol{g} \in \mathbb{R}^{\mathrm{n}}$ and the symmetric matrix $\boldsymbol{H} \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}}$ are the unknown parameters of $M(\boldsymbol{x})$. These parameters are determined by interpolating the yield at $m = 2n + 1$ points using the interpolating conditions

$$M(\boldsymbol{x}_k) = Y(\boldsymbol{x}_k), \qquad k = 1,2,\ldots,m \tag{13}$$

where $\boldsymbol{x}_k$ are the interpolation points. The freedom in *M* is taken up by minimizing the Frobenius norm of the change in the Hessian matrix $\boldsymbol{H}$, i.e., $\left\|\boldsymbol{H}^{new} - \boldsymbol{H}^{old}\right\|_{\mathrm{F}}$. Let *t* is the index of the interpolation point which provides the best yield value such that: $\mathrm{Y}(\boldsymbol{x}_t) \geq \mathrm{Y}(\boldsymbol{x}_{\mathrm{k}}),\ k = 1,2,\ldots,m$.

The model in Eq. (12) is then maximized, instead of the yield function, over a current trust region. Once *M* has been identified, a step $\boldsymbol{s}$ has to be added to $\boldsymbol{x}_t$ by solving the following trust region sub-problem

$$\max_{\boldsymbol{s}} M(\boldsymbol{x}_t + \boldsymbol{s}), \ \text{subject to } \|\boldsymbol{s}\| \leq \Delta \tag{14}$$

where Δ is the current trust region radius. This radius Δ is revised based on the agreement between the model and the actual function at the new point $(\boldsymbol{x}_t + \boldsymbol{s})$ measured by the

$$r = \frac{f(\boldsymbol{x}_t) - f(\boldsymbol{x}_t + \boldsymbol{s})}{M(\boldsymbol{x}_t) - M(\boldsymbol{x}_t + \boldsymbol{s})} \tag{15}$$

Thus, the trust region radius Δ has a lower bound ρ in the interval $[\rho_{\mathrm{fin}}, \rho_{\mathrm{ini}}]$. The parameter ρ is utilized to maintain enough distance between the interpolation points (Powell, 2006), where ρ_{ini} and ρ_{fin} are user-defined initial and final radii respectively. The value of Δ is revised nearly at each iteration. Let Δ_{old} and Δ_{new} be the old and the new values of Δ. The choice of Δ_{new} depends on the ratio in Eq. (15), and the Euclidean length of the step $\boldsymbol{s}$.

The algorithm is terminated when the trust region radius reaches the lower bound ρ_{fin} that fixes the final accuracy required in the parameters (Powell, 2007).

2.4. VARIANCE REDUCTION TECHNIQUE (VRT)

The main objective in space filling methods is to spread the sample data points as evenly as possible around the interior design space. This is in contrast to classical methods where samples are focused at the center and the extreme limits of design space. These methods are model independent, i.e., they assume no information is available about the functional behavior of the "true" model. Space filling methods include Monte Carlo sampling design, Latin hypercube sampling design and Orthogonal Array design.

2.4.1. *Monte Carlo Sampling (MCS) design*

The basic Monte Carlo sampling design was developed in (Metropolis and Ulam, 1949). This method simply selects the samples by generating pseudo random numbers within the range interval of each design variable. Figure 1.a) shows an example of the basic Monte Carlo sampling design for two design variables. The main drawback of the basic Monte Carlo sampling design is that the generated samples may leave large regions of the design space not explored.

2.4.2. *Latin Hypercube Sampling (LHS) design*

This design provides a more accurate estimate of the mean value of function than the Monte Carlo method (Giunta *et. al.*, 2003). The LHS involves dividing the design pace into equiprobable sub-regions. Then N samples are selected such that all sub-regions are sampled. Figure 1.b) shows an example of LHS sampling for two design variables with four samples. Latin Hypercube method requires a fewer number of runs to achieve the same level of confidence than the number required for the Monte Carlo approach because the entire probability range will be explored.

An advantage of LHS design is that there is no restriction on the total number of samples. This makes the LHS design suitable for constructing models for computationally expensive function evaluations with small number of samples.

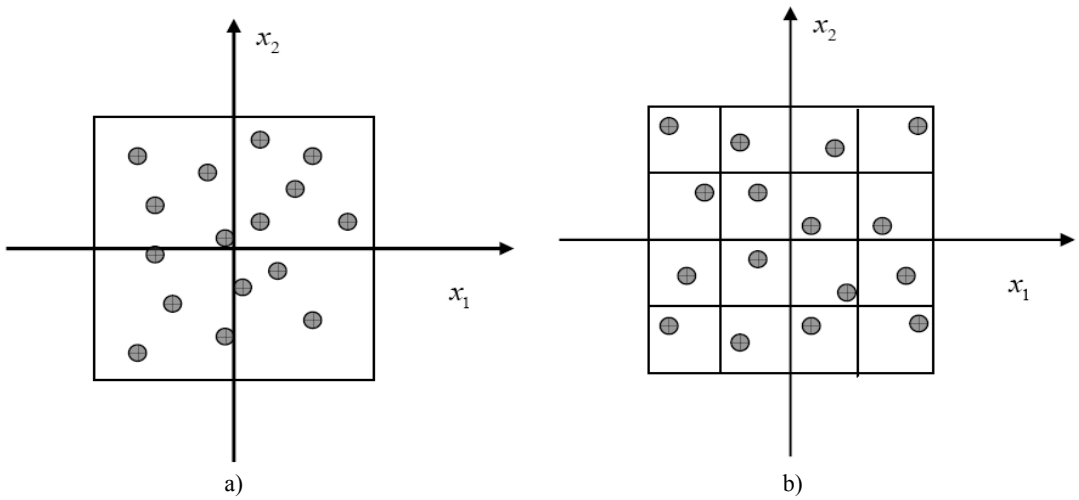


Figure 1. a) Basic Monte Carlo sampling and b) Latin Hypercube sampling in two dimensions

3. Methodology

A novel technique is being proposed here to solve the yield maximization problem. The technique is based on the availability of a pre-constructed surrogate, e.g., utilizing the SM technology along with a minimax optimizer. Then, the design center optimal solution can be obtained by employing NEWUOA algorithm with the pre-constructed surrogate. This technique is favorable for the microwave designers who are using SW packages to model and optimize the microwave circuit.

In our technique, we assume the pre-constructed surrogate is obtained by employing the GSM technology in addition to a minimax optimizer. This surrogate may suffer from locality (Bakr *et. al*., 2000 and 2001; Bandler *et. al*., 2004). However, any well-constructed surrogate can be efficiently utilized with NEWUOA algorithm in the yield optimization process.

4. Algorithm

Step 1) Set $k = 0$, ρ_{ini}, ρ_{fin}, $\boldsymbol{x}_{initial}$ to NEWUOA, $\boldsymbol{G}$ (a given covariance matrix), ε (SM termination criterion) and δ (a yield termination criterion).

Step 2) Evaluate the yield Y of the fine model at $\boldsymbol{x}_{initial}$ using LHS technique with the $\boldsymbol{G}$ matrix.

Step 3) For the same initial point $\boldsymbol{x}_{initial}$, apply parameter extraction (PE) to get $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{c}$, $\boldsymbol{d}$, $\boldsymbol{E}$ using the least square method utilizing Eq.(8)–(11) and construct the surrogate $\boldsymbol{R}_s^1(\boldsymbol{x})$ given by Eq. (7). Set $k = 1$.

Step 4) Apply minimax optimization to the surrogate in Step 3) until getting a better point $\boldsymbol{x}^k$.

Step 5) Reconstruct the surrogate $\boldsymbol{R}_s^{k+1}(\boldsymbol{x})$ at the point $\boldsymbol{x}^k$, by PE, then use minimax optimizer to obtain $\boldsymbol{x}^{k+1}$.

Step 6) If $\left\|\boldsymbol{x}^{k+1} - \boldsymbol{x}^k\right\| \leq \varepsilon$, go to Step7), else set $k = k + 1$ and go to Step5)

Step 7) Apply NEWUOA to the surrogate in Step 5) until getting a better point $\boldsymbol{x}_0^k$ with yield value Y $\left(\boldsymbol{x}_0^k\right)$.

Step 8) Estimate the yield of the fine model at the final point obtained $\boldsymbol{x}_0^k$ using LHS, then stop the algorithm.

5. Examples

5.1. Two-Section Capacitively Loaded Transmission Line Impedance Transformer

We consider the two-section transmission line impedance transformer (Bakr *et. al*., 2000). It is a two-dimensional problem, where the feasible region is constrained by the magnitude of the reflection coefficients $|S_{11}|$ at 11 frequency points {0.5, 0.6 ,…, 1.5 GHz} and defined by

$f_i\left(\boldsymbol{R}_f(\boldsymbol{x})\right) = |S_{11}(\boldsymbol{x}, \omega_i)| \leq 0.5 \quad 0.5 \text{ GHz} \leq \omega_i \leq 1.5 \text{ GHz}$

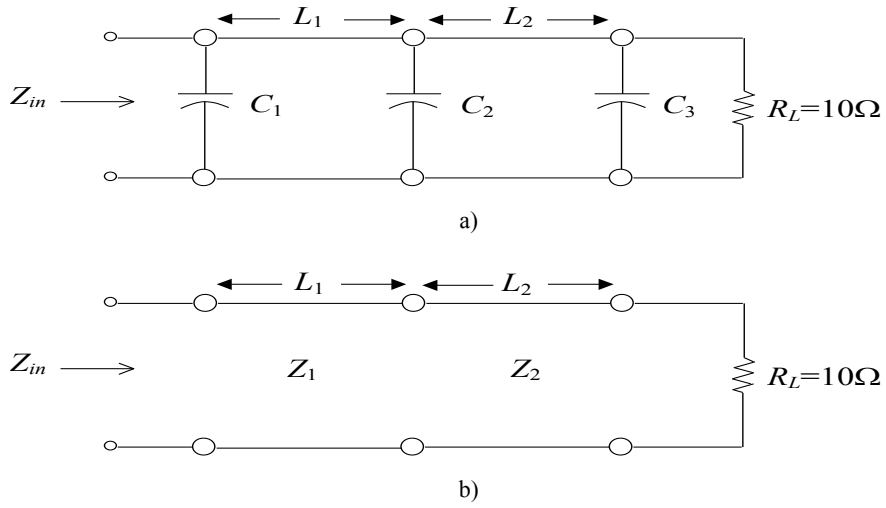


Figure 2. The fine a) and coarse b) models for capacitively loaded impedance transformer.

The coarse model is an ideal two-section transmission line (TL), where the fine model is a capacitively loaded TL with capacitors C = 10 pF as shown in figure 2. Design parameters are the characteristic impedances $\boldsymbol{x} = [Z_1 \quad Z_2]^T$ while the normalized lengths L_1 and L_2, w.r.t. the quarter-wave length at the center frequency 1GHz, are kept fixed at $[L_1 \;\; L_2]^T = [0.9333 \;\; 0.8001]^T$. The proposed technique is applied, starting from $[1.7 \;\; 4.36]^T$ ohm, which is an infeasible point w.r.t. the fine model parameter space. The yield values are estimated via Latin Hypercube sampling method with 200 sample points assuming normally distributed parameters. The results are shown in Table I for the independent parameters case.

Table I. Results for the two-section TL transformer with normally distributed independent parameters

Parameter Spreads	Initial Yield		Final Yield	
	surrogate	fine	surrogate	fine
(0.2,0.4)	11.5%	10.5%	47.5%	47.0%
(0.2,0.4)/2	2%	1.5%	85.5%	85.0%
(0.2,0.4)/3	0%	0.5%	97%	96.5%

5.2. SEVEN-SECTION CAPACITIVELY LOADED TRANSMISSION LINE IMPEDANCE TRANSFORMER

The seven-section transmission line (TL) capacitively loaded impedance transformer example is described in (Bakr *et. al*., 2001; Bandler *et. al*., 2004). We consider a "coarse" model as an ideal seven-section TL, where the "fine" model is a capacitively-loaded TL with capacitors $C_1 = \cdots \quad 0.025$ pF, see figure3. Design parameters are $\boldsymbol{x}_f = \left[L_1 \; L_2 \; L_3 \; L_4 \; L_5 \; L_6 \; L_7\right]^T$, which are the normalized lengths w.r.t. the quarter-wave length L_q at the center frequency 4.35 GHz. Design specifications are $f_i\left(\boldsymbol{R}_f(\boldsymbol{x})\right) = |S_{11}(\boldsymbol{x}, \omega_i)| \leq 0.07 \quad 1\text{GHz} \leq \omega_i \leq 7.7\text{GHz},$ with 68 points per frequency sweep.

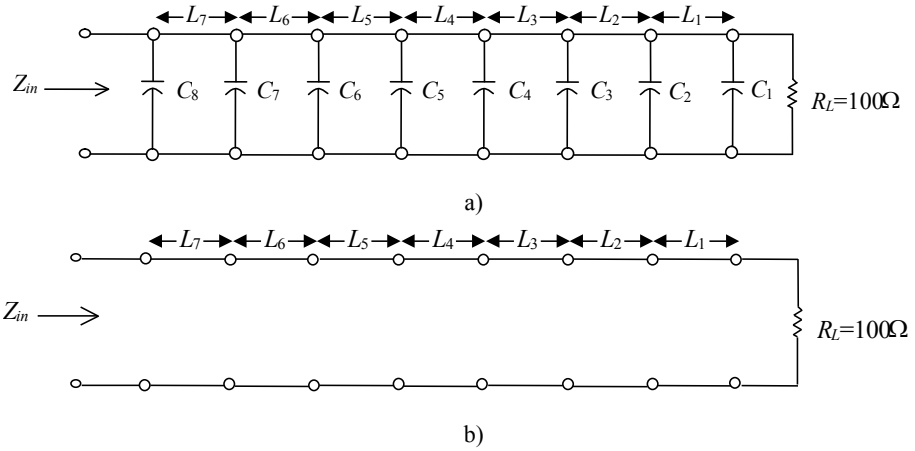


Figure 3. Seven-section capacitively-loaded impedance transformer: a) fine model, b) coarse model.

An initial infeasible point $[0.892\ 0.993\ 0.989\ 0.981\ 0.996\ 0.99\ 0.891]^T$ is considered. The yield values are estimated via LHS method with 200 sample points assuming normally distributed parameters. The results assuming correlated parameters with covariance matrices $\boldsymbol{G}$ and $\boldsymbol{G}/9$ (Abdel-Malek *et. al*., 2006) are shown in Table II, where

$$\boldsymbol{G}=\begin{bmatrix} 0.193 & -0.055 & -0.002 & 0.003 & -0.003 & 0.014 & -0.15 \\ -0.055 & 0.194 & -0.077 & -0.009 & 0.035 & -0.149 & 0.11 \\ -0.002 & -0.077 & 0.145 & -0.031 & -0.115 & 0.091 & -0.02 \\ 0.003 & -0.009 & -0.031 & 0.046 & 0.01 & -0.018 & 0.02 \\ -0.003 & 0.035 & -0.115 & 0.01 & 0.155 & -0.059 & -0.01 \\ 0.014 & -0.149 & 0.091 & -0.018 & -0.059 & 0.239 & -0.1 \\ -0.15 & 0.11 & -0.02 & 0.02 & -0.01 & -0.1 & 0.38 \end{bmatrix}$$

Table II. Results for the seven-section TL transformer with normally distributed correlated parameters

Covariance matrix	Initial Yield		Final Yield	
	surrogate	fine	surrogate	fine
$\boldsymbol{G}$	4%	3.4%	47.2%	20.4%
$\boldsymbol{G}/9$	0%	0.2%	100%	99.2%

5.3. SIX-SECTION H-PLANE WAVEGUIDE FILTER

We apply our technique to the six-section H-plane waveguide filter (Matthaei *et. al.*, 1964). A waveguide with a width 3.485 cm is used. The propagation mode is TE_{10} with a cutoff frequency of 4.3 GHz. The six-waveguide sections are separated by seven H-plane septa (as shown in figure 4) which have a finite thickness of 0.6223 mm. The design parameters $\boldsymbol{x}$ are the three waveguide-section lengths L_1, L_2 and L_3 and the septa widths W_1, W_2, W_3 and W_4. The feasible region is constrained by the magnitude of the reflection

coefficients $|S_{11}|$ at 44 frequency points {5.2, 5.3,…, 9.5 GHz}. These magnitudes have to satisfy the upper and lower design specifications given by

$$f_i\left(\boldsymbol{R}_f(\boldsymbol{x})\right)=\begin{cases}|S_{11}(\boldsymbol{x},\omega_i)|\geq 0.85 & \omega_i \leq 5.2 \text{ GHz}\\ |S_{11}(\boldsymbol{x},\omega_i)|\leq 0.16 & 5.4 \text{ GHz} \leq \omega_i \leq 9.0 \text{ GHz}\\ |S_{11}(\boldsymbol{x},\omega_i)|\geq 0.5 & \omega_i \geq 9.5 \text{ GHz}\end{cases}$$

An empirical coarse model with lumped inductances and dispersive transmission line sections is utilized. The formulas due to Marcuvitz (Marcuvitz, 1951), for the inductive susceptances corresponding to the H-plane septa, are simplified. The model is implemented and simulated in the Matlab environment. The simulation of the fine model is performed using High Frequency Structure Simulator (HFSS). We take $[L_1\ L_2\ L_3\ W_1\ W_2\ W_3\ W_4]^T$ = [16.1614 16.1899 16.6975 13.3376 12.0823 11.7456 11.5212]^T mm as a starting point. The yield values are estimated via LHS method with 200 sample points. The initial and final yield results assuming independent parameters at the surrogate and fine model are shown in Table III.

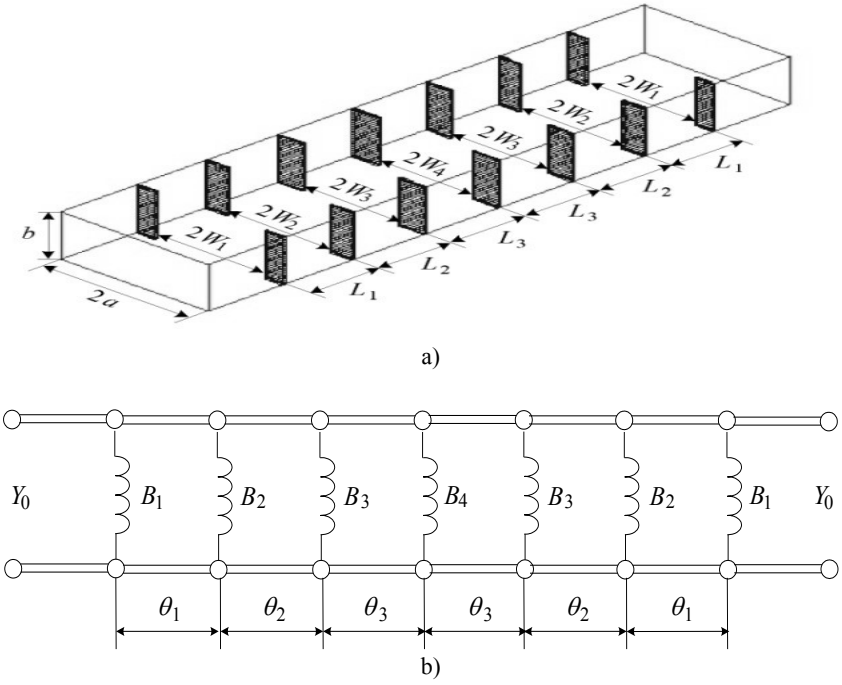


Figure 4. The six-section H-plane waveguide filter: a) the 3D view. b) the equivalent empirical circuit model.

Table III. Results for the six-section H-plane waveguide filter assuming independent parameters

Parameter Spreads*	Initial Yield		Final Yield	
	surrogate	fine	surrogate	fine
$3\sigma/2$	26.2%	27%	90%	90%
σ	23.5%	26%	99.5%	99%

σ = [0.3814 0.3833 0.3988 0.3163 0.2925 0.2794 0.2758]10^{-4} m.

6. Conclusion

In this paper, a new efficient statistical technique for yield maximization is introduced. The technique is successfully applied to different microwave circuits, and all the considered circuit examples show significant increases in the yield values. The proposed techniques combine a non-derivative trust region optimization algorithm, NEWUOA, with surrogates constructed by the SM technique in addition to LHS. This integration dramatically reduces the number of fine model simulations required in the design centering process. Hence, the consumed time and effort are also decreased.

References

Abdel-Malek, H.L., A.S.O. Hassan, E. A. Soliman, and S.A. Dakroury. The ellipsoidal technique for design centering of microwave circuits exploiting space-mapping interpolating surrogates, *IEEE Trans. Microwave Theory Tech.*, 54(10):3731–3738, 2006.

Bakr, M.H., J.W. Bandler, K. Madsen, J.E. Rayas-Sánchez and J. Søndergaard. Space mapping optimization of microwave circuits exploiting surrogate models," *IEEE Trans. Microwave Theory and Tech*., 48: 2297–2306, 2000.

Bakr, M.H., J. W. Bandler, K. Madsen and J. Søndergaard. An introduction to the space mapping technique, *Optimization and Engineering*, 2:369–384, 2001.

Bandler, J.W., R.M. Biernacki, S.H. Chen, P.A.Grobelny, and S.Ye. Yield-driven electromagnetic optimization via multilevel multidimensional models, *IEEE Trans. Microwave Theory Tech.*, 41:2269–2278, 1993.

Bandler, J.W., Q.S. Cheng, S.A. Dakroury, A.S. Mohamed, M.H. Bakr, K. Madsen, and J. Søndergaard. Space mapping: the state of the art, *IEEE Trans. Microwave Theory Tech*., 52(1): 337–361, 2004.

Bandler, J.W., J.E. Rayas-Sánchez and Q.J. Zhang. Yield-driven electromagnetic optimization via space mapping-based neuromodels, *Int. J. RF and Microwave CAE*, (12): 79–89, 2002..

Cheng, Q. S., J. W. Bandler, S. Koziel,M. H. Bakr and S. Ogurtsov. The state of the art of microwave CAD: EM-based optimization and modeling, *Inc. Int. J. RF and Microwave CAE* (20): 475–491, 2010.

Giunta, A. A., Jr. Wojtkiewicz, S. F. and M. S. Eldred. Overview of modern design of experiments methods for computational simulations, *Proceedings of the 41st AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, AIAA-2003-0649, 2003.

Hassan, A.S.O., H.L. Abdel-Malek and A.A. Rabie. Non-derivative design centering algorithm using trust region optimization and variance reduction, *Eng. Optim*., 38: 37–51, 2006.

Hassan, A.S.O., A.S. Mohamed, and A.Y. El-Sharabasy. Statistical microwave circuit optimization via a non-derivative trust region approach and space mapping surrogates , *IEEE MTT-S Int. Microw. Symp. Dig*., Baltimore, USA, 2011.

Koziel, S. ,J. W. Bandler, and K. Madsen. Space mapping optimization algorithms for engineering design, *IEEE MTT-S Int. Microw. Symp. Dig.*, San Francisco, CA, 2006.

Koziel, S., J.W. Bandler, and K.Madsen. A space-mapping framework for engineering optimization: Theory and implementation, *IEEE Trans. Microwave Theory Tech.*, 54(10):3721–3730, 2006.

Marcuvitz, N., *Waveguide Handbook*, 1st ed. New York, NY: McGraw-Hill, p.221, 1951.

Matthaei, G.L., L. Young and E.M.T. Jones, *Microwave Filters, Impedance-Matching Networks, and Coupling Structures*, 1st ed. New York, NY: McGraw-Hill, 1964.

McKay, M. D., R. J. Beckman and W. J. Conover A comparison of three methods for selecting values of input variables in analysis of output from a computer code, *Technometrics*, 21(2):239–245, 1979.

Metropolis, N. and S. Ulam. The Monte Carlo Method, *Journal of the American Statistical Association*, 44(247): 335-341, 1949.

Powell, M.J.D. The NEWUOA software for unconstrained optimization without derivatives, in *Large-Scale Nonlinear Optimization*, Springer, 255–297, 2006.

Powell, M.J.D. Development of NEWUOA for minimization without derivatives, *IMA Journal of numerical Analysis,* 28(4): 649–664, 2007.

Stein, M. L. Large sample properties of simulations using Latin Hypercube sampling, *Technometrics*, 29: 143–151, 1987.

On the Tolerance Approach to Possibilistic Nonlinear Regression over Interval Data

Milan Hladík and Michal Černý
First author: Charles University, Faculty of Mathematics and Physics, Department of Applied Mathematics, Malostranské nám. 25, 11800, Prague, Czech Republic; Department of Econometrics, University of Economics, Prague; Náměstí Winstona Churchilla 4, CZ13067 Prague 3, Czech Republic, milan.hladik@matfyz.cz.
Second author: Department of Econometrics, University of Economics, Prague; Náměstí Winstona Churchilla 4, CZ13067 Prague 3, Czech Republic, cernym@vse.cz.

Abstract. We study the tolerance-based approach to possibilistic nonlinear regression models with interval data. We provide a method for determination of interval regression parameters of the model for the crisp input – interval output case and for the interval input – interval output case. We define two classes of nonlinear regression models for which efficient algorithms exist. We illustrate the theory by examples.

Keywords: interval regression, nonlinear regression, possibilistic regression, tolerance quotient

1. Introduction

Nonlinear regression is a fundamental tool in data analysis. In this text we address the tolerance approach to possibilistic nonlinear regression, which is a natural generalization of the same concept used in linear regression (Hladík and Černý, 2010; Hladík and Černý, 2011). Possibilistic interval regression was pioneered by (Tanaka, Uejima and Asai, 1987; Tanaka, 1987) in the field of linear regression, and later extended to nonlinear regression (Hao, 2009; Hwang, Hong and Seok, 2006; Jen, Chuang and Su, 2003; Lingras and Butz, 2011; Xu, Luo, Xu and Zhang, 2009), mostly by means of support vector machine. Possibilistic regression was successfully applied in economic forecasting (Lin, Hung and Wu, 2011), system identification (Kaneyoshi, Tanaka, Kamei and Furuta, 1990), speech learning systems (Liu, 2009), or analytic hierarchy process (Entani and Inuiguchi, 2010), among others.

In this text we propose a very general framework for classification of nonlinear regression models which allows us to construct algorithms for computing their possibilistic interval regression parameters. This is useful in particular in case when data to be modeled are of interval nature.

The paper is organized as follows. First we review the notion of possibilistic regression, used in linear regression, and provide a formulation suitable for nonlinear regression models with both crisp input and crisp output data. Then we continue to models involving interval data; in particular, we distinguish crisp input – interval output models and interval input – interval output models. In Section 1.1 we review some examples of nonlinear regression functions widely used applications and in Section 2 we provide a certain general classification framework for nonlinear regression functions. Finally, in Section 3, we state the main problem and design the tolerance-based procedure for computation of interval regression parameters for the classes of functions defined in Section 2.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

In this text, a *regression function* is simply a continuous function

$$f(x, \theta) = f(x_1, \dots, x_m; \theta_1, \dots, \theta_p),$$

where $x_1, \dots, x_m$ are *data variables* and $\theta_1, \dots, \theta_p$ are *parameters*.

Given a dataset of n observations of the form

$$(y_i; x_{i*}) = (y_i; x_{i1}, \dots, x_{im}), \quad i = 1, \dots, n,$$

the *possibilistic regression* seeks for interval parameters

$$[\underline{\theta}, \overline{\theta}] = [\underline{\theta_1}, \overline{\theta_1}], \dots, [\underline{\theta_p}, \overline{\theta_p}] \tag{1}$$

such that

$$\forall i \in \{1, \dots, n\} \; \exists \theta \in [\underline{\theta}, \overline{\theta}] \text{ s.t. } f(x_{i*}, \theta) = y_i.$$

If the condition is satisfied for a given i, we say that *i-th observation is covered*. (Outside data analysis, the problem is sometimes referred to simply as "covering problem" or "envelopment problem".)

Of course, the problem of "finding the interval parameters (1)" must be stated more precisely. Usually we want the find the intervals *as narrow as possible* such that all observations are covered. But that is a multi-criteria optimization problem. The *tolerance approach* is a natural (but not the only possible) method of conversion of the multi-criteria problem to a single-criterion problem. Details of the approach will be discussed in Section 3. As shown in (Hladík and Černý, 2011) (where possibilistic linear regression is studied), the approach has several interesting theoretical properties.

Before we turn into theory, we review some examples of nonlinear regression functions useful in various fields of science and engineering.

1.1. Examples of useful nonlinear regression functions

We sketch only a few examples; more on applications of nonlinear models can be found in (Ratkowski, 1988; Seber and Wild, 2003).

Example. Nonlinear regression functions are often solutions to differential equations describing processes in physics, chemistry or biology. An interesting example is the class of growth curves describing the growth of populations. The Richard's Growth Equation ((Seber and Wild, 2003), p. 332) has a solution

$$f(x; \theta_1, \theta_2, \theta_3, \theta_4) = \theta_1 \cdot (1 + (\theta_4 - 1)e^{-\theta_2(x-\theta_3)})^{1/(1-\theta_4)},$$

which is known as the Richard's Curve. This model has interesting special cases: setting $\theta_4 = 2$ we get the logistic curve

$$f(x; \theta_1, \theta_2, \theta_3) = \frac{\theta_1}{1 + e^{-\theta_2(x-\theta_3)}}, \tag{2}$$

the limit case $\theta_4 \to 1$ yields the Gompertz Curve

$$f(x; \theta_1, \theta_2, \theta_3) = \theta_1 \cdot e^{-e^{-\theta_2(x-\theta_3)}} \tag{3}$$

and the special case with $\theta_4 = 0$ is the model of growth with exponential slow-down

$$f(x; \theta_1, \theta_2, \theta_3) = \theta_1 \cdot (1 - e^{-\theta_2(x-\theta_3)}). \tag{4}$$

Example. The submodel of the Logistic Model (2)

$$f(x;\theta_2,\theta_3)=\frac{1}{1+e^{-\theta_2(x-\theta_3)}},\quad [\theta_1\equiv 1] \tag{5}$$

(and similarly with other growth models) is often used when y is interpreted as the probability of an event, where the probability grows with x. For example, we can consider $x=$ pressure and $y=$ probability that the device or material under investigation will be damaged by the pressure. Another example: y can measure the response of a patient to the quantity of drug x.

Example. Another interesting example is the problem of estimation of the degree of polynomial:

$$f(x;\theta_1,\theta_2,\theta_3)=\theta_1+\theta_2x+\theta_3x^{\theta_4}. \tag{6}$$

Example. Berry's Model (Berry, 1967) describes the crop yield as a function of density of planting (or, equivalently, the area available to each plant). Let x_1 be the distance between plants in a row and x_2 the distance between rows of plants. Berry used a model of the form

$$f(x;\theta_1,\theta_2,\theta_3,\theta_4)=\left(\theta_1+\theta_2\left(\frac{1}{x_1}+\frac{1}{x_2}\right)+\frac{\theta_3}{x_1x_2}\right)^{-\theta_4}. \tag{7}$$

Example. In physics, the simple oscillation model is important:

$$y=\theta_1e^{-\theta_2x}\cos(\theta_3x). \tag{8}$$

Example. An important class of nonlinear models is the class of models involving a structural change. The basic example is continuous connection of two lines:

$$f(x;\theta_1,\theta_2,\theta_3,\theta_4)=\begin{cases}\theta_1+\theta_2x & \text{for } x\le\theta_4,\\ \theta_1+\theta_4(\theta_2-\theta_3)+\theta_3x & \text{for } x>\theta_4.\end{cases} \tag{9}$$

1.2. NOTATION

Let A^* denote the closure of a set $A\subseteq\mathbb{R}^n$. Given an interval $\boldsymbol{a}\subseteq\mathbb{R}\cup\{\pm\infty\}$, the numbers $\underline{a}$ and $\overline{a}$ denote its lower and upper boundary points, respectively, and a^c and a^Δ denote its center and radius, respectively. That is, $\boldsymbol{a}^*=[\underline{a},\overline{a}]=[a^c-a^\Delta,a^c+a^\Delta]$. Given a function f and a set A, the symbol $f(A)$ denotes the image of A under f. In particular, $f(\boldsymbol{a})$ stands for the image of an interval $\boldsymbol{a}$.

2. Classes of Nonlinear Regression Models

In order to solve interval nonlinear regression problems, we have to know how to compute image of a function over intervals. Formally, we consider a class of functions equipped by algorithms for determining their images.

Definition 1. *Let*

$$(f_1,f_1^L,f_1^U),\ldots,(f_K,f_K^L,f_K^U) \tag{10}$$

be a set of triples, where for all $k=1,\ldots,K$:

- $f_k : \mathbb{R} \to \mathbb{R}$ *is a continuous function,*
- $f_k^L(\underline{x}, \overline{x})$ *is an algorithm computing* $\underline{f_k([\underline{x}, \overline{x}])^*}$*,*
- $f_k^U(\underline{x}, \overline{x})$ *is an algorithm computing* $\overline{f_k([\underline{x}, \overline{x}])^*}$*.*

(a) The set (10) *is called* ***basis****.*

(b) The smallest class of functions (of any number of variables)

- *containing constants and the identity function,*
- *containing the functions* $f_1, \ldots, f_k$ *and* $+, -, \times, \div$ *and*
- *being closed under composition and restriction of domain*

is called ***functional universum*** *and is denoted as* $\mathcal{U}$*.*

Determining the image of a function is a fundamental problem of interval analysis (Moore, Kearfott and Cloud, 2009) and by far not trivial. Indeed, only for certain functions we can do it efficiently.

Arithmetic expressions

Interval arithmetic is defined naturally as an image of values over interval domains (Moore, Kearfott and Cloud, 2009). Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be real intervals. Then

$$\begin{aligned}
\boldsymbol{a} + \boldsymbol{b} &= [\underline{a} + \underline{b}, \overline{a} + \overline{b}], \\
\boldsymbol{a} - \boldsymbol{b} &= [\underline{a} - \overline{b}, \overline{a} - \underline{b}], \\
\boldsymbol{a} \cdot \boldsymbol{b} &= [\min(\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}), \max(\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b})], \\
\boldsymbol{a} \div \boldsymbol{b} &= [\min(\underline{a} \div \underline{b}, \underline{a} \div \overline{b}, \overline{a} \div \underline{b}, \overline{a} \div \overline{b}), \max(\underline{a} \div \underline{b}, \underline{a} \div \overline{b}, \overline{a} \div \underline{b}, \overline{a} \div \overline{b})].
\end{aligned}$$

Given an aritmetic expression $\mathcal{E}$ for a function f, we can evaluate $\mathcal{E}$ by using interval arithmetic. As long as each interval parameter appears at most once in $\mathcal{E}$, then the result equals the image of f. Otherwise, we obtain only an enclosure (a superset) of the image. For example, consider the function

$$f(x, y) = xy - 2x$$

with $x \in [1, 2]$ and $y \in [3, 4]$. Evaluating by interval arithemtic leads to the enclosure

$$f(\boldsymbol{x}, \boldsymbol{y}) \subseteq [1, 2][3, 4] - 2[1, 2] = [-3, 6].$$

However, f can be expressed in other ways. In the form

$$f(x, y) = x(y - 2)$$

each parameter appears just once, so the interval evaluation is exact, i.e.

$$f(\boldsymbol{x}, \boldsymbol{y}) = [1, 2]([3, 4] - 2) = [1, 4].$$

BASIC FUNCTIONS

For some basic function, computing their images over intervals is a simple task. For instance, $\exp(\boldsymbol{x}) = [\exp(\underline{x}), \exp(\overline{x})]$ as the exponential is increasing. Similarly for the functions $\log, \arctan, \ldots$ Some non-monotone functions are tractable, too, e.g. $\sin, \cos, x^n, \ldots$ Polynomials, however, are hard to evaluate exactly in general.

MONOTONICITY

The assumption that each interval parameter should appear at most once in a given expression is restrictive. Moreover, f may be expressed by other basic functions and operations than $+, -, \times, \div$. A strong tool in such a case is to utilize monotonicity. If $f(x) = f(x_1, \ldots, x_m)$ is monotone with respect to the kth parameter x_k, then we are able to get rid of one interval domain. Provided $f(x)$ is non-decreasing at x_k, $\overline{f(\boldsymbol{x})}$ is attained at $\overline{x}_k$, and $\underline{f(\boldsymbol{x})}$ is attained at $\underline{x}_k$. Similarly for the non-increasing case. In this way, the problem of determining $f(\boldsymbol{x})$ is reduced to the problem of determining $\overline{f(\boldsymbol{x})}$ and $\underline{f(\boldsymbol{x})}$ with smaller number of intervals. Hopefully, the sub-problems are of the previous types so that we can calculate the exact values.

For example, let

$$f(x, y) = \frac{x^2 + 6 - y}{y^2}$$

with $x \in [-1, 2]$ and $y \in [1, 2]$. The function is decreasing with respect to y on the interval domains, so in order to compute the lower limit $\underline{f(\boldsymbol{x}, \boldsymbol{y})}$ we fix $y = \overline{y}$, and calculate

$$f(\boldsymbol{x}, \overline{y}) = \frac{\boldsymbol{x}^2 + 5 - \overline{y}}{\overline{y}^2} = \frac{[-1, 2]^2 + 6 - 2}{2^2} = [1, 2].$$

Analogously, to compute the upper limit $\overline{f(\boldsymbol{x}, \boldsymbol{y})}$ we fix $y = \underline{y}$, and calculate

$$f(\boldsymbol{x}, \underline{y}) = \frac{\boldsymbol{x}^2 + 6 - \underline{y}}{\underline{y}^2} = \frac{[-1, 2]^2 + 6 - 1}{1^2} = [5, 9].$$

Putting together, we conclude $f(\boldsymbol{x}, \boldsymbol{y}) = [1, 9]$.

2.1. CLASS OF SUITABLE FUNCTIONS

From the above considerations it is clear that for a well-defined class of function we can determine their images over intervals effectively. For the purpose of interval nonlinear regression, we define the following classes.

Definition 2. *We define the classes of functions $\mathcal{A}$ and $\mathcal{B}$ as follows. Let*

$$f(x, \theta) = f(x_1, \ldots, x_m; \theta_1, \ldots, \theta_p) \in \mathcal{U}.$$

The function $f(x, \theta)$ belongs to the class $\mathcal{A}$ if the function can be analytically expressed such that

– *the expression consists of operations* $+, -, \times, \div$ *and basic functions that are easy to evaluate over intervals, and*

– f *is monotone with respect to the parameters* $\theta_1, \ldots, \theta_p$ *that appear more than once in the expression.*

The function $f(x, \theta)$ *belongs to the class* $\mathcal{B}$ *if the function can be analytically expressed such that*

– *the expression consists of operations* $+, -, \times, \div$ *and basic functions that are easy to evaluate over intervals, and*

– f *is monotone with respect to* $\theta_1, \ldots, \theta_p$, $x_1, \ldots, x_m$ *that appear more than once in the expression.*

We say that the function f *is of* ***type*** $\mathcal{A}$ *and* $\mathcal{B}$*, respectively.*

The significance of the Definition will be clarified in Section 3.1. If a nonlinear regression model is of type $\mathcal{A}$, then there is an efficient method for the tolerance-based possibilistic regression in the crisp input – interval output model. Observe that the crisp input – crisp output model is a special case, hence we also get an algorithmic method for this case as well. If a nonlinear regression model is of type $\mathcal{B}$, then there is an efficient method for the tolerance-based possibilistic regression in the interval input – interval output model.

2.2. EXAMPLES

Whenever we find out that a particular nonlinear regression function belongs to some of the classes $\mathcal{A}, \mathcal{B}$, we know that the nonlinear tolerance approach can be applied to it.

Consider the basis (10) containing $\exp$ and $\ln$. In both cases, the corresponding algorithms f^L and f^U are trivial.

Example. The growth curves (2), (3) and (4) are $\mathcal{B}$-functions.

Example. The regression function (6) can be written in the form

$$y = \theta_1 + \theta_2 x + \theta_3 e^{\theta_4 \ln x},$$

and hence it is an $\mathcal{A}$-type function. If we admit the logarithmic transformation of data $x' := \ln x$, we arrive at the form

$$y = \theta_1 + \theta_2 e^{x'} + \theta_3 e^{\theta_4 x'},$$

and in this form it is an $\mathcal{A}$-type function even if we do not have $\ln$ in the basis. But note that in general, the results of the tolerance-based approach procedure of estimation of interval regression parameters is invariant neither under reparametrization of the model nor under data transformations.

Example. This example shows that a suitable reparametrization of a nonlinear regression function might improve its classification. The Logistic Function is often written in the form

$$f(x; \theta_2, \theta_3) = \frac{e^{\theta_2 (x - \theta_3)}}{1 + e^{\theta_2 (x - \theta_3)}},$$

where both the variable x and the parameters θ_2, θ_3 occur twice, and hence in this form it is not an $\mathcal{A}$-type function; but its equivalent form (5) is an $\mathcal{B}$-type function.

Example. Berry's Model (7) is an $\mathcal{A}$-function.

Example. The model (9) is an $\mathcal{A}$-type function e. g. under the restriction $\theta_1\theta_3 \geq 0$.

3. The possibilistic interval nonlinear regression

3.1. Setting the problem

Here, we formulate the interval regression problem. Tha aim is to find interval domains for parameters such that all observations are covered by some realization of intervals:

> Find the minimal interval domains for parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p)$ such that for every $i = 1, \ldots, n$ one has

$$\boldsymbol{y}_i \subseteq f(\boldsymbol{x}_{i*}, \boldsymbol{\theta}). \tag{11}$$

This formulation covers also problems with crisp input or crisp output as special cases. The minimality means that there is no other interval vector $\boldsymbol{\theta}' \subsetneqq \boldsymbol{\theta}$ satisfying (11). Nevertheless, there may exist other interval vector, or typically many of them, that is also minimal with respect to inclusion. So there are many degrees of freedom which minimal solution to consider. To obtain good interval parameters, the following properties should be more or less satisfied:

- The radii of interval parameters, $\theta_1^\Delta, \ldots, \theta_p^\Delta$ are balanced. It is undesirable when some interval is very narrow, or even crisp, while another is very wide.
- The interval parameters follow the so called *central tendency*. That is, their centers more or less fit the data with respect to traditional goodness-of-fit measures.
- The method is not much sensitive to outliers.

In order to fulfill these requirements for interval linear regression models, the authors proposed in (Hladík and Černý, 2010; Hladík and Černý, 2011) a two level method. In the first step, we calculate crisp estimation $\theta^c = (\theta_1^c, \ldots, \theta_p^c)$ to the nonlinear regression model. In the second step, we minimally extend the parameters to intervals such that they cover all observations. This basic idea is usable for nonlinear regression as well; we do it in the next section.

3.2. Methodology

As indicated in the previous section, we calculate interval parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p)$ in two steps:

(a) Compute the centers $\theta^c = (\theta_1^c, \ldots, \theta_p^c)$;

(b) Compute the radii $\theta^\Delta = (\theta_1^\Delta, \ldots, \theta_p^\Delta)$.

Centers are determined by any traditional method for nonlinear regression. In case of interval input or output, we take the centers of the intervals. Thus, we have a standard nonlinear regression model with crisp data and

can apply any method to compute θ^c. This makes our approach flexible, since it doesn't rely on one concrete algorithm. Next, the property on central tendency is fulfilled, too.

In the second step, we calculate the radii θ^Δ. In order that the resulting interval parameters are balanced with respect to their radii, we introduce *tolerance rates* as a non-negative vector $c^\Delta = (c_1^\Delta, \dots, c_p^\Delta)$. The radii of interval parameters are then considered in the form $\theta^\Delta = \delta c^\Delta$, or

$$(\theta_1^\Delta, \dots, \theta_p^\Delta) = (\delta c_1^\Delta, \dots, \delta c_p^\Delta),$$

where $\delta > 0$ is the unknown *tolerance quotient*. The aim is to determine the minimal tolerance quotient such that the corresponding interval parameters cover all observations. A tolerance quotient satisfying the coverage condition is called feasible.

The tolerance rates are usually set up as $c^\Delta = |\theta^c|$ or $c^\Delta = (1, \dots, 1)$. The former corresponds to relative perturbations, while the latter force all interval parameters to have the same width. If the kth interval parameter is desired to be crisp, so it suffices to put $c_k^\Delta = 0$.

Now, all we need is to compute the minimal feasible tolerance quotient $\delta > 0$. We employ the bisection method. Denote $\boldsymbol{\theta}^\delta := [\theta^c - \delta c^\Delta, \theta^c + \delta c^\Delta]$ the form of the resulting interval parameters. The basic algorithmic scheme is as follows:

1. Put $\delta = 1$ and loop the following command for a given number of iterations.
2. If $\boldsymbol{y}_i \subseteq f(\boldsymbol{x}_{i*}, \boldsymbol{\theta}^\delta)$ for every $i = 1, \dots, n$, then decrease δ. Otherwise, increase δ.

Denote by δ^* the return value of δ. Notice that provided the amount of decrease and increase of δ is halved, the iterations converge exponentially fast to the optimum. Thus, for practical purposes, 5 to 15 iterations are usually enough to provide us with a sufficiently accurate approximation. For a model of type $\mathcal{A}$ or $\mathcal{B}$, the evaluation of the image $f(\boldsymbol{x}_{i*}, \boldsymbol{\theta}^\delta)$ is fast, therefore, the overall time complexity of the algorithm is mild.

If the last iteration was the decrease of δ, we increase δ^* correspondingly in order to obtain a feasible δ^*. However, it may still happen that δ^* is not feasible. We indicate it easily be observing that δ was never decreased in the run of the algorithm. This situation happens rarely, but cannot be excluded. For example, consider the Gompertz Curve of the form

$$f(x, \theta_1) = e^{-e^{x-\theta_1}}.$$

It is easily seen that for any θ_1, $f((-\infty, \infty), \theta_1) = (0, 1)$. If our dataset contains, say, a point $(x = 0, y = 1)$, that point cannot be covered, which implies that the algorithm tends to increase δ up to infinity. In general, from the algorithmic point of view, the problem whether a given point can be covered (with a possibly huge value of δ), is undecidable; hence we cannot do anything else than terminating the algorithm when the value of δ exceeds limits in which the value δ has reasonable interpretation for the regression model under consideration.

3.3. Properties of the model

Here we only sketch some properties of the model, which have been investigated in (Hladík and Černý, 2011) in the case of linear regression models.

– The method is flexible by utilizing any traditional method for the parameteric centers.

– If the input model is of type $\mathcal{A}$ or $\mathcal{B}$, then the optimal tolerance quotient, and thus the minimal interval parameters, are computed efficiently with a given precision.

– The interval parameters have balanced widths, proportional to the apriori given rates.

– Outliers can be handled.

Concerning outliers, they can be managed in many ways, depending on the purposes of decision maker. For instance, the method is easily adapted to the model, in which only a fraction, say 90%, of observations should be covered. Another possibility is to calculate the tolerance quotient δ^* such that the corresponding interval parameters cover e.g. 80% observations, and then consider as outliers all observations that are not covered by the tolerance quotient $1.1\delta^*$.

3.4. EXAMPLES

Example 1. Assume that we measure reliability of a material (y) as a function of time (x) for which the material is exposed to unfavorable conditions (such as unfavourable temperature or pressure). Of course it can be expected that the longer the exposition is, the higher level of disruption. Assume that the level of disruption is measured on a discrete scale $0, \ldots, 10$, where 0 means "no damage", 1 means "very mild damage", ..., and 10 means "totally damaged". Assume further that the values of y are determined by experts (say, by visual inspection of constructions where the material has been used). Due to a certain subjectivity of experts, it is appropriate to consider the grade $y \in \{1, \ldots, 9\}$ as an interval, say of the form

$$[\underline{y}, \overline{y}] = [y - 0.5, y + 0.5] \tag{12}$$

We model the dependence of y on x using the Gompertz curve

$$y = 10e^{-e^{-\theta_1(x-\theta_2)}}, \tag{13}$$

where θ_1 measures slope of the curve (that is, the speed of worsening of the condition of the material) and θ_2 measures the shift of the curve. The shift measures whether the process of wearing of the material starts earlier or later.

Assume that we have data from Table I. Using nonlinear least squares on the data $(x_1, y_1), \ldots, (x_{30}, y_{30})$, we fit

$$\widehat{\theta}_1 = 0.795, \quad \widehat{\theta}_2 = 4.887. \tag{14}$$

This curve describes "average" behavior of the material with respect to x.

Now we would like to extend the estimated crisp values $\theta_1^c = \widehat{\theta}_1$ and $\theta_2^c = \widehat{\theta}_2$ to interval values covering all observations, taking into account the fact that it is more appropriate to handle an observation y as an interval (12) rather than a fixed value.

We observe that the points $y \in \{0, 10\}$ can never be covered with the Gompertz curve. We take the following step. We divide data into three categories:

– A: material is unaffected by the unfavorable conditions;

– B: phase of wearing;

– C: phase of total weariness.

We assume that the phase B starts when the first mild defect is encountered (i.e. the first time with $y \geq 1$) and that the phase C starts when the first total damage is encountered (i.e. first time with $y = 10$). The division of data is also shown in Table I.

The main purpose of the Gompertz curve is modeling the dynamics of the wearing process, which corresponds to the phase B. Hence it makes sense to take into account only B-data and apply the tolerance method to them. (Observe that the data point $(x_{24}, [\underline{y}_{24}, \overline{y}_{24}])$, being a C-point, need not be covered.)

As a first example, we set $c^\Delta = \left(\begin{smallmatrix} 0.795 \\ 4.887 \end{smallmatrix}\right)$ (i.e., relative tolerances). We arrive at the value

$$\delta^* = 0.183.$$

Hence we conclude that it suffices to perturb the values $\widehat{\theta}_1, \widehat{\theta}_2$ by no more that 18.3% in order all intervals be covered. We can roughly say that "the truth" is covered by the intervals $[(1-0.183)\cdot 0.795, (1+0.183)\cdot 0.795]$ and $[(1-0.183)\cdot 4.887, (1+0.183)\cdot 4.887]$ for θ_1 and θ_2, respectively. The resulting data enclosure is plotted in Figure 1 with a dotted line. We can also say that the "pessimistic scenario" for the speed of weariness (measured by θ_1) is $(1+0.183)\cdot 0.795 = 0.94$ and that the "pessimistic scenario" for the shift is $(1-0.183)\cdot 4.887 = 3.99$.

As a second example, we set $c^\Delta = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ (i.e., absolute tolerances). We arrive at the value

$$\delta^* = 0.360.$$

Now the data are covered by the intervals $[0.795 - 0.36, 0.795 + 0.36]$ and $[4.887 - 0.36, 4.887 + 0.36]$ for θ_1 and θ_2, respectively. The resulting data enclosure is plotted in Figure 1 with a dashed-dotted line.

As a third example, we set $c^\Delta = \left(\begin{smallmatrix} 0 \\ 4.887 \end{smallmatrix}\right)$. This models the situation that the speed of worsening is kept constant and we can perturb only the shift θ_2 to cover the data. (Hence we seek for an interval for θ_2 only.) We arrive at the value

$$\delta^* = 0.254.$$

Now the data are covered by the interval $[(1-0.36)\cdot 4.887, (1+0.36)\cdot 4.887]$ for θ_2. The resulting data enclosure is plotted in Figure 1 with a dashed line. Now we can say: if we know that the speed of wearing is $\theta_1 = 0.795$, then the pessimistic scenario for θ_2 is $(1-0.36)\cdot 4.887 = 3.13$.

Example 2. In Example 1 we used the fact that the Gompertz function (13) is $\mathcal{A}$-type function. Using the fact that it is also the $\mathcal{B}$-type function, we can extend the example to the case where x-data are of interval nature. This corresponds to the situation that we do not know exactly the times in which the measurements were made. Again we use the data from Table I and for each of the observations we assume that its x-value is an interval

$$[x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}].$$

We set the values $(\theta_1^c, \theta_2^c) = (\widehat{\theta}_1, \widehat{\theta}_2)$ from (14). Using the tolerance method for covering the B-phase data, we arrive at the results

– $c^\Delta = \left(\begin{smallmatrix} 0.795 \\ 4.887 \end{smallmatrix}\right)$: $\delta^* = 0.250$,

– $c^\Delta = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$: $\delta^* = 0.61$,

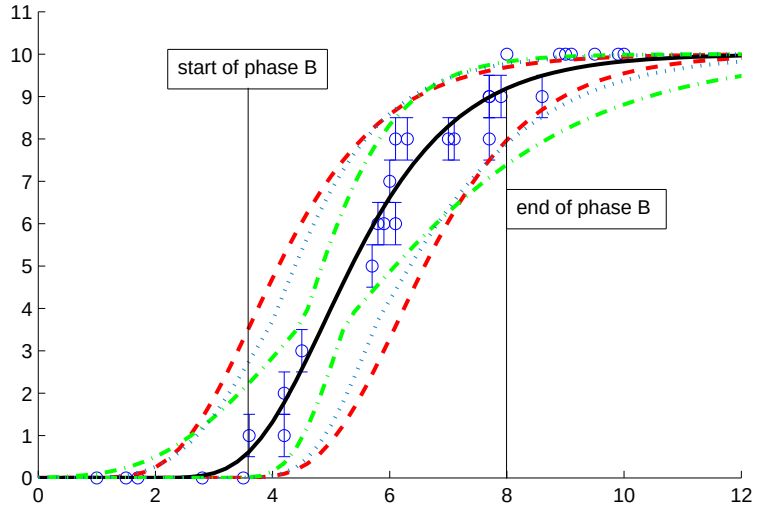


Figure 1. Data from Table I, the crisp model (13) with $\widehat{\theta}_1 = 0.795$ and $\widehat{\theta}_2 = 4.887$ and the enclosures with $c^{\Delta} = (0.79, 4.887)^T$ (relative tolerances, dotted), $c^{\Delta} = (1, 1)^T$ (absolute tolerances, dashed-dotted) and $c^{\Delta} = (0, 4.887)^T$ (only perturbation of θ_2 allowed, dashed).

Table I. Source data for the Example.

phase	i	x_i	y_i	$\underline{y}_i$	$\overline{y}_i$	phase	i	x_i	y_i	$\underline{y}_i$	$\overline{y}_i$
A	1	1.0	0	—	—	B	16	6.3	8	7.5	8.5
A	2	1.5	0	—	—	B	17	7.0	8	7.5	8.5
A	3	1.7	0	—	—	B	18	7.1	8	7.5	8.5
A	4	2.8	0	—	—	B	19	7.7	9	8.5	9.5
A	5	3.5	0	—	—	B	20	7.7	8	7.5	8.5
B	6	3.6	1	0.5	1.5	B	21	7.7	9	8.5	9.5
B	7	4.2	2	1.5	2.5	B	22	7.9	9	8.5	9.5
B	8	4.2	1	0.5	1.5	C	23	8.0	10	—	—
B	9	4.5	3	2.5	3.5	C	24	8.6	9	8.5	9.5
B	10	5.7	5	4.5	5.5	C	25	8.9	10	—	—
B	11	5.8	6	5.5	6.5	C	26	9.0	10	—	—
B	12	5.9	6	5.5	6.5	C	27	9.1	10	—	—
B	13	6.0	7	6.5	7.5	C	28	9.5	10	—	—
B	14	6.1	8	7.5	8.5	C	29	9.9	10	—	—
B	15	6.1	6	5.5	6.5	C	30	10.0	10	—	—

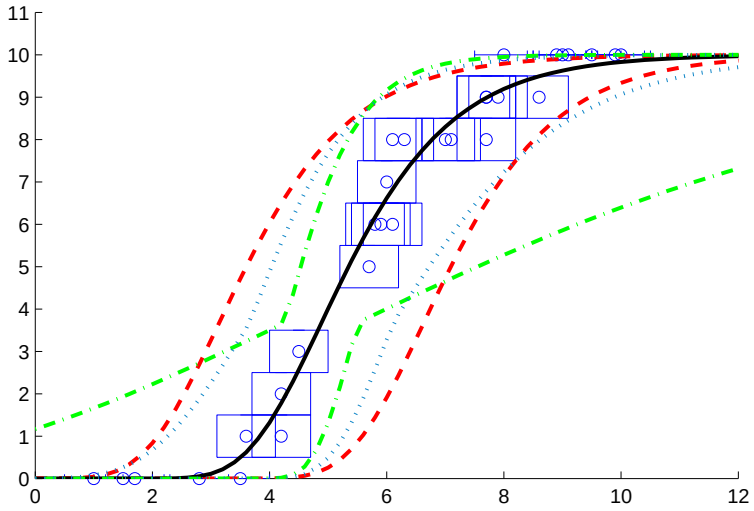


Figure 2. Data from Table I with interval-valued x_i's in the form $\boldsymbol{x}_i = [x_i - \frac{1}{2}, x_i + \frac{1}{2}]$, the crisp model (13) with $\widehat{\theta}_1 = 0.795$ and $\widehat{\theta}_2 = 4.887$ and the enclosures with $c^\Delta = (0.79, 4.887)^T$ (relative tolerances, dotted), $c^\Delta = (1, 1)^T$ (absolute tolerances, dashed-dotted) and $c^\Delta = (0, 4.887)^T$ (only perturbation of θ_2 allowed, dashed).

– $c^\Delta = \left(\begin{smallmatrix} 0 \\ 4.887 \end{smallmatrix}\right)$: $\delta^* = 0.357$,

with the resulting enclosures depicted in Figure 2. Recall that the data point $([x_{24} - \frac{1}{2}, x_{24} + \frac{1}{2}], [\underline{y}_{24}, \overline{y}_{24}])$, being a C-point, need not be covered.

4. Conclusions

In this text we extended the tolerance-based approach, originally designed for possibilistic linear regression, for a particular class of nonlinear regression models. The method provides a covering of either crisp or interval data of the model and for that class of models it can be computed by an efficient algorithm (provided that the algorithms f^U and f^L for the basic functions are efficient). For the class of non-$\mathcal{A}$-type models, the method provides only lower bound on the optimal tolerance rate δ^* in general. The interesting question for further research is whether and under which conditions the method could be adapted for a wider of nonlinear models to yield the optimal value.

Acknowledgements

The authors were partially supported by the Czech Science Foundation Grant P403/12/1947.

References

Berry, G. A Mathematical Model Relating Plant Yield with Arrangement for Regularly Spaced Crops. *Biometics* 23: 505–515, 1967.

Entani, T. and M. Inuiguchi. Group Decisions in Interval AHP Based on Interval Regression Analysis. In Huynh, V.-N. et al. (editors), *Integrated Uncertainty Management and Applications*, vol. 68 of Advances in Soft Computing, 269–280, Springer, 2010.

Hao, P.-Y. Interval Regression Analysis Using Support Vector Machines. *Fuzzy Sets and Systems* 160 (17): 2466–2485, 2009.

Hladík, M. and M. Černý. New Approach to Interval Linear Regression. In Kasımbeyli et al. (eds.), *24th Mini-EURO Conference on Continuous Optimization and Information-Based Technologies in the Financial Sector MEC EurOPT 2010. Selected Papers,* 197–171, Technika, Vilnius, 2010.

Hladík, M. and M. Černý. Interval Regression by Tolerance Analysis Approach. *Fuzzy Sets and Systems* (in press), 2011, DOI 10.1016/j.fss.2011.07.009.

Hwang, C., D. H. Hong and K. Ha Seok. Support Vector Interval Regression Machine for Crisp Input and Output Data. *Fuzzy Sets and Systems* 157 (8): 1114–1125, 2006.

Jen J.-T., C.-C. Chuang and S.-F. Su. Support Vector Interval Regression Networks for Interval Regression Analysis. *Fuzzy Sets and Systems* 138 (2): 283–300, 2003.

Kaneyoshi, M., H. Tanaka, M. Kamei and H. Furuta. New System Identification Technique Using Fuzzy Regression Analysis. In *Proceedings of the First International Symposium on Uncertainty Modeling and Analysis*, 258–533, Baltimore, USA, 1990.

Lin, K.-P., K.-C. Hung and M.-C. Wu. Long-term Business Cycle Forecasting Using Intuitionistic Fuzzy Least-squares Suppert Vector Regression. In *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2495–2499, Taipei, Taiwan, 2011.

Lingras, P. and C. Butz. Conservative and Aggressive Rough SVR Modeling. *Theoretical Computer Science* 412 (42): 5885–5901, 2011.

Liu, P. Study on a Speech Learning Approach Based on Interval Support Vector Regression. In *Proceedings of 41th International Conference on Computer Science & Education*, 1009-1012, Nanning, China, 2009.

Moore, R. E., R. B. Kearfott and M. J. Cloud. *Introduction to Interval Analysis.* SIAM, Philadelphia, 2009.

Ratkowski, D. A. *Handbook of Nonlinear Regression Models.* Marcel Dekker, 1988.

Seber, G. A. F. and C. J. Wild. *Nonlinear Regression.* Wiley, 2003.

Tanaka, H. Fuzzy Data Analysis by Possibilistic Linear Models. *Fuzzy Sets and Systems* 24 (3): 363–375, 1987.

Tanaka, H., S. Uejima and K. Asai. Linear Regression Analysis with Fuzzy Model. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-12 (6): 903–907, 1982.

Xu, S.-Q., Q.-Y. Luo, Xu G.-H. and L. Zhang. Assymetrical Interval Regression using Extended ε-SVM with Robust Algorithm. *Fuzzy Sets and Systems* 160 (7): 988–1002, 2009.

Accuracy of Concrete Creep Predictions Based on Extrapolation of Short-Time Data

Milan Jirásek and Svatopluk Dobruský
Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, 166 29 Prague, Czech Republic

Abstract. The paper evaluates the accuracy of predictions obtained with various creep models after updating of their parameters based on short-time data. The models considered in this comparative study include relatively simple formulae recommended by major design codes (ACI, *fib*) as well as more sophisticated models developed by researchers (B3, GL). Appropriate error measures are defined and two updating strategies are examined. Predictions of the models are checked against basic creep data from a comprehensive database. The dependence of the error on the load duration after which the update is performed is described. Finally, preliminary conclusions and recommendations regarding the choice of the model and updating strategy are formulated.

Keywords: concrete, creep, updating

1. Introduction

Concrete exhibits creep already at low stress levels and normal temperatures, and long-time measurements on laboratory samples as well as on concrete structures indicate that the growth of strain at constant stress continues even after many decades, see e.g. (Brooks, 2005) and (Bažant et al., 2010). Problems with excessive deflections caused by creep have been reported for many large-span prestressed concrete bridges, and comparative numerical simulations based on design codes and advanced models have revealed the essential role played by a good predictive creep model (Bažant et al., 2010; Bažant et al., 2011). Unfortunately, empirical formulae for determination of creep model parameters based exclusively on the fundamental properties (such as compressive strength, concrete mix composition, size and shape of the member, environmental conditions and curing) have a very limited accuracy and often lead to gross errors. It is essential to update the model parameters based on laboratory tests or measurements of the early response of the real structure.

The present paper compares updated predictions obtained with the following creep models:

- the ACI model, recommended by the permanent committee TC 209 of the American Concrete Institute;
- the *fib* model, recommended in the Model Code 2010 of the International Federation for Structural Concrete;
- the B3 model, developed at Northwestern University by Bažant and coworkers;
- the GL2000 model, developed at the University of Ottawa by Gardner and coworkers.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

Two updating strategies are considered: the standard one is based on simple least-square fitting, while its modification introduces weight factors that emphasize the influence of measured data directly preceding the interval of extrapolation. Absolute and relative error measures are defined and accuracy of the initial "blind" as well as updated predictions is checked against data from a comprehensive creep database (Bažant and Li, 2008), with focus on the dependence of the error on the load duration after which the update is performed.

2. Creep Models

Concrete creep at low and moderate stress levels is usually handled within the framework of aging linear viscoelasticity. Based on the superposition principle, the strain history corresponding to a given continuous and differentiable stress history can be computed using the integral formula

$$\varepsilon(t) = \int_{t_0}^{t} J(t,t')\dot{\sigma}(t')\,\mathrm{d}t' \tag{1}$$

For discontinuous stress histories, additional terms that reflect the influence of stress jumps can be added. In formula (1), ε is the strain, σ is the stress, t_0 is the time at the onset of loading, t is the current time, and J is the compliance function that can be determined from a creep test at constant stress. The value of $J(t,t')$ corresponds to the strain at time t in a creep test started at time t', divided by the stress level at which the test takes place. The time is measured from the set of concrete, i.e., it corresponds to the age of the material. For a non-aging viscoelastic material, the compliance function would depend only on the elapsed time $t-t'$, but for an aging material such as concrete it depends on t and t' separately. In all the models presented here, the values of time variables are supposed to be substituted in days.

For simplicity, the stress-strain relation (1) has been presented in a scalar format, valid for uniaxial stress. In a general extension to multiaxial stress, the volumetric and deviatoric parts of the response could be treated separately. In the absence of more precise data, it is usually assumed that all compliance coefficients are proportional to one single compliance function, which is equivalent to the assumption that the Poisson ratio remains constant and is not affected by creep.

2.1. ACI 209 Model

The model recommended by the permanent committee TC 209 Creep and Shrinkage in Concrete of the American Concrete Institute (ACI) was first adopted in 1971. Its most recent version, labeled as 209R-92, was published in 1992 (ACI, 1992) and again reapproved in 2008. The compliance function has the form

$$J(t,t') = \frac{1}{E_C}\sqrt{b+\frac{a}{t'}}\left[1+\frac{2.35\gamma}{(t')^m}\frac{(t-t')^{0.6}}{10+(t-t')^{0.6}}\right] \tag{2}$$

with time variables t and t' substituted in days. Parameter E_C is the conventional elastic modulus of concrete, measured at age 28 days. Parameters a, b and m depend on the type of cement and type of curing. For moist-cured concrete and cement of type I, their recommended values are $a = 4$, $b = 0.85$ and $m = 0.118$. Parameter γ is the product of six partial factors that depend on the type of curing, environmental humidity, volume-surface ratio of the concrete member, slump, mass fractions of fine and total aggregate and on the air content.

2.2. *fib* MODEL CODE

The *fib* Model Code 2010 (*fib*, 2010), accepted in 2011 by the International Federation for Structural Concrete (in French "fédération internationale du béton, *fib*"), is a successor of CEB Model Codes 1990 and 1999, developed by the Euro-International Committee for Concrete (CEB). The compliance function has the form

$$J(t,t') = \frac{1}{E_C}\exp\left(-\frac{s}{2}\left[1-\sqrt{\frac{28}{t'}}\right]\right) + \frac{\phi_{RH}\beta_f}{E_C}\frac{1}{0.1+t'^{0.2}}\left(\frac{t-t'}{\beta_H\beta_T+t-t'}\right)^{0.3} \tag{3}$$

Parameter E_C is the conventional elastic modulus, parameter s depends on the strength class of cement and hardening characteristics (e.g., $s = 0.25$ for normal cement of strength class 42.5 or for rapidly hardening cement of strength class 32.5), parameters ϕ_{RH} and β_f express the influence of environmental humidity and mean compressive strength, parameter β_H depends on humidity and strength as well as on the notional member size, and parameter β_T reflects the influence of temperature and is equal to 1 at room temperature.

2.3. B3 MODEL

Model B3 (Bažant and Baweja, 1995; Bažant and Baweja, 2000) covers creep and shrinkage of concrete, including their coupling. The compliance function has the general form

$$J(t,t') = q_1 + q_2 Q(t,t') + q_3 \ln[1+(t-t')^n] + q_4 \ln\left(\frac{t}{t'}\right) + J_{\mathrm{d}}(t,t') \tag{4}$$

where $n = 0.1$, q_1 is the inverse of the asymptotic elastic modulus, the terms containing parameters q_2, q_3 and q_4 represent the aging viscoelastic compliance, non-aging viscoelastic compliance and flow compliance, respectively, and $J_{\mathrm{d}}(t,t')$ is the additional compliance due to drying. Here we consider only basic creep, i.e., creep of sealed specimens, not affected by drying, and thus $J_{\mathrm{d}}(t,t')$ can be omitted. Function Q is not available in a closed form and is defined by the integral formula

$$Q(t,t') = \int_{t'}^{t} \frac{n s^{-m}}{(s-t')+(s-t')^{1-n}}\,\mathrm{d}s \tag{5}$$

where $m = 0.5$. Its specific values can be obtained by numerical integration or approximated using an explicit formula given in (Bažant and Baweja, 1995) and (Bažant and Baweja, 2000). Parameters q_i, $i = 1,2,3,4$, can be estimated based on composition of the concrete mix and mean compressive strength of concrete using empirical formulae.

2.4. GL MODEL

The model proposed by (Gardner and Lockman, 2001) and denoted as the GL2000 Model is a modification of the earlier Atlanta97 Model (or GZ Model) of (Gardner and Zhao, 1993). The compliance function has

the form

$$J(t,t') = \frac{1}{3.5 + (E_C - 3.5)\exp\left(\frac{s}{2}\left[1 - \sqrt{\frac{28}{t'}}\right]\right)} + \\ + \frac{\Phi}{E_C}\left[\frac{2(t-t')^{0.3}}{(t-t')^{0.3}+14} + \sqrt{\frac{7(t-t')}{t'(t-t'+7)}} + c_h\sqrt{\frac{t-t'}{t-t'+0.12(V/S)^2}}\right] \tag{6}$$

Parameter E_C is the conventional elastic modulus, parameter s depends on the type of cement, parameter Φ is different from 1 only if the first loading is preceded by drying and, if this is the case, depends on the drying time before loading and on the volume-surface ratio V/S, and parameter c_h depends on the environmental humidity.

3. Updating of Model Parameters

Parameters of creep models presented in the previous section can be estimated from the basic characteristics of the concrete mix, curing procedure and environmental conditions. In this sense, the models can be considered as predictive and used already in the design stage. However, the dependence of model parameters on the basic characteristics described by empirical equations has a limited accuracy. To get a better agreement between the model and the real behavior, it would be advisable to perform creep tests of samples made of the specific concrete intended for the designed structure. Due to the long-term nature of the creep process, it is impossible to run the complete tests before construction. A compromise consists in continuous updating of the model parameters from measurements on the real structure or on companion specimens kept under the same environmental conditions. In the design stage, the parameters can be estimated from composition and corrected based on short-term tests. During construction and even after completion of the structure, the parameters can be continuously updated as more and more measured data become available. For this purpose, it is essential to know how the accuracy of predictions of the future behavior of the structure evolves depending on the growing amount of available information describing the past behavior (of the structure or of a specimen made of the same concrete).

In the present preliminary study, we restrict attention to basic creep, so that the effect of environmental humidity on the compliance function is eliminated. Measured values of the compliance function are taken from a comprehensive creep database assembled at Northwestern University (Bažant and Li, 2008). The database contains a wide range of creep tests run in the past in many laboratories around the globe under a variety of conditions. For our purpose, only sufficiently long tests (at least 1000 days of loading) performed under sealed conditions are considered. Furthermore, tests at extremely high or low temperatures (below 5°C or above 50°C) are excluded. The study is limited to concrete with mean compressive strength at 28 days lower than 82 MPa, loaded at stress levels not exceeding 45% of the strength. By applying these criteria, 40 tests from 12 laboratories have been extracted from the database.

In principle, the updating procedure could be applied to all parameters of each model. However, this would result into complicated problems of nonlinear regression, with multiple local minima of the error function and with a danger of extremely high sensitivity to the unavoidable scatter of experimental data, especially during early stages of the response when only a few measured values are available. For this

reason, it seems preferable to consider the updated compliance function as the original compliance function transformed in a linear fashion, i.e., by vertical scaling and shifting. Mathematically, we can write

$$J_u(t,t') = p_1 + p_2 J_o(t,t') \tag{7}$$

where J_o is the original compliance function with parameters estimated from composition and J_u is the updated compliance function, adjusted such that the early part of the measured response up to the updating time t_m is reproduced with the minimum possible error. Coefficients p_1 and p_2 are obtained by minimizing the function

$$F(p_1,p_2) = \sum_{i=1}^{m} \left[p_1 + p_2 J_o(t_i,t_0) - J_e(t_i,t_0)\right]^2 \tag{8}$$

where J_e denotes the experimentally determined compliance function, t_0 is the age of concrete at load application, and $t_1 < t_2 < \ldots < t_m$ are the times at which individual measurements were taken, up to the selected updating time t_m. For each specific test, t_0 is fixed but t_m can have an arbitrary value between t_1 and the age at the end of the test, $t_{\max}$. Therefore, coefficients p_1 and p_2 and the resulting updated function J_u depend on the time t_m at which the updating is performed.

Function F defined in (8) is quadratic in terms of the variables p_1 and p_2, and the stationarity conditions lead to two linear equations,

$$p_1 m + p_2 \sum_{i=1}^{m} J_o(t_i,t_0) = \sum_{i=1}^{m} J_e(t_i,t_0) \tag{9}$$

$$p_1 \sum_{i=1}^{m} J_o(t_i,t_0) + p_2 \sum_{i=1}^{m} J_o^2(t_i,t_0) = \sum_{i=1}^{m} J_e(t_i,t_0) J_o(t_i,t_0) \tag{10}$$

from which the optimal values of p_1 and p_2 are easily computed.

As an example, consider the data on Water Tower Place concrete (Russell and Burg, 1996). The concrete mix consisted of $c = 501.7$ kg/m^3 of cement, $w = 195.7$ kg/m^3 of water and $a = 1676$ kg/m^3 of aggregates, and the mean compressive strength at 28 days was $\bar{f}_c = 63$ MPa. From these data, parameters of the B3 model can be estimated as follows:

$$q_1 = 126.77\, \bar{f}_c^{-0.5} = 15.97 \quad [10^{-6}/\text{MPa}] \tag{11}$$

$$q_2 = 185.4\, c^{0.5} \bar{f}_c^{-0.9} = 99.75 \quad [10^{-6}/\text{MPa}] \tag{12}$$

$$q_3 = 0.29(w/c)^4 q_2 = 0.669 \quad [10^{-6}/\text{MPa}] \tag{13}$$

$$q_4 = 20.3(a/c)^{-0.7} = 8.727 \quad [10^{-6}/\text{MPa}] \tag{14}$$

The cement was of type R (rapid hardening) according to the CEB classification, and the mix also contained 11.8 kg/m^3 of fly ash. The experiments were performed on standard 6-inch cylinders (152 mm in diameter and 305 mm in height) at room temperature (23°C) and stress level 15.5 MPa (i.e., 25% of the mean strength). The specific test considered here (test C_078_05 from the database) started at age $t_1 = 28$ days and was run under sealed conditions for 6768 days, i.e., 18.5 years.

The "blind" prediction based on the parameter values (11)–(14) is plotted as the dash-dotted curve in Figure 1a. If the data measured during the first 143 days of loading are taken into account, equations (9)–(10) lead to $p_1 = -16.77 \times 10^{-6}$/MPa and $p_2 = 1.542$. The updated compliance function (7) is plotted in

Figure 1a as the solid curve; it corresponds to the B3 model with modified parameters $q_1^* = p_1 + p_2 q_1 = 7.85$, $q_2^* = p_2 q_2 = 153.78$, $q_3^* = p_2 q_3 = 1.031$ and $q_4^* = p_2 q_4 = 13.454$ (all in 10^{-6}/MPa). The result is somewhat disappointing. The original blind prediction underestimates the compliance after 18.5 years of loading by 14.4% and the updated prediction overestimates it by 14.7%. If the update is performed already after 14 days of loading, the results get even worse, and the extrapolated compliance after 18.5 years of loading is then overestimated by 45.8%; see the dashed curve in Figure 1a.

The reason for the poor accuracy of updated predictions is that the update optimizes the fit of the entire initial period of loading up to time t_m while the main purpose should be an improved accuracy of the extrapolation to longer times. Therefore, it makes sense to reduce the influence of early measurements and emphasize those that are closer to the updating time and thus also to the intended extrapolation. This can be achieved by introducing weight factors that have larger values for measurements at later times. The simplest choice is to take the time elapsed from the first loading up to the given measurement as the weight factor. The definition of the function F to be minimized is then changed from (8) to

$$F(p_1, p_2) = \sum_{i=1}^{m} (t_i - t_0) \left[p_1 + p_2 J_o(t_i, t_0) - J_e(t_i, t_0) \right]^2 \tag{15}$$

and equations (9)–(10) are adjusted accordingly. This modified updating approach leads to a substantial improvement, as shown in Figure 1b. The update after 14 days still does not lead to an improvement (but at least it is not as bad as for the standard updating method), but the update after 143 days gives a very nice prediction of the future evolution of compliance, with the value after 18.5 years overestimated by only 3.6%.

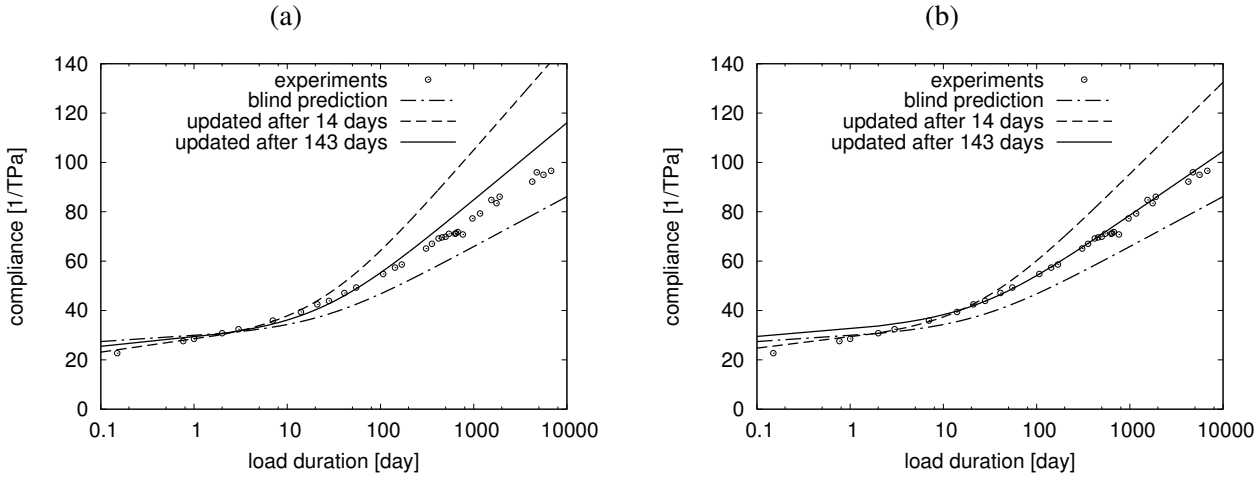


Figure 1. Compliance function of Water Tower Place concrete: (a) standard updating with equal weights of all measured points, (b) modified updating with the weight of each measured point proportional to the duration of loading.

4. Accuracy of Updated Predictions

To compare the accuracy of the original "blind" prediction and the updated predictions with different updating times, we can start from the so-called absolute residual error, which is considered as the root-mean-square (RMS) deviation of the prediction from the experimental data, averaged over the time interval from the updating time t_m up to time $t_{\max}$ at the end of the test. The averaging is done with respect to the logarithmic scale of the load duration, and so the error is defined as

$$e_u^{(m)} = \sqrt{\frac{\sum_{i=m+1}^{n} \ln \frac{t_{i+1}-t_0}{t_{i-1}-t_0} \left[J_u(t_i,t_0) - J_e(t_i,t_0)\right]^2}{\sum_{i=m+1}^{n} \ln \frac{t_{i+1}-t_0}{t_{i-1}-t_0}}} \tag{16}$$

where n is the total number of measurements, and t_{n+1} is set to t_n. The error of the updated prediction can be evaluated for $m \geq 2$, because the updating procedure needs at least the first two measurements, at times t_1 and t_2, for determination of two parameters, p_1 and p_2. Note that the deviations at times preceding or equal to the updating time t_m are not taken into account, and so the interval over which the RMS error is computed diminishes with increasing updating time (this is why the error is called "residual").

The error measure defined in (16) has the dimension of compliance and it can be used for comparison of the relative accuracy of individual models applied to the same test. For evaluation of the average accuracy in the set of 40 tests considered here, it is better to use the normalized error, defined as the absolute error according to (16) divided by the reference compliance value, which is taken as $J(t_0 + 1000, t_0)$ (i.e., as the compliance corresponding to the load duration of 1000 days). The normalized error is dimensionless and its value of 0.1 corresponds to 10% deviation with respect to the reference compliance value.

The dependence of the normalized residual error on the time elapsed from load application to the updating time is graphically presented for individual creep models in Figure 2. All the graphs still refer to one single test of the Water Tower Place concrete. The dashed curves show the error of the prediction based on standard updating and the solid curves refer to the modified updating. The first points of both curves always coincide because they correspond to updating after the second measurement, when two measured values uniquely determine parameters p_1 and p_2, independently on whether weighting is used or not. Later on, both curves in general differ and the modified update typically leads to higher accuracy, with some exceptions in the range from 4 to 80 days for the ACI model and from 1 to 13 days for the GL model.

For comparison, the graphs also contain dash-dotted curves that correspond to the blind prediction, with no updating. For the blind prediction, the model parameters remain fixed and the prediction does not evolve in time. However, to be able to compare directly the error of the blind and updated predictions, the error of the blind prediction is also evaluated over the interval that starts at the current updating time t_m. This residual error is defined by a formula similar to (16), with J_u replaced by J_o. As seen in Figure 2, the residual error of the blind prediction typically increases in time with increasing t_m, which means that the blind prediction is usually more accurate for short load durations than for long ones. Another interesting observation is that the early updates are in some cases less accurate than the blind prediction, which means that updating has an adverse effect on accuracy. This is particularly striking for model B3; see Figure 2a. Here, the blind prediction leads to normalized error (over the entire tested time interval) close to 0.1 while the update based on a few measurements up to load duration of 1 day gives a much higher error, about 0.5. To get improved accuracy as compared to the blind prediction, one needs to take into account measurements from at least 82 days of loading for the standard update and from at least 28 days of loading for the modified

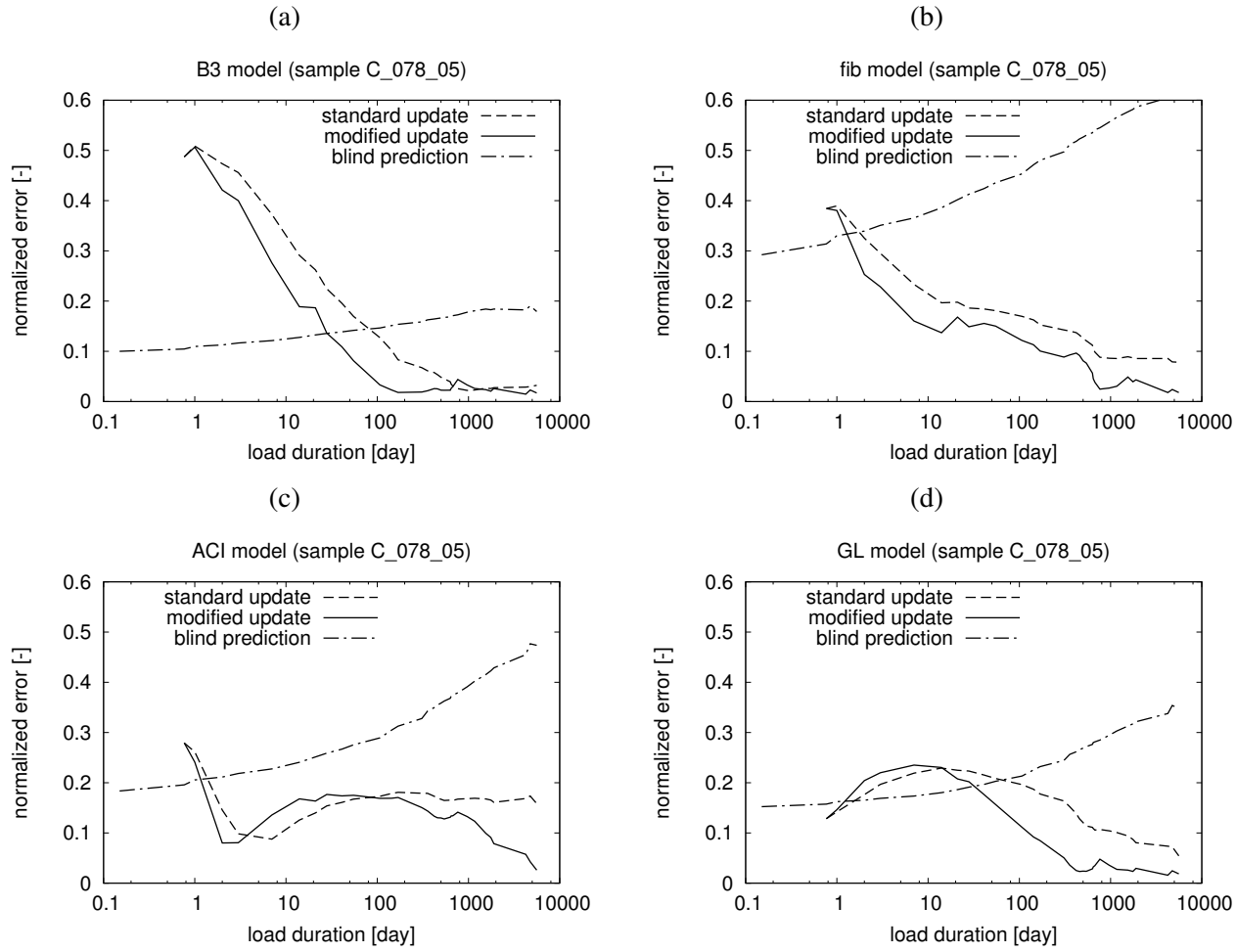


Figure 2. Normalized residual error $e_u^{(m)}/J(t_0+1000,t_0)$ as a function of the load duration up to the updating time, $t_m - t_0$, for Water Tower Place concrete: comparison of the blind prediction with the standard and modified updates for (a) B3 model, (b) *fib* model, (c) ACI model, (d) GL model.

update. A similar phenomenon, albeit less dramatic, can be observed for the other models as well. The initial error of the blind prediction using the *fib* model is much higher than for the B3 model, about 0.3, and the updated prediction becomes more accurate already after less than 2 days; see Figure 2b. For the ACI model, the behavior is similar, with a lower error of the initial blind prediction, about 0.2; see Figure 2c. Finally, for the GL model, the initial accuracy of the blind and updated predictions is comparable, about 0.16, but then the error of the updated predictions grows and remains above the error of the blind prediction up to 31 days for the modified update and up to 64 days for the standard update. Let us emphasize that all these observations refer to one single experimental test and cannot be considered as general statements. For instance, the fact that the blind prediction with the ACI model is more accurate than with the *fib* model is rather an exception. Nevertheless, this specific example illustrates the methodology and brings our attention

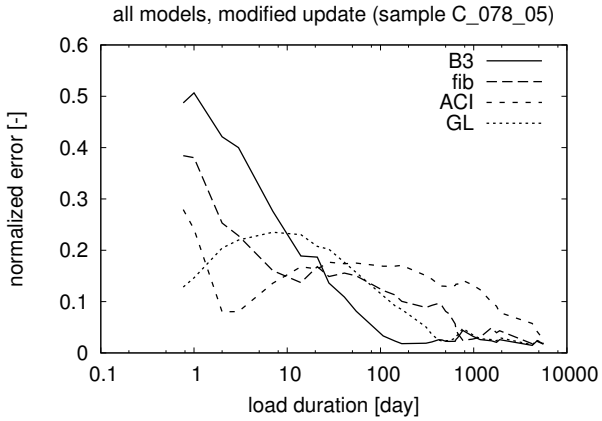


Figure 3. Normalized residual error $e_u^{(m)}/J(t_0+1000,t_0)$ as a function of the load duration up to the updating time, $t_m - t_0$, for Water Tower Place concrete: comparison of modified updates for individual models

to certain unexpected trends. For comparison, the evolution of the residual error based on the modified updating procedure for all creep models is plotted in Figure 3.

As already mentioned, the results in Figures 2 and 3 refer to one single test, and for other tests they can be quite different. To get an idea about the overall performance of individual creep models and updating procedures, it is necessary to take into account all the available tests and perform some averaging. Figure 4 shows the normalized residual error of the updated predictions based on the B3 model for all 40 tests considered in the present study. One can see that the modified updates (Figure 4b) are in general more accurate than standard ones (Figure 4a). After 100 days almost all the individual error points corresponding to the modified updates (perhaps with 2 or 3 exceptions) are below 0.2 and most of them are actually much lower.

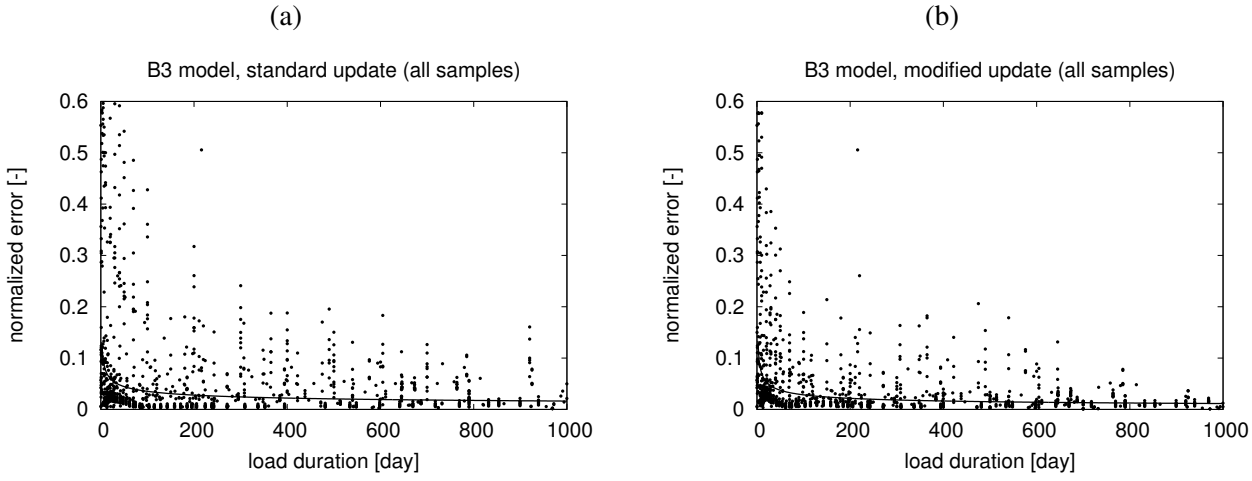


Figure 4. Normalized residual error of updated B3 model for all tests considered: (a) standard updating, (b) modified updating.

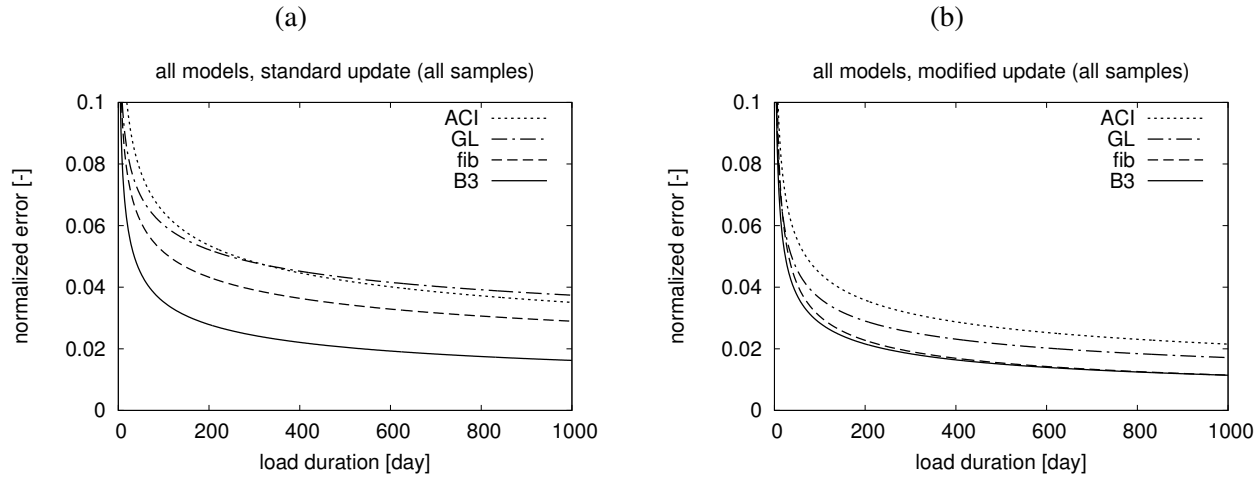


Figure 5. Comparison of power-law curves approximating the average normalized residual error for different models: (a) standard updating, (b) modified updating.

Table I. Average values of load duration after which updating leads to an improved prediction

model	*load duration [day]*
B3	48.7
fib	59.1
ACI	58.0
GL	18.6

For this overall comparison, all the tests have been truncated at load duration of 1000 days, otherwise the meaning of the residual error would be different for tests of different total durations. For easier comparison, the dependence of the normalized residual error on the updating time has been fitted (in the least-square sense) by a power law. The corresponding smooth curves that represent the average errors for individual models are plotted in Figure 5. It is confirmed that the modified updating procedure gives in general better results than the standard one. This effect is particularly strong for the *fib* model and the GL model. With standard updating, model B3 gives by far the highest accuracy, while the *fib* model is the second best. With modified updating, the average performance of the B3 model and the *fib* model is comparable. The GL model gives higher errors of the updated predictions, and the worst results are obtained with the ACI model, which is no surprise because the original version of this model is more than 40 years old.

It is also interesting to compare the typical load durations after which the modified updating procedure leads to an improvement. The average times needed to get at least the same accuracy as with the original blind prediction are summarized in Table I. It turns out that updating is beneficial for the GL model already after 19 days of loading while the other models require between 49 and 59 days. This means that if, for

instance, the B3 model is used, it does not seem to be a good idea to adjust the parameters based on measurements that cover only a few weeks of loading, and one should use data covering at least several months.

5. Conclusions

It should be emphasized that the results presented in this short paper are only partial, and a more detailed evaluation remains to be finished. An extension to drying creep represents another important step to be taken before the final conclusions can be drawn. Nevertheless, the preliminary findings lead to the following recommendations:

- Updating of creep predictions based on short-time measured data exhibit higher accuracy if the updating procedure incorporates weight factors that reduce the influence of very early stages of the response.
- The updating procedure can be expected to provide better accuracy than predictions based on concrete mix composition and similar data, provided that the measured response covers a certain minimum period of time, which is in the order of a couple of weeks for the GL model and a couple of months for the other models considered in this study.
- If the updating is based on measurements covering a sufficiently long loading period, the B3 model and the *fib* model seem to provide the most accurate predictions.

Acknowledgements

Financial support of the Czech Science Foundation under project P105/10/2400 is gratefully acknowledged.

References

ACI Committee 209 Prediction of creep, shrinkage and temperature effects in concrete structures. American Concrete Institute, Farmington Hills, Michigan, 1992.

Bažant, Z. P. and Baweja, S. Creep and shrinkage prediction model for analysis and design of concrete structures – model B3. *Materials and Structures*, 28:357–365, 1995.

Bažant, Z. P. and Baweja, S. Creep and shrinkage prediction model for analysis and design of concrete structures: Model B3. *Adam Neville Symposium: Creep and Shrinkage—Structural Design Effects*, 1–83, 2000.

Bažant, Z. P. and Li, G.-H. Comprehensive database on concrete creep and shrinkage. *ACI Materials Journal*, 106:635–638, 2008.

Bažant, Z. P., Yu, Q., Li, G.-H., Klein, G. and Křístek, V. Excessive deflections of record-span prestressed box girder: Lessons learned from the collapse of the Koror-Babeldaob bridge in Palau. *ACI Concrete International*, 32:44–52, 2010.

Bažant, Z. P., Hubler, M. and Yu, Q. Excessive creep deflections: An awakening. *ACI Concrete International*, 33:44–46, 2011.

Brooks, J. J. 30-year creep and shrinkage of concrete. *Magazine of Concrete Research*, 57:545–556, 2005.

Fédération internationale de béton (*fib*). Model Code 2010, first complete draft. Lausanne, Switzerland, 2010.

Gardner, N. J. and Lockman, M. J. Design provisions for drying shrinkage and creep of normal strength. *ACI Materials Journal*, 98:159–167, 2001.

Gardner, N. J. and Zhao, J. W. Creep and shrinkage revisited. *ACI Materials Journal*, 90:236–246, 1993.

Russell, H. and Burg, R. Test data on Water Tower Place Concrete. *International Workshop on High Performance Concrete*, American Concrete Institute, Farmington Hills, Michigan, 1996.

Toward complex risk assessment and management based on multisource data statistics of natural and technological disasters

Yuriy V. Kostyuchenko[1)], I. Kopachevsky[1)], P. Zlateva[2)], D. Solovyov[3)], and P. Akymenko[4)]
[1)]*Scientific Centre for Aerospace Research of the Earth of National Academy of Sciences of Ukraine; 55-b, O. Honchar street, Kiev, 01601, Ukraine; e-mail: yvk@casre.kiev.ua*
[2)]*Institute of System Engineering and Robotics, Bulgarian Academy of Sciences, Acad. G.Bonchev Str., Block 2, P.O.Box 79, Sofia, 1113, Bulgaria; e-mail: plamzlateva@abv.bg*
[3)]*Marine Hydrophysical Institute of National Academy of Sciences of Ukraine; 2, Kapitanskaya street, Sevastopol, 99011, Ukraine; e-mail: solmit@dvs.net.ua*
[4)]*Aerospace Society of Ukraine, 8, Moskovskaya street, Kiev, 01010, Ukraine, e-mail: akymenko@ukr.net*

Abstract: Long-term regional statistics of disasters distributions with consideration of climatic parameters and variations of economic activity have been analyzed. Suitable techniques of observation data regularization for normalization of data reliability are proposed. The technique proposed based on utilization of modified kernel-based nonlinear principal component analysis. Problem-oriented analysis of regional climatic parameters trends has been done also. Correlation matrices for sub-regional distribution of several types of natural and technological disasters in connection with dynamics of climatic parameters for observation period have been constructed. Form of indexes for analysis of impact of climatic parameters change to disaster emergency is proposed. As the result of analysis the stable correlations between analyzed distributions have been identified. Basing these correlations the assumptions about driving forces and triggers of most hazardous disasters types were formulated. In view of obtained results the suitable form of copula for disaster risk analysis and mitigation has been proposed.

Keywords: natural & technological disasters; multisource observations; data regularization; nonlinear component analysis; disaster driving forces; risk analysis; copula.

Problem of construction of correct techniques of risk assessment, first of all disaster – induced socio-ecological risks, is often connected with lack of reliable long-term series of catastrophic events observations, socio-ecological parameters, and natural systems state (Bartell, Gardner, O'Neill, 1992; O'Neill, Gardner, Barnthouse, et al., 1982). Correct and regular events statistics is important for construction of adequate risk function and also for risk management strategies development (US EPA, 2008; White, Maurice, Mysz, et al., 2008). Demonstration of way of observation data regularization for normalization of data reliability is the purpose of this paper.

Correct statistical analysis requires the set of data $\boldsymbol{x}_i$ with controlled reliability, which reflects distribution of investigated parameters over study area during whole observation period (taking into account variances of reliability of observation and archive data $\boldsymbol{x}_t$). Set of observation data $\boldsymbol{x}_t$ ($\mathbf{x}_t \in R^m$) consists of multi-source data: historical records, archives, observations, measurements, etc., including data with sufficient reliability $\boldsymbol{x}_j$ ($\mathbf{x}_j \in R^m$), where $j = 1, \ldots. N$. Problem of determination of controlled quality and

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

reliability spatial-temporal distribution of investigated parameters might be solved in framework of tasks of multivariate random processes analysis and multidimensional processes regularization (Raiffa, Schlaifer, 1968).

Required regularization may be provided by different ways. If we able to formulate stable hypothesis on distribution of reliability of regional archives data in the framework of defined problem we may to propose relatively simple way to determine investigated parameters distributions $\boldsymbol{x}_t^{(x,y)}$ towards distributions on measured sites $\boldsymbol{x}^m{}_t$ basing on Fowler, Kilsby, O'Connell (2003,):

$$x_t^{(x,y)} = \sum_{m=1}^{n} w_{x,y}(\tilde{x}_t^m) x_t^m \,, \tag{1}$$

where weighting coefficients $w_{x,y}(\tilde{x}_t^m)$ determined as:

$$\min\{\sum_{m=1}^{n} \sum_{x_t^m \in R^m} w_{x,y}(\tilde{x}_t^m)(1-\frac{x_t^m}{\tilde{x}_t^m})^2\} \tag{2}$$

according to Cowpertwait (1995). Here m – number of records/points of measurements or observations; n – number of observation series; $\boldsymbol{x}^m{}_t$ – distribution of observations data; R^m – set (aggregate collection) of observations; $\tilde{x}_t^m$ – mean distribution of measured parameters.

This is the simple way to obtain a regular spatial distribution of analyzed parameters over the study area, on which we can apply further analysis, in particular temporal regularization.

Further regularization should take into account both observation distribution temporal non-linearity (caused by imperfection of available statistics) and features of temporal-spatial heterogeneity of data distribution caused by systemic complexity of studied phenomena – natural and technological disasters. According to (Mudelsee, Börngen, Tetzlaff, 2001; Lee, Yoo, Choi, et al, 2004; Villez, Ruiz, Sin, et al, 2008) the kernel based non-linear approaches are quite effective for analysis of such types of distributions.

Proposed method is based on modified kernel principal component analysis (KPCA) (Scheolkopf, Smola, Muller, 1998; Mika, Scheolkopf, Smola, et al., 1999; Romdhani, Gong, Psarrou, 1999). In the framework of this approach the algorithm of non-linear regularization might be described as following rule:

$$x_i = \sum_{i=1}^{N} \alpha_i^k \tilde{k}_t(x_i, x_t) \tag{3}$$

In equation (3) the coefficients α selected according to optimal balance of relative validation function and covariance matrix, for example as (Lee, Yoo, Choi, et al, 2004):

$$C^F v = \frac{1}{N} \sum_{j=1}^{N} \Phi(x_j)\Phi(x_j)^T \cdot \sum_{i=1}^{N} \alpha_i \Phi(x_i) \,, \tag{4}$$

where non-linear mapping function of input data distribution Φ determined as (Scheolkopf, Smola, Muller, 1998):

$$\sum_{k=1}^{N} \Phi(x_k) = 0 \,, \tag{5}$$

and $\widetilde{k}_t$ – is mean values of kernel-matrix $\mathbf{K} \in R^N$ ($[\mathbf{K}]_{ij} = [k(\mathbf{x}_i, \mathbf{x}_j)]$). Vector components of matrix determined as $\mathbf{k_t} \in R^N$; $[\mathbf{k_i}]_j = [k_t(\mathbf{x_t}, \mathbf{x_j})]$. Matrix calculated according to modified rule of (Christianini, Shawe-Taylor, 2000) as:

$$\mathbf{k}_t(\mathbf{x}_i, \mathbf{x}_t) = \left\langle \rho_{j,t}^{x_j} (1 - \rho_{j,i})^{x_j} \right\rangle \tag{6}$$

Here ρ – empirical parameters, selected according to the classification model of study phenomena (Villez, Ruiz, Sin, et al, 2008).

Using described algorithm it is possible to obtain regularized spatial-temporal distribution of investigated parameters over whole observation period with rectified reliability (Mudelsee, Börngen, Tetzlaff, 2001).

Proposed regularization algorithm has been applied to analysis of disaster statistics and obtaining of smoothed distributions of frequency of disasters, climatic and socio-economic parameters for period 1960–2010 over Southern part of Ukraine adjacent to Black Sea – Black Sea Lowland. Study region includes Odesa, Mykolayiv, Kherson and Zaporizhyia administrative regions, and Northern – West shelf of Black Sea.

Black Sea Lowland is a flat plain, slightly sloping to the south. Heights from 5 (near the estuary Kuyal'nik liman), an average of 90–150 m. The Black Sea lowland composed of Paleogene and Neogene marine sediments (limestone, sand, clay) overlain by loess and loess-like loams. Low plains crossed by wide river valleys of the Dnieper, Southern Bug, Dniester Rivers and other watersheds are flat, characterized by sinkholes. The coastline is predominantly steep. Near the sea there are many deep estuaries and limans (Dnieper, Dniester, and others) and sand braid. The steppe landscapes with southern chernozem and dark chestnut soils are dominated. Much of the steppe is under cultivation and used as agricultural land.

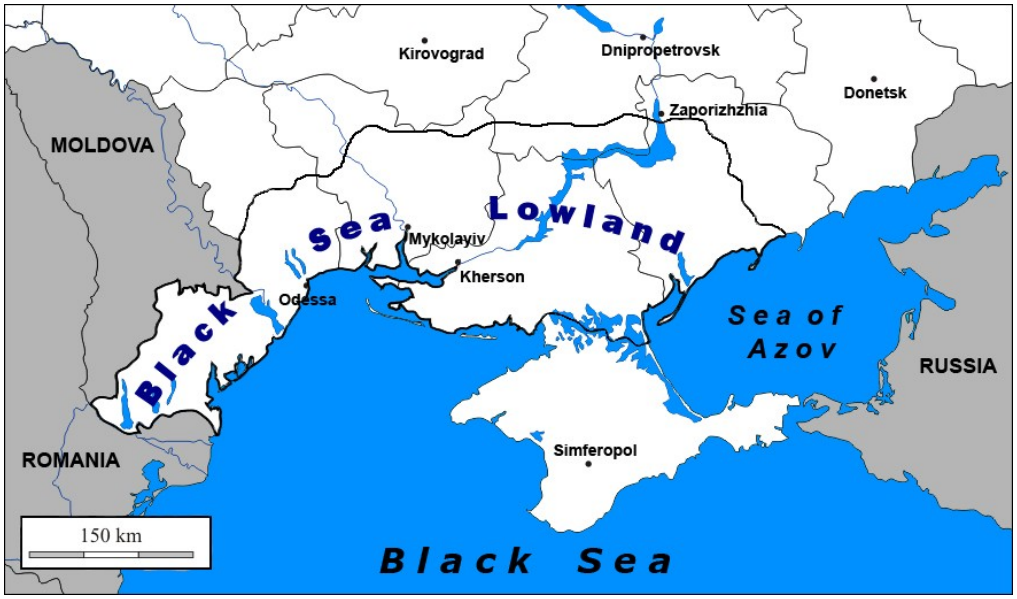


Figure 1. Study area: Black Sea Lowland

For the area studied for period 1960–2010 has been analyzed wide group of natural and technological disasters. Was used international classification of disasters according to (Guha-Sapir, Vos, Below, 2010) with some minor variations caused by national classification features and data availability. Final set of disaster analyzed includes the following types of events:

a) *natural*: climatological (& meteorological): wind storms, squalls, tornadoes, dust storms, hail-storms, heavy rains, heavy snowfalls, glazed frosts, heavy snow blasts, heavy frosts, heavy heat, heat and cold waves, droughts, forest, grassland fires, peatmoor fires; hydrological: high water level (floods, freshets), mudflows, underfloodings, inundations, avalanches; geophysical (and geological): earthquakes, mud volcanoes eruptions, landslides, rockslides, surface subsidence, karst caverns rockfalls; epidemiological (biological, connected with human diseases): individual cases of exotic and specific dangerous infectious diseases, group cases of dangerous infectious diseases, epidemic outbreak of dangerous infectious diseases, epidemics, pandemics, human infectious diseases of uncleared aetiology; epizootical (biological, connected with animal diseases): individual cases of exotic and specific dangerous infectious diseases, group cases of dangerous infectious diseases, epizootic outbreak of dangerous infectious diseases, epizootics, enzootics, animal infectious diseases of uncleared aetiology; and additionally was analyzed separately the landscape fires: forest fires, grassland fires, cropland fires, peatmoor fires;
b) *technological*: abrupt destroying of structures, buildings and communications; transport accidents including connected with threat of pollution; accidents on energy systems and structures: accidents on hydroelectric power plants, nuclear power plants emergencies, accidents on steam power plants, other power plants types emergencies, breakdowns of power lines; accidents on life-support systems: accidents on drainage systems with massive contamination, accidents on heating and hot water supply systems, accidents on water supply systems, accidents on gas pipelines, oil pipelines breakdowns, hydrodynamic accidents, dams breaks, technical reservoirs emergency draw-downs; anthropogenic derived environmental contamination: ground contamination, air contamination, surface water contamination, drinking-water contamination, ground-water contamination; fires and explosions: fires and explosions of civil and industrial structures, buildings and communications, mines, and dangerous objects.

For analysis were used data from international surveys and national reports (GSOD; USSR National Economy, 1969, 1980; State Budget of USSR, 1981–1985, 1989; National report, 2004, 2006, 2009; Guha-Sapir, Vos, Below, 2010).

On Figures 2 and 3 presented calculated according the described algorithm resulting distributions of mean probability of various types of disasters per year per 1.000 km^2 during observation period [$<P>/10^3 km^2 year$].

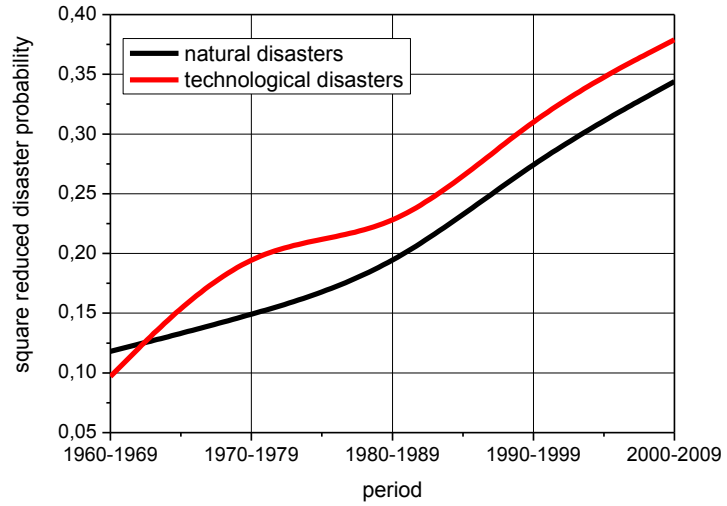


Figure 2. Natural and technological disasters probability distribution over study area

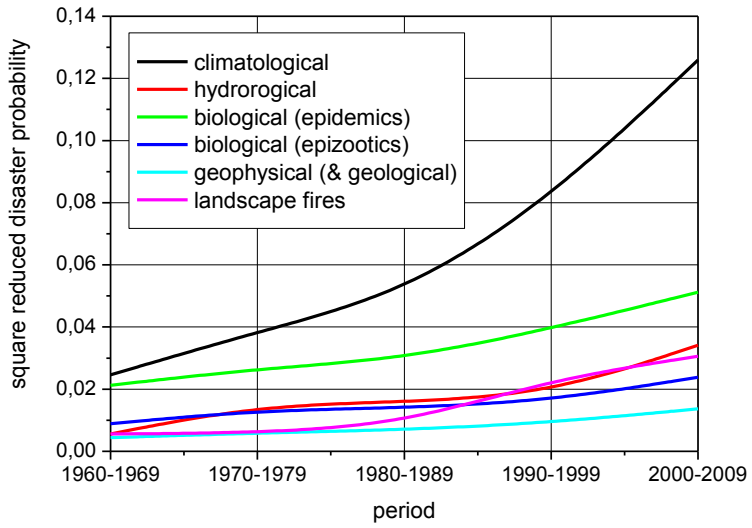


Figure 3. Distribution of probability of separate types of natural disasters

Basing on geographical distribution (location, water regime, topography) and in accordance with distribution of long-term meteorological trends the area studied was divided to tree conventional zones: preferentially sea impacted zone, zone of mixed impact and continental sites. Distribution of disasters over these zones might be calculated separately as it is shown on Figures 4 and 5.

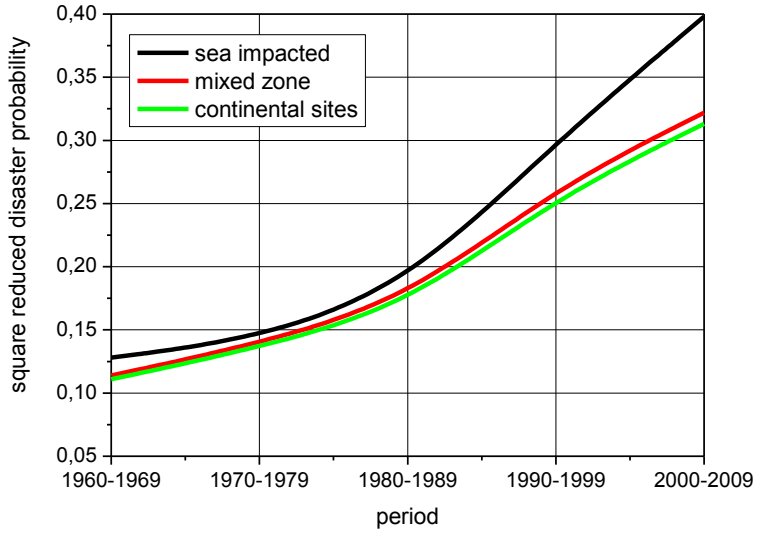


Figure 4. Distribution of probability of natural disasters according to regional features

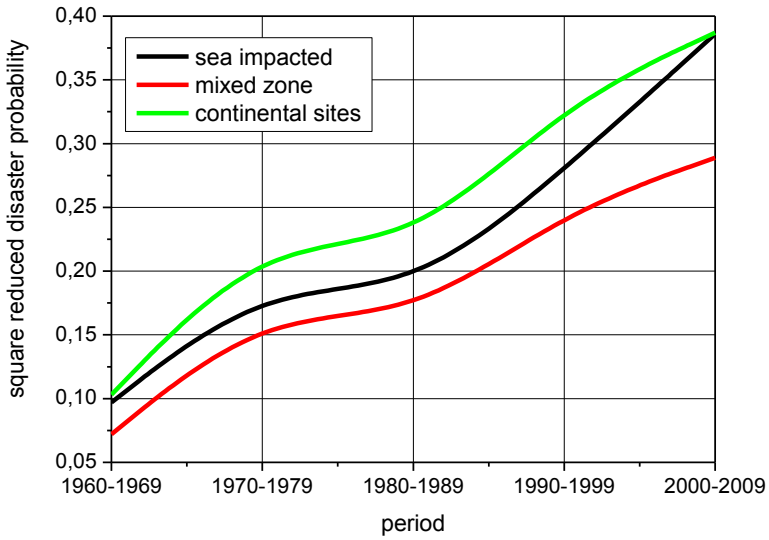


Figure 5. Distribution of probability of technological disasters according to regional features

Obtained distributions (Figures 2 – 5) demonstrate good corresponding with impact of climate-forming factors (sea influence), and reflect the distribution of regional geo-morphological features, such as terrain, erosion, neo-tectonics, landslides, subsidences, peculiarities of local water regime. Also the spatial-temporal changes of land-use are reflected on disaster probability distribution.

Important direction of further studies is analysis of correlations of disasters probability and regional ecosystems reactions: changes of biomass, variations of horizontal and vertical distribution of vegetation,

which might be registered via satellite observation (analysis of integrated spectral reflectance indexes as indicators of vegetation state, ecological stresses, and water load variations).

On Figures 2 – 5 presented the smoothed distributions of investigated parameters calculated using proposed algorithm. Non-smoothed output sets of parameters have been examined to correlations. Obtained correlations presented in Table 1.

Table I. Correlations between analyzed disaster distribution and some climatic parameters

disasters types 1960-2010		climat disasters	hydrolog disasters	bio (epid)	bio (epizoo)	geophys & geolog	landscape fire	mean natural disasters	power accident	water supply	environ contamin	construct destruct	industrial fires	transport accidents	mean techno disasters	temperature mean	temperature max	temperature reduced
climat disasters	corr coef	1	0,96237	0,99828	0,97971	0,99817	0,97408	0,98765	0,96984	0,99344	0,98104	0,84018	0,90555	0,88946	0,94715	0,85675	0,90196	0,94527
	sig		0,00871	8,54E-05	0,00346	9,40E-05	0,00499	0,00165	0,00626	6,37E-04	0,00313	0,07483	0,03435	0,04338	0,01447	0,06367	0,03631	0,01524
hydrolog disasters	corr coef	0,96237	1	0,9611	0,99706	0,97567	0,87989	0,98765	0,91745	0,97117	0,99346	0,84018	0,90555	0,87337	0,94715	0,83305	0,79006	0,83623
	sig	0,00871		0,00915	1,91E-04	0,00454	0,04906	0,0256	0,02812	0,00585	6,34E-04	0,04033	0,02952	0,05306	0,0198	0,07981	0,11176	0,07757
bio (epid)	corr coef	0,99828	0,9611	1	0,97937	0,99512	0,97512	0,99237	0,98187	0,99723	0,98356	0,8615	0,92725	0,91437	0,9637	0,83035	0,9124	0,95321
	sig	8,54E-05	0,00915		0,00355	4,09E-04	0,00469	7,99E-04	0,00292	1,75E-04	0,00252	0,06057	0,0233	0,02969	0,00826	0,08171	0,03071	0,01206
bio (epizoo)	corr coef	0,97971	0,99706	0,97937	1	0,98854	0,91332	0,94889	0,94264	0,98626	0,99862	0,8922	0,92608	0,8926	0,95074	0,83939	0,83153	0,87487
	sig	0,00346	1,91E-04	0,00355		0,00147	0,03023	0,01376	0,01635	0,00193	6,15E-05	0,04179	0,02386	0,04157	0,01303	0,07538	0,08088	0,05212
geophys & geolog	corr coef	0,99817	0,97567	0,99512	0,98854	1	0,95946	0,97735	0,95909	0,99256	0,98769	0,84655	0,90322	0,88061	0,94277	0,86783	0,87981	0,92432
	sig	9,40E-05	0,00454	4,09E-04	0,00147		0,00974	0,00408	0,00987	7,70E-04	0,00164	0,07047	0,03561	0,04863	0,01629	0,05652	0,0491	0,02471
landscape fires	corr coef	0,97408	0,87989	0,97512	0,91332	0,95946	1	0,98737	0,96769	0,96298	0,92104	0,75747	0,86885	0,8789	0,92174	0,8235	0,965	0,98266
	sig	0,00499	0,04906	0,00469	0,03023	0,00974		0,0017	0,00694	0,0085	0,02632	0,13805	0,05588	0,04966	0,02597	0,08662	0,00782	0,00273
mean natural disasters	corr coef	0,98765	0,92249	0,99237	0,94889	0,97735	0,98737	1	0,99118	0,98439	0,95754	0,84799	0,92193	0,92054	0,96033	0,78444	0,93509	0,98298
	sig	0,00165	0,0256	7,99E-04	0,01376	0,00408	0,0017		9,93E-04	0,00234	0,01044	0,0695	0,02588	0,02657	0,00943	0,11618	0,01966	0,00266
power accident	corr coef	0,96984	0,91745	0,98187	0,94264	0,95909	0,96769	0,99118	1	0,98104	0,95655	0,89134	0,9614	0,9636	0,98556	0,72164	0,93541	0,97419
	sig	0,00626	0,02812	0,00292	0,01635	0,00987	0,00694	9,93E-04		0,00313	0,0108	0,04229	0,00905	0,00829	0,00208	0,16874	0,93541	0,00496
water supply	corr coef	0,99344	0,97117	0,99723	0,98626	0,99256	0,96298	0,98439	0,98104	1	0,99119	0,87916	0,94494	0,93057	0,97507	0,82443	0,9089	0,93717
	sig	6,37E-04	0,00585	1,75E-04	0,00193	7,70E-04	0,0085	0,00234	0,00313		9,91E-04	0,0495	0,01538	0,02173	0,00471	0,08595	0,03255	0,01872
environ contamin	corr coef	0,98104	0,99346	0,98356	0,99862	0,98769	0,92104	0,95754	0,95655	0,99119	1	0,9061	0,94382	0,91489	0,96569	0,81934	0,84868	0,88928
	sig	0,00313	6,34E-04	0,00252	6,15E-05	0,00164	0,02632	0,01044	0,0108	9,91E-04		0,03405	0,01585	0,02942	0,00759	0,08964	0,06904	0,04349
construct destruct	corr coef	0,84018	0,89476	0,8615	0,8922	0,84655	0,75747	0,84799	0,89134	0,87916	0,9061	1	0,96026	0,92844	0,93527	0,50393	0,69397	0,78792
	sig	0,07483	0,04033	0,06057	0,04179	0,07047	0,13805	0,0695	0,04229	0,0495	0,03405		0,00945	0,02273	0,01958	0,38667	0,19363	0,11344
industrial fires	corr coef	0,90555	0,91469	0,92725	0,92608	0,90322	0,86885	0,92193	0,9614	0,94494	0,94382	0,96026	1	0,99276	0,99264	0,6142	0,85538	0,87952
	sig	0,03435	0,02952	0,0233	0,02386	0,03561	0,05588	0,02588	0,00905	0,01538	0,01585	0,00945		7,38E-04	7,58E-04	0,27039	0,06457	0,04928
transport accidents	corr coef	0,88946	0,87337	0,91437	0,8926	0,88061	0,8789	0,92054	0,9636	0,93057	0,91489	0,92844	0,99276	1	0,98775	0,58602	0,89093	0,89666
	sig	0,04338	0,05306	0,02969	0,04157	0,04863	0,04966	0,02657	0,00829	0,02173	0,02942	0,02273	7,38E-04		0,00162	0,29908	0,04253	0,03926
mean techno disasters	corr coef	0,94715	0,93477	0,9637	0,95074	0,94277	0,92174	0,96033	0,98556	0,97507	0,96569	0,93527	0,99264	0,98775	1	0,68939	0,9007	0,92296
	sig	0,01447	0,0198	0,00826	0,01303	0,01629	0,02597	0,00943	0,00208	0,00471	0,00759	0,01958	7,58E-04	0,00162		0,19783	0,037	0,02537
temperature mean	corr coef	0,85675	0,83305	0,83035	0,83939	0,86783	0,8235	0,78444	0,72164	0,82443	0,81934	0,50393	0,6142	0,58602	0,68939	1	0,74362	0,70956
	sig	0,06367	0,07981	0,08171	0,07538	0,05652	0,08662	0,11618	0,16874	0,08595	0,08964	0,38667	0,27039	0,29908	0,19783		0,1497	0,17948
temperature max	corr coef	0,90196	0,79006	0,9124	0,83153	0,87981	0,965	0,93509	0,93541	0,9089	0,84868	0,69397	0,85538	0,89093	0,9007	0,74362	1	0,94975
	sig	0,03631	0,11176	0,03071	0,08088	0,0491	0,00782	0,01966	0,93541	0,03255	0,06904	0,19363	0,06457	0,04253	0,037	0,1497		0,01342
temperature reduced	corr coef	0,94527	0,83623	0,95321	0,87487	0,92432	0,98266	0,98298	0,97419	0,93717	0,88928	0,78792	0,87952	0,89666	0,92296	0,70956	0,94975	1
	sig	0,01524	0,07757	0,01206	0,05212	0,02471	0,00273	0,00266	0,00496	0,01872	0,04349	0,11344	0,04928	0,03926	0,02537	0,17948	0,01342	

Presented figures are the solid base for analysis of driving forces and causes of disasters. In particular, such dangerous technological hazard as environmental contamination, which usually considering in connection with transport accidents, or water and energy systems collapse, is closely and more hard connected with natural hydrological disasters (floods, inundations). And drivers of hydrological threats are climatological and hydro-meteorological disasters.

Drastic change of frequency of climatological and hydro-meteorological disasters caused by increasing of droughts and heat waves in period 1985–2010, but it also connected with regional agricultural and land-use practices, not adopted to altered climatic stress and lead to additional losses registered as the disaster events.

Also it should be noted that such disasters as destroying of structures and buildings, fires and explosions of civil and industrial sector, and transport accidents obey the separate relationships, and so should be analyzed as the separate statistical group.

The special issue is the analysis of climate impact to disaster probability. As the index for climate – disaster correlation analysis the "reduced temperature maximum factor" T_{red} was proposed in form:

$$T_{red} = (1 - \frac{\frac{1}{N}\sum_{n=1}^{N} T_n}{T^{\max}})(1 - \frac{1}{T^{\max} - \frac{1}{N}\sum_{n=1}^{N} T^{\max}}), \quad (7)$$

where N – number of meteorological observations, T_n – is the observed temperature, T^{max} – maximal registered temperature during observation period. Mean correlations of this index with disaster distributions are higher than other climatic parameters such as mean or maximal temperature during the investigated period.

Technically, formula (7) describes the variations of maximal registered temperatures relative to long-term observed changes of maximal temperature (mean maximal temperature), reduced to regional trend of mean temperature. Used data indicates that T_{red} index is optimal correlation parameter for spatial-temporal variability of disasters frequency toward climate parameters variations on long time periods.

So using proposed algorithm it is possible to obtain spatially-temporally regularized distributions of disasters and to analyze the corresponding risks features.

Analyzing the results obtained it is possible to propose for the further risks assessment and disaster connected damage relief strategies development to use multivariate cumulative distribution function taking into account identified internal interrelations in structure of disasters distribution. If observation distribution describing as vector (X_1, X_2,..., X_n) with multivariate cumulative distribution function $F_{X1, X2,..., Xn}(x_1, x_2,..., x_n)$ and marginal distributions functions $F_{X1}(x_1)$, $F_{X2}(x_2)$,..., $F_{Xn}(x_n)$, according to (Sklar, 1996), it is possible to determine distribution function C on hypercube $[0,1]^n$ with uniform marginal distributions on interval [0,1] and $F_{X1,X2,...,Xn}(x_1, x_2,..., x_n) = C(F_{X1}(x_1), ..., F_{Xn}(x_n))$. Proposed regularization algorithm allows to produce distributions, which satisfy such conditions, and additionally distribution F_{Xi} might be presented as continuous for all $i = 1, 2, ..., n$ and so according to (Sklar, 1996), this distribution function C, or copula, will be determined uniquely. For our case following to (Rachev, Martin, Racheva, Stoyanov, 2009) the copula might be presented in form:

$$C_{v,\mathbf{c}}(u_1,u_2,...,u_n) = \frac{v\sqrt{|\mathbf{c}|(v\pi)^n}}{v+n} \cdot \frac{\prod_{n=1}^{N} \frac{(1+\frac{1}{n})^{\frac{v+n}{2}}}{1+\frac{v+n}{2n}}}{\prod_{n=1}^{N} \frac{(1+\frac{1}{n})^{\frac{v}{2}}}{1+\frac{v}{2n}}} \cdot \int_{\mathbf{x}_1}^{(u_1)} \dots \int_{\mathbf{x}_n}^{(u_n)} \left(1 + \frac{\mathbf{x}_i^{(x,y)}}{\mathbf{c}v}\right) \mathrm{d}x_t^{(x,y)} \quad (8)$$

where $\boldsymbol{c}$ – is the correlation matrix obtained through analysis of regularized data.

Using of such type view of function for description of dependencies between disaster risks corresponding to results of insurance analysts, in particular (Bradley and Taqqu, 2003; Rachev, Menn and Fabozzi, 2005). Proposed equation of copula might be recognized with some assumptions as the particular case of general formula proposed by (Rachev, Martin, Racheva, Stoyanov, 2009) or (Rachev, Menn and

Fabozzi, 2005). Proposed view of copula reflects that risks in complex multi-component systems can not be adequately described by linear superposition of scalar correlations on large time periods (Ermoliev, Hordijk, 2006; Ermoliev, Winterfeldt, 2010). More complicated interdependencies reflect complex systemic interrelations, spatial-temporal heterogeneities and study phenomena immanent uncertainties should be used in this case (Ermoliev, Winterfeldt, 2010). Besides it might be supposed that proposed way to analysis of multidimensional distributions of multivariate correlations could be quite successful used in systemic tasks of optimal control (Warga, 1972). This is the field for further applications.

References

Bartell S.M., Gardner R.H., O'Neill R.V. Ecological risk estimation. Boca Raton, FL: Lewis Publishers, 1992

Bradley B.O. and Taqqu M.S. Financial Risk and Heavy Tails // in *Handbook of Heavy Tailed Distributions in Finance*, edited by S. T. Rachev, Elsevier/North-Holland, - 2003

Christianini N., Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, UK, - 2000, 212 p.

Cowpertwait P.S.P. A generalized spatial–temporal model of rainfall based on a clustered point process // Proc. R. Soc. Lond. Ser. A 450, 1995, 163–175

Ermoliev Yuri, Winterfeldt Detlof von. Risk, Security and Robust Solutions // IIASA Interim Report, IR-10-013, - 2010, IIASA, 41 p.

Ermoliev, Y., Hordijk, L. (2006). Global Changes: Facets of Robust Decisions, in. K. Marti, Y. Ermoliev, M. Makowski, G. Pflug, (eds.): Coping with Uncertainty, Modeling and Policy Issues. Springer-Verlag, Berlin, Germany, pp. 4-28

Fowler H.J., Kilsby C.G., O'Connell P.E. Modeling the impacts of climatic change and variability on the reliability, resilience and vulnerability of a water resource system // Water Resour. Res. **39**, 2003, 1222

GSOD (Global Summary of the Day) // National Climatic Data Center of U.S. Department of Commerce, courtesy of NOAA Satellite and Information Service of National Environmental Satellite, Data, and Information Service (NESDIS) - www.ncdc.noaa.gov

Guha-Sapir D., Vos F., Below R., with Ponserre S. Annual Disaster Statistical Review 2010: The Numbers and Trends. Brussels: CRED; 2011, - 50 p.

Lee J.-M., Yoo C. K., Choi S. W., Vanrolleghem P. A., Lee I.-B. Nonlinear process monitoring using kernel principal component analysis / Chemical Engineering Science, 2004, **59**, pp. 223 – 234

Mika S., Scheolkopf B., Smola A.J., MJuller K.-R., Scholz M., Ratsch G.,. Kernel PCA and de-noising in feature spaces // Advances in Neural Information Processing Systems 11, 1999, pp. 536–542

Mudelsee M., Börngen M., Tetzlaff G. On the estimation of trends in the frequency of extreme weather and climate events. // in: Raabe A., Arnold K., Eds., Wissenschaftliche Mitteilungen, Institut für Meteorologie der Universität Leipzig, Institut für Troposphärenforschung e. V. Leipzig, Leipzig, v. 22, 2001, p. 78 - 88

National report "On Technogenic and Natural Security in Ukraine in 2003". Kiev., 2004, 435 p.; National report "On Technogenic and Natural Security in Ukraine in 2005". Kiev., 2006, 375 p.

National report "On Technogenic and Natural Security in Ukraine in 2009". Kiev., 2010, 252 p.

O'Neill R.V., Gardner R.H., Barnthouse L.W., Suter G.W. II, Hildebrand S.G., Gehrs C.W. Ecosystem risk analysis: a new methodology. // Environ Toxicol Chem 1982, - 1:167-177

Rachev S., Martin R. D., Racheva B., Stoyanov S. Stable Expected Tail Loss (ETL) Optimal Portfolios & Extreme Risk Management // FinAnalytica Foundation Paper, – 2009, 34 p.

Rachev S., Menn C. and Fabozzi F. J. *Fat Tailed and Skewed Asset Return Distributions: Implications for Risk*, Wiley-Finance, 2005

Raiffa Howard, Schlaifer Robert, Applied Statistical Decision Theory. MIT Press, Massachusetts Institute of Technology, Cambridge, Massachusetts and London, England, - 1968,- 356 p.

Romdhani S., Gong S., Psarrou A. A multi-view nonlinear active shape model using kernel PCA // Proceedings of BMVC, Nottingham, UK, - 1999, pp. 483–492

Scheolkopf B., Smola A.J., Muller K. Nonlinear component analysis as a kernel eigenvalue problem // Neural Computation, **10**, #5, 1998, pp. 1299–1399

Sklar A. Random Variables, Distribution Functions, and Copulas – a Personal Look Backward and Forward // in Ruschendorff et. al. (Eds.) *Distributions with Fixed Marginals and Related Topics*, Institute of Mathematical Sciences, Hayward , CA, - 1996

State Budget of USSR in 1981 - 1985. The Statistical Book. URSS Ministry of Finances. Moscow, "Finances and Statistics" Publishing, - 1987, 215 p.

State Budget of USSR in 1989. Brief Statistical Book. URSS Ministry of Finances, Central Department of State Budget. Moscow, "Finances and Statistics" Publishing, - 1989, 161 p.

US EPA (Environmental Protection Agency). Advice to EPA on the science and application of ecological risk assessment in environmental decision making: a report of the U.S. EPA Science Advisory Board. Science Advisory Board, Washington, 2008, DC: EPA-SAB-08-002

USSR National Economy in 1969. The Statistical Book. Central Administration of Statistics at Cabinet of Ministries of USSR., Moscow, "Statistics" Publishing, - 1970, 840 p.

USSR National Economy in 1980. The Statistical Book. Central Administration of Statistics at Cabinet of Ministries of USSR., Moscow, "Finances and Statistics" Publishing, - 1981, 577 p.

Villez K., Ruiz M., Sin G., Colomer J., Rosen C. and Vanrolleghem P. A. Combining multiway principal component analysis (MPCA) and clustering for efficient data mining of historical data sets of SBR processes // Water Science & Technology, **57**, 2008, # 10, pp. 1659 - 1666

Warga J. Optimal Control of Differential and Functional Equations. Academic Press, New York and London, 1972, - 531 p.

White M.L., Maurice C.G., Mysz A. et al. The critical ecosystem assessment model (CrEAM) -identifying healthy ecosystems for environmental protection planning. // In: Campbell, JC; Jones, KB; Smith, H. eds. North America Land Cover Summit, Chapter 12. Washington, DC: Association of American Geographers; 2008, - pp. 181−213

A RSM Method for Nonlinear Probabilistic Analysis of the NPP Reinforced Concrete Structures

Juraj Králik
Department of Structural Mechanics, University of Technology in Bratislava, Bratislava 813 68, Slovakia, juraj.kralik@stuba.sk

Abstract: This paper describes the reliability analysis of a concrete containment for NPP under a high internal overpressure. There is showed summary of calculation models and calculation methods for the probability analysis of the structural integrity considering degradation effects and high internal overpressure. The probabilistic safety assessment (PSA) level 3 aims at an assessment of the probability of the concrete structure failure under the excessive overpressure. The non-linear analysis of the concrete structures was considered. The constitutive model presented is a further extension of the smeared crack model. Following the experimental results of Vecchio, Collins and Červenka and Kupfer a new concrete cracking layered finite shell element was developed and implemented into the ANSYS system. In this model the Kupfer's bidimensional failure criterion of concrete is considered. The uncertainties of the loads level (long-time temperature and dead loads), the material model (concrete cracking and crushing, behavior of the reinforcement and liner), degradation effects and other influences following from the inaccuracy of the calculated model and numerical methods were taken into account in the response surface method (RSM). The results of the reliability analysis of the NPP structures are presented.

Keywords: Nuclear Power Plant; Concrete Failure; Probability; RSM; ANSYS.

1. Introduction

The International Atomic Energy Agency set up a program to give guidance to its member states on many aspects of the safety of nuclear power reactors. The resistance of the building structure must be checked for extreme steam pressure in the case of small or medium-sized accidents, such as a LOCA (Loss of Coolant Accident) or a HELB (High Energy Line Break) or a SBLA (Steam Line Break Accident) on the different primary loop piping system. A complex high confidence and low probability analysis of power plant buildings with a WWER 230 reactor meets all the requirements of the IAEA. Compliance with the IAEA (IAEA, 2008) and Eurocode 2 (Eurocode, 1990) will be considered in three load combinations: NOC (Normal Conditions), DBA (Design Basic Accident) and BDBA (Beyond Design Basic Accident). During the nineties the requirements for service safety of the NPP Type V1 increased. In 1994 the Slovak Nuclear Inspection Authority (Kralik, 2009) defined the conditions for the general reconstruction of the NPP V1 hermetic zone. This project was realized in the frame of the program PHARE "Reconstruction project of NPP V1" in cooperation with companies SIEMENS KWU, VÚEZ Levice and VÚJE Trnava (Kralik, 2002). Moreover, the principal tasks for checking the capacity and integrity of the hermetic zone after an accident involving the pipeline cooling system were defined. The object of NPP V1 is rectangular in plan (Figure 1). The part in the modul V-G/10-12 has dimension 13.7/39.2 m and in part V-G/12-17 has 25.0/39.2 m.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

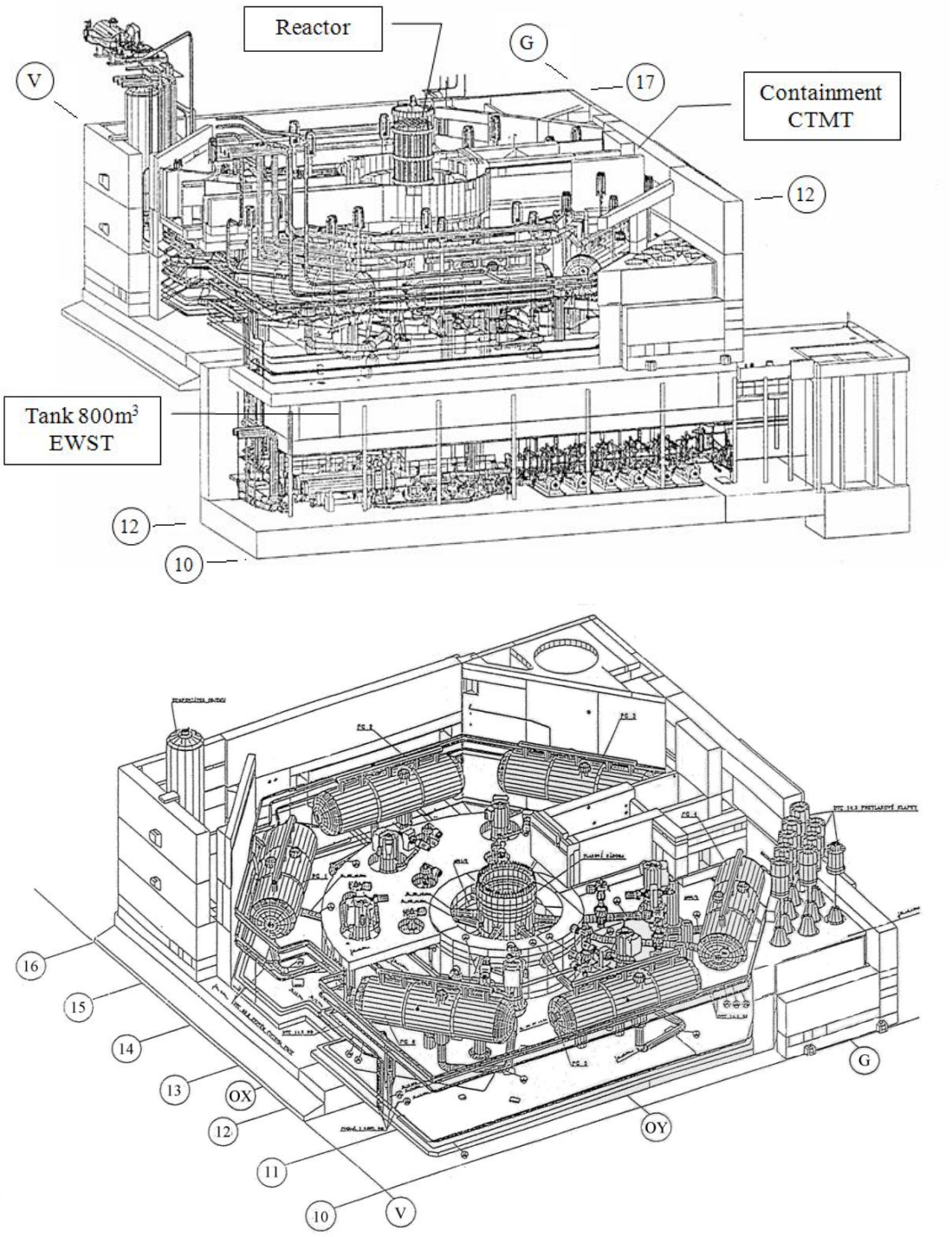


Figure 1. Containment and emergency tank 800 m^3 of NPP V1.

The floor of the steam box is at level 10.5 m. The exterior walls are with thickness about 1.0−1.7 m. The foundation conditions under the NPP V1 structure are also complicated. The greater part of the structure (modulus V-D/12-20) is located on a reinforced concrete foundation plate at depths of − 4.2 m and – 7.3 m under the surface level. The reactor itself and the building sector in the modulus V-G/10-12 under the emergency basin is located at depths of – 12.3 m and – 13.8 m. The complicated wall configuration inside the hermetic zone provides more possibilities for the occurrence of local peak strain into the spoil liaison wall and panel.

The general purpose of the reconstruction of NPP V1 (Králik, 2002) was to increase the reliability and safety of service in the hermetic zone by establishing a new condensation system using the space of an 800 m^3 tank with its cooling water for cutting down the radioactivity of escaped steam and its chilling from any accident that occurred and depressurization in the hermetic zone (Králik, 2002). The new cooling system includes piping systems and backward shutters that were installed in the newly-built holes of the tank ceiling plate in modulus 10/G (Figure 1). Therefore, we have to check the resistance and reliability of the upgraded structures through calculations and an experimental strength test. From 1995−96, the new technology system of the bearing structures of the hermetic zone were tested, and in accordance with a reconstruction project from 1999−2000, the bearing frames of the hermetic space have been checked after the application of a new safety system for accidents and after reconstruction of the bearing structure.
In the case of the analysis PSA 3 level it is necessary to determine the probability of the concrete structure failure under higher overpressure (Králik, 2009).

The general purpose of the probabilistic analysis of the containment integrity failure was to define the critical places of the structure elements and to estimate the structural collapse (Králik, 2005, 2009, 2010).

Two critical structures are in the NPP object – the containment (CTMT) and the emergency water safety tank (EWST). The foundation conditions under the NPP V1 structure are also complicated.

The complicated wall configuration inside the hermetic zone provides more possibilities for the occurrence of local peak stress along the contact of walls and plates. The computational FEM model of the NPP V1 structures is presented in Figure 2.

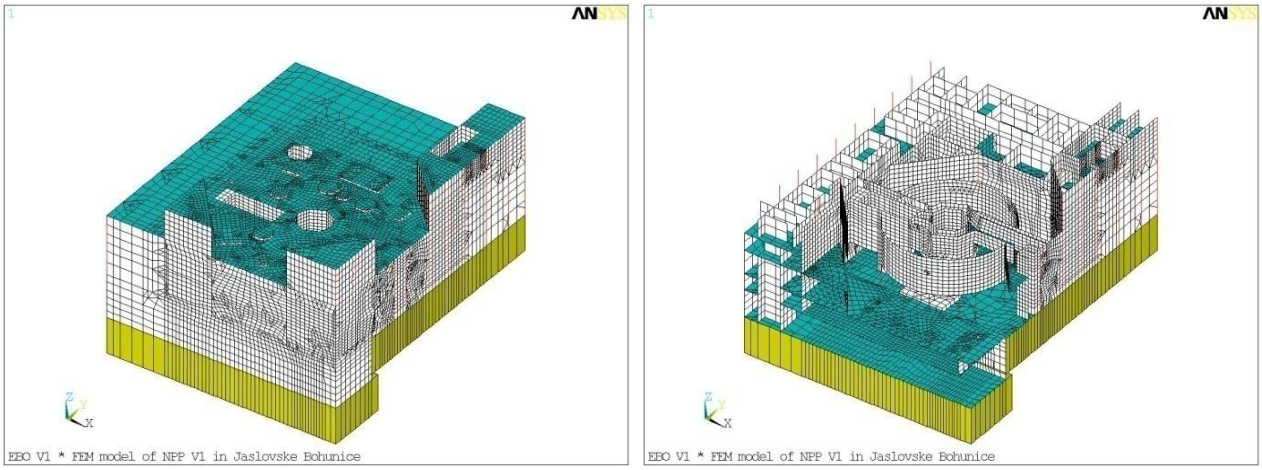


Figure 2. Computational model of the NPP building with layered shell elements.

For a complex analysis of the reinforced concrete structure of the hermetic zone for different kinds of loads, the ANSYS software and the CRACK program (Králik, 2009, 2010) were provided to solve this task. The building of the power block was idealized with a discrete model consisting of 26 923 elements with 325 036 DOF. The link finite elements and the infinite layered space elements developed by the author (Králik, 2009) were used to model the soil. The link finite elements for the model of the thin soil layer under the power block building loaded by a steam pressure are accurate enough, and thus create a more realistic model. Recently the soil under the foundation plate has been consolidated.

The accident scenario was defined by SIEMENS KWU, VÚEZ Tlmače and VÚJE Trnava within the Phare program and "The NPP V1 Reconstruction Project" (Králik, 2002).

2. Nonlinear Solution of Concrete Cracking and Crushing

The probabilistic analysis of the containment integrity failure is based on the nonlinear analysis of the concrete structures due to the accident of the coolant system and under the high level of the overpressure into the box of the steam generator (Králik, 2009, 2010).

The theory of large strain and rate independent plasticity were proposed during the high overpressure loading using the SHELL91 or the SHELL281 layered shell element from the ANSYS library (Kohnke, 2008).

The vector of the deformation parameters $\{r\}$ of this element (Figure 3) with the corner nodes "1, 2, 3, 4" and midside nodes "5, 6, 7, 8" is defined in the form

$$\{r\}=\{r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8\}^{\mathrm{T}}, \qquad \{r_i\}=\{u_{xi}, u_{yi}, u_{zi}, \theta_{xi}, \theta_{yi}\}^{\mathrm{T}} \qquad \forall\, i = 1..8 \tag{1}$$

The vector of the displacement of the l-th shell layer $\left\{u^l\right\}=\left\{u_x^l, u_y^l, u_z^l\right\}^T$ is approximated by the quadratic polynomial (Králik, 2009) in the form

$$\left\{u^l\right\}=\begin{Bmatrix} u_x^l \\ u_y^l \\ u_z^l \end{Bmatrix}=\sum_{i=1}^{8} N_i.\begin{Bmatrix} u_{xi} \\ u_{yi} \\ u_{zi} \end{Bmatrix}+\sum_{i=1}^{8} N_i.\frac{\zeta.t_i}{2}.\begin{bmatrix} a_{1,i} & b_{1,i} \\ a_{2,i} & b_{2,i} \\ a_{3,i} & b_{3,i} \end{bmatrix}.\begin{Bmatrix} \theta_{xi} \\ \theta_{yi} \end{Bmatrix} \tag{2}$$

where N_i is the shape function for i-th node of the 8-node quadrilateral shell element, u_{xi}, u_{yi}, u_{zi} are the motion of i-th node, ζ is the thickness coordinate, t_i is the thickness at i-th node, $\{a\}$ is the unit vector in x direction, $\{b\}$ is the unit vector in plane of element and normal to $\{a\}$, θ_{xi} or θ_{yi} are the rotations of i-th node about vector $\{a\}$ or $\{b\}$.

The linear strain vector $\left\{\varepsilon^l\right\}$ for the l-th layer is related to the nodal displacement vector by

$$\left\{\varepsilon^l\right\}=\left[B^l\right]\left\{u^l\right\}, \tag{3}$$

where $\left[B^l\right]$ is the strain-displacement matrix based on the element shape functions.

In the case of the elastic state the stress-strain relations for the l-th layer are defined in the form

$$\left\{\sigma^l\right\}=\left[D_e^l\right]\left\{\varepsilon^l\right\} \tag{4}$$

where $\left\{\varepsilon^l\right\}^T=\left\{\varepsilon_x,\varepsilon_y,\gamma_{xy},\gamma_{yz},\gamma_{zx}\right\}$ and $\left\{\sigma^l\right\}^T=\left\{\sigma_x,\sigma_y,\tau_{xy},\tau_{yz},\tau_{zx}\right\}$ and the matrix of the material stiffness

$$\left[D_e^l\right]=\begin{bmatrix} B^l E_x^l & B^l \mu_{xy}^l E_x^l & 0 & 0 & 0 \\ B^l \mu_{xy}^l E_x^l & B^l E_y^l & 0 & 0 & 0 \\ 0 & 0 & G_{xy}^l & 0 & 0 \\ 0 & 0 & 0 & \dfrac{G_{yz}^l}{k_s} & 0 \\ 0 & 0 & 0 & 0 & \dfrac{G_{zx}^l}{k_s} \end{bmatrix}$$

where $B^l=\dfrac{E_y^l}{E_y^l-\left(\mu_{xy}^l\right)^2 E_x^l}$, E_x^l (versus E_y^l) is Young modulus of the *l*-th layer in the direction *x* (versus *y*), G_{xy}^l, G_{yz}^l, G_{zx}^l are shear moduli of the *l*-th layer in planes XY, YZ and ZX; k_s is the coefficient of the effective shear area ($k_s=1+0.2\dfrac{A}{25t^2}\geq 1.2$), *A* is the element area, *t* is the element thickness.

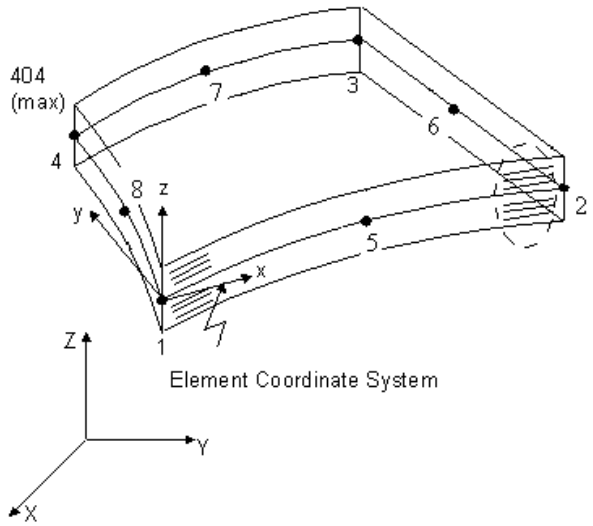


Figure 3. The shell element with 8 nodes.

2.1. GEOMETRIC NONLINEARITY

If the rotations are large but the mechanical strains (those that cause stresses) are small, then a large rotation procedure can be used. A large rotation analysis is performed in a static analysis in the ANSYS program [10].

The strain in the *n*-step of the solution can be computed from the relations

$$\{\varepsilon_n\} = [B_o][T_n]\{u_n\}, \tag{5}$$

where $\{u_n\}$ is the deformation displacement, $[B_o]$ is the original strain-displacement relationship, $[T_n]$ is the orthogonal transformation relating the original element coordinates to the convected (or rotated) element coordinates.

The convected element coordinate frame differs from the original element coordinate frame by the amount of rigid body rotation. Hence $[T_n]$ is computed by separating the rigid body rotation from the total deformation $\{u_n\}$ using the polar decomposition theorem. A corotational (or convected coordinate) approach is used in solving large rotation/small strain problems (Kohnke, 2008).

2.2. MATERIAL NONLINEARITY

The presented constitutive model is a further extension of the smeared crack model (Bažant et al, 2007; Červenka, 1985; Hinton and Owen, 1984; Hughes, 1984; Meskouris et al, 1997; Oñate et al, 1993), which was developed in (Králik, 2000). Following the experimental results (Červenka, 1985; Kupfer, 1969; Jerga and Križma, 2006) a new concrete cracking layered finite shell element was developed and incorporated into the ANSYS system (Králik, 2009). The layered approximation and the smeared crack model of the shell element are proposed.

The processes of the concrete cracking and crushing are developed during the increasing of the load. The concrete compressive strength f_c, the concrete tensile strength f_t and the shear modulus G are reduced after the crushing or cracking of the concrete (Kolmar, 1986).

In this model the stress-strain relation is defined (Figure 4) following ENV 1992-1-1 (1991)

- Loading in the compression region $\varepsilon_{cu} < \varepsilon^{eq} < 0$

$$\sigma_c^{ef} = f_c^{ef}.\frac{k.\eta - \eta^2}{1.+(k-2).\eta}, \quad \eta = \frac{\varepsilon^{eq}}{\varepsilon_c} \quad (\varepsilon_c \doteq -0.0022, \quad \varepsilon_{cu} \doteq -0.0035) \tag{6}$$

- Softening in the compression region $\varepsilon_{cm} < \varepsilon^{eq} < \varepsilon_{cu}$

$$\sigma_c^{ef} = f_c^{ef}.\left(1 - \frac{\varepsilon^{eq} - \varepsilon_c}{\varepsilon_{cm} - \varepsilon_{cu}}\right) \tag{7}$$

- The tension region $\varepsilon_t < \varepsilon^{eq} < \varepsilon_m$

$$\sigma_c^{ef} = f_t.\exp(-2.(\varepsilon^{eq} - \varepsilon_t)/\varepsilon_{tm}) \qquad (\varepsilon_t \doteq 0.0001, \quad \varepsilon_{tm} \doteq 0.002) \tag{8}$$

In the case of the plane state the strength function in tension f_t and in compression f_c were considered equivalent values f_t^{eq} and f_c^{eq}.

In the plane of principal stresses (σ_{c1}, σ_{c2}) the relation between the one and bidimensional stresses state due to the plasticity function by Kupfer (see Figure 5) can be defined as follows:

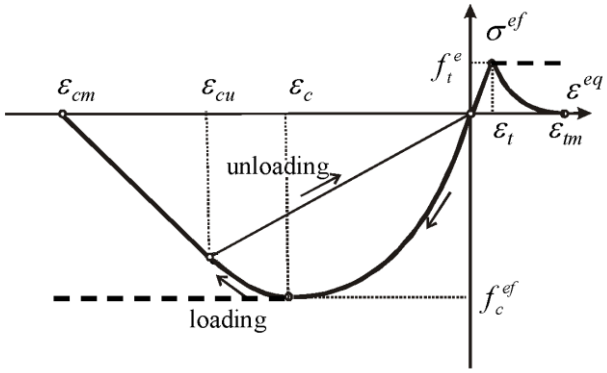


Figure 4. The concrete stress-strain diagram.

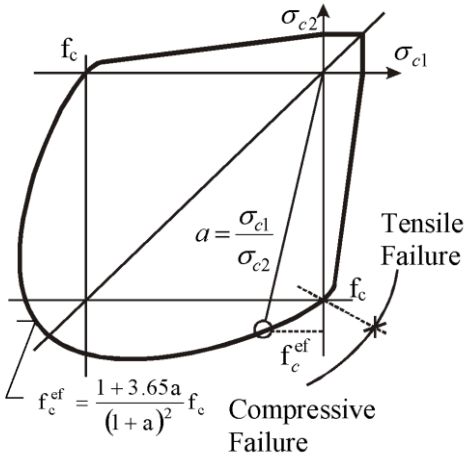


Figure 5. Kupfer's plasticity function.

- Compression-compression

$$f_c^{ef}=\frac{1+3.65.a}{(1+a)^2}.f_c, \qquad a=\frac{\sigma_{c1}}{\sigma_{c2}} \tag{9}$$

- Tension-compression

$$f_c^{ef}=f_c.r_{ec}, \quad r_{ec}=\left(1+5.3278\frac{\sigma_{c1}}{f_c}\right), \quad r_{ec}\geq 0.9 \tag{10}$$

- Tension-tension

$$f_t^{ef}=f_t.r_{et}, \quad r_{et}=\frac{A+(A-1).B}{A.B}, \quad B=K.x+A, \quad x=\sigma_{c2}/f_c, \tag{11}$$

$$r_{et}=1. \Leftrightarrow x=0, \quad r_{et}=0.2 \Leftrightarrow x=1.$$

The shear concrete modulus G was defined for cracking concrete by Kolmar (Kolmar, 1986) in the form

$$G=r_g.G_o, \quad r_g=\frac{1}{c_2}\ln\left(\frac{\varepsilon_u}{c_1}\right), \quad c_1=7+333(p-0.005), \quad c_2=10-167(p-0.005), \tag{12}$$

where G_o is the initial shear modulus of concrete, ε_u is the strain in the normal direction to crack, c_1 and c_2 are the constants dependent on the ratio of reinforcing, p is the ratio of reinforcing transformed to the plane of the crack ($0<p<0.02$).

It is proposed that the crack in the one layer of shell element is oriented perpendicular to the orientation of principal stresses. The membrane stress and strain vector depends on the direction of the principal stress and strain in one layer

$$\{\varepsilon_{cr}\}=[T_{\varepsilon}]\{\varepsilon\}, \qquad \{\sigma_{cr}\}=[T_{\sigma}]\{\sigma\}, \tag{13}$$

where $[T_{\varepsilon}]$, $[T_{\sigma}]$ are transformation matrices for the principal strain and stress in the direction θ in the layer.

$$[T_{\varepsilon}]=\begin{bmatrix} \cos^2\theta & \sin^2\theta & \sin\theta\cos\theta & 0 & 0 \\ \sin^2\theta & \cos^2\theta & -\sin\theta\cos\theta & 0 & 0 \\ -2\sin\theta\cos\theta & 2\sin\theta\cos\theta & \cos 2\theta & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & 0 & -\sin\theta & \cos\theta \end{bmatrix} \tag{14}$$

$$[T_{\sigma}]=\begin{bmatrix} \cos^2\theta & \sin^2\theta & 2\sin\theta\cos\theta & 0 & 0 \\ \sin^2\theta & \cos^2\theta & -2\sin\theta\cos\theta & 0 & 0 \\ -\sin\theta\cos\theta & \sin\theta\cos\theta & \cos 2\theta & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & 0 & -\sin\theta & \cos\theta \end{bmatrix}$$

The strain-stress relationship in the Cartesian coordinates can be defined in dependency on the direction of the crack (in the direction of principal stress, versus strain)

$$[\sigma_{cr}]=[D_{cr}]\{\varepsilon_{cr}\} \quad \text{and therefore} \quad [\sigma]=[T_{\sigma}]^{\mathrm{T}}[D_{cr}][T_{\varepsilon}]\{\varepsilon\} \tag{15}$$

For the membrane and bending deformation of the reinforced concrete shell structure the layered shell element, on which a plane state of stress is proposed on every single layer, was used.
The stiffness matrix of the reinforced concrete for the *l*-th layer can be written in the following form

$$\left[D_{cr}^{l}\right]=\left[T_{c.\sigma}^{l}\right]^{\mathrm{T}}\left[D_{cr}^{l}\right]\left[T_{c.\varepsilon}^{l}\right]+\sum_{s=1}^{N_{rein}}\left[T_{s}^{l}\right]^{\mathrm{T}}\left[D_{s}^{l}\right]\left[T_{s}^{l}\right] \tag{16}$$

where $[T_{c.\sigma}]$, $[T_{c.\varepsilon}]$, $[T_s]$ are the transformation matrices for the concrete and the reinforcement separately, N_{rein} is the number of the reinforcements in the *l*-th layer.

After cracking the elasticity modulus and Poisson's ratio are reduced to zero in the direction perpendicular to the cracked plane, and a reduced shear modulus is employed. Considering 1 and 2 two principal directions in the plane of the structure, the stress-strain relationship for the concrete *l*-th layer cracked in the 1-st direction, is

$$\begin{Bmatrix}\sigma_1\\ \sigma_2\\ \tau_{12}\\ \tau_{13}\\ \tau_{23}\end{Bmatrix}_l = \begin{bmatrix}0 & 0 & 0 & 0 & 0\\ 0 & E & 0 & 0 & 0\\ 0 & 0 & G_{12}^{cr} & 0 & 0\\ 0 & 0 & 0 & G_{13}^{cr} & 0\\ 0 & 0 & 0 & 0 & G_{23}^{cr}\end{bmatrix}_l \begin{Bmatrix}\varepsilon_1\\ \varepsilon_2\\ \gamma_{12}\\ \gamma_{13}\\ \gamma_{23}\end{Bmatrix}_l \tag{17}$$

where the shear moduli are reduced by the coefficient of the effective shear area k_s and parameter r_{g1} by Kolmar (12) as follows: $G_{12}^{cr} = G_o.r_{g1}$, $G_{13}^{cr} = G_o.r_{g1}$, $G_{23}^{cr} = G_o / k_s$

When the tensile stress in the 2-direction reaches the value f_t', the latter cracked plane perpendicular to the first one is assumed to form, and the stress-strain relationship becomes :

$$\begin{Bmatrix}\sigma_1\\ \sigma_2\\ \tau_{12}\\ \tau_{13}\\ \tau_{23}\end{Bmatrix}_l = \begin{bmatrix}0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & G_{12}^{cr}/2 & 0 & 0\\ 0 & 0 & 0 & G_{13}^{cr} & 0\\ 0 & 0 & 0 & 0 & G_{23}^{cr}\end{bmatrix}_l \begin{Bmatrix}\varepsilon_1\\ \varepsilon_2\\ \gamma_{12}\\ \gamma_{13}\\ \gamma_{23}\end{Bmatrix}_l \tag{18}$$

where the shear moduli are reduced by the parameter r_{g1} and r_{g2} by Kolmar (12) as follows: $G_{12}^{cr} = G_o.r_{g1}$, $G_{13}^{cr} = G_o.r_{g1}$, $G_{23}^{cr} = G_o r_{g2}$. The cracked concrete is anisotropic and these relations must be transformed to the reference axes XY. The simplified averaging process is more convenient for finite element formulation than the singular discrete model. A smeared representation for the cracked concrete implies that cracks are not discrete but distributed across the region of the finite element.

The smeared crack model (Červenka, 1985; Oñate, 1993), used in this work, results from the assumption, that the field of more micro cracks (not one local failure) brought to the concrete element will be created. The validity of this assumption is determined by the size of the finite element, hence its characteristic dimension $L_c = \sqrt{A}$, where A is the element area (versus integrated point area of the element). For the expansion of cracking the assumption of constant failure energies G_f = *const* is proposed in the form

$$G_f = \int_0^{\infty} \sigma_n(w)\,\mathrm{d}w = A_G.L_c, \qquad w_c = \varepsilon_w.L_c, \tag{19}$$

where w_c is the width of the failure, σ_n is the stress in the concrete in the normal direction, A_G is the area under the stress-strain diagram of concrete in tension. Concrete modulus for descend line of stress strain diagram in tension (crushing) can be described according to Oliver (Červenka, 1985; Oñate, 1993) in dependency on the failure energies in the form

$$E_{c,s} = \frac{E_c}{1-\lambda_c}, \qquad \lambda_c = \frac{2G_f E_c}{L_c.\sigma_{\max}^2}, \tag{20}$$

where E_c is the initial concrete modulus elasticity, σ_{max} is the maximal stress in the concrete tension. From the condition of the real solution of the relation (20) it follows, that the characteristic dimension of element must satisfy the following condition

$$L_c \le \frac{2G_f E_c}{\sigma_{\max}^2}, \tag{21}$$

The characteristic dimension of the element is determined by the size of the failure energy of the element. The theory of a concrete failure was implied and applied to the 2D layered shell elements SHELL91 and SHELL281 in the ANSYS element library (Kohnke, 2008).

The limit of damage at a point is controlled by the values of the so-called crushing or total failure function F_u. The modified Kupfer's condition (Kupfer, 1969) for the l-th layer of section is following

$$F_u^l = F_u^l\left(I_{\varepsilon 1}; I_{\varepsilon 2}; \varepsilon_u\right) = 0, \qquad F_u^l = \sqrt{\beta\left(3J_{\varepsilon 2}\right) + \alpha I_{\varepsilon 1}} - \varepsilon_u = 0, \tag{22}$$

where $I_{\varepsilon 1}$, $J_{\varepsilon 2}$ are the strain invariants, and ε_u is the ultimate total strain extrapolated from uniaxial test results ($\varepsilon_u = 0.002$ in the tension domain, or $\varepsilon_u = 0.0035$ in the compression domain), α, β are the material parameters determined from the Kupfer's experiment results ($\beta = 1.355,\ \alpha = 0.355\varepsilon_u$).

The failure function of the whole section will be obtained by the integration of the failure function through to the whole section in the form

$$F_u = \frac{1}{t}\int_0^t F_u^l\left(I_{\varepsilon 1}; I_{\varepsilon 2}; \varepsilon_u\right)\mathrm{d}z = \frac{1}{t}\sum_{l=1}^{N_{lay}} F_u^l\left(I_{\varepsilon 1}; I_{\varepsilon 2}; \varepsilon_u\right) t_l \tag{23}$$

where t_l is the thickness of the l-th shell layer, t is the total shell thickness and N_{lay} is the number of layers. This failure condition is determined by the maximum strain ε_s of the reinforcement steel in the tension area ($\max\left(\varepsilon_s\right) \le \varepsilon_{sm} = 0.01$) and by the maximum concrete crack width w_c ($\max\left(w_c\right) \le w_{cm} = 0.3\mathrm{mm}$).

3. Degradation of Reinforced Concrete Structure

The safety of nuclear power plants could be affected by the age related degradation of structures (ACI, 1989) if it is not detected prior to the loss of the functional capability and if timely corrective action is not taken.

The reduction even or the loss of functional capability of the key plant components could reduce the plant safety. Mild steel reinforcing bars are provided to control the extent of cracking and the width of cracks at operating temperatures, to resist tensile stresses and computed compressive stresses for elastic design, and to provide the structural reinforcement where required. Potential causes of degradation of the reinforcing steel would be corrosion, exposure to elevated temperatures and irradiation (ACI, 1989; Králik, 2009; Li, 2005; Meskouris and Wittek, 1997; Teplý et al., 2010).

So-called „uniform" or general corrosion consists of approximately uniform loss of metal over the whole exposed surface of the reinforcing bar. Faraday's law indicates that a corrosion current density of

i_{corr} = 1 μA cm^{-2} corresponds to a uniform corrosion penetration of 11.6 μm year^{-1}. Thus the reduction of the diameter of a corroding bar, ΔD, at time T, can be estimated directly (in mm) from i_{corr} as

$$\Delta D(T) = 0.0232 \int_{T_o}^{T} i_{corr}(t)\,\mathrm{d}t\,, \tag{24}$$

where T is the actual time, T_o is the time of corrosion initiation (in years). If a constant annual corrosion rate is assumed, equation (24) reduces to the following equation

$$\Delta D(T) = 0.0232(T - T_o) i_{corr}$$

The net cross-sectional area of a reinforcing bar, A_r at time T, is then equal to

$$A_r(T) = \begin{cases} \dfrac{\pi D_o^2}{4}, & T \le T_o \\ \dfrac{\pi\left[D_o - \Delta D(T)\right]^2}{4}, & T > T_o \end{cases}, \tag{25}$$

where D_o is the initial diameter of the reinforcing bar (in mm).

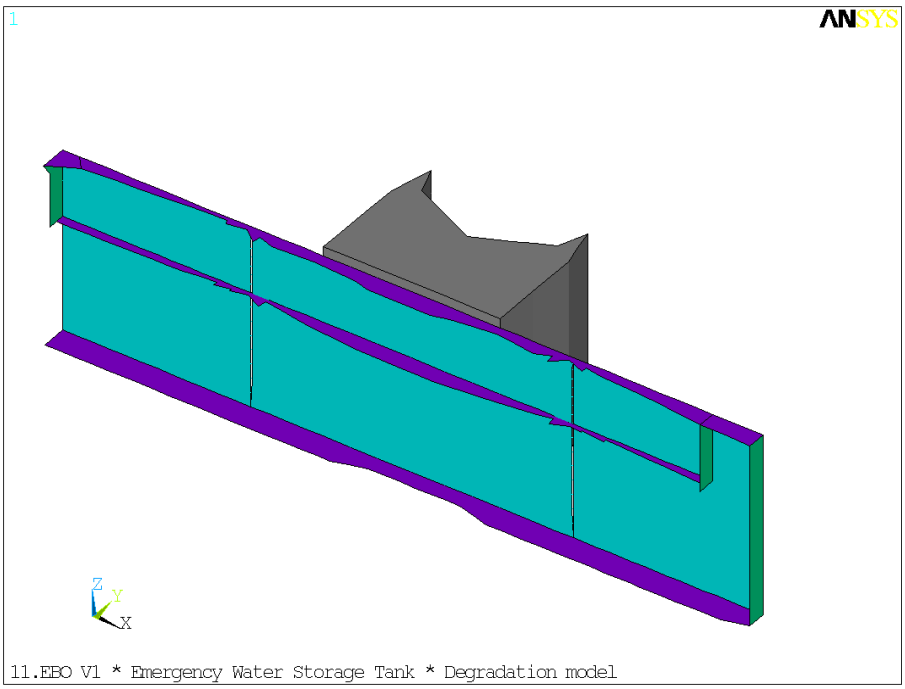


Figure 6. Computational model with two vertical cracks.

The limit concrete strength and its Young modulus are practically not changed during the corrosion process. The corrosion process can be considered as the reduction of the reinforced steel cross section. The

state of the NPP reinforced concrete structures has been periodically monitored in the frame of the IAEA requirements for safety and reliability of the NPP performance.

Special attention was paid to the EWST structure. After 10 years of the operation, two vertical cracks in the wall in axis 12, near the contact of the reactor corps and the basin wall, were identified (Figure 6). Also, hence the degradation effects at the bottom of the emergency water storage tank were identified (Figure 7). These effects were the consequence of the corrosion process of the reinforced steel in the basin plate.

Figure 7. Degradation effects at the basin bottom.

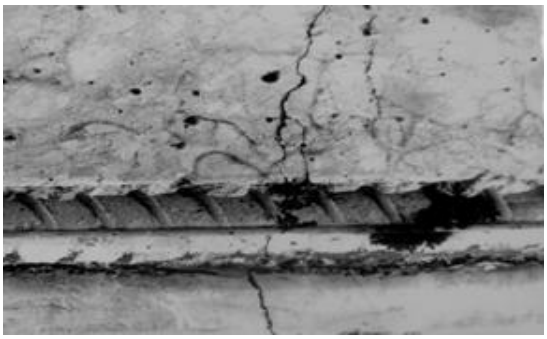


Figure 8. Detail of reinforcement and the concrete crack.

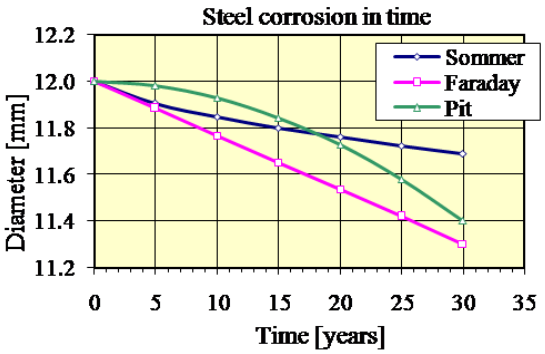


Figure 9. Steel corrosion depending on time.

There is proposed conservatively that this corrosion process had been acting from 10 to 20 years. In regard to all uncertainties to define the influence of the corrosion effects and to identify cracks the state of the NPP structure resistance is investigated using the probabilistic methodology and conservative proposition of the degradation processes.

The corrosion effects of the reinforcement at the basin bottom and in the concrete cracks were considered according to Faraday's law (Figure 9) using the uniform corrosion penetration (of 11.6 μm.year^{-1}). This assumption is also conservative from the point of view of the higher safety level (ACI, 1989).

4. Nonlinear Deterministic Analysis

The critical sections of the structure were determined on the base of the nonlinear analysis due to the monotone increasing of overpressure inside the hermetic zone (Králik and Cesnak, 2000; Králik, 2002). The resistance of these critical sections was considered taking into account the design values of the material characteristics and the load. The combination load and design criteria were considered for the BDBA state (IAEA, 2008).

The critical areas were identified in the connection walls and plate of the hermetic zone at level + 10.5 m near the hole in the modulus „10" and „V" (Figure 10). The tension forces and the bending moments were concentrated between two outside large walls.

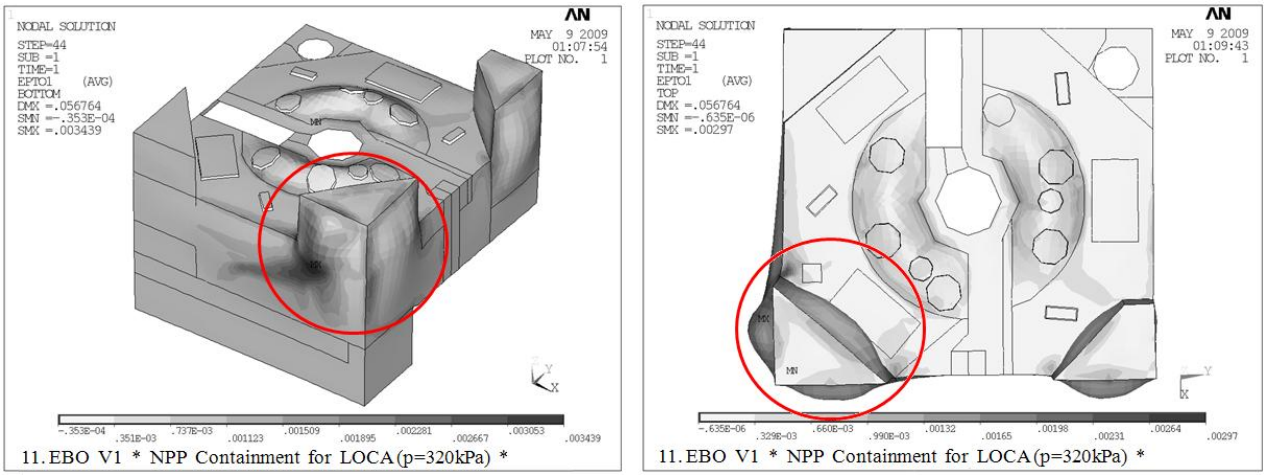


Figure 10. Critical area of containment structure loading at overpressure 320 kPa.

On the base of the nonlinear analysis of the containment resistance for median values of the material properties and failure function (22) the critical overpressure was equal to 309 kPa without the degradation effect and 287 kPa (versus 283 kPa) with degradation effects during 10 years (versus 20 years) (Figure 11a).

The capacity of the concrete wall of the EWST structure is equal 45% without considering the degradation effect (versus 50% with the degradation effect after 20 years) under overpressure 300kPa in the

CTMT structure (Figure 11b). The temperature of the water (about 55°C) in the emergency tank (EWST) affects the compression in the tank wall as the prestressed. This effect is opposite to the overpressure effects in the CTMT. The degradation effect occurs during overpressure higher than 150 kPa (Figure 11).

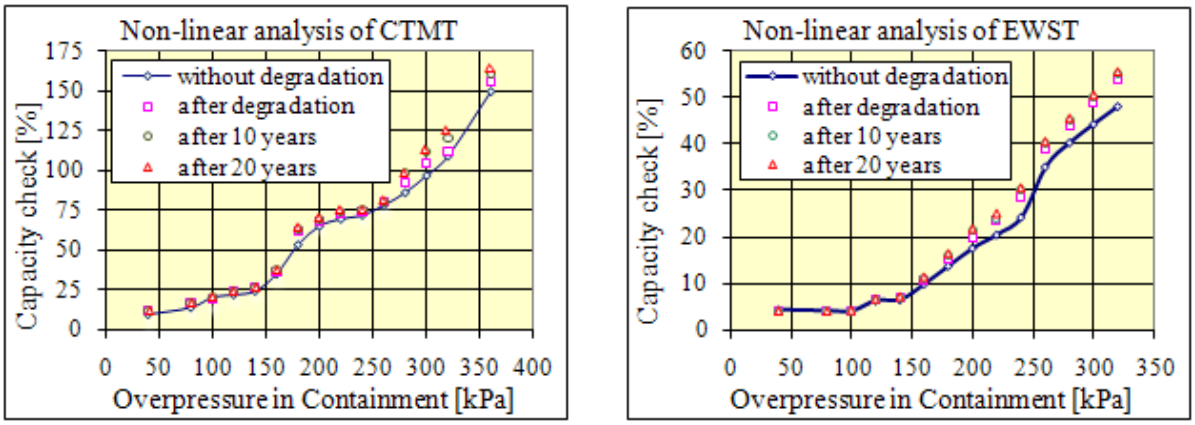


Figure 11. The capacity utilization of the reinforced concrete containment (CTMT) and emergency water storage tank (EWST) due to overpressure with and without degradation effects.

5. Probabilistic Analysis of the Structure Reliability

Recent advances and the general accessibility of information technologies and computing techniques give rise to assumptions concerning the wider use of the probabilistic assessment of the reliability of structures through the use of simulation methods in the world (Haldar and Mahadevan, 2000; Králik, 2009; Lenkei and Györgyi, 1999; Vejvoda, 2003;). A great attention should be paid to using the probabilistic approach in an analysis of the reliability of structures (Bažant et al., 2005; Holický and Marková, 2000; Kala, 2011; Melcher et al., 2004; Vořechovský and Chudoba, 2006).

Most problems concerning the reliability of building structures are defined today as a comparison of two stochastic values, loading effects E and the resistance R, depending on the variable material and geometric characteristics of the structural element. The variability of those parameters is characterized by the corresponding functions of the probability density $f_R(r)$ and $f_E(e)$. In the case of a deterministic approach to the design the deterministic (nominal) attributes of those parameters R_d and E_d are compared.
The deterministic definition of the reliability condition is of the form

$$R_d \geq E_d \tag{26}$$

and in the case of the probabilistic approach it is of the form

$$RF = g(R,E) = R - E \geq 0 \tag{27}$$

where $g(R,E)$ is the reliability function.

The probability of failure can be defined by the simple expression

$$P_f = P\left[R < E\right] = P\left[\left(R - E\right) < 0\right] \tag{28}$$

The reliability function *RF* can be expressed generally as a function of the stochastic parameters X_1, X_2 to X_n, used in the calculation of *R* and *E*.

$$RF = g(X_1, X_2, ..., X_n) \tag{29}$$

The failure function $g(\{X\})$ represents the condition (reserve) of the reliability, which can be either an explicit or implicit function of the stochastic parameters and can be single (defined on one cross-section) or complex (defined on several cross-sections, e.g., on a complex finite element model).
In the case of simulation methods the failure probability is calculated from the evaluation of the statistical parameters and theoretical model of the probability distribution of the reliability function $Z = g(X)$. The failure probability is defined as the best estimation on the base of numerical simulations in the form

$$p_f = \frac{1}{N}\sum_{i=1}^{N} I\left[g\left(X_i\right) \le 0\right] \tag{30}$$

where *N* in the number of simulations, $g(.)$ is the failure function, $I[.]$ is the function with value 1, if the condition in the square bracket is fulfilled, otherwise is equal 0.

The RSM method was chosen for the PSA analysis of the containment safety. It is based on the assumption that it is possible to define the dependency between the variable input and the output data through the approximation functions in the following form:

$$Y = c_o + \sum_{i=1}^{N} c_i X_i + \sum_{i=1}^{N} c_{ii} X_i^2 + \sum_{i=1}^{N-1}\sum_{j>i}^{N} c_{ij} X_i X_j \tag{31}$$

where c_o is the index of the constant member; c_i are the indices of the linear member and c_{ij} the indices of the quadratic member, which are given for predetermined schemes for the optimal distribution of the variables or for using the regression analysis after calculating the response. Approximate polynomial coefficients are given from the condition of the error minimum, usually by the "Central Composite Design Sampling" (CCD) method or the "Box-Behnken Matrix Sampling" (BBM) method (Kohnke, 2008).

The computation efficiency of the experimental design depends on the number of design points, which must be at least equal to the number of the unknown coefficients. In the classical design approach, a regression analysis is carried out to formulate the response surface after calculating the responses at the sampling points. These points should have at least 3 levels for each variable to fit the second-order polynomial, leading to 3^k factorial design. This design approach becomes inefficient with the increasing of the number of random variables. More efficient is the central composite design, which was developed by Box and Wilson (Kohnke, 2008).

The central CCD method is composed of (Figure 13a) :

1. Factorial portion of design − a complete 2^k factorial design (equal − 1, + 1)
2. Center point − n_o center points, $n_o \ge 1$ (generally $n_o = 1$)
3. Axial portion of design − two points on the axis of each design variable at distance α from the design center

Then the total number of design points is $N = 2^k + 2k + n_o$, which is much more than the number of the coefficients $p = (k + 1)(k + 2)/2$. The graphical representation for $k = 3$ and the matrix form of the coded values are represented in Figure 13.

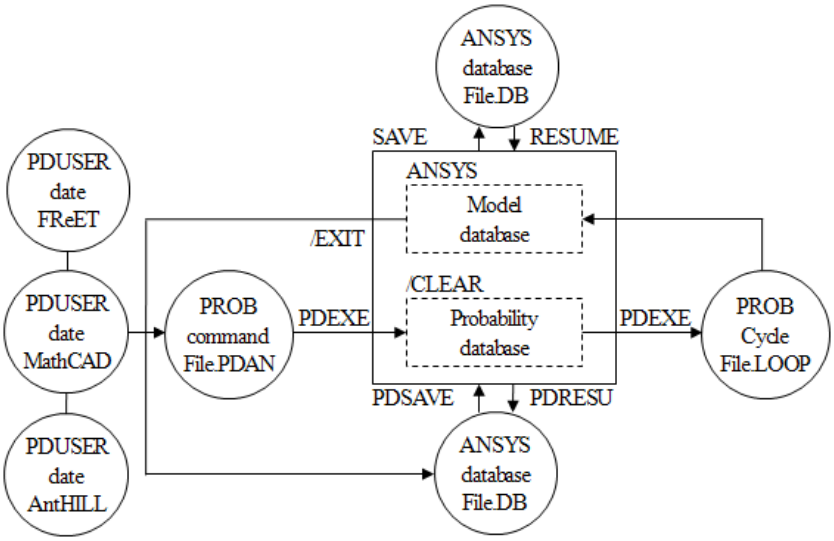
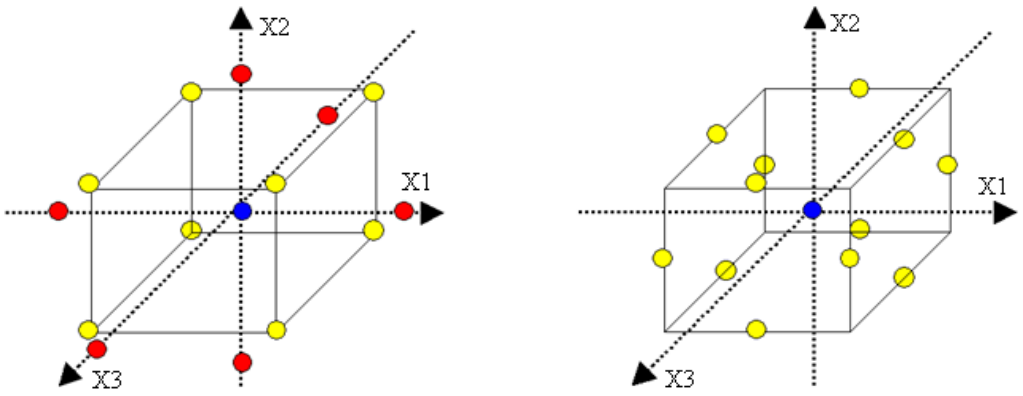


Figure 12. A procedural diagram of the probabilistic calculations using the ANSYS software system.

It is advisable to use the displacement-based FEM for reliability analysis of the complicated structures with one of the defined simulation methods. In this work the ANSYS licensed program [10] with a probabilistic postprocessor was utilized for the probability analysis of the reliability of the NPP structures for various action effects. In Figure 13, the procedural diagram sequence is presented from the structure of the model through the calculations, up to the evaluation of the probability of the structural failure.



a) The Central Composite Design b) The Box-Behnken Matrix Design

Figure 13. Distribution schemes of the stochastic numbers of the RSM method for three input variables

The postprocessor for the probabilistic design of structures enables to define the random variables using the standard distribution functions (normal, lognormal, exponential, beta, gamma, Weibull, etc.), or externally (user-defined sampling) using other statistical programs as the AntHILL or the FREET. The probabilistic calculation procedures are based on the Monte Carlo simulations (DS, LHS, user-defined sampling) and the approximation RSM method (CCD, BBM, user-defined sampling) [15 and 17]. The RSM

method generates the explicit performance function for the implicit or complicated limit state function. This method is very effective for robust and complicated tasks.

On the base of experimental design, the unknown coefficients are determined due to the random variables selected within the experimental region. The uncertainty in the random variables can be defined in the model by varying in the arbitrary amount producing the whole experimental region.

6. PSA Level 3 Analysis of Containment Failure

The methodology of the probabilistic analysis of integrity of reinforced concrete structures of containment results from requirements (IAEA, 2008) and experience from their applications (Králik, 2009, 2010 and 2011; Lenkei and Györgyi, 1999, Melcher et al., 2004; Rosowsky, 1999; Vejvoda, 2003; Vořechovský and Chudoba, 2006).

The probability of containment failure is calculated from the probability of the reliability function *RF* in the form,

$$P_f = P(RF < 0) \tag{32}$$

where the reliability condition *RF* is defined depending on a concrete failure condition (30)

$$RF = 1 - F_u(J_{\varepsilon 1}; J_{\varepsilon 2}; \varepsilon_u)/\ \varepsilon_u, \tag{33}$$

where the failure function $F_u(.)$ was considered in the form (23).

The previous design analyses, calculations and additions include various uncertainties, which determine the results of probability bearing analysis of containment structural integrity are presented in Table 1. Due to the mentioned uncertainties of the input data for the probabilistic analysis of reinforced concrete containment structures loss of integrity the mean values and standard deviations, the variable parameters for normal, lognormal and beta distribution were determined.

Table I Variable parameters of the input data

	Soil	Material		Loads			Model	
	Stiffness	Young Modulus	Dead load	Live load	Pres-sure	Tempe-rature	Action uncertaint	Resist. uncertaint
Characteristic value	kz_k	E_k	G_k	Q_k	P_k	T_k	Te_k	Tr_k
Variable	kz_{var}	e_{var}	g_{var}	q_{var}	p_{var}	t_{var}	Te_{var}	Tr_{var}
Histogram type	N	LN	N	BETA	N	BETA	N	N
Mean value μ	1	1	1	0.643	1	0.933	1	1
Deviation σ [%]	5	11.1	10	22.6	8	14.1	5	5
Minimum value	0.754	0.649	0.621	0.232	0.662	0.700	0.813	0.813
Maximum value	1.192	1.528	1.376	1.358	1.301	1.376	1.206	1.206

On the base of the RSM simulations the increment vector of the deformation parameters $\{\Delta r_s\}$ in the FEM is defined for the *s*-th simulation in the form

$$\{\Delta r_s\} = \left[K_{GN}\left(E_s, kz_s, F_\sigma\right)\right]^{-1}\left\{\Delta F\left(G_s, Q_s, P_s, T_s\right)\right\} \tag{34}$$

and the strain vector increment

$$\{\Delta\varepsilon_s\} = [B_s]\{\Delta r_s\} \tag{35}$$

where $[K_{GN}]$ is the nonlinear stiffness matrix depending on the variable parameters E_s, kz_s and F_σ, F_σ is the Kupfer's yield function defined in the stress components, $\{\Delta F\}$ is the increment vector of the general forces depending on the variable parameters G_s, Q_s, P_s and T_s for the *s*-th simulation. The total strain vector is defined as the sum of the strain increments.

Resulting from the variability of the input quantity 25 simulation steps on the base of the RSM method under the ANSYS-CRACK system were realized (Králik, 2005 and 2009). The probability of loss containment structure integrity was calculated from 10^6 Monte Carlo simulations for 25 steps of the RSM approximation method on the full structural FEM model. The probability analysis was considered for the structural model without (model V1_10) and with (model V1_11) the wall cracking and the corrosion effects below the emergency tank (Figure 1).

$$\{\varepsilon_s\} = \sum_{istep=1}^{Nstep} \{\Delta\varepsilon_s\}_{istep} \tag{36}$$

The evaluation of the probabilistic sensitivities was calculated from the correlation coefficients between all random input variables and a particular random output parameter by Spearman. These analyses show (Figure 14) that the variability of the overpressure and the structure stiffness has the fundamental impact upon the reliability of the containment. The effects of the variability of the concrete stiffness are dominant in the model V1_10 without the degradation effect (Figure 14a), and on the other hand, the variability of the overpressure is dominant in the model V1_11 with the degradation effect (Figure 14b).

The probability of the concrete structure failure in accordance with the relation (33) under overpressure 320 kPa is less than 10^{-6} in the model without cracking effects (original status). If the influence of the tank wall cracking and the corrosion effects are considered the probability of failure is equal $1.375.10^{-4}$ for overpressure 320 kPa. The histograms of the reliability function *RF* under overpressure 320 kPa for two models without (V1_10) and with (V1_11) the degradation effect, respectively, are presented in Figure 15.

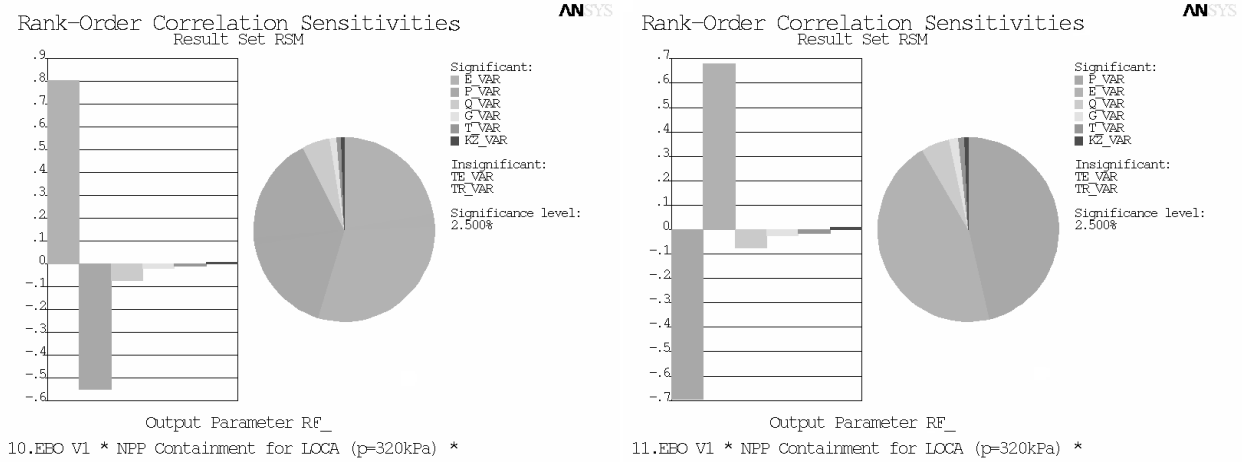


Figure 14. Sensitivity of reliability function RF.

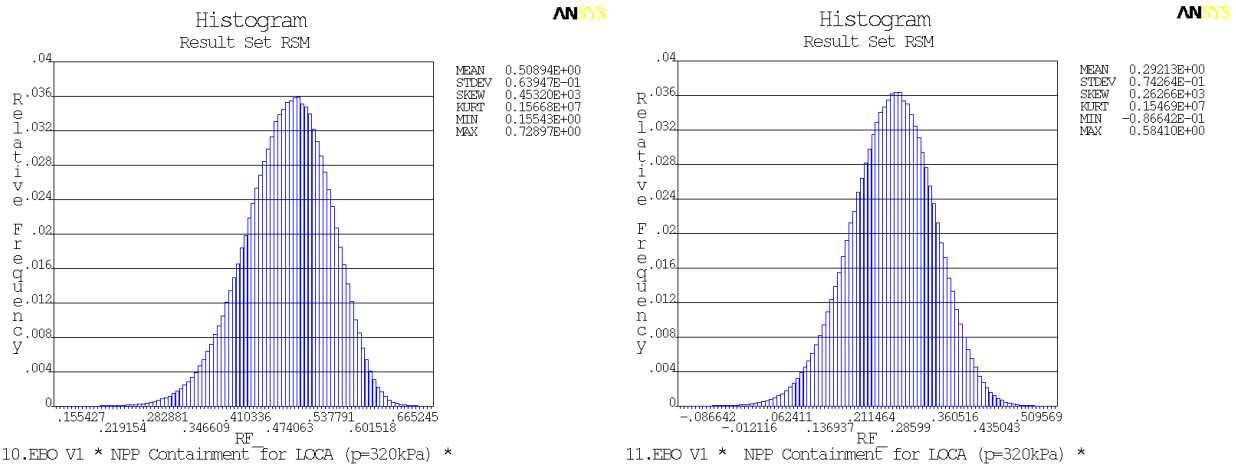


Figure 15. Reliability function RF under overpressure 320 kPa for two models: a) without the degradation effect, b) with the degradation effect

7. Conclusions

The probability analysis of the loss of the concrete containment integrity was made for the overpressure loads from 40 kPa to 320 kPa using the nonlinear solution of the static equilibrium considering the geometric and material nonlinearities of the reinforced concrete shell layered elements. The nonlinear analyses were performed in the CRACK program, which was developed by the author and implemented into the ANSYS system (Králik and Cesnak, 2000; Králik, 2009). The uncertainties of the loads level (longtime temperature and dead loads), the material model of the composite structure (concrete cracking and crushing, reinforcement, and liner), the degradation effects (carbonization and reinforcement corrosion) and other influences following from the inaccuracy of the calculated model and the numerical methods were taken into account in the Monte Carlo simulations on the base of the RSM method (Králik, 2009). The reliability function RF was defined in dependency on the failure function $F_u(I_{\varepsilon 1}; J_{\varepsilon 2}; \varepsilon_u)$ for requirements of the PSA analysis in the form (23). The probability of the loss of the concrete containment integrity is less than 10^{-6} for the original structural model. In the case of the degradation effects of the concrete structure under the emergency tank the probability of the containment failure is equal to $1{,}375.10^{-4}$ for the overpressure 320 kPa. The theory of the nonlinear analysis using the RSM method was developed in the framework of the VEGA grant project (Králik, 2009).

Acknowledgements

The project was performed with the financial support of the Grant Agency of the Slovak Republic (VEGA). The project registration number is VEGA 1/0849/08 and VEGA 1/1039/12.

References

Bažant, Z. P. Pang, S. D. Vořechovský, M. and Novák. D. Energetic-statistical size effect simulated by SFEM with stratified sampling and crack band model. *International Journal for Numerical Methods in Engineering* (Wiley), 71(11):1297-1320, 2007. also a Report, 2005, Dept. of Civil Engrg., Northwestern University, Evanston, Illinois.

ACI Committee E 222. *Corrosion of Metals in Concrete*, ACI 222R-89, American Concrete Institute, Detroit, MI. 1989.

Červenka, V. Constitutive Model for Cracked Reinforced Concrete, *ACI Journal 82* (1985) 877.

Eurocode 2. *Design of Concrete Structures, Part 1*, CEC 1990.

Haldar A. and Mahadevan S. *Probability, Reliability and Statistical Methods in Engineering Design*, John Wiley & Sons, New York, 2000.

Hinton E. and Owen D. R. J. *Finite Element Software for Plates and Shells*, DCE University College Swansea U. K., Pineridge Press Swansea, 1984.

Holický M. and Marková J. *Base of reliability theory and risk evaluation*, ČVUT Praha, 2005. (in Czech)

Hughes, T. J. R. *Numerical Implementation of Constitutive Models: Rate-Independent Deviatoric Plasticity, Theoretical Foundation for Large-Scale Computations for Nonlinear Material Behavior*, eds. S. Nemat-Nasser, R. J. Asaro, G. A. Hegemier, Martinus Nijhoff Publishers, Dordrecht, 1984.

IAEA. *Development and Application of Level 2 Probabilistic Safety Assessment for Nuclear Power Plants, Draft Safety Guide* DS393, Draft 6, 2008.

Jerga J. and Križma M.: Assessment of Concrete Damage, *Building Research Journal 54* (2006) 211.

Kala, Z. Sensitivity analysis of steel plane frames with initial imperfections, Engineering Structures, 33, 8, pp.2342-2349, 2011.

Kohnke P. *ANSYS, Theory*, SAS IP Inc. Canonsburg, 2008.

Kolmar, W. *Beschreibung der Kraftuebertragung ueber Risse in nichtlinearen Finite-Element-Berechnungen von Stahlbeton-tragwerken*, PhD Thesis, T.H. Darmstadt, Darmstadt, 1986.

Králik J., Cesnak J.: Nonlinear Analysis of Power Plant Buildings with the VVER 230 Reactor After a Loss off Coolant Accident, *Slovak Journal of Civil Engineering 8* (2000) 18.

Králik, J. Experimental and Numerical Analysis of Reinforced Concrete Containment after Loss of Coolant Accident, In: *International Symposium on Mechanics of Composites*, ed. M. Černý, CTU in Prague, Klokner Institute, Prague 2002, p. 91.

Králik, J. Probability Nonlinear Analysis of Reinforced Concrete Containment Damage due to High Internal Overpressure, Engineering Mechanics 12 , pp.113, 2005.

Králik, J. Comparison of Probabilistic Methods Efficiency to Solve the Structure Reliability in FEM, In: *IX. Int. Conf. on Reliability of Structures*, pp.41, Prague 2008.

Králik, J. *Reliability Analysis of Structures Using Stochastic Finite Element Method*, Edition STU Bratislava 2009.

Králik, J. *Safety and Reliability of Nuclear Power Buildings in Slovakia. Earthquake-Impact-Explosion*, Edition STU Bratislava 2009.

Králik, J. A RSM Method for Probabilistic Nonlinear Analysis of Reinforced Concrete Bubbler Tower Structure Integrity. In proc. *European Safety and Reliability Conference, ESREL 2009, Reliability, Risk and Safety, Theory and Applications*, CRC Press/A.Balkema Book, Taylor&Francis Group, Prague, Czech Republic, 7-10 September, Vol.2, p.1369-1372, 2009.

Králik, J. Probabilistic Safety Analysis of the Nuclear Power Plants in Slovakia. In: *Journal of KONBiN, Safety and Reliability Systems*, Ed. VERSITA Central European Science Publishers, Warszawa, No 2,3 (14, 15) 2010, pp. 35-48.

Králik, J. A RSM Method for Nonlinear Probabilistic Analysis of the Reinforced Concrete Structure Failure of a Nuclear Power Plant – Type VVER 440, In: *Engineering Mechanics*, Ed. AEM Prague, Vol. 18, No. 1, p.3-22, 2011.

Kupfer, H. Hilsdorf, H. K. Ruesch, H. Behavior of Concrete Under Biaxial Stresses, *Journal ACI 66*, pp.656,1969.

Lenkei, P. and Györgyi, J. *Probabilistic Structural Analysis of Nuclear Containment under LB LOCA*. Report JE Paks, N-Quad Ltd. 1999, Pécs.

Li, C. Q. Time Dependent Reliability Analysis of the Serviceability of Corrosion Affected Concrete Structures, International *Journal of Materials&Structural Reliability 3*, pp.105, 2005.

Melcher, J., Kala, Z., Holický, M., Fajkus, M. & Rozlívka, L. Design characteristics of structural steels based on statistical analysis of metallurgical products. *Journal of Constructional Steel Research*, 60, 3-5, pp.795-808, 2004.

Meskouris, K. and Wittek, U. *Aspects in Modern Computational Structural Analysis*, A.A. Balkema Publishers, Rotterdam, 1997.

Oñate, E. Oller, S. Oliver, J. Lubliner, J. A Constitutive Model for Cracking of Concrete Based on the Incremental Theory of Plasticity, *Engineering Computation* 5, pp. 309, 1993.

Rosowsky, D.V. Structural Reliability. Part of "*Structural Engineering Handbook*", ed. Chen Wai-Fah, Boca Raton, CRC Press LLC, 1999.

Teplý, B. Chromá, M. Rovnaník, P. Durability Assessment of Concrete Structures: Reinforcement Depassivation due to Carbonation, *Structure and Infrastructure Engineering*, pp. 317, 6, 2010.

Vejvoda, S. Keršner, Z. Novák, D. Teplý, B. Probabilistic Safety Assessment of the Steam Generator Cover, In*: SMiRT17, 17th International Conference on Structural Mechanics in Reactor Technology*, Prague 2003, CD-ROM.

Vořechovský, M. and Chudoba, R. Stochastic modeling of multi-filament yarns: II. Random properties over the length and size effect. *International Journal of Solids and Structures* (Elsevier), 43(3-4):435-458, 2006.

Application of the Direct Optimized Probabilistic Calculation

Martin Krejsa
Department of Structural Mechanics, Faculty of Civil Engineering, VSB - Technical University Ostrava, Ludvika Podeste 1875/17, 708 33 Ostrava - Poruba, Czech Republic, martin.krejsa@vsb.cz .

Abstract. The Direct Optimized Probabilistic Calculation method - DOProC - deals with probabilistic tasks where certain input quantities are of a random nature. DOProC is typically used in probabilistic reliability assessment of load carrying structures. DOProC can be also employed in probabilistic designs of structural elements with the specified reliability. In many cases, this calculation method is very efficient and provides accurate estimates of resulting probabilities.

DOProC has been successfully applied, among others, in the probabilistic calculation of fatigue cracks in steel structures and bridges which are subject to cyclical loads. The software used for that purpose, FCProbCalc, makes it possible to monitor efficiently and operatively development of fatigue damage to the structure and to specify times for service inspection. This means, the structure is compliant and well suited for operation in terms of fatigue damage. The methods and application can considerably improve estimation of maintenance costs for the structures and bridges subject to cyclical loads.

Keywords: Direct Optimized Probabilistic Calculation; DOProC; Safety Margin; Probability of Failure; Fatigue Crack Propagation; Inspection of Structure; Random Variable.

1. Introduction

Many calculation methods exist now for the designing and reliability assessment of load carrying structures and elements with the specified reliability. Those methods are based on the probability theory and mathematic statistics. They have been becoming more and more popular. The methods which are referred to as probabilistic or stochastic make it possible to analyze safety margin Z defined by a calculation model where at least some input quantities are of a random nature (Rackwitz et al., 1978). The calculation procedures contribute to a qualitatively higher level of the reliability assessment and, in turn, higher safety of those who use the buildings and facilities (Melchers, 1994).

The most frequently used and most numerous group of the computational method comprises the simulation methods which are based on the popular simulation technique - Monte Carlo (Bjerager, 1988; Hurtado et al., 1998) or any advanced or stratified simulation methods, for instance, Adaptive Sampling (Bucher, 1988), LHS (Helton et al., 2003; Olsson et al., 2003) and Importance Sampling. Eurocodes which are in force now mention the application of approximation methods FORM and SORM types (Der Kiureghian et al., 1998; Zhao et al., 2001) which are used mostly for calibration of partial factors. Those methods are also used in rather complex software applications based on Finite Elements Methods (FEM) (Vanmarcke et al., 1986; Reh et al., 2006; Stefanou, 2009).

The probabilistic approach to the assessment and designing of the structures has started appearing in practice recently only (Bergmeister et al., 2009). The pre-requisite is, however, a sufficient database of input

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

quantities including the experience from practical operation because many input quantities cannot be based on models and laboratory measurements only (this being the case, for instance, of geotechnics). Those computational methods are used, in particular, when designing the load-carrying systems for civil engineering structures and bridges (Kala, 2005; Krivy et al., 2007), where degradation processes in structures can be also taken into account (Konecny et al., 2007; Straub, 2009). It is possible to carry out the Performance-Based Design of structures which consider utility values such as durability, fire resistance, insulation or seismic resistance (Kralik et al., 2009; Teply et al., 2010). The probabilistic approach is used also in risk engineering (Kubecka, 2010). Stochastic models are being developed which describe interaction of building structures with subsoil or which taken into account undermining effects (Marschalko et al., 2011).

This paper describes the use of the original method and method which is under development now: the Direct Optimized Probabilistic Calculation (in short "DOProC") uses a purely numerical approach without any simulation techniques. This provides more accurate solutions to probabilistic tasks, and, in some cases, to considerably faster completion of computations. Such solution entails a small numerical error only and minor inaccuracies, the reason being discretizing of input and output quantities. In case of the probabilistic assessment of the reliability of structures, DOProC expresses directly the probability of failure p_f, which can be compared with the design value of nominal failure probability pd, defined in standards and regulations in force. Where the failure probability p_f is zero (the structure is excessively reliable) or equal to one (the input quantities in any combination result in a failure), DOProC estimates the result immediately and no probabilistic calculations are needed anymore. In this case, DOProC method represents a very suitable and highly efficient solution (Krejsa, 2011).

Theoretical background of DOProC was described in detail in many publications (Janas et al., 2009). DOProC can be used now to solve efficiently a number of probabilistic computations. It has been used, for instance, in probabilistic assessment of combined load, reliability of cross-sections and systems consisting of statically determined or undetermined load carrying constructions, in probabilistic assessment of load carrying constructions which are subject to impact, in probabilistic analyses of steel-fibre reinforced concrete mixtures or in probabilistic assessment of reliability of anchored reinforcement or arc reinforcement in underground and long mine works with a special focus on anti-slipping properties.

DOProC has proved to be a good solution, among others, in probabilistic analyses of fatigue crack propagation in constructions subject to cyclical loads. Detailed methods with examples of the probabilistic assessment for a construction subject to fatigue load are available, a particular attention being paid to cracks from the edge and those from surface. Similarly to other probabilistic analyses, this information is used as a basis for proposing a system of inspections of the cyclic load construction (Moan, 2004; Chen et al., 2011; Li et al., 2011). In order to improve quality of probabilistic calculations, a special software - FCProbCalc - was developed. Using this software, the task can be solved flexibly in a user-friendly environment.

2. Direct Optimized Probabilistic Calculation

The Direct Optimized Probabilistic Calculation ("DOProC") has been under development since 2002. The calculation procedure for a certain task in DOProC is clearly determined by its algorithm, while Monte Carlo simulation methods generate calculation data for simulation on a random basis. The term in the name of the method - "the optimized" - is based on following facts: the number of variables that enter calculation of the failure probability p_f, computation is, however, limited by capabilities of the software to process the

application numerically. If there are too many random variables, the application is extremely time demanding - even if high-performance computers are used. Therefore, efforts have been made to optimize calculations in order to reduce the number of operations, keeping, at the same time, reliable calculation results. Currently, the DOProC along with the optimizing steps can address well several probabilistic tasks.

2.1. BASICS OF DOPROC METHOD

Similarly as with the other probabilistic methods, input random quantities in DOProC (such as the load, geometry, material properties, or imperfections) are described using the non-parametric (empirical) distribution in histograms. This technique can be also used for parametric divisions. The distribution is typically based on observations, being often long-lasting ones. A computational procedure is being developed now, the aim being to implement into DOProC the statistic dependence of input parameters, such as (Vorechovsky et al., 2009).

The basic computation algorithm of DOProC is based on general terms and procedures used in probabilistic theories. Let the histogram B be an arbitrary function f of histograms A_j where j ranges from 1 to n. Then:

$$B = f(A_1, A_2, A_3, \ldots, A_j, \ldots, A_n) \; . \tag{1}$$

Each histogram A_j consists of i_j interval where each interval is limited with $a_{j,i}$ from below and $a_{j,i+1}$ from above. This means, that for the interval $i_j = 1$, the values will be as follows:

$$a_{j,1} \leq a_j < a_{j,2} \; , \tag{2}$$

where

$$a_{j,2} = a_{j,1} + \Delta a_j \; , \tag{3}$$

where

$$\Delta a_j = \frac{a_{j,\max} - a_{j,\min}}{i_j} \; . \tag{4}$$

In i_j, following formula is valid:

$$a_{j,i} \leq a_j < a_{j,i+1} \; . \tag{5}$$

Let us express a_j in that interval as $a_j^{(ij)}$. Similar relations are valid for the B histogram. If there are i intervals, the values of the histogram in the i^{th} interval range from b_i to b_{i+1}, this means $b^{(i)}$. They can be expressed as follows:

$$b^{(i)} = f\left(a_1^{(i1)}, a_2^{(i2)}, \ldots, a_j^{(ij)}, \ldots, a_n^{(in)}\right) \tag{6}$$

for the specific combination of arguments: $a_1^{(i1)}, a_2^{(i2)}, \ldots, a_j^{(ij)}, \ldots, a_n^{(in)}$. The same value - $b^{(i)}$ - can be derived for other values too (or at least for some values too) - $a_j^{(ij)}$. If the potential combination of values $a_j^{(ij)}$ is marked as l, the following general formula can be derived:

$$b^{(i)} = f\left(a_1^{(i1)}, a_2^{(i2)}, \ldots, a_j^{(ij)}, \ldots, a_n^{(in)}\right)_l \; . \tag{7}$$

The probability p_{bl}^{i} of occurrence of $b^{(i)}$ is the product of $p_{aj}^{(ij)}$ (probabilities of occurrence of a_j^{ij} values). Then:

$$p_{bl}^{i} = \left(p_{aj}^{(i1)} \cdot p_{aj}^{(i2)} \cdot p_{aj}^{(i3)} \cdot \ \ldots \ \cdot p_{aj}^{(ij)} \cdot \ \ldots \ \cdot p_{aj}^{(in)} \right) \ . \tag{8}$$

The probability of occurrence of all potential combinations $\left(a_1^{i1}, a_2^{i2}, \ldots, a_j^{ij}, \ldots, a_n^{in} \right)_l$ of f with the result of $b^{(i)}$ is:

$$p_b^{(i)} = \sum_{l=1}^{l} p_{bl}^{(i)} \ . \tag{9}$$

The number of intervals i_j in each histogram A_j can vary similarly as the number of i intervals in the histogram B. The number of intervals is of utter importance for the number of needed numerical operations and required computing time. On top of this, the accuracy of the calculation depends considerably on the number of intervals.

Fig. 1 shows the numerical operations in the probabilistic calculations with two random quantities expressed in a histogram using the basic computational DOProC algorithm. In this case, two load components are combined or a sum of two histograms is used.

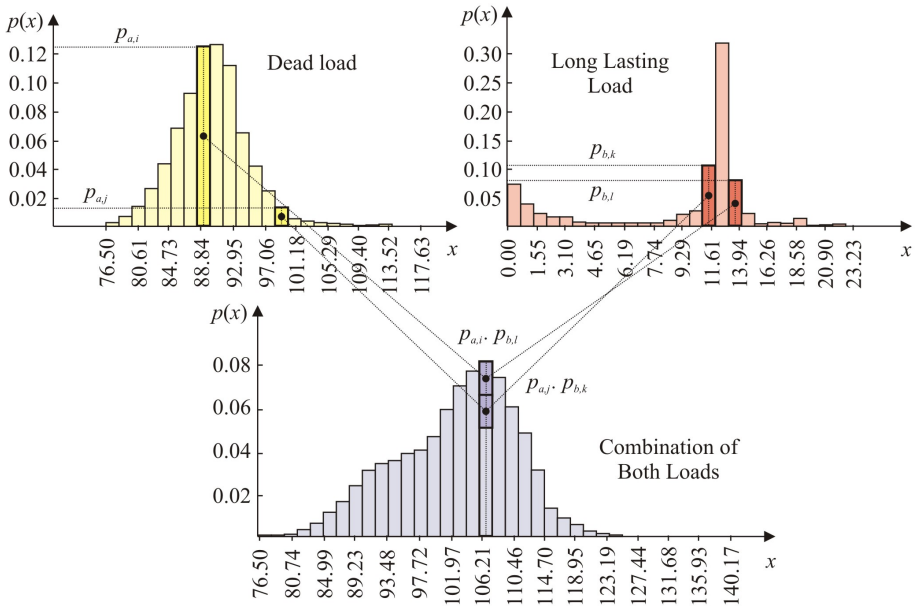


Figure 1. Principles of numerical operations with two histograms (the combination of dead load and long lasting load).

DOProC method is possible to use in ProbCalc (Janas et al., 2009; Janas et al., 2012) - software application which is still under development. It is rather easy and simple to implement quite a complicated analytical transformation model of a probabilistic task defined using a text-oriented editor, similar to Nessus software (Thacker et al., 2006) or Proban software (Tvedt, 2006). In more complex numerical calculation models,

there is a chance to use the procedure programmed by the user as DLL (with a dynamic library extension). More advanced user knowledge is required then to enter the probabilistic tasks in ProbCalc. It is essential to know, at least, general basics of algorithms because this influences the way of defining the computational model and selection of the best optimizing procedure. This weakness is removed if the application software is customized for a specific probabilistic task, this being, for instance, the case of FCProbCalc which is described in Chapter 3.3.

The computational complexity of DOProC is given, in particular, by:

- the number of random input quantities $i = 1 \ldots N$,
- the number of histogram classes (intervals) for each random input quantity n_i,
- complexity of the task (computational model),
- the algorithm used in the probabilistic calculation (the method used for definition of the computational model - in a ProbCalc text mode or using a dynamic library or application software).

2.2. USING DOPROC FOR CALCULATION OF FAILURE PROBABILITY

The construction should be designed in such as way so that the structural resistance R, would be higher than the load effects S. Considering all random phenomena in the load, manufacturing and installation inaccuracies and inaccuracies where the construction is used, the structural resistance R, and load effect S, should be regarded as random quantities - see Fig. 2. The both quantities need to be of the same dimension.

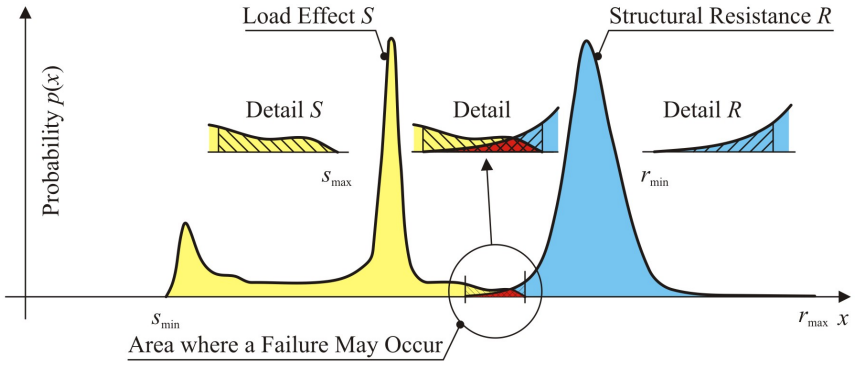


Figure 2. Probability density curves - load effect S, structural resistance R, and the area where a failure may occur.

The probabilistic reliability assessment is based on the reliability condition which can be expressed as follows:

$$Z = R - S \geq 0 \,, \tag{10}$$

where Z is safety margin, R is the structural resistance and S is the load effect. If the reliability condition is not fulfilled, such situation is undesirable in terms of reliability - it is a failure when the load effect S exceeds the magnitude of the structural reliability R. The area where a failure may occur is shown in Fig. 2.

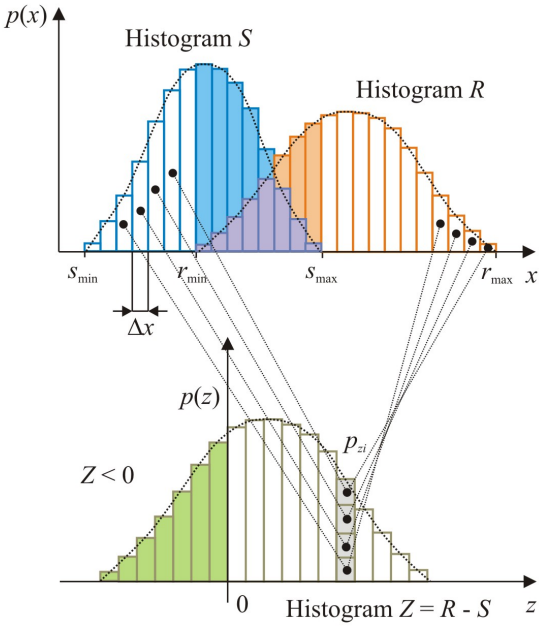


Figure 3. Basic approach to the calculation of the safety margin histogram Z, for two random variables using DOProC.

Fig. 3 shows the calculation of the safety margin Z, for two random quantities using the DOProC algorithm. The probability $p_{Z,i}$ in the i-class is the sum of products of the $p_{s,i}$ probabilities for s_i in i-classes of the S histogram and $p_{r,i}$ probability of r_i in i classes for the R histogram:

$$p_{Z,i} = \sum p_{s,i} \cdot p_{r,i} \ . \tag{11}$$

This results in the histograms for the safety margin Z, the final part of which gives the resulting probability failure p_f, which is compared then with the nominal probability of failure p_d.

If the Z histogram comprises n classes (intervals) with the Δz width, the resulting probability of failure p_f, is calculated then as the sum of probabilities $p_z^{(i=1\ldots j)}$ in individual intervals (classes) where the safety margin is $Z < 0$ (this results from (10)). In the interval where the boundary values of the j class of the Z reliability histograms are within $z_j < 0 < z_{j+1}$, the distribution of probability $p_z^{(j)}$ should be divided proportionally into two parts. This means, the final probability of failure p_f is determined using Fig. 4 and equation:

$$p_f = \sum_{i=1}^{j-1} p_z^{(i)} + p_z^{(j)} \cdot \left(1 - \frac{z_j + \frac{\Delta z}{2}}{\Delta z}\right) = \sum_{i=1}^{j-1} p_z^{(i)} + p_z^{(j)} \cdot \left(\frac{1}{2} - \frac{z_j}{\Delta z}\right) \ . \tag{12}$$

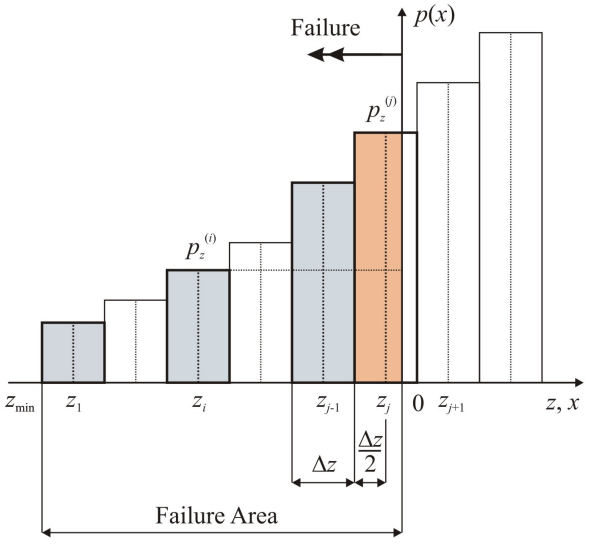


Figure 4. Calculation of the probability of failure p_f, from the histogram for the safety margin Z.

2.3. DOProC optimizing techniques

The purpose of the DOProC optimizing techniques is to minimize the computing time since the algorithm is limited to a certain extent, in particular, for extensive applications where too many simulations exist. If the optimizing techniques are used in DOProC, the failure probability p_f, can be determined in a real time. On top of this, results are reliable and accurate enough even in relatively demanding probabilistic tasks.

The optimizing techniques include:

- **Grouping of variable input quantities:** Grouping of the input quantities makes it possible to eliminate the number of input variable histograms. If possible, the resulting histogram is determined on the basis of the required mathematical operation. Then, the histogram is used for the probabilistic calculation of the model. This can considerably reduce the number of computational operations. This optimizing technique is used most frequently in calculations of the combined load or in a summary histogram which expresses impacts of wind loading by means of a ”wind rose”. If the grouping (this means, the creation of joint histograms of the input quantities) is possible and reliable, it is a very efficient and reasonable optimizing technique which reduces dramatically the number of computational operations in the probabilistic calculation.
- **Interval optimizing:** The objective of the interval optimizing is to minimize the number of classes used in the input quantity histograms. This reduces the number of computation operations and minimizes the machine time needed for the probabilistic calculation. A mandatory condition for this optimizing technique is the maintaining of sufficient accuracy of the required results. For this optimizing technique it is essential to make a sensitivity analysis and to check the influence of such reduction onto the result.
- **Zone optimizing:** In the zone analysis, each input quantity interval is divided into three zones. The first zone is always involved in creation of the probability of failure p_f, irrespective of values in other

histograms (the first zone is involved there, whatever combination of interval of the remaining input quantities is). The second zone may, but does not need to, be involved in the process (it is involved only in some combinations of intervals of the other input quantities), while the third zone is never involved there (when determining the probability of failure p_f, it is possible to omit this part of the histogram). If the zones are known, it is possible to calculate the probability of failure p_f very efficiently. Detailed information is available about the zone optimizing and practical aspects of this approach.

- **Trend optimizing:** Trend optimizing can be used as a supplement to the zone optimizing in the probabilistic calculations. In the zone optimizing technique is used, the calculation is carried out only for the zone #2. If a trend is found for the random variable (this means that the resulting positive value of the safety margin Z, increases with changes in the random variable) it does not make any sense to introduce other computational combinations. For such a quantity, the safety margin Z, cannot reach negative values and cannot influence the failure probability p_f. This means, it is possible to eliminate computational combinations and to keep only those which are really needed.
- **Grouping of partial computation results:** The purpose of the grouping of partial computational results is to decrease the number of computational operations during the assessment of the histograms of the quantities which are the result of the computational model. In case of the probabilistic reliability assessment, this group is defined by the safety margin Z, where the values entered pursuant to (10) are the calculated reliability of the structure R, and loading impacts S. In some cases, it is possible to enter directly the input quantity histogram into this group. (Such quantity can be the strength characteristic of the used material if the reliability assessment is done for the tension and the quantity is not involved in the computational model, or a limit deflection of the reliability assessment is based on the ultimate state of usability).
- **Computation parallelization:** The computation is carried out in several processors or core at the same time. The basic algorithm of DOProC is an optimum solution for the parallelism: partial results reached by multi core/processor computation are summed up in the final phase of the probabilistic calculation.
- **Combination of the aforementioned optimizing techniques:** Below is the recommended sequence of the optimizing techniques in DOProC:
  - Grouping: It should be used always, if possible.
  - Interval optimizing: It is recommended to minimize the number of histogram groups, particularly, when debugging the computational algorithm. Then, the number of the histogram classes should be optimized for specific results.
  - Other optimizing techniques which should be used, if possible and feasible in terms of complexity.

The optimizing techniques have been described in detail and implemented into ProbCalc and can be combined in the probabilistic calculation.

3. Using DOProC to calculate propagation of fatigue cracks

Probabilistic calculation of steel structures and bridges using DOProC method, leads to the probabilities of three basic random events in dependence on years of bridge's operation and fatigue crack propagation. On the basis of that calculation for each individual year, determined by analysis of reliability function, the

dependence of the failure probability on time of the bridge's operation is specified. When the limit reliability is known, it is possible to determine times of the structure's inspections (Krejsa, 2011).

3.1. Basics of propagation of fatigue cracks

Reliability of the load-bearing structure has been significantly influenced by degradation resulting, in particular, from the fatigue of the basic materials. Whler's curves are used when designing such structures. The service life can be limited until a failure occurs. The failure is, however, very difficult to determine. For purposes of the modeling, the amplitude oscillation is considered to be constant, and a certain number of load cycles is taken into account. The method has been developed to provide procedures describing real conditions, all this making the work of design engineers easier. As fatigue cracks appear randomly on existing structures (in crane rails and bridges), it is believed that the designing method is imperfect to a certain extent (Fisher et al., 1998). Methods are under development that would be able to reveal potential defects and damage resulting from initiation cracks that accelerate considerably the propagation of fatigue cracks (Giner et al., 2008). Linear fracture mechanics is among alternative methods. Machinery experts have been dealing with such issues for many years. Results have been gradually taken over and implemented into designs of the loading structures in buildings. This approach is typically used for the determination of times of inspection and analyses of inspection results. If cracks are not found, a conditional probability exists that they might appear later on.

Attention is paid to fatigue damage of building steel structures and bridges where the acceptable fatigue (Anderson, 2005) crack size is assessed. The acceptable crack size plays a key role in degradation of an element dimensioned for an extreme loading combination that is exposed to variable operation loads. It represents a possible degradation of an element in an ultimate limit state that can be still monitored.

The outcome is procedures that should clarify currently acceptable methods used for the designing of the fatigue crack in the context of the safe service life and acceptable failure rate. A flange of the composite reinforced concrete bridge has been chosen for applications of the theoretical solution. This tension is exposed, in particular, to tension. Depending on location of an initial crack, the crack may propagate from the edge or surface. Regarding the frequency, weight and concentration of stresses, those locations rank among those with the major hazard of fatigue cracks appearing in the steel structures and bridges.

3.2. Probabilistic approach to propagation of fatigue cracks

Occurrence of initiation cracks and crack propagation in structures subject to fatigue load has been known for a long time. The process is closely connected with fabrication of the steel structures and, in particular, with creation of details which tend to be damaged by fatigue. The key difference is between initiation of cracks resulting from steelmaking inclusions and those created during fabrication of structural details. Regarding the former, it takes a long time until it reaches the surface, while the latter is at the surface from the beginning of the loading. Standardized approaches of previous EC standards suppose that surface cracks were not present there. The acceptable damage method which is described in the new standard admits random occurrence of surface cracks. The major difference is that a fatigue crack might not be fragile, but could be ductile. In real components of steel structures and bridges, the latter is more frequent that the former which is used in experimental measurements in processed small test-pieces. This fact is not a new phenomenon. It has been known for a long time and has been mentioned, for instance, by (Anderson, 2005).

During the designing, fabrication and processing of details, nobody, however, paid attention to random occurrence of initiation cracks from surface areas (from the surface or from the edge).

Three sizes are important for the characteristics of the propagation of fatigue cracks. These are the initiation size, the detectable size and the final size which occurs prior to failure caused by a fragile or ductile crack. The fatigue crack damage depends on a number of stress range cycles. This is a time factor in the course of reliability for the entire designed service life. In the course of time, the failure rate increases, while the reliability drops.

The topic is discussed in two levels that affect each other: the probabilistic solution to the propagation of the fatigue crack and uncertainties in determination of quantities used in the calculation. When investigating into the propagation, the fatigue crack that deteriorates a certain area of the structure components is described with one dimension only: a. In order to describe the propagation of the crack, the linear elastic fracture mechanics is typically used. It is based on the Paris-Erdogan law (Sanford, 2003):

$$\frac{\mathrm{d}a}{\mathrm{d}N} = C \cdot (\triangle K)^m \ , \tag{13}$$

where C, m are material constants (Carpinteri et al., 2007), a is the crack size and N is the number of loading cycles.

The initial assumption is that the primary design should take into account the effects of the extreme loading resulting from the ultimate state of carrying capacity method. Then, the fatigue resistance should be assessed. This means, the reliability margin in the technical probability method is:

$$Z_{(R,S)} = RF = R - S \ , \tag{14}$$

where R is the random resistance of the element and S represents random variable effects of the extreme load.

When using (13), the condition for the acceptable crack length a_{ac} is:

$$N = \frac{1}{C} \int_{a_0}^{a_{ac}} \frac{\mathrm{d}a}{\triangle K^m} > N_{tot} \ , \tag{15}$$

where N is the number of cycles needed to increase the crack from the initiation size a_0 to the acceptable crack size a_{ac}, and N_{tot} is the number of cycles throughout the service life.

The equation for the propagation of the crack size (13) needs to be modified for this purpose. The state of stress near the crack face is described using ΔK (the stress intensity coefficient) which depends on the loading (bending, tension), size and shape of the fatigue crack, and geometry of the load-bearing component. If the $\Delta\sigma$ stress range and axial stress-load of the flange are constant, the following relation applies:

$$\Delta K = \Delta\sigma \cdot \sqrt{\pi a} \cdot F_{(a)} \ , \tag{16}$$

where $F_{(a)}$ is the calibration function which represents the course of propagation of the crack. After the change of the number of cycles from N_1 to N_2, the crack will propagate from the length a_1 to a_2. Having modified (13) and using (16), the following formula will be achieved:

$$\int_{a_1}^{a_2} \frac{\mathrm{d}a}{(\sqrt{\pi a} \cdot F_{(a)})^m} = \int_{N_1}^{N_2} C \cdot (\triangle\sigma)^m \, \mathrm{d}N \ . \tag{17}$$

If the length of the crack a_1 equals to the initial length a_0 (this is the assumed size of the initiation crack in the probabilistic approach) and if a_2 equals to the final acceptable crack length a_{ac} (this is the acceptable crack size which replaces the critical crack size a_{cr} if the crack results in a brittle fracture), the left-hand side of the equation (17) can be regarded as the resistance of the structure $R_{(a_{ac})}$:

$$R_{(a_{ac})} = \int_{a_0}^{a_{ac}} \frac{\mathrm{d}a}{(\sqrt{\pi a} \cdot F_{(a)})^m} \ . \tag{18}$$

If the upper integration limit a_d is used, the resistance of the structure $R_{(a_{ad})}$ can be specified similarly. Similarly, it is possible to define the cumulated effect of loads that is equal to the right side (randomly variable effects of the extreme load) (17):

$$S = \int_{N_0}^{N} C \cdot (\triangle\sigma)^m \, \mathrm{d}N \ = C \cdot (\triangle\sigma)^m \cdot (N - N_0) \ , \tag{19}$$

where N is the total number of oscillations of stress peaks ($\Delta\sigma$) for the change of the length from a_0 to $a_{a_{ac}}$, and N_0 is the number of oscillations in the time of initialization of the fatigue crack (typically, the number of oscillations is zero).

It is possible to define a reliability function RF:

$$RF_{(\mathbf{X})} = R_{(a_{ac})} - S_{(N)} \ . \tag{20}$$

where $\mathbf{X}$ is a vector of random physical properties such as mechanical properties, geometry of the structure, load effects and dimensions of the fatigue crack.

The analysis of the reliability function (20) gives a failure probability p_f:

$$p_f = P(RF_{(\mathbf{X})} < 0) = P(R_{(a_{ac})} < S_{(N)}) \ . \tag{21}$$

3.3. Applying theoretical approach to propagation of fatigue cracks in FCProbCalc

A tension flange has been chosen for applications of the theoretical solution suggested in the studies (Tomica et al., 2007). Depending on location of an initial crack, the crack may propagate from the edge or from the surface (see Fig. 5). Regarding the frequency, weight and stress concentration, those locations rank among those with the major hazard of fatigue cracks appearing in the steel structures and bridges.

A flange without stress concentration is used for confronting the both cases depending on the location of the crack initiation. The cases are different in calibration functions $F_{(a)}$ - and in weakened surfaces which are appearing during the crack propagation.

3.3.1. *Probabilistic calculation of fatigue cracks propagating from the edge*

For the crack propagating from the edge, the calibration function is:

$$F_{(a)} = 1.12 - 1.39 \cdot \frac{a}{b} + 7.32 \cdot \left(\frac{a}{b}\right)^2 - 13.8 \cdot \left(\frac{a}{b}\right)^3 + 14.0 \cdot \left(\frac{a}{b}\right)^4 \ , \tag{22}$$

where a is the length of the crack and b is the width of the flange (Janssen et al., 2002); (see Fig.5).

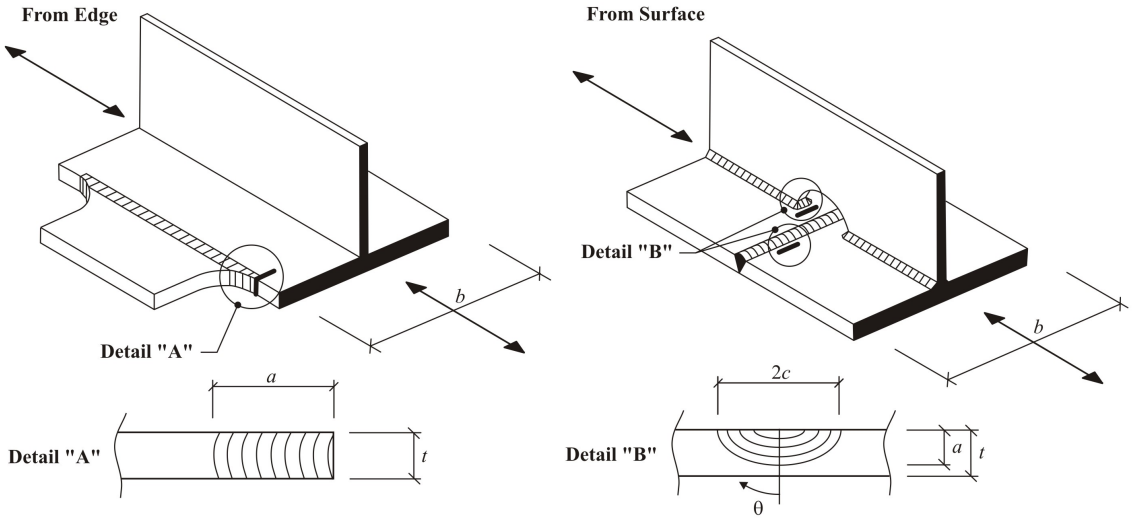


Figure 5. Characteristic propagation of cracks from the outer edge (left) and from the surface (right).

The acceptable crack size a_{ac} can be described then by a formula resulting from the deduced weakening of the cross-section area of the flange:

$$a_{ac} = b \cdot \left(1 - \frac{\sigma_{\max}}{f_y}\right) \ . \tag{23}$$

3.3.2. *Probabilistic calculation of fatigue cracks propagating from the surface*

A similar approach can be used to determine the acceptable size of a crack propagating from the surface. The bending component can be neglected for welded steel two-axis symmetric I-profiles where the fatigue crack appears in the lower tension flange. The flange is loaded only by the normal stress resulting from the axial load - tension: $\sigma_m = \sigma$.

It is rather difficult to deduce analytically the acceptable size of the crack propagating from the surface. In accordance with (Krejsa et al., 2010), the shape is replaced with a semi-elliptic curve where the ellipsis axes are a (the crack depth) and c (a half of the crack width) - see Fig. 5. The area of the surface crack depends on the number of N loading cycles and is described by the following formula:

$$A_{cr(N)} = \frac{1}{2} \cdot \pi \cdot a_N \cdot c_N \ . \tag{24}$$

During propagation of the fatigue crack from the surface, it is not enough to monitor only one crack size (which would be sufficient, for instance, for a crack propagating from the edge). In that case, the crack size needs to be analyzed for directions of the both semi-axes: a and c. The propagation of the fatigue crack from the surface in the a direction depends on the propagation in the c direction. Crack velocity propagation is

described by (13). In (Krejsa et al., 2010) there is a formula for calculation of the crack depth Δa as a result of an increased width of the Δc crack:

$$\Delta a = \left\{ \frac{1}{\left[1.1 + 0.35 \cdot \left(\frac{a}{t}\right)^2 \cdot \sqrt{\frac{a}{c}}\right]} \right\}^m \cdot \Delta c \ . \tag{25}$$

The crack sizes for a and c are during the propagation limited by upper limit values:

$$2 \cdot c \leq 0,4 \cdot b_f \quad a \leq 0,8 \cdot t_f \ , \tag{26}$$

If these upper limit values are exceeded, the fatigue crack propagates differently. (Krejsa, 2011) gives also the formula for the mutual dependence of the sizes in a and c:

$$c = 0.3027 \cdot \frac{a^2}{t} + 1.0202 \cdot a + 0.00699 \cdot t \ . \tag{27}$$

When determining the acceptable crack size, a modified relation (24) using (25) and (27), should be taken as a basis. After modification:

$$\sigma_{\max} \cdot \frac{b_f t_f}{b_f t_f - \frac{1}{2} \cdot \pi a \cdot \left(0.3027 \cdot \frac{a^2}{t_f} + 1.0202 \cdot a + 0.00699 \cdot t_f\right)} \leq f_y \ , \tag{28}$$

It is difficult to describe the a crack size directly explicitly. In order to calculate the acceptable crack size a_{ac}, it is necessary to use a numerical iteration approach where restrictions resulting from (28) should be taken as a basis.

3.3.3. *Determination of inspections of structures subject to fatigue*

Because it is not certain in the probabilistic calculation whether the initiation crack exists and what the initiation crack size is and because other inaccuracies influence the calculation of the crack propagation, a specialized inspection is necessary to check the size of the measureable crack in a specific period of time. The acceptable crack size influences the time of the inspection. If no fatigue cracks are found, the analysis of inspection results gives conditional probability during occurrence.

While the fatigue crack is propagating, it is possible to define following random phenomena that are related to the growth of the fatigue crack and may occur in any time t during the service life of the structure. Then:

- $U_{(t)}$ **phenomenon:** No fatigue crack failure has not been revealed within the t-time and the fatigue crack size $a_{(t)}$ has not reached the detectable crack size a_d . This means:

$$a_{(t)} < a_d \ , \tag{29}$$

- $D_{(t)}$ **phenomenon:** A fatigue crack failure has been revealed within the t-time and the fatigue crack size $a_{(t)}$ is still below the acceptable crack size a_{ac}. This means:

$$a_d \leq a_{(t)} < a_{ac} \ , \tag{30}$$

– $F_{(t)}$ **phenomenon:** A failure has been revealed within the t-time and the fatigue crack size $a_{(t)}$ has reached the acceptable crack size a_{ac}. This means:

$$a_{ac} < a_{(t)} \ . \tag{31}$$

If the crack is not revealed within the t-time, this may mean that there is not any fatigue crack in the construction element. This might be an initiative phase of nucleation of the fatigue crack (when a crack appears in the material) and this phenomenon is not taken into account in the fracture mechanics. Even if the fatigue crack is not revealed it is likely that it exists but the fatigue crack size is so small that it cannot be detected under existing conditions.

Using the phenomena above, it is possible to define probability for their occurrence in any t-time. Those three phenomena cover the complete spectrum of phenomena that might occur in the t-time. This means:

$$P(U_{(t)}) + P(D_{(t)}) + P(F_{(t)}) = 1 \ . \tag{32}$$

The probabilistic calculation is carried out in time steps where one step typically equals to one year of the service life of the construction. When the failure probability $P(F_{(t)})$ reaches the nominal failure probability p_d, an inspection should be carried out in order to find out fatigue cracks, if any, in the construction element. The inspection provides information about real conditions of the construction. Such conditions can be taken into account when carrying out further probabilistic calculations. The inspection in the t time may result in any of the three mentioned phenomena. Using the inspection results for the t time, it is possible to define the probability of the mentioned phenomena in another times: $T > t_I$. For that purpose, the conditional probability should be taken into consideration.

3.3.4. *Using FCProbCalc for the probabilistic calculation of fatigue cracks propagating*

FCProbCalc (Fig. 6) was developed using the aforementioned techniques. By means of FCProbCalc, it is possible to carry out the probabilistic calculation of propagation of fatigue cracks in a user friendly environment. The cracks propagate from edges or surface and the goal of the probabilistic calculation is to determine the time for the first inspection which focuses on damage to the structure.

Both deterministic and stochastic approaches are used for input values in the probabilistic calculation. In (Krejsa et al., 2010), the probabilistic assessment was carried out for a detail of a highway bridge made from steel/concrete which tends to suffer from fatigue damage. Real input values were used there: the geometric shape in the specified place, the yield stress f_y, the nominal designed stress of extreme impacts σ, material constants m and C, as well as constant stress oscillation $\Delta\sigma$. The source of the oscillation value was measurements of the response in regular operation. Other input data include the random quantities - they are expressed by means of the parametric distribution and were rather inaccurate if used as the input values. These values include the expected length of the detectable crack $a_d = 10$ mm, the number of load cycles per year $N = 1.10^6$ and, in particular, the size and exact location of the initiation crack a_0. Considering the detail of connection of the flange plate, it was decided to choose the mean value of $a_0 = 0.2$ mm with lognormal distribution. For all input data see Table I (the random quantities with variable values) and Table II (the deterministic quantities). The required reliability is expressed in the technical practice as a reliability index $\beta = 2$, that corresponds to the failure rate of $p_d = 0.02277$.

Using FCProbCalc it is possible to specify for a certain time interval the load effect S (Fig. 7), resistance of the structure $R(a_d)$ (Fig. 8) and $R(a_{ac})$ (Fig. 9), as well as probability of elementary phenomena U, D

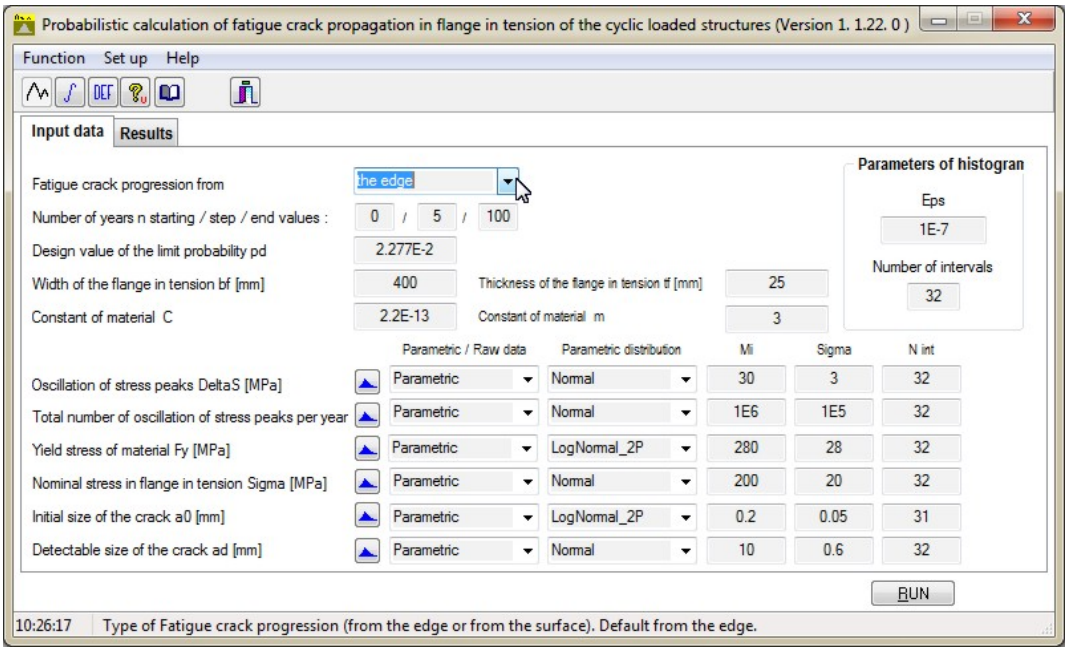


Figure 6. FCProbCalc desktop - entry of input quantities.

Table I. Overview of variable input quantities expressed in a histogram with parametric distribution of probabilities.

Quantity	Type	Mean value	Standard deviation
Oscillation of stress peaks $\Delta\sigma$ [MPa]	Normal	30	3
Total number of oscillation of stress peaks per year N [-]	Normal	10^6	10^5
Yield stress f_y [MPa]	Lognormal	280	28
Nominal stress in the flange plate σ [MPa]	Normal	200	20
Initial size of the crack a_0 [mm]	Lognormal	0.2	0.05
Smallest detectable size of the crack a_d [mm]	Normal	10	0.6

and F (Fig. 10) which are the source information for determination of the time of inspection which focuses on fatigue damage to the construction (Fig. 11).

The probabilistic calculation in FCProbCalc has proved, among others, that the propagation of the fatigue crack from the surface is considerably slower than that from the edge. The calculated time for the first inspection of the bridge is the 55^{th} year of operation for the fatigue crack propagating from the edge and 113^{th} year of operation for the fatigue crack propagating from the surface. This means the former propagation rate is approximately twice slower than the latter one.

Table II. Overview of input quantities expressed in a deterministic way.

Quantity	Value
Material constant m	3
Material constant C	$2.2 \cdot 10^{13}$
Width of the flange plate b_f [mm]	400
Thickness of the flange plate t_f [mm]	25
Nominal probability of failure p_d	0.02277

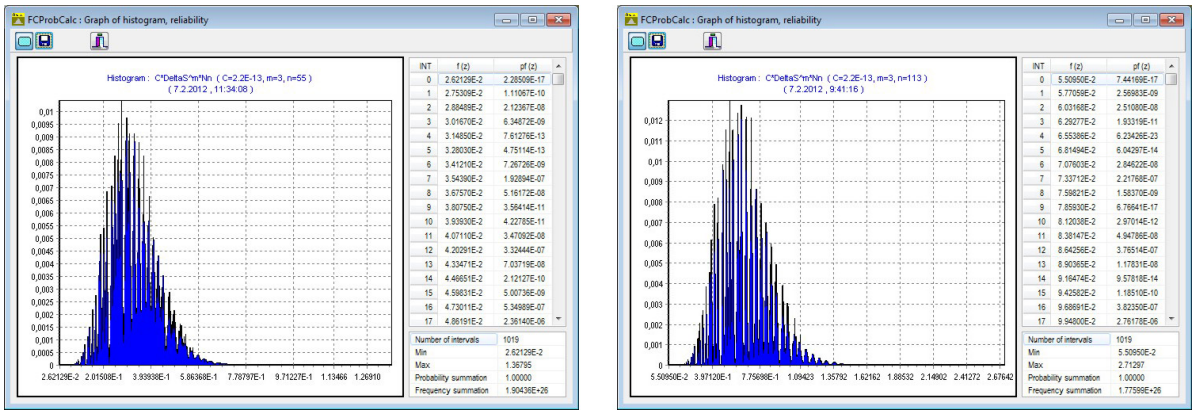


Figure 7. FCProbCalc program output: histograms for the load effects S after 55 years (left) and 113 years (right) of operation.

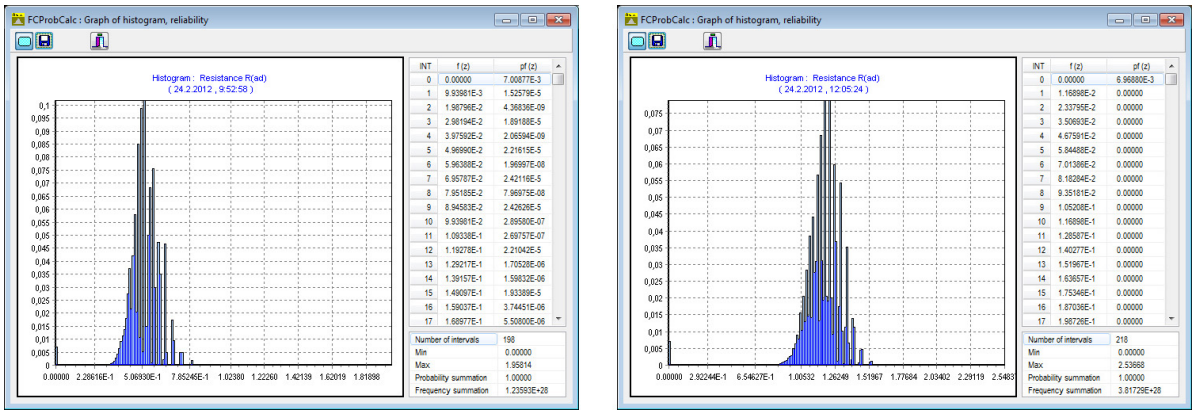


Figure 8. FCProbCalc program output: Resulting histogram of the structural resistance $R(a_d)$ for propagation of fatigue crack from the edge (left) and from the surface (right).

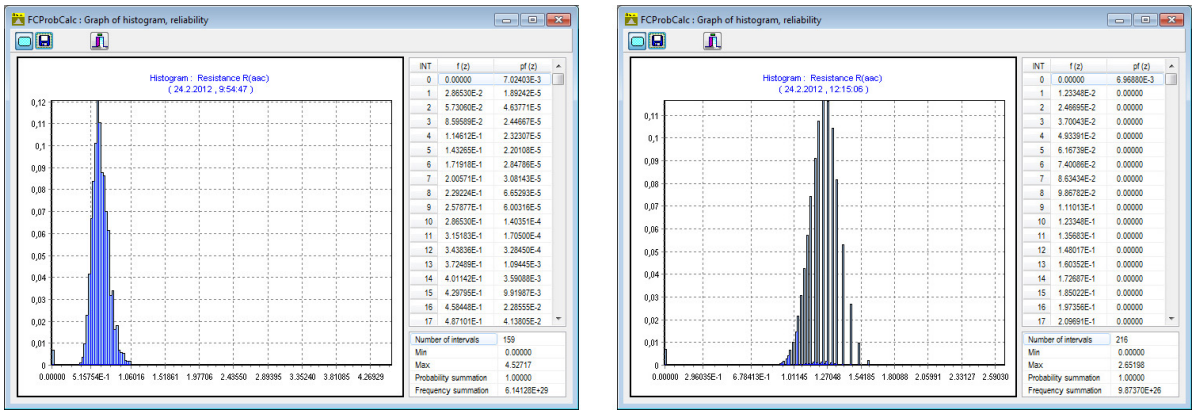


Figure 9. FCProbCalc program output: Resulting histogram of the structural resistance $R(a_{ac})$ for propagation of fatigue crack from the edge (left) and from the surface (right).

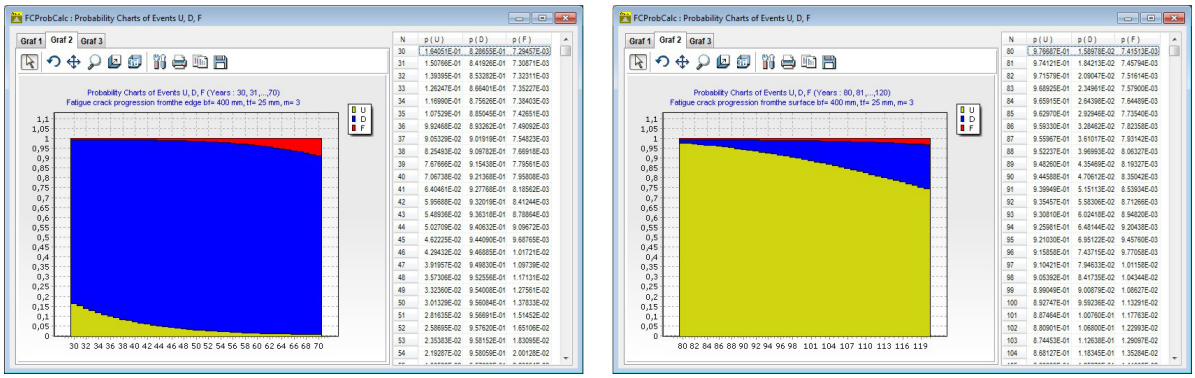


Figure 10. FCProbCalc program output: Probabilities of the phenomena U, D and F for the propagation of fatigue crack from the edge (30 to 70 years of operation, left) and for the propagation of fatigue crack from the surface (80 to 120 years of operation, right).

4. Conclusions

This paper discusses development of probabilistic methods and application of the probabilistic methods in assessment of reliabilities of structures. The basics of this work are a detailed overview of the Direct Optimized Probabilistic Calculation (DOProC) which can be used now in many probabilistic calculations. DOProC appears to be a very efficient tool that results in the solution affected by a numerical error and by an error resulting from the discretizing of the input and output quantities only. The biggest weakness of DOProC is a considerable increase in the machine time for probabilistic operations and rather many random variables in the computational model. The maximum number of the random variables depends on complexity of the computational model. What is also important is whether it is possible to use any of the described optimized steps.

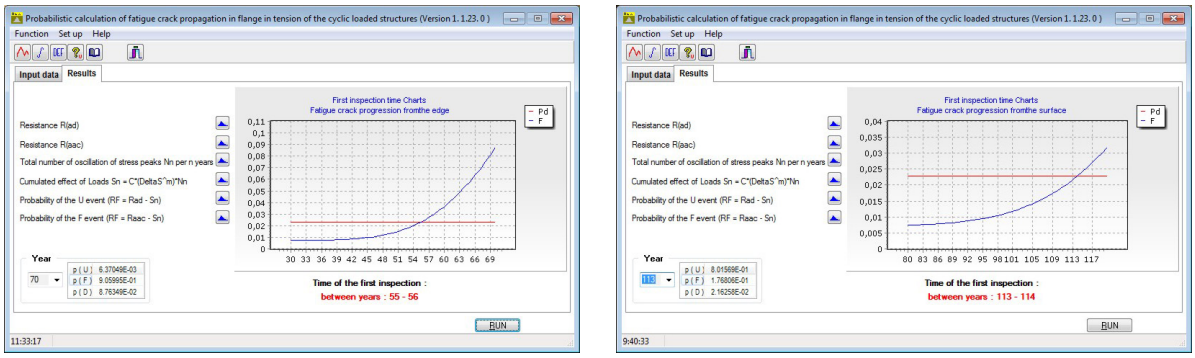


Figure 11. FCProbCalc program output: failure probability p_f , depending on the years of operation for the propagation of fatigue crack from the edge (30 to 70 years of operation, left) and for the propagation of fatigue crack from the surface (80 to 120 years of operation, right).

Examples of applications of the probabilistic method DOProC described in specialized papers and mentioned in this work should provide general information about this probabilistic method. DOProC seems to be a good choice not only for reliability assessment tasks but also for other probabilistic calculations. For instance, theoretical information and practical guidelines are available to the probabilistic assessment of propagation of fatigue cracks from the surface and edge, a particular attention being paid to the maximum permissible dimension and proposed system of regular inspections of the structure.

FCProbCalc was used for the probabilistic assessment of fatigue damage to a bridge structure where cracks were propagating from both the surface and edge. Times were specified for inspections of the bridge structure, where the purpose was to monitor occurrence of certain fatigue cracks. The comparison proved that velocity of propagation of the fatigue crack from the surface is considerably slower than that from the edge.

A relatively complex algorithm in DOProC requires good theoretical knowledge and practical computing skills of the user. It is essential to know, at least, general basics of algorithms because this influence the way of defining the computational model and selection of the best optimizing procedure. This weakness is removed if the application software is customized for a specific probabilistic task, this being, for instance, the case of FCProbCalc.

It should be pointed out that DOProC still provides many other options to be used. What is worth being investigated further is the use of statistically dependent input quantities with direct entries in the computational algorithm, assessment of reliability of structural systems and development of numerical procedures which will make the application of DOProC in matrix calculations more efficient.

Appendix

For a lite version of FCProbCalc and for other software products based on DOProC method please visit web pages **http://www.fast.vsb.cz/popv** (Janas et al., 2012).

Acknowledgements

The paper was published thanks to the financial support granted to the project "Mathematics for the 21^{st} century engineers – innovating the teaching of mathematics in technical schools in the context of a fast developing information and technical society". The project registration number is CZ.1.07/2.2.00/07.0332. Universities responsible for the project are the Technical University of Ostrava and West Bohemia University in Pilsen.

References

Anderson, T. L. *Fracture mechanics: fundamentals and applications*. Third edition, CRC Press, Taylor & Francis Group, Boca Raton, Florida, 2005. ISBN 0-8493-1656-1.

Bergmeister, K., Novak, D., Pukl, R. and V. Cervenka V. Structural and Reliability Analysis for Existing Engineering Structures, Theoretical Background. *Structure and Infrastructure Engineering*, vol. 5, issue 4, pp 267–275 (9 p), DOI 10.1080/15732470601185612, 2009.

Bjerager, P. Probability Integration by Directional Simulation. *Journal of Engineering Mechanics – ASCE*, vol. 114, issue 8, pp 1285–1302 (18 p), 1988.

Bucher, C. G. Adaptive Sampling – an Iterative Fast Monte-Carlo Procedure. *Structural Safety*, vol. 5, issue 2, pp 119–126 (8 p), DOI 10.1016/0167-4730(88)90020-3, 1988.

Carpinteri, A. and M. Paggi. Self-similarity and crack growth instability in the correlation between the Paris' constants. *Engineering Fracture Mechanics*, vol. 74, issue 7, pp 1041–1053 (13 p), DOI 10.1016/j.engfracmech.2006.12.007, May 2007.

Chen, N. Z., Wang, G. and C. G. Soares. Palmgren-Miner's rule and fracture mechanics-based inspection planning. *Engineering Fracture Mechanics*, vol. 78, issue 18, pp 3166–3182 (17 p), DOI 10.1016/j.engfracmech.2011.08.002, December 2011.

Der Kiureghian, A. and T. Dakessian. Multiple Design Points in First and Second-Order Reliability. *Structural Safety*, vol. 20, issue 1, pp 37–49 (13 p), DOI 10.1016/S0167-4730(97)00026-X, 1998.

Fisher, J. W., Kulak, G. L. and I. F. C. Smith. *A Fatigue Primer for Structural Engineers*. National Steel Bridge Allience, U.S.A., May 1998.

Giner, E., Sukumar, N., Denia, F. D. and F. J. Fuenmayor. Extended finite element method for fretting fatigue crack propagation. *International Journal of Solids and Structures*, vol. 45, issue 22–23, pp 5675–5687 (13 p), DOI 10.1016/j.ijsolstr.2008.06.009, November 2008.

Helton, J. C. and F. J. Davis. Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems. *Reliability Engineering & System Safety*, vol. 81, issue 1, pp 23–69 (47 p), DOI 10.1016/S0951-8320(03)00058-9, 2003.

Hurtado, J. E. and A. H. Barbat. Monte Carlo Techniques in Computational Stochastic Mechanics. *Archives of Computational Methods in Engineering*, vol. 5, issue 1, pp 3–29 (27 p), DOI 10.1007/BF02736747, 1998.

Janas, P., Krejsa, M. and V. Krejsa. *ProbCalc Software and DOProC Method*. Web pages. [on-line]. ¡http://www.fast.vsb.cz/popv¿. VSB – Technical University Ostrava, 2012.

Janas, P., Krejsa, M. and V. Krejsa. Structural Reliability Assessment using a Direct Determined Probabilistic Calculation. *Proceedings of the* 12^{th} *International Conference on Civil, Structural and Environmental Engineering Computing*, Civil–Comp Press, Stirlingshire, Scotland. Paper 79, 2009, ISBN 978-1-905088-31-7.

Janas, P., Krejsa, M. and V. Krejsa. Using the Direct Determined Fully Probabilistic Method for determination of failure. *Proceedings of the European Safety and Reliability Conference, Esrel 2009*. Taylor & Francis Group, London, 2010, ISBN 978-0-415-55509-8 (set of 3 volumes + CD-ROM), ISBN 978-0-203-85975-9 (e-book).

Janssen, M., Zuidema, J. and R. J. H. Wanhill. *Fracture Mechanics*. Second edition, Delft University Press, 2002. ISBN 90-407-2221-8.

Kala, Z. Sensitivity Analysis of the Stability Problems of Thin–Walled Structures. *Journal of Constructional Steel Research*, vol. 61, issue 3, pp 415–422 (8 p), DOI 10.1016/j.jcsr.2004.08.005, 2005.

Konecny, P., Tikalsky, P. J. and D. G. Tepke. Performance Evaluation of Concrete Bridge Deck Affected by Chloride Ingress. *Transportation Research Record*, issue 2028, pp 3–8 (6 p), DOI 10.3141/2028-01, 2007.

Kralik, J. and J. Kralik. Seismic Analysis of Reinforced Concrete Frame–Wall Systems Considering Ductility Effects in Accordance to Eurocode. *Engineering Structures*, vol. 31, issue 12, pp 2865–2872 (8 p), DOI 10.1016/j.engstruct.2009.07.029, 2009.

Krejsa, M. and V. Tomica. Calculation of Fatigue Crack Propagation Using DOProC Method. *Transactions of the VSB – Technical University of Ostrava*, no.1, 2010, vol.X, Civil Engineering Series, paper #11 (9 p), DOI 10.2478/v10160-010-0011-6. Publisher Versita, Warsaw, ISSN 1213-1962 (Print) ISSN 1804-4824 (Online).

Krejsa, M. and V. Tomica. Determination of Inspections of Structures Subject to Fatigue. *Transactions of the VSB – Technical University of Ostrava*, no.1, 2011, vol.XI, Civil Engineering Series, paper #7 (11 p), DOI 10.2478/v10160-011-0007-x. Publisher Versita, Warsaw, ISSN 1213-1962 (Print) ISSN 1804-4824 (Online).

Krejsa, M. *Using DOProC Method for Structural Reliability Assessment*. Habilitation thesis, VSB – TU Ostrava, 2011. ISBN 978-80-248-2385-0.

Krivy, V. and P. Marek. Probabilistic Design of Steel Frame Structures. *Stahlbau*, vol. 76, issue 1, pp 12–20 (13 p), DOI 10.1002/stab.200710003, 2007.

Kubecka, K. Utilisation of Risk Analysis Methods in Decision – Making Process on Fitness of Rehabilitation. *Stavebnictv*, pp 26–31 (6 p), 2010. ISSN 1802-2030, EAN 977180220300501.

Li, J., Peng, Y. B. and Chen J. B. Probabilistic Criteria of Structural Stochastic Optimal Controls. *Probabilistic Engineering Mechanics*, vol. 26, issue 2, pp 240–253 (14 p), DOI 10.1016/j.probengmech.2010.07.011, 2011.

Marschalko, M., Yilmaz, I., Bednarik, M. and K. Kubecka. Variations in the building site categories in the underground mining region of Doubrava (Czech Republic) for land use planning. *Engineering Geology*, vol. 122, issue 3–4, pp 169–178 (10 p), DOI 10.1016/j.enggeo.2011.05.008, October 2011.

Melchers, R. E. Structural System Reliability Assessment Using Directional Simulation. *Structural Safety*. vol. 16, issue 1–2, pp 23–37 (15 p), DOI 10.1016/0167-4730(94)00026-M, 1994.

Moan, T. Reliability–based management of inspection, maintenance and repair of offshore structures. *Structure and Infrastructure Engineering*, vol. 1, issue 1, pp 33–62 (30 p), DOI 10.1080/15732470412331289314, Taylor & Francis Ltd., March 2005.

Olsson, A., Sandberg, G. and O. Dahlblom. On Latin Hypercube Sampling for Structural Reliability Analysis. *Structural Safety*, vol. 25, issue 1, pp 47–68 (22 p), Article Number: PII S0167-4730(02)00039-5, DOI 10.1016/S0167-4730(02)00039-5, 2003.

Rackwitz, R. and B. Fiessler. Structural Reliability Under Combined Random Load Sequences. *Computers & Structures*, vol. 9, issue 5, pp 489–494 (6 p), DOI 10.1016/0045-7949(78)90046-9, 1978.

Reh, S., Beley, J. D., Mukherjee, S. and E. H. Khor. Probabilistic Finite Element Analysis Using ANSYS. *Structural Safety*, vol. 28, issue 1–2, pp 17–43 (27 p), DOI 10.1016/j.strusafe.2005.03.010, 2006.

Sanford, R. J. *Principles of Fracture Mechanics*. Pearson Education, Inc., U.S.A., 2003. ISBN 0-13-092992-1.

Stefanou, G. The Stochastic Finite Element Method: Past, Present and Future. *Computer Methods in Applied Mechanics and Engineering*, vol. 198, issue 9–12, pp 1031–1051 (21 p). DOI 10.1016/j.cma.2008.11.007, 2009.

Straub, D. Stochastic Modeling of Deterioration Processes through Dynamic Bayesian Networks. *Journal of Engineering Mechanics*. 2009. DOI 10.1061/(ASCE)EM.1943-7889.0000024.

Teply, B., Vorechovska, D. and Z. Kersner. Performance-Based Design of Concrete Structures: Durability Aspects. *Structural Engineering and Mechanics*, vol. 35, issue 4, pp 535–538 (4 p), 2010.

Thacker, B. H., Riha, D. S., Fitch, S. H. K., Huyse, L. J. and J. B. Pleming. Probabilistic Engineering Analysis Using the NESSUS software. *Structural Safety*, vol. 28, issue 1–2, pp 83–107 (25 p), 2006, DOI 10.1016/j.strusafe.2004.11.003.

Tomica, V. and M. Krejsa. Optimal Safety Level of Acceptable Fatigue Crack. *Proceedings of* 5^{th} *International Probabilistic Workshop*, Ghent, Belgium, 2007. Edited by L.Taerwe and D.Proske. (12 p). ISBN 978-3-00-022030-2.

Tvedt, L. Proban - Probabilistic Analysis. *Structural Safety*, vol. 28, issue 1-2, pp 150-163 (14 p), 2006, DOI 10.1016/j.strusafe.2005.03.003.

Vanmarcke, E., Shinoyuka, M., Nakagiri, S., Schueller, G. I. and M. Gtigoriu. Random – Fields and Stochastic Finite – Elements. *Structural Safety*, vol. 3, issue 3-4, pp 143–166 (24 p), DOI 10.1016/0167-4730(86)90002-0, 1986.

Vorechovsky, M. and D. Novak. Correlation Control in Small-Sample Monte Carlo type simulations I: A Simulated Annealing Approach. *Probabilistic Engineering Mechanics*. vol. 24, issue 3, pp 452–462 (11 p), DOI 10.1016/j.probengmech.2009.01.004, 2009.

Zhao, Y. G. and T. Ono. Moment Methods for Structural Reliability. *Structural Safety*, vol. 23, issue 1, pp 47–75 (29 p), DOI 10.1016/S0167-4730(00)00027-8, 2001.

Determination of statistical material parameters of concrete using fracture test and inverse analysis based on FraMePID-3PB tool

David Lehký, Zbyněk Keršner, Drahomír Novák
Brno University of Technology, Faculty of Civil Engineering
Institute of Structural Mechanics,Brno, Czech Republic,
lehky.d@fce.vutbr.cz, kersner.z@fce.vutbr.cz, novak.d@fce.vutbr.cz

Abstract: The knowledge of fracture mechanical parameters values is fundamental for virtual failure modeling of elements and structures made of concrete. A key parameter of nonlinear fracture mechanics modeling is certainly specific fracture energy of concrete and its variability. Within this paper experimental results from three-point bending tests on notched-beam specimens are analyzed. Two basic approaches are applied to determine fracture mechanical parameters from these tests: (i) effective crack model / work-of-fracture method, (ii) inverse analysis using artificial neural networks and virtual stochastic simulations. In order to automate the whole time consuming process of inverse analysis a FraMePID-3PB software tool has been developed. The paper is focused on the determination of statistical fracture-mechanical parameters values of four different concrete types. It is a part of complex methodology for statistical and reliability analyses of concrete structures.

Keywords: Fracture parameters, concrete, inverse analysis, artificial neural networks, nonlinear analysis, fracture mechanics, FraMePID-3PB software

1. Introduction

The stochastic nonlinear computational mechanics faces in real-world application problems a fundamental obstacle – the lack of the knowledge of basic random variables involved in the problem. The direct experimental testing, often performed as compression test on cubic specimens, provides incomplete information about mechanical and fracture parameters and the lack of information is often substituted by an engineering judgment or by the information from literature. One possibility is to get parameters of computational model indirectly – based on combination of fracture test with inverse analysis. The paper describes a methodology to get such parameters using experimental data of three-point bending tests used in inverse analysis based on combination of artificial neural networks and stochastic analysis (Novák and Lehký, 2006). Since the whole procedure of inverse analysis is time consuming and complicated from data handling and artificial neural network training point of view a software tool FraMePID-3PB has been developed to automate fully the whole task.

A key parameter of nonlinear fracture mechanics modeling is certainly specific fracture energy of concrete and its variability, which is a subject of research of many authors, e.g. Bažant and Planas (1998). Other important parameters of concrete are modulus of elasticity, tensile and compressive strength. Crack propagation resistivity is described by e.g. effective crack elongation, effective fracture toughness etc. (Karihaloo, 1995). Determination of parameters values was done using two techniques – (i) direct

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

evaluation from experimental load–deflection diagram by effective crack model and work-of-fracture method; (ii) inverse analysis using artificial neural network based method.

Depending on sample size of statistical set, statistical characteristics of material parameters being identified can be determined using two approaches: (i) "One by one" approach – parameters of each specimen are identified separately and final statistics are calculated from the set of all values for each parameter. (ii) "Direct approach" – in case of larger statistical set it is more efficient not to identify each specimen one by one but to identify the whole statistical set for all specimens together based on random response of fracture tests (Lehký and Novák, 2011). The first approach was used in this paper.

2. Laboratory tests

Laboratory experiments and evaluation of fracture-mechanical parameters were performed using four sets of specimens of different concrete types: I (C30/37 H), II (C25/30 B3), III (C25/30 XC1 GK16), and IV (C20/25 XC1 GK16) prepared and casted in co-operation with Bautechnische Prüf- und Versuchsanstalt GmbH and University of Natural Resources and Life Sciences in Vienna, Austria. Specimens were tested in laboratory at Faculty of Civil Engineering, Brno University of Technology in Brno, Czech Republic in following ages: 91 days (set I), 87 days (set II), 67 days (set III) and 66 days (set IV). Each set consists of 9 specimens except of set IV which consists of 8 specimens. Nominal sizes of specimens were 100×100×400 mm. In the center of the beam the edge notch of the depth about 1/3 of the depth of the specimen was cut using diamond blade saw. Specimens were tested in three-point bending (3PB) configuration. Loading span was equal to 300 mm. Example of the tested specimen is in Figure 1.

Testing was performed using mechanical press Heckert FPZ 100/1. Loading of specimen was applied continuously with constant increment of displacement 0.1 mm/min in the middle of the span (300 mm). Midspan deflections were recorded using inductive sensor with accuracy of 0.001 mm. Result of measurement is diagram load vs. midspan deflection (*l*–*d* diagram).

For enlarging the set of material parameters with compressive strength which is not obtained from 3PB test the compression tests were performed too. It was carried out using two broken parts obtained after each three-point bending test. Broken parts were cut to nominal size 100×100×100 mm using diamond blade saw (Figure 2).

3. Evaluation of material parameters

3.1. Effective crack approach, work-of-fracture method

Recorded *l*–*d* diagram serves as a basis for evaluation of effective crack elongation and effective fracture toughness (or effective toughness) using models of equivalent elastic crack (Karihaloo, 1995). Then, using work-of-fracture method, a fracture work or specific fracture energy are assessed. As was already mentioned specific fracture energy is basic parameter of cohesive crack models which are used for prediction of fracture behavior of structures made of quasi-brittle materials (Stibor, 2004; Veselý, 2004).

a) b)

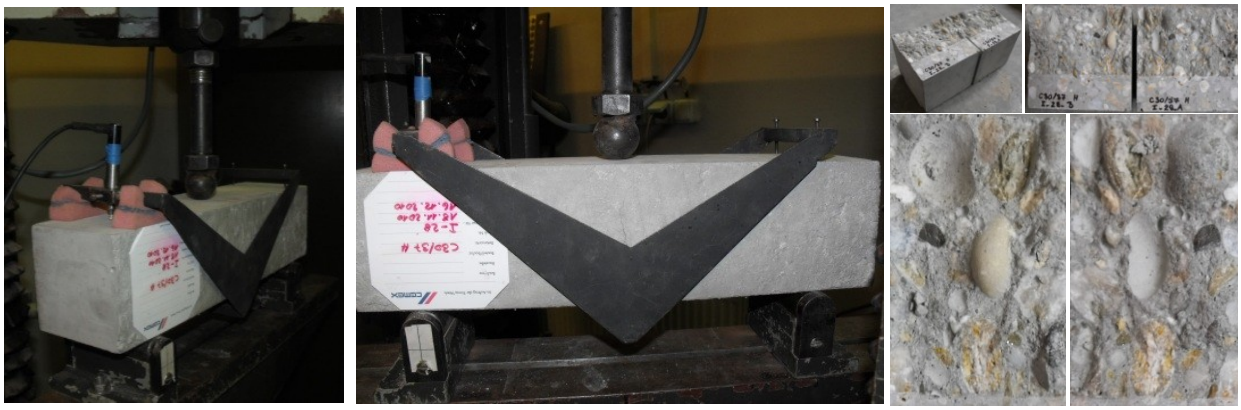


Figure 1. a) Selected specimen tested in three-point bending configuration and b) the fracture parts/surfaces after test [photo: B. Kucharczyková].

a) b)

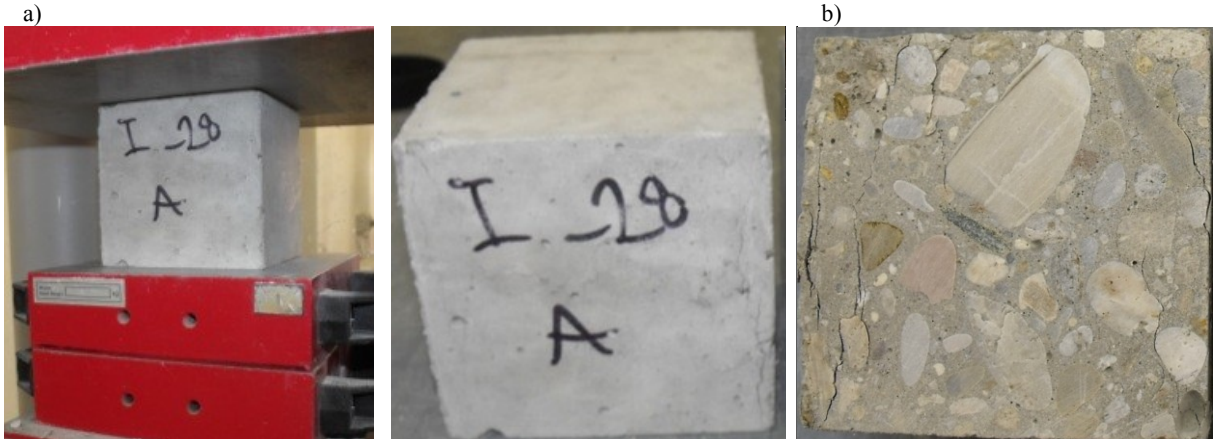


Figure 2. a) Compression test on broken parts of specimens and b) the shape after destruction [photo: B. Kucharczyková].

Important step before parameters evaluation from obtained *l–d* diagram is to recognize the origination of "catastrophe" in measured data (Frantík and Keršner, 2006). In time series of deflections in loading point an irregularity of loading speed and sudden increase of displacement can occur. Then, the time derivative of deflection is a useful criterion to detect the origin and range of the so-called fold catastrophe. The catastrophe is recognizable as extreme values of loading speed. The corrected *l–d* diagram and fold catastrophe have such properties which can help to discover the probable development of the diagram in the catastrophic part.

3.2. INVERSE ANALYSIS

Along with classical fracture mechanical parameters evaluation from fracture tests, parameters identification using artificial neural network based inverse method was carried out; see details in Novák and

Lehký (2006). The basis of inverse analysis is finite element method (FEM) model which is used for numerical simulation of three-point bending fracture test (Figure 3). FEM model was created in ATENA software (Červenka et al., 2007); material model 3D Nonlinear Cementitious 2 with rotational cracks was used. Subject of identification were following three parameters of concrete: modulus of elasticity, tensile strength and fracture energy. Other parameters of material model mentioned above, e.g. compressive strength, were omitted from identification based on sensitivity analysis. Here, Spearman's nonparametric rank-order correlation coefficient was used (Novák et al., 1993).

The material model parameters are considered as random variables described by a probability distribution, rectangular distribution is a "natural choice" as the lower and upper limits represent the bounded range of physical existence. The variables are then simulated randomly based on the Monte Carlo type simulation; the small-sample simulation Latin Hypercube Sampling (LHS) is recommended (McKay et al., 1979). A multiple calculation of deterministic computational model using random realizations of material model parameters is performed and a statistical set of the virtual structural response is obtained. Random realizations and the corresponding responses from the computational model serve as the basis for the training of an appropriate neural network (Cichocki and Unbehauen, 1993). After the training the neural network is ready to solve the main task: To provide the best material parameters in order the numerical simulation will result in the best agreement with experiment. This is performed by means of the simulation of network using measured response as an input. It results in a set of identified material parameters. The last step is results verification – calculation of computational model using identified parameters. A comparison with experiment will show to what extend the inverse analysis was successful. More details about structure of artificial neural network, training set, etc. are described in section 4.

To obtain statistical characteristics of material parameters inverse analysis is performed for each specimen (*l*–*d* diagram) individually. The set of identified values is obtained as the result of individual identification and can be assessed statistically as it is usually done for experiments.

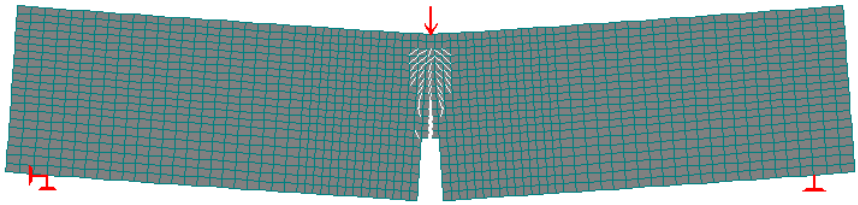


Figure 3. Scheme of nonlinear FEM computational model of three-point bending test.

4. FraMePID-3PB software tool

The methodology of artificial neural network based inverse analysis is general and can be used for any inverse task, which is its advantage. On the other hand it is very time consuming. In order to automate the whole difficult process of material parameters identification a FraMePID-3PB software tool has been developed. The whole system is based on standardized fracture test of beam with central edge notch in three-point bending configuration described in section 2. Finite element computational model implemented in FraMePID-3PB is created in ATENA software (Červenka et al., 2007). 3D Nonlinear Cementitious 2

material model for concrete is used. Softening of concrete is modeled using model according to Hordijk (1991).

Previous identifications using various types of concrete mixtures and ages showed that the structure of artificial neural network used for identification in this testing configuration is in all cases almost the same. Thank to that and using standardized test the time needed for inverse analysis can be significantly reduced because only one neural network is created, trained, tested and implemented within FraMePID-3PB system. Therefore, time consuming training set preparation using stochastic nonlinear analysis and training of the network using suitable optimization technique is performed only once. Structure of neural network implemented within FraMePID-3PB system is as follows (see Figure 5): 1 hidden layer with 5 nonlinear neurons (hyperbolic tangent transfer function), output layer with 3 linear neurons (linear transfer function) and 3 inputs of the network. Three output neurons correspond to three material parameters which are being identified (modulus of elasticity, tensile strength and specific fracture energy), three inputs correspond to three parameters extracted from *l*–*d* diagram.

During training set preparation for artificial neural network material parameters are randomized. Here, purposely large variability was used in order to create rather general network which will be able to identify parameters of concretes of various strengths and ages. Mean values were 40 GPa for modulus of elasticity, 4.5 MPa for tensile strength and 200 J/m^2 for fracture energy. Coefficients of variation were 0.2 for modulus of elasticity, 0.25 for tensile strength and 0.4 for fracture energy. Training set was generated using 100 simulations of Latin Hypercube Sampling method. Training of the network was carried out using Levenberg–Marquardt (Singh et al., 2007) and genetic algorithms (Schwefel, 1991) optimization methods.

Procedure of material parameters identification using FraMePID-3PB tool can be itemized as follows:

1. *L*–*d* diagram obtained from experiment is loaded into FraMePID-3PB. Curve is analyzed and inputs of inverse analysis are extracted and prepared for neural network (Figure 4).
2. Input signal is transmitted through the neural network and from the output layer of the network the best set of material parameters is obtained. This step is possible because neural network is trained in advance and remains the same for parameters identification of various concretes (Figure 5). Emphasize, that there is no new nonlinear fracture mechanics calculations to train network – the network is ready to use and implemented in FraMePID-3PB.
3. Verification of identification is performed. Obtained material parameters are used in the computational model and numerical analysis is carried out. Here, ATENA software is linked to FraMePID-3PB for data transfer. Resulting *l*–*d* diagram is compared with experimental one which will show to what extent the inverse analysis was successful (Figure 6).

At present, FraMePID-3PB software operates with "basic" configuration of experiment and model as was mentioned above. But, it was designed more generally with respect to next future extension for other configurations, e.g. specimens with various notch depths, other softening models of concrete (linear, multilinear, etc.), additional testing configurations (compressive test, wedge splitting test), etc. This will help with routine material parameters identification for various research and practical tasks.

5. Results

Values of selected parameters of all 35 specimens of four sets obtained using both above mentioned methods were statistically evaluated and their mean values and coefficients of variation (COV) can be

found in Table 1. There is a significant advantage in case of using inverse analysis – value of tensile strength of concrete can also be determined.

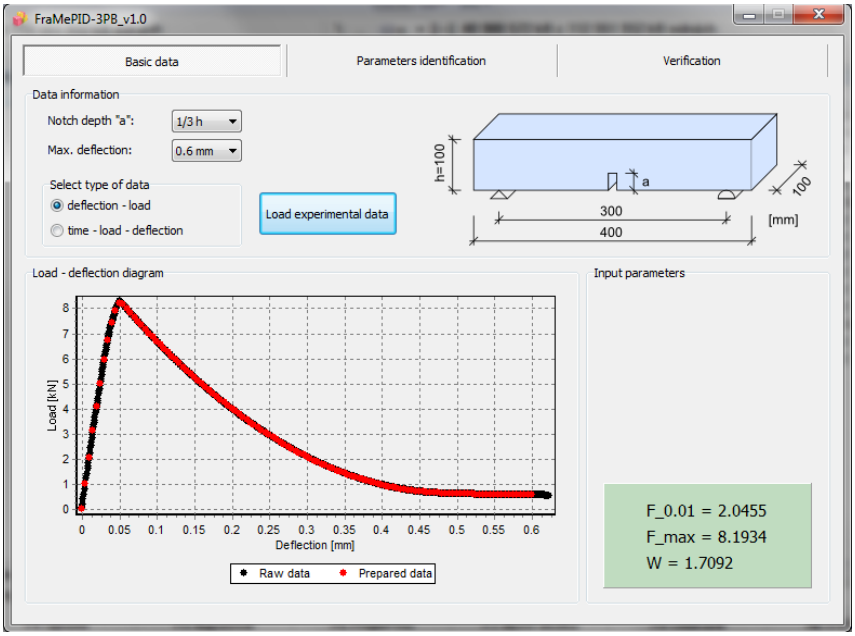


Figure 4. FraMePID-3PB tool panel – experimental data loading and preparation of input signal for neural network.

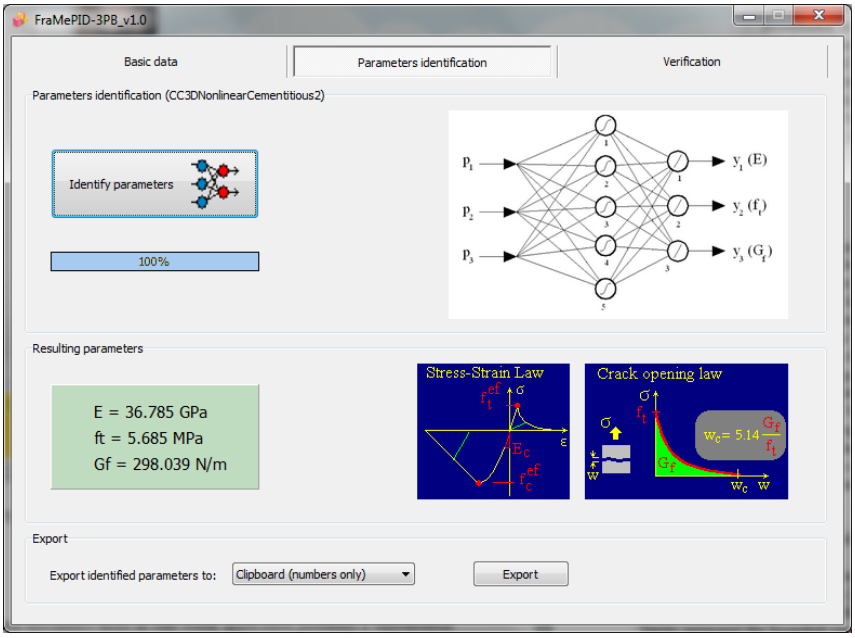


Figure 5. FraMePID-3PB tool panel – structure of neural network and material parameters identification.

Table I. Selected statistical parameters of set I (C30/37 H) obtained from experiment and inverse analysis.

Parameter of concrete / model	Experiment		Inverse analysis		Ratio (inv. analysis / experiment)
	Mean value	COV [%]	Mean value	COV [%]	
Modulus of elasticity [GPa]	35.5	7.4	40.3	16.6	1.14
Tensile strength [MPa]	–	–	5.0	14.3	–
Compressive strength [MPa]	58.5	8.0	–	–	–
Specific fracture energy [J/m^2]	235.9	18.6	281.5	19.5	1.19
Effective crack elongation [mm]	9.5	21.9	–	–	–
Effective fracture toughness [MPa.m$^{1/2}$]	1.489	9.9	–	–	–
Effective toughness [J/m2]	62.3	13.7	–	–	–
Volume density [kg/m^3]	2341.8	0.7	–	–	–

Table II. Selected statistical parameters of set II (C25/30 B3) obtained from experiment and inverse analysis.

Parameter of concrete / model	Experiment		Inverse analysis		Ratio (inv. analysis / experiment)
	Mean value	COV [%]	Mean value	COV [%]	
Modulus of elasticity [GPa]	30.8	8.6	35.0	8.2	1.14
Tensile strength [MPa]	–	–	4.1	17.2	–
Compressive strength [MPa]	47.3	5.4	–	–	–
Specific fracture energy [J/m^2]	188.9	11.5	211.8	18.1	1.12
Effective crack elongation [mm]	12.5	23.5	–	–	–
Effective fracture toughness [MPa.m$^{1/2}$]	1.406	8.0	–	–	–
Effective toughness [J/m2]	65.2	21.8	–	–	–
Volume density [kg/m^3]	2286.2	1.5	–	–	–

Table III. Selected statistical parameters of set III (C25/30 XC1 GK16) obtained from experiment and inverse analysis.

Parameter of concrete / model	Experiment		Inverse analysis		Ratio (inv. analysis / experiment)
	Mean value	COV [%]	Mean value	COV [%]	
Modulus of elasticity [GPa]	35.4	5.6	40.4	9.5	1.14
Tensile strength [MPa]	–	–	4.2	12.1	–
Compressive strength [MPa]	53.4	5.2	–	–	–
Specific fracture energy [J/m^2]	183.3	5.5	214.0	6.0	1.17
Effective crack elongation [mm]	12.4	22.7	–	–	–
Effective fracture toughness [MPa.m$^{1/2}$]	1.405	9.0	–	–	–
Effective toughness [J/m2]	56.3	19.9	–	–	–
Volume density [kg/m^3]	2326.9	0.9	–	–	–

Table IV. Selected statistical parameters of set IV (C20/25 XC1 GK16) obtained from experiment and inverse analysis.

Parameter of concrete / model	Experiment		Inverse analysis		Ratio (inv. analysis / experiment)
	Mean value	COV [%]	Mean value	COV [%]	
Modulus of elasticity [GPa]	31.2	4.3	34.8	5.3	1.12
Tensile strength [MPa]	–	–	3.1	15.6	–
Compressive strength [MPa]	39.8	5.4	–	–	–
Specific fracture energy [J/m^2]	146.2	13.3	166.8	15.4	1.14
Effective crack elongation [mm]	13.0	14.3	–	–	–
Effective fracture toughness [MPa.m$^{1/2}$]	1.131	10.6	–	–	–
Effective toughness [J/m2]	41.4	20.7	–	–	–
Volume density [kg/m^3]	2292.2	0.6	–	–	–

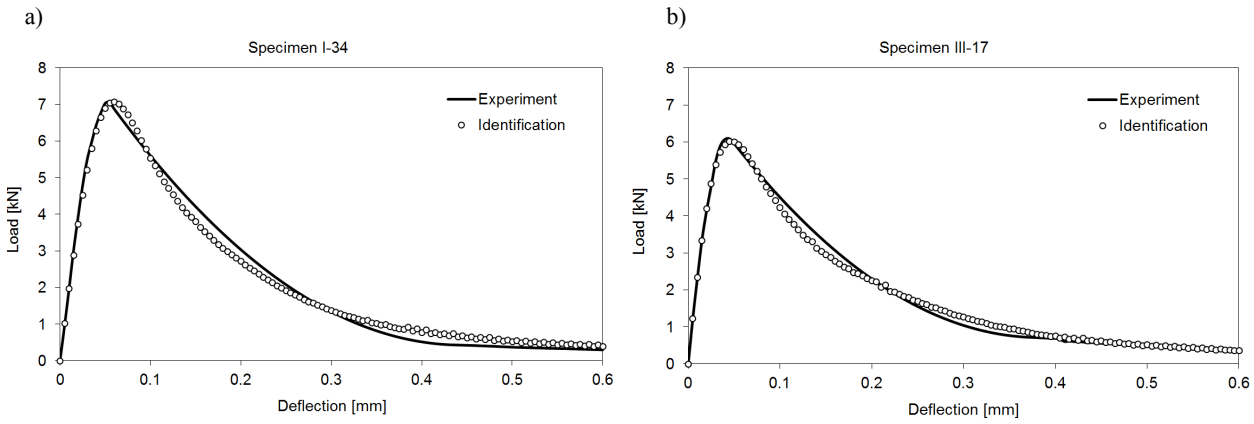


Figure 6. Comparison of selected experimental and numerically simulated load–deflection diagrams with material parameters obtained from identification: a) specimen I-34 (C30/37 H) and b) specimen III-17 (C25/30 XC1 GK16).

From the presented results it is possible to conclude recommended values of mechanical–fracture parameters for deterministic and stochastic nonlinear FEM analyses of beam/structures made of all four analyzed concretes, see Tables I–IV. Two-parametric lognormal probability distribution function is suggested for all three identified parameters (modulus of elasticity, tensile strength and fracture energy) and all four tested concrete types based on curve fitting tests carried out using FReET software (Novák et al., 2011) and JCSS Probabilistic Model Code recommendations (2001). Detailed results of all parameters for every single specimen and comparison of experimental and numerical *l–d* diagrams can be found in Keršner et al. (2011).

6. Conclusions

The proposed paper describes fracture tests and consequent evaluation of fracture mechanical parameters of specimens made of four different concrete types. Determination of values of these parameters was done using two techniques – (i) direct evaluation of parameters from experimental *l–d* diagram by effective crack model and work-of-fracture method; (ii) inverse analysis using artificial neural network based method. Results were compared; both techniques provided results which are close to each other including basic information on variability (COV). The inverse analysis technique provided additionally values of tensile strength of concretes. *L–d* diagrams from numerical simulations of all six specimens with identified parameters shows very good agreement with experimental ones. Results can serve efficiently as input data for stochastic nonlinear simulation of studied concretes.

Acknowledgements

The presented research was funded by research project No. TA01011019 of Technological Agency of Czech Republic (TACR) "SimSoft - Simulation software tools for virtual reliability testing and lifetime analysis of concrete structures" and partially by the project No. MSM0021630519 from the Ministry of Education of the Czech Republic. Authors gratefully acknowledge this support.

References

Bažant, Z.P. and Planas, J. *Fracture and Size Effect in Concrete and other Quasibrittle Materials*. CRC Press, Boca Raton, Florida, 1998.

Cichocki, A. and Unbehauen, R. *Neural networks for optimization and signal processing*. John Wiley & Sons Ltd. & B.G. Teubner, Stuttgart, Germany, 1993.

Červenka, V., Jendele, L., and Červenka, J. *ATENA Program Documentation – Part 1: Theory*. Cervenka Consulting, Prague, Czech Republic, 2007.

Frantík, P. and Keršner, Z. 2006. Vyhodnocení lomového experimentu s katastrofickou ztrátou stability (Evaluation of fracture test with catastrophic loss of stability). In: *Staticko-konštrukčné a stavebno-fyzikálne problémy stavebných konštrukcií*, Štrbské Pleso, Slovakia, 141–142, 2006, (in Czech).

Joint Committee on Structural Safety. *JCSS Probabilistic Model Code – Part II*. 2001, http://www.jcss.ethz.ch/publications/publications_pmc.html

Hordijk, D.A. *Local Approach to Fatigue of Concrete*. PhD thesis, Delft University of Technology, 1991.

Karihaloo, B.L. *Fracture mechanics of concrete*. Longman Scientific & Technical, New York, 1995.

Keršner, Z., Lehký, D., Novák, D., Kucharczyková, B., Frantík, P. and Řoutil, L. *Fracture energy tests: specimens I (C30/37 H), II (C25/30 B3), III (C25/30 XC1 GK16), IV (C20/25 XC1 GK16)*. Report No. 20110727, Institute of Structural Mechanics and Institute of Building Testing, Faculty of Civil Engineering, Brno University of Technology, Brno, Czech Republic, 2011.

Lehký, D. and Novák, D. 2011. Statistical material parameters identification of fibre-reinforced concrete. In: *The 9th International Symposium on High Performance Concrete – Design, Verification & Utilization (HPC 2011)*, Rotorua, New Zealand, p. 38 + CD (8 pages), 2011.

McKay, M.D., Conover, W.J. and Beckman, R.J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21: 239–245, 1979.

Novák, D., Teplý, B. and Shiraishi, N. Sensitivity analysis of structures: a review. *5th International Conference on Civil and Structural Engineering Computing*, Edinburgh, Scotland, 201–207, 1993.

Novák, D. and Lehký, D. ANN Inverse Analysis Based on Stochastic Small-Sample Training Set Simulation. *Engineering Application of Artificial Intelligence*, 19: 731-740, 2006.

Novák, D., Vořechovský, M. and Rusina, R. *FReET v.1.5 – program documentation. User´s and Theory Guides*. Brno/Červenka Consulting, Czech Republic, 2011, http://www.freet.cz.

Schwefel, H.P. *Numerical optimization for computer models*. Wiley, Chichester, 1991.

Singh, V., Gupta, I. and Gupta, H.O. ANN-based estimator for distillation using Levenberg-Marquardt approach. *Engineering Applications of Artificial Intelligence*, 20: 249–259, 2007.

Stibor, M. *Lomové parametry betonu a jejich určování (Fracture parameters of concrete and their evaluation)*. Ph.D. thesis, Faculty of Civil Engineering, Brno University of Technology, Brno, Czech Republic, 2004, (in Czech).

Veselý, V. *Parametry betonu pro popis lomového chování (Parameters of concrete for description of its fracture behavior)*. Ph.D. thesis, Faculty of Civil Engineering, Brno University of Technology, Brno, Czech Republic, 2004, (in Czech).

Reliability Analysis of High-Rise Buildings under Wind Loads

Ming-Yi Liu[1] and Yu-Jie Li[2]
[1]*Department of Civil Engineering, Chung Yuan Christian University, Jhongli City, Taoyuan County 32023, Taiwan, myliu@cycu.edu.tw*
[2]*Department of Civil Engineering, Chung Yuan Christian University, Jhongli City, Taoyuan County 32023, Taiwan, g9772001@cycu.edu.tw*

Abstract: The objective of this paper is to conduct the reliability analysis of high-rise buildings under wind loads. Numerical examples are provided to capture the dynamic effects of structures with eccentricity between the elastic and mass centers. The framework of this research consists of two stages. The first stage includes two parts: the deterministic analysis of wind-induced acceleration for a variety of attack angles, i.e., the demand, and the determination of allowable acceleration based on the occupant comfort criteria for wind-excited buildings, i.e., the capacity. According to the results obtained in the first stage, the reliability analysis is conducted in the second stage, which can predict the probability of dissatisfaction with occupant comfort criteria for a variety of probability distributions of the structural eccentricity. The findings indicate that, compared to the lognormal and type I extreme value distributions, the normal distribution can be used to more conservatively simulate the uncertainties of the eccentricity between the elastic and mass centers. Furthermore, the probability of dissatisfaction with occupant comfort criteria of the torsionally coupled system is relatively higher than that of the torsionally uncoupled system for each attack angle due to the coupled mode effects.

Keywords: Reliability analysis, High-rise building, Wind load, Elastic center, Attack angle

1. Introduction

Traditionally, structural analysis is based on deterministic approaches, i.e., each parameter of analytical model is considered to be a certain value. In fact, uncertainties exist in design, construction, operation and maintenance of real structures. Consequently, traditional analysis is not able to effectively capture structural properties. On the basis of probabilistic approaches, reliability analysis is used to simulate probability distribution of each parameter, implying that uncertainties can be reasonably modeled by such method. This fact indicates that reliability analysis is a more appropriate tool than traditional analysis. In Taiwan, both structural safety and occupant comfort of high-rise buildings have become important due to frequent typhoons, implying that wind hazard is a significant factor for design purposes. Uncertainties of both wind loads and high-rise buildings have to be considered for structural design. From the above description, reliability analysis is useful for exploring the problem of high-rise buildings under wind loads.

The objective of this paper is to conduct the reliability analysis of high-rise buildings under wind loads. Numerical examples are provided to capture the dynamic effects of structures with eccentricity between the elastic and mass centers. The framework of this research consists of two stages, as shown in Figure 1. The first stage includes two parts: the deterministic analysis of wind-induced acceleration for a variety of attack angles, i.e., the demand, and the determination of allowable acceleration based on the occupant comfort

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

criteria for wind-excited buildings, i.e., the capacity. According to the results obtained in the first stage, the reliability analysis is conducted in the second stage, which can predict the probability of dissatisfaction with occupant comfort criteria for a variety of probability distributions of the structural eccentricity.

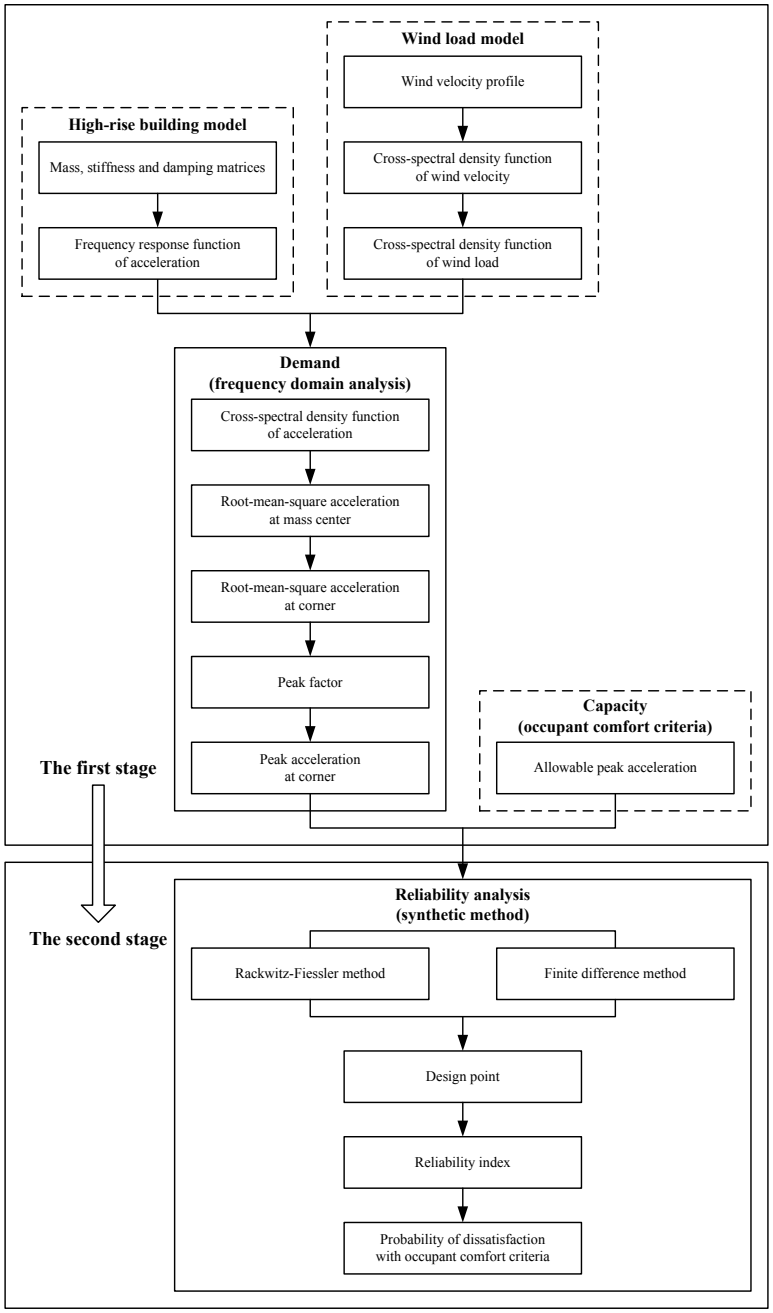


Figure 1. Framework of the research.

2. High-Rise Building Model

An N-story torsionally coupled system is used to simulate a high-rise building, and its corresponding three-dimensional configuration and top view of the i th floor are illustrated in Figure 2(a) and 2(b), respectively[1], where x, y, z and θ are the coordinates of the system; D_i, B_i, H_i and Z_i are the depth, breadth, height and elevation, respectively; MC_i, EC_i and AC_i are the mass, elastic and aerodynamic centers, respectively; Ex_i and Ey_i are the eccentricities between EC_i and MC_i in the x and y axes, respectively; Ax_i and Ay_i are the eccentricities between AC_i and MC_i in the x and y axes, respectively. Several assumptions are adopted in this study: (1) each rigid diaphragm with three degrees of freedom x_i, y_i and θ_i is characterized by the mass M_i and the moment of inertia I_i about MC_i; (2) each massless column is characterized by Kx_i, Ky_i and $K\theta_i$, which individually denote the stiffnesses in the x, y and θ axes referred to EC_i; (3) wind loads are applied at AC_i; (4) MC_i, EC_i and AC_i are non-coincident, and MC_i is located in the centroid of the diaphragm; (5) the Rayleigh damping with the mass-related coefficient A_0 and the stiffness-related coefficient A_1 is used. The model in Figure 2 can be simplified to an N-story torsionally uncoupled system when Ex_i and Ey_i both equal zero.

The procedure for the modeling of high-rise buildings is summarized in Figure 1. Based on the parameters mentioned above, the mass, stiffness and damping matrices of the system can be generated, and the frequency response function of acceleration can therefore be obtained (Kan and Chopra, 1977; Yang et al., 1981; Samali et al., 1985; Kareem, 1985; Kareem, 1992; Wu and Yang, 2000; Liu et al., 2008).

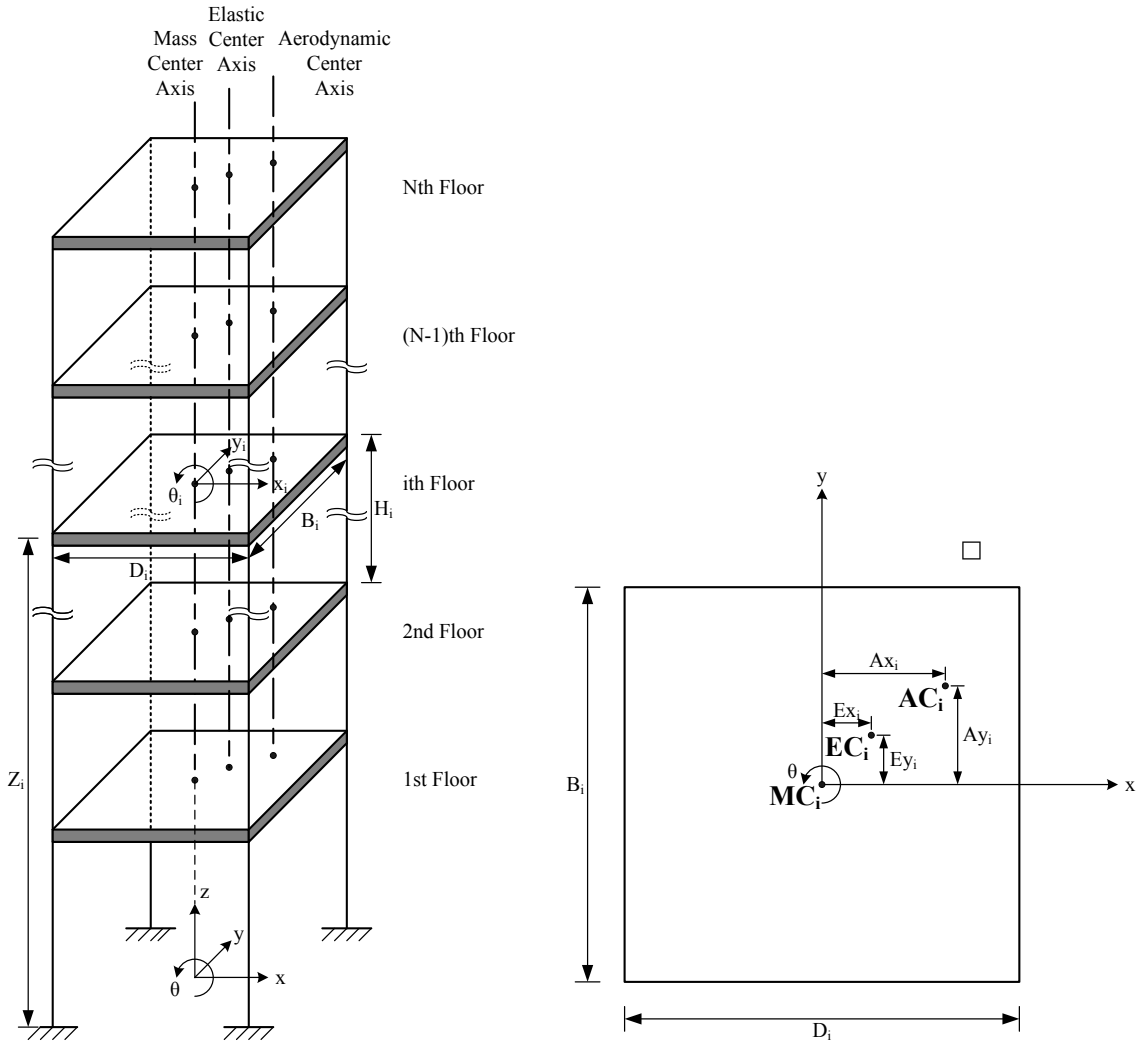
3. Wind Load Model

Wind loads can be decomposed into an average, aerodynamic damping and fluctuation terms. The fluctuation term is considered and the other two terms are neglected, which can be used to appropriately conduct the dynamic analysis under the assumption of small deformation theory. The wind load components including the drag, lift and torque are illustrated in Figure 3, where the attack angle ϕ is defined as the angle between the wind direction and the x axis. The drag and lift both act through AC_i, where the former and the latter are parallel and perpendicular to the wind direction, respectively. The torque is due to the eccentricity between AC_i and MC_i. The drag, lift and torque can be written as a function of ϕ (Yang et al., 1981; Samali et al., 1985; Wu and Yang, 2000; Simiu and Scanlan, 1996; Peng, 2005).

The procedure for the computation of wind loads is summarized in Figure 1. According to the power law, the wind velocity profile showing the variations in the mean wind velocity over the elevation can be expressed as a function of the exponent α, the gradient height Z_G and the gradient wind velocity V_G (Simiu and Scanlan, 1996). Based on the wind velocity profile, the reference mean wind velocity at 10 m above the ground V_R, the ground roughness coefficient K_0 and the exponential decay coefficient C_1 are used to calculate the cross-spectral density function of wind velocity between two elevations (Davenport,

[1] The subscript i in Figure 1 represents the parameter of the i th floor in this research.

1968). By combining both the wind velocity profile and cross-spectral density function of wind velocity, the cross-spectral density function of wind load between two axes can be obtained by the air density ρ, the windward side area of floor, the mean wind velocity, the drag coefficient C_D, the lift coefficient C_L, Ax_i and Ay_i for a variety of ϕ (Yang et al., 1981; Samali et al., 1985; Wu and Yang, 2000).



(a) Three-dimensional configuration

(b) Top view of the i th floor

Figure 2. N -story torsionally coupled system.

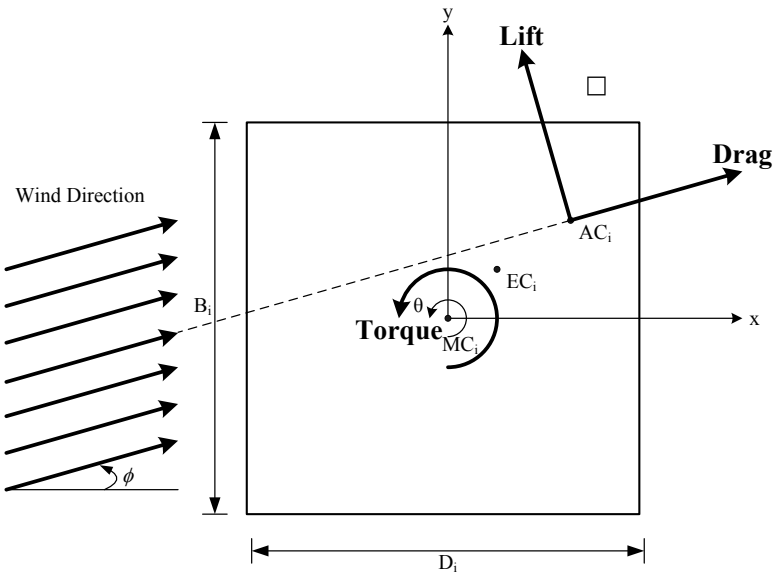


Figure 3. Wind load components.

4. Demand and Capacity

As shown in Figure 1, the deterministic analysis of wind-induced acceleration of structures with eccentricity between EC_i and MC_i for a variety of ϕ is conducted from the high-rise building and wind load models. The computational procedure is based on the frequency domain analysis. By combining both the frequency response function of acceleration and cross-spectral density function of wind load, the cross-spectral density function of acceleration between two axes, and the corresponding root-mean-square acceleration at mass center and that at corner can be calculated (Kareem, 1985; Kareem, 1992). The peak acceleration at corner, i.e., the demand, then can be obtained by multiplying the response at mass center by the peak factor (Melbourne, 1977).

The allowable peak acceleration of structures, i.e., the capacity, can be determined based on the occupant comfort criteria for wind-excited buildings. The threshold can be written as a function of the frequency of structural oscillation F, the duration of wind velocity T and the return period of wind velocity R (Melbourne and Palmer, 1992).

5. Reliability Analysis

According to the demand and capacity for different attack angles obtained in the first stage, the reliability analysis of high-rise buildings under wind loads based on the synthetic method combining both the Rackwitz-Fiessler and finite difference methods is conducted in the second stage, which can predict the

design point, reliability index and probability of dissatisfaction with occupant comfort criteria for a variety of probability distributions of the structural eccentricity, as shown in Figure 1.

The basic variables X_i ($i=1,2,\cdots,n$) and the corresponding limit state function $Z=g(X_1,X_2,\cdots,X_n)$ are the essences of reliability analysis. X_i ($i=1,2,\cdots,n$) can be used to simulate the uncertainties of n parameters in a system. Z is the standard to judge whether each performance criterion is satisfied in the system. Figure 4(a) illustrates the relationships between Z and X_i ($i=1,2,\cdots,n$) in the original coordinate system. The limit state ($Z=0$) is the boundary between the safe region ($Z>0$) and the unsafe region ($Z<0$). X_i ($i=1,2,\cdots,n$) can be transformed to the standard normal variables X_i' ($i=1,2,\cdots,n$), respectively, and the corresponding limit state function $Z=g(X_1',X_2',\cdots,X_n')$ can therefore be determined. The relationships between Z and X_i' ($i=1,2,\cdots,n$) in the transformed coordinate system are illustrated in Figure 4(b). The safe region ($Z>0$) and the unsafe region ($Z<0$) are divided by the limit state ($Z=0$) similar to Figure 4(a). The design point $(x_1'^*,x_2'^*,\cdots,x_n'^*)$[2] is located in $Z=0$ closest to the origin. The reliability index β with the distance between the design point and the origin is a useful index for assessing the system reliability. The probability of dissatisfaction with occupant comfort criteria p_f can be expressed as a function of β (Haldar and Mahadevan, 2000a; Haldar and Mahadevan, 2000b).

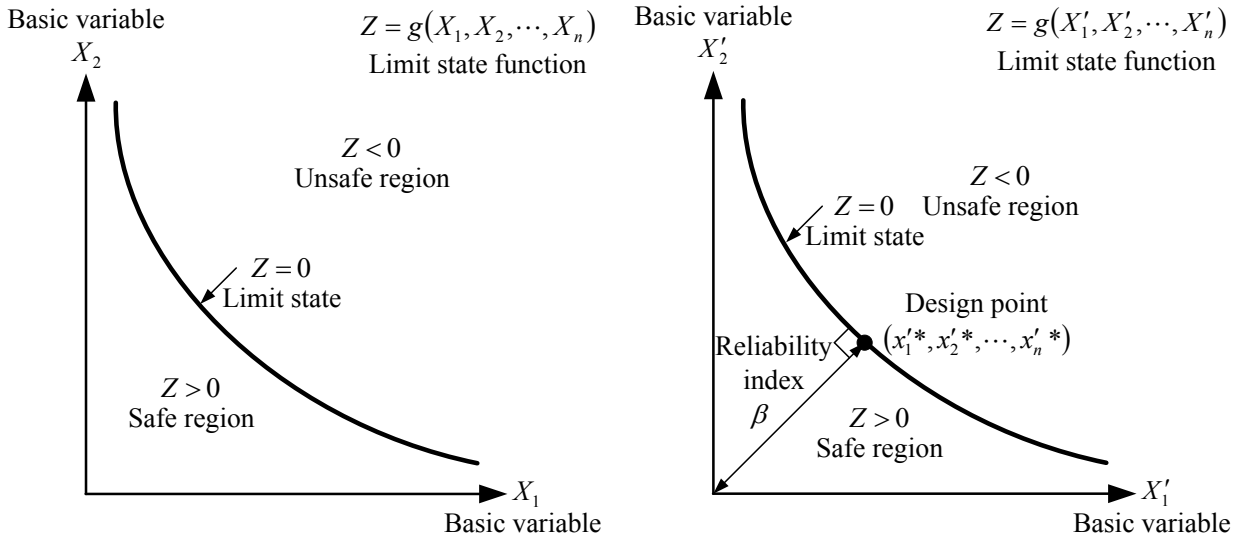


Figure 4. Relationships between the limit state function and the basic variables.

The Rackwitz-Fiessler method contains the parameters in both the original and transformed coordinate systems. The algorithm is formulated as follows (Rackwitz and Fiessler, 1978):

[2] The parameters marked with asterisk represent the ones based on the design point in this research.

Step 1. Z is defined.

Step 2. The initial value for each component of the design point in the original coordinate system $x_i^*(i=1,2,\cdots,n)$ is given. These values are assumed to be $\mu_{X_i}(i=1,2,\cdots,n)$ representing the mean values of $X_i(i=1,2,\cdots,n)$, respectively. The corresponding initial value of $g(x_1^*,x_2^*,\cdots,x_n^*)$ can therefore be determined.

Step 3. For the non-normal variables in $X_i(i=1,2,\cdots,n)$, both their mean values and standard deviations of the equivalent normal variables, i.e., $\mu_{X_i}^N$ and $\sigma_{X_i}^N(i=1,2,\cdots,n)$, respectively, can be estimated based on the Rackwitz-Fiessler transformation (Rackwitz and Fiessler, 1976). $x_i'^*(i=1,2,\cdots,n)$ then can be calculated as

$$x_i'^* = \frac{x_i^* - \mu_{X_i}^N}{\sigma_{X_i}^N}. \tag{1}$$

Step 4. $\left(\frac{\partial g}{\partial X_i}\right)^*(i=1,2,\cdots,n)$ are calculated.

Step 5. $\left(\frac{\partial g}{\partial X_i'}\right)^*(i=1,2,\cdots,n)$ can be calculated as

$$\left(\frac{\partial g}{\partial X_i'}\right)^* = \left(\frac{\partial g}{\partial X_i}\right)^* \sigma_{X_i}^N. \tag{2}$$

Step 6. $x_i'^*(i=1,2,\cdots,n)$ can be modified by the recursive formula

$$New(x_1'^*,x_2'^*,\cdots,x_n'^*) = \left[\frac{\sum_{i=1}^{n}\left(\frac{\partial g}{\partial X_i'}\right)^* x_i'^* - g(x_1^*,x_2^*,\cdots,x_n^*)}{\sum_{i=1}^{n}\left(\frac{\partial g}{\partial X_i'}\right)^{2^*}}\right]\left[\left(\frac{\partial g}{\partial X_1'}\right)^*,\left(\frac{\partial g}{\partial X_2'}\right)^*,\cdots,\left(\frac{\partial g}{\partial X_n'}\right)^*\right]. \tag{3}$$

Step 7. β can be calculated as

$$\beta = \sqrt{\sum_{i=1}^{n}\left(x_i'^*\right)^2}, \tag{4}$$

indicating that its value is equal to the distance between $(x_1'^*,x_2'^*,\cdots,x_n'^*)$ and the origin. The convergence tolerance for β is assigned.

Step 8. Eq. (1) can be rewritten as

$$x_i^* = \mu_{X_i}^N + \sigma_{X_i}^N x_i'^*, \tag{5}$$

implying that x_i^* ($i = 1,2,\cdots,n$) are modified and $g(x_1^*, x_2^*, \cdots, x_n^*)$ can therefore be redetermined. The convergence tolerance for $g(x_1^*, x_2^*, \cdots, x_n^*)$ is assigned. Steps 3 to 8 are repeated until the convergence tolerances for both β and $g(x_1^*, x_2^*, \cdots, x_n^*)$ are achieved.

Eqs. (1), (2) and (5) will also be valid for the normal variables in X_i ($i = 1,2,\cdots,n$) if $\mu_{X_i}^N$ and $\sigma_{X_i}^N$ are replaced by their mean values μ_{X_i} and standard deviations σ_{X_i} ($i = 1,2,\cdots,n$), respectively. p_f can be approximately estimated based on the convergent β.

The finite difference method used herein is a perturbation-based approach. The computational procedure is summarized as follows (Haldar and Mahadevan, 2000b):

Step 1. The initial values of X_i, i.e., X_i^0 ($i = 1,2,\cdots,n$), are assumed to be μ_{X_i} ($i = 1,2,\cdots,n$), respectively. The corresponding value of Z before perturbation can therefore be determined as

$$Z_0 = g(X_1^0, X_2^0, \cdots, X_n^0). \tag{6}$$

Step 2. The small and positive ΔX_1 with respect to the perturbation of X_1 is given. ΔX_1 is assumed to be proportional to σ_{X_1}. X_1^0 is replaced by $X_1^0 + \Delta X_1$ and X_i^0 ($i = 2,3,\cdots,n$) remain the previous values. The corresponding value of Z after perturbation can therefore be determined as

$$Z_1 = g(X_1^0 + \Delta X_1, X_2^0, \cdots, X_n^0). \tag{7}$$

Step 3. The difference of Z before and after perturbation can be calculated as

$$\Delta Z = Z_1 - Z_0. \tag{8}$$

The derivative of Z with respect to X_1 can be approximately estimated as $\frac{\Delta Z}{\Delta X_1}$.

Step 4. Similarly, the derivatives of Z with respect to X_i ($i = 2,3,\cdots,n$) can also be approximately obtained as $\frac{\Delta Z}{\Delta X_i}$ ($i = 2,3,\cdots,n$), respectively, by repeating Steps 2 and 3.

By individually substituting $g\left(X_1^0, X_2^0, \cdots, X_n^0\right)$ and $\frac{\Delta Z}{\Delta X_i}\,(i=1,2,\cdots,n)$ of the finite difference method for $g(x_1^*, x_2^*, \cdots, x_n^*)$ and $\left(\frac{\partial g}{\partial X_i}\right)^*\,(i=1,2,\cdots,n)$ of the Rackwitz-Fiessler method, these two methods can be combined (Haldar and Mahadevan, 2000b).

6. Numerical Examples

To illustrate the computational procedure in Figure 1, two numerical examples, i.e., the torsionally uncoupled and coupled systems, are provided to conduct the reliability analysis of high-rise buildings under wind loads for a variety of attack angles. The results can be used to capture the dynamic effects due to the structural eccentricity.

Four types of parameters: the high-rise building model, wind load model, occupant comfort criteria and reliability analysis, are considered in this study. All parameters of the two numerical examples are the same except the eccentricity between the elastic and mass centers. For the parameters of the high-rise building model, a 40-story building ($N=40$) with a height of 160 m is used. The geometric configuration and dynamic properties of each floor are assumed to be identical, as shown in Table I. Table II summarizes the parameters of the wind load model, where C_D and C_L are a function of ϕ (Peng, 2005). Table III lists the parameters of the occupant comfort criteria, where F is selected from the natural frequency of the first mode of each system. The parameters of the reliability analysis are illustrated in Table IV, where X_1 and X_2 are employed to simulate Ex_i and Ey_i, respectively. For both the torsionally uncoupled and coupled systems, $\mu_{X_1}=Ex_i$ and $\mu_{X_2}=Ey_i$ are given, and the probability distribution of X_1 and that of X_2 are assumed to be identical. Three types of probability distributions: the normal, lognormal and type I extreme value distributions, are used to model the uncertainties of both X_1 and X_2.

The acceleration at the top floor corner is the target of both the deterministic and reliability analyses. This is because such response is the maximum throughout the system. The relationships between the peak acceleration at corner of the 40th floor, i.e., the demand, and ϕ for the torsionally uncoupled and coupled systems are shown in Figure 5(a) and 5(b), respectively. The allowable peak acceleration independent of ϕ, i.e., the capacity, is displayed in the figures. Both the figures illustrate that the demand is comparatively lower than the capacity for each ϕ. Consequently, the occupant comfort criteria are satisfied in the two numerical examples from the viewpoint of deterministic approaches. These two figures also show that the maximum peak acceleration occurs when the wind direction is parallel to the x axis, i.e., $\phi=0^\circ$ or 180°. The peak acceleration of the torsionally coupled system is relatively higher than that of the torsionally uncoupled system for each ϕ due to the coupled mode effects.

Table I. Parameters of the high-rise building model.

D_i	[m]	24
B_i	[m]	48
H_i	[m]	4
Ex_i	[m]	0 (TUS)
		− 1.2 (TCS)
Ey_i	[m]	0 (TUS)
		− 2.4 (TCS)
Ax_i	[m]	− 0.5
Ay_i	[m]	4.8
M_i	[kg]	10^6
I_i	[kg-m^2]	2.4×10^8
Kx_i	[N/m]	7×10^9
Ky_i	[N/m]	1.05×10^{10}
$K\theta_i$	[N/rad]	3.5×10^{12}
A_0	[s^{-1}]	0.2
A_1	[s]	0

TUS: Torsionally uncoupled system
TCS: Torsionally coupled system

Table II. Parameters of the wind load model.

α	[None]	0.36
Z_G	[m]	500
V_G	[m/s]	27.4
V_R	[m/s]	6.5
K_0	[None]	0.025
C_1	[None]	7.7
ρ	[kg/m^3]	1.23
C_D	[None]	(Peng, 2005)
C_L	[None]	(Peng, 2005)

Table III. Parameters of the occupant comfort criteria.

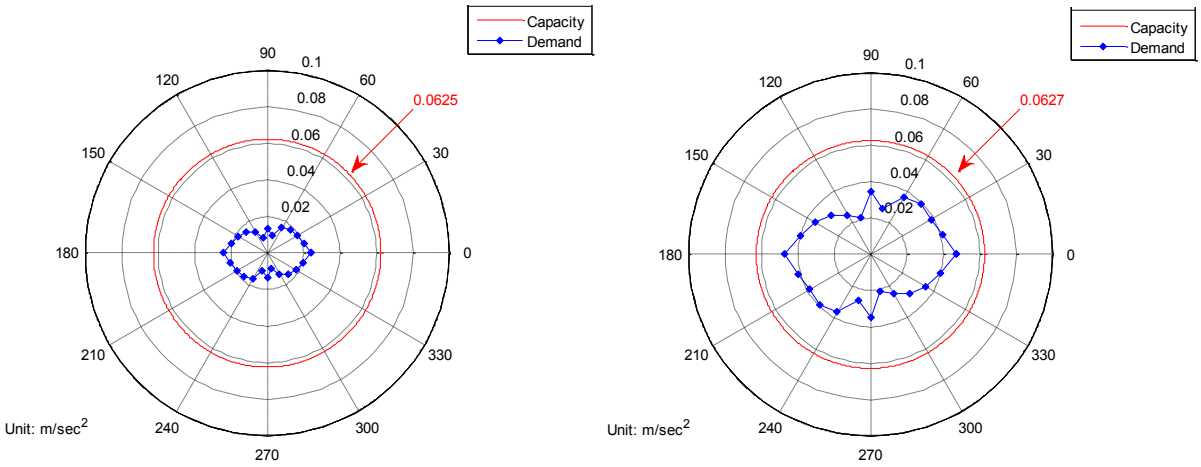
F	[Hz]	0.5164 (TUS)
		0.5106 (TCS)
T	[s]	600
R	[yr]	0.5

TUS: Torsionally uncoupled system
TCS: Torsionally coupled system

Table IV. Parameters of the reliability analysis.

μ_{X_1}	[m]	0 (TUS)
		− 1.2 (TCS)
μ_{X_2}	[m]	0 (TUS)
		− 2.4 (TCS)
σ_{X_1}	[m]	0.8
σ_{X_2}	[m]	0.8

TUS: Torsionally uncoupled system
TCS: Torsionally coupled system

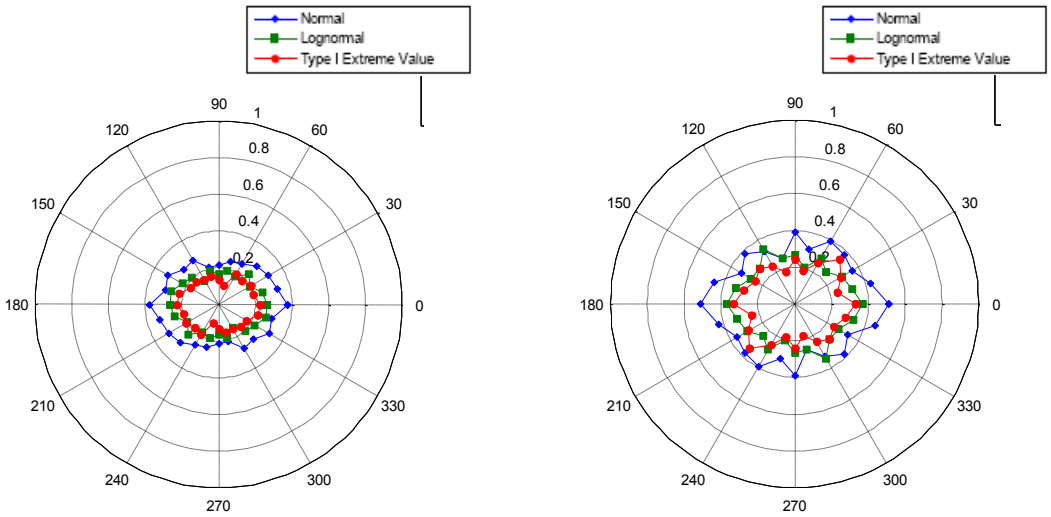


(a) Torsionally uncoupled system (b) Torsionally coupled system

Figure 5. Relationships between the peak acceleration at corner of the 40th floor and the attack angle.

The reliability analysis is conducted based on the demand and capacity obtained by the deterministic analysis. The relationships between the probability of dissatisfaction with occupant comfort criteria for three types of probability distributions of the structural eccentricity, i.e., the normal, lognormal and type I extreme value distributions, and ϕ for the torsionally uncoupled and coupled systems are shown in Figure 6(a) and 6(b), respectively. Both the figures illustrate that the probability for the case of the normal distribution is relatively higher than that for the other two cases for each ϕ. Furthermore, the probability for the case of the lognormal distribution is close to that of the type I extreme value distribution for each ϕ. The findings indicate that, compared to the lognormal and type I extreme value distributions, the normal distribution can be used to more conservatively simulate the uncertainties of the eccentricity between EC_i and MC_i in the two numerical examples from the viewpoint of probabilistic approaches. These two figures also show that the maximum probability occurs when the wind direction is parallel to the x axis, i.e., $\phi = 0^\circ$ or 180˚. The probability of the torsionally coupled system is relatively higher than that of the torsionally

uncoupled system for each ϕ due to the coupled mode effects. The results are in agreement with those obtained by the deterministic analysis.



(a) Torsionally uncoupled system (b) Torsionally coupled system

Figure 6. Relationships between the probability of dissatisfaction with occupant comfort criteria and the attack angle.

7. Conclusions

The objective of this paper is to conduct the reliability analysis of high-rise buildings under wind loads. Two numerical examples, i.e., the torsionally uncoupled and coupled systems, are provided to capture the dynamic effects of structures with eccentricity between the elastic and mass centers. The framework of this research consists of two stages. The first stage includes two parts: the deterministic analysis of wind-induced acceleration for a variety of attack angles, i.e., the demand, and the determination of allowable acceleration based on the occupant comfort criteria for wind-excited buildings, i.e., the capacity. According to the results obtained in the first stage, the reliability analysis is conducted in the second stage, which can predict the probability of dissatisfaction with occupant comfort criteria for three types of probability distributions of the structural eccentricity, i.e., the normal, lognormal and type I extreme value distributions.

In the first stage, both the examples illustrate that the demand is comparatively lower than the capacity for each attack angle. Consequently, the occupant comfort criteria are satisfied in the two numerical examples from the viewpoint of deterministic approaches. These two examples also show that the maximum peak acceleration occurs when the wind direction is parallel to the x axis. The peak acceleration of the torsionally coupled system is relatively higher than that of the torsionally uncoupled system for each attack angle due to the coupled mode effects.

In the second stage, both the examples illustrate that the probability for the case of the normal distribution is relatively higher than that for the other two cases for each attack angle. Furthermore, the

probability for the case of the lognormal distribution is close to that of the type I extreme value distribution for each attack angle. The findings indicate that, compared to the lognormal and type I extreme value distributions, the normal distribution can be used to more conservatively simulate the uncertainties of the eccentricity between the elastic and mass centers in the two numerical examples from the viewpoint of probabilistic approaches. These two examples also show that the maximum probability occurs when the wind direction is parallel to the *x* axis. The probability of the torsionally coupled system is relatively higher than that of the torsionally uncoupled system for each attack angle due to the coupled mode effects. The results are in agreement with those obtained by the deterministic analysis.

Acknowledgements

The authors would like to acknowledge the support of the National Science Council, Taiwan, under Grant No. NSC 98-2218-E-033-001.

References

Davenport, A. G. The Dependence of Wind Load upon Meteorological Parameters. *Proceedings of the International Research Seminar on Wind Effects on Buildings and Structures*, University of Toronto Press, Toronto, Canada, 1968.

Haldar, A. and S. Mahadevan. *Probability, Reliability, and Statistical Methods in Engineering Design*. John Wiley & Sons, Inc., New York City, New York, USA, 2000a.

Haldar, A. and S. Mahadevan. *Reliability Assessment Using Stochastic Finite Element Analysis*. John Wiley & Sons, Inc., New York City, New York, USA, 2000b.

Kan, C. L. and A. K. Chopra. Elastic Earthquake Analysis of Torsionally Coupled Multistorey Buildings. *Earthquake Engineering and Structural Dynamics*, 5(4):395-412, 1977.

Kareem, A. Lateral-Torsional Motion of Tall Buildings to Wind Loads. *Journal of Structural Engineering*, *ASCE*, 111(11):2479-2496, 1985.

Kareem, A. Dynamic Response of High-Rise Buildings to Stochastic Wind Loads. *Journal of Wind Engineering and Industrial Aerodynamics*, 42(1-3):1101-1112, 1992.

Liu, M. Y., W. L. Chiang, J. H. Hwang, and C. R. Chu. Wind-Induced Vibration of High-Rise Building with Tuned Mass Damper Including Soil-Structure Interaction. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(6-7):1092-1102, 2008.

Melbourne, W. H. Probability Distributions Associated with the Wind Loading of Structures. *Civil Engineering Transactions, Institution of Engineers, Australia*, CE19(1):58-67, 1977.

Melbourne, W. H. and T. R. Palmer. Accelerations and Comfort Criteria for Buildings Undergoing Complex Motions. *Journal of Wind Engineering and Industrial Aerodynamics*, 41(1-3):105-116, 1992.

Peng, M. H. *Investigation of Wind Forces on Objects in Gusty Flow Using Hilbert-Huang Transform*. Master thesis, National Central University, Taiwan, 2005.

Rackwitz, R. and B. Fiessler. Note on Discrete Safety Checking when Using Non-Normal Stochastic Models for Basic Variables. *Load Project Working Session*, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, 1976.

Rackwitz, R. and B. Fiessler. Structural Reliability under Combined Random Load Sequences. *Computers and Structures*, 9(5):489-494, 1978.

Samali, B., J. N. Yang, and C. T. Yeh. Control of Lateral-Torsional Motion of Wind-Excited Buildings. *Journal of Engineering Mechanics*, *ASCE*, 111(6):777-796, 1985.

Simiu, E. and R. H. Scanlan. *Wind Effects on Structures: Fundamentals and Applications to Design*. Third Edition, John Wiley & Sons, Inc., New York City, New York, USA, 1996.

Wu, J. C. and J. N. Yang. LQG Control of Lateral-Torsional Motion of Nanjing TV Transmission Tower. *Earthquake Engineering and Structural Dynamics*, 29(8):1111-1130, 2000.

Yang, J. N., Y. K. Lin, and B. Samali. Coupled Motion of Wind-Loaded Multi-Story Building. *Journal of Engineering Mechanics Division*, *ASCE*, 107(EM6):1209-1226, 1981.

Interval forms of a sixth-order class of modified Ostrowski methods

Taher Lotfi, Parisa Bakhtiari, Katayoun Mahdiani
Department of Applied Mathematics and Computer Sciences , Hamedan Branch,
Islamic Azad University, Hamadan, Iran
Email: lotfi@iauh.ac.ir, bakhtiari@iauh.ac.ir, mahdiani@iauh.ac.ir

Abstract. In this paper, interval extension of a sixth-order class of the classic modified Ostrowski methods which improves the order of convergence of Ostrowski method from four to six for solving nonlinear equations is introduce. Also, error analysis and convergence will be discussed. Some implemented examples with INTLAB are also included to illustrate the validity and applicability of the scheme. The results of proposed method are compared with the results of the interval Ostrowski method and the interval Newton method.

Keywords: interval analysis, nonlinear equations, a sixth-order class of the modified Ostrowski method.

1. Introduction

1.1. Background

Interval analysis was formally introduced by Moore (Moore, 1962; Moore, 1966). In practice, interval analysis provides rigorous enclosure of solutions to the given model equations. In fact, interval algorithms are designed to automatically provide rigorous bounds on accumulated rounding errors, approximation errors, and propagated uncertainties in initial data during the process of the computation (Moore et al., 2009; Neumaier, 1990).

There are many iterative methods to find a simple root of a nonlinear equation. The classical Newton method is one of the most important these methods with convergence of quadratic. To improve the local order of convergence and efficiency index, many modified methods have been proposed in (Grau And Barreo, 2006; Liu And Wanh, 2010; Bi et al., 2009; Bi et al., 2009; Thukral, 2008; Kou, 2007; Ham, 200; Kou, 2010; Sharma et al., 2010; Sharma et al., 2007; Ostrowski, 1973). One of these important and basic methods is the Ostrowski method (Ostrowski, 1973). This method is a basic for many modified methods. The order of convergence of this method increases by at least two at the expense of additional function evaluation at another point iterated by the Newton method.

There are a class of the modified ostrowski methods that improves the order of convergence of the Ostrowski method with an additional function evaluation (Chun And Ham, 2007). The local order of convergence of this method is 6 (Chun And Ham, 2007).

An interval Newton method has been developed for solving nonlinear equations. This verified approach enables us to compute interval enclosures for the exact values of the solution with sharp bounds (Moore et al., 2009).

In the present article, using the interval extension of the Newton method in (Moore et al., 2009), an interval extension of a sixth-order class of the classic modified Ostrowski methods (Chun And Ham, 2007) is introduced to find the enclosure roots of nonlinear equations. Convergence rate of the proposed method is

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

also examined. Moreover, error bound and comparison of this method with interval Newton and Ostrowski methods are given. Applicability and reliability of this algorithm will be investigated and justified through some examples implemented by using INTLAB, which is free to use (Rump, 199).

1.2. PRELIMINARIES: NOTATIONS AND RESULTS

We first introduce some basic notations and properties of interval arithmetic from (Moore et al., 2009; Neumaier, 2001). An interval number is a closed set in $\mathbb{R}$ that includes the possible range of an unknown real number, where $\mathbb{R}$ denotes the set of real numbers. Therefore, a real interval is a set of the form $\boldsymbol{x} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}]$, where $\underline{\boldsymbol{x}}$ and $\overline{\boldsymbol{x}}$ are the lower and upper bounds (end-points) of the interval number $\boldsymbol{x}$, respectively. The set of compact real intervals is denoted by $\mathbb{IR} = \{\boldsymbol{x} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}] | \quad \underline{\boldsymbol{x}}, \overline{\boldsymbol{x}} \in \mathbb{R}, \quad \underline{\boldsymbol{x}} \leq \overline{\boldsymbol{x}}\}$.

A real number x is identified with a point interval $\boldsymbol{x} = [x, x]$ and is called *degenerated interval*. The quality of interval analysis is measured by the width of the interval results, and a sharp enclosure for the exact solution is desirable. The *mid-point* and the *width* of an interval $\boldsymbol{x}$ are denoted by $\mathrm{mid}\,(\boldsymbol{x}) = (\underline{\boldsymbol{x}}+\overline{\boldsymbol{x}})/2$, and $\mathrm{wid}\,(\boldsymbol{x}) = \overline{\boldsymbol{x}} - \underline{\boldsymbol{x}}$, respectively.

Considering $|\boldsymbol{x}| = max\{|\underline{\boldsymbol{x}}|, |\overline{\boldsymbol{x}}|\}$ for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}$ and $a, b \in \mathbb{R}$ we can conclude that (Moore et al., 2009):

$$\mathrm{wid}\,(a\boldsymbol{x} + b\boldsymbol{y}) = |a|\mathrm{wid}\,(\boldsymbol{x}) + |b|\mathrm{wid}\,(\boldsymbol{y}),$$

$$\mathrm{wid}\,(\boldsymbol{x}\boldsymbol{y}) \leq |\boldsymbol{x}|\mathrm{wid}\,(\boldsymbol{y}) + |\boldsymbol{y}|\mathrm{wid}\,(\boldsymbol{x}),$$

Definition 1.1. We say that $\boldsymbol{f}$ is an *interval extension* of f, if for degenerate interval arguments, $\boldsymbol{f}$ agrees with f, i.e. $\boldsymbol{f}([x, x]) = f(x)$.

It should be noted that in general $\boldsymbol{f}$ is not the set image of f. Generally $f(\boldsymbol{x}) \subseteq \boldsymbol{f}(\boldsymbol{x})$. Besides, when $\boldsymbol{f}$ is an inclusion function of f, then we can directly obtain lower and upper bounds of $\boldsymbol{f}$ over any interval $\boldsymbol{x}$ within the domain of $\boldsymbol{f}$ just by taking $\underline{\boldsymbol{f}}(\boldsymbol{x})$ and $\overline{\boldsymbol{f}}(\boldsymbol{x})$, respectively.

Definition 1.2. An interval extension $\boldsymbol{f}$ is said to be *Lipschitz* in $\boldsymbol{x}^{(0)}$ if there is a constant L such that $\mathrm{wid}\,(\boldsymbol{f}(\boldsymbol{x})) \leq L\,\mathrm{wid}\,(\boldsymbol{x})$ for every $\boldsymbol{x} \subseteq \boldsymbol{x}^{(0)}$.

Hence, the width of $\boldsymbol{f}(\boldsymbol{x})$ approaches zero at least *linearly* with the width of $\boldsymbol{x}$.

Lemma 1.3 (See (Moore et al., 2009))**.** *If $\boldsymbol{f}$ is a natural interval extension of a real rational function with $\boldsymbol{f}(\boldsymbol{x})$ defined for $\boldsymbol{x} \subseteq \boldsymbol{x}^0$, where $\boldsymbol{x}$ and $\boldsymbol{x}^{(0)}$ are intervals, then $\boldsymbol{f}$ is Lipschitz in $\boldsymbol{x}^{(0)}$; In other words:*

$$\mathrm{wid}\,(\boldsymbol{f}(\boldsymbol{x})) \leq L\,\mathrm{wid}\,(\boldsymbol{x}), \tag{1}$$

Definition 1.4. An interval sequence $\{\boldsymbol{x}^{(k)}\}$ is *nested* if $\boldsymbol{x}^{(k+1)} \subseteq \boldsymbol{x}^{(k)}$ for all k.

Lemma 1.5 (See (Moore et al., 2009))**.** *Suppose $\{\boldsymbol{x}^{(k)}\}$ is such that there is a real number $x \in \boldsymbol{x}^{(k)}$ for all k. Define $\{\boldsymbol{y}^{(k)}\}$ by $\boldsymbol{y}^{(1)} = \boldsymbol{x}^{(1)}$ and $\boldsymbol{y}^{(k+1)} = \boldsymbol{x}^{(k+1)} \cap \boldsymbol{y}^{(k)}$ for all $k = 1, 2, \cdots$. Then $\boldsymbol{y}^{(k)}$ is nested with limit $\boldsymbol{y}$, and*

$$x \in \boldsymbol{y} \subseteq \boldsymbol{y}^{(k)} \quad \forall k.$$

Lemma 1.6 (See (Moore et al., 2009))**.** *Every nested sequence $\{\boldsymbol{x}^{(k)}\}$ converges and has the limit $\cap_{k=1}^{\infty}\boldsymbol{x}^{(k)}$.*

1.2.1. *Interval Newton method*

The Newton's method is the well-known iterative method for finding a simple zero of function. Let f be a real-valued function of a real variable x, and suppose that f is continuously differentiable.

Let $\boldsymbol{f}'(\boldsymbol{x})$ be an inclusion monotonic interval extension of $f'(x)$ and consider the algorithm

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} \cap \boldsymbol{N}(\boldsymbol{x}^{(k)}), \qquad (k = 0, 1, 2, \cdots), \tag{2}$$

where

$$\boldsymbol{N}(\boldsymbol{x}) = \operatorname{mid}(\boldsymbol{x}) - \frac{f(\operatorname{mid}(\boldsymbol{x}))}{\boldsymbol{f}'(\boldsymbol{x})}. \tag{3}$$

This is well-known as the interval Newton method (Moore et al., 2009).

Theorem 1.7 (See (Moore et al., 2009))**.** *If an interval $\boldsymbol{x}^{(0)}$ contains a zero x^* of $f(x)$, then so does $\boldsymbol{x}^{(k)}$ for all $k = 0, 1, 2, \cdots$, defined by (2). Furthermore, the intervals $\boldsymbol{x}^{(k)}$ form a nested sequence converging to x^* if $0 \notin \boldsymbol{f}'(\boldsymbol{x}^{(0)})$.*

The interval Newton method (2) is asymptotically error squaring.

Theorem 1.8 (See (Moore et al., 2009))**.** *Given a real rational function f of a single real variable x with rational extensions $\boldsymbol{f}$, $\boldsymbol{f}'$ of f, f', respectively, such that f has a simple zero x^* in an interval $\boldsymbol{x}^{(0)}$ for which $\boldsymbol{f}(\boldsymbol{x}^{(0)})$ is defined and $\boldsymbol{f}'(\boldsymbol{x}^{(0)})$ is defined and does not contain zero i.e. $0 \notin \boldsymbol{f}'(\boldsymbol{x}^{(0)})$. Then there is a positive real number C such that*

$$\operatorname{wid}(\boldsymbol{x}^{(k+1)}) \leq C \left(\operatorname{wid}(\boldsymbol{x}^{(k)})\right)^2.$$

If $0 \notin \boldsymbol{f}'(\boldsymbol{x}^{(0)})$, then $0 \notin \boldsymbol{f}'(\boldsymbol{x}^{(k)})$ for all k and $\operatorname{mid}(\boldsymbol{x}^{(k)})$ is not contained in $\boldsymbol{N}(\boldsymbol{x}^{(k)})$, unless $f(\operatorname{mid}(\boldsymbol{x}^(k)) = 0$. So, convergence of the sequence follows (Moore et al., 2009; Neumaier, 2001).

1.2.2. *Classic Ostrowski method*

The Ostrowski method is an important and basic method for finding a simple root of a nonlinear equation. This method is given by

$$x_{n+1} = S(x_n, y_n), \tag{4}$$

$$S(x_n, y_n) = y_n - \frac{f(y_n)}{f(x_n) - 2f(y_n)} \frac{f(x_n)}{f'(x_n)},$$

$$y_n = x_n - \frac{f(x_n)}{f'(x_n)}.$$

This method is an improvement of the Newton method with the order of convergence equal to 4. A detailed describe of this method is given in (Ostrowski, 1973).

1.3. A SIXTH-ORDER CLASS OF THE CLASSIC MODIFIED OSTROWSKI METHODS

There is a class of the modified Ostrowski methods that uses the same information of the Ostrowski method i.e., three evaluations of function f and one evaluation of derivative f' per step. This method is an improvement of the Ostrowski method. The local order of convergence of this method is 6. This iteration method is given by

$$x_{n+1} = G(x_n, y_n, z_n), \tag{5}$$

$$G(x_n, y_n, z_n) = z_n - H(u_n)\frac{f(z_n)}{f'(x_n)}, \quad u_n = \frac{f(y_n)}{f(x_n)},$$

$$z_n = y_n - \frac{f(x_n)}{f(x_n) - 2f(y_n)}.\frac{f(y_n)}{f'(x_n)},$$

$$y_n = x_n - \frac{f(x_n)}{f'(x_n)},$$

where $H(t)$ represents any real-value function that satisfies the properties $H(0) = 1,\ H'(0) = 2,\ |H''(0)| < \infty$.

A detailed describe of this method is given in (Chun And Ham, 2007).

2. Main results

In this section a new interval method of *a class of the modified Ostrowski methods* is introduced for computing a simple zero of a nonlinear equation. Our proposed approach is based on interval analysis which was first invented by R. Moore (Moore, 1962; Moore, 1966; Moore et al., 2009) and a sixth-order class of modified Ostrowski methods (Chun And Ham, 2007). This method is an improvement of its previous method (Chun And Ham, 2007) which increases the local order of convergence. It is proved that the order of convergence of the proposed method is 6 while the order of convergence of the Ostrowski method is 4.

2.1. INTERVAL FORM OF A SIXTH-ORDER CLASS OF THE MODIFIED OSTROWSKI METHODS

Let $\boldsymbol{f'}(\boldsymbol{x})$ is inclusion monotonic interval extension of $f'(x)$. We seek a solution of the equation $f(x) = 0$, on interval $\boldsymbol{x} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}]$. Interval extension of (5) is introduced as

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} \cap \boldsymbol{G}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{z}^{(k)}), \quad k = 0, 1, 2, \ldots, \tag{6}$$

where

$$\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \text{mid}\,(\boldsymbol{z}) - H(\text{mid}\,(\boldsymbol{u}_n))\frac{f(\text{mid}\,(\boldsymbol{z}))}{\boldsymbol{f'}(\boldsymbol{x})}, \quad \boldsymbol{u}_n = \frac{\boldsymbol{f}(\boldsymbol{y})}{f(\text{mid}\,(\boldsymbol{x}))}, \tag{7}$$

$$\boldsymbol{z}^{(k)} = \boldsymbol{x}^{(k)} \cap \boldsymbol{S}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}), \tag{8}$$

$$\boldsymbol{S}(\boldsymbol{x}, \boldsymbol{y}) = \text{mid}\,(\boldsymbol{y}) - \frac{f(\text{mid}\,(\boldsymbol{x}))}{f(\text{mid}\,(\boldsymbol{x})) - 2f(\text{mid}\,(\boldsymbol{y}))}.\frac{f(\text{mid}\,(\boldsymbol{y}))}{\boldsymbol{f'}(\boldsymbol{x})}, \tag{9}$$

$$\boldsymbol{y}^{(k)} = \boldsymbol{x}^{(k)} \cap \boldsymbol{N}(\boldsymbol{x}^{(k)}), \tag{10}$$

$$\boldsymbol{N}(\boldsymbol{x}) = \operatorname{mid}(\boldsymbol{x}) - \frac{f(\operatorname{mid}(\boldsymbol{x}))}{\boldsymbol{f}'(\boldsymbol{x})}, \tag{11}$$

where $H(t)$ represents any real-value function that satisfies the properties $H(0) = 1,\ H'(0) = 2$. Moreover, It is necessary to say that interval extension of the classic Ostrowski method is given in (8)–(10).

Now a computational algorithm for finding inclosure roots of a given nonlinear equation based on our proposed method by using above relations is introduced.

Algorithm
To summarize the previous development, the following computational algorithm is produced : To find an inclosure solution to $f(x) = 0$ given an initial interval $\boldsymbol{x}^{(0)}$ included one root:

INPUT initial interval $\boldsymbol{x}^{(0)}$;
tolerance TOL;
maximum number of iteration J;
functions f, f', $\boldsymbol{f}'$.

for k=0: J-1
Compute $\boldsymbol{N}(\boldsymbol{x}^{(k)})$ from (10).
$\boldsymbol{y}^{(k)} := \boldsymbol{N}(\boldsymbol{x}^{(k)}) \cap \boldsymbol{x}^{(k)}$.
Compute $\boldsymbol{S}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)})$ from (9).
$\boldsymbol{z}^{(k)} := \boldsymbol{S}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}) \cap \boldsymbol{x}^{(k)}$.
Compute $\boldsymbol{G}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{z}^{(k)})$ from (7).
$\boldsymbol{x}^{(k+1)} := \boldsymbol{G}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{z}^{(k)}) \cap \boldsymbol{x}^{(k)}$.
If wid $(\boldsymbol{x}^{(k+1)}) \leq TOL$, then go to OUTPUT STEP
end

OUTPUT ($\boldsymbol{x}^{(k+1)}$); (The procedure was successful.)

2.2. Convergence analysis and error bounds

In this section, we'll deal with the convergence and error bound for interval method (6). Unlike the classic form of this method, the interval version always displays a very regular behavior. To begin with, we will assume that $f : \boldsymbol{x} \to \mathbb{R}$ is a continuously differentiable function, and $x^* \in \boldsymbol{x}$ is a root of f. We also assume that an interval extension of f' exist and satisfies $0 \notin \boldsymbol{f}'(\boldsymbol{x})$. In particular, this implies that $f'(x) \neq 0$ for all $x \in \boldsymbol{x}$. The sequence of the proposed interval method has some nice properties.

Theorem 2.1 (Interval form of a sixth-order class of the modified Ostrowski methods)**.** *Assume $f \in C(\boldsymbol{x}^{(0)})$ and $0 \notin \boldsymbol{f}'(\boldsymbol{x}^{(k)})$ for $k = 0, 1, 2, \cdots$. If $\boldsymbol{x}^{(0)}$ contains a root x^* of f, then so do all intervals $\boldsymbol{x}^{(k)},\ k = 1, 2, \cdots$, generated by (6). Besides, the intervals $\boldsymbol{x}^{(k)}$ form a nested sequence converging to x^*.*

Proof. By induction, since $0 \notin \boldsymbol{f}'(\boldsymbol{x}^{(k)})$, if $x^* \in \boldsymbol{x}^{(0)}$ then $x^* \in \boldsymbol{x}^{(k)}$ for $k = 1, 2, \cdots$. Also, Lemma (1.5) leads that the intervals $\boldsymbol{x}^{(k)}$, generated by (6), form a nested sequence. Therefore, since for $k = 0, 1, 2, \cdots$ we have $x^* \in \boldsymbol{x}^{(k)}$ then $x^* \in \cap_k \boldsymbol{x}^{(k)}$ or $\lim_{n\to\infty} \cap_{k=0}^{n} \boldsymbol{x}^{(k)} = x^*$ and the proof is completed. ■

One of the most useful properties of the interval operator $\boldsymbol{G}$ of this method is that we are provided with a means of detecting when a region does not contain a root of f. As this is a common situation, it is important that we can quickly discard a set on the grounds of it containing no roots. Another important contribution from the properties of $\boldsymbol{G}$ is a simple verifiable condition that guarantees the existence of a unique root within an interval.

Theorem 2.2. *Suppose $f \in C^2(\boldsymbol{x}^{(0)})$ and $0 \notin \boldsymbol{f}'(\boldsymbol{x}^{(k)})$ for $k = 0, 1, 2, \cdots$.*

(1) If $x^* \in \boldsymbol{x}^{(0)}$ and $\boldsymbol{G}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}\, \boldsymbol{z}^{(k)}) \subseteq \boldsymbol{x}^{(k)}$, then $\boldsymbol{x}^{(k)}$ contains exactly one zero of f.

(2) If $\boldsymbol{x}^{(k)} \cap \boldsymbol{G}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{z}^{(k)}) = \emptyset$, then $\boldsymbol{x}^{(k)}$ does not contain any zero of f.

Proof. First, part one is proved. Since $0 \notin \boldsymbol{f}'(\boldsymbol{x}^{(k)})$, then $f'(x) \neq 0$ for all $x \in \boldsymbol{x}^{(k)}$ and therefore f is monotonic on $\boldsymbol{x}^{(k)}$. In other words, it has at most one zero in $\boldsymbol{x}$. Hence, it is sufficient to find a zero $x^* \in \boldsymbol{x}^{(k)}$. Since $\boldsymbol{G}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{z}^{(k)}) \subseteq \boldsymbol{x}^{(k)}$, using the Theorem (2.1) it clear that the f has exactly one root in $\boldsymbol{x}^{(k)}$.

To establish part (2), suppose x^* is a zero of f and $x^* \in \boldsymbol{x}^{(0)}$, then Theorem (2.1) results $x^* \in \boldsymbol{G}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}\, \boldsymbol{z}^{(k)})$. Consequently $x^* \in \boldsymbol{x}^{(k)} \cap \boldsymbol{G}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{z}^{(k)})$ which is contradiction. So the proof is completed. ■

If we start with an $\boldsymbol{x}^{(0)}$ such that

$$\boldsymbol{G}(\boldsymbol{x}^{(0)}, \boldsymbol{y}^{(0)}, \boldsymbol{z}^{(0)}) \subseteq \boldsymbol{S}(\boldsymbol{x}^{(0)}, \boldsymbol{y}^{(0)}) \subseteq \boldsymbol{N}(\boldsymbol{x}^{(0)}) \subseteq \boldsymbol{x}^{(0)},$$

then Theorems (2.1–2.2) guarantee a nested sequence of intervals $\{\boldsymbol{x}^{(k)}\}$ convergent to an interval $\boldsymbol{x}^*$ such that $x^* \in \boldsymbol{x}^{(0)}$ and $\boldsymbol{x}^* = \boldsymbol{G}(\boldsymbol{x}^*, \boldsymbol{x}^*, \boldsymbol{x}^*)$ and $\boldsymbol{x}^* \subseteq \boldsymbol{x}^{(k)}$ for all $k = 0, 1, 2, \cdots$. On a computer, the procedure can be stopped when $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)}$ or $\operatorname{wid}(\boldsymbol{x}^{(k)}) \leq \varepsilon$; using IA, interval arithmetic, at a specific number of digits, this yields the narrowest possible interval containing $\boldsymbol{x}^*$, inclusion solution.

Remark 2.3. In much the same way of Theorems (2.1–2.2) it is possible to state the interval Ostrowski method (8–10) convergent conditions. We just show it's convergence rate in Theorem (2.4).

The sequence (6) converges to x^* if the assumptions of the Theorems (2.1–2.2) are hold. Under conditions similar to those of Theorem (1.8), it is possible to show that the convergence rate is 6. First we need to show that the convergence rate of interval Ostrowski method, (8), is 4. A formal statement of this property ia as follows.

Theorem 2.4. *Assume that $f \in C(\boldsymbol{x}^0)$ with $0 \notin \boldsymbol{f}'(\boldsymbol{x}^{(0)})$, and f has a unique simple root $x^* \in \boldsymbol{x}^{(0)}$. Then, if $\boldsymbol{S}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}) \subseteq \boldsymbol{x}^{(k)}$, the sequence (8) has convergent rate four, i.e., there exists a constant K such that*

$$\operatorname{wid}(\boldsymbol{z}^{(k)}) \leq K \left(\operatorname{wid}(\boldsymbol{x}^{(k)})\right)^4. \tag{12}$$

Proof. By Mean Value Theorem we have

$$f(\operatorname{mid}(\boldsymbol{x}^{(k)})) = f'(\xi)[\operatorname{mid}(\boldsymbol{x}^{(k)}) - x^*],$$

with ξ is between $\mathrm{mid}\,(\boldsymbol{x}^{(k)})$ and x^*. Since $\boldsymbol{S}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}) \subseteq \boldsymbol{x}^{(k)}$, thus from (8), we have

$$\boldsymbol{z}^{(k)} = \mathrm{mid}\,(\boldsymbol{y}^{(k)}) - \frac{[\mathrm{mid}\,(\boldsymbol{x}^{(k)}) - x^*]f'(\xi_1)}{f(\mathrm{mid}\,(\boldsymbol{x}^{(k)})) - 2f(\mathrm{mid}\,(\boldsymbol{y}^{(k)}))} \frac{[\mathrm{mid}\,(\boldsymbol{y}^{(k)}) - x^*]f'(\xi_2)}{\boldsymbol{f}'(\boldsymbol{x}^{(k)})}, \tag{13}$$

therefore

$$\mathrm{wid}\,(\boldsymbol{z}^{(k)}) = \frac{|\mathrm{mid}\,(\boldsymbol{x}^{(k)}) - x^*||f'(\xi_1)||\mathrm{mid}\,(\boldsymbol{y}^{(k)}) - x^*||f'(\xi_2)|}{|f(\mathrm{mid}\,(\boldsymbol{x}^{(k)})) - 2f(\mathrm{mid}\,(\boldsymbol{y}^{(k)}))|} \mathrm{wid}\,\left(\frac{1}{\boldsymbol{f}'(\boldsymbol{x}^{(k)})}\right). \tag{14}$$

It is clear that

$$|\mathrm{mid}\,(\boldsymbol{x}^{(k)}) - x^*| \leq \mathrm{wid}\,(x^{(k)}). \tag{15}$$

Also, from Lemma (1.3) we have

$$\mathrm{wid}\,\left(\frac{1}{\boldsymbol{f}'(\boldsymbol{x}^{(k)})}\right) \leq \mathrm{wid}\,(\boldsymbol{x}^{(k)}). \tag{16}$$

Furthermore, since $\boldsymbol{y}^{(k)}$ is generated from (10), Theorem (1.8) leads

$$|\mathrm{mid}\,(\boldsymbol{y}^{(k)}) - x^*| \leq \mathrm{wid}\,(\boldsymbol{y}^{(k)}) \leq \left(\mathrm{wid}\,(\boldsymbol{x}^{(k)})\right)^2. \tag{17}$$

Let $|f'(\xi_1)| \leq K_1$, $|f'(\xi_2)| \leq K_2$ and $|f(\mathrm{mid}\,(\boldsymbol{x}^{(k)})) - 2f(\mathrm{mid}\,(\boldsymbol{y}^{(k)}))| \leq K_3$, so, considering (14–17), we have the following error bound

$$\mathrm{wid}\,(\boldsymbol{z}^{(k)}) \leq \frac{K_1K_2}{K_3}\left(\mathrm{wid}\,(\boldsymbol{x}^{(k)})\right)^4 = K\left(\mathrm{wid}\,(\boldsymbol{x}^{(k)})\right)^4, \tag{18}$$

where $K = K_1K_2/K_3$, and the proof is completed. ■

The interval form of this class of modified Ostrowski methods, (6), improves the local order of convergence of the interval Ostrowski method, (8), with an additional evaluation of the function.

Theorem 2.5. *Assume that $f \in C(\boldsymbol{x}^0)$ with $0 \notin \boldsymbol{f}'(\boldsymbol{x}^{(0)})$, and f has a unique simple root $x^* \in \boldsymbol{x}^{(0)}$. Then, if $\boldsymbol{G}(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{z}^{(k)}) \subseteq \boldsymbol{x}^{(k)}$, the sequence (6) has convergent rate **6**, i.e., there exists a constant K such that*

$$\mathrm{wid}\,(\boldsymbol{x}^{(k+1)}) \leq K\left(\mathrm{wid}\,(\boldsymbol{x}^{(k)})\right)^6. \tag{19}$$

`Proof.` By Mean Value Theorem we have

$$f(\mathrm{mid}\,(\boldsymbol{x}^{(k)})) = f'(\xi)\left[\mathrm{mid}\,(\boldsymbol{x}^{(k)}) - x^*\right],$$

with ξ is between $\mathrm{mid}\,(\boldsymbol{x}^{(k)})$ and x^*. Therefore, from (6), we have

$$\boldsymbol{x}^{(k+1)} = \mathrm{mid}\,(\boldsymbol{z}^{(k)}) - \frac{\left[\mathrm{mid}\,(\boldsymbol{u}_n^{(k)}) - x^*\right] H'(\xi_1)\left[\mathrm{mid}\,(\boldsymbol{z}^{(k)}) - x^*\right] f'(\xi_2)}{\boldsymbol{f}'(\boldsymbol{x}^{(k)})}, \tag{20}$$

and

$$\operatorname{wid}\left(\boldsymbol{x}^{(k+1)}\right)=\left|\operatorname{mid}\left(\boldsymbol{u}_n^{(k)}\right)-x^*\right|\left|H'(\xi_1)\right|\left|\operatorname{mid}\left(\boldsymbol{z}^{(k)}\right)-x^*\right|\left|f'(\xi_2)\right|\operatorname{wid}\left(\frac{1}{\boldsymbol{f}'(\boldsymbol{x}^{(k)})}\right). \tag{21}$$

It is clear that

$$\left|\operatorname{mid}\left(\boldsymbol{z}^{(k)}\right)-x^*\right|\leq\operatorname{wid}\left(\boldsymbol{z}^{(k)}\right),\ \ \left|\operatorname{mid}\left(\boldsymbol{u}_n^{(k)}\right)-x^*\right|\leq\operatorname{wid}\left(\boldsymbol{u}_n^{(k)}\right). \tag{22}$$

Also, since $\boldsymbol{z}^{(k)}$ is generated by (8), from Theorem (2.4) there is positive constant $\widetilde{K}$ such that

$$\operatorname{wid}\left(\boldsymbol{z}^{(k)}\right)\leq\widetilde{K}\left(\operatorname{wid}\left(x^{(k)}\right)\right)^4, \tag{23}$$

and from Lemma (1.3) we have

$$\operatorname{wid}\left(\boldsymbol{u}_n^{(k)}\right)=\operatorname{wid}\left(\frac{\boldsymbol{f}(\boldsymbol{y}^{(k)})}{f(\operatorname{mid}\left(\boldsymbol{x}^{(k)}\right))}\right)\leq\operatorname{wid}\left(\boldsymbol{x}^{(k)}\right). \tag{24}$$

Now, let $|H'(\xi_1)|\leq K_1,\ |f'(\xi_2)|\leq K_2$ and using the relations (16), (22)–(24) in (21), we have the following error bound as

$$\operatorname{wid}\left(\boldsymbol{x}^{(k+1)}\right)\leq K\left(\operatorname{wid}\left(\boldsymbol{x}^{(k)}\right)\right)^6, \tag{25}$$

where $K=K_1K_2\widetilde{K}$. So the local order of convergence of the proposed interval method in this paper is 6 and the proof is completed. ■

3. Numerical implementation

In this section, we want to apply our method to solve three examples. Also we compare the computed results and justify the accuracy and applicability of the mentioned algorithm and theorem in the previous section. In fact, we illustrate three examples by applying the interval form of a class of modified Ostrowski methods (First algorithm). In these examples, we use the functions H as

$$H(t)=1+2t+\mu t^2+\gamma t^3, \tag{26}$$

$$H(t)=\frac{\gamma+(\beta+2\gamma)t}{\gamma+\beta t}, \tag{27}$$

$$H(t)=1+\frac{4t}{1+\sqrt{1-4t}}, \tag{28}$$

where $\mu,\gamma,\beta\in\mathbb{R}$ and compare the results with the results of the interval Ostrowski method, (Second algorithm) and the Newton interval algorithm (Third algorithm) proposed by Moore in (Moore et al., 2009). Numerical results are computed by using INTLAB toolbox created by Rump (Rump, 199).

3.1. Examples

Example 3.1. Suppose we want to solve $f(x)=0$ with $f(x)=x^3+4x^2-10$. The solution of this nonlinear equation is $x=1.36523001341409684$.

First algorithm (Interval form of a class of the modified Ostrowski methods)
We take $\boldsymbol{x}^{(0)} = [1, 2]$ and $\mu = \gamma = \beta = 1$.
Using (26), we obtain

$$\boldsymbol{x}^{(1)} = [\underline{1.36}40535118270273, \underline{1.36}57596756771437],$$
$$\boldsymbol{x}^{(2)} = [\underline{1.3652300134140966}, \underline{1.3652300134140968}].$$

Using (27), we obtain

$$\boldsymbol{x}^{(1)} = [\underline{1.36}40173134450457, \underline{1.36}56675343411908],$$
$$\boldsymbol{x}^{(2)} = [\underline{1.3652300134140966}, \underline{1.3652300134140971}].$$

Using (28), we obtain

$$\boldsymbol{x}^{(1)} = [\underline{1.36}40626978019838, \underline{1.36}57830581588510],$$
$$\boldsymbol{x}^{(2)} = [\underline{1.3652300134140966}, \underline{1.3652300134140971}].$$

Second algorithm (Interval Ostrowski method)
Taking $\boldsymbol{x}^{(0)} = [1, 2]$, we obtain

$$\boldsymbol{x}^{(1)} = [\underline{1.3}571456435997661, \underline{1.3}687534033070793],$$
$$\boldsymbol{x}^{(2)} = [\underline{1.3652}299985894846, \underline{1.3652}300283811582],$$
$$\boldsymbol{x}^{(3)} = [\underline{1.3652300134140967}, \underline{1.3652300134140969}].$$

Third algorithm (Interval Newton method)
Taking $\boldsymbol{x}^{(0)} = [1, 2]$, we obtain

$$\boldsymbol{x}^{(1)} = [\underline{1}.2840909090909089, \underline{1}.4151785714285716],$$
$$\boldsymbol{x}^{(2)} = [\underline{1.36}43820799441290, \underline{1.36}64268517584581],$$
$$\boldsymbol{x}^{(3)} = [\underline{1.3652}298533439375, \underline{1.3652}302030363476],$$
$$\boldsymbol{x}^{(4)} = [\underline{1.3652300134140944}, \underline{1.3652300134140995}],$$
$$\boldsymbol{x}^{(5)} = [\underline{1.3652300134140968}, \underline{1.3652300134140971}].$$

The results show that the proposed interval method in first algorithm is faster than the interval Ostrowski and Newton methods. For $\boldsymbol{f}(\boldsymbol{x}^{(2)})$ of three methods, we have:
for the first method from (26)

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = 1.0e - 014 \times [-0.35527136788006, 0.17763568394003],$$

and from (27) and (28)

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = 1.0e-014 \times [-0.35527136788006, 0.53290705182008],$$

for the second method

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = 1.0e-006 \times [-0.24480473648226, 0.24715706103962],$$

and for the third method

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = [-0.01399644365940, 0.01977546736885].$$

Example 3.2. Suppose we want to solve $f(x) = 0$ with $f(x) = -\sin^2(x) + 3\cos(x) + x\exp(x^2) + 5$. The solution of this nonlinear equation is
$x = -1.20764782713091892$.

First algorithm (Interval form of a class of the modified Ostrowski methods)
We take $\boldsymbol{x}^{(0)} = [-1.5, -1]$ and $\mu = \gamma = \beta = 1$.
Using (26), we obtain

$$\boldsymbol{x}^{(1)} = [\underline{-1.20}64954394368772, \underline{-1.20}82267483787090],$$
$$\boldsymbol{x}^{(2)} = [\underline{-1.2076478271309}178, \underline{-1.2076478271309}201].$$

Using (27), we obtain

$$\boldsymbol{x}^{(1)} = [\underline{-1.20}64583312305396, \underline{-1.20}80334360657845],$$
$$\boldsymbol{x}^{(2)} = [\underline{-1.2076478271309}172, \underline{-1.2076478271309}205].$$

Using (28), we obtain

$$\boldsymbol{x}^{(1)} = [\underline{-1.20}65037762572525, \underline{-1.20}82701783956324],$$
$$\boldsymbol{x}^{(2)} = [\underline{-1.2076478271309}178, \underline{-1.2076478271309}198].$$

Second algorithm (Interval Ostrowski method)
Taking $\boldsymbol{x}^{(0)} = [-1.5, -1]$, we obtain

$$\boldsymbol{x}^{(1)} = [\underline{-1.2}106610890531674, \underline{-1.2}015072023630468],$$
$$\boldsymbol{x}^{(2)} = [\underline{-1.2076478}827088639, \underline{-1.2076477}721947516],$$
$$\boldsymbol{x}^{(3)} = [\underline{-1.2076478271309}190, \underline{-1.2076478271309}187].$$

Third algorithm (Interval Newton method)
Taking $\boldsymbol{x}^{(0)} = [-1.5, -1]$, we obtain

$$\boldsymbol{x}^{(1)} = [\underline{-1}.2336804731371404, \underline{-1}.1649846948946225],$$
$$\boldsymbol{x}^{(2)} = [\underline{-1.20}87043535177426, \underline{-1.20}6878139234288 8],$$
$$\boldsymbol{x}^{(3)} = [\underline{-1.2076}482734756950, \underline{-1.2076}47441617257 1],$$
$$\boldsymbol{x}^{(4)} = [\underline{-1.207647827130}9605, \underline{-1.207647827130}8799],$$
$$\boldsymbol{x}^{(5)} = [\underline{-1.2076478271309}190, \underline{-1.2076478271309}187].$$

The results of this example also show that the proposed interval method in first algorithm is faster than the interval Ostrowski method and interval Newton method. For $\boldsymbol{f}(\boldsymbol{x}^{(2)})$ of three methods, we have:
for the first method from (26),

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = 1.0e-013 \times [-0.28421709430405, 0.25757174171304],$$

from (27),

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = 1.0e-013 \times [-0.37303493627406, 0.38191672047106],$$

from (28),

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = 1.0e-013 \times [-0.23092638912204, 0.25757174171304],$$

for the second method

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = 1.0e-005 \times [-0.11286442926917, 0.11156112638134],$$

and for the third method

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = [-0.02148941450246, 0.01561231591470].$$

Example 3.3. Suppose we want to solve $f(x) = 0$ with $f(x) = \exp(-x) + \cos(x)$. The solution of this nonlinear equation is $x = 1.746139530408012285$.

First algorithm (Interval form of a class of the modified Ostrowski methods)
We take $\boldsymbol{x}^{(0)} = [1, 2]$ and $\mu = \gamma = \beta = 1$.
Using (25), we obtain

$$\boldsymbol{x}^{(1)} = [\underline{1.746}0718336173990, \underline{1.746}3268259266115],$$
$$\boldsymbol{x}^{(2)} = [\underline{1.7461395304080122}, \underline{1.7461395304080125}].$$

Using (26), we obtain

$$\boldsymbol{x}^{(1)} = [\underline{1.746}0778555890572, \underline{1.746}3311262322414],$$

$$\boldsymbol{x}^{(2)} = [\underline{1.7461395304080122}, \underline{1.7461395304080125}].$$

Using (27), we obtain

$$\boldsymbol{x}^{(1)} = [\underline{1.746}0700894634930, \underline{1.746}3255804217692],$$
$$\boldsymbol{x}^{(2)} = [\underline{1.7461395304080122}, \underline{1.7461395304080125}].$$

Second algorithm (Interval Ostrowski method)
Taking $\boldsymbol{x}^{(0)} = [1, 2]$, we obtain

$$\boldsymbol{x}^{(1)} = [\underline{1.74}51497239632181, \underline{1.74}87777462087110],$$
$$\boldsymbol{x}^{(2)} = [\underline{1.7461395303}511697, \underline{1.7461395304}645352],$$
$$\boldsymbol{x}^{(3)} = [\underline{1.7461395304080122}, \underline{1.7461395304080125}].$$

Third algorithm (Interval Newton method)
Taking $\boldsymbol{x}^{(0)} = [1, 2]$, we obtain

$$\boldsymbol{x}^{(1)} = [\underline{1}.7148342558350511, \underline{1}.8008450820083625],$$
$$\boldsymbol{x}^{(2)} = [\underline{1.74}59521461460445, \underline{1.74}62654180910368],$$
$$\boldsymbol{x}^{(3)} = [\underline{1.7461395}288164853, \underline{1.7461395}317156374],$$
$$\boldsymbol{x}^{(4)} = [\underline{1.7461395304080122}, \underline{1.7461395304080125}].$$

The results of this example also show that the proposed interval method in first algorithm is faster than all the others. For $\boldsymbol{f}(\boldsymbol{x}^{(2)})$ of three methods, we have:
for the first method from (26), (27) and (28)

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = 1.0e-015 \times [-0.19428902930941, 0.24980018054067],$$

for the second method

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = 1.0e-010 \times [-0.65516481129180, 0.65887156841527],$$

and for the third method

$$\boldsymbol{f}(\boldsymbol{x}^{(2)}) = 1.0e-003 \times [-0.14591526185559, 0.21720562386582].$$

4. Conclusion

In this paper, a new enclosure method was introduced to find the interval solution of a given nonlinear equation. This method has the local order of convergence equal to 6. Moreover, necessary and sufficient conditions about the convergency were discussed in details. Also, error bound and convergence rate were studied. To verify the theory, this algorithm was then tested using some examples via INTLAB. Furthermore, the suggested method was compared with interval Ostrowski method and interval Newton method. As expected, according to the discussed theory, this method was better than both of the other methods.

5. Acknowledgement

The authors has been supported by Islamic Azad University of Hamedan.

References

W. Bi, H. Ren, Qingbiao Wu, Three-step iterative methods with eighth-order convergence for solving nonlinear equations, J. Comput. Appl. Math. 225 (2009) 105–112.

W. Bi, Qingbiao Wu, H. Ren, A new family of eighth-order iterative methods for solving nonlinear equations, Appl. Math. Comput. 214 (2009) 236–245.

C. Chun, Y. Ham, Some sixth-order variants of Ostrowski root-finding methods, Appl. Math. Comput. 193 (2007) 389–394.

M. Grau, J. L. Diaz-Barrero, An improvement to Ostrowski root-finding method, Appl. Math. Comput., vol. 173, PP. 450–456, 2006.

Y. Ham, C. Chun, Sang-Gu Lee, Some higher-order modifications of Newton's method for solving nonlinear equations, J. Comput. Appl. Math. 222 (2008) 477–486.

J. Kou, Some new sixth-order methods for solving non-linear equations, Appl. Math. Comput. 189 (2007) 647–651.

J. Kou, X. Wang, Some improvements of Ostrowski's method, Applied Mathematics Letters 23 (2010) 92–96.

L. Liu, X. Wang, Eighth-order methods with high efficiency index for solving nonlinear equations, Appl. Math. Comput., vol. 215, pp. 3449–3454, 2010.

R.E. Moore, Interval arithmetic and automatic error analysis in digital computing, Ph.D. dissertation, Department of Mathematics, Standford University, Standford, CA, 1962.

R.E. Moore, Interval Analysis, Prentice-Hall, Englewood Cliff, NJ, 1966.

R. E. Moore, R. B. Kearfott, M. J. Cloud, Introduction to Interval Analysis, SIAM, 2009.

A. Neumaier, Interval Methods for Systems of Equations, Cambridge University Press, 1990.

A. Neumaier, Introduction to Numerical Analysis, Cambridge University Press, 2001.

A.M. Ostrowski, Solution of Equations in Euclidean and Banach Spaces, third ed, Academic Press, New York, 1973.

S. Rump, INTLAB-INTerval LABoratory, in: T.Csendes (Ed.), Developments in Reliable Computing, Kluwer Academic Publishers, Dordrecht, 1999.

J.R. Sharma, R. Sharma, A new family of modified Ostrowski's methods with accelerated eighth order convergence, Numerical Algorithms 54 (2010) 445–458.

J.R. Sharma, R.K. Guha, A family of modified Ostrowski methods with accelerated sixth order convergence, Appl. Math. Comput. 190 (2007) 111–115.

R. Thukral, Introduction to a Newton-type method for solving nonlinear equations, Appl. Math. Comput. 195 (2008) 663–668.

Sensitivity Measures for Minimizing Model Uncertainty in Probabilistic Analysis

M. Mahsuli[1] and T. Haukaas[2]
[1] *PhD Candidate, Department of Civil Engineering, University of British Columbia, Vancouver, Canada, mahsuli@gmail.com*
[2] *Associate Professor, Department of Civil Engineering, University of British Columbia, Vancouver, Canada, terje@civil.ubc.ca*

Abstract: This paper presents a new set of reliability sensitivity measures. The purpose with these measures is to identify the optimal manner in which to reduce model uncertainty in order to improve risk estimates. In particular, two sensitivity measures are presented. One identifies the buildings, or other components within a region, that should be subjected to more refined modeling. The other sensitivity measure identifies model types that should be subjected to further research to improve the model form. The developments in this paper are presented in the context of a region with 622 buildings that are subjected to seismic hazard. A comprehensive seismic risk analysis is conducted with approximately 300 random variables, more than 30 different model types, and more than 3,000 individual model instances. All models are probabilistic and emphasis is placed on explicit characterization of epistemic uncertainty, *i.e.*, reducible uncertainty. The models are available in a new computer program called Rt, which is tailored for reliability analysis with multiple probabilistic models. The primary result from the analysis is risk estimates, presented in the form of loss probability curves. However, focus in this paper is on the development and evaluation of sensitivity measures, in order to guide efforts to reduce the model uncertainty and thus improve the risk estimates. For the considered region it is found that concrete shear wall buildings, and structural response models for such buildings, rank highest according to both sensitivity measures. As described in this paper, this means that allocating resources for detailed analysis and improved models for this type of building has the greatest impact on the risk estimates.

Keywords: Probabilistic models, model uncertainty, sensitivity analysis, risk, reliability

1. Introduction

The primary objective in this paper is to identify the optimal course of action to reduce model uncertainty. Context is provided by seismic risk analysis, where multiple probabilistic models are employed for hazard, infrastructure, and impacts. In this paper, reliability methods are utilized in conjunction with a library of probabilistic models to make predictions about potential future seismic losses. The analysis is carried out with a new reliability-based risk analysis program, called Rt, which is specifically developed for multi-model reliability and optimization analysis. Rt is freely available online at www.inrisk.ubc.ca. The library of probabilistic models is implemented in Rt, and certain models are devoted particular attention in this paper. However, although the study is focused on seismic risk, the methods and models are generic. In fact, the developments in this paper are intended as universal techniques for the recognition and subsequent reduction of epistemic uncertainty, *i.e.*, reducible uncertainty. To this end, two questions are asked and

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

addressed in this study: 1) Which infrastructure components should be subjected to more detailed modeling to reduce the epistemic uncertainty, in order to improve the quality of the present risk analysis? 2) Which models in the library of probabilistic models should be prioritized in research efforts to reduce the epistemic uncertainty to improve the quality of future risk analyses? The vehicle for the developments is reliability analysis with the limit-state function

$$g = g(\boldsymbol{\theta}, \mathbf{x}, \mathbf{v}) = l_0 - l(\boldsymbol{\theta}, \mathbf{x}, \mathbf{v}) \tag{1}$$

where l_t = loss threshold, $l(\boldsymbol{\theta},\mathbf{x},\mathbf{v})$ = loss due to earthquake damage, $\boldsymbol{\theta}$ = vector of epistemic random variables, $\mathbf{x}$ = vector of aleatory random variables, and $\mathbf{v}$ = vector of decision variables that are at the discretion of the decision maker. The focus in this paper is on the uncertainty described by $\boldsymbol{\theta}$, while the $\mathbf{v}$ is omitted in the following. Reliability methods, such as the first-order and second-order reliability methods and importance sampling, estimates the probability that $g < 0$. As a result, reliability analysis with Eq. (1) yields the probability that the cost l exceeds l_o. In other words, the result is a point on the loss exceedance probability curve, hereafter called loss curve. Loss curve results are presented later in this paper and they appear prominently in several areas of seismic risk analysis. They are particularly popular in the insurance industry and in modern performance-based earthquake engineering.

It is emphasized in this paper that many interacting probabilistic models are required to evaluate $l(\boldsymbol{\theta},\mathbf{x},\mathbf{v})$ in Eq. (1). In fact, a significant effort is made to develop or improve models for all facets of the hazards, infrastructure, and impacts associated with seismic risk. In turn, the models are implemented in Rt to facilitate the communication between the models at run-time. The new object-oriented software architecture to accomplish this is described by Mahsuli and Haukaas (2012). From a broader perspective, Rt is intended as a continuously growing framework of predictive probabilistic models, with explicit characterization of epistemic uncertainty. This is intended to promote targeted future efforts to reduce that uncertainty. In fact, the framework provides a rational basis for allocating resources to gather data and build better models, which ultimately yields improved risk mitigation decisions. This motivates the developments in this paper.

2. Models

The approach adopted in this paper has two components: probabilistic models and reliability methods. In contrast with many contemporary seismic risk analysis approaches, the present approach circumvents conditional probability models in favour of simulation-type models that produce scalars or vectors of physical responses. This is necessary in order to evaluate Eq. (1). In particular, the models that are utilized in this study employ random variables to discretize the uncertainty. A simple but instructive model is the linear regression model

$$y = \theta_1 + \theta_2 \cdot h_2(\mathbf{x}) + \theta_3 \cdot h_3(\mathbf{x}) + \cdots + \varepsilon \tag{2}$$

where y = model response, θ_i = model parameters, $h_i(\mathbf{x})$ = explanatory functions, and ε = zero-mean normally distributed model error. In the Bayesian approach to linear regression for this model, the parameters θ_i and the standard deviation of ε are random variables. This approach is adopted here, where the model parameters are categorized as epistemic random variables, *i.e.*, $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \sigma_\varepsilon\}$. Furthermore, their probability distribution is affected by model improvement, typically by data gathering efforts. The

statistical inference for the random variables in **θ** is carried out in accordance with Box and Tiao (1992), Gardoni *et al.* (2002), and others. In fact, the Bayesian modeling philosophy is adopted throughout this study, although the model forms may vary.

In Rt's framework of models, each input variable in **x** to Eq. (2) is either provided as a random variable by the analyst or as a response from an "upstream" model. In turn, the response, *y*, may serve as input to a "downstream" model. For example, one variable in **x** may be an earthquake magnitude predicted by another model, while *y* may be a site-specific ground shaking intensity that serves as input to a building response model.

The specific set of models considered in this paper simulate the occurrence of hazards, building responses, damage, and cost for 622 buildings on the campus of the University of British Columbia (UBC) in Vancouver, Canada. Figure 1 displays a map of the region with the UBC campus identified in reference to downtown Vancouver. The dots in the zoomed map of the UBC campus identify the 622 building on campus. The second author's research group surveyed each building to gather data about building type, building height, footprint area, *etc*.

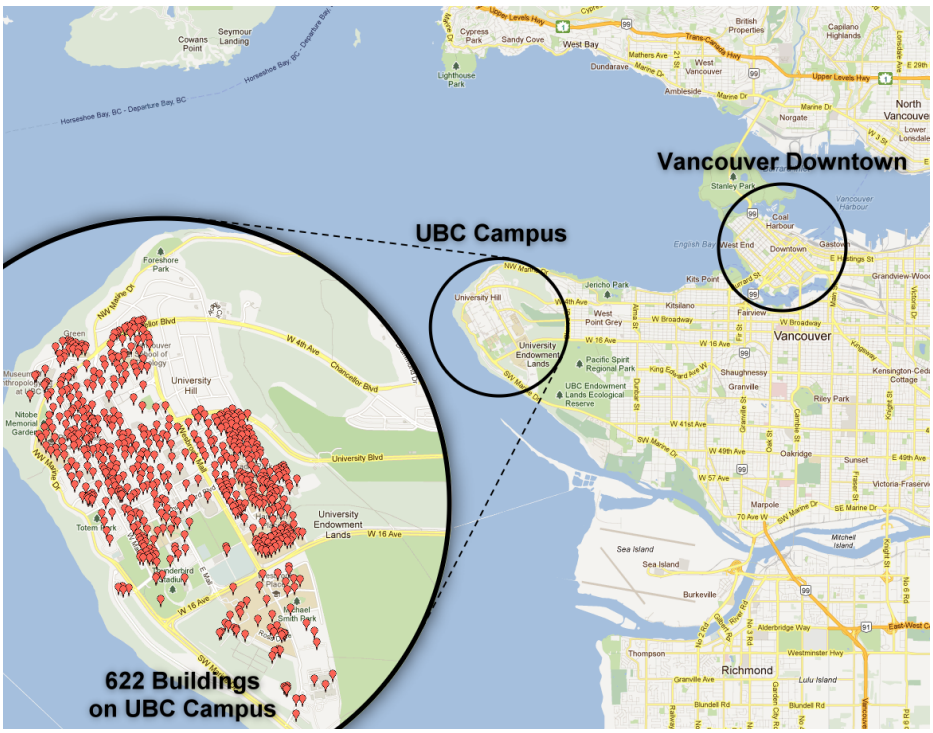


Figure 1. Map of the UBC campus and the 622 buildings that are modeled in this study.

Table I displays some of the information that was gathered for each of the 622 buildings at the UBC campus. For brevity, only some of the buildings are presented. These particular buildings were selected for this table because they appear prominently in the rankings that are presented later in this paper.

Table I: Information for a few selected buildings among the 622 buildings on the UBC campus.

Building Name	Footprint Area (m^2)	Number of Stories	Year Erected	Mean Total Value ($)	Code Level	Longitude	Latitude
Animal Science Main Sheep Unit	552	1	1976	671,109	Moderate	-123.2341	49.2509
Chan Centre for Performing Arts	3,315	1	1997	7,107,360	Moderate	-123.2551	49.2698
Morris & Helen Belkin Art Gallery	1,105	2	1995	4,738,240	Moderate	-123.2562	49.2682
Power House Meter Station	90	1	1960	206,730	Low	-123.2545	49.265
St. Mark Chapel	524	2	1997	2,406,909	Moderate	-123.249	49.2722
University Centre Addition	230	1	1987	528,310	Moderate	-123.2568	49.2691
Vanier Pump Station	16	1	1986	37,277	Moderate	-123.2603	49.2648
Village Shops 1	922	2	1980	4,234,872	Moderate	-123.2428	49.2666
Village Shops 2	1,171	1	1980	2,689,424	Moderate	-123.2434	49.2664
Wesbrook Animal Care Unit	596	1	1981	1,369,012	Moderate	-123.2489	49.2652
…	…	…	…	…	…	…	…

The UBC campus is subjected to three sources of seismicity: Shallow crustal earthquakes, deep subcrustal earthquakes, and megathrust subduction earthquakes. The first two occur within area sources, while subduction earthquakes originate from a faultline that runs under the ocean outside the coastline of the Pacific Northwest.

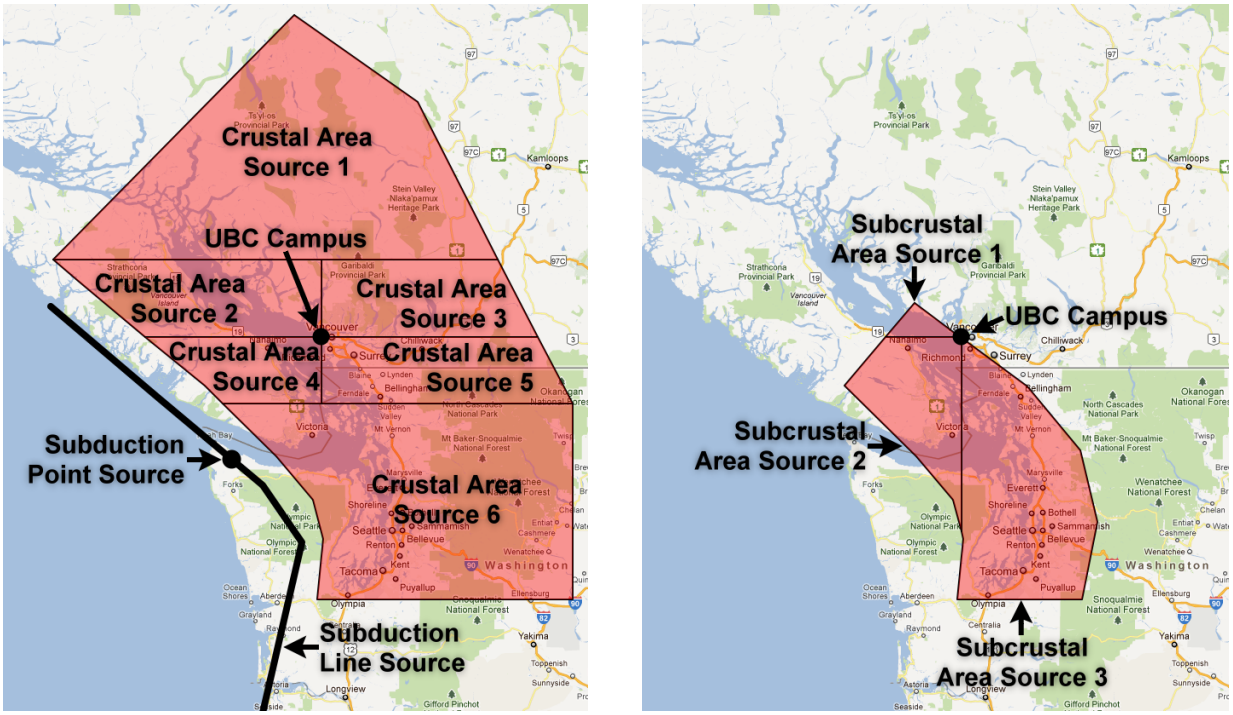


Figure 2. Sources of earthquakes affecting the UBC campus.

Table II: Overview of models employed in the analysis.

Model Name	Formulation	Instances	Size($\mathbf{x}$)	K = Size($\boldsymbol{\theta}$)
Concrete Frame with Masonry Infill Wall Structural Damage	Nonlinear Regression	10	7	9
Concrete Frame with Masonry Infill Wall Structural Response	Linear Regression	10	4	25
Concrete Moment Frame Structural Damage	Nonlinear Regression	22	8	10
Concrete Moment Frame Structural Response	Linear Regression	22	4	25
Concrete Shear Wall Structural Damage	Nonlinear Regression	134	7	9
Concrete Shear Wall Structural Response	Linear Regression	134	4	25
Crustal Intensity	Algorithm	1	4	1
Non-Structural Acceleration Damage	Nonlinear Regression	622	2	3
Non-Structural Drift Damage	Nonlinear Regression	622	2	3
Precast Concrete Structural Damage	Nonlinear Regression	11	8	10
Precast Concrete Structural Response	Linear Regression	11	4	25
Reinforced Masonry Structural Damage	Nonlinear Regression	58	8	10
Reinforced Masonry Structural Response	Linear Regression	58	4	25
Steel Braced Frame Structural Damage	Nonlinear Regression	5	7	9
Steel Braced Frame Structural Response	Linear Regression	5	4	25
Steel Frame with Concrete Shear Wall Structural Damage	Nonlinear Regression	6	6	8
Steel Frame with Concrete Shear Wall Structural Response	Linear Regression	6	4	25
Steel Frame with Masonry Infill Wall Structural Damage	Nonlinear Regression	2	6	8
Steel Frame with Masonry Infill Wall Structural Response	Linear Regression	2	4	25
Steel Light Frame Structural Damage	Nonlinear Regression	22	6	8
Steel Light Frame Structural Response	Linear Regression	22	4	25
Steel Moment Frame Structural Damage	Nonlinear Regression	4	7	9
Steel Moment Frame Structural Response	Linear Regression	4	4	25
Subcrustal Intensity	Algorithm	1	4	1
Subduction Intensity	Algorithm	1	4	1
Unreinforced Masonry Structural Damage	Nonlinear Regression	14	6	8
Unreinforced Masonry Structural Response	Linear Regression	14	4	25
Wood Large Frame Structural Damage	Nonlinear Regression	128	5	7
Wood Large Frame Structural Response	Linear Regression	128	4	25
Wood Light Frame Structural Damage	Nonlinear Regression	206	5	7
Wood Light Frame Structural Response	Linear Regression	206	4	25

It is noted that Figure 2 divides the area sources into several sub-areas. Specifically, the crustal earthquake source is divided into six area sources, while the subcrustal area source is divided into three area sources. This is done for practical reasons that relate to the reliability analysis. In particular, the first-order reliability

method (FORM) is employed, and this type of analysis requires a continuously differentiable limit-state function that is relatively linear in the space of the random variables. This is achieved by the subdivision of the area sources in Figure 2. Furthermore, the subduction source is divided into a point source and a line source. This is done for physical reasons. Specifically, a certain range of magnitudes of subduction earthquakes is associated with rupture of the entire fault line. The location of such earthquakes is therefore known and thus modeled by a point source. In contrast, subduction earthquakes of lower magnitudes are associated with partial rupture of the fault. The unknown location of this type of earthquakes is modeled by the line source shown in Figure 2.

In addition to the models required to simulate the earthquake hazard, an array of other models were utilized in this analysis. Table II provides an overview of these models. It is important to note that each model conforms to the following format: It takes random variables and other parameters as input, and it produces a physical measurable scalar or vector as output. For example, each of the earthquake location models described above takes the realization of a few random variables as input and produces the corresponding hypocenter location as output. Table II shows the number of instances of each model in the analysis. It also shows the number of random variables that each model takes. Specifically, the last two columns in Table II displays the number of aleatory random variables, size($\mathbf{x}$), and the number of epistemic random variables, size($\boldsymbol{\theta}$), respectively, in each model. It is emphasized that the models for building response and building damage, *i.e.*, the models in Table II that contain the word "Damage" or "Response" are simplified models rather than detailed finite element models. This is discussed later in the context of refining the building response and damage models.

3. Analysis

Given the sub-division of area sources in Figure 2 it is understood that there are 11 sources of earthquakes in the reliability analysis (Crustal + Subcrustal + Subduction = 6 + 3 + 2 = 11). As a result, a multi-hazard analysis is necessary. Several multi-hazard analysis options are available; one is the load coincidence method proposed by Wen (1990). However, matters simplify because the probability of coincidence of two earthquakes is negligible in this particular application. To address the presence of multiple hazards, let $i = \{1, 2, \ldots, N\}$, where $N = 11$ = number of hazards, and let β_i denote the reliability index associated with the limit-state function in Eq. (1) for each hazard. It is emphasized that each hazard is analyzed separately. From the theory of FORM reliability analysis it is know that the associated probability, *i.e.*, the point on the loss curve is

$$p_i = \Phi(-\beta_i) \tag{3}$$

where Φ is the standard normal cumulative distribution function. Provided the Poisson process is valid for each hazard, with rates λ_i, the rate of loss exceedance associated with each hazard is $\lambda_i p_i$. The combined rate including all hazards is the sum of the individual rates, and the Poisson distribution provides the probability of loss exceedance within a time period, T:

$$p = 1 - p(0) = 1 - \exp\left(-T \cdot \sum_{i=1}^{N} \lambda_i p_i\right) \tag{4}$$

In the context of FORM analysis is common to employ the reliability index instead of the probability. For this purpose, the generalized reliability index associated with p is obtained by inversion of the standard normal cumulative distribution function:

$$\beta = -\Phi^{-1}(p) \tag{5}$$

where β is employed in the following as a surrogate measure for the exceedance probability when all hazards are considered.

4. Loss Curve Results

As mentioned earlier, an important objective in this study is to compute loss curves. To illustrate the concept, Figure 3 shows two loss curves obtained by Monte Carlo sampling with 100,000 samples. The black solid line displays the loss curve that is obtained when all random variables are included. To highlight the significance of epistemic uncertainty, the grey line in Figure 3 shows the loss curve that is obtained if all the epistemic random variables, *i.e.*, $\boldsymbol{\theta}$, are set equal to their mean values. Naturally, this results in an underestimation of the probability of high losses, *i.e.*, a "slimmer" tail of the loss curve. In fact, particular focus in this study is on the tail of the loss curve because of its importance in risk mitigation decisions. Unfortunately, although Monte Carlo sampling is a robust analysis approach, it yields less accurate results in the tail than around the mean of the loss. In contrast, FORM has two advantages that are explored in this study. First, it is tailored to estimate small probabilities, *i.e.*, it addresses the tail of the loss curve. Second, it facilitates the computation of sensitivity measures that are employed in the following.

The loss curves in Figure 3 are plotted from zero to $100 million, and the figure reveals that there is roughly a 5% chance that this loss threshold will be exceeded. A better estimate is obtained by running FORM analysis with the following limit-state function:

$$g = 100,000,000 - l(\boldsymbol{\theta}, \mathbf{x}) \tag{6}$$

FORM analysis for the individual hazards yield the reliability indices shown in Table III. The table shows that subduction earthquakes are associated with the lowest reliability indices. This implies that these earthquake sources produce the highest loss exceedance probabilities. However, it is also observed in Table III that subduction earthquakes are associated with low occurrence rates. This means that their overall influence on the seismic risk must be investigated further, which is done in the following. The results in Table III are substituted into Eq. (4) to compute the probability of exceeding a $100 million loss considering all earthquake sources. This yields $p = 0.076$, *i.e.*, a 7.6% chance of exceeding that loss.

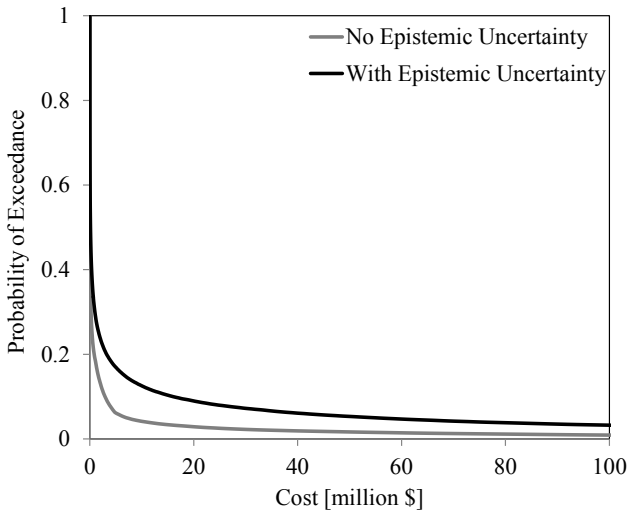


Figure 3. Loss curve with and without epistemic uncertainty.

TableIII: Occurrence rates and reliability indices for individual hazards.

Source	Occurrence Rate	Reliability Index
Crustal Area Source 1	0.062	3.5
Crustal Area Source 2	0.019	2.6
Crustal Area Source 3	0.017	2.5
Crustal Area Source 4	0.010	2.4
Crustal Area Source 5	0.017	2.6
Crustal Area Source 6	0.063	3.5
Subcrustal Area Source 1	0.0029	1.6
Subcrustal Area Source 2	0.027	2.0
Subcrustal Area Source 3	0.063	2.7
Subduction Line Source	0.0010	1.3
Subduction Point Source	0.0013	1.3

5. Sensitivity with respect to Model Refinement Decisions

In the context of the regional seismic risk analysis of the UBC campus, suppose it is contemplated to refine some of the building models to reduce the epistemic uncertainty. In particular, the analyst may seek to replace simple building response models with detailed finite element models. Clearly, only the most important buildings can be addressed due to the time it takes to establish a detailed finite element model and the added computational cost. This section provides guidance for the analyst to prioritize between buildings.

First, it is recognized that the objective is to reduce epistemic uncertainty. Provided that the epistemic uncertainty has been properly included in the models, a more detailed model will produce results with less uncertainty. Second, it is understood that it is the effect on the overall loss curve that must guide the decision to replace a model. In other words, the model whose $\boldsymbol{\theta}$ has the largest influence on p in Eq. (4) should be replaced with a better model. In particular, for a model with only one epistemic random variable, it is the sensitivity $\partial p/\partial \sigma$, where σ is the standard deviation of that epistemic random variable that should guide the prioritization. In general, each model has several epistemic random variables. To this end, a sensitivity measure that represents the derivative of p with respect to the standard deviation of all the epistemic random variables in a model is sought.

Suppose a model, such as the one in Eq. (2), is generically written as $y = y(\boldsymbol{\theta},\mathbf{x})$, where $\boldsymbol{\theta}$ and $\mathbf{x}$ remain the vectors of epistemic and aleatory random variables, respectively. Furthermore, let K denote the number of epistemic random variables in the model. Next, consider the well-known first-order approximation of the variance of the response from this model with respect to the epistemic random variables:

$$\sigma^2 = \nabla_{\boldsymbol{\theta}} y^T \cdot \boldsymbol{\Sigma}_{\boldsymbol{\theta}\boldsymbol{\theta}} \cdot \nabla_{\boldsymbol{\theta}} y = \sum_{i=1}^{K}\sum_{j=1}^{K}\left(\frac{\partial y}{\partial \theta_i} \cdot \frac{\partial y}{\partial \theta_j} \cdot \rho_{ij} \cdot \sigma_i \cdot \sigma_j \right) \tag{7}$$

where $\nabla_{\boldsymbol{\theta}} y$ = gradient vector of y with respect to $\boldsymbol{\theta}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}\boldsymbol{\theta}}$ = covariance matrix of $\boldsymbol{\theta}$, ρ_{ij} = correlation coefficient between the components of $\boldsymbol{\theta}$, and σ_i = standard deviation of the components of $\boldsymbol{\theta}$. For models that are linear with respect to the epistemic random variables, Eq. (7) provides exact results; otherwise, it is an approximation. In order to study the influence of epistemic uncertainty on β, the following derivative is sought and evaluated by the chain rule of differentiation:

$$\frac{\partial \beta}{\partial \sigma} = \sum_{i=1}^{N}\left(\frac{\partial \beta}{\partial p} \cdot \frac{\partial p}{\partial p_i} \cdot \frac{\partial p_i}{\partial \beta_i} \cdot \frac{\partial \beta_i}{\partial \sigma} \right) \tag{8}$$

where N = number of hazards. The derivatives in the right hand side of Eq. (8) are addressed separately in the following. The first derivative is obtained by differentiating Eq. (5):

$$\frac{\partial \beta}{\partial p} = -\frac{1}{\varphi(\beta)} \tag{9}$$

The second derivative is obtained by differentiating Eq. (4):

$$\frac{\partial p}{\partial p_i} = T \cdot \lambda_i \cdot e^{-T\cdot\sum_{i=1}^{N}\lambda_i p_i} \tag{10}$$

The third derivative is obtained by differentiating Eq. (3):

$$\frac{\partial p_i}{\partial \beta_i} = -\varphi(\beta_i) \tag{11}$$

The fourth derivative is obtained by adding contributions from all the epistemic random variables in the model. The chain rule of differentiation yields:

$$\frac{\partial \beta_i}{\partial \sigma} = \sum_{j=1}^{K}\left(\frac{\partial \beta_i}{\partial \sigma_j} \cdot \frac{\partial \sigma_j}{\partial \sigma} \right) \tag{12}$$

where the first derivative in the right-hand side is a well-known reliability sensitivity measure, see for example Der Kiureghian (2005) for details, while the last derivative in the right-hand side of Eq. (12) is obtained by differentiating Eq. (7):

$$\frac{\partial \sigma}{\partial \sigma_j} = \frac{1}{\sigma} \cdot \frac{\partial y}{\partial \theta_j} \cdot \sum_{i=1}^{K} \left(\frac{\partial y}{\partial \theta_i} \cdot \rho_{ij} \cdot \sigma_i \right) \tag{13}$$

In the following, Eq. (8) is evaluated and compared for the 622 buildings at the UBC campus. Table IV displays values for $\partial\beta/\partial\sigma$ for the 10 highest ranked buildings. In other words, Table IV identifies the models for which a reduction in the epistemic uncertainty would have the greatest impact on the reliability index. Naturally, an increase in epistemic uncertainty, *i.e.*, an increase in σ, increases the probability of exceeding a \$100 million loss, which in turn reduces the reliability index; hence the minus sign in Table IV. It is observed in Table IV that the highest ranked buildings are mostly concrete shear wall buildings, which may indicate that the structural model for this type of building has the greatest potential for improvement. This point is brought up later in this paper.

Table V shows the ranking of magnitude models. It reveals that the magnitude model for subcrustal area source 2 is the model for which a reduction in the epistemic uncertainty would have the greatest impact on the loss probability. Similarly, Table VI shows the ranking of ground motion intensity models for an arbitrarily selected building. This ranking of the intensity models were observed for 430 of the buildings, while the intensity models for subduction and subcrustal earthquakes switch places for the other 192 buildings. It is reemphasized that these results provide a basis for selecting models to be refined if proper resources are available.

It is of interest to investigate the value of $\partial\beta/\partial\sigma$ if the entire collection of models is considered as one model. For this case, the evaluation of $\partial\beta/\partial\sigma$ is disaggregated into two parts:

$$\frac{\partial \beta}{\partial \sigma} = \sum_{i \in H} \frac{\partial \beta}{\partial \sigma_i} \cdot \frac{\partial \sigma_i}{\partial \sigma} + \sum_{i \in B} \frac{\partial \beta}{\partial \sigma_i} \cdot \frac{\partial \sigma_i}{\partial \sigma} \tag{14}$$

where the first sum is taken over epistemic random variables associated with hazard models, while the second sum is taken over epistemic random variables associated with building models. The analysis reveals that the first sum equals $-8.54 \cdot 10^{-9}$ and the second equals $-7.20 \cdot 10^{-7}$. This shows that, in the context of the epistemic uncertainty that is modeled in this study, it is far more effective to reduce the epistemic uncertainty in the building models rather than the hazard models.

6. Sensitivity with respect to Model Improvement Decisions

In the long run, researchers seek to improve the library of models that are available. This effort to reduce epistemic uncertainty in generic models addresses a different problem than that addressed in the previous section. In particular, the objective in the previous section was to identify, *e.g.*, the building that should be subjected to more detailed modeling. In contrast, this section identifies which generic models should be prioritized for further research and data gathering. To make such decisions, it is necessary to assess the cost of long-term model improvement, and how those efforts will improve the assessment of risk. To this end,

TableIV: Top 10 building models according epistemic uncertainty.

Building Name	Building Type	$\partial\beta/\partial\sigma$
University Centre Addition	Concrete Shear Wall	-0.76
Village Shops 2	Concrete Shear Wall	-0.57
Power House Meter Station	Concrete Shear Wall	-0.48
Vanier Pump Station	Steel Light Frame	-0.40
Wesbrook Animal Care Unit	Concrete Shear Wall	-0.27
St. Mark Chapel	Concrete Shear Wall	-0.11
Chan Centre for Performing Arts	Concrete Shear Wall	-0.11
Animal Science Main Sheep Unit	Concrete Moment Frame	-0.11
Morris & Helen Belkin Art Gallery	Concrete Shear Wall	-0.023
Village Shops 1	Concrete Shear Wall	-0.022

Table V: Ranking of magnitude models according to epistemic uncertainty.

Model	$\partial\beta/\partial\sigma$
Subcrustal Area Source 2 Magnitude	-0.097
Subcrustal Area Source 3 Magnitude	-0.062
Subcrustal Area Source 1 Magnitude	-0.023
Crustal Area Source 2 Magnitude	-0.019
Crustal Area Source 3 Magnitude	-0.019
Crustal Area Source 5 Magnitude	-0.018
Crustal Area Source 4 Magnitude	-0.017
Crustal Area Source 6 Magnitude	-0.0039
Crustal Area Source 1 Magnitude	-0.0034

Table VI: Ranking of intensity models according to epistemic uncertainty.

Model	$\partial\beta/\partial\sigma$
Crustal Intensity	-2.81
Subcrustal Intensity	-1.42
Subduction Intensity	-0.80

the sensitivity of the reliability index, β, with respect to the cost of modeling, c, *i.e.*, $\partial\beta/\partial c$, is sought. Chain rule of differentiation yields:

$$\frac{\partial \beta}{\partial c} = \sum_{i=1}^{N} \left(\frac{\partial \beta}{\partial \beta_i} \cdot \frac{\partial \beta_i}{\partial c} \right) \tag{15}$$

where the first derivative is expanded and explained earlier, and the second derivative is obtained by the chain rule, adding contributions over the K epistemic random variables of the model:

$$\frac{\partial \beta_i}{\partial c} = \sum_{j=1}^{K} \left(\frac{\partial \beta_i}{\partial \sigma_j} \cdot \frac{\partial \sigma_j}{\partial n} \cdot \frac{\partial n}{\partial c} \right) \tag{16}$$

where n is introduced to quantify the number of observations that are used to develop the model. The first derivative in the right-hand side of Eq. (16) is addressed in the previous section. The second derivative represents the change in the standard deviation of a model parameter due to a change in the number of observations that are employed to build the model. In the following, this derivative, $\partial\sigma_j/\partial n$, is expressed for three types of models: Linear regression models, nonlinear regression models, and generic models. For linear regression models, according to Box and Tiao (1992), the variance of the model parameter θ_j is

$$\sigma_j^2 = \frac{1}{n-k} \left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\theta}} \right)^T \left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\theta}} \right) \left(\left(\mathbf{X}^T \mathbf{X} \right)^{-1} \right)_{jj} \tag{17}$$

where k = number of θ-parameters in the model, $\mathbf{y}$ = vector of observed results, $\mathbf{X}$ = matrix of observations, $\hat{\boldsymbol{\theta}}$ = mean vector of model parameters, and $(\,)_{jj}$ identifies the j^{th} diagonal component. The variance of the model error is

$$\sigma_\varepsilon{}^2 = \frac{1}{n-k} \left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\theta}} \right)^T \left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\theta}} \right) \tag{18}$$

The derivative of Eq. (17) with respect to n is

$$\frac{\partial \sigma_j}{\partial n} = \frac{-\sigma_j}{2(n-k)} \tag{19}$$

and, similarly, the derivative with respect to Eq. (18) is:

$$\frac{\partial \sigma_\varepsilon}{\partial n} = \frac{-\sigma_\varepsilon}{2(n-k)} \tag{20}$$

Similar derivatives are obtained for nonlinear regression models, where the expressions analogous to Eqs. (17) and (18) are made available by Seber and Wild (2003). For generic models, such as finite element models, it is argued that the epistemic uncertainty is primarily present in the random variables that are input to the model. This is assumed here, although some efforts have been made by Haukaas and Gardoni (2011) and others to incorporate epistemic uncertainty into finite element models. To this end, it is of interest to identify how the epistemic uncertainty in a physical random variable say, concrete strength, is affected by inclusion of more information. As a starting point, consider the well-known expression for variance:

$$\sigma_j^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \mu_j \right)^2 \tag{21}$$

where $x_i = i^{\text{th}}$ observation and μ_j = mean of observations. The derivative is:

$$\frac{\partial \sigma_j}{\partial n} = \frac{-\sigma_j}{2(n-1)} \tag{22}$$

It is emphasized that the expressions for the derivative $\partial\sigma_j/\partial n$ are derived under the assumption that the mean model is unaffected by added observations. Clearly, this may not be the case in practical circumstances, where one added observation may precipitate an increase in the variance of the epistemic random variables. However, if the fundamental model form is correct, then over time, new data will serve to reduce the epistemic uncertainty. The previously presented formulas are derived on this basis.

It is observed that the minus sign in equations above for $\partial\sigma_j/\partial n$ correctly implies that that the standard deviation of the model parameter is reduced when n is increased. Moreover, it is observed that the reduction in the standard deviation is smaller when n is large than when n is small. In other words, a model that is based on a large number of observations will benefit less from a few more observations. Furthermore, it is observed that the reduction in the standard deviation is greater when k is large than when k is small. In other words, a model with more parameters, *i.e.*, a more complex model, will benefit more from new observations.

The last derivative in the right-hand side of Eq. (16) is the inverse of the cost of obtaining one data point. Naturally, the quantification of this cost is challenging. In fact, some observations are readily obtained, while others come at a significant cost. Examples of typical engineering observations that are counted by n include: 1) Testing of building on a shake table, which can be used to calibrate the building response, damage, and repair cost models; 2) Analysis of a highly refined numerical building model, which can be used to calibrate building response models; 3) Survey of buildings damaged in earthquakes, which can be used to calibrate building damage models; and 4) Claims reports from insurance companies, which can be used to calibrate building damage models. Although the cost of obtaining such data vary, it is assumed in this study that each observation will take two to three days of paid work and cost around \$500, *i.e.*, $\partial c/\partial n = 500$.

The sensitivity measure $\partial\beta/\partial c$ is now evaluated for the models that were employed in the regional risk analysis for the 622 buildings at the UBC campus. The model types were presented in Table II. Table VII identifies the five models with highest value of $\partial\beta/\partial c$. This means that allocating resources for improving these models has the greatest impact on the reliability index. In particular, gathering data to improve the concrete shear wall structural response model has the greatest effect on the reliability index per dollar spent.

The positive sign of the $\partial\beta/\partial c$-values indicates that the reliability index increases when resources are allocated to data gathering. This makes sense, because the resulting model improvement reduces the uncertainty, which in turn reduces the probability of exceeding a \$100 million loss. This decrease in probability is reflected by the increase in the reliability index, which is correctly captured by the positive sign of the $\partial\beta/\partial c$-values.

Table VIII identifies the five models with lowest value of $\partial\beta/\partial c$. In other words, these are the models for which data gathering and model improvement would not have significant impact on the reliability index. While structural response models rank highest according to $\partial\beta/\partial c$, structural damage models rank lowest. In other words, in the context of the models employed in this study, it appears worthwhile to focus attention on reducing the epistemic uncertainty in the structural response models, *i.e.*, models for building displacement and acceleration, rather than the damage models. However, an important remark is made in regards to these results: The ranking according to $\partial\beta/\partial c$ depends on the number of instances of a model type in the analysis. For example, in the present analysis, almost 54% of the building value is associated with concrete shear wall buildings, *i.e.*, buildings of the type that ranked first in Table VII. On one hand, this skews the results

towards higher $\partial\beta/\partial c$-values for this building type. Although this is not the only reason for the observed result, it should be duly noted when applying these results to regions with other compositions of the building stock. On the other hand, the ranking in Table VII and Table VIII are still valuable for the considered region. For researchers who seek to improve the risk assessment for this particular region, the results in Table VII and Table VIII are valid as measures to guide the allocation of resources for model improvement.

An additional remark about the results in Table VII, particularly the high rank of concrete shear wall structural response models, is made in regards to the results in the previous section. There it was noted that several of the concrete shear wall buildings are primary candidates for more refined structural analysis to reduce epistemic uncertainty. It is interesting to note that this type of building consistently ranks high according to both sensitivity measures.

Table VII: Top five model types according to cost of model improvement.

Model Type	$\partial\beta/\partial c$ [$\cdot 10^{-6}$]
Concrete Shear Wall Structural Response	46.7
Wood Large Frame Structural Response	4.4
Wood Light Frame Structural Response	1.0
Concrete Frame with Masonry Infill Wall Structural Response	0.52
Concrete Moment Frame Structural Response	0.50

Table VIII: Bottom five model types according to cost of model improvement.

Model Type	$\partial\beta/\partial c$ [$\cdot 10^{-12}$]
Wood Light Frame Structural Damage	0.2
Steel Light Frame Structural Damage	0.09
Steel Braced Frame Structural Damage	0.06
Steel Moment Frame Structural Damage	0.03
Steel Frame with Masonry Infill Wall Structural Damage	0.02

7. Conclusions

The overarching vision behind this paper is twofold. First, it is sought to identify and characterize epistemic uncertainty in a comprehensive manner. This is important because epistemic uncertainty, such as model uncertainty, is reducible and has significant influence on risk estimates. Second, it is sought to allocate resources in an optimal manner to reduce the epistemic uncertainty. The first goal is achieved by utilizing a library of probabilistic models that contain random variables that represent epistemic uncertainty. These models are implemented in a computer program, called Rt, dedicated to multi-model reliability analysis. Rt is employed in this paper to conduct risk analysis for a region in Vancouver, Canada that comprises 622 buildings. The second goal is addressed in this paper by the development of two new reliability sensitivity measures. These are implemented in Rt and evaluated in the regional risk analysis. The results show that the

epistemic uncertainty associated with the models for concrete shear wall buildings is the most cost-effective to address for this region.

Acknowledgements

The financial support from the Natural Science and Engineering Research Council of Canada (NSERC) through the Strategic Project Grant number 336498-06 is gratefully acknowledged. The first author also thanks the Killam Trusts for support in the form of a Killam Doctoral Fellowship. The authors also thank Mr. Aprizal Kristijanto for surveying buildings at the campus of the University of British Columbia in Vancouver.

References

Box, G. E. P., and G. C. Tiao. *Bayesian inference in statistical analysis*. New Jersey: Wiley-Interscience, 1992.

Der Kiureghian, A. First-and second-order reliability methods. In *Engineering design reliability handbook.*, eds. E. Nikolaidis, D. M. Ghiocel and S. Singhal. Boca Raton, FL: CRC Press, 2005.

Gardoni, P., A. Der Kiureghian, and K. M. Mosalam. Probabilistic capacity models and fragility estimates for reinforced concrete columns based on experimental observations. *Journal of Engineering Mechanics* 128 (10): 1024-38, 2002.

Mahsuli, M., and T. Haukaas. A computer program for multi-model reliability and optimization analysis. *Journal of Computing in Civil Engineering,* In Press, 2012.

Seber, G. A. F., and C. J. Wild. *Nonlinear regression*. New York: Wiley-IEEE, 2003.

Wen, Y. K. *Structural load modeling and combination for performance and safety evaluation*. Amsterdam, Netherlands: Elsevier, 1990.

Reliable Dynamic Analysis of Structures Using Imprecise Probability

Mehdi Modares and Joshua Bergerson
Department of Civil, Architectural, and Environmental Engineering,
Illinois Institute of Technology, Chicago, IL, 60616, USA,
Emails: mmodares@iit.edu, jbergers@iit.edu

Abstract: Dynamic analysis of a structure is an essential procedure in designing a structure subjected to a system of suddenly applied loads such as wind or earthquake excitations. However, throughout conventional dynamic analyses, the existence of uncertainties in mechanical properties of a structure has not been considered. One method to quantify the presence of uncertainty in a physical system is to use an imprecise probability framework based on the concept of p-boxes.

In this work, a new method for reliable dynamic analysis of structures using p-box based imprecise probability is developed. This method establishes a framework for handling incomplete information in structural dynamics. Using this method, the reliability and robustness in dynamic analysis are achieved. Example problem that illustrates the developed algorithm is presented.

Keywords: Structural Dynamics; Uncertainty; Imprecise Probability; P-Box.

1. Introduction

Throughout conventional dynamic analysis, structural parameters in a system (including geometric, material, and load properties) are treated as deterministic values. Although the true values for structural properties generally have small variations from the deterministic values used in analysis, these uncertainties affect the system's dynamic response significantly.

There are two distinct types of uncertainty: aleatoric and epistemic. Aleatoric uncertainty is due to inherent randomness in the system. This type of uncertainty is irreducible with higher precision or more data acquisition. Epistemic uncertainty is due to lack of knowledge, modeling errors, and/or insufficient data. This uncertainty, also known as reducible uncertainty, can be reduced with further investigations, more data, and updated models.

A method for handling uncertainty in a system with no assumptions of a well-defined Probability Density Function (PDF) is imprecise probability. This method involves setting bounds on the Cumulative Distribution Function (CDF), between which the true CDF exists. Williamson and Downs (1990) developed a method for combining imprecise probabilities using the four basic arithmetic operations, known as dependency bounds convolutions. These methods are enveloping methods in order to retain robustness in calculations, while avoiding overestimation on bounds due to dependency between different imprecise probabilities. Ferson and Donald (1998) developed a method for handling imprecise probability, referred to as Probability Bounds Analysis (PBA). Another method for handling imprecise probability is to discretize the upper and lower CDFs into probability boxes (p-boxes), each of equal probability mass. (Ferson et al. 1998)

A much older framework for handling imprecise probability is the Dempster-Shafer approach. While

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

the development of the Dempster-Shafer (DS) structures has been confirmed several times, the methods proposed for combining multiple DS structures has shown complexities. Ferson et al. (2003) suggested that the concept of Dempster-Shafer structures and p-boxes are in fact highly related and outlined procedures for converting from one to the other. They also expanded methods for combining multiple DS structures or p-boxes, for operations including: enveloping, intersection, unweighted and weighted mixtures, averaging, and logarithmic pooling. Moreover, they outlined the various desired properties of any method of combining multiple p-boxes or DS structures, and summarized which methods satisfied these properties. From this, they concluded that the best method for combining is enveloping, also used in this work.

Modares et al. (2006) showed a method of dynamic analysis using finite element method for mechanical uncertainties defined by interval variables, which yields exact bounds on a system's natural frequencies. This method utilizes the monotonic behavior of eigenvalues of symmetric matrices subject to non-negative definite perturbations. Zhang et al. (2010) outlined a method for extending imprecise probability methods to finite element analysis with p-box uncertainty in loading in static problems.

In this work, a new method for reliable dynamic analysis of structures using p-box based imprecise probability is developed which establishes a framework for handling incomplete information in structural dynamics. Independent uncertainties in stiffness defined by p-boxes are considered. This method applies the method for bounding natural frequencies, developed by Modares et al. (2006), to analysis of systems with independent interval uncertainty for each level of p-box discretization. Using this method, the reliability and robustness in dynamic analysis of an uncertain system are achieved. Although this work represents p-box uncertainties in stiffness properties, the extension of the developed method for uncertainties in the mass matrix is straightforward. Numerical example that illustrates the developed algorithm is presented, as well as comparison to interval Monte-Carlo simulation.

2. Review of Deterministic Frequency Analysis

The equation of motion for an undamped system is:

$$[M]\{\ddot{x}\} + [K]\{x\} = \{0\}, \tag{1}$$

where [M] is the global mass matrix, [K] is the global stiffness matrix, $\{x\}$ is the displacement vector, and $\{\ddot{x}\}$ is the acceleration vector. Assuming a harmonic motion for the displacement

$$\{x\} = \{u\}e^{i\omega t} \tag{2}$$

the equation of motion becomes a generalized eigenvalue problem as:

$$([K] - (\omega^2)[M])\{u\} = \{0\} \tag{3}$$

or:

$$[K]\{u\} = (\omega^2)[M]\{u\}, \tag{4}$$

where ω is the set of circular natural frequencies of the system and $\{u\}$ is the set of their corresponding mode shapes.

3. Review of Dynamic System with Independent Interval Stiffness Uncertainties

Considering a dynamic system with independent interval uncertainty in material properties, the interval eigenvalue problem is:

$$(\sum([l_i,u_i])[\bar{K}_i])\{u\} = (\tilde{\omega}^2)[M]\{u\} \quad (5)$$

where, $[l_i, u_i]$ are the element interval of uncertainty, $[\bar{K}_i]$ is the deterministic contribution of the element stiffness matrix to the global structural stiffness, and $\tilde{\omega}^2$ are the interval natural circular frequencies.

For this system, it is proven that the solutions of two deterministic problems are sufficient to bound all natural frequencies (Modares et al., 2006):

$$(\sum(u_i)[\bar{K}_i])\{u\}=(\omega^2_{max})[M]\{u\} \quad (6)$$

$$(\sum(l_i)[\bar{K}_i])\{u\}=(\omega^2_{min})[M]\{u\} \quad (7)$$

4. Problem Statement

In this work, a method for finite-element-based dynamic analysis of a system with independent uncertainty in the stiffness of each member defined by p-boxes is developed. The steps for the developed method are described in the next sections.

4.1 P-box Quantification of Input Uncertainty

Let $\overline{F}_i(x)$ and $\underline{F}_i(x)$ represent the upper and lower bounds of the CDF of the stiffness for member *i*, respectively. Thus, $\overline{F}_i(x)$ and $\underline{F}_i(x)$ are the bounding functions for the p-box describing the stiffness of every member *i* (Figure 1).

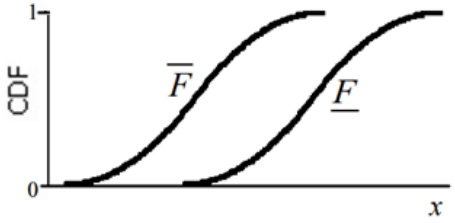


Figure 1. A general p-box.

A p-box can be discretized into *z* intervals, each with equal probability mass of 1/z. It is worth noting that by conducting this transformation, it is assumed all uncertainty exists in the value of the stiffness and not in the cumulative probability value. Let $\gamma \in [0,1]$ represent the value of the CDF functions $\overline{F}_i(x_L)$ and $\underline{F}_i(x_U)$. The discretization is done by selecting values of $\overline{F}_i(x_L)$ corresponding to $\gamma = (q-1)/z$ and

selecting values of $\underline{F}_i(\mathrm{x}_U)$ corresponding to $\gamma = q/z$ for every member *i* for $q = 1, 2, \ldots, z$.

This may be expressed using the inverse functions $\overline{F}_i^{-1}$ and $\underline{F}_i^{-1}$ as:

$$\mathrm{x}_{iqL} = \overline{F}_i^{-1}\left(\frac{q-1}{z}\right), \tag{8}$$

$$\mathrm{x}_{iqU} = \underline{F}_i^{-1}\left(\frac{q}{z}\right). \tag{9}$$

These bounds can be combined for each value of *q* and written in interval form for member *i* as:

$$\tilde{\mathrm{x}}_{iq} = [\mathrm{x}_{iqL}, \mathrm{x}_{iqU}] \tag{10}$$

with each $\tilde{\mathrm{x}}_{iq}$ having a probability mass of (1/*z*) for $i = 1, 2, \ldots, N$, where *N* is the number of members in the system.

One issue with this discretization process is that when evaluating $\overline{F}_i^{-1}(0)(q=1)$, many common distributions produce infinite tails. This problem also exists when evaluating $\underline{F}_i^{-1}(1)(q=z)$. Therefore, the tails must be truncated. Ferson et al. (2003) details the effects of truncating infinite tails and concluded that while the decision of where to truncate can have an effect on analysis, it is generally so small that it is negligible compared with the magnitude of the p-box or compared to the effect of discretizing the p-box even when *z* is sufficiently large. Thus, the infinite tails of $\overline{F}_i(\mathrm{x}_L)$ and $\underline{F}_i(\mathrm{x}_U)$ are truncated at $\gamma = \nu$ and $\gamma = (1-\nu)$, respectively, where $\nu << (1/z)$. Thus, for $\overline{F}_i(\mathrm{x}_L)$ and $\underline{F}_i(\mathrm{x}_U)$ with infinite tails:

$$\mathrm{x}_{i1L} = \overline{F}_i^{-1}(\nu) \tag{11}$$

$$\mathrm{x}_{izU} = \underline{F}_i^{-1}(1-\nu) \tag{12}$$

4.2 P-BOX BASED FREQUENCY ANALYSIS

The monotonicity concept for natural frequencies proven by Modares et al. (2006) is used for bounding each p-box based eigenvalue problem (Eqs. (6) & (7)) as follows:

$$(\mathrm{x}_{1(q_1)U} \cdot [\overline{\mathrm{K}}_1] + \mathrm{x}_{2(q_2)U} \cdot [\overline{\mathrm{K}}_2] + \ldots + \mathrm{x}_{N(q_N)U} \cdot [\overline{\mathrm{K}}_N])\{u\} = (\omega_U^2)[\mathrm{M}]\{u\} \tag{13}$$

$$(\mathrm{x}_{1(q_1)L} \cdot [\overline{\mathrm{K}}_1] + \mathrm{x}_{2(q_2)L} \cdot [\overline{\mathrm{K}}_2] + \ldots + \mathrm{x}_{N(q_N)L} \cdot [\overline{\mathrm{K}}_N])\{u\} = (\omega_L^2)[\mathrm{M}]\{u\} \tag{14}$$

where q_i are independently defined as possibilities of discretization levels for element *i*: $q_i \in [1, 2, \ldots, z]$. The results are arranged and condensed back down to *z* values for each ω_L and ω_U following procedure for such condensation outlined by Williamson and Downs (1990). These *z* intervals are then used to construct p-boxes for each natural frequency.

5. Numerical Example

The numerical example obtains the natural circular frequencies of a 2D three-element truss with 3 active degrees of freedom, with independent uncertainty in the modulus of elasticity of each member. The uncertainty in each member is represented by a p-box, with lower and upper bounds following normal distribution with given properties:

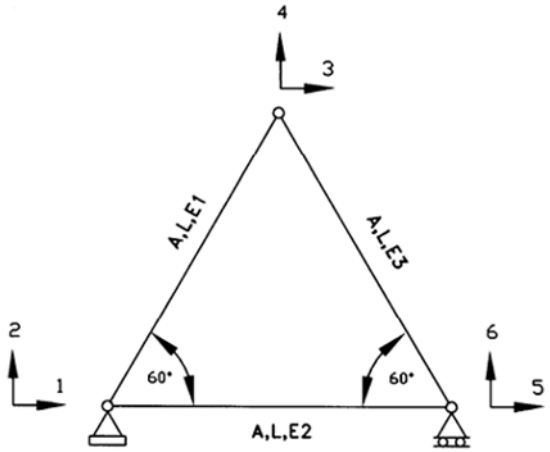


Figure 2. 2D truss system

The structure's properties are as follows:

$$\text{E1: } \tilde{\mu} = [.95,\ 1.05] \cdot E,\ \sigma = .0194 \cdot E$$
$$\text{E2: } \tilde{\mu} = [.85,\ 1.15] \cdot E,\ \sigma = .0582 \cdot E$$
$$\text{E3: } \tilde{\mu} = [.90,\ 1.10] \cdot E,\ \sigma = .0388 \cdot E$$

The infinite tails are truncated at CDF levels of ν and $1-\nu$, where $\nu = .005$. Figures 3a, 3b, and 3c depict the p-boxes as defined above for each modulus of elasticity.

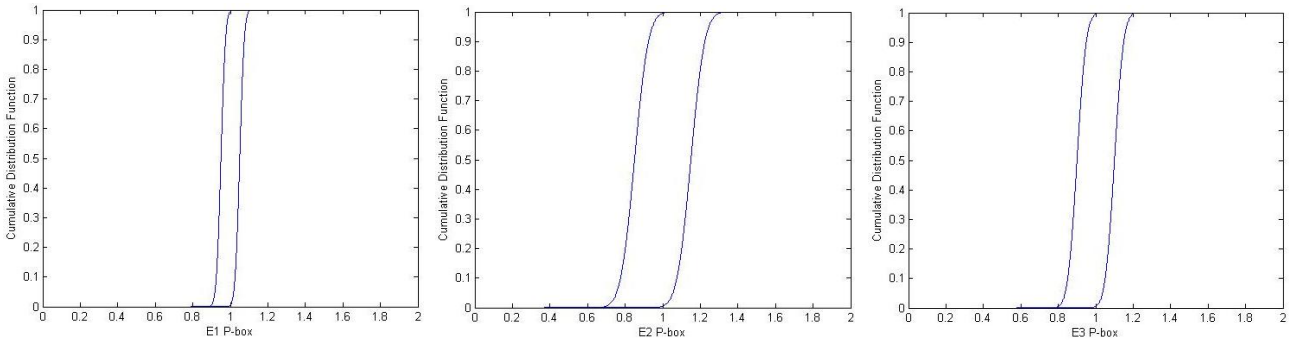


Figure 3a, 3b, & 3c. P-boxes E1, E2, & E3, respectively.

5.1 SOLUTION USING DEVELOPED METHOD

The above three p-boxes are discretized into ten intervals. The results are given in the Table I in the form of the interval of Eq. (10).

Table I. Coordinates for the discretized p-boxes of E1, E2, and E3.

$\tilde{E}1$/E		$\tilde{E}2$/E		$\tilde{E}3$/E	
x_{1qL}	x_{1qR}	x_{2qL}	x_{2qR}	x_{3qL}	x_{3qR}
0.9000	1.0251	0.7001	1.0754	0.8001	1.0503
0.9251	1.0337	0.7754	1.1010	0.8503	1.0673
0.9337	1.0398	0.8010	1.1195	0.8673	1.0797
0.9398	1.0451	0.8195	1.1353	0.8797	1.0902
0.9451	1.0500	0.8353	1.1500	0.8902	1.1000
0.9500	1.0549	0.8500	1.1647	0.9000	1.1098
0.9549	1.0602	0.8647	1.1805	0.9098	1.1203
0.9602	1.0663	0.8805	1.1990	0.9203	1.1327
0.9663	1.0749	0.8990	1.2246	0.9327	1.1497
0.9749	1.1000	0.9246	1.2999	0.9497	1.1999

Figures 4a, 4b, and 4c show the discretized p-boxes for each modulus of elasticity.

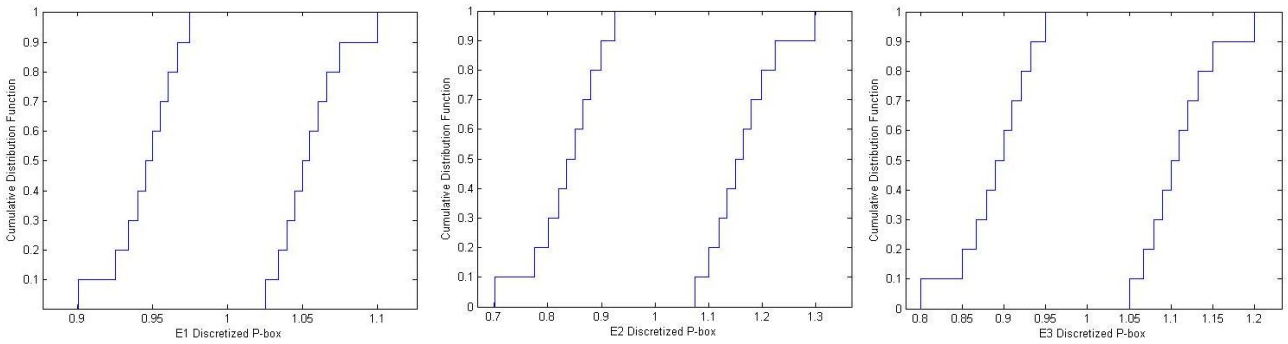


Figure 4a, 4b, & 4c. Discretized p-box for E1, E2, & E3, respectively.

Using Eqs. (13) and (14), lower and upper bounds on each natural frequency of the system are obtained, and the data is condensed to *z* intervals. Results are given in Table II and depicted in Figures 5a, 5b, and 5c.

Table II. Condensed p-box coordinates for all natural frequencies of the system

$\omega_{1a}\cdot L/\sqrt{(E/\rho)}$		$\omega_{2a}\cdot L/\sqrt{(E/\rho)}$		$\omega_{3a}\cdot L/\sqrt{(E/\rho)}$	
$\omega_{1aL}\cdot L/\sqrt{(E/\rho)}$	$\omega_{1aU}\cdot L/\sqrt{(E/\rho)}$	$\omega_{2aL}\cdot L/\sqrt{(E/\rho)}$	$\omega_{2aU}\cdot L/\sqrt{(E/\rho)}$	$\omega_{3aL}\cdot L/\sqrt{(E/\rho)}$	$\omega_{3aU}\cdot L/\sqrt{(E/\rho)}$
0.5661	0.6620	0.8911	1.0345	1.2189	1.4091
0.5845	0.6646	0.9138	1.0395	1.2542	1.4157
0.5915	0.6666	0.9286	1.0437	1.2672	1.4205
0.5957	0.6684	0.9352	1.0474	1.2743	1.4251
0.5986	0.6702	0.9399	1.0510	1.2796	1.4297
0.6011	0.6720	0.9445	1.0545	1.2846	1.4342
0.6035	0.6742	0.9488	1.0589	1.2891	1.4394
0.6059	0.6767	0.9535	1.0638	1.2937	1.4463
0.6084	0.6804	0.9577	1.0750	1.2990	1.4579
0.6117	0.6964	0.9639	1.0935	1.3056	1.4896

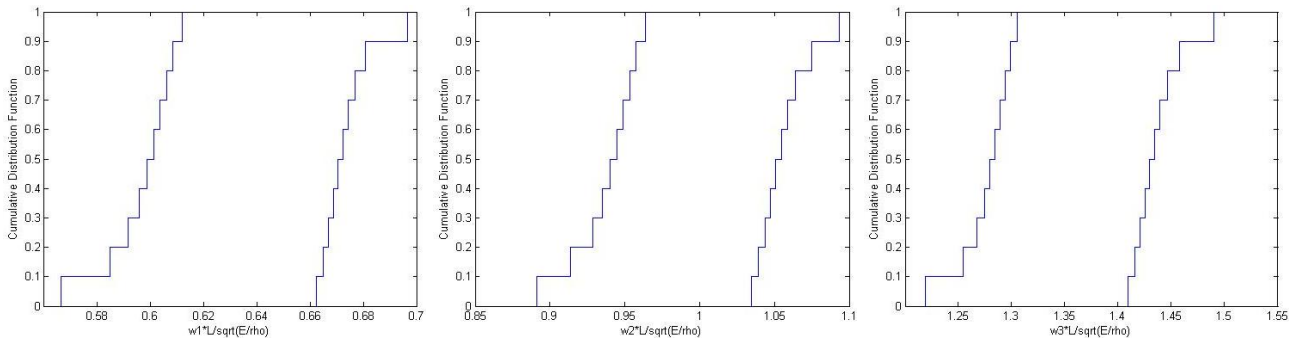


Figure 5a,5b, & 5c. Condensed p-boxes for ω_1, ω_2, & ω_3, respectively.

5.2 ALTERNATE SOLUTION USING INTERVAL MONTE-CARLO SIMULATION

Alternatively, the problem is solved using 10^4 interval Monte-Carlo simulation. One Monte-Carlo simulation is completed by randomly selecting a CDF value for each member in the system, where the CDF is bounded by the interval $[\nu,\ 1-\nu]$, or [.005, .995]. This CDF value is then used for its corresponding lower and upper values of the modulus of elasticity for each member as:

$$\mathrm{E}_{i\mathrm{L}} = \mu_{i\mathrm{L}} + \sigma_i \cdot \varphi^{-1}(CDF_i)$$

$$\mathrm{E}_{i\mathrm{U}} = \mu_{i\mathrm{U}} + \sigma_i \cdot \varphi^{-1}(CDF_i)$$

where φ^{-1} is the quantile function. The bounding values are used in Eqs. (13) and (14) to solve for the

lower and upper natural frequencies. The results for lower and upper natural frequencies are sorted and condensed to ten discretizations. The results are summarized below in Table III.

Table III. Condensed p-box coordinates from Monte-Carlo simulation for all natural frequencies

$\omega_{1MC}\cdot L/\sqrt{(E/\rho)}$		$\omega_{2MC}\cdot L/\sqrt{(E/\rho)}$		$\omega_{3MC}\cdot L/\sqrt{(E/\rho)}$	
$\omega_{1LMC}\cdot L/\sqrt{(E/\rho)}$	$\omega_{1UMC}\cdot L/\sqrt{(E/\rho)}$	$\omega_{2LMC}\cdot L/\sqrt{(E/\rho)}$	$\omega_{2UMC}\cdot L/\sqrt{(E/\rho)}$	$\omega_{3LMC}\cdot L/\sqrt{(E/\rho)}$	$\omega_{3UMC}\cdot L/\sqrt{(E/\rho)}$
0.5719	0.6595	0.8969	1.0299	1.2303	1.4035
0.5923	0.6623	0.9262	1.0361	1.2668	1.4107
0.5962	0.6644	0.9335	1.0409	1.2745	1.4163
0.5990	0.6663	0.9391	1.0447	1.2802	1.4208
0.6015	0.6680	0.9437	1.0484	1.2850	1.4249
0.6038	0.6697	0.9482	1.0517	1.2894	1.4293
0.6060	0.6715	0.9521	1.0555	1.2938	1.4335
0.6083	0.6736	0.9566	1.0597	1.2984	1.4388
0.6110	0.6765	0.9617	1.0655	1.3040	1.4461
0.6145	0.6902	0.9687	1.0885	1.3118	1.4777

Figures 6a, 6b, and 6c show the results from both the developed method (solid lines) and the Monte-Carlo simulation (dashed lines). As observed, the bounds on natural frequency from Monte-Carlo simulation are inner-bounds of the developed method.

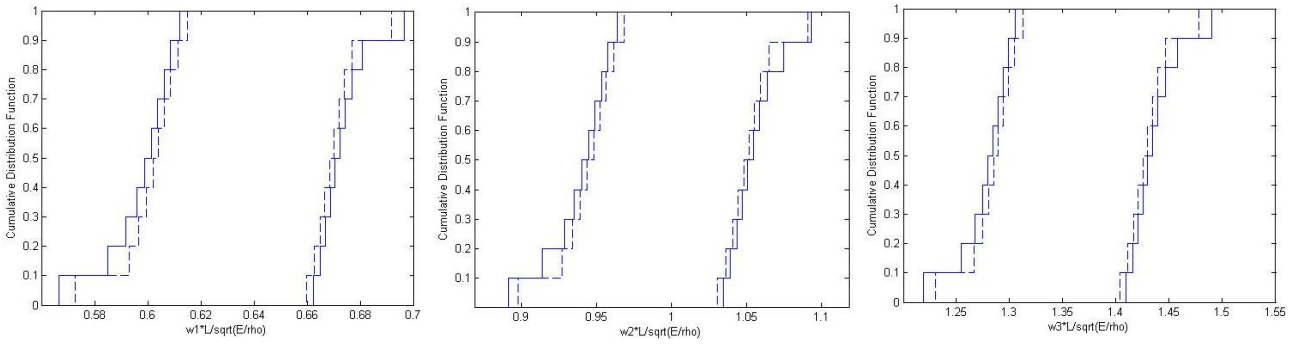


Figure 6a,6b, & 6c. Condensed p-boxes for ω_1, ω_2, & ω_3, respectively, with Monte-Carlo results represented by dashed lines.

6. Conclusion

A new method for dynamic analysis of structures with uncertain properties defined by independent p-boxes is developed. The developed method allows for uncertainty in the stiffness matrix, but can be shown to handle uncertainty in the mass matrix under the same framework. The developed method also allows for an uncertain dynamic analysis, yielding best-possible, discrete p-boxes for all natural frequencies of the dynamic system. The results obtained by this method are upperbounds of the results obtained by interval Monte-Carlo simulation procedures.

References

Augenbaugh, J. M., and Paredis, C. J. J. The Value of Using Imprecise Probabilities in Engineering Design, *Journal of Mechanical Design*, 128(4), 969-979, 2006.

Bellman, R. *Introduction to Matrix Analysis*. New York: McGraw-Hill Book Company, 1970. Print.

Berleant, D., and Goodman-Strauss, C. Bounding the Results of Arithmetic Operations on Random Variables of Unknown Dependency Using Intervals, *Reliable Computing*, 4(2), 147-165, 1998.

Ferson, S., and Donald, S. Probability Bounds Analysis, *Probabilistic Safety Assessment and Management*, Mosleh, A., and Bari, R.A. Eds., Springer-Verlag, New York, NY, 1203-1208, 1998.

Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D., S., Sentz, K. Construction probability boxes and Dempster-Shafer structures, Tech. Rep. SAND2002-4015, Sandia National Laboratories 2003.

Modares, M., Mullen, R., L., Muhanna, R., L. Natural Frequencies of a Structure with Bounded Uncertainty, *Journal of Engineering Mechanics*, ASCE, December ed., 1363-1371, 2006.

Muhanna, R. L., and Mullen, R. L. Uncertainty in Mechanics Problems-Interval-Based Approach. *Journal of Engineering Mechanics*, June-2001, 557-566, 2001.

Oberkampf, W., Helton, J. and Sentz, K. "Mathematical Representations of Uncertainty," *AIAA Non-Deterministic Approaches Forum, AIAA 2001-1645*, Seattle, WA, April 16-19, 2001.

Williamson, R., and Downs, T. Probabilistic Arithmetic I: Numerical Methods for Calculating Convolutions and Dependency Bounds, *International Journal of Approximate Reasoning*, 4, 89-158, 1990.

Zhang, H., Mullen, R., L., Muhanna, R., L. Finite Element Structural Analysis using Imprecise Probabilities Based on P-Box Representation, *4th International Workshop on Reliable Engineering Computing*, 211-225, 2010.

Partial safety factor concepts for shear loaded existing reinforced concrete structures

[1)]Thomas Moser, [2)]Alfred Strauss
[1)]*Schimetta Consulting, 1190 Vienna, Austria, thomas.moser@schimetta.co.at*
[2)]*Department of Civil Engineering and Natural Hazards, University of Natural Resources and Life Science, 1190 Vienna, Austria, alfred.strauss@boku.ac.at*

Abstract: General probabilistic methods for reliability assessment allow uncertainties in the planning, measurement, implementation and maintenance of new and existing reinforced concrete bridges to be taken into consideration. The implementation of probabilistic verification procedures permitted in recognised standards for assessing the reliability of existing reinforced concrete structures has so far only had limited application and needs to be optimised accordingly. The objective of this paper is to present and discuss a method of adapting partial safety coefficients based on recorded stochastic models representing the properties of existing and new structures and a possible reduction in reliability level. The method will be demonstrated on a typical three-span slab bridge used by Austrian Federal Railways.

Keywords: Shear capacity, Existing Concrete structures; Partial Safety factors

1. Introduction

Due to the constant increase in traffic, the re-calculation of existing structures plays a more and more important role in civil engineering design work.

One method for the efficient assessment of existing bridges, which has a high level of flexibility regarding the variable requirements of structures, is the probabilistic / reliability based method, which is becoming available in the new European design standards for engineering. However, they only have limited acceptance due to the additional measurements and tests needed for calculations and the definition of responsibility regarding residual risk. The objective of this paper is therefore (a) to demonstrate the efficient assessment of existing reinforced steel structures using a methodical procedure for probabilistic analysis; especially regarding the verification of shear capacities of reinforced concrete, (b) to explain the rules in the standard regulations and the opportunities to adapt the reliability level in relation to the remaining service life, and (c) to discuss a method of adapting partial safety coefficients of the semi-probabilistic verification procedure based on measured and specified stochastic models and the required reliability level (Strauss et al., 2008). For the design of special structures, naturally also nonlinear finite element analysis is applied and even called for in respective guidelines such as the upcoming German guideline for jointless bridges (Bundesministerium für Verkehr Bau und Stadtentwicklung 2011), which demands nonlinear design, if soil conditions and span length warrant it. Within this contribution linear design assumptions concerning the soil structure interaction of concrete frame bridges without expansion joints will be investigated. In particular, the linearity between temperature and structural expansion will be studied, as will be the linearity of the strain field in the soil body and pavement in the transition area between concrete bridge deck and soil dam.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

2. Reliability of new and existing structures

Austrian standard ÖNORM EN 1990 permits a division into three consequence classes (*CC*1, *CC*2 and *CC*3) for the determination of a required reliability level, see Table I columns 1 & 2. The levels are laid down for the objects currently being planned and are defined for the reference period of one year.

Table I: Reliability index for planning new structures (JCSS, 2000; EN 1990, 2003)

Consequence class (*CC*)	Consequence on failure	Reference period 1 year	
		EN 1990	JCSS
CC1	Low	$\beta = 4.2$	$\beta = 3.1–4.2$
CC2	Medium	$\beta = 4.7$	$\beta = 3.3–4.4$
CC3	High	$\beta = 5.2$	$\beta = 3.7–4.7$

In addition to defining the consequence class, the "Probabilistic Model Code" of the Joint Committee on Structural Safety also differentiates costs incurred through the definition of the reliability index. This is shown in Table I column 4, specified as the range of β.

The assessment of existing bridges can normally be performed using probabilistic analysis (Strauss, 2008). These analysis methods allow information to be taken into account that has been recorded from the existing structures after many years, or even decades, of service and enable a shortened planned service life to be taken into account. A shortened remaining service life or a shortened observation period is a persuasive argument for a reduction of the maximum values for changing influences originally applied to the reference period (wind force, snow load, etc.) (Petschacher, 2010).

3. Reliability assessment of existing structures

Assessments of the safe capacity and usability of existing structures generally lead to the following results:

- Values below a minimum reliability level $\beta \leq \beta_l$ (l = low) regarding safe capacity and usability.
- Values below a critical threshold of the reliability level in safe capacity and usability $\beta_l \leq \beta \leq \beta_r$ (r = repair). Traffic over the structures and use of the structure must be restricted.
- There is no reduction in the defined reliability level regarding safe capacity or usability $\beta \geq \beta_r$.

Limit specifications for the above mentioned reliability level are documented for new structures in Austrian standard ÖNORM EN 1990. For existing structures the reliability thresholds β_l and β_r are to be established based on safety-relevant and economic factors. The JCSS Model Code includes such recommendations; see Table I, column 4. The following suggestions are made for consequence class *CC*2, for example:

$$\beta_r \geq 4.2 \quad \text{and} \quad \beta_l \geq 3.3 \quad \text{for the reference period of one year} \tag{1}$$

An alternative definition of the upper and lower threshold of the reliability indices β_l and β_r according to Steenbergen et al. (2011) is:

$$\beta_l = \beta - 1.5 \tag{2}$$

$$\beta_r = \beta - 0.5 \tag{3}$$

4. Shear force verification for existing structures

The design and dimensioning of new structures in line with specifications stipulated by European standards (ÖNORM EN 1992-2, 2007) involves the semi-probabilistic safety concept. This calculation method is based on partial safety factors in relation to influences and resistances, which have been calibrated according to probabilistic observations (Rackwitz, 2006). As mentioned in the introduction, applying the European standard specifications and the related partial safety concepts guarantees a reliability index in structural safety of $\beta \geq 3.8$ for a reference period of 50 years and a probability of failure of $p_f = 10^{-6}$ for a reference period of one year. The objective of the following observations is the verification of the probabilistic-based concept mentioned in the introduction to adapt partial safety factors for existing structures taking into consideration the β_r and β_l reliability level (Moser et al., 2010). The shear resistance is characterised by three threshold observations that have partial interdependence in the described dimensions. This section deals with the limit thresholds that depend on shear force reinforcement. For the other two thresholds, refer to Moser et al. (2010). The measurement value of the reinforcement-specific shear force resistance at design level according to ÖNORM EN 1992-2 (2007) is:

$$V_{Rd,s} = \frac{A_{sw}}{s} \cdot z \cdot f_{ywd} \cdot (cot\theta + cot\alpha) \cdot sin\alpha \tag{4}$$

where A_{sw} = surface of the shear force reinforcement, s = mean value of distance between the stirrups, z = mean value of the inside cantilever ($0.9d$), f_{ywd} = rated value of the yield point of the shear reinforcement, θ = mean value of the angle of the concrete pressure strut (in accordance with ÖNORM B 1992-2 (2008) $31 \leq \theta \leq 45$), α = mean value of the angle of the lateral reinforcement.

A probabilistic observation of shear force verification for existing structures generally allows a more flexible approach to existing uncertainties. If probabilistic analysis is limited to the loading of the structure itself g_1, g_2 and vertical traffic loading (ÖNORM EN 1991-2, 2004) the limit condition function works like this:

$$\tilde{G} = \Theta_R \cdot R - (\Theta_G \cdot G + \Theta_Q \cdot Q) \tag{5}$$

Where R = resistance, Θ_R = model fuzzy resistance, G = body weight, Θ_G = model fuzzy own weight, Q = change in load, and Θ_Q = model fuzzy change in load. The base variables of Eq (5) are described by their mean value μ, their standard deviation σ and their probability type. The effect/load models in the form of probability density functions (PDFs) that are contained within the limit condition function Eq. (5) can, in part, be deduced from standard-specific nominal value specifications. For railway load model 71 referred to in the following, the estimation of the mean value assumes the standard-specific nominal value of the load as a fractile value of 95%. The assumption of the distribution $CoV = 0.10$ and 0.20, for example, leads to the mean value of the railway load model 71 shown in Table II.

Table II: Mean values of railway load model 71 assuming the standard-specific nominal value of the load as a fractile value of 95%

Variation coefficient	Mean value of concentrated load P	Mean value of distributed load p
0.10	214.59 kN	68.67 kN/m
0.20	187.97 kN	37.59 kN/m

4.1 INVERSE DETERMINATION OF PARTIAL SAFETY FACTORS

Observing the measurement model and at the same time the limit condition functions, as shown in Figure 1 is the basis for inverse determination of partial safety factors of the resistance and the effects. These are deduced in the following steps:

- Development of the probabilistic limit condition function from the measurement model
- Description of a starting dimension (e.g. geometric size) of the measurement model as a function of the partial safety coefficient and the sought after dimension of the measurement model.
- Formulation of the probabilistic limit condition function depending on the partial safety factors – substitution of the distributed starting dimensions due to its formulation based on partial safety factors.
- Definition of a target reliability index of the observed limit condition function
- Iterative adjustment of the partial safety coefficient through to fulfilment of the target reliability specification
- Final semi–probabilistic verification with the updated partial safety coefficients.

This procedure is applied later on for determining the partial safety factors in connection with the shear capacity of existing reinforced concrete bridges.

The following equation can be deduced from the maximum shear force resistance in equation (4):

$$\frac{A_{sw}}{s} \cdot z \cdot \frac{f_{yw,k}}{\gamma_s} \cdot (cot\theta + cot\alpha) \cdot sin\alpha = \left(\gamma_G \cdot V_{G,k} + \gamma_Q \cdot V_{Q,k}\right) \tag{6}$$

and with:

$$\tau_{y,k} = \frac{A_{sw}}{s} \cdot f_{yw,k} \cdot (cot\theta + cot\alpha) \cdot sin\alpha \tag{7}$$

formula (6) can be simplified in the following way:

$$\tau_{y,k} \cdot z \cdot \frac{1}{\gamma_s} = \left(\gamma_G \cdot V_{G,k} + \gamma_Q \cdot V_{Q,k}\right) \tag{8}$$

Through conversion we arrive at the inside cantilever as a function of the effect of the partial safety factor γ_s and the expression τ_{yk}:

$$z = \left(\gamma_G \cdot V_{G,k} + \gamma_Q \cdot V_{Q,k}\right) \cdot \frac{\gamma_s}{\tau_{y,k}} \tag{9}$$

The probabilistic limit condition function $\tilde{G}(V_{R,S})$ of the shear force resistance of the shear reinforcement is thus:

$$\tilde{G}(V_{R,S}) = \Theta_R \cdot \tau_y \cdot z - \left(\Theta_G \cdot V_G + \Theta_Q \cdot V_Q\right) \tag{10}$$

with:

$$\tau_y = \frac{A_{sw}}{s} \cdot f_{yw} \cdot (cot\theta + cot\alpha) \cdot sin\alpha \tag{11}$$

The variables f_{yw} and s are distributed, while surface A_{sw} and the angles α and θ are deterministic dimensions. The limit condition function $\tilde{G}(V_{R,S})$ is a result of applying equation (9) to equation (10).

$$\tilde{G}(V_{R,S}) = \Theta_R \cdot \tau_y \cdot \left(\gamma_G \cdot V_{G,k} + \gamma_Q \cdot V_{Q,k}\right) \cdot \frac{\gamma_s}{\tau_{y,k}} - \left(\Theta_G \cdot V_G + \Theta_Q \cdot V_Q\right) \tag{12}$$

5. Case Study: Ringstraßenbridge in Krems

The design of the bridge took place based on the standards valid at the time (ÖNORM B4000-2, 1952; ÖNORM B4003-1, 1956; ÖNORM B4200-4, 1957). The historical verification procedure for the shear force resistance resulted in a fraction of the dimensions required according to the current standards and directives. This is generally a problem that applies to the assessment of existing railway bridges in the critical sectors of supports and pillars.

Table III shows the parameters and related stochastic model for determining the shear force resistance of the Kremser Ringstraßenbrücke which contribute directly or indirectly to the iterative adaptation of the partial safety factors.

5.1 ITERATIVE DETERMINATION OF PARTIAL SAFETY FACTORS

Formula (12) derived above serves as the basis for iterative determination of the partial safety factors. The following assessments are based on an assumed CoV_{VQ} of the variable effect of 0.10 and 0.20. For the iterative determination of the partial safety coefficient, the static software program FReET (Novák et al., 2005) was used to define the limit condition functions, and, with the help of an expanded MonteCarlo method, the Latin Hyper Cube Sampling Method (Novák et al., 1997), it was possible to determine the reliability level for every step of iteration.

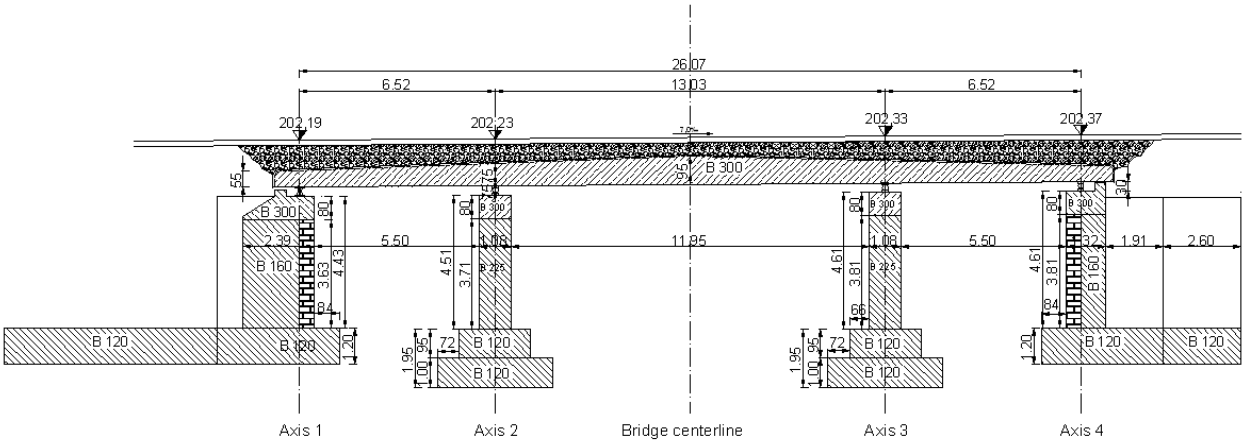


Figure 1. Ringstraßenbridge longitudinal section of bridge axis

Table III: Stochastic model for determining the partial safety coefficient of the lateral force resistance of the stress reinforcement of the Kremser Ringstraßenbridge

Characteristics	PDF	μ	CoV	Ref.
Model correction values				
Effect	LN	1	0.10	Braml, 2010
Resistance	LN	1.1	0.10	Braml, 2010
Reinforcement				
Yield strength f_y [MPa]	N	436	0.05	
Stress reinforcement A_{sw} [m^2/m]	Det.	5.24e-4		
Bow distance s [m]	N	0.45	0.067	
Effect				
Lateral force due to structure V_G [kN/m]	N	30.84	0.05	
Lateral force due to payload V_Q [kN/m]	N	44.90	0.1	
Geometry				
Width [m]	Det.	1.0		
Θ [rad]	Det.	0.54		
α [rad]	Det.	0.78		

Table IV: Required partial safety coefficients $\gamma_{G,cal}$ and $\gamma_{Q,cal}$ for the limit condition $\tilde{G}(V_{R,s})$ and the target reliability level $\beta = 4.30$; $\beta_r = 3.80$ and $\beta_l = 2.90$; with $\gamma_s = 1.15$ and $CoV_{VQ} = 0.10$ (0.20)

β_{erf}	Reference period	Partial safety factor of the effect $\gamma_{G,cal} = \eta\,\gamma_G$	$\gamma_{Q,cal} = \eta\,\gamma_Q$	β_{cal}
4.30	6	1.55	1.67	≈ 4.26 (4.38)
3.80	6	1.35	1.45	≈ 3.72 (3.84)
2.90	6	1.08	1.16	≈ 2.83 (3.00)

The determination of the required partial safety factor for a specified minimum reliability level (e.g. specified in the reliability index β) took place using simultaneous change to the partial safety factors γ_G and γ_Q using calibration factor η. The partial safety factor of the resistance γ_s was kept constant. The minimum reliability level for a reference period of 6 years was defined as follows (Moser et al., 2010):

$$\beta = 4.30, \qquad \beta_r = 3.80, \qquad \beta_l = 2.90.$$

This minimum reliability level produces partial safety factors shown in Table IV for a distribution in the effect model of $CoV_{VQ} = 0.10$ and $CoV_{VQ} = 0.20$. As a result, verification with $\gamma_{G,cal} = 1.08$, $\gamma_{Q,cal} = 1.16$ means that limited operation of the bridge is possible.

6. Results

A reliability index of $\beta \approx 3.8$ was determined for the limit condition $\tilde{G}(V_{R,S})$ using the partial safety factors specified according to ÖNOMR EN 1992-2 and ÖNORM EN 1990 which corresponds with the reliability index defined in ÖNORM EN 1990 for an observation period of 50 years.

Figure 2 shows the development of the calibration factor η for the partial safety factors γ_G and γ_Q compared with the required reliability level of β for a variation coefficient of the effect side of $CoV = 0.10$ and 0.20.

Example: The target reliability level $\beta = 3.8$ for a planned service life of 50 years yields $\beta_l = 2.8$ for a residual service life of 12 years. For this β Figure 2 shows the calibration factor $\eta = 0.82$ for the limit condition $\tilde{G}(V_{R,S})$ and thus $\gamma_{G,cal} = 0.82 \cdot 1.35 = 1.11$ and $\gamma_{Q,cal} = 0.82 \cdot 1.45 = 1.19$. The partial safety factors on the resistance side are kept as constants for this observation.

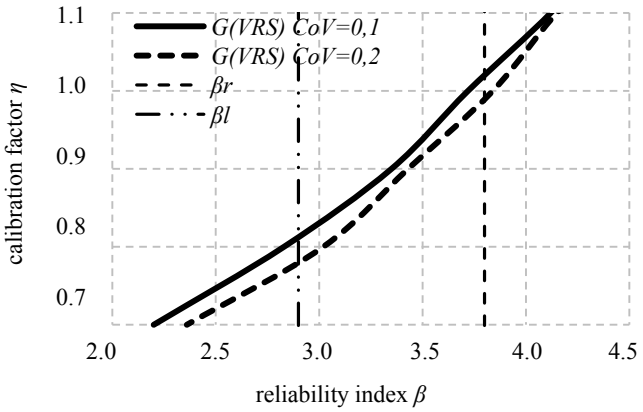


Figure 2. Development of the reliability index β as a function of calibration factor η for the partial safety factors γ_G and γ_Q

7. Summary

This article deals with engineering calculations of the shear capacity of existing reinforced concrete structures. Using the approach of inverse probabilistic calculation, we have presented a method for adapting the partial safety factors of the semi-probabilistic partial safety concept for a desired reliability level. The practical application of the method was demonstrated using an existing bridge, where during this optimisation and adaptation of the partial safety factors the effect side and resistance side were treated as variable dimensions. An extension of this method opens up the possibility of multiple simultaneous optimisations/adaptations in the form of Pareto-optimisation of the partial safety factors. Due to the calibration of the partial safety factors the semi-probabilistic verification routines specified in the standards and currently applied in practice can be kept so that there is no need to dispense with the opportunity to reduce the reliability index.

Acknowledgements

The research activities were financed by the FFG and ÖBB – Infrastruktur AG in the course of the ProbaS project – "Probabilistische Beurteilung von Strukturen".

This research was performed with the financial support of ASFINAG and BMVIT as part of the research project "Monitoring und Analyse von integralen Brückenbauwerke" and with the financial support of FFG during the EUREKA Eurostar research project "Risk Lifetime Assessment of Concrete Structures". The opinions and results published in this article are those of the authors and do not necessarily reflect those of the supporting organisations.

References

Braml T. *Zur Beurteilung der Zuverlässigkeit von Massivbrücken auf Grundlage der Ergebnisse von Überprüfungen am Bauwerk*. München: Dissertation an der Universität der Bundesweh, 2010.

Joint Committee on Structural Safety (JCSS). *Probabilistic Model Code 12th draft.* http://www.jcss.ethz.ch, 2000

Moser T., Strauss A., Bergmeister K. *Teilsicherheitsbeiwerte für bestehende Stahlbetontragwerke – Nachweis der Schubtragfähigkeit*; Paper in Review, 2011

Novák D., Teplý B., Keršner Z. *The role of Latin Hypercube Sampling method in reliability engineering*; In: Structural Safety and Reliability: 7th International Conference on Structural Safety and Reliability. Kyoto, 1997

Novák D. et al Revision 6/2005. *FREET – Multipurpose Probabilistic Software for Statistical, Sensitivity and Reliability Analysis – Program Documentation*

ÖNORM B4000-2. 1952. *Berechnung und Ausführung der Tragwerke; allgemeine Grundlagen; Raumgewichte von Baustoffen und Lagergut*

ÖNORM B4003-1. 1956. *Berechnung und Ausführung der Tragwerke; allgemeine Grundlagen; Eisenbahnbrücken*

ÖNORM B4200-4 1957 *Stahlbetontragwerke; Berechnung und Ausführung*

ÖNORM EN 1990. 2003. *Eurocode Grundlagen der Tragwerksplanung*

ÖNORM EN 1991-1-1. 2003. *Eurocode 1: Einwirkung auf Tragwerke Teil 1-1: Allgemeine Einwirkungen – Wichten, Eigengewicht, Nutzlasten im Hochbau*

ÖNORM EN 1991-2. 2004. *Eurocode 1: Einwirkung auf Tragwerke Teil 2: Verkehrslasten auf Brücken*

ÖNORM EN 1992-2. 2007. *Eurocode 2: Bemessung und Konstruktion von Stahlbeton und Spannbeton Teil 2: Betonbrücken - Bemessungs- und konstruktionsregeln*

ÖNORM B 1992-2. 2008. *Eurocode 2: Bemessung und Konstruktion von Stahlbeton- und Spannbeton Teil2: Betonbrücken - Bemessungs- und konstruktionsregeln; Nationale Festlegungen zu ÖNORM EN 1992-2, nationale Erläuterungen und nationale Ergänzungen*

Austrian Railway Federation. *Ausführungsplanung für den Bau einer Bahnbrücke über die Ringstraße und zwei Rad- bzw. Fußwege auf der Strecke Krems – Grein bei km 1,130,* 1959.

Petschacher M. Evaluating remaining lifetime of bridges by means of BWIM; Bridge Maintenance, Safety, Management and Life – Cycle Optimization – Frangopol, Sause & Kusko (eds). London: Taylor & Francis Group, 2010

Rackwitz R. *Zuverlässigkeit und Lasten im konstruktiven Ingenieurbau, Teil I: Zuverlässigkeitstheoretische Grundlagen*. München: Technische Universität ,2006

Steenbergen R.D.J.M., de Boer A., van der Veen C. *Safety assessment of existing concrete slab bridges of shear capacity*; ICASP; Zürich: *in print,* 2011

Strauss A., Bergmeister K., Hoffmann S., Pukl R., Novák D. *Advanced life – cycle analysis of existing concrete bridges*. Journal of Materials in Civil Engineering. 20 (1); Page 9 - 19, 2008.

Variance-based sensitivity analysis in the presence of correlated input variables

Thomas Most
Dynardo – Dynamic Software and Engineering GmbH, Weimar, Germany
thomas.most@dynardo.de

Abstract. In this paper we propose an extension of the classical Sobol' estimator for the estimation of variance based sensitivity indices. The approach assumes a linear correlation model between the input variables which is used to decompose the contribution of an input variable into a correlated and an uncorrelated part. This method provides sampling matrices following the original joint probability distribution which are used directly to compute the model output without any assumptions or approximations of the model response function.

Keywords: Sobol' indices, sensitivity analysis, correlation, regression

1. Introduction

In the assessment of computational engineering models sensitivity analysis has played an increasing role during the last two decades. For probabilistic models variance based sensitivity methods are very common. Based on the work of Sobol' (Sobol'1993) first order indices have been extended to total indices capturing higher order coupling effects between the input variables (Homma and Saltelli1996). Based on this early work a large number of extensions and improvements have been developed. A precise overview of existing variance based methods and their background is given in (Saltelli et al.2008). As stated in (Saltelli et al.2010) the total effect sensitivity index has become a prestigious measure to quantify higher order effects of input variables which is necessary for model simplification by variable fixing. However, the basic formulation of the variance decomposition and the common estimation procedures require so far independent input variables.

In practical applications often input variables are correlated, therefore, an increased interest in extending the classical methods for correlated inputs is marked in the recent years. In several investigations, assumptions about the model output or approximations up to a certain degree are necessary (Oakley and O'Hagan2004), (Xu and Gertner2008), (DaVeiga et al.2009). Complete model independent approaches for the estimation of total effect indices are not available so far. In (Jacques et al.2006) classical estimators have been used by grouping the correlated inputs which results in independent groups of parameters. However, even this approach is not useful for more complex models where almost all variables are coupled with each other.

In this paper we propose an extension of the classical Sobol' estimator. The basics of this method have been developed in (Most et al.2010), where a variable decomposition similar to (Xu and Gertner2008) was realized. The general idea has been adapted in (Mara and Tarantola2012) for normally distributed random variables. In our approach we assume a linear correlation model between the input variables, which is used

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

to decompose the samples into a correlated and an uncorrelated part with respect to a certain variable. This was realized in (Mara and Tarantola2012) by an orthogonalization schemes which may lead to non-unique solution depending on the variable order. In our approach we decompose the contribution of the input variables directly by using the correlation information of the standard normal space. This decomposition results in a unique solution. For non-normal distributions the Nataf model (Nataf1962) is applied to transform the original space and the correlation information to the standard normal space. The proposed method finally provides modified sampling matrices following the original joint probability distribution while keeping the samples of a certain variable unchanged. These samples are used directly to compute the model output without any assumptions or approximations of the model response function which makes the proposed method model independent. Additionally, reduced polynomial models are presented which are used to estimate the variance contribution of single variables.

Alternatively to the decomposition approach, in (Most et al.2010) a reordering approach has been proposed. In this method the investigated variable has been taken as the first one in the variables set for the Cholesky decomposition in the standard normal space. With this approach the samples required for the estimation of the sensitivity indices can be directly determined. This method has been adapted in (Kucherenko et al.2012). In our paper we do not discuss this approach, since the implementation is much more complicated compared to the decomposition approach.

2. Variance based sensitivity analysis

2.1. First order and total effect sensitivity indices

Assuming a model with a scalar output Y as a function of a given set of m random input parameters X_i

$$Y = f(X_1, X_2, \ldots, X_m), \tag{1}$$

the first order sensitivity measure was introduced as (Sobol'1993)

$$S_i = \frac{V_{X_i}(E_{\mathbf{X}_{\sim i}}(Y|X_i))}{V(Y)}, \tag{2}$$

where $V(Y)$ is the unconditional variance of the model output and $V_{X_i}(E_{\mathbf{X}_{\sim i}}(Y|X_i))$ is named the *variance of conditional expectation* with $\mathbf{X}_{\sim i}$ denoting the matrix of all factors but X_i. $V_{X_i}(E_{\mathbf{X}_{\sim i}}(Y|X_i))$ measures the first order effect of X_i on the model output.

Since first order sensitivity indices measure only the decoupled influence of each variable an extension for higher order coupling terms is necessary. For this purpose total effect sensitivity indices have been introduced (Homma and Saltelli1996)

$$S_{Ti} = 1 - \frac{V_{\mathbf{X}_{\sim i}}(E_{X_i}(Y|\mathbf{X}_{\sim i}))}{V(Y)}, \tag{3}$$

where $V_{\mathbf{X}_{\sim i}}(E_{X_i}(Y|\mathbf{X}_{\sim i}))$ measures the first order effect of $\mathbf{X}_{\sim i}$ on the model output which does not contain any effect corresponding to X_i.

2.2. SAMPLING BASED ESTIMATES

One of the most simple procedures to estimate the first order indices is to sample a set of n samples of the input parameter set $\mathbf{X} = [X_1, X_2, \ldots, X_m]$ according to their joint probability distribution and compute for each sample $\mathbf{x}_j$ the model output y_j. A scatter plot is obtained if the model output values are plotted against the values of a single variable as shown in Figure 1. Then the sample values are sorted according to a single

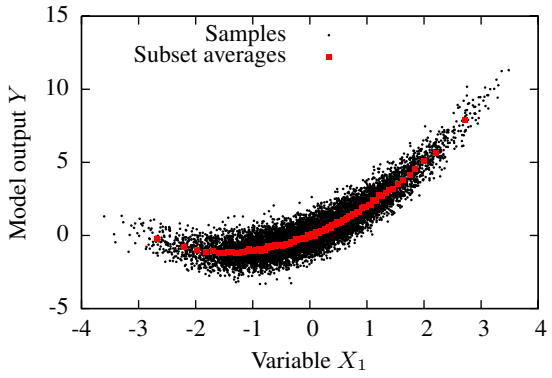


Figure 1. Scatter plot of a single input variable with the model output including subset averages of sorted sample subsets

input variable X_i and subdivided in a number of subsets. For each subset the average value is computed which is equivalent to $E_{\mathbf{X}_{\sim i}}(Y|X_i)$. By calculating the variance of the subset averages, an estimate of $V_{X_i}(E_{\mathbf{X}_{\sim i}}(Y|X_i))$ is obtained, which can directly be used to estimate the first order sensitivity index. This procedure has the advantage that it gives suitable estimates for independent and dependent input parameters. However, an extension for the total sensitivity indices does not exist so far since a sorting concerning $\mathbf{X}_{\sim i}$ seems not possible.

In order to compute the first order and total sensitivity indices using sampling methods, a matrix combination approach is very common in sensitivity analysis. In this procedure two independent sampling matrices $\mathbf{A}$ and $\mathbf{B}$ are generated according to the joint probability density function of the input parameters and recombined in a matrix $\mathbf{C}_i$, where $\mathbf{C}_i$ contains the entries of matrix $\mathbf{B}$ except the i-th column which is taken from $\mathbf{A}$. The estimates of the sensitivity indices can be obtained following (Saltelli et al.2008) as

$$\hat{S}_i = \frac{\mathbf{y_A}^T \mathbf{y_{C_i}} - n(\bar{y}_\mathbf{A})^2}{\mathbf{y_A}^T \mathbf{y_A} - n(\bar{y}_\mathbf{A})^2}, \quad \hat{S}_{T_i} = 1 - \frac{\mathbf{y_B}^T \mathbf{y_{C_i}} - n(\bar{y}_\mathbf{B})^2}{\mathbf{y_B}^T \mathbf{y_B} - n(\bar{y}_\mathbf{B})^2}, \tag{4}$$

where $\mathbf{y_A}$, $\mathbf{y_B}$ and $\mathbf{y_{C_i}}$ are vectors containing the model outputs of the sampling matrices and $\bar{y}_\mathbf{A}$ and $\bar{y}_\mathbf{B}$ are the corresponding mean value estimates. Instead of the estimators in Eq. (4) other approaches exist as discussed in (Saltelli et al.2010). However all of these methods are based on the generation of the sampling matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}_i$. Due to the combination of independent columns of $\mathbf{A}$ and $\mathbf{B}$ the matrices $\mathbf{C}_i$ follow the joint probability density function of $\mathbf{A}$ and $\mathbf{B}$ only if the input variables are independent. In case of dependent parameters X_k and X_l the k-th and l-th columns of the matrices $\mathbf{C}_k$ and $\mathbf{C}_l$ are independent due to the independence of $\mathbf{A}$ and $\mathbf{B}$. Correlations between X_i and the remaining variables are lost in $\mathbf{C}_i$, see the first two plots in Figure 2. For the correlated case only a grouping of the dependent parameters according to

(Jacques et al.2006) would lead to equivalent probability density functions, but this would limit the method to special cases where the grouping is possible.

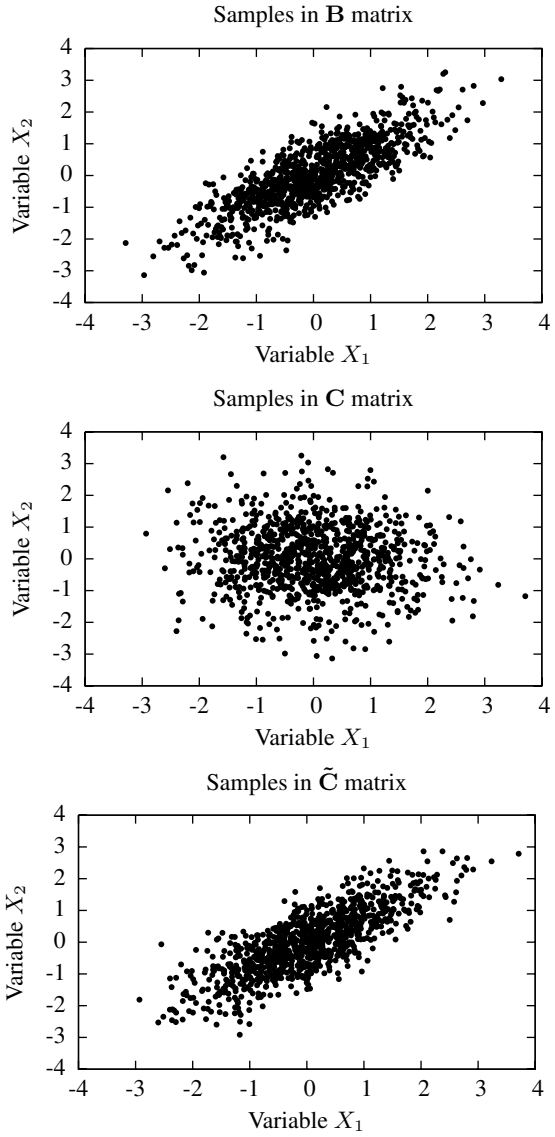


Figure 2. Original and modified joint probability distributions of two correlated variables using the original and the extended matrix combination approach

2.3. REGRESSION BASED ESTIMATES

In order to decrease the numerical effort of a probabilistic analysis often surrogate models are used to approximate the model output instead of evaluating the real sophisticated model. For this purpose, the model

output is replaced mainly by a continuous function, which can be evaluated quite fast compared to a real model call. A very common method is polynomial regression, where the model response Y is generally approximated by a polynomial basis function

$$\mathbf{p}_{\mathbf{X}}^{T}(\mathbf{x}) = \left[1\ x_1\ x_2\ x_3\ \ldots\ x_1^2\ x_2^2\ x_3^2\ \ldots\ x_1x_2\ x_1x_3\ \ldots\ x_2x_3\ \ldots\right] \tag{5}$$

of linear or quadratic order with or without linear coupling terms. The model output y_j for a given sample $\mathbf{x}_j$ of the input parameters $\mathbf{X}$ can be formulated as the sum of the approximated value $\hat{y}_j$ and an error term ϵ_j

$$y(\mathbf{x}_j) = \hat{y}_j(\mathbf{x}_j) + \epsilon_j = \mathbf{p}_{\mathbf{X}}^{T}(\mathbf{x}_j)\boldsymbol{\beta}_Y + \epsilon_j, \tag{6}$$

where $\boldsymbol{\beta}_Y$ is a vector containing the unknown regression coefficients. These coefficients are generally estimated from a given set of sampled support points by assuming independent errors with equal variance at each point. By using a matrix notation the resulting least squares solution reads (Myers and Montgomery2002)

$$\hat{\boldsymbol{\beta}}_Y = (\mathbf{P}_{\mathbf{X}}^{T}\mathbf{P}_{\mathbf{X}})^{-1}\mathbf{P}_{\mathbf{X}}^{T}\mathbf{y}, \tag{7}$$

where $\mathbf{P}_{\mathbf{X}}$ is a matrix containing the basis polynomials of the support point samples.

In order to verify the approximation model the coefficient of determination is often used

$$R^2 = 1 - \frac{\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}{\sum_{j=1}^{n}(y_j - \bar{y})^2}, \quad \bar{y} = \sum_{j=1}^{n} y_j. \tag{8}$$

Generally, R^2 is interpreted as the fraction of the variance of the true model represented by the approximation model. This can be used to estimate total sensitivity indices based on the multivariate regression model (Bucher2009)

$$\hat{S}_{T_i}^{R} = R_{\mathbf{X}}^2 - R_{\mathbf{X}_{\sim i}}^2, \tag{9}$$

where $R_{\mathbf{X}}^2$ is obtained using the full parameter set to build up the regression model and $R_{\mathbf{X}_{\sim i}}^2$ is estimated for a regression model with a reduced parameter set $\mathbf{X}_{\sim i}$.

First order sensitivity indices can be estimated by using only these polynomial basis terms which belong to the investigated variable X_i. The resulting one-dimensional coefficient of determination is a direct estimate of first order indices

$$\hat{S}_i^{R} = R_{X_i}^2. \tag{10}$$

3. Correlated input variables

3.1. Representation of non-normal joint distribution functions

In probabilistic models the dependency between input variables is often modeled by a linear correlation model based on the Pearson correlation coefficient representing pairwise linear correlations

$$\rho(X_i, X_j) = \frac{E[(X_i - \bar{X}_i)(X_j - \bar{X}_j)]}{\sqrt{V(X_i)V(X_j)}}, \tag{11}$$

with the mean values $\bar{X}_i$ and $\bar{X}_j$.

In our study the Nataf model (Nataf1962) is used to generate multivariate joint distribution function of non-normally distributed random variables. In this model the marginal distributions of the individual random variables are transformed to standard normal densities

$$Z_i = \mathcal{T}_{X_i}(X_i), \tag{12}$$

and the resulting multivariate distribution is assumed to be jointly normal

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{\sqrt{(2\pi)^m |\mathbf{C_{ZZ}}|}} \exp\left[-\frac{1}{2}\mathbf{z}^T \mathbf{C}_{\mathbf{ZZ}}^{-1}\mathbf{z}\right], \tag{13}$$

where m is the number and $\mathbf{C_{ZZ}}$ is the covariance matrix of the random variables which contains in case of standard normal variables the correlation coefficients $C_{ij} = \rho(Z_i, Z_j)$. If the correlation coefficients are given in the standard normal space, the correlation coefficients in the original space can be obtained as follows

$$\rho(X_i, X_j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{x_i - \bar{X}_i}{\sigma_{X_i}} \frac{x_j - \bar{X}_j}{\sigma_{X_j}} f_{Z_i Z_j}(z_i, z_j) dz_i dz_j, \tag{14}$$

where $f_{Z_i Z_j}(z_i, z_j)$ is a two-dimensional standard normal probability density function. Often the correlation coefficients are given in the original space. In this case the correlation coefficients in the standard normal space can be derived by solving Eq. (14) iteratively.

3.2. Decomposition of input variables

If input variables are correlated with each other, a single variable can be represented as the sum of a function of the remaining inputs (correlated) and an independent (uncorrelated) part

$$X_i = f(X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots, X_m) + X_i^{U,\mathbf{X}_{\sim i}} = X_i^{C,\mathbf{X}_{\sim i}} + X_i^{U,\mathbf{X}_{\sim i}}. \tag{15}$$

By using a linear correlation model, the correlated part of each input variable can be represented by a linear combination of the remaining variable set $\mathbf{X}_{\sim i}$

$$X_i^{C,\mathbf{X}_{\sim i}} = \beta_{X_i,0} + \sum_{j=1, j\neq i}^{m} \beta_{X_i,j} X_j. \tag{16}$$

In case of standard normal and linearly correlated random variables Z_i, the coefficients

$$Z_i^{C,\mathbf{Z}_{\sim i}} = \beta_{Z_i,0} + \sum_{j=1, j\neq i}^{m} \beta_{Z_i,j} Z_j, \tag{17}$$

can be directly determined from the correlation coefficients

$$\beta_{Z_i,0} = 0, \quad \boldsymbol{\beta}_{Z_i} = \mathbf{C}_{\mathbf{Z}_{\sim i}\mathbf{Z}_{\sim i}}^{-1} \boldsymbol{\rho}_{Z_i}, \tag{18}$$

where $\boldsymbol{\beta}_{Z_i} = [\beta_{Z_i,1}, \ldots, \beta_{Z_i,i-1}, \beta_{Z_i,i+1}, \ldots, \beta_{Z_i,m}]$ contains the coefficients belonging to $\mathbf{Z}_{\sim i}$, $\mathbf{C}_{\mathbf{Z}_{\sim i}\mathbf{Z}_{\sim i}}$ is the correlations matrix of $\mathbf{Z}_{\sim i}$ and $\boldsymbol{\rho}_{Z_i} = [\rho(Z_i, Z_1), \ldots, \rho(Z_i, Z_{i-1}), \rho(Z_i, Z_{i+1}), \ldots, \rho(Z_i, Z_m)]$ contains

the correlation coefficients of $\mathbf{Z}_{\sim i}$ with respect to Z_i. A derivation of Eq. (18), which is valid only in the standard normal space, can be found in the appendix.

For a given set of discrete samples of the standard normal parameter set $\mathbf{Z}$ arranged in matrix $\mathbf{M}$, the uncorrelated parts of the jth column of $\mathbf{M}$ with respect $\mathbf{Z}_{\sim i}$ can be determined as follows

$$\mathbf{M}_{(j)}^{U,\mathbf{Z}_{\sim i}} = \mathbf{M}_{(j)} - \mathbf{M}_{(j)}^{C,\mathbf{Z}_{\sim i}} \tag{19}$$

with

$$\mathbf{M}_{(j=i)}^{C,\mathbf{Z}_{\sim i}} = \sum_{k=1,k\neq i}^{m} \beta_{Z_i,k}\mathbf{M}_{(k)}, \quad \mathbf{M}_{(j\neq i)}^{C,\mathbf{Z}_{\sim i}} = \mathbf{M}_{(j)}. \tag{20}$$

For later use the correlated and uncorrelated parts of a single random variable with respect to only one other variable is introduced. The correlated part of X_j with respect to X_i can be formulated as follows

$$X_j^{C,X_i} = \beta_{X_j,0}^{X_i} + \beta_{X_j,1}^{X_i} X_i. \tag{21}$$

In the standard normal space we obtain

$$Z_j^{C,Z_i} = \beta_{Z_j,0}^{Z_i} + \beta_{Z_j,1}^{Z_i} Z_i, \quad \beta_{Z_j,0}^{Z_i} = 0, \quad \beta_{Z_j,i}^{Z_i} = \rho(Z_i, Z_j). \tag{22}$$

For a given sampling matrix $\mathbf{M}$ the decomposition reads

$$\mathbf{M}_{(j)}^{U,Z_i} = \mathbf{M}_{(j)} - \mathbf{M}_{(j)}^{C,Z_i} = \mathbf{M}_{(j)} - \rho(Z_i, Z_j)\mathbf{M}_{(i)}. \tag{23}$$

In the second decomposition the corresponding ith column of $\mathbf{M}^{C,Z_i}$ is equal to $\mathbf{M}_{(i)}$ and the values of the ith column of $\mathbf{M}^{U,X_i}$ vanish.

As discussed in (Saltelli et al.2004) the conditional variance $V_{\mathbf{X}_{\sim i}}(E_{X_i}(Y|\mathbf{X}_{\sim i}))$ may decrease by introducing dependence between X_i and $\mathbf{X}_{\sim i}$. This may result in total effect indices smaller than first order indices. Furthermore, if we use a linear correlation model and increase the absolute value of the correlation coefficient between two input variables close to one, the total effect indices of both variables approach to zero. For factor fixing this relation requires a step-wise fixing of a single factor and a revaluation of the indices of the remaining variables. If only the sensitivity indices of the full model are used for factor fixing, this may result in a mistaken removing of important variables. One possibility to overcome this, is to evaluate the sensitivity indices on one side for the uncorrelated part of a single variable X_i^U and on the other side for this variable including all correlated portions of the other variables $X_j^{C,X_i}, j \neq i$. This concept was introduced for first order indices of linear models in (Xu and Gertner2008). In our study we introduce first order indices S_i^C and S_i^U as well as total effect indices $S_{T_i}^C$ and $S_{T_i}^U$ for the correlated and uncorrelated parts of a single variable X_i. Per definition S_i^C is equivalent to the original first order index and $S_{T_i}^U$ to the original total effect index. The additional measures S_i^U and especially $S_{T_i}^C$ should give us additional information for factor fixing without requiring a step-wise revaluation of the sensitivity indices.

4. Estimation of sensitivity indices for correlated input variables

4.1. Extension of the matrix combination approach

In the original matrix combination approach presented in section 2.2 it is intended to modify a single input variable X_i by a second random sampling set while keeping the other variables unchanged in order to add or to remove the influence of X_i from the model.

In order to calculate the correlated first order and total effect indices, not only the samples of X_i itself, but also all correlated parts of the remaining variables with respect to X_i have to be modified. By assuming a linear correlation model in the standard normal space as presented in section 3.2, a decomposition of the matrices $\mathbf{A}$ and $\mathbf{B}$ in an uncorrelated and correlated part with respect to X_i is performed. For this purpose both matrices are transformed to the correlated standard normal space by the marginal transformations

$$\mathcal{A}_{kl} = \Phi[F_{X_l}^{-1}(A_{kl})], \quad \mathcal{B}_{kl} = \Phi[F_{X_l}^{-1}(B_{kl})], \tag{24}$$

where $\Phi(\cdot)$ is cumulative distribution function of a standard normal random variable and $F_{X_l}^{-1}(\cdot)$ is the inverse cumulative distribution function of X_l. By using the decomposition proposed in Eq. (23) we obtain the columns of the correlated and uncorrelated sampling matrices as follows

$$\boldsymbol{\mathcal{A}}_{(j)}^{C,Z_i} = \rho(Z_i,Z_j)\boldsymbol{\mathcal{A}}_{(i)}, \quad \boldsymbol{\mathcal{A}}_{(j)}^{U,Z_i} = \boldsymbol{\mathcal{A}}_{(j)} - \boldsymbol{\mathcal{A}}_{(j)}^{C,Z_i}, \tag{25}$$

$$\boldsymbol{\mathcal{B}}_{(j)}^{C,Z_i} = \rho(Z_i,Z_j)\boldsymbol{\mathcal{B}}_{(i)}, \quad \boldsymbol{\mathcal{B}}_{(j)}^{U,Z_i} = \boldsymbol{\mathcal{B}}_{(j)} - \boldsymbol{\mathcal{B}}_{(j)}^{C,Z_i}. \tag{26}$$

Now a modified matrix $\tilde{\boldsymbol{\mathcal{C}}}_i^C$ can be obtained, which contains the uncorrelated part of $\boldsymbol{\mathcal{B}}$ with respect to Z_i and the correlated part of $\boldsymbol{\mathcal{A}}$

$$\tilde{\boldsymbol{\mathcal{C}}}_i^C = \boldsymbol{\mathcal{C}}_i^{C,Z_i} = \boldsymbol{\mathcal{B}}^{U,Z_i} + \boldsymbol{\mathcal{A}}^{C,Z_i}. \tag{27}$$

Finally the entries of matrix $\tilde{\boldsymbol{\mathcal{C}}}_i^C$ are transformed to the original space by the individual marginal transformations

$$\tilde{C}_{i,kl}^C = F_{X_l}[\Phi^{-1}(\tilde{\mathcal{C}}_{i,kl}^C)]. \tag{28}$$

Since the correlated part with respect to X_i is modified simultaneously for all variables, $\tilde{\mathbf{C}}_i^C$ follows the original joint probability distribution of $\mathbf{X}$ as shown in Figure 2.

In order to obtain the uncorrelated first order and total effect indices we use the decomposition with respect to $\mathbf{X}_{\sim i}$. For this purpose the matrices $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ are decomposed following Eq. (19)

$$\boldsymbol{\mathcal{A}}_{(j)}^{U,\mathbf{Z}_{\sim i}} = \boldsymbol{\mathcal{A}}_{(j)} - \boldsymbol{\mathcal{A}}_{(j)}^{C,\mathbf{Z}_{\sim i}}, \quad \boldsymbol{\mathcal{A}}_{(j=i)}^{C,\mathbf{Z}_{\sim i}} = \sum_{k=1,k\neq i}^{m} \beta_{Z_i,k}\boldsymbol{\mathcal{A}}_{(k)}, \quad \boldsymbol{\mathcal{A}}_{(j\neq i)}^{C,\mathbf{Z}_{\sim i}} = \boldsymbol{\mathcal{A}}_{(j)}, \tag{29}$$

$$\boldsymbol{\mathcal{B}}_{(j)}^{U,\mathbf{Z}_{\sim i}} = \boldsymbol{\mathcal{B}}_{(j)} - \boldsymbol{\mathcal{B}}_{(j)}^{C,\mathbf{Z}_{\sim i}}, \quad \boldsymbol{\mathcal{B}}_{(j=i)}^{C,\mathbf{Z}_{\sim i}} = \sum_{k=1,k\neq i}^{m} \beta_{Z_i,k}\boldsymbol{\mathcal{B}}_{(k)}, \quad \boldsymbol{\mathcal{B}}_{(j\neq i)}^{C,\mathbf{Z}_{\sim i}} = \boldsymbol{\mathcal{B}}_{(j)}. \tag{30}$$

The resulting $\tilde{\boldsymbol{\mathcal{C}}}_i^U$ contains the uncorrelated part of Z_i from matrix $\boldsymbol{\mathcal{A}}$ and the correlated part with respect to $\mathbf{Z}_{\sim i}$ from matrix $\boldsymbol{\mathcal{B}}$

$$\tilde{\boldsymbol{\mathcal{C}}}_i^U = \boldsymbol{\mathcal{C}}_i^{U,\mathbf{Z}_{\sim i}} = \boldsymbol{\mathcal{A}}^{U,\mathbf{Z}_{\sim i}} + \boldsymbol{\mathcal{B}}^{C,\mathbf{Z}_{\sim i}}. \tag{31}$$

Analogous to the correlated part, $\tilde{\mathbf{C}}_i^U$ is obtained by transforming the entries of $\tilde{\boldsymbol{\mathcal{C}}}_i^U$ back to the origin space.

4.2. EXTENSION FOR REGRESSION BASED INDICES

The regression based indices presented in section 2.3 can be directly applied for correlated input variables. However, the one-dimensional coefficient of determination is an estimate of the first order index of a single variable including all correlated parts of the other variables. Using a matrix $\boldsymbol{\mathcal{Z}}$ containing the support points of the regression in the standard normal space, the first order index reads

$$\hat{S}_i^{R,C} = R^2_{\boldsymbol{\mathcal{Z}},Z_i}. \tag{32}$$

The estimate for the total effect index given in Eq. (9) quantifies the total variance contribution of the uncorrelated part Z_i^U of variable Z_i

$$\hat{S}_{T_i}^{R,U} = R^2_{\boldsymbol{\mathcal{Z}},\mathbf{Z}} - R^2_{\boldsymbol{\mathcal{Z}},\mathbf{Z}_{\sim i}}. \tag{33}$$

The correlated part of Z_i can be represented by the other variables and thus the explained variation of the reduced model is decreased by the contribution of Z_i^U only. The total effect of the variable Z_i including all correlated parts of the other variables can be estimated analogously to the matrix combination approach: The sampling matrix $\boldsymbol{\mathcal{Z}}$ is decomposed in a correlated and uncorrelated part with respect to Z_i according to Eq. (31). Then the uncorrelated part of the samples $\boldsymbol{\mathcal{Z}}^{U,Z_i}$ is used for the reduced model within the estimate of the total effect index

$$\hat{S}_{T_i}^{R,C} = R^2_{\boldsymbol{\mathcal{Z}},\mathbf{Z}} - R^2_{\boldsymbol{\mathcal{Z}}^{U,Z_i},\mathbf{Z}_{\sim i}}. \tag{34}$$

In order to estimate the first order index of the uncorrelated part of the variable Z_i with respect to all other variables, Eq. (19) is used to calculate the uncorrelated part of ith column of the sample matrix $\boldsymbol{\mathcal{Z}}$ and the corresponding first order sensitivity index is obtained

$$\hat{S}_i^{R,U} = R^2_{\boldsymbol{\mathcal{Z}}^{U,\mathbf{Z}_{\sim i}},Z_i}. \tag{35}$$

With the four measures $\hat{S}_i^{R,U}$, $\hat{S}_i^{R,C}$, $\hat{S}_{T_i}^{R,U}$ and $\hat{S}_{T_i}^{R,C}$ regression based sensitivity indices are introduced, which give similar results as the model independent measures proposed in section 4.1, but only if the regression model can represent the underlying investigated model. If this is not the case, e.g. if complex nonlinearities and interactions describe the model response Y in terms of the input variables $\mathbf{X}$, the polynomial based measures quantify only the contribution which is represented by the basis function. The total contribution of higher order and unexplained dependencies, which are not represented by the regression model, can be estimated using the full regression basis

$$\hat{S}_{T_i}^{Unexplained} = 1 - R^2_{\boldsymbol{\mathcal{Z}},\mathbf{Z}}. \tag{36}$$

Since the introduced sensitivity measures investigate the influence of a single variable including all correlated parts of the other variables, it can not be distinguished, if the influence of this variable is caused by its contribution within the model or if it is caused by its correlation with other important variables. This fact shall be clarified by a simple example: A purely additive model with three inputs is given

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3. \tag{37}$$

The input variables are normally distributed, have zero mean and the covariance

$$\Gamma = \begin{bmatrix} \sigma_{X_1}^2 & 0 & 0 \\ 0 & \sigma_{X_2}^2 & \rho\sigma_{X_2}\sigma_{X_3} \\ 0 & \rho\sigma_{X_2}\sigma_{X_3} & \sigma_{X_3}^2 \end{bmatrix}, \tag{38}$$

where ρ is the linear correlation between X_2 and X_3.

The first order and total effect indices of this problem can be derived analytically as follows: X_3 can be formulated with respect to X_2 and vice versa as

$$\begin{aligned} X_3 &= \rho\frac{\sigma_{X_3}}{\sigma_{X_2}}X_2 + \sqrt{1-\rho^2}X_3^U, \\ X_2 &= \rho\frac{\sigma_{X_2}}{\sigma_{X_3}}X_3 + \sqrt{1-\rho^2}X_2^U, \end{aligned} \tag{39}$$

with X_2^U and X_3^U having the same variances as X_2 and X_3, respectively. Since X_2 and X_3^U as well as X_3 and X_2^U are independent, we can calculate the variance contribution of the inputs as follows

$$\begin{aligned} V_{X_1^C}(Y) &= V_{X_1^U}(Y) = \beta_1^2\sigma_{X_1}^2, \\ V_{X_2^C}(Y) &= (\beta_2\sigma_{X_2} + \rho\beta_3\sigma_{X_3})^2, \\ V_{X_2^U}(Y) &= (1-\rho^2)\beta_2^2\sigma_{X_2}^2, \\ V_{X_3^C}(Y) &= (\beta_3\sigma_{X_3} + \rho\beta_2\sigma_{X_2})^2, \\ V_{X_3^U}(Y) &= (1-\rho^2)\beta_3^2\sigma_{X_3}^2. \end{aligned} \tag{40}$$

With help of the total variance

$$V(Y) = V_{X_1}(Y) + V_{X_2^C}(Y) + V_{X_3^U}(Y) = V_{X_1}(Y) + V_{X_2^U}(Y) + V_{X_3^C}(Y), \tag{41}$$

the sensitivity indices can be determined

$$S_i^C = S_{T_i}^C = \frac{V_{X_i}^C(Y)}{V(Y)}, \quad S_i^U = S_{T_i}^U = \frac{V_{X_i}^U(Y)}{V(Y)}. \tag{42}$$

From Eq. (40) it follows, that the variance contribution of the uncorrelated parts of X_2 and X_3 vanish if the correlation coefficient tends to one or minus one. On the other hand it can be seen, that if one of the variable factors β_2 or β_3 is zero, the corresponding variance contribution and thus the sensitivity index of the correlated part is not zero, due to the remaining influence of the other variable.

In order to distinguish between the variance contribution due to the correlation and the contribution by the model parameters, an effective variance contribution is defined

$$V_{X_i}^{eff}(Y) = \beta_i^2\sigma_{X_i}^2. \tag{43}$$

For the general polynomial regression case, the following procedure is proposed: First the full regression model is build up in the standard normal space and the regression coefficients $\hat{\boldsymbol{\beta}}$ are determined. Then the

first order effects are estimated by using only the linear and quadratic coefficients of the full model, which belong to the investigated variable Z_i,

$$\hat{S}_i^{\boldsymbol{\beta}} = \frac{V_{\boldsymbol{\beta}_{Z_i}}(Y)}{V(Y)}, \tag{44}$$

where $V_{\boldsymbol{\beta}_{Z_i}}(Y)$ is the resulting variance if only the first order regression terms of Z_i are considered. Total effects can be estimated similarly

$$\hat{S}_{T_i}^{\boldsymbol{\beta}} = 1 - \frac{V_{\boldsymbol{\beta}_{\mathbf{Z}_{\sim i}}}(Y)}{V(Y)}, \tag{45}$$

where $V_{\boldsymbol{\beta}_{\mathbf{Z}_{\sim i}}}(Y)$ is estimated by using the regression model without all coefficients, first order and interaction, which belong to Z_i. The definition of $S_i^{\boldsymbol{\beta}}$ and $S_{T_i}^{\boldsymbol{\beta}}$ is not following the general definition of the first order and total effect sensitivity indices, where the variance of conditional expectation is used. Nevertheless, this measures may be used to give additional information about the unknown underlying model.

However, for strongly correlated random variables the accuracy of the estimated regression coefficients may be very poor. In such a case the estimated sensitivity measures $\hat{S}_i^{\boldsymbol{\beta}}$ and $\hat{S}_{T_i}^{\boldsymbol{\beta}}$ may have very low accuracy. In the first numerical example, this problem is investigated.

5. Numerical examples

5.1. Additive linear model

In the first example a purely additive model is investigated

$$Y = X_1 + X_2 + X_3. \tag{46}$$

The input variables are normally distributed, have zero mean and the covariance

$$\Gamma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2\rho \\ 0 & 2\rho & 4 \end{bmatrix}. \tag{47}$$

Different cases with $\rho = 0.0, 0.8, 0.99999, -0.8, -0.5$ are investigated with the sampling based approach and the regression based approach. The sampling based sensitivity indices are calculated using the subset averaging and the proposed matrix recombination method with 10.000 Latin Hypercube samples. In Table I the estimated sensitivity indices are compared to the analytical solution obtained with Eq. (40). The table indicates a very good agreement of all estimated indices with the reference values for the uncorrelated and the correlated cases. In the uncorrelated case, the sensitivity indices of an additive model sum up to one. By considering correlated inputs this is not the case, since coupling terms caused by the correlations are introduced indirectly into the model. This can result in a sum larger than one (correlated formulation with $\rho = 0.8$) and also smaller than one (correlated formulation with $\rho = -0.5$, uncorrelated formulation with $\rho = 0.8$). Interestingly the influence of a variable seems to vanish looking only on the correlated or uncorrelated formulation of the total effect indices which is observed for variable X_2 in the correlated formulation with $\rho = -0.5$ and X_2 and X_3 in the uncorrelated formulation with $\rho = 1.0$. However, if both

Table I. Sensitivity indices of the additive model assuming uncorrelated and correlated inputs

	Analytical		Subset	Matrix combination				Regression					
	$S_i^C, S_{T_i}^C$	$S_i^U, S_{T_i}^U$	$\hat{S}_i$	$\tilde{S}_i^C$	$\tilde{S}_{T_i}^C$	$\tilde{S}_i^U$	$\tilde{S}_{T_i}^U$	$\hat{S}_i^{R,C}$	$\hat{S}_{T_i}^{R,C}$	$\hat{S}_i^{R,U}$	$\hat{S}_{T_i}^{R,U}$	$\hat{S}_i^{\boldsymbol{\beta}}$	$\hat{S}_{T_i}^{\boldsymbol{\beta}}$
$\rho = 0.0$													
X_1	0.167	0.167	0.177	0.157	0.168	0.157	0.166	0.155	0.155	0.168	0.168	0.168	0.155
X_2	0.167	0.167	0.177	0.154	0.164	0.153	0.163	0.168	0.168	0.169	0.168	0.168	0.168
X_3	0.667	0.667	0.674	0.667	0.673	0.669	0.672	0.670	0.670	0.672	0.672	0.672	0.670
$\rho = 0.8$													
X_1	0.109	0.109	0.118	0.117	0.108	0.118	0.107	0.110	0.110	0.109	0.108	0.109	0.110
X_2	0.735	0.039	0.742	0.738	0.727	0.050	0.039	0.735	0.735	0.040	0.039	0.109	0.456
X_3	0.852	0.157	0.861	0.858	0.848	0.169	0.157	0.853	0.853	0.156	0.156	0.434	0.783
$\rho = 0.99999$													
X_1	0.100	0.100	0.110	0.088	0.103	0.088	0.101	0.099	0.099	0.101	0.100	0.100	0.099
X_2	0.900	0.000	0.907	0.900	0.916	0.000	0.000	0.900	0.900	0.001	0.000	0.100	0.500
X_3	0.900	0.000	0.907	0.900	0.916	0.000	0.000	0.900	0.900	0.001	0.000	0.401	0.800
$\rho = -0.8$													
X_1	0.357	0.357	0.368	0.361	0.359	0.359	0.359	0.362	0.362	0.354	0.353	0.353	0.362
X_2	0.129	0.129	0.137	0.141	0.130	0.136	0.139	0.134	0.134	0.127	0.127	0.353	-0.788
X_3	0.514	0.514	0.523	0.527	0.513	0.520	0.512	0.521	0.521	0.509	0.509	1.420	0.300
$\rho = -0.5$													
X_1	0.250	0.250	0.256	0.240	0.250	0.243	0.249	0.241	0.241	0.254	0.253	0.253	0.241
X_2	0.000	0.188	0.011	0.005	0.000	0.179	0.188	0.000	0.000	0.189	0.188	0.253	-0.266
X_3	0.563	0.750	0.573	0.571	0.570	0.755	0.752	0.564	0.564	0.759	0.759	1.023	0.495

the correlated and uncorrelated formulation of the total effect indices are considered, no important variable can be misinterpreted or mistakenly removed from the model. This is not the case if either only the correlated or the uncorrelated formulation is used.

Additionally the regression based indices by using 1000 Latin Hypercube samples as support points are given in Table I. The estimates $\hat{S}_i^{R,U}$, $\hat{S}_i^{R,C}$, $\hat{S}_{T_i}^{R,U}$ and $\hat{S}_{T_i}^{R,C}$ agree very well with the analytical values. The introduced measures $\hat{S}_i^{\boldsymbol{\beta}}$ and $\hat{S}_{T_i}^{\boldsymbol{\beta}}$, which analyze the variance contribution by coefficient removing, are equivalent for the uncorrelated case. In the correlated case, the results differ in that way, that these indices indicate how the total variance is decreased or eventually increased, if the variable X_i is removed completely from the model formulation. For this reason negative indices or values larger as one are possible in contrast to classical sensitivity indices. This measures should be considered as additional information in order to detect imaginary variable importance caused only by input correlations.

However, the accuracy of the coefficient based indices $\hat{S}_i^{\boldsymbol{\beta}}$ and $\hat{S}_{T_i}^{\boldsymbol{\beta}}$ may be very low for highly correlated variables, if the regression model can not represent the true model perfectly. This situation is analyzed by adding to the deterministic model a random noise ϵ which is independent of the inputs

$$Y = X_1 + X_2 + X_3 + \epsilon, \tag{48}$$

where ϵ is a zero-mean normally distributed random variable with standard deviation $\sigma_\epsilon = 0.5$. In Figure 3 the standard deviation of the estimated regression coefficients of the tree inputs are shown for a set of 100 support point samples. The figure indicates an increasing error in the regression coefficients of both correlated variables if the correlation coefficient is increased, while the error in the coefficient of the uncorrelated variable is almost unchanged. This results in an inaccurate estimate of the sensitivity measures $\hat{S}_i^{\boldsymbol{\beta}}$ and $\hat{S}_{T_i}^{\boldsymbol{\beta}}$

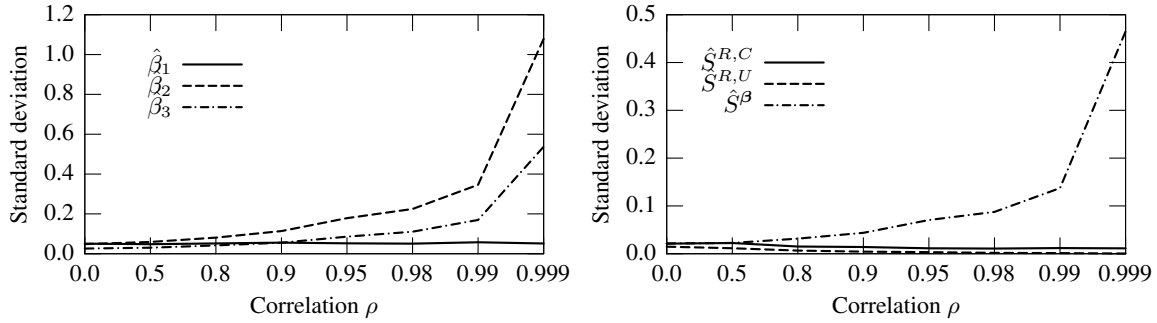


Figure 3. Standard deviation of the regression coefficients (left) and the estimated total effect sensitivity indices of X_2 (right) dependent on the correlation for the additive model including random noise

for highly correlated variables. In Figure 3 the standard deviation of the estimated index $\hat{S}_{T_2}^{\boldsymbol{\beta}}$ is displayed additionally. Similar to the regression coefficient itself, the accuracy of the sensitivity index decreases with increasing correlation. Nevertheless, if the variables are correlated with $\rho \leq 0.9$ the error in the estimate seems acceptable. In contrast to the measure based on single regression coefficients, the accuracy of the sensitivity indices estimated as the difference of the coefficient of determination of the full and a reduced regression model is not decreased with increasing correlation. This is plotted additionally in Figure 3. Using only 100 samples, the accuracy of the estimated regression based measures seems very good.

5.2. COUPLED NONLINEAR MODEL

In the second example a nonlinear model is investigated, which contains linear, quadratic and interaction terms

$$Y = X_1 + 2X_1^2 + X_2 + X_3 + X_2X_3. \tag{49}$$

The covariance is taken analogous to the previous example. In Figure 4 the scatter plots are shown for the first two variables. In the figure a significant quadratic influence of X_1 and a significant influence of the coupling term X_2X_3 can be remarked in the uncorrelated case. If X_2 and X_3 are assumed to be correlated, the scatter plot of X_2 with Y indicates also a quadratic behavior, while in the model formulation only a linear and a coupling term appears. This quadratic influence can be explained, if we rewrite Eq. (49) by considering the decomposition of X_3 analogous to Eq. (39) as follows

$$Y = X_1 + 2X_1^2 + X_2 + \rho\frac{\sigma_{X_3}}{\sigma_{X_2}}X_2 + \sqrt{1-\rho^2}X_3^U + \rho\frac{\sigma_{X_3}}{\sigma_{X_2}}X_2^2 + \sqrt{1-\rho^2}X_2X_3^U. \tag{50}$$

In Table II the corresponding sensitivity indices are given. The table indicates a good agreement of the first order indices estimated with the extended matrix combination approach and the regression based method with the indices obtained with the subset averaging. The effect seen in the scatter plots, that the influence of X_2 dominated in the uncorrelated case by the coupling term and in the correlated case by quadratic dependencies, can be observed also in the sensitivity indices. In the uncorrelated case the difference between first order and total effect indices of X_2 is more than 20%, but in the correlated case this difference decreases significantly. This example clarifies, that the proposed measures quantifying the total effect influence give useful information for model interpretation in the case of correlated inputs.

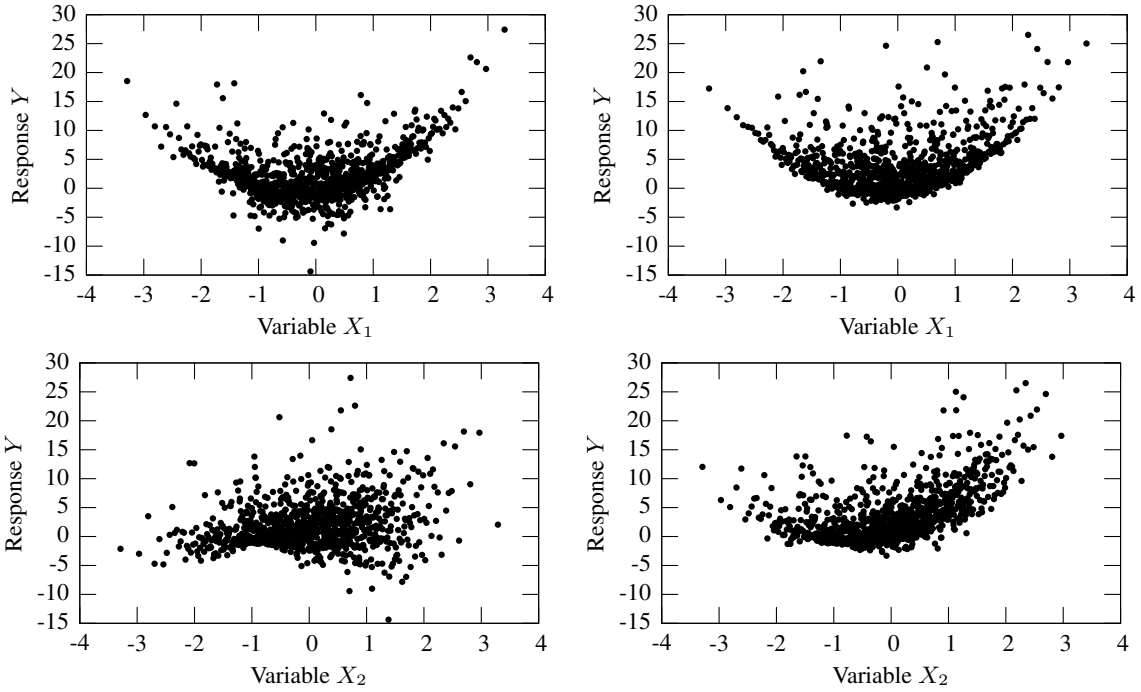


Figure 4. Anthill plots of the response of the coupled nonlinear function with respect to X_1 and X_2 by assuming $\rho = 0.0$ (left) and $\rho = 0.8$ (right)

Table II. Sensitivity indices of the coupled nonlinear model assuming uncorrelated and correlated inputs

	Subset	Matrix combination				Regression					
	$\hat{S}_i$	$\tilde{S}_i^C$	$\tilde{S}_{T_i}^C$	$\tilde{S}_i^U$	$\tilde{S}_{T_i}^U$	$\hat{S}_i^{R,C}$	$\hat{S}_{T_i}^{R,C}$	$\hat{S}_i^{R,U}$	$\hat{S}_{T_i}^{R,U}$	$\hat{S}_i^{\boldsymbol{\beta}}$	$\hat{S}_{T_i}^{\boldsymbol{\beta}}$
$\rho = 0.0$											
X_1	0.484	0.489	0.518	0.490	0.518	0.491	0.511	0.491	0.510	0.512	0.491
X_2	0.070	0.036	0.256	0.034	0.259	0.060	0.292	0.059	0.291	0.058	0.274
X_3	0.231	0.217	0.452	0.216	0.454	0.204	0.441	0.204	0.440	0.231	0.436
$\rho = 0.8$											
X_1	0.382	0.405	0.362	0.405	0.361	0.373	0.367	0.371	0.365	0.365	0.373
X_2	0.501	0.528	0.540	0.055	0.071	0.506	0.562	0.014	0.072	0.041	0.469
X_3	0.551	0.548	0.592	0.090	0.123	0.565	0.615	0.064	0.120	0.165	0.593

5.3. ISHIGAMI FUNCTION

In the final example the well-known Ishigami function (Ishigami and Homma1990) is investigated. This function was defined for independent uniformly distributed variables

$$-\pi \leq X_i \leq \pi, \quad i = 1, 2, 3, \tag{51}$$

as follows

$$Y = \sin(X_1) + a\sin^2(X_2) + bX_3^4\sin(X_1), \tag{52}$$

with $a = 7.0$ and $b = 0.1$. The analytical first order and total sensitivity indices were derived in (Homma and Saltelli1996) and are given in Table III. The values indicate a vanishing first order influence of X_3 but a significant coupling term of X_1 and X_3 as indicated in the total effect indices.

In our analysis again the subset averaging and the matrix combination approach are applied using 10.000 Latin Hypercube samples. The regression based approach is not used due to the highly nonlinear functional behavior, which can not be represented by low order polynomials. The results of the analyses assuming uncorrelated and even correlated inputs are given additionally in Table III.

The correlation has been assumed in the second case between X_1 and X_3 as $\rho_{23} = 0.5$. In the Nataf model this correlation coefficient is transformed to $\tilde{\rho}_{23} = 0.518$ in the correlated Gaussian space using the iteration in Eq. (14). In the third case additional correlations are introduced ($\rho_{12} = 0.3$, $\rho_{23} = 0.8$, $\tilde{\rho}_{12} = 0.313$, $\tilde{\rho}_{23} = 0.814$). For the uncorrelated case the indices obtained with the different methods agree

Table III. Sensitivity indices of the Ishigami test function assuming uncorrelated and correlated inputs

	Analytical		Subset	Matrix combination			
	S_i	S_{T_i}	$\hat{S}_i$	$\tilde{S}_i^C$	$\tilde{S}_{T_i}^C$	$\tilde{S}_i^U$	$\tilde{S}_{T_i}^U$
$\rho_{23} = 0.0$							
X_1	0.314	0.557	0.324	0.330	0.570	0.330	0.570
X_2	0.442	0.442	0.461	0.456	0.429	0.456	0.429
X_3	0.000	0.244	0.008	0.008	0.251	0.008	0.251
$\rho_{13} = 0.5$							
X_1	-	-	0.305	0.312	0.453	0.065	0.346
X_2	-	-	0.479	0.484	0.473	0.484	0.473
X_3	-	-	0.174	0.172	0.444	0.071	0.200
$\rho_{12} = 0.3, \rho_{13} = 0.5, \rho_{23} = 0.8$							
X_1	-	-	0.310	0.306	0.799	0.057	0.354
X_2	-	-	0.576	0.559	0.785	0.022	0.454
X_3	-	-	0.213	0.198	0.927	0.022	0.117

very well with the theoretical values. Interestingly for the first correlated case, the first order index of X_3 increases significantly due to the correlation.

5.4. CONCLUSIONS

In the presented paper an extension of the classical Sobol' estimator for correlated input variables has been proposed. In this method the matrix recombination is modified by changing not only the variable itself but also its coupling terms with other variables due to the correlations. For this purpose a decomposition of the sampling matrices in an uncorrelated and a correlated part is proposed. This decomposition is based on a linear correlation model between the input variables. Additionally a regression based approach is presented, which is much more efficient, if the model behavior can be represented by the regression basis. However, attention is required in both methods in order to perform the decomposition. For non-normally distributed inputs a transformation to the Gaussian space is necessary before decomposing the samples.

Appendix

In the standard normal space the correlated part of a single variable Z_i can be represented by the remaining variable set $\tilde{\mathbf{Z}} = \mathbf{Z}_{\sim i}$ as follows

$$Z_i^{C,\tilde{\mathbf{Z}}} = \beta_{Z_i,0} + \sum_{j=1}^{m-1} \beta_{Z_i,j} \tilde{Z}_j. \tag{53}$$

The intercept $\beta_{Z_i,0}$ is zero, since all variables are normally distributed with zero mean. An estimate of the coefficients $\boldsymbol{\beta}_{Z_i}$ can be obtained from a discrete sample matrix $\boldsymbol{\mathcal{Z}}$ by using linear regression

$$\hat{\boldsymbol{\beta}}_{Z_i} = (\mathbf{P}_{\tilde{\mathbf{Z}}}^T \mathbf{P}_{\tilde{\mathbf{Z}}})^{-1} \mathbf{P}_{\tilde{\mathbf{Z}}}^T \boldsymbol{\mathcal{Z}}_{(i)}, \tag{54}$$

where $\boldsymbol{\mathcal{Z}}_{(i)}$ is the ith column of $\boldsymbol{\mathcal{Z}}$ and the level matrix $\mathbf{P}_{\tilde{\mathbf{Z}}}$ of the reduced variable set contains all columns of $\boldsymbol{\mathcal{Z}}$ without the ith column. The entries of $\mathbf{P}_{\tilde{\mathbf{Z}}}^T \mathbf{P}_{\tilde{\mathbf{Z}}}$ can be obtained from the reduced sample matrix $\tilde{\boldsymbol{\mathcal{Z}}}$ as follows

$$(\mathbf{P}_{\tilde{\mathbf{Z}}}^T \mathbf{P}_{\tilde{\mathbf{Z}}})_{kl} = \sum_{t=1}^{n} \tilde{\mathcal{Z}}_{tk} \tilde{\mathcal{Z}}_{tl}. \tag{55}$$

In the case of standard normal variables, the sum in Eq. (55) can be formulated in terms of the estimator of the correlation coefficients as follows

$$(\mathbf{P}_{\tilde{\mathbf{Z}}}^T \mathbf{P}_{\tilde{\mathbf{Z}}})_{kl} = (n-1)\hat{\rho}(\tilde{Z}_k, \tilde{Z}_l). \tag{56}$$

Analogously, the entries of $\mathbf{P}_{\tilde{\mathbf{Z}}}^T \boldsymbol{\mathcal{Z}}_{(i)}$ can be formulated as

$$(\mathbf{P}_{\tilde{\mathbf{Z}}}^T \boldsymbol{\mathcal{Z}}_{(i)})_k = \sum_{t=1}^{n} \tilde{\mathcal{Z}}_{tk} \mathcal{Z}_{ti} = (n-1)\hat{\rho}(\tilde{Z}_k, Z_i), \tag{57}$$

which finally results in the estimates of the regression coefficients

$$\hat{\boldsymbol{\beta}}_{Z_i} = (\hat{\mathbf{C}}_{\tilde{\mathbf{Z}}\tilde{\mathbf{Z}}})^{-1} \hat{\boldsymbol{\rho}}_{\tilde{\mathbf{Z}},Z_i}, \tag{58}$$

where $\mathbf{C}_{\tilde{\mathbf{Z}}\tilde{\mathbf{Z}}}$ is the correlation matrix of the reduced variable set $\tilde{\mathbf{Z}}$. In order to generate samples of the full variable set $\mathbf{Z}$ the input correlation values have to be defined. If this is the case, the predefined values instead of the estimators can be used and the regression coefficients $\boldsymbol{\beta}_{Z_i}$ can be exactly determined.

Acknowledgements

The author would like to thank H. Keitel and A. Saltelli for helpful discussions on this topic.

References

Bucher, C. (2009). *Computational Analysis of Randomness in Structural Mechanics*. London: CRC Press, Taylor & Francis Group.

DaVeiga, S., F. Wahl, and F. Gamboa (2009). Local polynomial estimation for sensitivity analysis for models with correlated inputs. *Technometrics 51*, 439–451.

Homma, T. and A. Saltelli (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety 52*, 1–17.

Ishigami, T. and T. Homma (1990). An importance quantification technique in uncertainty analysis. In *Proc. ISUMA '90, 1st Intern. Symp. on Uncertainty Modelling and Analysis, University of Maryland, 3-5 December*, pp. 398–403.

Jacques, J., C. Laverge, and N. Devictor (2006). Sensitivity analysis in presence of model uncertainty and correlated inputs. *Reliability Engineering and System Safety 91*, 1126–1134.

Kucherenko, S., S. Tarantola, and P. Annoni (2012). Estimation of global sensitivity indices for models with dependent variables. *Computer Physics Communications 183*, 937–946.

Mara, T. A. and S. Tarantola (2012). Variance-based sensitivity indices for models with dependent inputs. *Reliability Engineering and System Safety, in press*.

Most, T., A. Saltelli, and M. Ratto (2010). Estimating total effect sensitivity indices in the presence of correlated input variables. Internal report, GRK 1462, Bauhaus-Universität Weimar.

Myers, R. and D. C. Montgomery (2002). *Response Surface Methodology* (2 ed.). John Wiley & Sons, Inc.

Nataf, A. (1962). Détermination des distributions de probabilités dont les marges sont données. *Comptes Rendus de l'Academie des Sciences 225*, 42–43.

Oakley, J. E. and A. O'Hagan (2004). Probabilistic sensitivity analysis of complex models: A Bayesian approach. *Journal of Royal Statistical Society B 66*, 751–769.

Saltelli, A. et al. (2004). *Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models*. John Wiley and Sons publishers.

Saltelli, A. et al. (2008). *Global Sensitivity Analysis. The Primer*. Chichester, England: John Wiley & Sons, Ltd.

Saltelli, A. et al. (2010). Variance based sensitivity analysis of model output: Design and estimator for the total sensitivity index. *Computer Physics Communications 181*, 259–270.

Sobol', I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiment 1*, 407–414.

Xu, C. and G. Z. Gertner (2008). Uncertainty and sensitivity analysis for models with correlated inputs. *Reliability Engineering and System Safety 93*, 1563–1573.

Robust Design Optimization in industrial virtual product development

Thomas Most & Johannes Will
Dynardo – Dynamic Software and Engineering GmbH, Weimar, Germany
thomas.most@dynardo.de, johannes.will@dynardo.de

Abstract. In this paper different strategies to search for a robust design are presented and investigated with respect to their efficiency and applicability to time consuming numerical models. After starting with deterministic optimization we introduce different measures to define the robustness of a design. An iterative Robust Design Optimization (RDO) is proposed where deterministic optimization is combined with variance-based robustness analysis and final reliability proof. The iterative procedure is compared to coupled RDO approaches, where the robustness or reliability measures are calculated for each optimization design. For such a procedure often global approximation models are used in order to enable the application for more complex problems.

Keywords: Robustness, reliability, optimization, sensitivity analysis

1. Introduction

Due to target-oriented, automatic optimization of virtual products new design possibilities are explored. However, highly optimized designs lead to high imperfection sensitivities and tend to loose robustness. Often the deterministic optimum is pushed to the boundaries of the feasible design space. As a result the optimized design, which was found by assuming deterministic model properties, may not be realizable in a production process. For this reason it is necessary to investigate, how the optimized design is affected by scattering model input variables, which could be e.g. geometry and material parameters, boundary conditions and loads. The scattering inputs are modeled int this paper by means of scalar random variables having a certain dependence between each other. Random variables have the advantage compared to other uncertainty models, that efficient methods of the well developed probability theory can be applied.

In this paper different strategies to search for a robust design are presented and investigated with respect to their efficiency and applicability to time consuming numerical models. After starting with deterministic optimization we introduce different measures to define the robustness of a design. An iterative Robust Design Optimization (RDO) is proposed where deterministic optimization is combined with variance-based robustness analysis and final reliability proof. This procedure is state-of-the-art in Dynardo's supported RDO projects (Roos and Hoffmann2008),(Roos et al.2009). The iterative procedure is compared to coupled RDO approaches, where the robustness or reliability measures are calculated for each optimization design. For such a procedure often global approximation models are used in order to enable the application for more complex problems. All presented methods are available in Dynardo's optiSLang software package (optiSLang2011), which supports a wide range of CAE solvers in order to perform a reliable optimization, sensitivity, robustness and reliability analysis as well as Robust Design Optimization.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

2. Deterministic optimization

For deterministic single-objective optimization problems the optimization task can be formulated by a single scalar-valued objective function

$$f\left(x_1, x_2, \ldots, x_k\right) \rightarrow \min, \tag{1}$$

which is often an implicit function of the design variables. The design variables can be defined as continuous variables with a lower and upper bound or as discrete variables which assume several discrete values. In unconstrained optimization problems only the bounds or values of the design variables limit the optimization space. The optimizer searches between these limits for the minimum value of the objective function $f(\mathbf{x})$.

In engineering problems often additional restrictions have to be fulfilled by the optimal design. With help of equality and inequality constraints

$$g_i(x_1, x_2, \ldots, x_k) = 0,\ i = 1 \ldots m_e, \quad h_j(x_1, x_2, \ldots, x_k) \geq 0,\ j = 1 \ldots m_u, \tag{2}$$

such restrictions can be formulated.

As optimization pre-processing a global sensitivity analysis may help to understand or to formulate the optimization problem and to possibly reduce the number of optimization variables, which enables the application of more efficient optimization strategies. In our analysis we perform variance based sensitivity analysis (Saltelli et al.2008). By representing continuous optimization variables by uniform distributions, variance based sensitivity analysis quantifies the contribution of each optimization variable to a possible improvement of the model responses. In contrast to local derivative based sensitivity methods, the variance based approach quantifies the contribution with respect to the defined variable ranges. Using the results of the sensitivity analysis the number of optimization variables may be reduced and suitable start points can be found for a following optimization.

Unfortunately, sufficiently accurate variance based methods require huge numerical effort due to the large number of simulation runs. Therefore, often meta-models are used to represent the model responses surrogate functions in terms of the model inputs. However, many meta-model approaches exist and it is often not clear which one is most suitable for which problem (Roos et al.2007). Another disadvantage of meta-modeling is its limitation to a small number of input variables. Due to the curse of dimensionality the approximation quality decreases for all meta-model types dramatically with increasing dimension. As a result, an enormous number of samples is necessary to represent high-dimensional problems with sufficient accuracy. In order to overcome these problems, Dynardo developed the Metamodel of Optimal Prognosis (Most and Will2008),(Most and Will2011). In this approach the optimal input variable subspace together with the optimal meta-model approach are determined with help of an objective and model independent quality measure, the Coefficient of Prognosis.

In Figure 1 the recommended flow of single-objective optimization procedure is shown: after the definition of the design variables and objective and constraint functions the design space is explored by sensitivity analysis. The obtained variable sensitivities may help to reduce the number of design variables. The best designs found in the sensitivity analysis could be used as start designs for the following optimization procedure which finally will determine an optimal design.

Figure 1. Flowchart of deterministic single-objective optimization

3. Robust Design Optimization

In Robust Design optimization the optimization task is formulated under the consideration of uncertainties. For this purpose we model the uncertainties with scalar random variables with a given distribution type and a possible correlations. In the RDO framework the optimization variables itself (e.g. geometry parameters of a structure) and even additional variables (e.g. material properties) may be assumed as random. This may result in pure optimization, pure stochastic and mixed optimization-stochastic variables. Additionally to the deterministic objective and constraint functions the robustness of a design is considered within the RDO procedure.

A robust design may be characterized intuitively in that way, that its performance is largely unaffected by random perturbations of the model inputs. A possible measure is the variance indicator, where the relative variations of the critical model responses are compared to the relative variation of the input variables. If certain model responses are limited with respect to an undesired performance, the safety margin can be quantified as the interval between the mean value of the model response and the limit. This is shown in

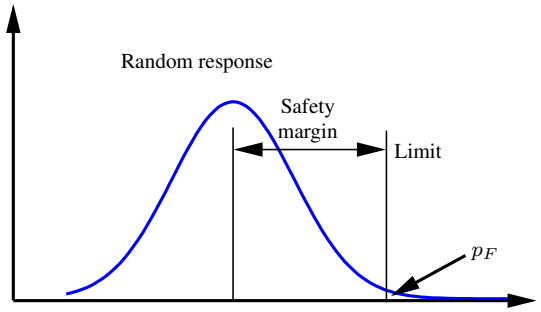


Figure 2. Random model response with given limit value and corresponding safety margin and failure probability p_F

Figure 2. The safety margin can be formulated in terms of the standard deviation of the model response. In the variance-based robustness analysis a specific safety margin $\alpha\sigma_Y$, which has to be defined by the designer, has to be proven for all critical responses

$$\|y_{limit} - \bar{Y}\| \leq \alpha\sigma_Y. \tag{3}$$

Alternatively the probability that a certain limit is exceeded can be quantified and proven to be less than an acceptable value. This probability indicator can be evaluated by the probability-based robustness analysis, which is called reliability analysis.

3.1. VARIANCE-BASED ROBUSTNESS ANALYSIS

In variance-based robustness analysis the variations of the critical model responses are investigated. In our study random sampling methods are used to generate discrete samples of the joined probability density function of the given random variables. Based on these samples the statistical properties of the model responses as mean value, standard deviation, quantiles and higher order stochastic moments are estimated. In order to obtain a sufficient quality of these estimates, it is required, that the sampling scheme represents the marginal distributions of the single random variables as well as the defined correlations between each other with high accuracy. Some very basic stochastic methods to generate sample sets are variants of the Monte-Carlo method. The simplest version is the so-called plain Monte-Carlo method (PMC). With this methods the natural scatter can be modeled quite well, but the statistical uncertainty is fairly large if the sample size is small. Therefore we utilize Latin Hypercube Sampling (LHS) with minimized correlation errors (Hungtington and Lyrintzis1998), where the marginal distributions and the predefined input correlations are represented with a small number of samples.

Based on the estimates of the mean value and the standard deviation the safety margin can be estimated by using Eq. (3) for the responses where a performance limit is given. However, by using variance-based robustness analysis only safety margins up to two sigma can be proven with a small number of samples. For larger safety margins (e.g. six sigma) the true failure rate may be heavily vary for different distribution types of the output. Since the distribution of the output is not exactly known, an estimate of low failure probabilities by variance-based measures may be very inaccurate. Therefore, we recommend to prove safety margins larger than three sigma by reliability analysis.

3.2. RELIABILITY ANALYSIS

In reliability analysis the limit state function divides the random variable space into a safe domain $S = \{\mathbf{x} : g(\mathbf{x}) > 0\}$ and a failure domain $F = \{\mathbf{x} : g(\mathbf{x}) \leq 0\}$. The vector $\mathbf{x}$ denotes a position in the space spanned by the random variable vector $\mathbf{X} = [X_1, X_2, \ldots, X_m]$. The failure probability p_F is defined as the integral of the joint probability density function $f_{\mathbf{X}}(\mathbf{x})$ of the random variables with respect to the failure domain

$$p_F = P\left[\mathbf{X} : g(\mathbf{X}) \leq 0\right] = \int_{g(\mathbf{X})\leq 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}. \tag{4}$$

The computational challenge in determining the integral of Eq. (4) lies in the evaluation of the limit state function $g(\mathbf{x})$ at a specific position $\mathbf{x}$. In CAE-based analyses the limit state function is generally an implicit function of the input variables.

The most simple and robust method for the evaluation of Eq. (4) is the Monte Carlo Simulation (MCS) where the estimated failure probability is obtained from a set of n samples $\mathbf{x}_i$ as

$$\hat{p}_F^{MCS} = \frac{1}{n} \sum_{i=1}^{n} I\left(g(\mathbf{x}_i)\right), \tag{5}$$

where the indicator function $I\left(g(\mathbf{x}_i)\right)$ is one if $g(\mathbf{x}_i)$ is negative or zero and zero else. MCS can represent arbitrary types of LSFs including discontinuities and multiple design points. The disadvantage of this method is the large number of required samples, which increases dramatically with decreasing failure probability.

Thus for engineering problems, where we deal with small probabilities of failure, this method may be very inefficient.

Further well-known methods are directional sampling (Bjerager1988), which can be applied for small failure probabilities but is limited to a small number of random variables, and the First Order Reliability Method (FORM) (Hasofer and Lind1974), which may be very efficient with respect to the number of solver evaluations. However, FORM is limited to only one dominant failure region and to a smooth limit state function, if gradient-based methods are used for the design point search.

In our study we investigate a global approximation technique, the Metamodel of Optimal Prognosis (Most and Will2008), where a global or local polynomial is constructed on the samples obtained in the variance-based robustness analysis and an adaptive approximation technique (Roos and Adam2006), (optiSLang2011) where the regions around possible design points are adapted with new support points for the approximation. The adaptive method has been proven to be very efficient for industrial problems with nonlinear limit state functions and multiple design points (Roos and Hoffmann2008),(Roos et al.2009).

3.3. Iterative Robust Design Optimization

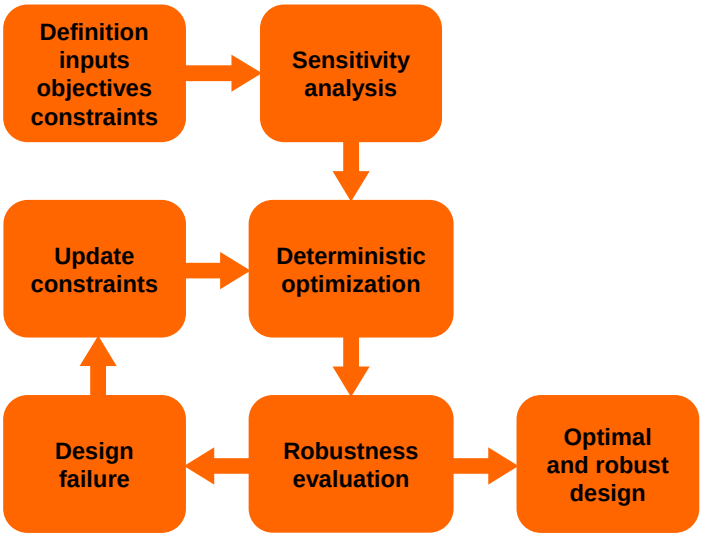


Figure 3. Flowchart of the iterative Robust Design Optimization

In iterative RDO procedure deterministic optimization is utilized by considering safety factors within the constraint conditions. These safety factors should be chosen in that way, that the robustness requirements are fulfilled. Generally the safety factors are not known a priori. In this case a suitable initial guess is specified and the initial deterministic optimization is performed. Additionally the robustness criteria are evaluated at the optimal design found by the optimizers. If the robustness requirements are not fulfilled, the optimization constraints are adjusted in a next step and the deterministic optimization procedure and the corresponding robustness analysis are performed again. This procedure is repeated until the robustness requirements are

fulfilled. In Figure 3 the flowchart of the iterative RDO procedure is shown. If a small probability of failure is required a final reliability analysis is performed.

In order to fulfill the required failure probability a suitable safety margin has to been chosen. Since the distribution of the investigated responses is unknown, this choice may be critical for the success of the iterative RDO procedure. In Table I the safety margin is given for different distribution types at different failure probabilities. The table indicates for $p_F = 10^{-2}$ a much smaller deviation between the different distributions as for $p_F = 10^{-6}$. This means that for small failure probabilities the safety margin used in the iterative RDO procedure should be taken by assuming a non-normal distribution. For example for $p_F = 10^{-6}$ a safety margin of 6σ may be a good choice.

Table I. Required safety margins to assure a given failure probability p_F

Distribution type	Required safety margin		
	$p_F = 10^{-2}$	$p_F = 10^{-3}$	$p_F = 10^{-6}$
Normal	2.32	3.09	4.75
Log-normal	3.37	5.70	14.90
Rayleigh	2.72	3.76	6.11
Weibull	2.67	3.66	5.88

3.4. Coupled Robust Design Optimization

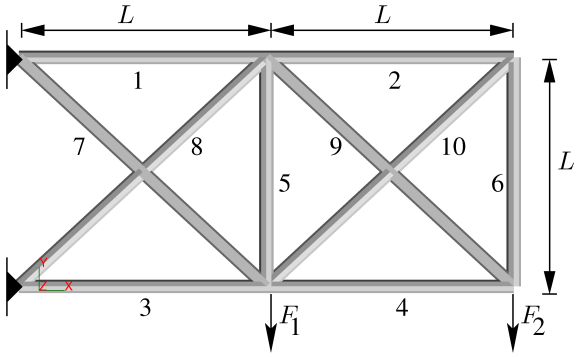
In the coupled RDO procedure an optimization is performed by considering robustness and reliability constraints directly. This means that the robustness and/or reliability measures have to been evaluated at every optimization (nominal) design. This leads to a nested double loop with the pure optimization procedure in the outer loop and the robustness analysis in the inner loop. This procedure may require a very large number of solver runs, especially if the optimization is coupled with reliability analysis. Such a strategy would limit the coupled procedure to simple and fast models. For more complex problems an improvement with respect to the number of solver runs is necessary.

One possibility could be to use a global approximation of the model responses with respect to the optimization and stochastic variables. For this purpose support points have to been generated in the mixed optimization-stochastic space which cover the possible variable values sufficiently. However, since the approximation is not exact a final robustness or reliability proof of the obtained design should be performed.

Another possibility to reduced the numerical effort of a coupled RDO procedure is to use an estimate of the safety margin, similar to the iterative approach, but with a reduced number of samples for the calculation of the mean values and standard deviations. However, in this case statistical errors may be significant and the corresponding objective and constraint function may contain additional noise. Therefore only optimization strategies should be applied which can handle such distortions. Again a more accurate robustness or reliability proof should by performed for the detected optimal design.

4. Application example

In the following example the proposed methodology is applied exemplary. For this purpose the ten-bar-truss structure shown in Figure 4 is investigated. This structure has been investigated e.g. in (Haftka and Gürdal1992). The optimization task is to minimize the mass of the truss structure. The absolute stresses of the each of the trusses should not exceed 30000 psi. Under the consideration of scattering cross sections (normally distributed with coefficient of variation of 5%), scattering material properties (Young's modulus is log-normally distributed with 5% variation) and scattering loads (normally distributed with 10% variation) the total probability of exceeding the stress limit in one of the trusses should be below 10^{-6}. The cross section areas a_i are taken as continuous optimization variables with the bounds given in Figure 4.



$$
\begin{aligned}
F_1 &= F_2 = 100000 \text{ lb}_f \\
E &= 10^7 \text{ psi} \\
\rho &= 0.1 \text{ lbs/in}^3 \\
L &= 360 \text{ in} \\
0.1 \text{ in}^2 &\leq a_i \leq 20 \text{ in}^2
\end{aligned}
$$

Figure 4. Investigated initial truss structure

4.1. Deterministic optimization

In a first step a pure deterministic optimization is performed in order to find a suitable truss topology. In the next section this optimal topology is optimized under the consideration of uncertainties. The limit of the maximum stress is reduced by a global safety factor of 1.2 to 25000 psi. Before starting the optimization task a sensitivity analysis is done. For this purpose 100 Latin Hypercube samples are generated uniformly distributed in the space of the optimization variables. Each design is evaluated by a finite element solver using geometrically linear truss elements. Using the MOP approach for sensitivity analysis the variable importance is quantified. The MOP approach indicates highly nonlinear dependencies between the optimization variables and the truss stresses as indicated in Figure 5. The number of optimization variables can not be reduced in this example since each cross section is the most important variable with respect to the belonging stress value.

The best design of the sensitivity analysis which fulfills the constraints is taken as start point for a gradient-based optimization. The mass of this start design is 3369.4 lbs. After 13 iterations with total 143 solver calls the optimizer found the optimal parameter set indicated in Figure 6. The parameter values agree excellent with the solution given in (Haftka and Gürdal1992). The results indicate, that the trusses 2,5,6 and 10 are set to its minimum value since they are not needed to carry the loads. The total mass of the optimized structure is 1593.2 lbs which is less than 50% of the start design.

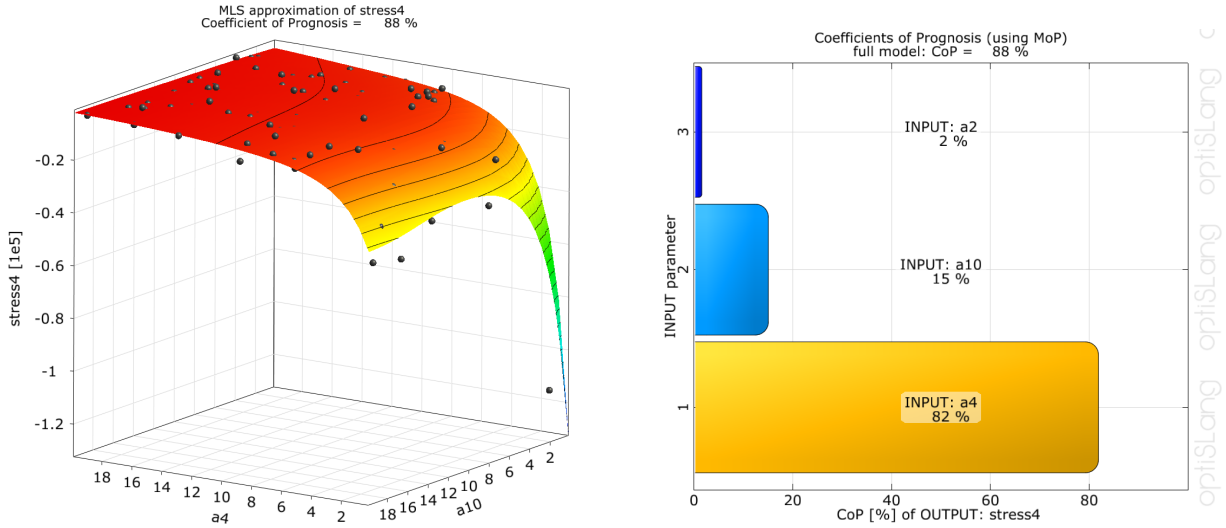


Figure 5. Approximation function and variable sensitivities obtained with the MOP approach using 100 Latin Hypercube samples as design exploration

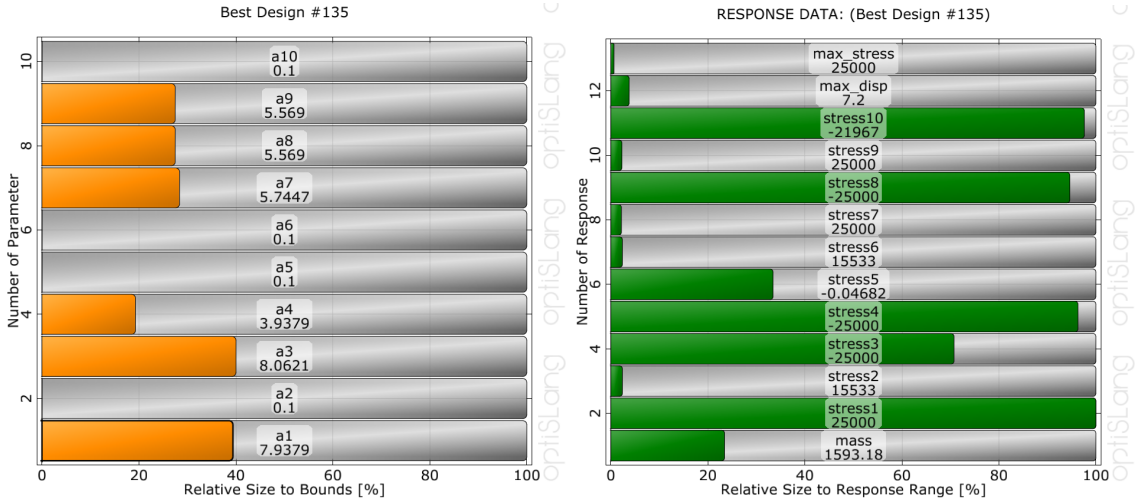
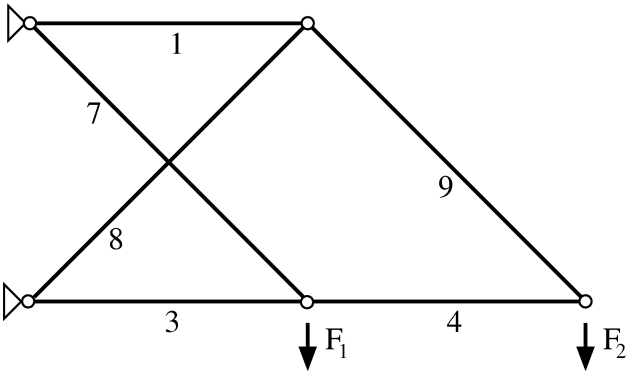


Figure 6. Results of deterministic gradient-based optimization of the full truss structure

The ten-bar-truss can be reduced by removing the unimportant trusses 2,5,6 and 10 from the structure. The stresses of the reduced structure can be simply calculated by using equilibrium equation of the forces at each of the truss nodes. The resulting stresses are given in Figure 7. Since the structure is statically determined the stresses in the trusses are independent of the Young's modulus. Thus it is not necessary to consider the Young's modulus in the optimization or robustness analysis anymore. In Figure 8 the results of a gradient-based optimization of the reduced truss structure are given. The figure indicates a slightly lower mass of the structure. Furthermore the stress limit is reached in all trusses.



$$stress1 = \frac{2F_2}{a_1} \qquad stress3 = -\frac{F_1 + F_2}{a_3}$$
$$stress4 = -\frac{F_2}{a_4} \qquad stress7 = \sqrt{2}\frac{F_1}{a_7}$$
$$stress8 = -\sqrt{2}\frac{F_2}{a_8} \qquad stress9 = \sqrt{2}\frac{F_2}{a_9}$$

Figure 7. Reduced truss structure with analytical stress values

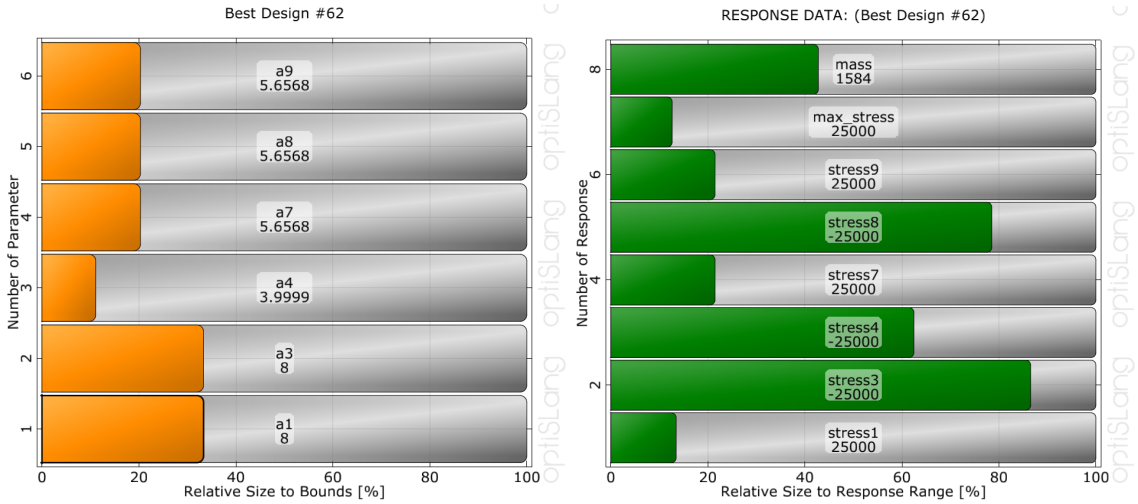


Figure 8. Results of deterministic gradient-based optimization of the reduced truss structure

4.2. Iterative Robust Design Optimization

The iterative Robust Design Optimization is performed by combining deterministic optimization using safety factors for the constraint conditions with variance-based robustness analysis. If the robustness analysis indicates a robust design the required failure probability is proven by reliability analysis. For this purpose the optimized reduced truss structure shown in Figure 7 is investigated by variance-based robustness analysis. The assumed statistical properties and distribution types of the scattering variables are used to generate 100 Latin Hypercube samples. Based on the solver evaluations the statistical properties of the truss stresses can be obtained. In Table II the results are given for the first optimization step including following robustness analysis. The table indicates for five of the six trusses a safety margin of about 1.75σ. Since we want to obtain a structure with a failure probability below 10^{-6} a safety margin of 6σ seems to be necessary in order to consider non-normal distributions of the output (see Table I). If we assume that the coefficient of variation of each stress values is constant, if the mean value is changed, we can estimate the required constraint for

Table II. Iterative Robust Design Optimization and final reliability proof of the reduced truss structure

Optimization Step 1	68 designs	Constraint	a_1	a_3	a_4	a_7	a_8	a_9	mass
		25000	8.000	8.000	4.000	5.657	5.657	5.657	1584.0
Robustness Step 1	100 samples	max. stress	stress1	stress3	stress4	stress7	stress8	stress9	
Mean value		27540	25070	-25060	-25060	25060	-25060	25060	
Standard deviation		2425	2853	2105	2825	2765	2788	2827	
Cov. of variation		0.088	0.114	0.084	0.113	0.110	0.111	0.113	
Safety margin		1.01σ	1.73σ	2.35σ	1.75σ	1.79σ	1.77σ	1.75σ	
Optimization Step 2	35 designs	Constraint	a_1	a_3	a_4	a_7	a_8	a_9	mass
		18000	11.111	11.111	5.555	7.857	7.857	7.857	2200.0
Robustness Step 2	100 samples	max. stress	stress1	stress3	stress4	stress7	stress8	stress9	
Mean value		19810	18050	-18050	-18040	18040	-18050	18040	
Standard deviation		1772	2044	1552	2000	1988	2060	1991	
Cov. of variation		0.089	0.113	0.086	0.111	0.110	0.114	0.110	
Safety margin		5.75σ	5.85σ	7.70σ	5.98σ	6.02σ	5.80σ	6.01σ	

Reliability analysis	Number of solver runs	Failure probability	Reliability index
Directional sampling	3674	$3.19 \cdot 10^{-7}$	4.98
FORM	225	-	9.70
MOP+DS	100 (from robustness)	$5.05 \cdot 10^{-7}$	4.89
ARSM+DS	101	$5.75 \cdot 10^{-7}$	4.86

the second iteration step by an extrapolation of the mean stress value

$$\begin{aligned} constraint_{step2} + 6 \cdot CV_{step1} \cdot constraint_{step2} &\leq 30000, \\ constraint_{step2} &= 30000/(1 + 6 \cdot CV_{step1}), \end{aligned}$$

which leads to a value of about 18000. The deterministic optimization is repeated with the new constraint value and the robustness of the optimized structure is assessed again by 100 Latin Hypercube samples. Table II indicates that the optimized structure of step 2 almost fulfills a safety margin of 6σ for all truss stresses.

Finally the failure probability is estimated more accurately by reliability analysis. For this purpose we investigate different methods with respect to their efficiency and accuracy. As reference solution directional sampling is used. First the First Order Reliability Method (FORM) is applied which converges to a minor important design point with a very low failure probability. Due to the individual stress limits in the six trusses the combined limit state function has several kinks and design points which lead to the wrong convergence point of FORM.

As second procedure we use a global approximation with the robustness samples as support points. For this purpose the Metamodel of Optimal Prognosis is built with these samples and an almost linear

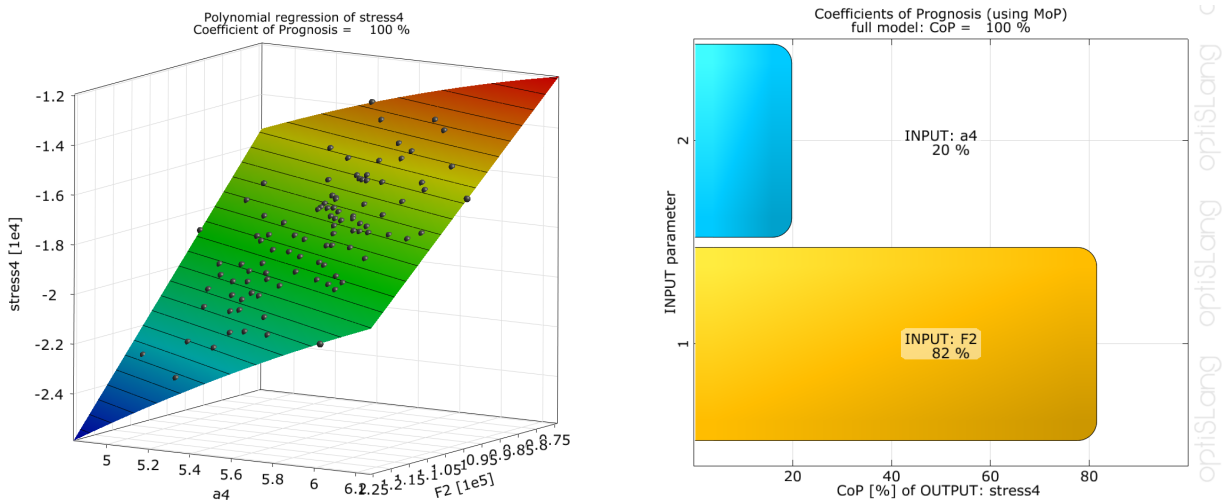


Figure 9. Approximation function and variables sensitivities obtained with the MOP approach using the 100 robustness samples of step 2 (CoP=100.00%)

dependence is indicated as shown in Figure 9. The individual stress values can be represented with very excellent approximation quality indicated by a Coefficient of Prognosis of 100.00%. However, since only 100 robustness samples are generated with the original distributions functions, the estimation of a very small failure probability requires an extrapolation of the approximation model, which may lead to a wrong estimate of the failure probability. Nevertheless, the calculated failure probability given in Table II shows good agreement with the reference from direction sampling. Since the robustness samples are available anyhow, this procedure requires no additional solver runs and should be investigated if the MOP indicates a good approximation quality.

The third investigated procedure is the Adaptive Response Surface Method according to (Roos and Adam2006). This methods uses an initial sampling scheme as support points which is stretched by factor three in order to cover a larger domain. With two additional adaption steps, where new sampling schemes are placed around the detected important regions, the method converges to a failure probability close to the reference solution. Since the number of solver evaluations is very small and since no extrapolation is used in the approximation, from our viewpoint this ARSM approach is the method of choice for reliability analysis of complex engineering problems. The proposed iterative Robust Design Optimization procedure including the ARSM reliability proof has been successfully applied to real industrial problems in (Roos and Hoffmann2008),(Roos et al.2009).

4.3. Coupled Robust Design Optimization

In a further analysis the coupled Robust Design Optimization approach is applied. For this purpose in a first step a global approximation model is used and in a second step direct solver runs are evaluated. The global approximation model is built by using a uniform distribution for all optimization and stochastic variables, where the lower and upper bounds are taken for the cross section areas as $2\,\text{in}^2 \leq a_i \leq 20\,\text{in}^2$ and for the pure stochastic forces the bounds are taken as mean value +/- 5σ. 500 Latin Hypercube samples

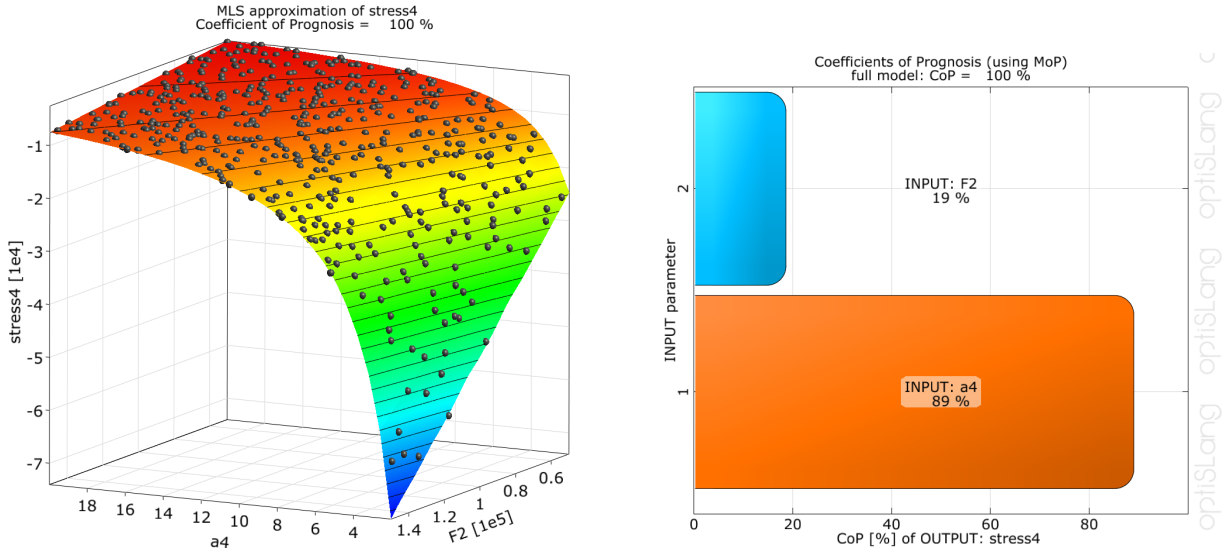


Figure 10. Approximation function and variables sensitivities obtained with the MOP approach using the 500 samples in the mixed design-stochastic space (CoP=99.99%)

are generated within this mixed optimization-stochastic space and the Metamodel of Optimal Prognosis is built. For all individual stress values and the mass an approximation quality better or equal 99.99% could be reached by the quadratic Moving Least Squares approximation, which is included in the MOP approach. In Figure 10 the approximation function and the variable sensitivities are shown exemplary. If we compare the variable sensitivities of Figure 10 with these of Figure 9, we notice, that the pure stochastic force variation is dominant in the local robustness problem but minor important in the mixed space. In many other applications we observed similar results, that the pure stochastic variables are minor dominant with respect to the optimization variables in the mixed space due to their smaller variation. This fact may lead to an inaccurate representation of the influence of the stochastic variables in an approximation model. As a consequence the estimated robustness measures may be inaccurate as well.

For our example we use an Evolutionary Algorithm (EA) running with the approximation model, where for each optimization design a variance-based robustness analysis is performed by using 100 Latin hypercube samples. The mass is taken as deterministic objective function and the stress constraints are formulated with respect to the statistical measures of the robustness analysis

$$mean\_stress_i + 6 \cdot sigma\_stress_i \leq 30000. \tag{6}$$

The results of this optimization are given in Table III. The obtained mass is almost similar to the mass obtained by the iterative procedure, but some cross section areas are different. The final optimum is investigated by a robustness analysis with direct solver calls which results in an estimated safety margin slightly larger as 6σ for all stress values. The reliability proof reports an failure probability below the required 10^{-6}. Again the ARSM approach is very efficient. In the investigated example the coupled RDO approach using a global approximation model gives satisfactory results with a relatively small number of solver runs. However, in cases where the approximation quality is not as excellent as in our example the global approximation may fail. In such cases the iterative approach should give the most efficient solution.

Table III. Coupled Robust Design Optimization using a global response surface with final robustness and reliability proof

RDO on global MOP	500 support points	a_1	a_3	a_4	a_7	a_8	a_9	mass
EA on MOP	591 nominal designs	11.306	10.303	5.640	8.049	8.131	7.980	2211.0
	(100 robustness samples each)							

Robustness analysis	100 samples	stress1	stress3	stress4	stress7	stress8	stress9
Mean value		17730	-19470	-17770	17610	-17470	17790
Standard deviation		1961	1683	1964	1977	1958	1993
Cov. of variation		0.111	0.086	0.111	0.110	0.112	0.112
Safety margin		6.26σ	6.26σ	6.23σ	6.27σ	6.40σ	6.13σ

Reliability analysis	Number of solver runs	Failure probability	Reliability index
Directional sampling	3366	$2.26 \cdot 10^{-7}$	5.05
ARSM+DS	84	$1.24 \cdot 10^{-7}$	5.16

Table IV. Coupled Robust Design Optimization using the direct solver with rough robustness estimates and final robustness and reliability proof

RDO	3822 solver calls	a_1	a_3	a_4	a_7	a_8	a_9	mass
ARSM	(182 nominal designs with	11.035	10.024	5.562	7.697	7.931	7.848	2153.5
	20 robustness samples each)							

Robustness analysis	100 samples	stress1	stress3	stress4	stress7	stress8	stress9
Mean value		18170	-20000	-18030	18420	-17880	18070
Standard deviation		2018	1716	2021	2010	2003	2020
Cov. of variation		0.111	0.086	0.112	0.109	0.112	0.112
Safety margin		5.86σ	5.83σ	5.92σ	5.76σ	6.05σ	5.91σ

Reliability analysis	Number of solver runs	Failure probability	Reliability index
Directional sampling	4444	$1.57 \cdot 10^{-6}$	4.66
ARSM+DS	121	$1.77 \cdot 10^{-6}$	4.64

Finally the coupled RDO procedure is performed by direct solver runs. In order to limit the number of solver evaluations the robustness analysis inside the RDO is performed using only 20 samples to estimate the statistical properties. The final results are again verified by a more accurate robustness analysis and a reliability proof. In Table IV the results of the direct RDO procedure are given. As optimizer an adaptive polynomial response surface approach is used, which can handle noisy model responses (optiSLang2011). Due to the small number of robustness samples the estimated mean values and standard deviations contain statistical errors which may lead to noticeable solver noise. The direct RDO approach leads to a truss struc-

ture having a smaller mass and safety margins for all stress values close to 6σ. The results of the reliability analysis indicate a small violation of the reliability requirement. Additionally, the number of necessary solver runs is much larger as by using the iterative procedure or the global approximation by MOP. For this reason an application of the direct RDO approach for industrial problems is not very attractive.

5. Conclusions

In this paper Robust Design Optimization concepts have been proposed, which are applicable for complex engineering problems, where the underlying structural model is often a very time consuming numerical simulation model. By means of an investigated truss structure different procedures have been analyzed. First, an iterative RDO procedure has been proposed. In this approach after each deterministic optimization the required safety margin is checked by variance-based robustness analysis. If the safety margin is not sufficient the deterministic optimization constraints are adapted. For a satisfactory safety margin the required failure probability is proven finally by efficient reliability methods.

The iterative method was compared with a global response surface method built up in the mixed optimization-stochastic space. If the approximation has very high accuracy, which was checked by the Coefficient of Prognosis, a coupled RDO procedure applied on the response surface may lead to sufficient results. However, since the global approximation has often low accuracy and since the numerical effort with respect to the number of solver runs is similar to the iterative procedure, we recommend the iterative procedure for practical applications. The iterative procedure has been proven to be very efficient and accurate for real product development in (Roos and Hoffmann2008),(Roos et al.2009).

References

Bjerager, P. (1988). Probability integration by directional simulation. *Journal of Engineering Mechanics 114*, 1288–1302.

Haftka, R. and Z. Gürdal (1992). *Elements of structural optimization*, Volume 11. Springer.

Hasofer, A. and N. Lind (1974). An exact and invariant first-order reliability format. *Journal of the Engineering Mechanics Division, ASCE 100*(6), 111–121.

Hungtington, D. E. and C. S. Lyrintzis (1998). Improvements to and limitations of Latin hypercube sampling. *Probabilistic Enginerring Mechanics 13*, 245–253.

Most, T. and J. Will (2008). Metamodel of Optimal Prognosis - an automatic approach for variable reduction and optimal metamodel selection. In *Proc. Weimarer Optimierungs- und Stochastiktage 5.0, Weimar, Germany, November 20-21, 2008*.

Most, T. and J. Will (2011). Sensitivity analysis using the Metamodel of Optimal Prognosis. In *Proc. Weimarer Optimierungs- und Stochastiktage 8.0, Weimar, Germany, November 24-25, 2011*.

optiSLang (2011). *The optimizing Structural Language, An Users' Manual* (Version 3.2.1 ed.). Weimar, Germany: Dynardo GmbH.

Roos, D. and U. Adam (2006). Adaptive moving least square approximation for the design reliability analysis. In *Proc. Weimarer Optimierungs- und Stochastiktage 3.0, Weimar, Germany*.

Roos, D., J. Einzinger, and V. Bayer (2009). Robust design optimization applied to structural, thermal and fluid analysis including manufacturing tolerances. In *Proc. Weimarer Optimierungs- und Stochastiktage 6.0, Weimar, Germany*.

Roos, D. and R. Hoffmann (2008). Successive robust design optimization of an electronic connector. In *Proc. Weimarer Optimierungs- und Stochastiktage 5.0, Weimar, Germany*.

Roos, D., T. Most, J. F. Unger, and J. Will (2007). Advanced surrogate models within the robustness evaluation. In *Proc. Weimarer Optimierungs- und Stochastiktage 4.0, Weimar, Germany, November 29-30, 2007*.

Saltelli, A. et al. (2008). *Global Sensitivity Analysis. The Primer*. Chichester, England: John Wiley & Sons, Ltd.

Nonlinear Interval Finite Elements for Structural Mechanics Problems

Rafi L. Muhanna[1], Robert L. Mullen[2], and M. V. Rama Rao[3]
[1]*School of Civil and Environmental Engineering, Georgia Institute of Technology Atlanta, GA 30332-0355 USA, rafi.muhanna@gtsav.gatech.edu*
[2]*School of Civil and Environmental Engineering, University of South Carolina, GA 30332-0355, USA rlm@cec.sc.edu*
[3]*Vasavi College of Engineering, Hyderabad - 500 031 INDIA dr.mvrr@gmail.com*

Abstract: Interval Finite Element Method (IFEM) has been developed to handle load, material, and geometric uncertainties that are introduced in a form of interval numbers defined by their lower and upper bounds. However, the scope of the previous methods was limited to linear problems. The present work introduces an IFEM formulation for problems involving material nonlinearity. The algorithm is based on the previously developed high accuracy interval solutions. Two approaches are introduced; an iterative method that generates successive approximations to the secant stiffness and a modified Newton-Raphson method based on deterministic/interval two face strategy that carries out the iteration successfully by identifying interval multipliers for each load throughout the iteration procedure. Examples are presented to illustrate the behavior of both formulations.

Keywords: Finite Elements; Interval; Nonlinear; Materials.

1. Introduction

Structural analysis without considering uncertainty in loading or material properties leads to an incomplete understanding of the structural performance. Structural analysis using interval variables has been used by several researchers to incorporate uncertainty into structural analysis (Koyluoglu, H. U., Cakmak, A. S., and Nielson, S. R. K. 1995, Muhanna, R. L. and Mullen, R. L. 1995, Nakagiri S. and Yoshikawa, N. 1996, Rao, S. S. and Sawyer, P. 1995, Rao, S. S. and Berke, L. 1997, Rao, S.S., and Chen Li 1998, Muhanna and Mullen, 2001, Neumaier and Pownuk 2007). To the authors' knowledge, applications of interval methods for the analysis of structures with material nonlinearity do not exist anywhere in literature.

In this paper, we present an initial investigation into the application of interval finite element methods to non-linear problems of structural mechanics. In this work, we will consider deformation theory of two dimensional truss structures with a plasticity model for the material response. Critical to our development is the computation of element strains with minimal possible overestimation. Usually, derived quantities in Interval Finite Element Method (IFEM) such as stresses and strains have additional overestimation in comparison with primary quantities such as displacements. This issue has plagued displacement-based IFEM for quite some time. The recent development of mixed/hybrid IFEM formulation by the authors (Rama Rao, Mullen and Muhanna, 2010) is capable of simultaneous calculation of interval strains and

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

displacements with the same accuracy. This opened the road for further progress in new application areas such as nonlinear analysis.

This work presents two approaches to the solution of interval finite elements with material nonlinearity, namely: the interval secant and interval modified Newton-Raphson methods. The paper is structured as follows. First, a short review of some interval concepts and an overview of linear IFEM are introduced. The interval secant and interval Newton-Raphson methods are then presented in sections 3 and 4. Examples are finally presented and discussed.

2. Linear Interval Finite Element Method

Finite element method is one of the most common numerical methods for solving differential and partial differential equations with enormous applications in different fields of science and engineering. Interval finite element methods have been developed to handle the analysis of systems for which uncertain parameters are described as intervals. A variety of solution techniques have been proposed for IFEM. A comprehensive review of these techniques can be found in (Muhanna *et al.*, 2007, Zhang, 2005, and Rama Rao, Mullen and Muhanna, 2010). Interval analysis concerns the numerical computations involving interval numbers. *All interval quantities will be introduced in non-italic boldface font*. The four elementary operations of real arithmetic, namely addition (+), subtraction (−), multiplication (×) and division (÷) can be extended to intervals. Operations $\circ \in \{+,-,\times,\div\}$ over interval numbers $\mathbf{x}$ and $\mathbf{y}$ are defined by the general rule (Moore, 1966; Neumaier, 1990)

$$\mathbf{x} \circ \mathbf{y} = [\min(x \circ y), \max(x \circ y)] \quad for \ \circ \in \{+,-,\times,\div\}, \tag{1}$$

in which x and y denote generic elements $x \in \mathbf{x}$ and $y \in \mathbf{y}$. Software and hardware support for interval computation are available such as (Sun microsystems, 2002; Knüppel, 1994, and INTLAB,1999). For a real-valued function $f(x_1,\ldots,x_n)$, the *interval extension* of $f(\ )$ is obtained by replacing each real variable x_i by an interval variable $\mathbf{x}_i$ and each real operation by its corresponding interval arithmetic operation. From the fundamental property of *inclusion isotonicity* (Moore, 1966), the range of the function $f(x_1,\ldots, x_n)$ can be rigorously bounded by its interval extension function

$$f(\mathbf{x}_1,..,\mathbf{x}_n) \supseteq \{f(x_1,..,x_n) \mid x_1 \in \mathbf{x}_1,..,x_n \in \mathbf{x}_n\} \tag{2}$$

Equation (2) indicates that $f(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ contains the range of $f(x_1,\ldots,x_n)$ for all $x_i \in \mathbf{x}_i$. A natural idea to implement interval FEM is to apply the interval extension to the deterministic FE formulation. Unfortunately, such a naïve use of interval analysis in FEM yields meaningless and overly wide results (Muhanna and Mullen, 2001; Dessombz *et al*., 2001). The reason is that in interval arithmetic each occurrence of an interval variable is treated as a different, independent variable. It is critical to the formulation of the interval FEM that one identifies the dependence between the interval variables and prevents the overestimation of the interval width of the results. In this paper, an element-by-element (EBE) technique is utilized for element assembly (Muhanna and Mullen, 2001; Zhang, 2005). The elements are detached so that there are no connections between elements, avoiding element coupling. The Lagrange multiplier method is then employed to impose constraints to ensure the compatibility. Then a mixed/hybrid

formulation is incorporated to simultaneously calculate the interval strains and displacements (Rama Rao, Mullen and Muhanna, 2010). This linear formulation results in the interval linear system of equations that has the following structure:

$$(K + B\ \mathbf{D}\ A)\ \mathbf{u} = a + F\ \mathbf{b}, \tag{3}$$

with interval quantities in **D** and **b** only. The term ($K + B\ \mathbf{D}\ A$) represents the interval structural stiffness matrix and the $a + F\ \mathbf{b}$ term, the structural loading. Any interval solver can be used to solve Eq. (3), however, the following iterative scheme that is developed by Neumaier (Neumaier and Pownuk, 2007) is superior for large uncertainty, defining:

$$C := (K + BD_0A)^{-1} \tag{4}$$

where D_0 is chosen in a manner that ensures its invertability (often D_0 is selected as the midpoint of **D**), the solution **u** can be written as:

$$\mathbf{u} = (Ca) + (CF)\mathbf{b} + (CB)\mathbf{d} \tag{5}$$

To obtain a solution with tight interval enclosure we define two auxiliary interval quantities,

$$\begin{aligned} \mathbf{v} &= A\mathbf{u} \\ \mathbf{d} &= (D_0 - \mathbf{D})\mathbf{v}, \end{aligned} \tag{6}$$

which, given an initial estimate for **u**, we iterate as follows:

$$\mathbf{v}^{k+1} = \{ACa) + (ACF)\mathbf{b} + (ACB)\mathbf{d}^k\} \cap \mathbf{v}^k, \quad \mathbf{d}^{k+1} = \{(D_{c0} - \mathbf{D}_c)\mathbf{v}^{k^{+1}} \cap \mathbf{d}^k, \tag{7}$$

until the enclosures converge, from which the desired solution **u** can be obtained in a straightforward manner.

In this paper the above mentioned iterative enclosure has been used for the solution of the linear interval system of Equation (3). The solution includes displacements, strains, and forces simultaneously with the same high level of accuracy.

3. Interval Secant Method

The first method chosen for solving the system of non-linear interval equations resulting from the interval finite element method is the secant method (Cook, 2002). Given a constitutive relationship, the secant method is an iterative approach that predicts the value of the secant modulus corresponding to a certain level of loading. If load uncertainty is given as an interval value, the resulting element strain will also be an interval quantity. This will lead to an interval value for the secant modulus with the bounds on secant modulus calculated from the bounds on the element strain. In the present work, we will introduce an

iterative algorithm that allows the prediction of the interval secant modulus and calculates relevant quantities such as stresses, strains, and displacements for nonlinear material problems

To illustrate the approach, we will assume that for each element in the structure the constitutive relationship is defined as a cubic function as shown below:

$$\sigma = a\varepsilon + b\varepsilon^3, \tag{8}$$

as shown in Figure 1, where σ, ε, a, and b are stress, strain, and constants respectively. The iteration process starts by taking the initial value of the secant modulus at zero strain. In subsequent iterations, a secant modulus is calculated from the current element strain using Eq. (8) as

$$Es(i) = \frac{\sigma(i)}{\varepsilon(i)}, \tag{9}$$

where $Es(i)$, $\sigma(i)$, and $\varepsilon(i)$ are the secant modulus, stress, and strain at iteration i respectively. The iterations continue until convergence with respect to the secant modulus is achieved. However, if the load is given as an interval value, the resulting secant modulus will also be an interval quantity.

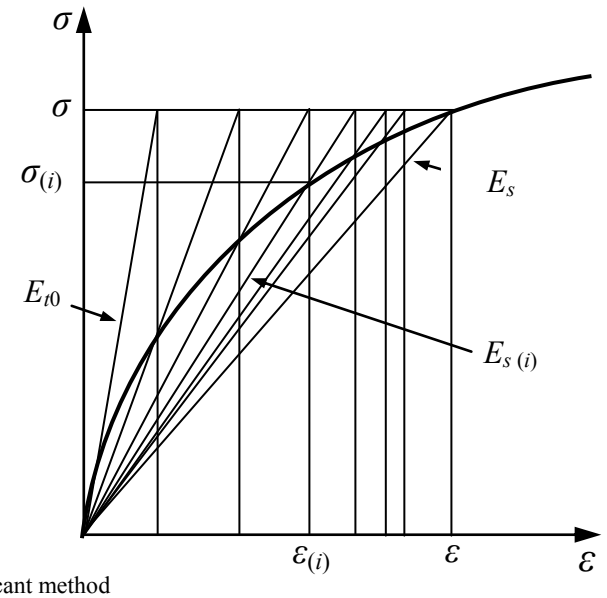


Figure 1. Stress-strain relationship, secant method

A direct calculation of the interval secant modulus from Eq. (9) will lead to an overestimation, however, considering the physics of the problem we can confirm that the lower and upper bounds of the stress in a given element correspond to lower and upper bounds of the strain respectively. Considering this dependency, the interval form of Equation (8) can be introduced as

$$[\underline{Es}(i), \overline{Es}(i)] = [\inf(\frac{\underline{\sigma}(i)}{\underline{\varepsilon}(i)}, \frac{\overline{\sigma}(i)}{\overline{\varepsilon}(i)}), \sup(\frac{\underline{\sigma}(i)}{\underline{\varepsilon}(i)}, \frac{\overline{\sigma}(i)}{\overline{\varepsilon}(i)})] \tag{10}$$

The IFEM formulation presented in section 2 (Rama Rao, Mullen and Muhanna, 2010) provides the strains along with displacements and forces. Solution of conventional interval finite elements provides interval bounds of displacements, the calculations of strains from displacements result in significant overestimation. In the current formulation, the strains are not calculated from displacements but are obtained simultaneously with the displacements and forces, thus avoiding any additional *overestimation*. These sharp bounds on interval strains thus obtained are used in the following iterative algorithm for calculating the updated interval values of the secants.

3.1. ALGORITHM FOR SECANT UPDATE

The following notations are used:

K : interval stiffness matrix
P : interval load vector
U : solution vector, includes stain and stress vectors
ε: : current strain
σ : current stress
inf: : infimum
sup: : supremum
$\mathbf{E}_s$: current secant
E_{t0} : initial secant modulus

for count = 1: countmax
  $\mathbf{K}_c(\mathrm{U})\ \mathbf{U} = \mathbf{P}$
  $\mathbf{U} = \mathbf{K}^{-1}(\mathrm{U})\ \mathbf{P}$: *Obtain solution based on algorithm given in section* 2.
    for e = 1: number of elements
      $\max(\sigma) = a \times \sup(\boldsymbol{\varepsilon}) + b \times (\sup(\boldsymbol{\varepsilon}))^3$
      $\max(E_s) = \max(\sigma) / \sup(\boldsymbol{\varepsilon}(\mathrm{e}))$
      $\min(\sigma) = a \times \inf(\boldsymbol{\varepsilon}) + b \times (\inf(\boldsymbol{\varepsilon}))^3$
      $\min(E_s) = \min(\sigma) / \inf(\boldsymbol{\varepsilon}(\mathrm{e}))$
      $\mathbf{E}_s(\mathrm{e}) = \mathrm{infsup}(\min(E_s), \max(E_s))$
    end : of loop on elements
  $\mathbf{K}_c$: update K with the new values of $\mathbf{E}_s$
end : of loop on count

For the stopping criterion the sum of the L_1 norms of the following relative change of the secant lower and upper bounds is required to be less than a specified small value

$$\left\| \frac{\underline{Es}(i+1) - \underline{Es}(i)}{\underline{Es}(i)} \right\|_1, \quad \left\| \frac{\overline{Es}(i+1) - \overline{Es}(i)}{\overline{Es}(i)] } \right\|_1 \tag{11}$$

4. Interval Modified Newton-Raphson Method

Newton-Raphson method and modified Newton-Raphson method are iterative methods to find the relation of load versus displacement based on a given constitutive relationship, or within the context of finite elements, to solve the following nonlinear system of equations

$$K(U)U = P, \tag{12}$$

where K, U, and P are the stiffness matrix, the displacement vector, and the load vector, respectively.

Modified Newton-Raphson method in finite element applications uses incremental tangent stiffness at each loading level to predict the displacement as summation of a number of incremental solutions using the out of balance load for each increment. For each increment (iteration) the balance of forces at each node is checked until equilibrium is attained that represents the convergence for that specific load level.

If the applied load is given as an interval value, the internal forces at each node will be intervals, checking the equilibrium at each node will represent a significant challenge. As a matter of fact, the dependency of internal forces on the applied load will result in an overestimation that will not allow equilibrium to be checked properly at each node. In the next section we will introduce a formulation in which we try to delay the use of interval multipliers as much as possible in a way that nodal equilibrium can be checked.

4.1. Formulation of Interval Newton-Raphson Method

The formulation of interval finite element introduced in a recent work by the authors (Rama Rao, Mullen and Muhanna, 2010), provides a solution vector that includes displacements, internal forces, and strains of the system. Using the tangent stiffness, K_t, in each of the iterations will convert the system of equations in Equation (12) to a linear system of equations of the form

$$K_t\mathbf{U} = \mathbf{P}, \tag{13}$$

this equation can be reintroduced in the form

$$K_t\mathbf{U} = M\mathbf{d}, \tag{14}$$

where M is a matrix with dimensions (No. degrees of freedom × number of loads) and $\mathbf{d}$ is a vector of load interval multipliers, (Mullen and Muhanna 1999). The solution of Equation (14) can be given in the following form

$$\mathbf{U} = K_t^{-1}M\mathbf{d}, \tag{15}$$

or

$$\mathbf{U} = M_1\mathbf{d}, \tag{15a}$$

where M_1 is a deterministic matrix with the dimensions (No. degrees of freedom × number of loads). The entries of this matrix are the system solution introduced per each load, or the solution Load-By-Load (LBL). For the clarity of formulation we will assume only two applied loads. This assumption will not

impose any restriction on the generality of the formulation. Since we order the unknowns in U as first displacements then element forces, ending with element strains, the entries of last rows of matrix M_1 are element strains introduced LBL and have the form

$$M_s = \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \\ \vdots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} \end{pmatrix}, \tag{16}$$

where ε_{ij} is the strain of the i^{th} element due to the j^{th} load. Truss structures will be considered for the rest of the paper, such consideration should not limit the generality of the formulation.

In Newton-Raphson iteration and for a given load level, using the tangent stiffness will result in a vector of internal forces different from the vector of applied forces and equilibrium will not be satisfied. The difference between the two vectors is used to compute the residual response until convergence, or equilibrium is attained. This will require the computation of the internal force vector and the comparison with the applied load vector for each of the iterations. In the current interval formulation, as mentioned above, we will try to delay operations on intervals as much as possible.

If we consider the nonlinear constitutive relationship in Equation (8) the internal force for any element can be given as

$$F_i = \sigma_i A_i = (a\varepsilon_i + b\varepsilon_i^3)A_i, \tag{17}$$

where A_i is the cross-sectional area of the i^{th} element. Substituting for strains from equation (16) and including load interval multipliers in Equation (15), we get

$$F_i = \{a\begin{pmatrix}\varepsilon_{i1} & \varepsilon_{i2}\end{pmatrix}\begin{pmatrix}\mathbf{d}_1 \\ \mathbf{d}_2\end{pmatrix} + b[\begin{pmatrix}\varepsilon_{i1} & \varepsilon_{i2}\end{pmatrix}\begin{pmatrix}\mathbf{d}_1 \\ \mathbf{d}_2\end{pmatrix}]^3\}A_i, \tag{17}$$

or

$$F_i = [a(\varepsilon_{i1}\mathbf{d}_1 + \varepsilon_{i2}\mathbf{d}_2) + b(\varepsilon_{i1}^3\mathbf{d}_1^3 + \varepsilon_{i2}^3\mathbf{d}_2^3 + 3\varepsilon_{i1}\varepsilon_{i2}^2\mathbf{d}_1\mathbf{d}_2^2 + 3\varepsilon_{i1}^2\varepsilon_{i2}\mathbf{d}_1^2\mathbf{d}_2)]A_i, \tag{18}$$

If n elements meet at node m, the x global component of resultant of element forces can be obtained as

$$\begin{aligned}\mathbf{F}_{mx} = [&a(\varepsilon_{11}A_1\cos\theta_1 + \cdots + \varepsilon_{n1}A_n\cos\theta_n)\mathbf{d}_1 + \\ &+ a(\varepsilon_{12}A_1\cos\theta_1 + \cdots + \varepsilon_{n2}A_n\cos\theta_n)\mathbf{d}_2 + \\ &+ b[(\varepsilon_{11}^3A_1\cos\theta_1 + \cdots + \varepsilon_{n1}^3A_n\cos\theta_n)\mathbf{d}_1^3 + \\ &+ \quad(\varepsilon_{12}^3A_1\cos\theta_1 + \cdots + \varepsilon_{n2}^3A_n\cos\theta_n)\mathbf{d}_2^3 + \\ &+ (3\varepsilon_{11}\varepsilon_{12}^2A_1\cos\theta_1 + \cdots + 3\varepsilon_{n1}\varepsilon_{n2}^2A_n\cos\theta_n)\mathbf{d}_1\mathbf{d}_2^2 + \\ &+ (3\varepsilon_{11}^2\varepsilon_{12}A_1\cos\theta_1 + \cdots + 3\varepsilon_{n1}^2\varepsilon_{n2}A_n\cos\theta_n)\mathbf{d}_1^2\mathbf{d}_2]\end{aligned} \tag{19}$$

and both components can be introduced in the following form

$$\begin{aligned}\mathbf{F}_{mx} &= c_{mx1}\mathbf{d}_1 + c_{mx2}\mathbf{d}_2 + c_{mx3}\mathbf{d}_1^3 + c_{mx4}\mathbf{d}_2^3 + c_{mx5}\mathbf{d}_1\mathbf{d}_2^2 + c_{mx6}\mathbf{d}_1^2\mathbf{d}_2 \\ \mathbf{F}_{my} &= c_{my1}\mathbf{d}_1 + c_{my2}\mathbf{d}_2 + c_{my3}\mathbf{d}_1^3 + c_{my4}\mathbf{d}_2^3 + c_{my5}\mathbf{d}_1\mathbf{d}_2^2 + c_{my6}\mathbf{d}_1^2\mathbf{d}_2\end{aligned} \tag{20}$$

or

$$\begin{pmatrix}\mathbf{F}_{mx}\\ \mathbf{F}_{my}\end{pmatrix} = \begin{pmatrix}c_{mx1} & c_{mx2}\\ c_{my1} & c_{my2}\end{pmatrix}\begin{pmatrix}\mathbf{d}_1\\ \mathbf{d}_2\end{pmatrix} + \begin{pmatrix}c_{mx3} & c_{mx4}\\ c_{my3} & c_{my4}\end{pmatrix}\begin{pmatrix}\mathbf{d}_1^3\\ \mathbf{d}_2^3\end{pmatrix} + \begin{pmatrix}c_{mx5} & c_{mx6}\\ c_{my5} & c_{my6}\end{pmatrix}\begin{pmatrix}\mathbf{d}_1\mathbf{d}_2^2\\ \mathbf{d}_1^2\mathbf{d}_2\end{pmatrix} \tag{21}$$

and the vector of internal forces, after including all nodes, can be introduced in the form

$$\mathbf{F} = \begin{pmatrix}\mathbf{F}_{1x}\\ \mathbf{F}_{1y}\\ \vdots\\ \mathbf{F}_{mx}\\ \mathbf{F}_{my}\end{pmatrix} = \begin{pmatrix}c_{1x1} & c_{1x2}\\ c_{1y1} & c_{1y2}\\ \vdots & \vdots\\ c_{mx1} & c_{mx2}\\ c_{my1} & c_{my2}\end{pmatrix}\begin{pmatrix}\mathbf{d}_1\\ \mathbf{d}_2\end{pmatrix} + \begin{pmatrix}c_{1x3} & c_{1x4}\\ c_{1y1} & c_{1y2}\\ \vdots & \vdots\\ c_{mx3} & c_{mx4}\\ c_{my3} & c_{my4}\end{pmatrix}\begin{pmatrix}\mathbf{d}_1^3\\ \mathbf{d}_2^3\end{pmatrix} + \begin{pmatrix}c_{1x5} & c_{1x6}\\ c_{1y5} & c_{1y6}\\ \vdots & \vdots\\ c_{mx5} & c_{mx6}\\ c_{my5} & c_{my6}\end{pmatrix}\begin{pmatrix}\mathbf{d}_1\mathbf{d}_2^2\\ \mathbf{d}_1^2\mathbf{d}_2\end{pmatrix} \tag{22}$$

where m is the total number of nodes in the system. Interval internal force vector in Equation (22) is introduced as a product of two separate parts; deterministic and interval. The interval part represents the original load multipliers. This form will allow the comparison of the deterministic values of the applied load (matrix M in Equation (13)) with the deterministic values of the internal forces during the iteration procedure. If we reintroduce Equation (22) in the form

$$\mathbf{F} = MM_1\begin{pmatrix}\mathbf{d}_1\\ \mathbf{d}_2\end{pmatrix} + MM_2\begin{pmatrix}\mathbf{d}_1^3\\ \mathbf{d}_2^3\end{pmatrix} + MM_3\begin{pmatrix}\mathbf{d}_1\mathbf{d}_2^2\\ \mathbf{d}_1^2\mathbf{d}_2\end{pmatrix} \tag{22}$$

and compare with Equation (13) for the case of two loads

$$P = M_1\begin{pmatrix}\mathbf{d}_1\\ \mathbf{d}_2\end{pmatrix}, \tag{23}$$

The '*out of balance'* force vector can now be introduced as

$$\delta\mathbf{F} = (M + MM_1)\begin{pmatrix}\mathbf{d}_1\\ \mathbf{d}_2\end{pmatrix} + MM_2\begin{pmatrix}\mathbf{d}_1^3\\ \mathbf{d}_2^3\end{pmatrix} + MM_3\begin{pmatrix}\mathbf{d}_1\mathbf{d}_2^2\\ \mathbf{d}_1^2\mathbf{d}_2\end{pmatrix} \tag{24}$$

To complete the iteration and update procedure we will introduce what is called deterministic/interval strategy. This strategy is based on introducing a deterministic out of balance Load-By-Load matrix in the form

$$\delta M = (M + MM_1) + MM_2 + MM_3 \tag{24}$$

and providing a deterministic solution in the form

$$\delta UM = K^{-1}\delta M \tag{25}$$

where the entries of δUM are the deterministic solution for a given iteration introduced LBL. The solution will be used to update the deterministic LBL element strains as follows

$$Ms = Ms + \delta UMs \tag{26}$$

where δUMs is a matrix of the dimension (number of elements × number of loads). This matrix is the bottom part of δUM that contains incremental values of element strains listed LBL. The updated value of Ms is used to update the values of internal forces.

On the interval side the incremental solution will be obtained from

$$\delta\mathbf{U} = K^{-1}\delta\mathbf{F} \tag{26}$$

and the interval solution update is

$$\mathbf{U} = \mathbf{U} + \delta\mathbf{U} \tag{27}$$

Crucial to the quality of the solution given in Equation (27) is the evaluation of $\delta\mathbf{F}$ used in Equation (26). For example, if we consider the x component of internal forces at node m as in Equation (20), the '*out of balance'* force can be given as

$$\delta\mathbf{F}_{mx} = (m_{mx} + c_{mx1})\mathbf{d}_1 + c_{mx2}\mathbf{d}_2 + c_{mx3}\mathbf{d}_1^3 + c_{mx4}\mathbf{d}_2^3 + c_{mx5}\mathbf{d}_1\mathbf{d}_2^2 + c_{mx6}\mathbf{d}_1^2\mathbf{d}_2 \tag{28}$$

The objective here is to obtain the tightest possible enclosure for $\delta\mathbf{F}_{mx}$. Due to the multiple occurrences of $\mathbf{d}_1$ and $\mathbf{d}_2$ in (28), a direct evaluation of the function will lead to overestimation due to what is called interval dependency. Special treatment is required to obtain a tight enclosure. This is done by using *inclusion isotonicity* property of Interval arithmetic (Moore *et*, *all*, 2009, Neumaier 1990).

In other words, given a function $f = f(x_1,\ldots, x_n)$ of several variables, the precise range of values taken by f as x_i varies through given intervals $\boldsymbol{x}_i$ is introduced in the form

$$f(\mathbf{x}_1,\cdots,\mathbf{x}_n) = \{f(x_1,\cdots,x_n) : x_1 \in \mathbf{x}_1,\cdots,x_n \in \mathbf{x}_n\} \tag{29}$$

Usually, centered forms are used to reduce overestimation due to dependency of the enclosure of f (Moore

1979, Neumaier 1990). In the present work, the boundary value form (Neumaier, 1990) is used to evaluate the enclosure of function given in Equation (24). The following boundary value form has been suggested by Professor Neumaier during private communication for an enclosure of the function in Equation (28)

$$\begin{aligned}\delta\mathbf{F}_{mx}(d_1,d_2) = {} & [(m_{mx}+c_{mx1})d_{01}+c_{mx2}d_{02}]+(c_{mx3}d_{01}+c_{mx6}d_{02})\mathbf{d}_1^2+\\ & (c_{mx5}d_{01}+c_{mx4}d_{02})\mathbf{d}_2^2+(\mathbf{d}_1-d_{01})[(m_{mx}+c_{mx1})+c_{mx3}\mathbf{d}_1+c_{mx5}\mathbf{d}_2^2]+\\ & (\mathbf{d}_2-d_{02})(c_{mx2}+c_{mx6}\mathbf{d}_1^2+c_{mx4}\mathbf{d}_2^2)\end{aligned} \tag{30}$$

taking the intersection of the two results computed:

a. with lower bounds in place of d_{01} and d_{02}
b. with upper bounds in place of d_{01} and d_{02}

4.2. STOPPING CRITERIA

Two stopping criteria can be used. The first criterion is deterministic and it is straightforward requiring that

$$\frac{\|\delta M\|_2}{\|M\|_2} \le \delta, \tag{31}$$

where δ is a small specified value. The second criterion is the new containment stopping criterion, which is intrinsic to interval arithmetic. In the deterministic version of modified Newton-Raphson method, the usual stopping criterion is to continue the iteration procedure until the resisting internal forces are equal to the applied loads at each node. In the interval version, the applied loads and internal resisting forces are both intervals and the goal of the iteration is that the interval resisting forces to evolve until become equal to the applied ones. Due to dependency and resulting overestimation it is very difficult to capture such moment of equality between the interval applied loads and interval internal forces. As a matter of fact, during the iteration procedure, the difference between the values of interval internal forces and interval applied loads continues to become smaller converging from one side until a certain stage where one bound of the interval internal forces switches side and contains the interval applied load. A verification of the results of the iteration when the containment occurs shows that the correct solution is indeed obtained. To observe that '*the solution is reached when the evolved interval internal resisting forces contain the interval applied loads'* makes a complete engineering sense. Figure 2 illustrates the containment stopping criterions presented in terms of stress-strain instead of load-displacement.

5 Example Problems

Two example problems are chosen to illustrate the applicability of the present interval approach to handle material nonlinearity in case of truss problems. These examples are chosen to demonstrate the ability of the

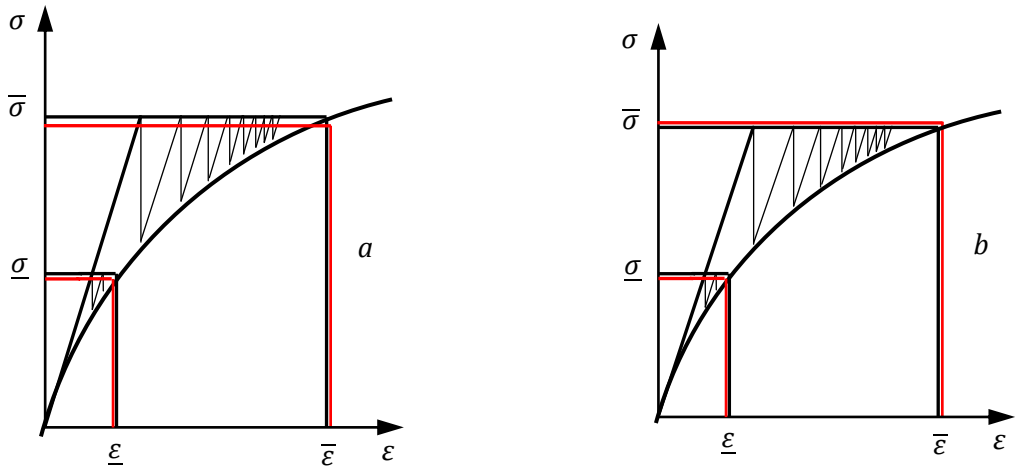


Figure 2. Stress-strain relationship, Modified Newton-Raphson method. *a*) before containment, *b*) after containment

current approach to obtain sharp bounds to the displacements and forces even in the presence of large uncertainties and large number of interval variables. It is assumed that for each element in the structure the constitutive relationship is given in the following cubic functional form for both examples

$$\sigma = E\varepsilon - 10^{13.6}\varepsilon^3, \tag{32}$$

where E is the modulus of elasticity. This relationship is shown in the Figure 3.

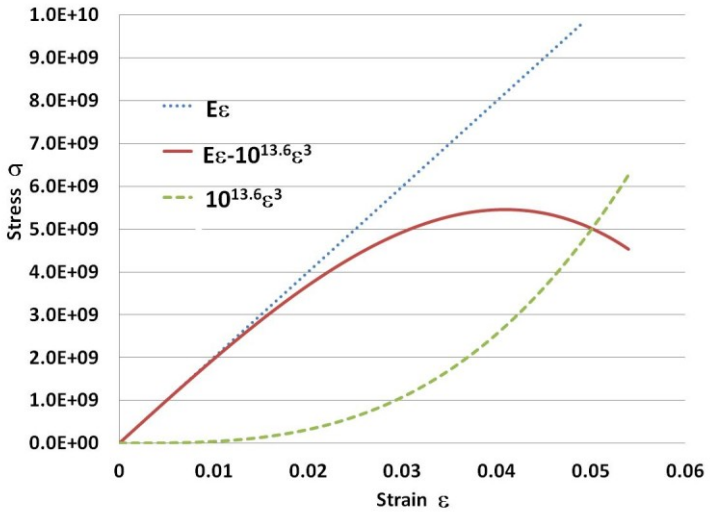


Figure 3. Stress-strain relationship given in Equation (32)

The two example problems are solved for various levels of interval widths of the loads centered at their nominal values. All interval variables are assumed to vary independently. Solution procedures outlined in sections 3 and 4 are based on the current values of member strains ε. These strains can be obtained using three different approaches viz. solution using modified Newton-Raphson method, the secant method and combinatorial approach. The first and second approaches compute member strains with same level of accuracy as displacements. In the third approach, member strains are computed combinatorially in each iteration.

The first example chosen is a five bar truss (Rama Rao, Mullen and Muhanna, 2011) as shown in Figure 4. The truss is subjected to a nominal point load of 200 kN at the node 2 in the horizontal direction to the right and a nominal point load of 200 kN at the node 3 in the vertically downward direction. The Young's modulus of each element is $E_i = 2 \times 10^{11}$ N/m^2, i = 1,5 while the cross sectional area is 1.0×10^{-4} m^2.

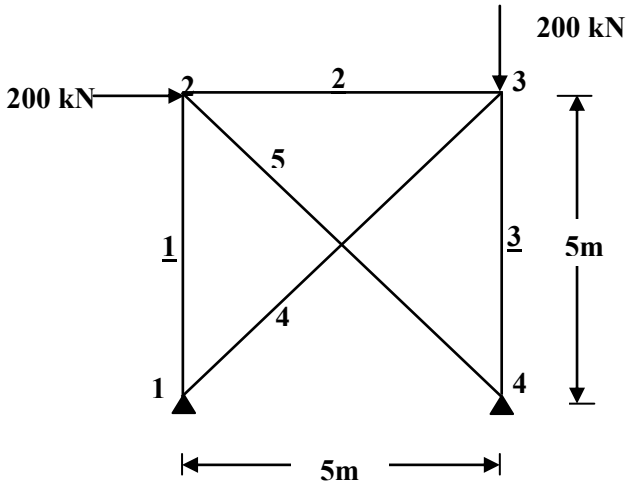


Figure 4. Five bar truss

Tables I, II and III show the computed values of selected displacements (horizontal displacement U_2 at node 2 and vertical displacement V_3 at node 3) and selected strains (strains ε_1 and ε_3 in elements 1 and 3) using the above approaches. Load uncertainties considered in Tables I, II and III are 1%, 10% and 25% (± 0.5%, ± 5% and ± 12.5% respectively about the mean value of load). Overestimation involved in results using the modified Newton-Raphson approach and secant approach is evaluated by comparing the corresponding solutions obtained with the combinatorial approach. Percentage error in the lower and upper bounds of the present solution is computed with reference to the corresponding bounds of the combinatorial solution. It is observed from these tables that error in bounds is quite small for displacements and strains (U_2, V_3, ε_1 and ε_3).

It is observed that the errors in strains (secondary unknowns) are numerically comparable with the error of displacements (primary unknowns). Thus, the present approach succeeds in obtaining the same level of sharpness for primary and derived quantities. This observation holds true at larger values of uncertainty as it will be seen in Tables II and III.

Table I Five bar truss − displacements for 1% uncertainty the load

Method	$U_2 \times 10^1$ (m)		$V_3 \times 10^2$ (m)		$\varepsilon_1 \times 10^3$		$\varepsilon_3 \times 10^2$	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Combinatorial	1.30644	1.32011	-6.92295	-6.84845	6.77652	6.84633	-1.38459	-1.36969
Secant	1.30643	1.32012	-6.92303	-6.84833	6.77554	6.84730	-1.38460	-1.36966
Error%	0.0008	0.0008	0.0012	0.0018	0.0145	0.0142	0.0007	0.0022
Newton	1.30541	1.32113	-6.93144	-6.83991	6.76784	6.85500	-1.38628	-1.36798
Error%	0.0784	0.0774	0.1227	0.1247	0.128	0.127	0.123	0.125

Table II Five bar truss − displacements for 10% uncertainty the load

Method	$U_2 \times 10^1$ (m)		$V_3 \times 10^2$ (m)		$\varepsilon_1 \times 10^3$		$\varepsilon_3 \times 10^2$	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Combinatorial	1.24516	1.38184	-7.26069	-6.51559	6.46303	7.16125	-1.45214	-1.30312
Secant	1.24475	1.38195	-7.26160	-6.51138	6.45189	7.17150	-1.45232	-1.30227
Error%	0.0329	0.0080	0.0125	0.0646	0.1724	0.1431	0.0124	0.0652
Newton	1.23489	1.39212	-7.34664	-6.43065	6.37602	7.24826	-1.46932	-1.28613
Error%	0.8247	0.7444	1.1838	1.3036	1.346	1.215	1.184	1.304

Table III Five bar truss − displacements for 25% uncertainty the load

Method	$U_2 \times 10^1$ (m)		$V_3 \times 10^2$ (m)		$\varepsilon_1 \times 10^3$		$\varepsilon_3 \times 10^2$	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Combinatorial	1.14376	1.48561	-7.83351	-5.96876	5.94287	7.68894	-1.56670	-1.19375
Secant	1.14159	1.48600	-7.83605	-5.94566	5.90929	7.71708	-1.56721	-1.18913
Error%	0.1897	0.0263	0.0324	0.3870	0.5650	0.3660	0.0326	0.3870
Newton	1.11797	1.51263	-8.05715	-5.75784	5.72444	7.91333	-1.61143	-1.15156
Error%	2.2544	1.8189	2.8550	3.5337	3.675	2.918	2.855	3.534

Figure 5 shows the computed interval values of horizontal displacement U_2 at node 2. The figure depicts the variation of the widths of the modified Newton-Raphson, the secant and the combinatorial approaches with the variation of load from its mean value. It is observed from this figure that the solutions computed using tangent and secant approaches enclose the combinatorial solution at all values of variation from 0 percent to 25 percent. A similar behavior is observed in the plot for variation of width of vertical displacement V_3 at node 3 in Figure 6. Figure 7 and 8 show the variation of strains in members 1 and 3 with the variation of uncertainty of load. *It is observed from these figures that the present solution using modified Newton-Raphson approach and secant approach enclose the combinatorial solution at all levels of uncertainty.*

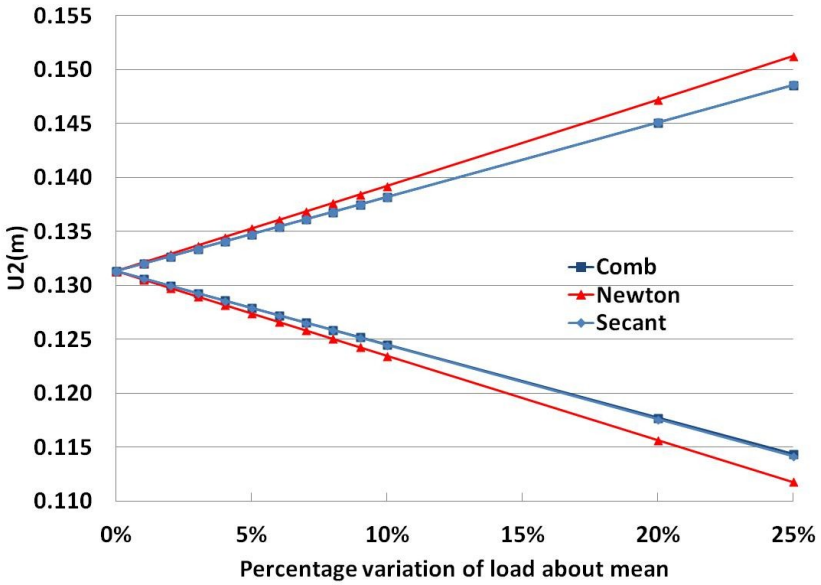


Figure 5. Five bar truss – variation of horizontal displacement at node 2 with uncertainty of load

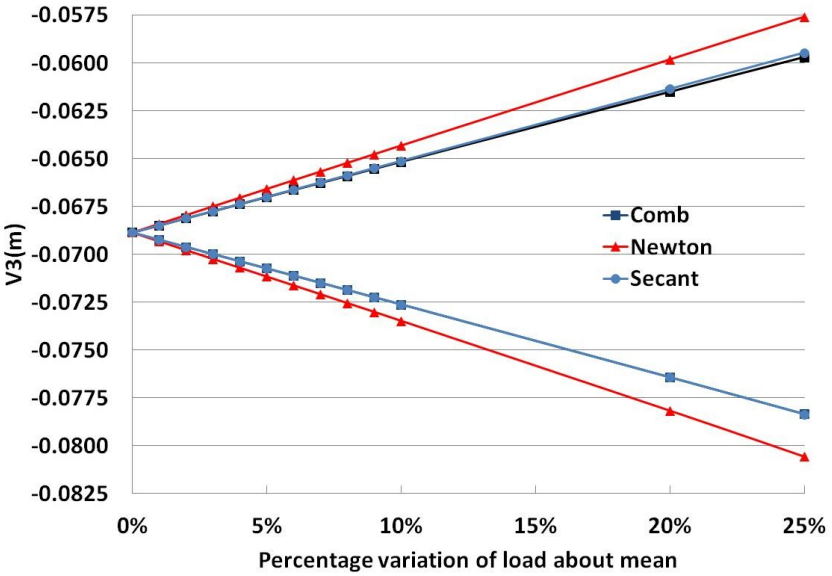


Figure 6. Five bar truss – variation of vertical displacement of node 3 with uncertainty of load

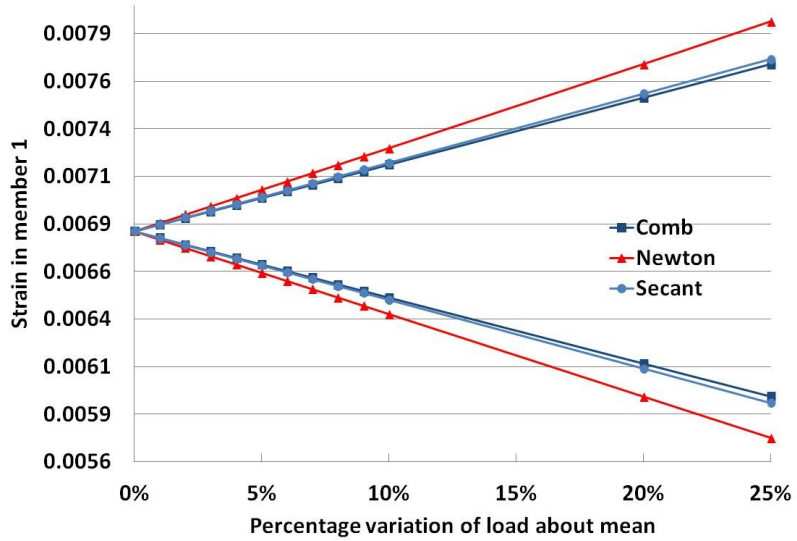


Figure 7. Five bar truss – variation of strain in member 1 with uncertainty of load

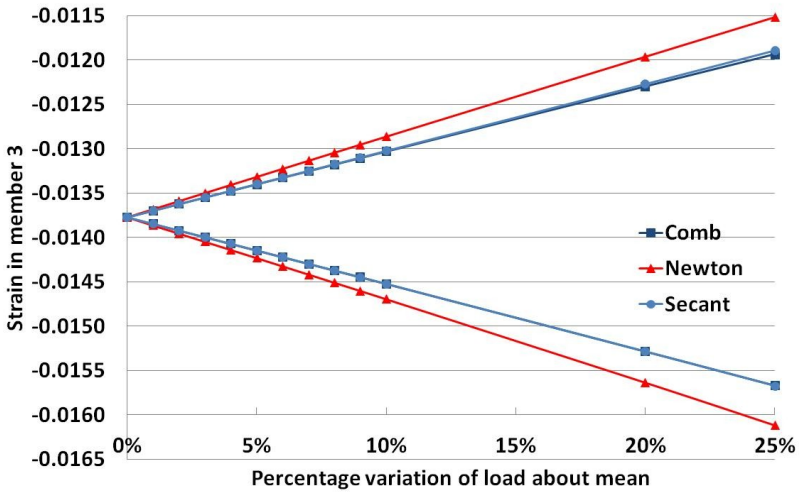


Figure 8. Five bar truss – variation of strain in member 3 with uncertainty of load

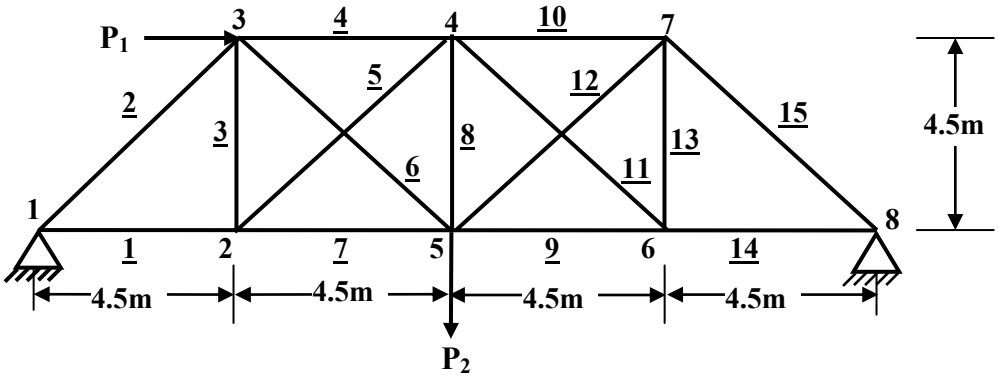


Figure 9 Fifteen-bar truss

The fifteen bar truss shown in Figure 9 is subjected to a horizontal point load $P_1 = 150$ kN acting to the right and vertical point load of $P_2 = 250$ kN acting downwards applied at the joints 3 and 5 respectively. Cross section areas of elements 1, 2, 3, 13, 14 and 15 are 10.0×10^{-5} m^2 while for the rest of the elements is the cross sectional area is 6.0×10^{-5} m^2. The deterministic value of Young's modulus of each element is $E_i = 2 \times 10^{11}$ N/m^2, $i = 1, 2, \ldots 15$. Results are computed using combinatorial approach, secant modulus approach and modified Newton-Raphson approach. Tables IV and V present the selected values of displacements and strains at load uncertainties of 1% and 10% respectively. It is observed from these tables that the displacements and strains computed using the modified Newton-Raphson approach and secant modulus approach are sharply enclosing the corresponding values computed using combinatorial solution at all levels of uncertainty. Figures 10 and 11 shows the plot of strain in members 2 and 8 computed for various levels of uncertainty from 0% to 10%. Figures 12 and 13 show the plots of horizontal displacement at node 3 and vertical displacement at node 5 respectively, computed for various levels of uncertainty from 0% to 10% .It is observed in all these figures that the solution computed using secant modulus approach as well as modified Newton-Raphson approach enclose the combinatorial solution at all levels of uncertainty.

Table IV Fifteen bar truss – Selected values of displacements and strains for 1% uncertainty in load

Method	U_3 (m)		$V_5 \times 10^1$ (m)		$\varepsilon_2 \times 10^3$		$\varepsilon_8 \times 10^2$	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Combinatorial	1.38087	1.39857	-3.52317	-3.48189	-6.29427	-6.17663	1.25047	1.26541
Secant	1.38079	1.39863	-3.52323	-3.48179	-6.29427	-6.17662	1.25026	1.26561
Error%	0.0058	0.0043	0.0017	0.0029	0.000	0.0002	0.0168	0.0002
Newton	1.37204	1.40475	-3.53113	-3.47009	-6.37471	-6.09618	1.24730	1.26716
Error%	0.6395	0.4419	0.2259	0.3389	1.2780	1.3025	0.2535	0.1383

Table V Fifteen bar truss – Selected values of displacements and strains for 10% uncertainty in load

Method	U_3 (m)		$V_5 \times 10^1$ (m)		$\varepsilon_2 \times 10^3$		$\varepsilon_8 \times 10^2$	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Combinatorial	1.30352	1.48091	-3.71321	-3.29977	-6.82498	-5.64853	1.18457	1.33427
Secant	1.29957	1.48158	-3.71391	-3.29394	-6.82499	-5.64678	1.18055	1.33637
Error%	0.3030	0.0452	0.0189	0.1767	0.0001	0.0310	0.3394	0.1574
Newton	1.22645	1.55402	-3.80985	-3.19791	-7.63374	-4.84061	1.15888	1.35789
Error%	5.9125	4.9368	2.6026	3.0869	11.8500	14.3032	2.1687	1.7703

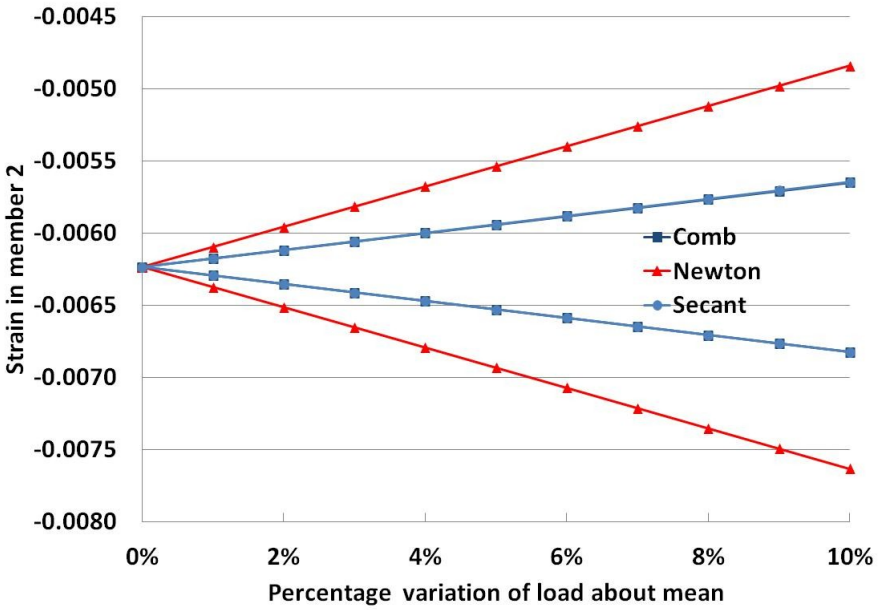


Figure 10. Fifteen bar truss – Variation of strain in member 2 with uncertainty of load

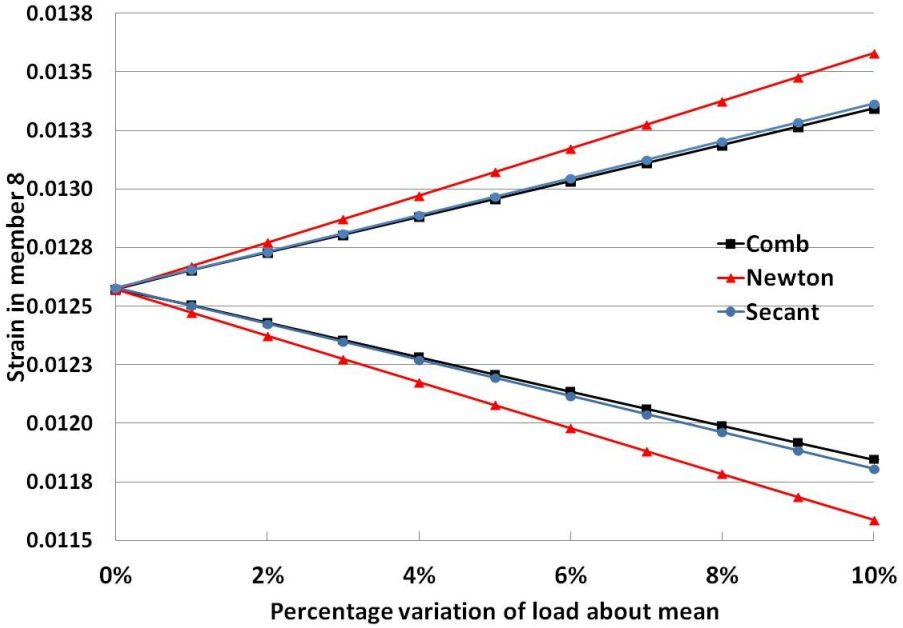


Figure 11. Fifteen bar truss – Variation of strain in member 8 with uncertainty of load

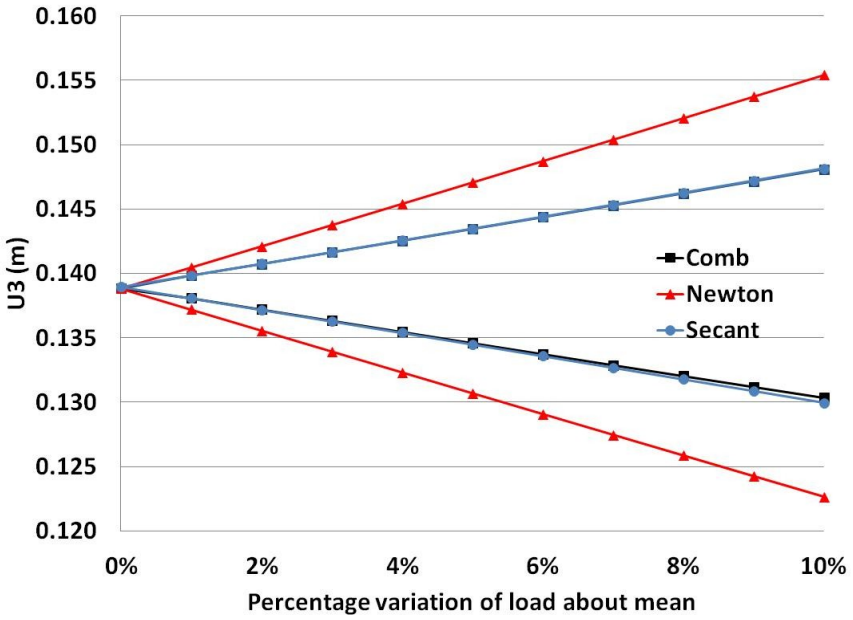


Figure 12. Fifteen bar truss – Variation of horizontal displacement at node 3 with uncertainty of load

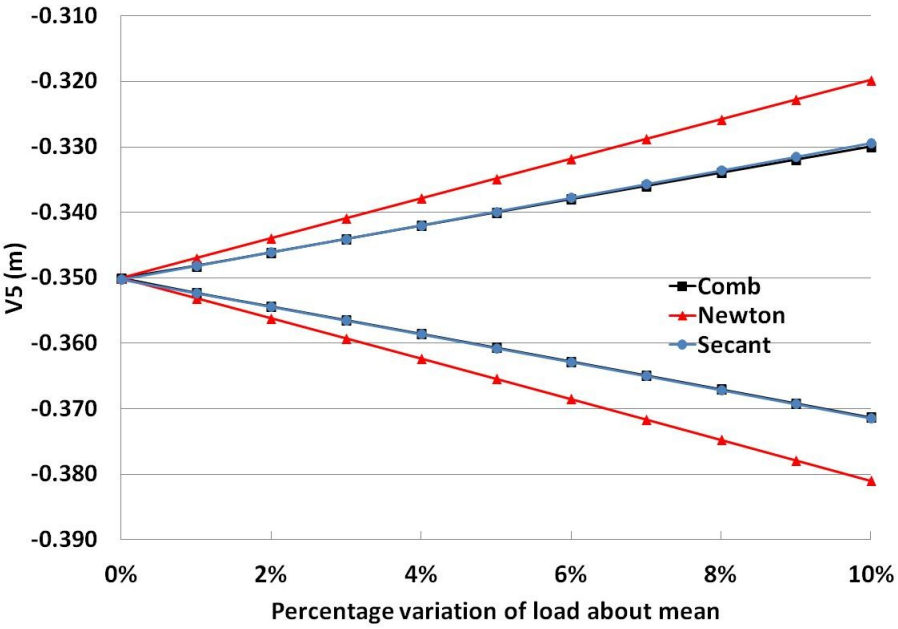


Figure 13. Fifteen bar truss – Variation of vertical displacement at node 5 with uncertainty of load

Conclusion

A Non-linear Interval Finite Element Method (NIFEM) for structural analysis is presented. Uncertainty in the applied load is represented as interval numbers and material nonlinearity is considered. The presented methods are an interval extension of the well known modified Newton-Raphson and the secant methods. Example problems illustrate the application of the methods to truss problems with large uncertainties. A new containment stopping criterion, which is intrinsic for intervals, has been introduced. The computational cost of the extension to interval numbers in both methods is comparable to the additional cost associated with introducing interval values into linear problems (Muhanna, *et al.*, 2007). Further study of non-linear interval finite elements methods for the refinement for different nonlinear material models is still needed to provide a more complete understanding of nonlinear interval finite element methods.

References

Cook, R. D., Malkus, D. S., Plesha, M. E., and Witt, R. J.2002. Concepts and Applications of Finite Element Analysis, 4th edition, Wiley.

Dessombz, O., Thouverez, F., Laîné, J.-P., and Jézéquel, L. 2001. Analysis of mechanical systems using interval computations applied to finite elements methods." *J. Sound. Vib.*, 238(5), 949-968.

Ganzerli, S. and Pantelides, C. P. 1999. Load and resistance convex models for optimum design, *Struct. Optim.*, 17, 259-268.

Köylüoglu, U., Cakmak, S., Ahmet, N., & Soren, R. K. 1995. Interval Algebra to Deal with Pattern Loading and Structural Uncertainty. *Journal of Engineering Mechanics* 121(11): 1149- 1157.

McWilliam, S. 2000. Anti-optimisation of uncertain structures using interval analysis, *Comput. Struct.*, 79, 421-430.

Möller, B. and Beer, M. 2008. Engineering computation under uncertainty capabilities of non-traditional models. *Comput. Struct.*, 86(10), 1024-1041.

Moore, R. E., 1979. *Methods and applications of interval analysis*, SIAM, Philadelphia.

Moore, R. E., Kearfott, R. B., Cloud, M. J., 2009. Introduction to Interval analysis, SIAM.

Muhanna, R. L. & Mullen, R. L. 1995. Development of Interval Based Methods for Fuzziness in Continuum Mechanics. In Proceedings of *ISUMA-NAFIPS'95*: *17-20 September,1995. IEEE* .

Muhanna, R. L. & Mullen, R. L. 2001. Uncertainty in Mechanics Problems—Interval-Based Approach, *Journal of Engineering Mechanics* 127 (6): 557–566.

Muhanna, R. L., Zhang, H., & Mullen, R. L. 2007. Interval finite element as a basis for generalized models of uncertainty in engineering mechanics. *Reliable Computing*, 13(2), 173–194.

Nakagiri, S. & Yoshikawa, N. 1996. Finite Element Interval Estimation by Convex Model. *In Proceedings of 7th ASCE EMD/STD Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability, WPI, MA, 7-9 August 7-9, 1996.*

Neumaier, A. ,1990. *Interval methods for systems of equations*, Cambridge University Press.

Neumaier, A. & Pownuk, A. 2007. Linear Systems with Large Uncertainties, with Applications to Truss Structures, *Reliable Computing*, 13(2): 149-172.

Pownuk, A., 2004. Efficient method of solution of large scale engineering problems with interval parameters." *Proc. NSF workshop on reliable engineering computing* (REC2004), R. L. Muhanna and R. L. Mullen, eds., Savannah, GA, USA.

Rama Rao, M. V., Mullen, R. L., Muhanna, R. L., 2011, "A New Interval Finite Element Formulation With the Same Accuracy in Primary and Derived Variables", International Journal of Reliability and Safety, Vol. 5, Nos. 3/4.

Qiu, Z. and Elishakoff, I. 1998. Antioptimization of structures with large uncertain-but-non-random parameters via interval analysis, *Compt. Meth. Appl. Mech. Engrg.*, 152, 361-372.

Rao, S. S. & Sawyer, P. 1995. Fuzzy finite element approach for analysis of imprecisely defined systems. *AIAA J.*, 33(12): 2364–2370.

Rao, S. S. and Berke, L. 1997. Analysis of uncertain structural systems using interval analysis. *AIAA J.,* 35(4): 727–735.

Rao, S. S. & Chen, Li. 1998. Numerical solution of fuzzy linear equations in engineering analysis. *International Journal of Numerical Methods in Engineering.*,43: 391–408.

Rump, S.M. INTLAB - INTerval LABoratory. In Tibor Csendes, editor, Developments in Reliable Computing, pages 77-104. Kluwer Academic Publishers, Dordrecht, 1999.

Sun microsystems 2002. *Interval arithmetic in high performance technical computing*. Sun microsystems. (A White Paper).

Zhang, H. 2005. Nondeterministic Linear Static Finite Element Analysis: An Interval Approach. *Ph.D. Dissertation, Georgia Institute of Technology, School of Civil and Environmental Engineering*.

Estimating load condition having caused structure failure and an optimal design taking account of the estimated result

T. Muromaki[1)], K. Hanahara[2)], Y. Tada[2)], S. Kuroda[3)] and T. Fukui[3)]
[1)] *Graduate School of Engineering, Kobe University, Kobe, 657-8501, Japan*
email: muromaki@opt.cs.kobe-u.ac.jp
[2)] *Graduate School of System Informatics, Kobe University, Kobe, 657-8501, Japan*
email: hanahara@cs.kobe-u.ac.jp, tada@cs.kobe-u.ac.jp
[3)] *Martec K. K., Kobe, 650-0046, Japan*
email: kuroda@martec.co.jp, fukui@martec.co.jp

Abstract: In order to improve a design of structure, it is important to know the actual load condition of failed structure. We develop an estimation method of loading conditions based on images of failed structures and an FEM analysis model. Preparing a database that consists of deformation data of the structure corresponding to various load conditions, our system is able to estimate the load conditions that caused structure failure based on the processed images of failed structure samples. Adopting elasto-plastic model of the structure, the magnitude of the load having caused the failure is also estimated in addition to the position and orientation of the critical load. We adopt the EM algorithm to obtain the distribution of the critical load. An optimal design problem that takes account of the distribution of the estimated critical load condition is formulated as a minimization problem with a multi-objective function; the stiffness and the structural weight are also adopted as the evaluation items that make up the objective function. The particle swarm optimization (PSO) is adopted as the optimization algorithm. The approach is applied to crane-hook. The result of estimated critical load distribution and the optimal design based on the load distribution are demonstrated.

Keywords: load estimation; optimal design; database; finite element analysis; EM algorithm; crane-hook.

1. Introduction

Avoiding failure of structure system is one of the most important missions for design engineer of structures. In order to improve an existing structure so that it does not fail, it is important to know the load condition that causes structure failure. Generally, the load condition is identified by integrating the information obtained from the sensor devices; continuous monitoring is essential. However, almost all structures themselves have no information about the load conditions during their service life. In this case, several failure detection methods proposed in the past are not applicable (Quek et al., 2009; Lam and Ng, 2009). It is necessary to estimate the load condition by means of another approach. We develop a load estimation system; this system is applicable to the failed structures having permanent deformation. The system inputs the digital images of failed structures and outputs the estimated probability of the load conditions that caused the failure. The information from the sensor devices is not required. We deal with the failed crane-hooks as a concrete example.

Crane-hook is one of the most useful equipments for suspension work. Recently, excavators having a

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

crane-hook are widely used in construction work sites. One reason is that there are work sites where the crane trucks for suspension work are not available because of the narrowness of the working site; an excavator has superior to a crane truck in general. Another reason is that such an excavator is convenient because they can perform the conventional digging tasks as well as the hanging works mentioned above. Figure 1 shows a sample of excavator with crane-hook and the close-up image of its bucket part where the crane-hook is attached.

Though it has such a convenience, there are cases that the crane-hooks are damaged during some kind of hanging works. Figure 2 shows a typical crane-hook and its damaged sample to be repaired. This type of hook can be used to suspend objects whose weight is up to 2.9t. In Fig. 2, we can see that the locking apparatus, called latch, is left open. From the view point of safety, such failure must be avoided. Improvement of the performance and the service life is important; the real conditions of such suspension tasks in practical environment are, however, still unclear. Therefore, the identification of the cause of failure is one of the key issues for the safety improvement.

Our previous work (Muromaki et al., 2012) gives the estimation of such load conditions in the form of probability distribution. In the study, however, we focused only on the estimation of the load applied position and the direction. We did not estimate the load magnitude. This is because the analysis model is based on the linear deformation theory and the analysis results do not well reflect the significant magnitude of the applied load that causes the permanent deformation as shown in Fig. 2. In the current study, the analysis model is improved so as to be applicable for the permanent deformation. By adding the information of the load magnitude, the failure estimation result becomes more meaningful.

In addition to the load estimation approach, we discuss an optimal design taking account of the estimated result. In order to improve the performance of crane-hook, we formulate the multi-objective optimization problem considering the structural weight and the stiffness. The evaluation of stiffness is performed in terms of the estimated load conditions. The optimization problem is solved by using the particle swarm optimization (PSO).

The outline of this paper is as follows. In section 2, we construct an FEM model of crane-hook and introduce the Load-Deformation (L-D) database. This database is prepared by using the FEM model; it is constructed as a collection of the applied load conditions and the corresponding deformed node positions. In section 3, we explain the image processing procedure to detect the feature points from the failed crane-hook images. In section 4, we develop the identification method of load condition based on our evaluation criterion. We introduce the EM algorithm for the arrangement of the identification results. In section 5, we apply our estimation approach to the actual failed crane-hooks. The estimation results are represented by the form of probability distribution. In section 6, we deal with a design approach that is based on estimated load condition obtained by means of examination of the failed structures. Some concluding remarks are expressed in section 7.

2. Crane-hook Model and Load-deformation Database

2.1. CRANE-HOOK MODEL

We construct a finite element model of the crane-hook based on one of its actual designs. Figure 3 shows the design drawing of the crane-hook adopted as the reference. Its cross-sectional shapes are illustrated by

Figure 1. Excavator with crane-hook and its close-up.

Figure 2. Typical crane-hook and failed sample.

the shaded area at two positions. One is the lowest center position "D" where the load is applied in the typical suspension work. The other is the position "C" where the largest stress occurs in the typical work. This area is usually called "critical section". These cross-sectional shapes, called "T shape", have been achieved by expert engineers empirically. Figure 4 shows the conceptual finite element model based on one dimensional beam element and the constructed model of crane-hook based on the actual design shown in Fig. 3. As indicated in Fig. 4, the latch part is omitted in the adopted model because it does not contribute to support the applied load. Each element is constructed by N_d layered as shown in Fig. 4(b). The height of cross-section is indicated by the variable "h". The height of each layer is assigned evenly. The width of each layer is specified by "b_i" ($i = 1, \cdots, N_d$). By changing these widths b_i, we can represent various cross-sectional shapes. The analysis model is constructed of N_e elements.

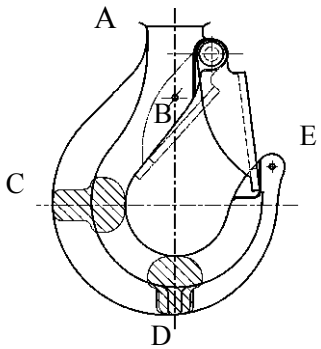


Figure 3. Design drawing of crane-hook.

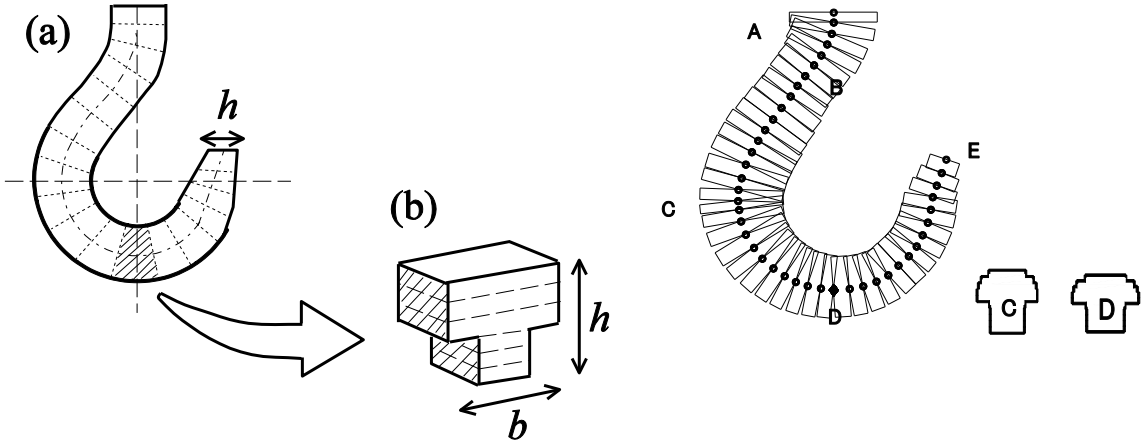


Figure 4. FE model based on 1-D beam element and constructed model referring the design drawing.

In the linear elastic deformation analysis, the equilibrium equation is obtained by means of the conventional analysis approach and expressed as

$$\boldsymbol{F} = \boldsymbol{K}\boldsymbol{U} \tag{1}$$

where $\boldsymbol{F}$, $\boldsymbol{K}$ and $\boldsymbol{U}$ are the external force vector, the stiffness matrix and the displacement vector. Given a specified external force vector and boundary conditions, the corresponding deformation of the structure is calculated on the basis of Eq. (1). In order to adapt the FEM model to the permanent deformation, we introduce an elasto-plastic deformation analysis. In our analysis, the stress-strain relationship of the material is approximated by a piecewise linear function as shown in Fig. 5. In this figure, E_1 is the Young's modulus, E_2 is the tangent modulus and $\bar{\sigma}$ is the yield stress. The dashed line indicates the relationship in the unloading process; the tangent modulus is assumed to be equal to the Young's modulus E_1. In order to calculate the displacement of finite element model, we utilize the incremental solution scheme (Crisfield, 1991). The incremental formulation is expressed as

$$\Delta\boldsymbol{F} = \boldsymbol{K}_t \Delta\boldsymbol{U} \tag{2}$$

where $\Delta\boldsymbol{F}$, $\boldsymbol{K}_t$ and $\Delta\boldsymbol{U}$ are the incremental force vector, the tangent stiffness matrix and the incremental displacement vector, respectively. The tangent stiffness matrix $\boldsymbol{K}_t$ takes over the role of the stiffness matrix in elastic analysis. It relates small change in force to small change in displacement. The matrix $\boldsymbol{K}_t$ takes the form

$$\boldsymbol{K}_t = \boldsymbol{K}_t(\boldsymbol{U}),\ \boldsymbol{K}_t(\boldsymbol{0}) = \boldsymbol{K} \tag{3}$$

where $\boldsymbol{K}$ is the elastic stiffness matrix used in the linear elastic analysis as Eq. (1). The total displacement is computed by the sum of the incremental displacements.

$$\boldsymbol{U} = \sum_j \Delta \boldsymbol{U}_j \tag{4}$$

In the assessment process of the yielding, we utilize the layered approach (Owen and Hinton, 1980). In this approach the beam element is subdivided into layers, as shown in Fig. 4. A layer element is determined to be in yield state as a whole in the case that the central stress of the layer reaches the material yield stress. The stiffness values of the elements are determined according to the relationship shown in Fig. 5.

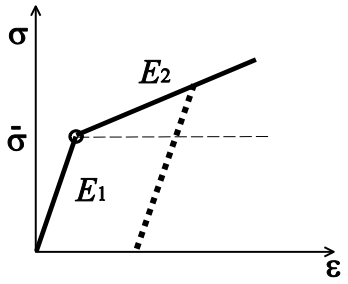


Figure 5. Adopted stress-strain relationship model.

2.2. ESTIMATION OF PHYSICAL PARAMETERS

As shown in Fig. 5, the stress-strain relationship has the three parameters: Young's modulus E_1, tangent modulus E_2 and yield stress $\bar{\sigma}$, which must be determined. We estimate them based on experimental data. The stretch experiment of crane-hooks conducted is as follows. A load is applied at the point "D" of crane-hook shown in Fig. 3. We measure the distance between points "B" and "E" at various load magnitude and calculate the enlargement of the distance. Figure 6 shows the obtained experimental result. The ordinate and the abscissa represent the magnitude of load [kN] and the enlargement of distance between point "B" and "E" [mm], respectively. On the basis of the obtained data, the plastic deformation begins to be observed at a load around 120 [kN]. The nominal load of the crane-hook dealt with in this study is 29 [kN], thus we can see that the nominal load is included in the elastic deformation area. In order to determine the material parameters based on the experimental data, we formulate a minimum square error problem as follows:

$$\text{Minimize} \sum_i (y_i - \hat{y}_i)^2 \quad \text{with respect to } E_1, E_2, \bar{\sigma} \tag{5}$$

In the above equation, y_i is the enlargement between "B" and "E" for the i th load magnitude in the result of stretch experiment. The symbol $\hat{y}_i$ indicates the enlargement obtained by the FEM analysis for the same load magnitude. Table I shows the range of these parameters and the results of the error-minimization. The results are obtained by the exhaustive search and are shown in the lowest row. In the following, we utilize these estimated values for the FEM analysis.

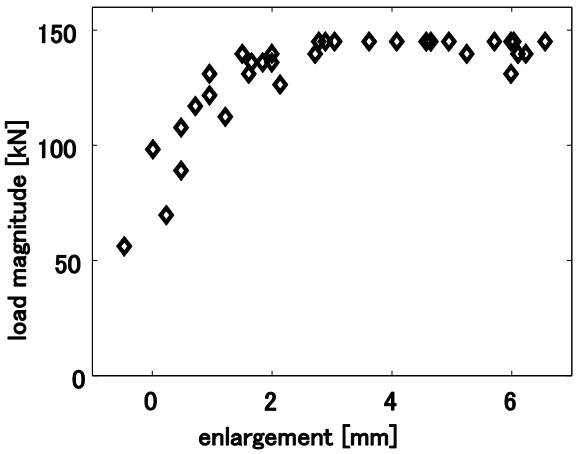


Figure 6. Relation between applied load and enlargement of displacement.

Table I. Range of parameters and results of error-minimization

	E_1 [GPa]	E_2 [GPa]	$\bar{\sigma}$ [MPa]
range	180 ~ 280	0.1 ~ 50	100 ~ 400
estimated	260	1	200

2.3. CONSTRUCTION OF L-D DATABASE

The Load-Deformation (L-D) database is a collection of data; the various load conditions and the corresponding information of deformation are recorded. In the current study, the L-D database is designed to have the following information obtained by the FEM analysis:

- analysis number
- applied load condition on the FEM model
  - load applied node
  - load magnitude
  - load direction
- deformed node positions

We explain the contents of L-D database concretely. Figure 7 shows the load applied nodes and load directions. The FEM model consists of 42 elements and 43 nodes. The load applied nodes are 9 nodes from 21th node to 37th node; the list of node is $[\,21, 23, 25, 27, 29, 31, 33, 35, 37\,]$. The load direction is considered 7 direction patterns from -180° (leftward) to 0° (rightward); -90° corresponds to the vertical direction. These direction vectors are defined in a global coordinate system. The load magnitude is prepared from 20[kN] to 140[kN] with an interval of 20[kN]. The ratio of these magnitudes to the nominal load is from 0.69 to 4.82. The load magnitude is 7 patterns and the list is [20, 40, 60, 80, 100, 120, 14] [kN]. The pattern of all combination is 441 and all calculation results are recorded in the L-D database as the form of deformed node positions. The estimation of the load condition is performed based on the deformed node positions in the L-D database.

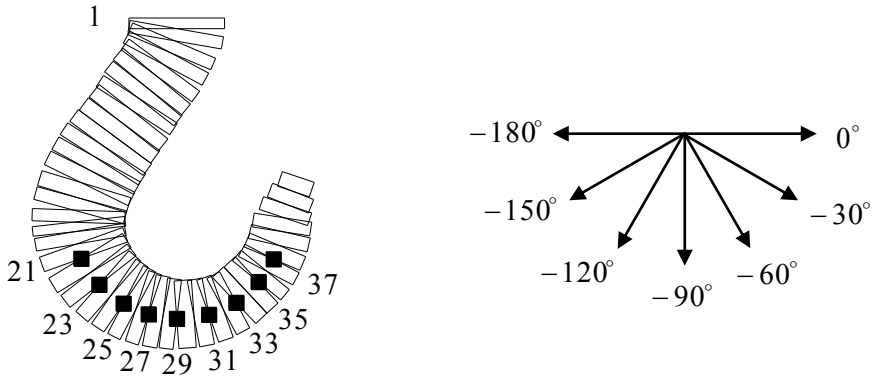


Figure 7. Load conditions; load applied nodes and load directions.

3. Feature Points Detection from Crane-hook Images

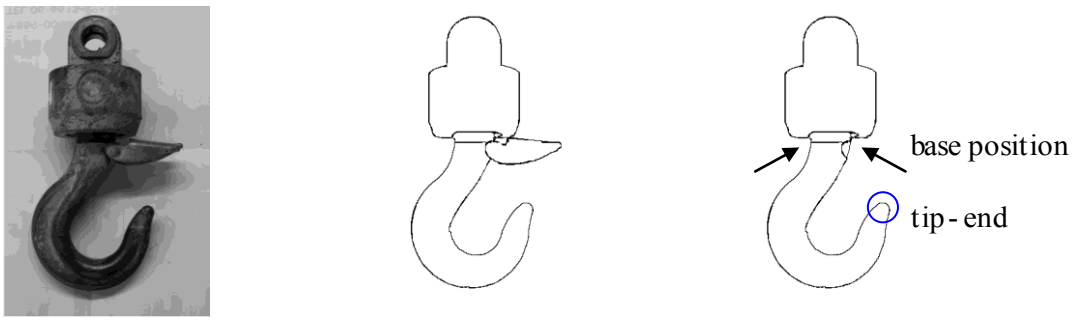
In this section, we explain the detection process of hook deformation from the failed hook image. The deformation of a failed crane-hook is represented based on the feature points detected from the failed image. The selection of feature points is performed by means of the digital image processing. This section consists of two parts. In the former part, we introduce the procedure of image processing used in this study. The outline of crane-hook is obtained from the failed hook image. In the latter part, we discuss the detection algorithm of feature points. The feature points are identified based on the obtained outline image.

3.1. Preprocessing of Crane-hook Image

The digital image of a failed hook is processed to obtain its shape outline. Figure 8(a) shows a typical example of failed hook image. This figure is displayed with a 256 step gray scale. Figure 8(b) is obtained by applying the digital image processing to Fig. 8(a). If there are some line gaps, we modify the outline through manual operations. Figure 8(c) shows the outline image without the latch part. The removing of latch part and the interpolation of line gap is performed by manual operations. In Fig. 8(c), the tip-end and the base positions are indicated by the circle and the arrow, respectively. The details are explained in our previous work (2012).

3.2. Detection Process of Feature Points

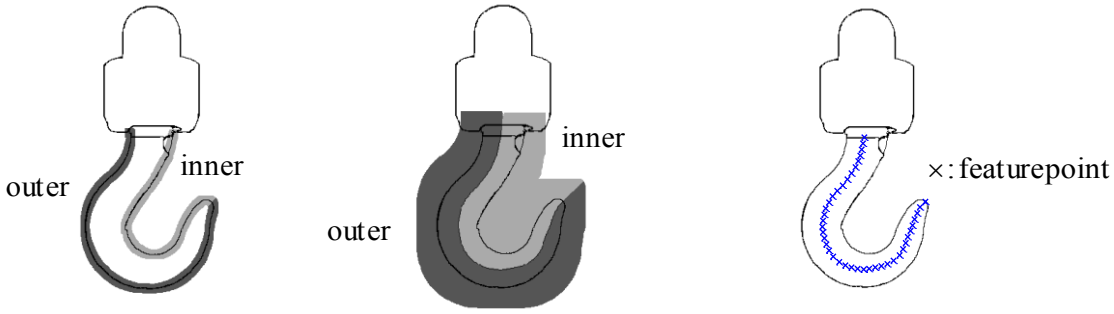
The detection process of feature points is consists of two steps. The first step is the determination of boundary line between the inner area and the outer area. The second step is the selection of the feature points on the boundary line. As the preparation of the first step, the outline is divided into two parts; one is the inner outline and the other is the outer outline. By following the outline from the tip-end point, the inner outline at the upper side and the outer outline at the lower side are determined. In order to determine the



(a) Gray scale image (b) Filtered outline image (c) Outline image (without latch part)

Figure 8. Processed images.

boundary line, we utilize the dilation process of the outlines. In each cycle of dilation process, the same label that distinguishes between the inner area and the outer area is assigned to the adjacent pixels of the outlines. Figure 9(a) shows the result of 5 cycles of dilation process. The outer outline is represented darker and the inner outline is represented lighter. By repeating the dilation process until the inner and outer areas collide, the boundary between the inner area and the outer area is obtained. Figure 9(b) shows the boundary line obtained by the dilation process. The feature points are detected on this boundary line. The boundary line is divided into N_f sections evenly. The division points are selected as the feature points and they are represented by the symbol '×' in Fig. 9(c).



(a) Dilation of outline (b) Boundary of inner and outer area (c) Detected feature points

Figure 9. Detection process of feature points.

4. Identification Method of Load Condition

The load condition of a failed crane-hook is identified by using the L-D database. In this section, we explain the criterion used in the identification process. In the latter part, we introduce the EM algorithm that is used in the load estimation.

4.1. GEOMETRIC MOMENT

Generally, the coordinate systems used in the FEM analysis do not necessarily coincide with the coordinates used in the image processing. The shape of crane-hook should be expressed in a shape representation that is insensitive to the coordinate transformation. In this study, we turn our attention to the geometric moment of the shape. The geometric moment plays important role in object recognition and shape analysis (Ghorbel et al, 2005). The shape of hook can be described quantitatively by using geometrical moments, such as the mean, variance, and higher-order moments. The mean corresponds to the geometrical center of the shape. *n*th moments are expressed as

$$m_{pq} = \frac{1}{M}\sum_{i=1}^{M}(x_i - \bar{x})^p (y_i - \bar{y})^q, \quad p+q=n, \tag{6}$$

where p and q are degrees of power. The pair of x_i and y_i represents the coordinates of the deformed node positions in the L-D database or those of the feature points in the failed hook image. These values are normalized in terms of the base length of crane-hook. The constant M is the number of FE nodes or the feature points on the boundary line. The quantities $\bar{x}$ and $\bar{y}$ are calculated as follows:

$$\bar{x} = \frac{1}{M}\sum_{i=1}^{M} x_i \quad , \quad \bar{y} = \frac{1}{M}\sum_{i=1}^{M} y_i \tag{7}$$

We represent the information of geometric moment as a vector form. The geometric moment vector is defined as

$$\boldsymbol{m} = \begin{bmatrix} m_{10} & m_{01} & \cdots & m_{0n} \end{bmatrix}^T. \tag{8}$$

The moment vector obtained from the node positions in the L-D database is represented as '${}^{LD}\boldsymbol{m}^{(i)}$', where the superscript (i) indicates the analysis number in the L-D database. The moment vector obtained from the feature points based on a failed hook image is represented as ${}^{IM}\boldsymbol{m}$.

4.2. IDENTIFICATION PROCESS OF LOAD CONDITION

The data search process in the L-D database is performed based on the following evaluation function:

$$e = ({}^{LD}\boldsymbol{m}^{(i)} - {}^{IM}\boldsymbol{m})^T \boldsymbol{W} ({}^{LD}\boldsymbol{m}^{(i)} - {}^{IM}\boldsymbol{m}) \tag{9}$$

where $\boldsymbol{W}$ is the weighting factor matrix. This function evaluates the similarity of the geometric moment between the deformed node positions and the feature points of the failed hook image. In this study, we evaluate the geometric moment up to the 3rd order. The weighting factor matrix $\boldsymbol{W}$ is expressed as

$$\boldsymbol{W} = \mathrm{diag}\begin{bmatrix} 10 & 10 & 1 & 1 & 1 & 0.1 & 0.1 & 0.1 & 0.1 \end{bmatrix}. \tag{10}$$

We attach importance to the lower order moment; these values are determined empirically. The

identification process is expressed as the following minimization problem:

$$\text{Minimize} \quad e \qquad \text{with respect to} \quad i, \tag{11}$$

where i is the analysis number of the L-D database. By finding the solution, we can know the applied load condition, that is, the load applied node, the load magnitude and the load direction.

4.3. EM ALGORITHM

A variety of load conditions can be obtained from the estimation process of load condition. In addition to the variety of load conditions, the quality of the estimation process can be uneven. In order to discuss the variety and the uncertainty together, we summarize the estimated results in the form of statistical representation. The representation is implemented in the form of probability distribution. In the implementation process, we utilize the EM algorithm (McLachan and Krishnan, 1997). EM (Expectation Maximization) is an iterative optimization method to estimate some unknown parameters $\boldsymbol{\Theta}$ from the given measurement data $\boldsymbol{\chi}$. However, we are not given some "hidden" nuisance variables $\boldsymbol{G}$, which need to be integrated out. We maximize the posterior probability of the parameters $\boldsymbol{\Theta}$ given the data $\boldsymbol{\chi}$, marginalizing over $\boldsymbol{G}$:

$$\boldsymbol{\Theta}^{*} = \arg\max_{\Theta} \sum_{\boldsymbol{G}} P(\boldsymbol{\Theta}, \boldsymbol{G} \mid \boldsymbol{\chi}) \tag{12}$$

We can search for a maximum of $P(\boldsymbol{\Theta}, \boldsymbol{G} \mid \boldsymbol{\chi})$ by means of the following algorithm:

1. step E: calculate $P(\boldsymbol{\Theta}^{\text{old}}, \boldsymbol{G} \mid \boldsymbol{\chi})$
2. step M: $\boldsymbol{\Theta}^{\text{new}} = \arg\max_{\Theta} \sum_{\boldsymbol{G}} P(\boldsymbol{\Theta}^{\text{old}}, \boldsymbol{G} \mid \boldsymbol{\chi}) \ln P(\boldsymbol{\Theta}, \boldsymbol{G} \mid \boldsymbol{\chi})$

In the current study, the probability density functions of the random variables are assumed to be a mixture of Gaussian distribution. This probability function consists of a linear combination of Gaussian distributions and is expressed as

$$P(\boldsymbol{\chi}) = \sum_{k=1}^{K} \pi_k N(\boldsymbol{\chi} \mid \boldsymbol{\mu}_k, \boldsymbol{V}_k), \quad 0 \le \pi_k \le 1, \quad \sum_{k=1}^{K} \pi_k = 1, \tag{13}$$

$$N(\boldsymbol{\chi} \mid \boldsymbol{\mu}, \boldsymbol{V}) = \frac{1}{(2\pi)^{D/2} |\boldsymbol{V}|^{1/2}} \exp\{-\frac{(\boldsymbol{\chi} - \boldsymbol{\mu})^{T} \boldsymbol{V}^{-1} (\boldsymbol{\chi} - \boldsymbol{\mu})}{2}\}, \tag{14}$$

where $\pi_k, \boldsymbol{\mu}, \boldsymbol{V}$ and D are the mixture weights, the mean, the variance and a number of dimension, respectively. The parameter vector $\boldsymbol{\Theta}$ is denoted by $\boldsymbol{\Theta} = [\pi_k, \boldsymbol{\mu}_k, \boldsymbol{V}_k]$. In this case, the EM algorithm is represented as follows:

1. initialize the parameters, π_k, $\boldsymbol{\mu}_k$ and $\boldsymbol{V}_k$
2. step E: calculate

$$\gamma(z_{nk}) = \frac{\pi_k N(\boldsymbol{\chi}_n \mid \boldsymbol{\mu}_k, \boldsymbol{V}_k)}{\sum_{j=1}^{K} \pi_j N(\boldsymbol{\chi}_n \mid \boldsymbol{\mu}_j, \boldsymbol{V}_j)} \tag{15}$$

3. step M: calculate

$$\boldsymbol{\mu}_k^{\text{new}} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) \boldsymbol{\chi}_n \tag{16}$$

$$\boldsymbol{V}_k^{\text{new}} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) (\boldsymbol{\chi}_n - \boldsymbol{\mu}_k^{\text{new}})(\boldsymbol{\chi}_n - \boldsymbol{\mu}_k^{\text{new}})^T \tag{17}$$

$$\pi_k^{\text{new}} = \frac{N_k}{N} \tag{18}$$

After the convergence of this algorithm, we can determine the probability distribution of load condition.

5. Estimation Results

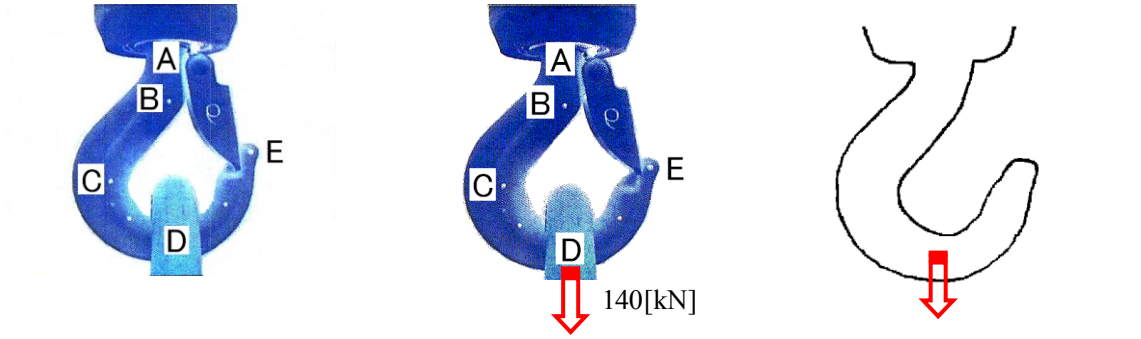
5.1. Identification of Load Condition Using Stretch Test Image

In order to confirm the effectiveness of the implemented estimation approach, we apply the method to a deformed hook image obtained by means of the stretch experiment. Figure 10(a) shows the initial state image of crane-hook. Figure 10(b) shows the deformed state image. A downward load (140 [kN]) is applied around the point 'D'. We can see that the contact point around the tip-end moves slightly leftward. Figure 10(c) shows the outline image of deformed hook. The latch part is excluded by manual operations. The feature points are detected from this image and the geometrical moment of this image is calculated based on the feature points. The solution of the minimization problem (11) is searched in the L-D database. The estimated load condition is shown in Fig. 11.

Comparing the estimated result with the actual applied load, the load applied point and the load direction of the estimated result shift leftward and the load magnitude is smaller than the applied load. Because of the width of loading device, the estimation problem has some difficulty for the proper identification of load condition. It should be noted that each of the estimated results is expressed to have the uncertainty to this level.

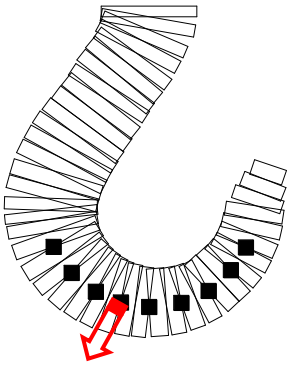
5.2. Sample Result of Load Estimation

In this part, we show a sample result of load identification. We use the failed hook shown in Fig. 8(a). The detected feature points are shown in Fig. 9(c). The identification result is shown in Fig. 12. The load applied position shifts rightward from the lower center. The load magnitude is three times greater than the nominal load. The load direction shifts rightward. The obtained result indicates that great rightward force is placed on the rightward position from the lower center.



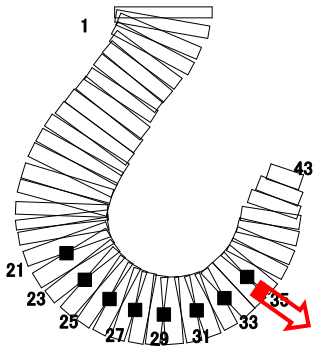
(a) Initial state (b) Deformed hook (c) Outline image of deformed hook

Figure 10. Crane-hook images of stretch experiment.



load applied node: 27
load magnitude: 100 [kN]
load direction: $\frac{4}{3}\pi$

Figure 11. Estimated loading condition using the stretch experiment image in Fig. 10.

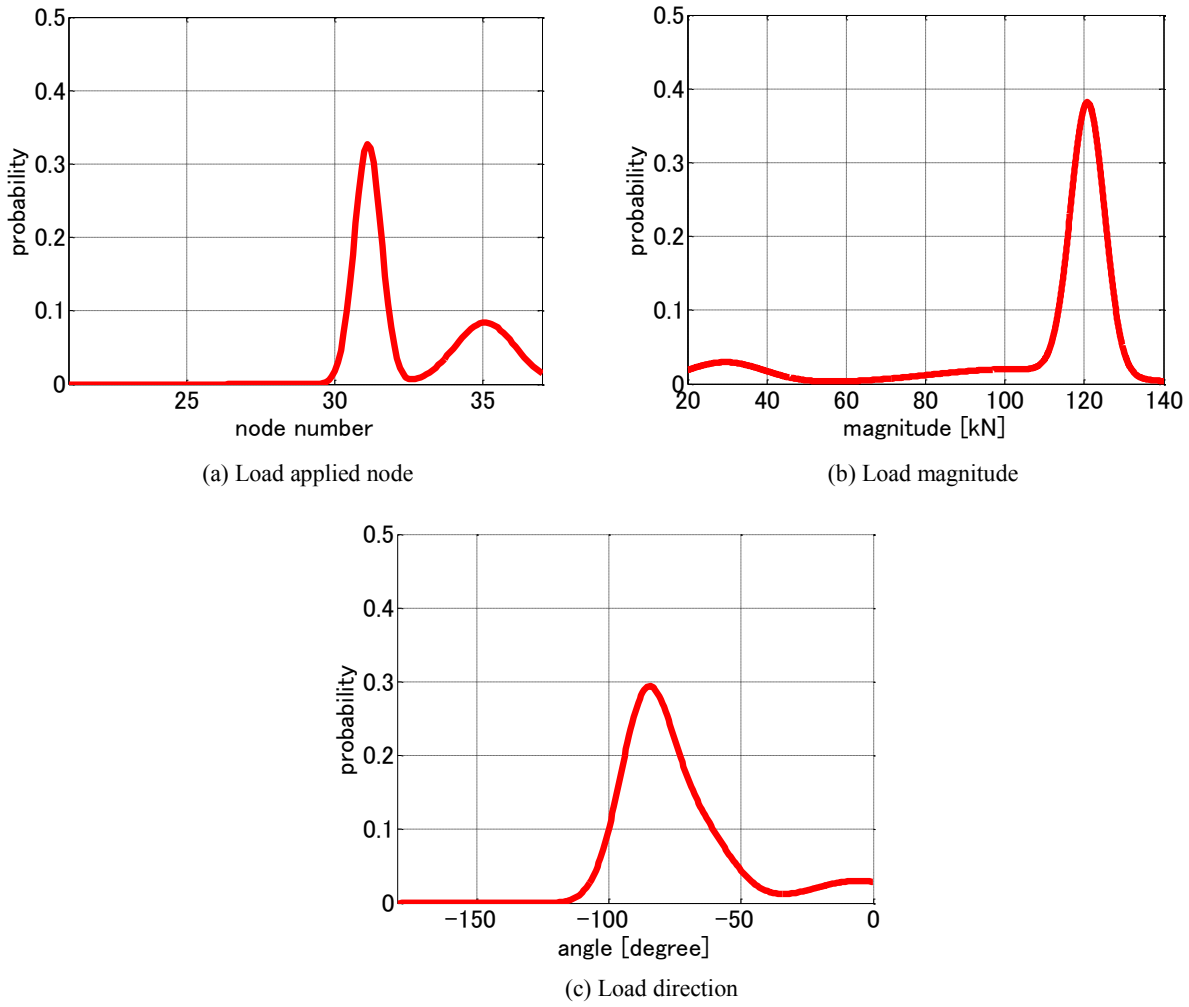


load applied node: 35
load magnitude: 100 [kN]
load direction: $\frac{5}{3}\pi$

Figure 12. Estimated loading condition using failed hook image (Fig. 8(a)).

5.3. IDENTIFICATION RESULT USING EM ALGORITHM

We apply the identification method to 12 failed samples. The probability distribution is obtained from the identified load conditions. In the current study, the number of samples is not enough to estimate the probability distribution on 3 load components (the load applied node, the load magnitude and the load direction) simultaneously. In this study, the probability distribution of each load component is treated independently. The probability distribution of each load component is expressed by a mixture of Gaussian distribution individually.



(a) Load applied node

(b) Load magnitude

(c) Load direction

Figure 13. Probability distribution of the mixture Gaussian distribution for each load component.

Figure 13 shows the obtained results. Figure 13(a) is the distribution of load applied node. The abscissa indicates the node number. This graph has two peaks. One is sharp peak around node 31 and the other is

dull curve around node 35. The distribution of load applied node has high probability at the two positions and the highest probability occurs at 31th node. Figure 13(b) is the distribution of load magnitude.

The abscissa indicates the load magnitude. This graph has one peak at 120[kN]. This value is 4 times greater than the nominal load of this hook. Figure 13(c) is the distribution of load angle. The abscissa indicates the load angle. This graph has a mild peak between -50° and -100°. The load direction tends to concentrate around downward direction. In the current study, because the load component is treated independently, we cannot show a correlation among the load components. Combining the individual results, we can see that a great load is applied at rightward position from lower center and the load direction is downward.

Figure 14 shows the estimated load applied position and the load direction. We illustrate the probability by gray-scale level based on the obtained probability distribution shown in Fig. 13(a) and (c). The high probability area is represented darkly, whereas the low probability area is represented lightly. The darkest region and arrow in Fig. 14 is the most probable area and the direction when crane-hooks are failed.

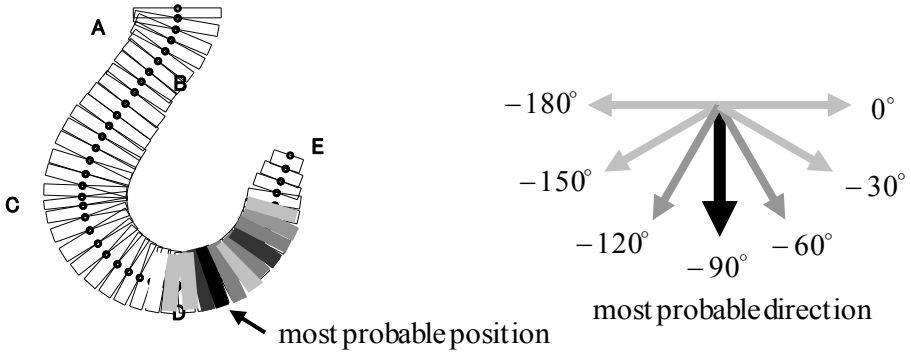


Figure 14. Estimated load applied position and load direction.

6. Optimal Design

6.1. Design Variables And Their Parametric Representation

In this study, the design variables of crane-hook are the parameters of cross-section of beam elements. As shown in Fig. 4, the parameters are the height and layer widths. These design variables are represented as functions of the local coordinate s attached at the center line of hook, as $h(s), b_i(s)\ (0 \le s \le L)$, where L is the length of the contour line. The coordinate s is indicated in Fig. 15. The start point of s is the base point "A" and the end point is the tip point "E". In the followings, $h(s)$ and $b_i(s)$ are called the shape functions.

We represent such shape functions as linear combination of the Gaussian function. The shape functions are then expressed as

$$h(s) = \sum_{j=1}^{N_h} \alpha_j^h \exp\left(-\frac{(s-\mu_j^h)^2}{2(\beta_j^h)^2} \right) \tag{19}$$

$$b_i(s) = \sum_{j=1}^{N_b} \alpha_j^{b_i} \exp\left(-\frac{(s-\mu_j^{b_i})^2}{2(\beta_j^{b_i})^2} \right) \tag{20}$$

where α is the scaling factor, μ and β are the location of the peak and the standard deviation. The constants N_h and N_b are the number of Gaussian functions representing h and b_i, respectively. By introducing this representation, the shape functions $h(s)$ and $b_i(s)$ are represented in terms of the coefficients $\alpha_j^h, \beta_j^h, \mu_j^h$ $(j=1,\cdots,N_h)$ and $\alpha_j^{b_i}, \beta_j^{b_i}, \mu_j^{b_i}$ $(j=1,\cdots,N_b)$.

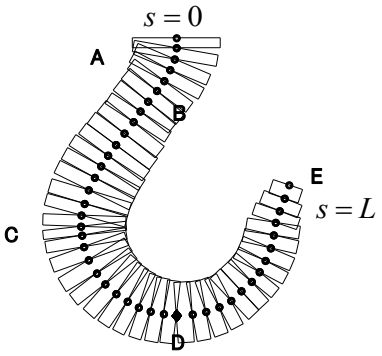


Figure 15. Local curvilinear coordinate s.

6.2. SETTING OF CRITERIA AND FORMULATION OF OPTIMIZATION PROBLEM

We explain the formulation of criteria that evaluate the goodness of crane-hook design. In order to improve the performance of crane-hook, we employ the following criteria:

- structural weight
- structural stiffness

The first criterion is selected for achievement of lightweight. The lightweight is important for the saving of material cost and the compactness. The structural weight J_1 is formulated as

$$J_1 = \sum_{i=1}^{N_e} \rho A_i l_i \tag{21}$$

where ρ is the material density, A_i is the cross-sectional area of i-th element and l_i is the element length.

The second criterion is selected for the evaluation of the robustness of structure against unspecified multiple load conditions. For this evaluation, we adopt the ratio between the norm of the global

displacement vector and the norm of the possible load vector. The robustness of the structure is evaluated in terms of the maximum value of the ratio. The possible load vector is represented by $\widetilde{\boldsymbol{F}}$. This vector specifies the possible load applied points and respective components based on the estimation result. We utilize the estimation result of the load applied node discussed in the previous section. The maximum ratio is expressed as

$$\max_{\widetilde{\boldsymbol{F}}\neq 0}\frac{\|\boldsymbol{U}\|}{\|\widetilde{\boldsymbol{F}}\|} \tag{22}$$

$$\boldsymbol{U}=\boldsymbol{K}^{-1}\boldsymbol{F} \quad \text{where} \quad \boldsymbol{F}=\boldsymbol{B}_v\widetilde{\boldsymbol{F}} \tag{23}$$

where $\|\bullet\|$ represents the vector norm and $\boldsymbol{K}$ is the elastic stiffness matrix in Eq. (1). The global force vector $\boldsymbol{F}$ is associated with the possible load vector $\widetilde{\boldsymbol{F}}$ by the weight factor matrix $\boldsymbol{B}_v$. The component value of the matrix $\boldsymbol{B}_v$ is specified according to the probability distribution of load applied node shown in Fig. 13(a). The maximum ratio is rewritten as the following form.

$$\max_{\widetilde{\boldsymbol{F}}\neq 0}\frac{\|\boldsymbol{U}\|}{\|\widetilde{\boldsymbol{F}}\|}=\max_{\widetilde{\boldsymbol{F}}\neq 0}\frac{\|\boldsymbol{K}^{-1}\boldsymbol{F}\|}{\|\widetilde{\boldsymbol{F}}\|}=\max_{\widetilde{\boldsymbol{F}}\neq 0}\frac{\|\boldsymbol{K}^{-1}\boldsymbol{B}_v\widetilde{\boldsymbol{F}}\|}{\|\widetilde{\boldsymbol{F}}\|} \tag{24}$$

The magnitude of $\|\widetilde{\boldsymbol{F}}\|$ is normalized to be 1. Here, instead of searching the maximum value directly, we utilize the matrix norm. According to the maximum principle of the eigenvalue, the maximum value of this function is calculated as the matrix norm induced by the Euclidean vector norm $\|\bullet\|_2$ (Roger and Charles, 1985). The second criterion is formulated as

$$J_2=\|\boldsymbol{K}^{-1}\boldsymbol{B}_v\|_2=\max\left\{\sqrt{\lambda}:\lambda \text{ is an eigenvalue of } (\boldsymbol{K}^{-1}\boldsymbol{B}_v)^T(\boldsymbol{K}^{-1}\boldsymbol{B}_v)\right\}. \tag{25}$$

This criterion represents a displacement-force ratio and the unit is [m/N].

We formulate the criteria to be minimized in the above. For this multi-objective optimization problem with the two items, an integrated evaluation function is introduced in terms of the weighting factors. The multi-objective optimal design problem of the crane-hook is then expressed as follows:

$$\begin{aligned} &\text{Minimize} && J=\gamma_1\frac{J_1}{\underline{J_1}}+\gamma_2\frac{J_2}{\underline{J_2}}\equiv\gamma_1\widetilde{J}_1+\gamma_2\widetilde{J}_2\\ &\text{with respect to} && \boldsymbol{x}\\ &\text{subject to} && h_L\le h(s)\le h_U,\ b_L\le b_i(s)\le b_U,\ \gamma_1+\gamma_2=1 \end{aligned} \tag{26}$$

where $\boldsymbol{x}=\left[\alpha_j^h \quad \beta_j^h \quad \mu_j^h \quad \alpha_j^{b_i} \quad \beta_j^{b_i} \quad \mu_j^{b_i}\right]$ $(i=1,\cdots,N_d,\ j=1,\cdots,N_h \text{ or } N_b)$. The constants $\underline{J_1}$ and $\underline{J_2}$ are the evaluation item values for normalization. We adopt the FEM model of crane-hook shown in Fig. 4,

called “reference design”, for this normalization. The values $\underline{J_1}$ and $\underline{J_2}$ are calculated for the reference design. The coefficients γ_1 and γ_2 are the weighting factors for the criteria. We conduct the optimal designs under the various combinations of the weighting factors. In the constraint conditions, $h_{L,}\, b_L$ and h_U, b_U are the lower and upper bounds of the height and width.

6.3. Setting of Parameters for Numerical Calculation

The optimization problem (26) is solved by means of the particle swarm optimization (PSO). The PSO is one of the population-based stochastic optimization techniques and has been successfully applied in many research and application areas (Behera and Choukiker, 2010; Mauro et.al., 2009). In the PSO, we need to specify the number of particles and the number of iterations. These parameters are shown in Table II(a). The values of the material parameters E_1, E_2 and $\overline{\sigma}$ are same as the determined values in section 2.2. The parameters of finite element model are shown in Table II(b). The constants in Eq. (26) are specified in Table II(c). The evaluation item values obtained based on the reference design are as follows:

- structural weight $\underline{J_1}$: 2.660 [kg]
- structural stiffness $\underline{J_2}$: 1.4083×10^{-7} [m/N].

The adopted weighting factors in Eq. (26) are specified as

$$\gamma_i = \{0.0, 0.05, 0.10, \cdots, 0.90, 0.95, 1.00\} \quad (i = 1, 2)\,. \tag{27}$$

Table II. Parameters for numerical calculation

	Item	Symbol	Value
(a) Finite element model	Material density	ρ	7.87 [g/cm^3]
	Number of elements	N_e	40
	Number of layers	N_d	10
(b) Optimization problem	Number of Gaussian	N_h, N_b	4
	Lower bound of size	h_L, b_L	5 [mm]
	Upper bound of size	h_U, b_U	40 [mm]
(c) PSO	Number of particles		1000
	Number of iteration		100

6.4. Obtained Optimal Designs

Figure 16 shows the distribution of the criterion function values of the solution of the optimization problem (26). The abscissa shows the structural weight $\widetilde{J}_1$ and the ordinate shows the structural stiffness $\widetilde{J}_2$. Each item value indicates the ratio to the value of reference design. It can be seen that the Pareto line is constructed by the obtained solutions. From the results in this figure, there is a trade-off relationship among the evaluation items. Respective shapes concerning to the solutions (A) and (B) on the Pareto line in Fig.16 are shown in Fig. 17. The evaluation item values are indicated in the caption. In Fig. 17(a) and (b), the left

part shows the height distribution of the elements. In the following, we call this part as “hook shape”. The right part shows the selected cross-section of the design solution; the upper is the section of the point “C” (17th element) and the lower is that of the lowest center point “D” (29th element). The top and bottom of the section shape correspond to the inner and outer surfaces of the hook, respectively. These two are important sections in the practical design scene.

The following features are observed for the obtained hook shape:

– hook shape becomes thinner toward the tip point “E” from the lowest center point “D”
– thickness of region around the point “B” is greater than any other region

If the attached importance on “the structural weight J_1” is larger, the first feature becomes more remarkable. Because the stresses on the surface of hook between the load applied point and the tip point “E” are equals 0, this part has no contribution to the strength. Tapering off around the tip point “E” is a rational shape. In both solutions, the thickest region is not the point “C” (critical section) but also around the point “B”. Because we specify the weight factor matrix $\boldsymbol{B}_v$ based on the estimation result, the possible load applied node is shifted rightward. As a result of this formulation, the most critical area shifts from the point “C” to the upward point “B”.

We discuss the feature of the cross-sectional shape. The section of the point “C” is the rectangular shape. If we attach importance to the weighting factor γ_1, the cross-sectional shape of the point “C” becomes thinner toward the bottom. The stress in the lower part is smaller than that of the upper part; thus the tapering off shape is good for the lightweight. At the point “D”, the width of section around the center part is thinner than both end side (upper and lower). In order to keep high stiffness in the unspecified load cases, it is better to thicken the bottom than to thicken the middle part.

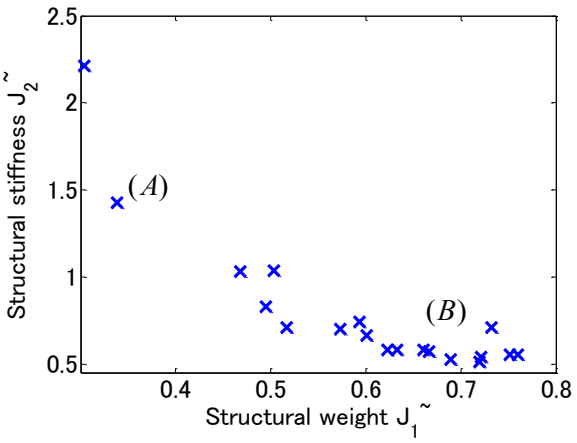
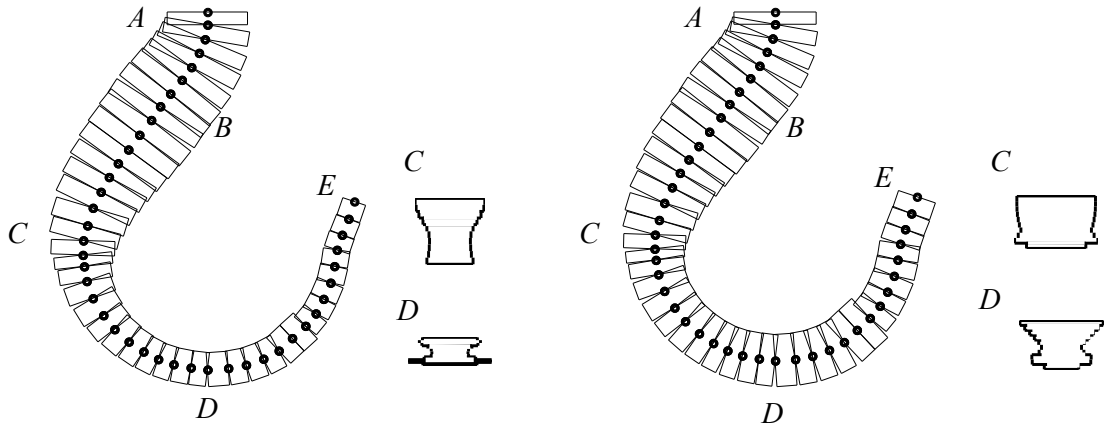


Figure 16. Distribution of objective function values for respective weighting factors.



(a) Hook shape and sections of solution (A) ($\tilde{J}_1 = 0.3324, \tilde{J}_2 = 1.4328$)

(b) Hook shape and sections of solution (B) ($\tilde{J}_1 = 0.6895, \tilde{J}_2 = 0.5263$)

Figure 17. Obtained designs of crane-hook.

The key points obtained from the observation of the solutions are as follows:

- hook shape is tapering off from the lowest center point
- the region around the point "B" is thicker than any other area
- cross-sectional shape of the point "C" is rectangular
- tapering-off shape of hook becomes conspicuous as the importance is attached to the lightweight

7. Conclusion

The estimation of load condition and the optimal design of crane-hook are presented and discussed. The objective of the estimation is to find out the load condition when crane-hooks are failed. In order to adapt the FEM model to the permanent deformation, we implement elasto-plastic deformation analysis. The Load-Deformation database that has the pairs of the applied load condition for the FEM model and their deformed node positions is constructed. The feature points are detected on the failed hook image in order to compare those feature points with the deformed node positions recorded in the L-D database. The applied load condition corresponding to the failed crane-hook image is then obtained by using a difference-minimization approach. The identified loading conditions summarized individually in the form of the mixture Gaussian distribution. In the parameter estimation process on the distribution, we utilize the EM algorithm. The result is that the load applied position lies between the lowest center point and the tip-end, the load magnitude is four times larger than the nominal load and the load direction is the downward.

In order to obtain a high-quality design, we formulate the multi-objective optimization problem taking account of the estimated results. The evaluation items are the structural weight and the structural stiffness. In the representation of the design variables, we utilize the Gaussian functions. The multi-objective problem is converted into the single objective problem by introducing the weighting factors. This problem is solved by means of the PSO. The obtained crane-hook shapes have a tapered shape similar to those of actual crane-

hook designs. However, there are different features on the crane-hook shape. The thickest area of the obtained shape shifts upper part than that of the actual design. The cross-sections do not have a T-shape that is implemented in the actual design but have "a rectangular shape" or "an hourglass shape". By introducing the Gaussian function to represent the design variables, we can reduce the number of design variables and represent the shape functions effectively.

In this study, the components of load condition are estimated individually. This is because the number of failed samples is not enough to calculate the correlation among the load components. Taking the correlation into account, the failure estimation result becomes more meaningful. This is our future work.

References

Quek, S. T., Tran, V. A., Hou, X. Y. and Duan W. H. Structural damage detection using enhanced damage locating vector method with limited wireless sensors. *Journal of sound and vibration*, 38(4-5):411-427, 2009.

Lam, H. F. and Ng, C. T. The selection of pattern features for structural damage detection using an extended Bayesian ANN algorithm. *Engineering Structures*, 30(10):2762-2770, 2009.

Muromaki, T., Hanahara, K., Tada, Y. and Nishimura, T. Estimation of loading conditions of failed crane-hook: an image-based approach with knowledge and simulation. *International Journal of Reliability and Safety*, 6(1/2/3):130-147, 2012.

Crisfield, M. A. *Non-linear Finite Element Analysis of Solid and Structures: Volume 1*, John Wiley & Sons, 1991.

Owen, D. R. J. and Hinton, E. *Finite Elements in Plasticity: Theory and Practice*. Pineridge Press Limited, Chapter 5, 1980.

Ghorbel, F., Derrode, S. Dhahbi, S. and Mezhoud, R. Reconstructing With Ceometric Moments. *Int. Conf. on Machine Intelligence (ACIDCA-ICMI'05)*, Tozeur, Tunisia, November 5-7, 2005.

McLachlan, G. J. and Krishnan, T. *The EM Algorithm and Extensions*, Wiley-Interscience, 1997.

Roger, A. H. and Charles, R. J. *Matrix Analysis*, Cambridge University Press 1985.

Behera, S. K. and Choukiker, Y. Design and Optimization of Dual Band Microstrip Antenna Using Particle Swarm Optimization Technique. *Infrared Milli Terahz Waves*, 31:1346-1354, 2010.

Mauro, A. S. S. R. et. al. Optimal Design of Shell-and-Tube Heat Exchanges Using Particle Swarm Optimization. *Industrial & Engineering Chemistry Research*, 48(6):2927-2935, 2009.

Frequency response functions of discretized structural systems with uncertain parameters

G. Muscolino[1)], R. Santoro[1)], A. Sofi[2)]
[1)]*Department of Civil Engineering, University of Messina, Messina 98166, Italy*
gmuscolino@unime.it; roberta.santoro@unime.it
[2)]*Department of Mechanics and Materials, University Mediterranea of Reggio Calabria, Reggio Calabria 89124, Italy, alba.sofi@unirc.it*

Abstract: A procedure for deriving in explicit approximate form the *frequency response function* (*FRF*) of linear discretized structures with uncertain stiffness properties is presented. The proposed procedure is based on the following main steps: *i*) to perform the spectral decomposition of the deviation of the stiffness matrix (with respect to its nominal value) so as to obtain a sum of rank-one matrices, each one associated to a single uncertain parameter; *ii*) to project the equations of motion in the modal subspace; *iii*) to introduce a novel series expansion of the *FRF* in the modal subspace which provides an approximate, but explicit, expression of the *FRF* of structural systems with uncertain parameters. The potential of the proposed series expansion are demonstrated in the context of the so-called *improved interval analysis* by determining the range of the modulus of the *FRF* of structures with uncertain-but-bounded parameters.

Keywords: Frequency response function; Uncertain parameters; Spectral decomposition; Improved interval analysis.

1. Introduction

In Structural Dynamics, the *frequency response function* (*FRF*), also called *transfer function*, is a complex function able to provide information about the behavior of a structure over a range of frequencies. For instance, the frequency domain response of a single-degree-of-freedom system (SDOF), i.e. an oscillator, is evaluated simply multiplying the *FRF* by the Fourier transform of the forcing function. For multi-DOF structural systems the *FRF* describes the relationship between a local excitation applied at one location on the structure and the resulting response at another and/or the same location. The frequency domain approach often gives information useful for structural design purposes that cannot be alternatively caught by the time domain approach. Moreover, it is sometimes more convenient to perform the analysis in the frequency domain; as an example, for structures with frequency dependent parameters or subjected to stationary random processes and so on. Indeed, in all these cases the evaluation of the *FRF* is required.

In practical engineering problems, material properties, geometry and boundary conditions of a structure may experience fluctuations, due to measurement and manufacturing errors or other factors, which may significantly affect the response. The uncertainties are usually described following two contrasting points of view, known as *probabilistic* and *non-probabilistic approaches*. The probabilistic approach requires a wealth of data, often unavailable, to define the probability distribution density of the uncertain structural

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

parameters. If available information is fragmentary or incomplete, non-probabilistic approaches, such as convex models, fuzzy set theory or interval models (Elishakoff and Ohsaki, 2010), can be alternatively applied.

Among non-probabilistic approaches, the *interval model* turns out to be the most suitable approach when only the upper and lower bounds of a non-deterministic property are well defined. Indeed, this model is derived from the interval analysis (Moore, 1966; Alefeld and Herzberger, 1983; Moore et al., 2009) in which the number is treated as an interval variable ranging between its lower and upper bounds. Unfortunately, the "ordinary" interval analysis (Moore, 1966) suffers from the so-called *dependency phenomenon* (Muhanna and Mullen, 2001; Moens and Vandepitte, 2005; Moore et al., 2009) which often leads to an overestimation of the interval width that could be catastrophic from an engineering point of view. This occurs when an expression contains multiple instances of one or more interval variables. Indeed, the ordinary interval arithmetic operations erroneously assume that the operand interval numbers are independent. To limit the catastrophic effects of the dependency phenomenon, the so-called *generalized interval analysis* (Hansen, 1975) and the *affine arithmetic* (Comba and Stolfi, 1993; Stolfi and De Figueiredo, 2003) have been introduced in the literature. In these formulations, each intermediate result is represented by a linear function with a small remainder interval (Nedialkov et al., 2004).

In the framework of probabilistic approaches, the *FRF* has been evaluated by Falsone and Ferro (2005, 2007) in explicit form by taking into account the properties of the natural deformation modes of the finite element discretized structure. In a non-probabilistic context, Moens and Vandepitte (2004) proposed a numerical procedure to efficiently calculate close outer approximations on the envelope *FRF* of structures with interval uncertainties. The *FRF* of systems with uncertain-but-bounded parameters was also evaluated by Manson (2005) employing both the *complex interval analysis* and the *complex affine arithmetic*.

In this paper, an alternative approach for the evaluation of the *FRF* of discretized structures with uncertain stiffness properties is presented. The proposed procedure requires the following preliminary steps: *i*) the spectral decomposition of the deviation of the stiffness matrix (with respect to its nominal value) to obtain a sum of rank-one matrices, each one associated to a single uncertain parameter; *ii*) the modal analysis to project the equations of motion in the modal subspace. In a second stage, a novel series expansion of the modal *FRF*, named *Rational Series Expansion* (*RSE*), which provides an approximate, but explicit, expression of the *FRF* of structural systems with uncertain parameters, is derived. Finally, the proposed series expansion together with the so-called *improved interval analysis* presented by Muscolino and Sofi (2011) is used to obtain the range of the modulus of the *FRFs* of structures with uncertain-but-bounded parameters.

Numerical applications performed on a truss structure and a portal frame with uncertain Young's moduli of the material have demonstrated the accuracy of the proposed explicit approximation of the *FRF*.

2. Preliminary concepts

2.1. EQUATIONS OF MOTION

Let us consider a quiescent *n*-DOF linear structural system with uncertain stiffness properties subjected to the forcing vector $\mathbf{f}(t)$. The equations of motion can be cast in the form:

$$\mathbf{M}\ddot{\mathbf{u}}(\boldsymbol{\alpha},t)+\mathbf{C}\dot{\mathbf{u}}(\boldsymbol{\alpha},t)+\mathbf{K}(\boldsymbol{\alpha})\mathbf{u}(\boldsymbol{\alpha},t)=\mathbf{f}(t) \tag{1}$$

where **M** and **C** are the $n\times n$ mass and damping matrices of the structure; $\boldsymbol{\alpha} = [\alpha_1, \alpha_2,..., \alpha_r]^T$ is the vector collecting the r dimensionless uncertain parameters α_i; $\mathbf{u}(\boldsymbol{\alpha},t)$ is the vector of nodal displacements and a dot over a variable denotes differentiation with respect to time t.

It is worth noting that the relationship between the stiffness matrix, $\mathbf{K}(\boldsymbol{\alpha})$, and the vector $\boldsymbol{\alpha}$ is often linear or, by applying a suitable variable transformation, it is always possible to make the stiffness matrix depend linearly on the new variables. Based on this concept, the stiffness matrix $\mathbf{K}(\boldsymbol{\alpha})$ is herein expressed as a linear function of the uncertain properties, i.e.:

$$\mathbf{K}(\boldsymbol{\alpha})=\mathbf{K}_0+\Delta\mathbf{K}(\boldsymbol{\alpha})=\mathbf{K}_0+\sum_{i=1}^{r}\mathbf{K}_i\,\alpha_i\,; \tag{2}$$

$$\mathbf{K}_0=\mathbf{K}(\boldsymbol{\alpha}_0);\quad \mathbf{K}_i=\frac{\partial}{\partial\alpha_i}\mathbf{K}(\boldsymbol{\alpha})\bigg|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \tag{3}$$

where $\mathbf{K}_0$ is the nominal value of the stiffness matrix, which is a positive definite symmetric matrix of order $n\times n$, $\mathbf{K}_i$ is a semi-positive definite symmetric matrix of order $n\times n$ and rank p_i and α_i is the i-th dimensionless uncertain parameter. In structural engineering problems, the fluctuating properties can be reasonably assumed to satisfy the condition $|\alpha_i| < 1$, with the symbol $|\bullet|$ denoting absolute value.

In the framework of the traditional modal analysis, the solution of the equations of motion (1) may be pursued by introducing the following coordinate transformation:

$$\mathbf{u}(\boldsymbol{\alpha},t)=\boldsymbol{\Phi}_0\,\mathbf{q}(\boldsymbol{\alpha},t) \tag{4}$$

where $\mathbf{q}(\boldsymbol{\alpha},t)$ is the vector gathering the first m modal coordinates $q_j(\boldsymbol{\alpha},t)$ ($j = 1,2,..., m \le n$); $\boldsymbol{\Phi}_0$ is the modal matrix, of order $n\times m$, pertaining to the nominal configuration in which $\mathbf{K}_0 = \mathbf{K}(\boldsymbol{\alpha}_0)$. Specifically, the modal matrix $\boldsymbol{\Phi}_0$, collecting the first m eigenvectors normalized with respect to the mass matrix **M**, is evaluated as solution of the following eigenproblem:

$$\mathbf{K}_0\boldsymbol{\Phi}_0=\mathbf{M}\boldsymbol{\Phi}_0\,\boldsymbol{\Omega}_0^2\,;\quad \boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{M}\boldsymbol{\Phi}_0=\mathbf{I}_m \tag{5}$$

where $\mathbf{I}_m$ denotes the identity matrix of order m; $\boldsymbol{\Omega}_0^{\,2} = \boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{K}_0\boldsymbol{\Phi}_0$ is the spectral matrix of the nominal structural system, say a diagonal matrix listing the squares of the natural circular frequencies of the structure, $\omega_{0,i}$, for the nominal values of the uncertain parameters; the apex T means transpose matrix. By applying the coordinate transformation (4), the equations of motion (1) can be projected in the modal space:

$$\ddot{\mathbf{q}}(\boldsymbol{\alpha},t)+\boldsymbol{\Xi}\,\dot{\mathbf{q}}(\boldsymbol{\alpha},t)+\boldsymbol{\Omega}^2(\boldsymbol{\alpha})\,\mathbf{q}(\boldsymbol{\alpha},t)=\mathbf{p}(t) \tag{6}$$

where $\boldsymbol{\Omega}^2(\boldsymbol{\alpha}) = \boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{K}(\boldsymbol{\alpha})\boldsymbol{\Phi}_0$; $\boldsymbol{\Xi} = \boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{C}\boldsymbol{\Phi}_0$ is the generalised damping matrix, which for classically damped systems is a diagonal one; $\mathbf{p}(t) = \boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{f}(t)$ is the modal forcing vector. Notice that by virtue of the decomposition (2) of the stiffness matrix, the following relationship holds:

$$\boldsymbol{\Omega}^2(\boldsymbol{\alpha}) = \boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{K}(\boldsymbol{\alpha})\boldsymbol{\Phi}_0 = \boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{K}_0\boldsymbol{\Phi}_0 + \sum_{i=1}^{r}\boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{K}_i\boldsymbol{\Phi}_0\alpha_i = \boldsymbol{\Omega}_0^2 + \sum_{i=1}^{r}\boldsymbol{\Omega}_i^2\alpha_i \tag{7}$$

where

$$\boldsymbol{\Omega}_i^2 = \boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{K}_i\boldsymbol{\Phi}_0 \tag{8}$$

is not a diagonal matrix.

2.2. FREQUENCY DOMAIN RESPONSE

In some cases, such as for structures with frequency dependent parameters or in presence of stochastic stationary excitations, it is more convenient to perform the analysis in the so-called frequency domain.

In the context of the frequency domain analysis, it is assumed that the loading is periodic and has been resolved into its discrete harmonic components by Fourier transformation. The corresponding harmonic components of the structural response can be derived by performing the Fourier transform of both sides of Eq. (6) (or Eq.(1)) obtaining the following set of algebraic frequency dependent equations:

$$\left[-\omega^2\,\mathbf{I}_m + \mathrm{i}\,\omega\boldsymbol{\Xi} + \boldsymbol{\Omega}^2(\boldsymbol{\alpha})\right]\mathbf{Q}(\boldsymbol{\alpha},\omega) = \mathbf{P}(\omega) \tag{9}$$

where $\mathbf{Q}(\boldsymbol{\alpha},\omega)$ and $\mathbf{P}(\omega)$ are the vectors collecting the Fourier transforms of $\mathbf{q}(\boldsymbol{\alpha},t)$ and $\mathbf{p}(t)$, respectively. The modal frequency response vector $\mathbf{Q}(\boldsymbol{\alpha},\omega)$, solution of Eq.(9), can be expressed as follows:

$$\mathbf{Q}(\boldsymbol{\alpha},\omega) = \mathbf{H}(\boldsymbol{\alpha},\omega)\mathbf{P}(\omega) \tag{10}$$

where

$$\mathbf{H}(\boldsymbol{\alpha},\omega) = \left[-\omega^2\,\mathbf{I}_m + \mathrm{i}\,\omega\boldsymbol{\Xi} + \boldsymbol{\Omega}^2(\boldsymbol{\alpha})\right]^{-1} = \left[\,\mathbf{H}_0^{-1}(\omega) + \sum_{i=1}^{r}\boldsymbol{\Omega}_i^2\alpha_i\,\right]^{-1} \tag{11}$$

is the *modal frequency response function (FRF) matrix* (referred to also as *transfer function matrix*) whose expression has been derived taking into account Eq. (7) and introducing the *FRF* matrix of the nominal structural system, given by:

$$\mathbf{H}_0(\omega) = \left[-\omega^2\,\mathbf{I}_m + \mathrm{i}\,\omega\boldsymbol{\Xi} + \boldsymbol{\Omega}_0^2\right]^{-1}. \tag{12}$$

It is worth noting that the *FRF* matrix $\mathbf{H}(\boldsymbol{\alpha},\omega)$ is not diagonal, while for classically damped systems the matrix $\mathbf{H}_0(\omega)$ is a diagonal one.

Once the modal frequency response $\mathbf{Q}(\boldsymbol{\alpha},\omega)$ is evaluated, the frequency response $\mathbf{U}(\boldsymbol{\alpha},\omega)$ in the nodal space can be obtained by performing the Fourier Transform of Eq.(4), i.e.:

$$\mathbf{U}(\boldsymbol{\alpha},\omega) = \boldsymbol{\Phi}_0\,\mathbf{Q}(\boldsymbol{\alpha},\omega). \tag{13}$$

To avoid the inversion of the parametric frequency dependent matrix in Eq.(11), the Neumann series expansion can be adopted which leads to the following expression:

$$\mathbf{H}(\boldsymbol{\alpha},\omega)=\left[\mathbf{I}_m+\mathbf{H}_0(\omega)\sum_{i=1}^{r}\boldsymbol{\Omega}_i^2\alpha_i\right]^{-1}\mathbf{H}_0(\omega)=\mathbf{H}_0(\omega)+\sum_{k=1}^{\infty}(-1)^k\left[\mathbf{H}_0(\omega)\sum_{i=1}^{r}\boldsymbol{\Omega}_i^2\alpha_i\right]^{k}\mathbf{H}_0(\omega). \quad (14)$$

The convergence of this series expansion is guaranteed if and only if the least square norm of the matrix in square brackets is less than one. In the next section, an alternative series expansion of the modal *FRF* matrix for structural systems with uncertain parameters is proposed.

3. Proposed explicit form of the FRF matrix

3.1. SPECTRAL DECOMPOSITION OF THE STIFFNESS MATRIX

As well known, in structural engineering the stiffness matrix is always a positive definite matrix. In the previous section, the stiffness matrix has been assumed to depend on r dimensionless uncertain parameters satisfying the conditions $|\alpha_i| < 1$, i.e. $\mathbf{K} = \mathbf{K}(\boldsymbol{\alpha})$. Furthermore, the stiffness matrix has been decomposed according to Eq.(2), where $\mathbf{K}_0$ is a positive definite symmetric matrix of order $n \times n$, while $\mathbf{K}_i$ is a semi-positive symmetric matrix of order $n \times n$ and rank p_i. As an example, in the case of truss structures and shear-type frames, the matrices $\mathbf{K}_i$ have rank $p_i = 1$. Instead, for flexible frames the matrix $\mathbf{K}_i$ has rank $p_i = 3$ and so on. The foregoing property can be exploited to perform the *spectral decomposition* (referred to also as *eigendecomposition*) of the matrices $\mathbf{K}_i$. To this aim, the following eigenproblems have to be solved:

$$\mathbf{K}_i\boldsymbol{\psi}_i^{(\ell)}=\lambda_i^{(\ell)}\mathbf{K}_0\boldsymbol{\psi}_i^{(\ell)}, \quad (i=1,2,...,r;\ \ell=1,2,...,p_i) \quad (15)$$

where $\lambda_i^{(\ell)}$ denote the eigenvalues which are real positive numbers, while $\boldsymbol{\psi}_i^{(\ell)}$ are the associated eigenvectors. Due to the semi-positivity of the matrix $\mathbf{K}_i$, among the n eigenvalues of the i-th eigenproblem in Eq. (15) only $p_i < n$ eigenvalues are different from zero. As an example, in the case of truss structures and shear-type frames only one eigenvalue different from zero is found for each uncertain parameter; for flexible frames, each eigenproblem (15) yields three eigenvalues different from zero and so on.

By imposing that the eigenvectors $\boldsymbol{\psi}_i^{(\ell)}$ satisfy the orthonormalization condition:

$$\boldsymbol{\Psi}_i^{\mathrm{T}}\mathbf{K}_0\boldsymbol{\Psi}_i=\mathbf{I}_{p_i}; \quad \boldsymbol{\Psi}_i=\left[\boldsymbol{\psi}_i^{(1)} \quad \boldsymbol{\psi}_i^{(2)} \quad \cdots \quad \boldsymbol{\psi}_i^{(p_i)}\right], \quad (16)$$

the following relationship holds:

$$\boldsymbol{\Psi}_i^{\mathrm{T}}\mathbf{K}_i\boldsymbol{\Psi}_i=\boldsymbol{\Lambda}_i; \quad \boldsymbol{\Lambda}_i=\mathrm{Diag}\left[\lambda_i^{(1)}, \quad \lambda_i^{(2)} \quad ,\cdots, \quad \lambda_i^{(p_i)}\right]. \quad (17)$$

Then, after very simple algebra, by applying the previously described spectral decomposition, the matrix $\mathbf{K}_i$ can be written as:

$$\mathbf{K}_i = \left(\mathbf{K}_0 \boldsymbol{\Psi}_i\right) \boldsymbol{\Lambda}_i \left(\boldsymbol{\Psi}_i^{\mathrm{T}} \mathbf{K}_0\right) = \sum_{\ell=1}^{p_i} \lambda_i^{(\ell)} \left(\mathbf{K}_0 \boldsymbol{\psi}_i^{(\ell)}\right)\left(\boldsymbol{\psi}_i^{(\ell)\mathrm{T}} \mathbf{K}_0\right) = \sum_{\ell=1}^{p_i} \lambda_i^{(\ell)} \mathbf{v}_i^{(\ell)}\, \mathbf{v}_i^{(\ell)\mathrm{T}} \tag{18}$$

where

$$\mathbf{v}_i^{(\ell)} = \mathbf{K}_0\, \boldsymbol{\psi}_i^{(\ell)}. \tag{19}$$

Substituting Eq.(18) into Eq. (2), the stiffness matrix $\mathbf{K}(\boldsymbol{\alpha})$ can be expressed as the superposition of $p = \sum_{i=1}^{r} p_i$ changes of rank-one, i.e.:

$$\mathbf{K}(\boldsymbol{\alpha}) = \mathbf{K}_0 + \Delta\mathbf{K} = \mathbf{K}_0 + \sum_{i=1}^{r} \alpha_i \left[\sum_{\ell=1}^{p_i} \lambda_i^{(\ell)} \mathbf{v}_i^{(\ell)}\, \mathbf{v}_i^{(\ell)\mathrm{T}}\right]. \tag{20}$$

Finally, upon introducing the spectral decomposition of the stiffness matrix given by Eq.(20) into Eq. (7), the matrix $\boldsymbol{\Omega}^2(\boldsymbol{\alpha})$, appearing in the *FRF* matrix (11), takes the following form:

$$\boldsymbol{\Omega}^2(\boldsymbol{\alpha}) = \boldsymbol{\Phi}_0^{\mathrm{T}} \mathbf{K}(\boldsymbol{\alpha}) \boldsymbol{\Phi}_0 = \boldsymbol{\Omega}_0^2 + \sum_{i=1}^{r} \boldsymbol{\Omega}_i^2 \alpha_i = \boldsymbol{\Omega}_0^2 + \sum_{i=1}^{r} \alpha_i \left[\sum_{\ell=1}^{p_i} \lambda_i^{(\ell)} \mathbf{w}_i^{(\ell)}\, \mathbf{w}_i^{(\ell)\mathrm{T}}\right] \tag{21}$$

where

$$\boldsymbol{\Omega}_i^2 = \sum_{\ell=1}^{p_i} \lambda_i^{(\ell)} \mathbf{w}_i^{(\ell)}\, \mathbf{w}_i^{(\ell)\mathrm{T}} \tag{22}$$

with

$$\mathbf{w}_i^{(\ell)} = \boldsymbol{\Phi}_0^{\mathrm{T}} \mathbf{K}_0\, \boldsymbol{\psi}_i^{(\ell)}. \tag{23}$$

3.2. APPROXIMATE MODAL FRF MATRIX FOR TRUSS STRUCTURES WITH UNCERTAIN PARAMETERS

In order to illustrate the proposed procedure for the derivation of an explicit approximate form of the *FRF* matrix, the simplest case of truss structures is first examined. In particular, recalling that for truss structures the *i*-th eigenproblem in Eq. (15) gives only one eigenvalue different from zero, i.e. $p_i = 1$, $\lambda_I = \lambda_i^{(1)}$ and $\boldsymbol{\psi}_i = \boldsymbol{\psi}_i^{(1)}$, the spectral decomposition of the matrix $\mathbf{K}_i$ outlined in the previous section reduces to:

$$\mathbf{K}_i = \lambda_i\, \mathbf{v}_i\, \mathbf{v}_i^{\mathrm{T}} \tag{24}$$

where $\mathbf{v}_i = \mathbf{v}_i^{(1)}$. Accordingly, the matrix $\boldsymbol{\Omega}_i^2(\boldsymbol{\alpha})$ in Eq. (22) takes the following form:

$$\mathbf{\Omega}_i^2 = \alpha_i \lambda_i \, \mathbf{w}_i \, \mathbf{w}_i^{\mathrm{T}} \tag{25}$$

where $\mathbf{w}_i = \mathbf{w}_i^{(1)}$. By substituting Eq. (25) into Eq. (14), the Neumann series expansion of the modal *FRF* matrix can be rewritten as:

$$\mathbf{H}(\boldsymbol{\alpha}, \omega) = \mathbf{H}_0(\omega) + \sum_{s=1}^{\infty} (-1)^s \left[\mathbf{H}_0(\omega) \sum_{i=1}^{r} \alpha_i \, \lambda_i \, \mathbf{w}_i \, \mathbf{w}_i^{\mathrm{T}} \right]^s \mathbf{H}_0(\omega). \tag{26}$$

In order to improve the convergence, the terms into square brackets in Eq.(26) are herein rewritten in explicit form obtaining the following expression of the *FRF* matrix, named *Rational Series Expansion* (*RSE*):

$$\begin{aligned} \mathbf{H}(\boldsymbol{\alpha}, \omega) = \left[\mathbf{H}_0^{-1}(\omega) + \sum_{i=1}^{r} \alpha_i \, \lambda_i \, \mathbf{w}_i \, \mathbf{w}_i^{\mathrm{T}} \right]^{-1} &\approx \mathbf{H}_0(\omega) - \sum_{i=1}^{r} \frac{\alpha_i \lambda_i}{1 + \alpha_i \lambda_i \, d_i \, (\omega)} \mathbf{D}_i(\omega) \\ &+ \sum_{i=1}^{r} \sum_{j=i+1}^{r} \frac{\alpha_i \alpha_j \lambda_j \lambda_i}{1 + \alpha_j \lambda_j d_{ij}(\omega)} d_{ij}(\omega) \left[\mathbf{D}_{ij}(\omega) + \mathbf{D}_{ij}{}^{\mathrm{T}}(\omega) \right] - \\ &- \sum_{i=1}^{r} \sum_{\substack{j=1 \\ j \neq i}}^{r} \sum_{\substack{k=1 \\ k \neq j}}^{r} \frac{\alpha_i \alpha_j \alpha_k \lambda_i \lambda_j \lambda_k}{1 + \alpha_k \lambda_k d_{jk}(\omega)} d_{ij}(\omega) d_{jk}(\omega) \mathbf{D}_{ik}(\omega) + \\ &+ \sum_{i=1}^{r} \sum_{\substack{j=1 \\ j \neq i}}^{r} \sum_{\substack{k=1 \\ k \neq j}}^{r} \sum_{\substack{\ell=1 \\ \ell \neq k}}^{r} \frac{\alpha_i \alpha_j \alpha_k \alpha_\ell \lambda_i \lambda_j \lambda_k \lambda_\ell}{1 + \alpha_\ell \lambda_\ell d_{k\ell}(\omega)} d_{ij}(\omega) d_{jk}(\omega) d_{k\ell}(\omega) \mathbf{D}_{i\ell}(\omega) - \cdots \end{aligned} \tag{27}$$

where only the first four terms have been retained and the following complex quantities have been introduced:

$$d_i(\omega) = \mathbf{w}_i^{\mathrm{T}} \mathbf{H}_0(\omega) \mathbf{w}_i \, ; \qquad \mathbf{D}_i(\omega) = \mathbf{H}_0(\omega) \mathbf{w}_i \, \mathbf{w}_i^{\mathrm{T}} \mathbf{H}_0(\omega); \tag{28}$$

$$d_{is}(\omega) = \mathbf{w}_i^{\mathrm{T}} \mathbf{H}_0(\omega) \mathbf{w}_s \, ; \qquad \mathbf{D}_{is}(\omega) = \mathbf{H}_0(\omega) \mathbf{w}_i \, \mathbf{w}_s^{\mathrm{T}} \mathbf{H}_0(\omega); \quad (s = j, k, \ell, m, \ldots). \tag{29}$$

Equation (27) holds if and only if the following conditions are satisfied:

$$\left\| \alpha_i \lambda_i d_i(\omega) \right\| < 1; \qquad \left\| \alpha_j \lambda_j d_{ij}(\omega) \right\| < 1; \ldots \tag{30}$$

where the symbol $\|\bullet\|$ denotes the modulus of $\bullet$.

Moreover, if $|\alpha_s| << 1$, the approximate modal *FRF* matrix can be accurately evaluated by retaining only first-order terms of the *RSE* in Eq.(27), i.e.:

$$\mathbf{H}(\boldsymbol{\alpha},\omega)=\left[\mathbf{H}_0^{-1}(\omega)+\sum_{i=1}^{r}\alpha_i\,\lambda_i\,\mathbf{w}_i\mathbf{w}_i^{\mathrm{T}}\right]^{-1}\approx\mathbf{H}_0(\omega)-\sum_{i=1}^{r}\frac{\alpha_i\lambda_i}{1+\alpha_i\lambda_i\,d_i\,(\omega)}\,\mathbf{D}_i(\omega). \tag{31}$$

It has to be emphasized that Eqs. (27) and (31) provide with different levels of accuracy closed form expressions of the modal *FRF* matrix of truss structures with uncertain parameters. This remarkable result can be exploited to derive explicit solutions for the frequency domain response of truss structures with fluctuating parameters.

3.3. APPROXIMATE MODAL FRF MATRIX FOR THE MOST GENERAL CASE OF DISCRETIZED STRUCTURES

In this section, an approximate closed form expression of the *FRF* matrix for the most general case of discretized structural systems is derived by applying the procedure described above for truss structures. Specifically, taking into account that in this case the *i*-th eigenproblem in Eq. (15) gives p_i eigenvalues different from zero, the spectral decomposition of the stiffness matrix leads to Eq. (22) for the matrix $\boldsymbol{\Omega}_i^2$. Substituting this expression into Eq.(14) and rewriting the terms of the Neumann series expansion according to Eq.(27), the modal *FRF* matrix can be approximated in explicit form by the following *RSE*:

$$\begin{aligned}\mathbf{H}(\boldsymbol{\alpha},\omega)\approx\mathbf{H}_0(\omega)&-\sum_{i=1}^{r}\sum_{\ell=1}^{p_i}\frac{\alpha_i\,\lambda_i^{(\ell)}}{1+\alpha_i\,\lambda_i^{(\ell)}\,b_{i\ell}(\omega)}\,\mathbf{B}_{i\ell}\,(\omega)+\\&+\sum_{i=1}^{r}\sum_{\ell=1}^{p_i}\sum_{\substack{j=1\\j\neq i}}^{r}\sum_{m=1}^{p_j}\frac{\alpha_i\alpha_j\lambda_i^{(\ell)}\lambda_j^{(m)}}{1+\alpha_j\lambda_j^{(m)}\,b_{ij\ell m}(\omega)}\,b_{ij\ell m}(\omega)\mathbf{B}_{ij\ell m}(\omega)+\\&-\sum_{i=1}^{r}\sum_{\ell=1}^{p_i}\sum_{\substack{j=1\\j\neq i}}^{r}\sum_{m=1}^{p_j}\sum_{\substack{k=1\\k\neq j}}^{r}\sum_{n=1}^{p_k}\frac{\alpha_i\alpha_j\alpha_k\lambda_i^{(\ell)}\lambda_j^{(m)}\lambda_k^{(n)}}{1+\alpha_k\lambda_k^{(n)}\,b_{jkmn}(\omega)}\,b_{jkmn}(\omega)b_{ij\ell m}(\omega)\mathbf{B}_{ik\ell n}(\omega)+\cdots\end{aligned} \tag{32}$$

where

$$b_{i\ell}(\omega)=\mathbf{w}_i^{(\ell)\mathrm{T}}\mathbf{H}_0(\omega)\,\mathbf{w}_i^{(\ell)};\qquad\mathbf{B}_{i\ell}(\omega)=\mathbf{H}_0(\omega)\,\mathbf{w}_i^{(\ell)}\,\mathbf{w}_i^{(\ell)\mathrm{T}}\mathbf{H}_0(\omega); \tag{33}$$

$$b_{ij\ell m}(\omega)=\mathbf{w}_i^{(\ell)\mathrm{T}}\mathbf{H}_0(\omega)\,\mathbf{w}_j^{(m)};\qquad\mathbf{B}_{ij\ell m}(\omega)=\mathbf{H}_0(\omega)\,\mathbf{w}_i^{(\ell)}\,\mathbf{w}_j^{(m)\mathrm{T}}\mathbf{H}_0(\omega); \tag{34}$$

$$b_{jkmn}(\omega)=\mathbf{w}_j^{(m)\mathrm{T}}\mathbf{H}_0(\omega)\,\mathbf{w}_k^{(n)};\qquad\mathbf{B}_{ik\ell n}(\omega)=\mathbf{H}_0(\omega)\,\mathbf{w}_i^{(\ell)}\,\mathbf{w}_k^{(n)\mathrm{T}}\mathbf{H}_0(\omega) \tag{35}$$

are complex quantities. Obviously, Eq.(32) holds if and only if the following conditions are satisfied:

$$\left\|\alpha_i\lambda_i^{(\ell)}\,b_{i\ell}(\omega)\right\|<1;\quad\left\|\alpha_j\lambda_j^{(m)}\,b_{ij\ell m}(\omega)\right\|<1;\quad\left\|\alpha_k\lambda_k^{(n)}\,b_{jkmn}(\omega)\right\|<1;\ldots \tag{36}$$

If the uncertain parameters satisfy the condition $|\alpha_s|<<1$, an accurate approximation of the modal *FRF* matrix can be obtained by retaining only first-order terms of the *RSE* in Eq.(32), i.e.:

$$\mathbf{H}(\boldsymbol{\alpha},\omega)=\left[\mathbf{H}_0^{-1}(\omega)+\sum_{i=1}^{r}\alpha_i\left(\sum_{\ell=1}^{p_i}\lambda_i^{(\ell)}\mathbf{w}_i^{(\ell)}\,\mathbf{w}_i^{(\ell)\mathrm{T}}\right)\right]^{-1}\approx\mathbf{H}_0(\omega)-\sum_{i=1}^{r}\sum_{\ell=1}^{p_j}\frac{\alpha_i\,\lambda_i^{(\ell)}}{1+\alpha_i\,\lambda_i^{(\ell)}\,b_{i\ell}(\omega)}\,\mathbf{B}_{i\ell}\,(\omega). \quad (37)$$

Equations (32) and (37) represent closed form expressions which approximate with different accuracy the modal *FRF* matrix of discretized structures with uncertain parameters. Such expressions are very useful to investigate the effects of the fluctuating properties on the frequency domain response of discretized structures, since the response can be derived in explicit form as well.

4. Uncertain-but-bounded parameters

4.1. PRELIMINARY DEFINITIONS: REAL AND COMPLEX INTERVAL VARIABLES

In this section, the r uncertain structural parameters α_i (i = 1,2,…,r) introduced in the above formulation are assumed independent and are modeled as interval variables. Then, according to the *"ordinary" interval analysis* (Moore, 1966; Alefeld and Herzberger, 1983; Neumaier, 1990; Moore et al., 2009), denoting by $\mathbb{IR}$ the set of all closed real interval numbers, the bounded set-interval vector of real numbers $\boldsymbol{\alpha}^I \triangleq [\underline{\boldsymbol{\alpha}},\overline{\boldsymbol{\alpha}}]\in\mathbb{IR}^r$, such that $\underline{\boldsymbol{\alpha}}\le\boldsymbol{\alpha}\le\overline{\boldsymbol{\alpha}}$, can be introduced. The apex I characterizes the interval variables, while $\underline{\boldsymbol{\alpha}}$ and $\overline{\boldsymbol{\alpha}}$ denote the vectors collecting the lower and upper bounds of the i-th uncertain parameter $\alpha_i^I\in\mathbb{IR}$, say $\underline{\alpha}_i$ and $\overline{\alpha}_i$.

Unfortunately, the "ordinary" interval analysis suffers from the so-called *dependency phenomenon* (Muhanna and Mullen, 2001; Moens and Vandepitte, 2005; Moore et al., 2009) which often leads to an overestimation of the interval width that could be catastrophic from an engineering point of view. This occurs when an expression contains multiple instances of one or more interval variables. To limit the catastrophic effects of the dependency phenomenon, the so-called *generalized interval analysis* (Hansen, 1975) and *affine arithmetic* (Comba and Stolfi, 1993; Stolfi and De Figueiredo, 2003) have been introduced in the literature. In these formulations, each intermediate result is represented by a linear function with a small remainder interval (Nedialkov et al., 2004). According to the philosophy of the *affine arithmetic*, Muscolino and Sofi (2011) proposed the so-called *improved interval analysis* based on the definition of the *extra symmetric unitary interval* (EUI) variable $\hat{e}^I\triangleq[-1,+1]$, ($i=1,2,\ldots,r$). The EUI is defined in such a way that the following properties hold:

$$\hat{e}_i^I-\hat{e}_i^I=0;\qquad \hat{e}_i^I\times\hat{e}_i^I=\left(\hat{e}_i^I\right)^2=[1,1]; \quad (38)$$

$$\hat{e}_i^I\times\hat{e}_j^I=[-1,+1],\quad i\neq j;\qquad \hat{e}_i^I/\hat{e}_i^I=[1,1]. \quad (39)$$

where the subscript i means that the EUI variable is associated to the i-th uncertain-but-bounded parameter. In the previous equations, [1,1] = 1 is the so-called unitary *thin interval*. It is useful to remember that a thin interval occurs when $\underline{\alpha}=\overline{\alpha}$ and it is defined as $\alpha^I\triangleq[\underline{\alpha},\underline{\alpha}]$, so that $\alpha\in\mathbb{R}$. Then, introducing the midpoint value (or mean), $\alpha_{0,i}$, and the deviation amplitude (or radius), $\Delta\alpha_i$, of the i-th real interval variable α_i^I:

$$\alpha_{0,i} = \frac{1}{2}\left(\underline{\alpha}_i + \overline{\alpha}_i\right); \quad \Delta\alpha_i = \frac{1}{2}\left(\overline{\alpha}_i - \underline{\alpha}_i\right), \tag{40}$$

the following *affine form* definition can be adopted:

$$\alpha_i^I = \alpha_{0,i} + \Delta\alpha_i \, \hat{e}_i^I, \quad \left(i = 1, 2, ..., r\right). \tag{41}$$

In the case of complex interval variables, within the framework of the affine arithmetic, Manson (2005) proposed an approach which allows to take into account the dependency between the real and imaginary components of the complex variables. Conversely, the "ordinary" complex interval analysis assumes that the real and imaginary components are independent. According to the philosophy of the affine arithmetic, a complex interval variable $z_i^I = x_i^I + \mathrm{i}\, y_i^I$ is herein defined as:

$$z_i^I = z_{0,i} + \Delta z_i \, \hat{e}_i^I = \left(x_{0,i} + \mathrm{i}\, y_{0,i}\right) + \left(\Delta x_i + \mathrm{i}\, \Delta y_i\right)\hat{e}_i^I \tag{42}$$

where $\mathrm{i} = \sqrt{-1}$ denotes the imaginary unit; $x_{0,i}$ and $y_{0,i}$ are the midpoint values (or means) and Δx_i and Δy_i are the deviation amplitude (or radius) of the real and imaginary part of the complex interval variable, respectively, given by:

$$x_{0,i} = \frac{1}{2}\left(\underline{x}_i + \overline{x}_i\right); \quad y_{0,i} = \frac{1}{2}\left(\underline{y}_i + \overline{y}_i\right); \quad \Delta x_i = \frac{1}{2}\left(\overline{x}_i - \underline{x}_i\right); \quad \Delta y_i = \frac{1}{2}\left(\overline{y}_i - \underline{y}_i\right). \tag{43}$$

4.2. INTERVAL STIFFNESS MATRIX

In structural engineering, the uncertain-but-bounded parameters can be reasonably assumed to posses symmetric deviation amplitude $\overline{\alpha}_i = -\underline{\alpha}_i \equiv \alpha_i$, so that the generic interval variable, according to the improved interval analysis, can be written in *affine form* as:

$$\alpha_i^I = \Delta\alpha_i \, \hat{e}_i^I \tag{44}$$

being $\alpha_{0,i} = 0$ and $\Delta\alpha_i > 0$.

Then, following the interval formalism above introduced, the stiffness matrix $\mathbf{K}(\boldsymbol{\alpha})$ can be expressed as a linear function of the interval variables, i.e.:

$$\mathbf{K}(\boldsymbol{\alpha}) = \mathbf{K}_0 + \Delta\mathbf{K}(\boldsymbol{\alpha}) = \mathbf{K}_0 + \sum_{i=1}^{r} \mathbf{K}_i \, \Delta\alpha_i \, \hat{e}_i^I, \quad \boldsymbol{\alpha} \in \boldsymbol{\alpha}^I = \left[\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}\right] \tag{45}$$

where the matrices $\mathbf{K}_0$ and $\mathbf{K}_i$, of order $n \times n$, have been defined in Eq. (3) and $\Delta\alpha_i$ is the dimensionless fluctuation of the i-th uncertain parameter. Furthermore, by virtue of the decomposition (45) of the stiffness matrix, the following relationship holds:

$$\boldsymbol{\Omega}^2(\boldsymbol{\alpha}) = \boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{K}_0\boldsymbol{\Phi}_0 + \sum_{i=1}^{r}\boldsymbol{\Phi}_0^{\mathrm{T}}\mathbf{K}_i\boldsymbol{\Phi}_0\Delta\alpha_i\,\hat{e}_i^I = \boldsymbol{\Omega}_0^2 + \sum_{i=1}^{r}\boldsymbol{\Omega}_i^2\Delta\alpha_i\,\hat{e}_i^I, \quad \boldsymbol{\alpha}\in\boldsymbol{\alpha}^I = \left[\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}\right] \tag{46}$$

where $\boldsymbol{\Omega}_i^2$ is the matrix defined in Eq.(8).

4.3. APPROXIMATE INTERVAL MODAL FRF MATRIX

In order to simplify interval computations, the attention is herein focused on small deviation amplitudes of the uncertain-but-bounded parameters, i.e. $|\alpha_i| \ll 1$. Under this assumption, based on the *RSE* in Eq.(37) the *interval modal FRF matrix*, in the most general case of discretized structural systems with uncertain-but-bounded stiffness properties, can be expressed in the following approximate explicit form:

$$\mathbf{H}(\boldsymbol{\alpha},\omega) \approx \mathbf{H}_0(\omega) - \sum_{i=1}^{r}\sum_{\ell=1}^{p_i}\frac{\Delta\alpha_i\,\hat{e}_i^I\,\lambda_i^{(\ell)}}{1+\Delta\alpha_i\,\hat{e}_i^I\,\lambda_i^{(\ell)}\,b_{i\ell}(\omega)}\,\mathbf{B}_{i\ell}(\omega), \quad \boldsymbol{\alpha}\in\boldsymbol{\alpha}^I = \left[\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}\right] \tag{47}$$

where $b_{i\ell}(\omega)$ and $\mathbf{B}_{i\ell}(\omega)$ are the complex functions defined in Eq.(33). Alternatively, the matrix $\mathbf{H}(\boldsymbol{\alpha},\omega)$ can be rewritten in a more suitable affine form, as follows:

$$\mathbf{H}(\boldsymbol{\alpha},\omega) = \mathbf{H}_0(\omega) + \sum_{i=1}^{r}\sum_{\ell=1}^{p_i}\left(a_{0,i\ell}(\omega) + \Delta a_{i\ell}(\omega)\,\hat{e}_i^I\right)\mathbf{B}_{i\ell}(\omega), \quad \boldsymbol{\alpha}\in\boldsymbol{\alpha}^I = \left[\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}\right] \tag{48}$$

where $a_{0,i\ell}(\omega)$ and $\Delta a_{i\ell}(\omega)$ are complex functions describing the midpoint and the deviation amplitude of the $i\ell$-th term in Eq.(47), given, respectively, by:

$$a_{0,i\ell}(\omega) = \frac{\left(\Delta\alpha_i\,\lambda_i^{(\ell)}\right)^2 b_{i\ell}(\omega)}{1-\left(\Delta\alpha_i\,\lambda_i^{(\ell)}b_{i\ell}(\omega)\right)^2}; \quad \Delta a_{i\ell}(\omega) = \frac{\Delta\alpha_i\,\lambda_i^{(\ell)}}{1-\left(\Delta\alpha_i\,\lambda_i^{(\ell)}b_{i\ell}(\omega)\right)^2}. \tag{49}$$

Equation (48) can be recast in the following form:

$$\mathbf{H}(\boldsymbol{\alpha},\omega) = \mathbf{N}_0(\omega) + \Delta\mathbf{N}(\boldsymbol{\alpha},\omega), \quad \boldsymbol{\alpha}\in\boldsymbol{\alpha}^I = \left[\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}\right] \tag{50}$$

where $\mathbf{N}_0(\omega)$ and $\Delta\mathbf{N}(\boldsymbol{\alpha},\omega)$ are the midpoint and the deviation matrices of the modal *FRF* defined in the context of the proposed *RSE*, respectively, as:

$$\mathbf{N}_0(\omega) = \mathbf{H}_0(\omega) + \sum_{i=1}^{r}\sum_{\ell=1}^{p_i}a_{0,i\ell}(\omega)\,\mathbf{B}_{i\ell}(\omega); \tag{51}$$

$$\Delta\mathbf{N}(\boldsymbol{\alpha},\omega)=\sum_{i=1}^{r}\sum_{\ell=1}^{p_i}\Delta a_{i\ell}(\omega)\hat{e}_i^I\,\mathbf{B}_{i\ell}(\omega),\quad \boldsymbol{\alpha}\in\boldsymbol{\alpha}^I=\left[\underline{\boldsymbol{\alpha}},\overline{\boldsymbol{\alpha}}\right]. \tag{52}$$

4.4. BOUNDS OF THE MODULUS OF THE NODAL INTERVAL FRF

The aim of this section is to determine the range of the modulus of the nodal interval *FRFs* of linear discretized structures with uncertain-but-bounded parameters. Once the modal *FRF* matrix is known, the square modulus of the *FRF* of the *p*-th DOF of the structural system can be defined as:

$$\left\|H_{N,pp}(\boldsymbol{\alpha},\omega)\right\|^2=\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\mathbf{H}^{*}(\boldsymbol{\alpha},\omega)\,\boldsymbol{\phi}_{0,p}\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\mathbf{H}(\boldsymbol{\alpha},\omega)\,\boldsymbol{\phi}_{0,p},\quad \boldsymbol{\alpha}\in\boldsymbol{\alpha}^I=\left[\underline{\boldsymbol{\alpha}},\overline{\boldsymbol{\alpha}}\right] \tag{53}$$

where $\boldsymbol{\phi}_{0,p}^{\mathrm{T}}$ is the *p*-th row of the modal matrix $\boldsymbol{\Phi}_0$ solution of the eigenproblem (5). Substituting Eq. (50) into Eq. (53), the following relationship is obtained:

$$\left\|H_{N,pp}(\boldsymbol{\alpha},\omega)\right\|^2=\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\left[\mathbf{N}_0^{*}(\omega)+\Delta\mathbf{N}^{*}(\boldsymbol{\alpha},\omega)\right]\boldsymbol{\phi}_{0,p}\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\left[\mathbf{N}_0(\omega)+\Delta\mathbf{N}(\boldsymbol{\alpha},\omega)\right]\boldsymbol{\phi}_{0,p},\quad \boldsymbol{\alpha}\in\boldsymbol{\alpha}^I=\left[\underline{\boldsymbol{\alpha}},\overline{\boldsymbol{\alpha}}\right]. \tag{54}$$

Aiming to evaluate the upper bound and the lower bound of the modulus of $H_{N,pp}(\boldsymbol{\alpha},\omega)$, Eq. (54) is rewritten as:

$$\left\|H_{N,pp}(\boldsymbol{\alpha},\omega)\right\|^2=\mathrm{mid}\left\|H_{N,pp}(\omega)\right\|^2+\mathrm{dev}\left\|H_{N,pp}(\boldsymbol{\alpha},\omega)\right\|^2,\quad \boldsymbol{\alpha}\in\boldsymbol{\alpha}^I=\left[\underline{\boldsymbol{\alpha}},\overline{\boldsymbol{\alpha}}\right] \tag{55}$$

where the symbols $\mathrm{mid}\|\bullet\|^2$ and $\mathrm{dev}\|\bullet\|^2$ denote the midpoint and the deviation of the square modulus of the interval nodal *FRF* defined in Eq.(54).

In order to simplify interval computations, higher-order terms are neglected, namely the term $\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\Delta\mathbf{N}^{*}(\boldsymbol{\alpha},\omega)\,\boldsymbol{\phi}_{0,p}\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\Delta\mathbf{N}(\boldsymbol{\alpha},\omega)\,\boldsymbol{\phi}_{0,p}$ in Eq. (54) is disregarded. According to this approximation, the midpoint and the deviation functions introduced in Eq.(55) can be written as:

$$\mathrm{mid}\left\|H_{N,pp}(\omega)\right\|^2\approx\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\mathbf{N}_0^{*}(\omega)\,\boldsymbol{\phi}_{0,p}\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\mathbf{N}_0(\omega)\,\boldsymbol{\phi}_{0,p}; \tag{56}$$

$$\mathrm{dev}\left\|H_{N,pp}(\boldsymbol{\alpha},\omega)\right\|^2\approx\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\mathbf{N}_0^{*}(\omega)\boldsymbol{\phi}_{0,p}\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\Delta\mathbf{N}(\boldsymbol{\alpha},\omega)\,\boldsymbol{\phi}_{0,p}+\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\Delta\mathbf{N}^{*}(\boldsymbol{\alpha},\omega)\,\boldsymbol{\phi}_{0,p}\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\mathbf{N}_0(\omega)\,\boldsymbol{\phi}_{0,p},\quad \boldsymbol{\alpha}\in\boldsymbol{\alpha}^I=\left[\underline{\boldsymbol{\alpha}},\overline{\boldsymbol{\alpha}}\right] \tag{57}$$

where $\mathbf{N}_0(\omega)$ and $\Delta\mathbf{N}(\boldsymbol{\alpha},\omega)$ are the midpoint and the deviation matrices introduced in Eqs.(51) and (52).

The lower bound, $\left\|\underline{H}_{N,pp}(\omega)\right\|^2$, and the upper bound, $\left\|\overline{H}_{N,pp}(\omega)\right\|^2$, of the square modulus of the nodal *FRF* of the *p*-th DOF can be evaluated, according to the philosophy of the affine arithmetic, as the minimum and maximum of the various combinations, i.e.:

$$\left\|\underline{H}_{N,pp}(\omega)\right\|^2 = \mathrm{mid}\left\|H_{N,pp}(\omega)\right\|^2 - \Delta\left\|H_{N,pp}(\omega)\right\|^2; \tag{58}$$

$$\left\|\bar{H}_{N,pp}(\omega)\right\|^2 = \mathrm{mid}\left\|H_{N,pp}(\omega)\right\|^2 + \Delta\left\|H_{N,pp}(\omega)\right\|^2. \tag{59}$$

In the previous equations the function $\Delta\left\|H_{N,pp}(\omega)\right\|^2$ is obtained upon substituting the matrix $\Delta\mathbf{N}(\boldsymbol{\alpha},\omega)$, defined in Eq.(52), into Eq.(57) and then deriving the maximum of the deviation $\mathrm{dev}\left\|H_{N,pp}(\boldsymbol{\alpha},\omega)\right\|^2$ according to the main properties of the interval analysis, i.e.:

$$\Delta\left\|H_{N,pp}(\omega)\right\|^2 = \sum_{i=1}^{r}\left|\sum_{\ell=1}^{p_i}\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\left[\mathbf{N}_0^*(\omega)\boldsymbol{\phi}_{0,p}\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\Delta a_{i\ell}(\omega)\mathbf{B}_{i\ell}(\omega) + \Delta a_{i\ell}^*(\omega)\mathbf{B}_{i\ell}^*(\omega)\boldsymbol{\phi}_{0,p}\boldsymbol{\phi}_{0,p}^{\mathrm{T}}\,\mathbf{N}_0(\omega)\right]\boldsymbol{\phi}_{0,p}\right|. \tag{60}$$

Notice that the function in square brackets is a real function and that the symbol $|\bullet|$ means absolute value. Obviously, the lower bound and the upper bound, $\left\|\underline{H}_{N,pp}(\omega)\right\|$ and $\left\|\bar{H}_{N,pp}(\omega)\right\|$, of the modulus of the nodal *FRF* of the *p*-th DOF can be obtained straightforwardly by taking the square root of Eqs. (58) and (59).

5. Numerical applications

5.1. TRUSS STRUCTURE WITH UNCERTAIN YOUNG'S MODULI

The first numerical application concerns the 24-bar truss structure depicted in Fig. 1. The Young's moduli of $r = 7$ bars are taken as uncertain parameters with fluctuations $|\alpha_i| < 1$ around the nominal value $E_0 = 2.1 \times 10^8$ kN/m^2, i.e. $E_i = E_0\,(1+\alpha_i)$, (i = 18, 19,..., 24). The cross-sectional areas of the bars are set equal to $A_i = 5 \times 10^{-4}$ m^2 while the lengths L_i (i = 1, 2,..., 24) can be deduced from Fig.1 where $L = 3$ m. Furthermore, each node possesses a lumped mass $M = 500$ Kg. Only the first $m = 8$ vibrations modes are retained in the modal analysis and the modal damping ratio has been assumed equal to $\zeta = 0.05$ for all the modes.

In Fig. 2, the exact *FRF* of the first modal coordinate, $H_{11}(\boldsymbol{\alpha},\omega)$, evaluated performing the inversion of the matrix into square brackets in Eq.(11) for $\alpha_i = \alpha = 0.05$, (i = 18, 19,..., 24) is compared with the corresponding approximate *FRF* obtained by applying the proposed *RSE* (Eq.(27)). Notice that a good matching of the exact *FRF* is achieved by retaining only the first-order terms in the *RSE*.

Figure 3 displays an analogous comparison for larger parameter fluctuations, say $\alpha_i = \alpha = 0.1$. As expected, in this case the proposed *RSE* truncated to first-order terms is less accurate, especially in the frequency range around the fundamental frequency of the system. Including second-order terms allows to improve the accuracy, as shown in the enlargement in Fig 3b.

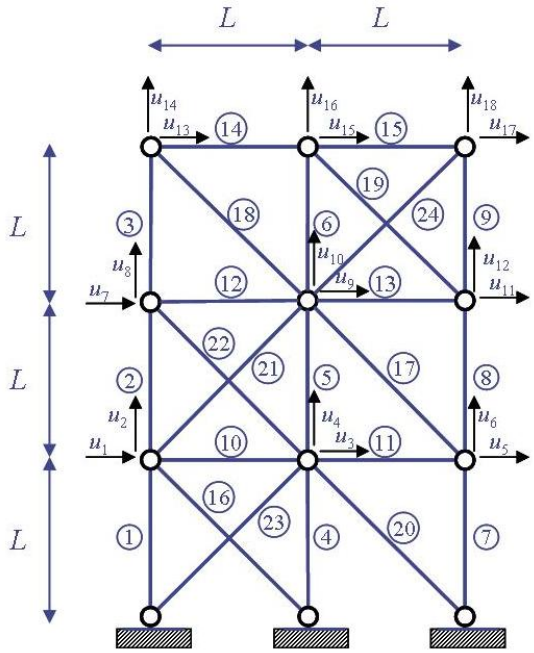


Figure 1. Truss structure with uncertain Young's moduli.

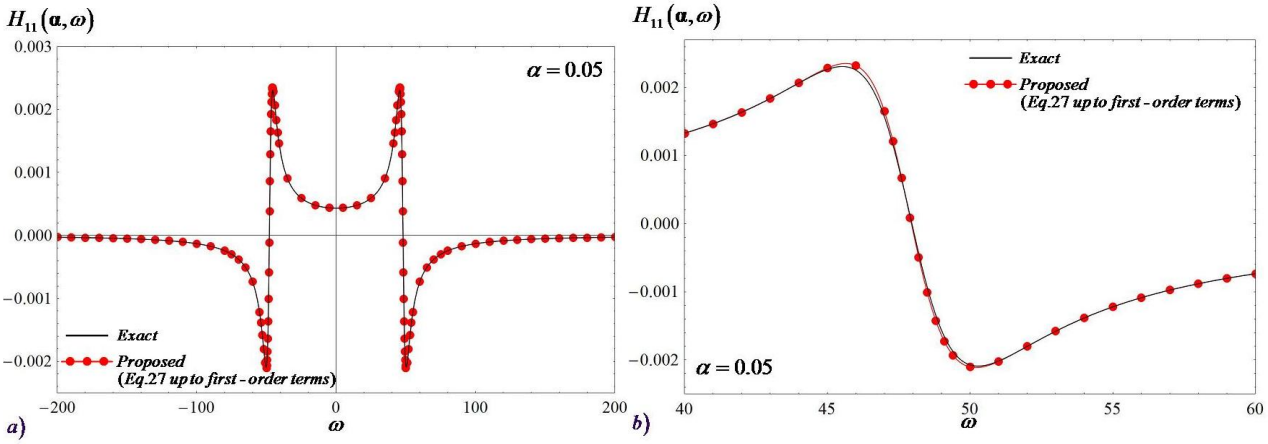


Figure 2. *FRF* of the first modal coordinate $H_{11}(\alpha,\omega)$: a) comparison between the exact *FRF* and the proposed *RSE* truncated to first-order terms; b) enlargement showing the convergence of the *RSE* close to the fundamental frequency of the structure ($\alpha = 0.05$).

In order to demonstrate the capability of the proposed explicit approximation of the *FRF* matrix to handle different uncertainty models, the fluctuating Young's moduli of the bars are now treated as interval variables i.e. $E_i^I = E_0(1+\Delta\alpha_i\,\hat{e}_i^I)$, ($i$ = 18, 19,..., 24), with symmetric deviations $\Delta\alpha_i = \Delta\alpha = 0.05$.

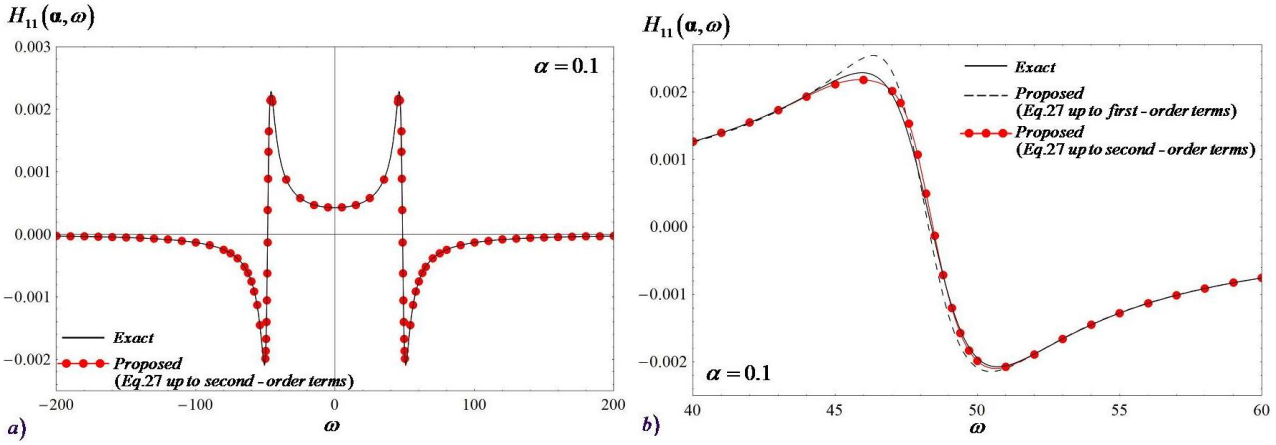


Figure 3. *FRF* of the first modal coordinate $H_{11}(\boldsymbol{\alpha},\omega)$: a) comparison between the exact *FRF* and the proposed *RSE* truncated to second-order terms; b) enlargement showing the convergence of the *RSE* close to the fundamental frequency of the structure ($\alpha = 0.1$).

In Figs.4 and 5, the upper bound and the lower bound of the modulus of the *FRF* of the nodal displacements u_1 and u_{13} of the truss (see Fig. 1), obtained by applying the proposed *RSE* truncated to first-order terms (see Eqs. (58) and (59)), are contrasted with the exact bounds. The latter are obtained following the philosophy of the *vertex method* (Muhanna and Mullen, 2001; Moens and Vandepitte, 2005), namely evaluating the modulus of the *FRF* for all the combinations of the bounds of the uncertain parameters and then taking at each frequency ω the maximum and minimum value among all the moduli of the *FRF* so obtained. Notice that the proposed estimates of the upper bound and lower bound of both $\|H_{N,11}(\boldsymbol{\alpha},\omega)\|$ and $\|H_{N,1313}(\boldsymbol{\alpha},\omega)\|$ are very close to the exact ones.

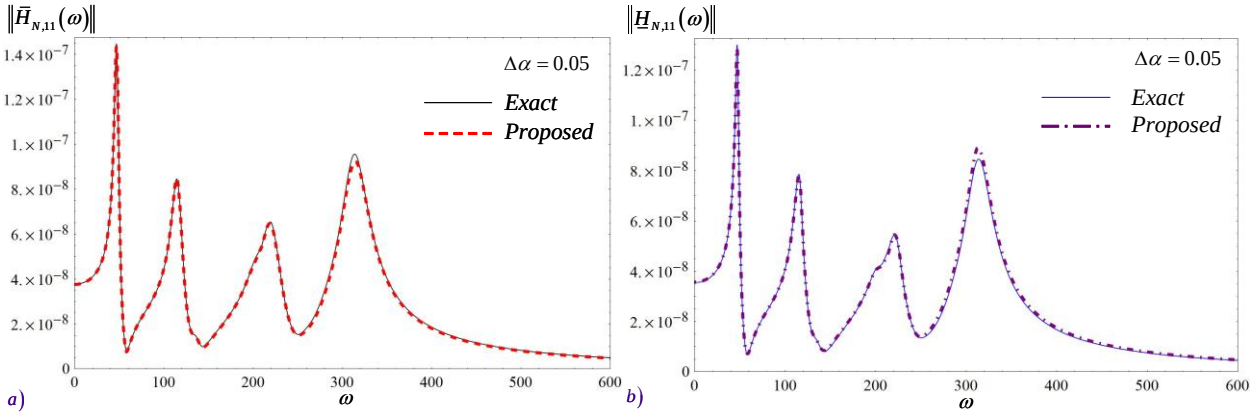


Figure 4. Comparison between the exact and proposed a) upper bound and b) lower bound of the modulus of the *FRF* of the nodal displacement u_1 ($\Delta\alpha = 0.05$).

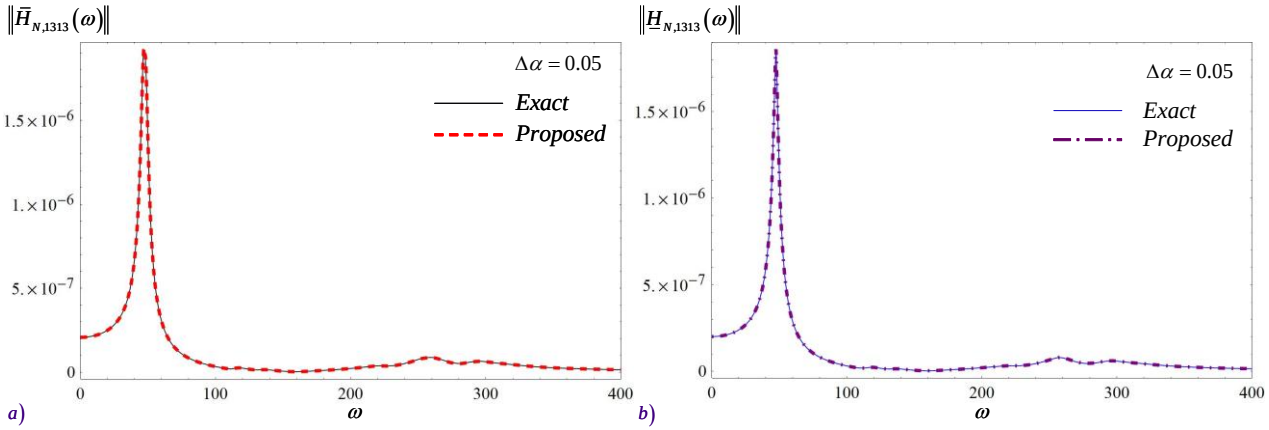


Figure 5. Comparison between the exact and proposed a) upper bound and b) lower bound of the modulus of the *FRF* of the nodal displacement u_{13} ($\Delta\alpha = 0.05$).

5.2. FLEXIBLE FRAME WITH UNCERTAIN YOUNG'S MODULI

As second application, a portal frame with uncertain Young's moduli is considered (see Fig. 6). It is assumed that the elastic moduli of the beam and columns exhibit fluctuations $|\alpha_i| < 1$ around the nominal value $E_0 = 2.85 \times 10^7$ kN/m^2, i.e. $E_i = E_0\,(1+\alpha_i)$, (i = 1, 2, 3). The geometrical properties of the portal frame are indicated in Fig.6, where $b = 0.30$ m, $h = 0.60$ m, $L = 3$ m and $H = 2$ m. Furthermore, each node possesses a lumped mass M=500 Kg. The modal damping ratio is set equal to $\zeta = 0.05$.

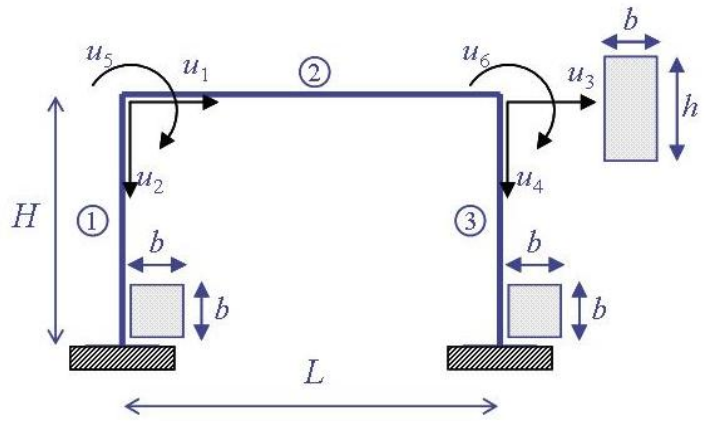


Figure 6. Portal frame with uncertain Young's moduli.

Figure 7 displays the comparison between the exact and approximate *FRFs* of the first modal coordinate, $H_{11}(\boldsymbol{\alpha},\omega)$, for $\alpha_i = \alpha = 0.05$, (i = 1, 2, 3). The convergence of the *RSE* can be detected by inspection of the enlargement in Fig. 6b, where different approximations obtained retaining terms up to the third-order are reported. It can be seen that the proposed *RSE* truncated to the third-order provides an accurate approximation of the *FRF* close to the fundamental frequency of the structure. The results pertaining to

larger parameter fluctuations, $\alpha_i = \alpha = 0.10$, shown in Figure 8, demonstrate the accuracy of the proposed *RSE* even for high uncertainty levels. Obviously, in this case higher-order terms of the *RSE* play an increasing important role.

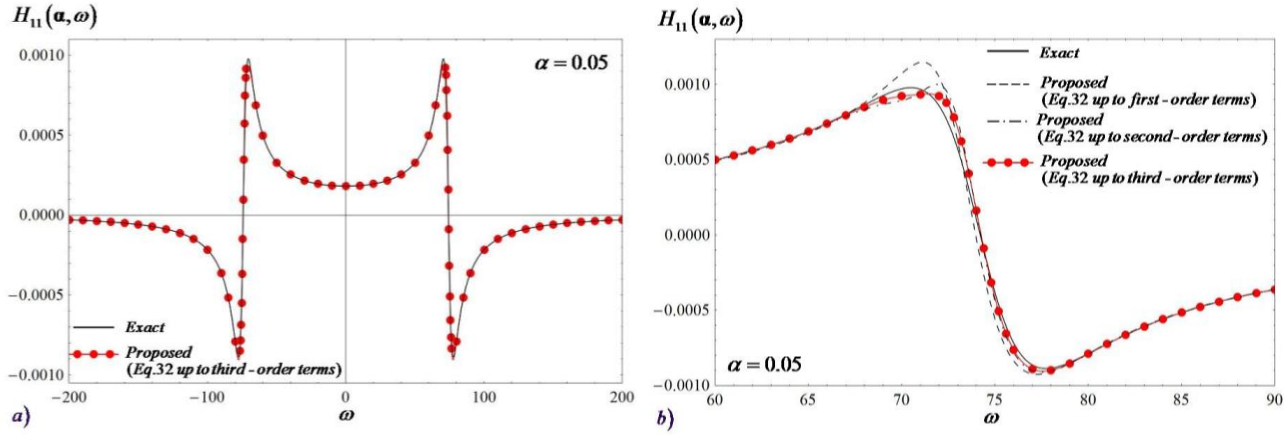


Figure 7. FRF of the first modal coordinate $H_{11}(\boldsymbol{\alpha},\omega)$: a) comparison between the exact *FRF* and the proposed *RSE* truncated to third-order terms; b) enlargement showing the convergence of the *RSE* close to the fundamental frequency of the structure ($\alpha = 0.05$).

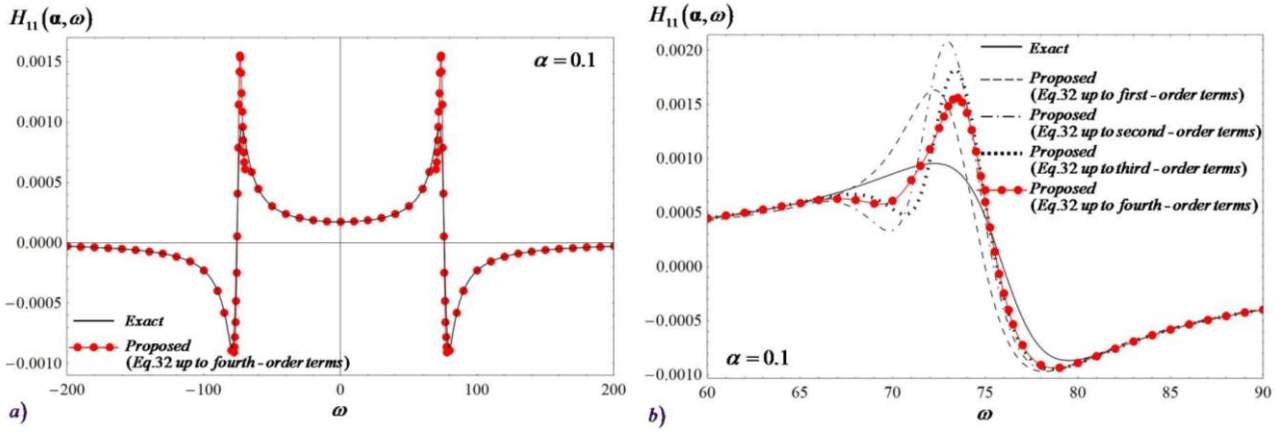


Figure 8. FRF of the first modal coordinate $H_{11}(\boldsymbol{\alpha},\omega)$: a) comparison between the exact *FRF* and the proposed *RSE* truncated to fourth-order terms; b) enlargement showing the convergence of the *RSE* close to the fundamental frequency of the structure ($\alpha = 0.10$).

Finally, the fluctuating Young's moduli of the beam and columns are modelled as uncertain-but-bounded parameters i.e. $E_i^I = E_0(1+\Delta\alpha_i\,\hat{e}_i^I)$, ($i$ = 1, 2, 3), with symmetric deviations $\Delta\alpha_i = \Delta\alpha = 0.05$. Figure 9 displays the comparison between the upper bound and the lower bound of the modulus of the *FRF* of the nodal displacement u_1, $\|H_{N,11}(\boldsymbol{\alpha},\omega)\|$, obtained by applying the proposed *RSE* truncated to first-order terms (see Eqs. (58) and (59)), and the exact bounds evaluated following the philosophy of the *vertex method*. It

can be observed that also in the case of flexible frames the *RSE* provides accurate estimates of the upper bound and the lower bound of the modulus of the *FRF*.

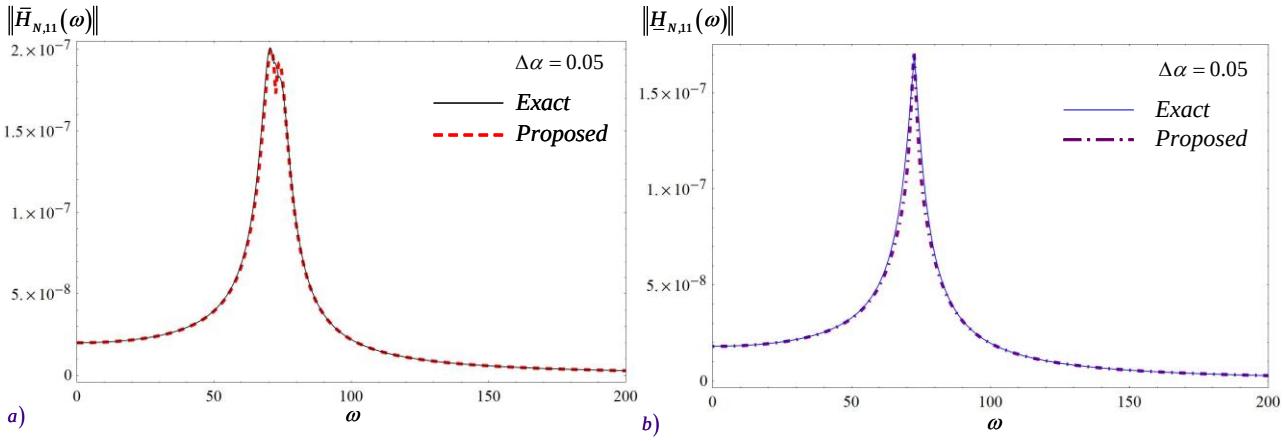


Figure 9. Comparison between the exact and proposed (a) upper bound and (b) lower bound of the modulus of the *FRF* of the nodal displacement u_1 ($\Delta\alpha = 0.05$).

6. Concluding remarks

The evaluation of the *frequency response function* (*FRF*) matrix of linear structures with uncertain stiffness properties has been addressed. Specifically, a procedure for deriving the *FRF* matrix in explicit approximate form has been presented. The proposed method relies on the spectral decomposition of the deviation of the stiffness matrix (with respect to its nominal value) which allows to obtain a sum of rank-one matrices, each one associated to a single uncertain parameter. Then, the equations of motion are projected in the modal subspace and, after some algebra, the Neumann series expansion of the *FRF* matrix is rewritten in an alternative explicit form, herein called *Rational Series Expansion* (*RSE*). The proposed *RSE* represents a useful tool for performing the frequency domain analysis of linear structures with uncertain parameters since it enables one to derive closed form expressions of the response and then investigate the effects of the fluctuating parameters. The latter can be modeled resorting either to probabilistic or non-probabilistic approaches depending on the available information on their variability.

The accuracy of the proposed *RSE* has been assessed by analyzing a truss structure and a portal frame with uncertain Young's moduli. Numerical results have shown that the estimates of the *FRF* provided by the *RSE* are very close to the exact ones even for large fluctuations of the uncertain parameters. The versatility of the proposed *RSE* has been demonstrated by modeling the fluctuating Young's moduli as uncertain-but-bounded parameters. The estimates of the upper bound and lower bound of the modulus of the *FRF* derived by applying the *RSE* in conjunction with the so-called *improved interval analysis* have been shown to be in good agreement with the exact bounds evaluated following the philosophy of the *vertex method*.

References

Alefeld, G. and J. Herzberger. *Introduction to Interval Computations*, Academic Press, New York, 1983.

Comba, J. L. D. and J. Stolfi. Affine Arithmetic and Its Applications to Computer Graphics, *Anais do VI Simposio Brasileiro de Computaao Grafica e Processamento de Imagens (SIBGRAPI'93*", Recife (Brazil), October, 9–18, 1993.

Elishakoff, I. and M. Ohsaki. *Optimization and Anti-Optimization of Structures under Uncertainties*, Imperial College Press, London, 2010.

Falsone, G. and G. Ferro. A Method for the Dynamical Analysis of FE Discretized Uncertain Structures in the Frequency Domain. *Computer Methods in Applied Mechanics and Engineering* 194: 4544–4564, 2005.

Falsone, G. and G.Ferro. An Exact Solution for the Static and Dynamic Analysis of FE Discretized Uncertain Structures. *Computer Methods in Applied Mechanics and Engineering* 196: 2390–2400, 2007.

Hansen, E. R. A Generalized Interval Arithmetic. In: K. Nicket Editor, *Interval Mathematics, Lect. Notes Comput. Sc.* 29: 7-18, 1975.

Manson, G. Calculating Frequency Response Functions for Uncertain Systems Using Complex Affine Analysis. *Journal of Sound and Vibration* 288: 487–521, 2005.

Moens, D. and D. Vandepitte. An Interval Finite Element Approach for the Calculation of Envelope Frequency Response Functions. *International Journal for Numerical Methods in Engineering* 61: 2480–2507, 2004.

Moens, D. and D. Vandepitte. A Survey of Non-Probabilistic Uncertainty Treatment in Finite Element Analysis. *Computer Methods in Applied Mechanics and Engineering* 194: 1527–1555, 2005.

Moore, R. E. *Interval Analysis*, Prentice-Hall, Englewood Cliffs, 1966.

Moore, R. E., R. B. Kearfott and M. J. Cloud. *Introduction to Interval Analysis*, SIAM, Philadelphia, USA, 2009.

Muhanna, R. L. and R. L. Mullen. Uncertainty in Mechanics: Problems-Interval-Based Approach. *Journal Engineering Mechanics ASCE* 127: 557-566, 2001.

Muscolino, G. and A. Sofi. Stochastic Analysis of Structures with Uncertain-But-Bounded Parameters via Improved Interval Analysis. *Probabilistic Engineering Mechanics*, doi:10.1016/j.probengmech.2011.08.011, in press.

Nedialkov, N. S., V. Kreinovich and S. A. Starks. Interval Arithmetic, Affine Arithmetic, Taylor Series Methods: Why, What Next? *Numerical Algorithms* 37: 325-336, 2004.

Neumaier, A. *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, UK, 1990.

Stolfi, J. and L.H. De Figueiredo. An Introduction to Affine Arithmetic, *TEMA Tend. Mat. Apl. Comput.*, 4: 297-312, 2003.

Reliable/reliability computing for concrete structures: Methodology and software tools

[1)]Drahomír Novák, [2)]Radomír Pukl
[1)]*Brno University of Technology, Brno, Czech Republic, novak.d@fce.vutbr.cz*
[2)]*Cervenka Consulting, Prague, Czech Republic, radomir.pukl@cervenka.cz*

Abstract: The paper presents a complex methodology for statistical and reliability analyses of concrete structures. It describes the virtual simulation concept and tools used on the way from assessment of experimental results to reliability analysis. The methodology is represented by consequent steps starting by material parameters identification based on artificial neural networks and finite element modelling. The aim is finally to perform advanced reliability assessment using appropriate stochastic finite element model. Software tools are briefly described. Selected examples of application illustrate the applicability of the approach.

Keywords: Reliability, concrete, inverse analysis, artificial neural networks, nonlinear analysis, fracture mechanics, simulation.

1. Introduction

Reliable computing for reliability assessment requires combination of advanced techniques to treat both nonlinearity and uncertainty. A large number of efficient stochastic analysis methods have been developed during last years. The common feature of all methods is the fact that they require a repetitive evaluation (simulations) of the response or limit state functions. The development of reliability methods is from the historical perspective certainly a struggle to decrease an excessive number of simulations. In spite of the increasing capabilities of computer hardware using a large number of simulations is still a problem when dealing with computationally demanding tasks and small-sample simulation is needed.

The objective of the contribution is to present methods and software for efficient statistical, sensitivity and reliability assessment implemented in FReET software (Novák et al. 2011). The attention is given to those techniques that are developed for analyses of computationally intensive problems like nonlinear FEM. Sensitivity analysis is based on nonparametric rank-order correlation. Statistical correlation is imposed by the simulated annealing. As software development is performed in a complex project and system for reliability assessment of concrete structures SIMSOFT, the full role of software FReET will be also shortly described – including degradation module FReET-D and methodology for inverse analysis and identification.

The paper presents briefly a complex methodology for statistical, reliability and risk analyses of concrete structures. But the methodology is valid generally, not only for concrete structures. It describes the virtual simulation concept and tool used on the way from assessment of experimental results to reliability analysis. The whole approach is based on small-sample randomization of nonlinear fracture mechanics finite element analysis of reinforced concrete structures. Efficient techniques of both nonlinear numerical analysis of concrete structures and stochastic simulation methods have been combined in order to offer an

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

advanced tool for assessment of realistic behaviour of concrete structures from reliability and risk points of view.

2. Methodology

The stochastic response requires repeated analyses of the structure with stochastic input parameters, which reflects randomness and uncertainties in the input values. The system uses the nonlinear computer simulation for realistic prediction of structural response and its resistance. Nonlinear fracture mechanics simulation utilizes state of art techniques including: damage mechanics, fracture mechanics and plasticity material models, smeared crack approach - fictitious crack, crack band method, softening of concrete in both tension and compression, combination of nonlinear concrete behavior with discrete and smeared reinforcement in reinforced concrete and pre-stressed structures. As the nonlinear structural analysis is computationally very demanding, a suitable technique of statistical sampling should be utilized, which allows relatively small number of simulations. Final results are: statistical characteristics of response (stresses, deflections, crack width etc.), information on dominating and non-dominating variables (sensitivity analysis) and estimation of reliability using reliability index and theoretical failure probability).
In order to use appropriate parameters of material laws in the computational model, an inverse analysis based on experiments in a laboratory or in situ has to be performed. A suitable technique for the inverse analysis is the stratified sampling scheme for the modeling of uncertain model parameters combined with artificial neural networks.

The procedure can be outlined as follows:

- experiment (laboratory, in situ);
- development of a deterministic computational model to capture the experiment;
- inverse analysis to obtain parameters of the computational model;
- deterministic computational model of a structure;
- stochastic model of a structure;
- statistical, sensitivity and reliability analyses of a structure.

3. Key soft computing methods

3.1. Material parameters identification

The basic step for efficient nonlinear FEM modeling is to solve the inverse problem: “Which material model parameters should be used to capture the experiment well?” The recently proposed identification strategy is based on a coupling of the stratified sampling in the nonlinear fracture mechanics analysis and in the artificial neural network (Novák & Lehký 2006). The fundamental scheme of the approach is shown in Fig. 1; the neural network is trained by the values of the load-deflection curve and the values of identified parameters (considered to be random variables) in a repeated stochastic way – the preparation of a training set for a neural network uses stratified simulation. A multiple calculation of a deterministic computational

model using random realizations of model parameters is performed resulting in a "bundle" of load-deflection curves (usually overlapping the experimental curve). Realizations of the load-deflection curves serve as a basis for the training of an appropriate artificial neural network. Such training can be called stochastic training due to the stochastic origin of the load-deflection curves. After the training procedure, the neural network is ready for the key task: to select the material model parameters which can capture the experimental load-deflection curve as closely as possible.

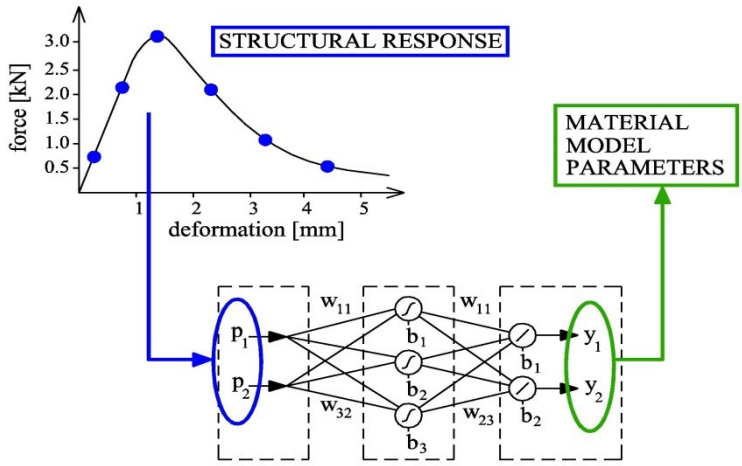


Figure 1. Scheme of inverse analysis.

3.2. UNCERTAINTIES SIMULATION

For time-intensive calculations such as those involving nonlinear fracture mechanics of concrete, small-sample simulation techniques based on stratified sampling of the Monte Carlo type represent a rational compromise between feasibility and accuracy. Therefore, Latin hypercube sampling (LHS) was selected as a key fundamental technique.

The method belongs to the category of stratified simulation methods (e.g. Mc Kay & Conover 1979, Novák et. al 1998). It is a special type of Monte Carlo simulation which uses the stratification of the theoretical probability distribution function of input random variables. It requires a relatively small (tens or hundreds) number of simulations (repetitive calculations of the structural response) to estimate the requested statistics of the response.

The basic feature of LHS is that the probability distribution functions for all random variables are divided into N_{Sim} equivalent intervals (N_{Sim} is the number of simulations); the values from the intervals are then used in the simulation process (random selection, middle of interval or mean value). This means that the range of the probability distribution function of each random variable is divided into intervals of equal probability. The samples are chosen directly from the distribution function based on an inverse transformation of the distribution function.

It has been proved that the best LHS strategy, which simulates the means and variances very well, is the approach suggested e.g. by Keramat & Kielbasa (1997) or Huntington & Lyrintzis (1998). The sample of each interval is chosen as the mean (Fig. 2):

$$x_{i,k} = \frac{\int_{y_{i,k-1}}^{y_{i,k}} x \cdot f_i(x)\,dx}{\int_{y_{i,k-1}}^{y_{i,k}} f_i(x)\,dx} = N_{Sim} \cdot \int_{y_{i,k-1}}^{y_{i,k}} x \cdot f_i(x)\,dx \quad , \tag{1}$$

where f_i is the PDF of variable X_i, and the integration limits are:

$$y_{i,k} = F_i^{-1}\left(\frac{k}{N_{Sim}}\right) . \tag{2}$$

A robust technique for imposing statistical correlation based on the stochastic method of optimization, called simulated annealing, has been proposed recently by Vořechovský & Novák (2009). The imposition of the prescribed correlation matrix into the sampling scheme can be understood as an optimization problem.

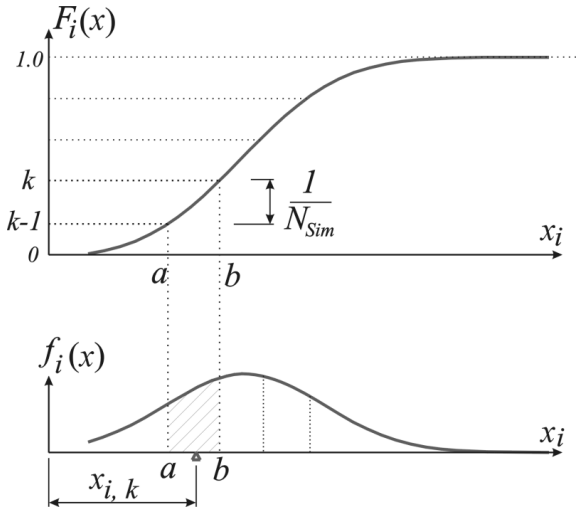


Figure 2. Illustration of sampling.

3.3. RELIABILITY ANALYSIS

In cases when we are constrained by a small number of simulations (tens, hundreds) it can be difficult to estimate the failure probability. The following approaches are therefore utilized here; they are

approximately ordered from elementary (an extremely small number of simulations, inaccurate) to more advanced techniques:

- Cornell´s reliability index – the calculation of a reliability index from the estimation of the statistical characteristics of the safety margin;
- curve-fitting approaches – based on the selection of the most suitable probability distribution of the safety margin;
- FORM approximation (Hasofer-Lind´s index);
- importance sampling techniques;
- response surface methods.

These approaches are well known in reliability literature, and also the provision of all details is beyond the aim of this paper. In spite of the fact that the calculation of the failure probability (and/or reliability index) using some of these techniques does not always belong to the category of very accurate reliability techniques (the first three in the list), they represent a feasible alternative in many practical cases.

4. Software tools

4.1. SARA – COMPLEX SOLUTION

The authors combined efficient techniques of both nonlinear numerical analysis of engineering structures and stochastic methods to offer an advanced tool for the reliability assessment of concrete structures. Within the framework of this complex system attention is also paid to the modeling of degradation phenomena, such as carbonation of concrete, corrosion of reinforcement, chloride attack, etc. The combination of all parts (structural analysis, reliability assessment, inverse analysis and degradation modeling) is presented together as the SARA software. The recently developed version of the SARA software is called RLACS or SARA Science; its structure is similar to the SARA system, but it incorporates an extended version of the ATENA NLFEM software: ATENA Science.

A representation of the program combination within SARA software is presented in Fig. 3. It includes: SARA (Bergmeister et al. 2004, Pukl et al. 2003a,b; Strauss et al. 2008; Novák et al. 2005) – a software shell which controls the communication between following individual programs: ATENA (Červenka et al. 2007) – FEM nonlinear analysis of concrete structures; FReET (Novák et al. 2011) – the probabilistic engine based on LHS simulation; DLNNET (Lehký 2011; Novák & Lehký 2006) – artificial neural network software; FReET-D (Teplý et al. 2011) – degradation module based on FReET. The fundamental version of ATENA is called ATENA Engineering; its native GUI is directly integrated into the SARA system as shown in Fig. 3. Recent development of ATENA represents ATENA Science package. Within the reliability analysis it is controlled by RLACS Studio through special commands in ATENA input files.

4.2. FReET – UNCERTAINTIES SIMULATION

The probabilistic software FReET (Novák et al. 2011) allows simulations of uncertainties of the analyzed problem basically at random variables level (typically in civil/mechanical engineering – material properties,

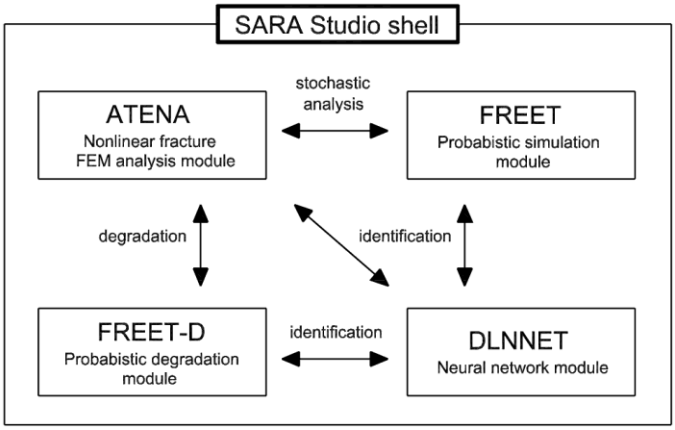


Figure 3. The program combination within SARA software.

loading, geometrical imperfections, environment). The attention is given to those techniques that are developed for analyses of computationally intensive problems; nonlinear FEM analysis being a typical example. Stratified simulation technique Latin hypercube sampling (LHS) is used in order to keep the number of required simulations at an acceptable level (Novák et al. 1998). This technique can be used for both random variables' and random fields' levels.

Statistical correlation is efficiently imposed by the stochastic optimization technique – the simulated annealing (Vořechovský & Novák 2009). Sensitivity analysis is based on nonparametric rank-order correlation coefficients and may serve e.g. for model reduction in subsequent analyses. State-of-the-art probabilistic algorithms are implemented to compute the probabilistic response and reliability generally, including durability limit states.

4.3. FReET-D – DEGRADATION SIMULATION

There are many predictive computational models for degradation modelling mainly carbonation of concrete, chloride ingress and corrosion of reinforcement at different sophistication levels. Frequently, heuristic models are employed using more or less simplified approaches and data. Common feature of all these models is that input data are very uncertain. There is a software implementation where all relatively well-known models are summarized within the framework of unified software environment. It is called FReET-D where a combination of analytical models and simulation techniques has been amalgamated to form specialized software for assessing the potential degradation of newly designed as well as existing concrete structures (Teplý et al. 2011, 2012; Veselý et al. 2010). Models implemented (mainly simple-to-use "point-in-space" probabilistic models) for carbonation, chloride ingress, corrosion of reinforcement and others which may serve directly in the durability assessment of concrete structures in the form of a durability limit states, i.e. the assessment of service life and the level of the relevant reliability measure. Several features are offered including parametric studies and Bayesian updating. Altogether, 32 models are implemented as pre-defined dynamic-link library functions. FReET-D actually represents a specialized module of FReET software (Novák et al. 2011), mentioned above.

4.4. ATENA – NONLINEAR SIMULATION

The ATENA software (Červenka et al. 2002, 2007) was developed for realistic simulation of reinforced concrete structures. It is based on the finite element method with non-linear material models, and utilized for analysis of beams and girders, plates and shells, bridges, tunnels, dams, composite structures, strengthening, structural details, fastenings, fibre reinforced structures and masonry structures etc.

The ATENA software consists of calculating core ensuring the non-linear numerical analysis, and a user-friendly graphical interface for an efficient communication between end-user and program core. The numerical core covers the finite element technology, non-linear material models and non-linear solution. The non-linear material models are based on the orthotropic damage theory and special concrete-related theory of plasticity. As one of the main features the non-linear fracture mechanics is employed for concrete cracking in tension. Based on the fracture energy approach the tensile cracks are modeled as smeared material damage which enables utilization of the continuum mechanics even for the damaged material. Objectivity of the solution is ensured using crack band method. The material law exhibits softening after reaching the tensile strength. The behavior of concrete in compression is defined by special theory of plasticity (three-parameter model), with non-associated plastic flow rule and softening. This material model for concrete can successfully reproduce also other important effect, such as volume change under plastic compression or compressive confinement. The native graphical user-interface supports all the specifics of reinforced concrete, e.g. input of discrete reinforcing bars, or evaluation of crack patterns in the damaged structural model.

The new ATENA software class ATENA Science (www.cervenka.cz) enables time-dependent (dynamic, fatigue) and temperature-dependent (fire resistance) nonlinear analysis of complex concrete structures. In the last version the ATENA Science is equipped with a new user friendly interface shell called ATENA Studio.

5. Examples of application

The complex methodology and software has been applied mainly for reliability analysis of concrete bridges, e.g. Pukl et al. 2003ab, Lehký et al. 2010, Podroužek et al. 2010, Strauss et al. 2008. Detailed description of particular application for deteriorated bridge structure can be found in fib bulletin 62 Structural concrete, section 9.19., that example can be regarded as the most elaborated one.

The interesting application is analysis of facade panels made of alternative FRC-material. As it represents a new facade system utilizing a new composite material which exhibits a large variability, computational analysis was desirable to address reliability issues connected with this special structure (Keršner et al. 1997).

A 3D FEM computational model has been developed using ATENA 3D Engineering nonlinear fracture mechanics software (Červenka et al. 2007). Wind intake was simulated by continuous loading. The Newton-Raphson solution method with a loading increment step of 1 kN/m^2 provided a non-linear solution to obtain ultimate load, cracks at final stage are depicted in Fig. 4. A 3D cementitious material model was used with material parameters identified by identification technique based on artificial neural networks. All of the input basic random variables involved and the particular set of their statistical parameters (mean

value, COV, probability distribution function (PDF)) are summarized in Table I for both the reference panel (R) and the degraded panel (D).

Statistical simulations were performed using parameters from Table 1, the resulting bundle of *l–d* curves is shown in Fig. 5 and ultimate load statistics were evaluated. Action of load – wind intake was considered deterministic at several levels up to 20 kN/m^2. The theoretical failure probability – the probability that the panel will not resist the load (wind intake) was calculated using mathematical model of a PDF. The results of this reliability study are shown in Fig. 6.

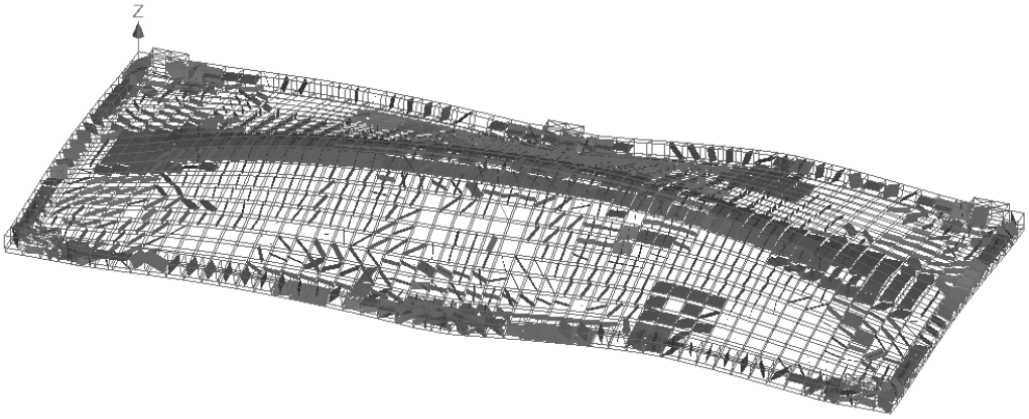


Figure 4.Cracks at failure of panel.

Table I. Basic random variables of fibre-reinforced concrete.

Variable	Unit	Mean		COV	PDF
Modulus of elasticity	GPa	10.1	R	0.195	Rayleigh
		7.8	D	0.199	Weibull min
Compressive strength	MPa	53.5	R	0.250	Log-normal
		31.5	D	0.250	Log-normal
Tensile strength	MPa	6.50	R	0.250	Weibull min
		3.81	D	0.250	Weibull min
Fracture energy	J/m^2	816.2	R	0.383	Weibull max
		195.8	D	0.418	Log-normal

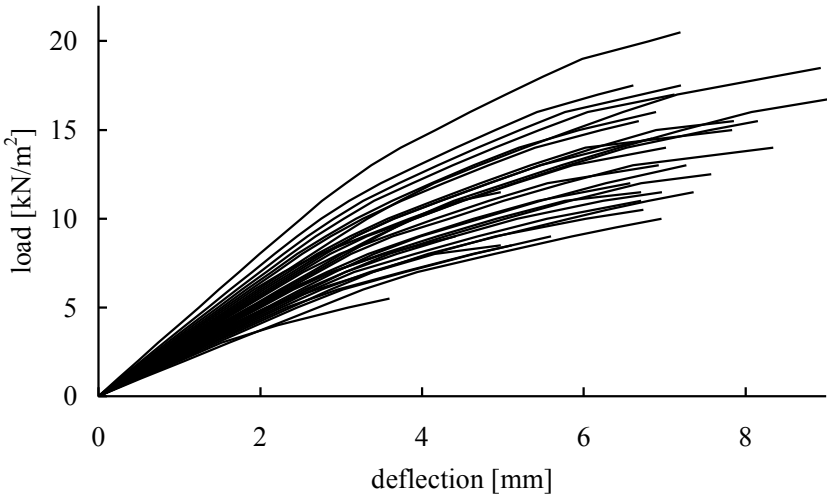


Figure 5. Random *l–d* curves – the reference facade panel.

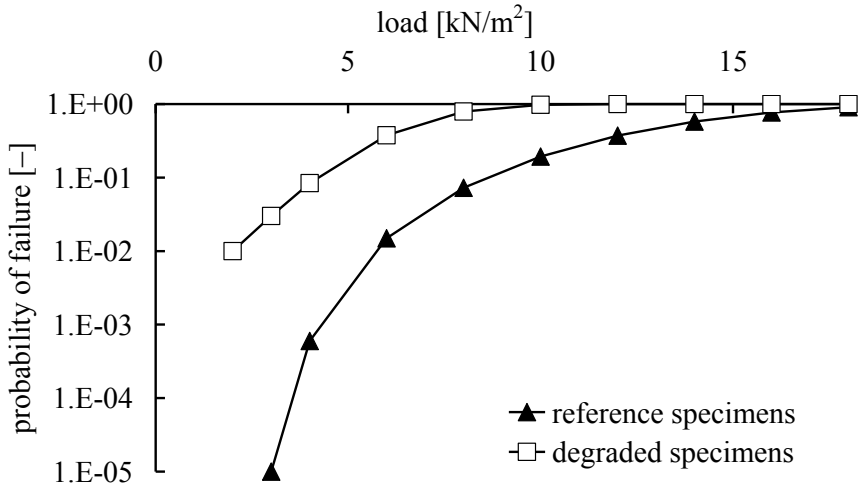


Figure 6. Theoretical failure probabilities for different levels of load – wind intake.

6. Conclusions

Virtual simulation concept and tools used on the way from assessment of experimental results to reliability analysis are briefly presented. The advanced methods for nonlinear, stochastic, reliability and degradation analysis were integrated into software package usable for complex reliability assessment of engineering

structures, which enables realistic simulation of the structural behaviour, damage and failure accounting uncertainties of input parameters, nonlinear material response and material deterioration. The presented tools and methods have been used in numerous practical applications of analysis, design, and life time assessment of concrete bridges, buildings, tunnels, power plants and other civil engineering structures.

Acknowledgements

The presented research is funded by research project No. TA01011019 of Technological Agency of Czech Republic "SimSoft" and by the Czech science foundation grant No. P105/10/1156 "Comocos". Authors gratefully acknowledge this support.

References

Bergmeister, K., Costa, C., Novák, D., Santa, U., Strauss, A., Teplý, B. and Wendner, R. Case study – Reliability assessment of a bridge subject to chloride induced corrosion and associated structural deterioration. Section 9.19 In: *Structural Concrete, Textbook on behavior, design and performance*, second edition, Volume 5, fib bulletin No. 62, 2012.

Bergmeister, K., Novák, D. and Pukl, R. An advanced tool for reliability assessment of concrete structures. *Progress in Structural Engineering, Mechanics and Computation, Zingoni (ed.), Proc. of 2nd int. conf*., Cape Town, South Africa, Taylor&Francis Group: 899-904, 2004.

Červenka, V., Jendele, L. and Červenka, J. ATENA Program Documentation – Part 1: Theory. Cervenka Consulting, Prague, Czech Republic, 2007.

Červenka, V., Červenka, J. and Pukl, R. ATENA – A tool for engineering analysis of fracture in concrete, *Sadhana* **27**(4): 485-492, 2002.

Huntington, D.E. and Lyrintzis, C.S. Improvements to and limitations of Latin hypercube sampling. *Probabilistic Engineering Mechanics* Vol. 13, No. 4: 245-253, 1998.

Keramat, M. and Kielbasa, R. Efficient Average Quality Index of Estimation of Integrated Circuits by Modified Latin Hypercube Sampling Monte Carlo (MLHSMC). *Proc. of IEEE Symp. on Circuits and Systems, Hong Kong,* 1997.

Keršner, Z., Novák, D., Řoutil, D. and Podroužek, J. Stochastic nonlinear Analysis of concrete structures – Part II: Application to fiber-reinforced concrete facade panels. *Proc. of ICASP '07 the Int. Conf. on App. of Safety and Probability*, Japan, 2007.

Lehký, D. DLNNET – program documentation. Theory and User's Guides, Brno, Czech Republic, (in preparation), 2011.

Lehký, D., Keršner, Z. and Novák, D. Jointless Bridge: Determination of fracture mechanical parameters values for nonlinear analysis. *Proc. of IABMASS*, Philadelphia, USA, 2010.

McKay, M.D., Conover, W.J. and Beckman, R.J. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* Vol. 21: 239-245, 1979.

Novák, D., Teplý, B. and Keršner, Z. The role of Latin Hypercube Sampling method in reliability engineering. *Proc. of ICOSSAR-97*, Kyoto, Japan, 403-409, 1998.

Novák, D., Vořechovský, M., Lehký, D., Rusina, R., Pukl, R. and Červenka, V. Stochastic nonlinear fracture mechanics finite element analysis of concrete structures. *Proc. of ICoSSaR '05 Int. Conf. on Struct. Safety and Reliability*, Italy, 781-788. 2005.

Novák, D. and Lehký, D. ANN inverse analysis based on stochastic small-sample training set simulation. *J. of Engineering Application of Artificial Inteligence*, 19, 731-740, 2006.

Novák, D., Vořechovský, M. and Rusina, R. FReET v.1.5 – program documentation. User´s and Theory Guides. Brno/Červenka Consulting, Czech Republic, http://www.freet.cz , 2011.

Podroužek, J., Novák, D. and Strauss, A. Jointless Bridge: Reliability assessment. *Proc. of IABMASS*, Philadelphia, USA, 2010.

Pukl, R., Červenka, V., Strauss, A., Bergmeister, K. and Novák, D. An advanced engineering software for probabilistic-based assessment of concrete structures using nonlinear fracture mechanics. *Proc. of 9th Int. Conf. on Applications of Statistics and Probability in Civil Engineering – ICASP* 9, San Francisco, USA, Rotterdam Millpress: 1165-1171, 2003a.

Pukl, R., Novák, D. and Bergmeister, K. Reliability assessment of concrete structures. *Proc. of Computational modelling of concrete structures (Euro-C2003),* N. Bicanic et al. (eds), St. Johann, Austria: 793-803. Lisse:Balkema, 2003b.

Strauss, A., Bergmeister, K., Hoffmann, S., Pukl, R. and Novák, D. Advanced life-cycle analysis of existing concrete bridges. *Journal of Materials in Civil Engineering*, Vol. 20, 9-19, 2008.

Teplý, B., Chromá, M., Rovnaník, P. and Novák, D. Role of modeling in probabilistic durability assessment of concrete structures. *IALCCE2012 Symposium*, Vienna, Austria (in press), 2012.

Teplý, B, Chromá, M, Vořechovská, D. and Rovnaník, P. FReET-D. Program documentation, Part 1: Theory, Brno/Červenka Consulting, Czech Republic, http://www.freet.cz, 2011.

Veselý, V., Teplý, B., Chromá, M., Vořechovská, D. and Rovnaník, P.: FReET-D. version 1.2, Program Documentation, Part 2, User Manual, Revision 08/2010, Brno/Červenka Consulting, Czech Republic, http://www.freet.cz, 2010.

Vořechovský, M. and Novák, D. Correlation control in small-sample Monte Carlo type simulations I: A simulated annealing approach. *Probabilistic Engineering Mechanics* 24, 452-462, 2009.

Importance sampling strategy for oscillatory stochastic processes

Jan Podrouzek
Centre for Water Resource Systems (CWRS), Vienna University of Technology, Austria,
podrouzek@waterresources.at

Abstract: This paper contributes to the structural reliability problem by presenting a novel approach that enables for identification of stochastic oscillatory processes as a critical input for given mechanical models. Identification development follows a transparent image processing paradigm completely independent of state-of-the-art structural dynamics, aiming at delivering a simple and wide purpose method. Validation of the proposed importance sampling strategy is based on multi-scale clusters of realizations of digitally generated non-stationary stochastic processes. Good agreement with the reference pure Monte Carlo results indicates a significant potential in reducing the computational task of first passage probabilities estimation, an important feature in the field of e.g. probabilistic seismic design or risk assessment generally.

Keywords: Stochastic process, Critical excitation, Reliability analysis, Importance sampling, Image processing, Pattern recognition, Identification problem

1. Introduction

The necessity for adopting probabilistic design concepts has become imperative among the structural static problems (Ang and Tang, 1990; Haldar and Mahadevan, 2000; Melchers 2001). On the other hand, structural dynamics is still far from practical utilizations of such concepts despite cheap contemporary computational costs. Among the main reasons is the uncertain nature of environmental loading that has to be modelled as a time-varying phenomena, represented in this paper by non-stationary stochastic oscillatory process as an analogy to earthquake event.

It is a well accepted fact that structures respond in a very uncertain manner to different ground motion events while there is very limited a priori knowledge on the structural behaviour. Same applying for models, an implication is the necessity to perform the structural analysis for each realization of the event separately, which makes the Monte-Carlo based reliability analysis computationally unfeasible for realistic assumptions, i.e. small probabilities and large sample sizes.

There have been several recent attempts to avoid such reliability problem in its full form. Moustafa (2011) proposed a framework for deriving optimal earthquake loads expressed as a Fourier series. More widely, critical excitation methodologists propose to identify critical frequency content of ground motions maximizing the mean earthquake energy input rate to structures, for details see e.g. (Takewaki, 2006). From a different perspective, Barbato et al. (2011) approximates the first passage problem by formulating exact closed form solutions for the spectral characteristics of random processes. Macke et al. (2002) presents an importance sampling technique for randomly excited dynamical systems.

The author of this paper attempt to, unlike the above, maintain the up-to-date most conceptually correct fully probabilistic concept (Ang and Tang, 2007) while reducing the number of required analyses by means of the proposed identification framework. It is based on a non-traditional assumption that there exists a

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

finite set of rules capable of classifying synthetic samples of stochastic processes according to their importance as a critical input for given mechanical model. Whether such set of rules could be formulated for arbitrary system remains an open problem for further research.

2. Development

The identification strategy development follows a transparent image processing paradigm completely independent of state-of-the-art structural dynamics, thus representing a non-traditional option in the field. Reason behind such premise is experimental, aiming at delivering simple and wide-purpose method. The goal can be formulated as follows: find the critical realization ($S_{T,Crit}$) of a stochastic process (S) from a target sample set S_T under defined critical response (C_r) criteria.

Proposed STS strategy steps:

1) Construct a training sample set S_t of size $S_t << S_T$.
2) Solve the mechanical model (i.e. carry out a structural dynamic analysis): S_t -> C_r, usually extremely computationally expensive, therefore the size of S_t should be as small as possible.
3) Select a proper graphical representation G of S_t (in time domain), which should serve for automatic feature extraction in the next step. There are two general options maintaining the physicality of S_t -> G S_t, transformation of S_t into evolutionary spectra (Priestley, 1965) or wavelet-vector coefficients based scalogram (Wolfram, 2011), both as 2D graphical arrays. The computational complexity of this task should be minimized, therefore small resolution is desired.
4) Find a finite set of rules R such that consistently maps R(G S_t) -> C_r, Narrow the search domain by ignoring pixels with constant or random-behaviour. Include pixels into R for which the difference of state values between upper and lower 5^{th} percentile of the ranked G S_t : C_r is maximized.
5) Obtain $S_{T,Crit}$ by applying R -> S_T.

In the broader context one should use the STS strategy to limit the number of necessary executions of numerical analysis of the mechanical model. It is assumed that mechanisms behind rules extracted from reasonably small samples are applicable to arbitrarily larger sample. Clearly, whenever using a black-box type of approach, there is a risk of extracting mechanisms that apply only to the training sample if its sample size is too small or in cases of "statistical bad luck". The determination of minimal size of a training set should be based on a requirement for STS's predictive confidence.

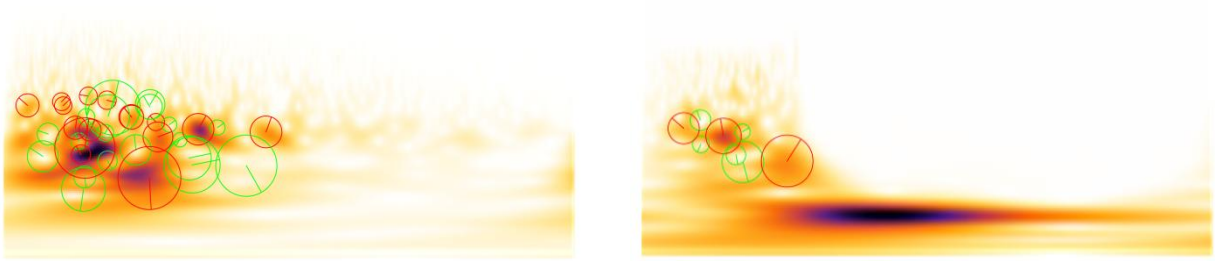


Figure 1. Graphical representation (G) of L1 (left) and L2 (right) in a form of Wavelet Scalogram and visualized detected keypoints (R) using their scale (radius of the circle), orientation and contrast sign (colour).

As stated before, the proposed STS strategy aims at general and automated feature extraction. It should be noted here however that rare instances where experienced when visual comparison of the ranked scalograms G itself enabled for formulation of identification rule R by comparing the number of regions with steep contrast gradient, i.e. image keypoints. For such feature a number of standardized algorithms exists, e.g. implemented SURF (Herbert et al 2008), numerically robust against translation, rotation and scale changes. Such approach can be interpreted as assessment of localized of energy in the time domain and proved to be consistent for configurations of SDOF oscillators loaded by stationary or amplitude modulated processes. In such instances a low number of detected keypoints indicates a critical process, i.e. G has minimal scatter of excitation energy, for example see Fig. 2.

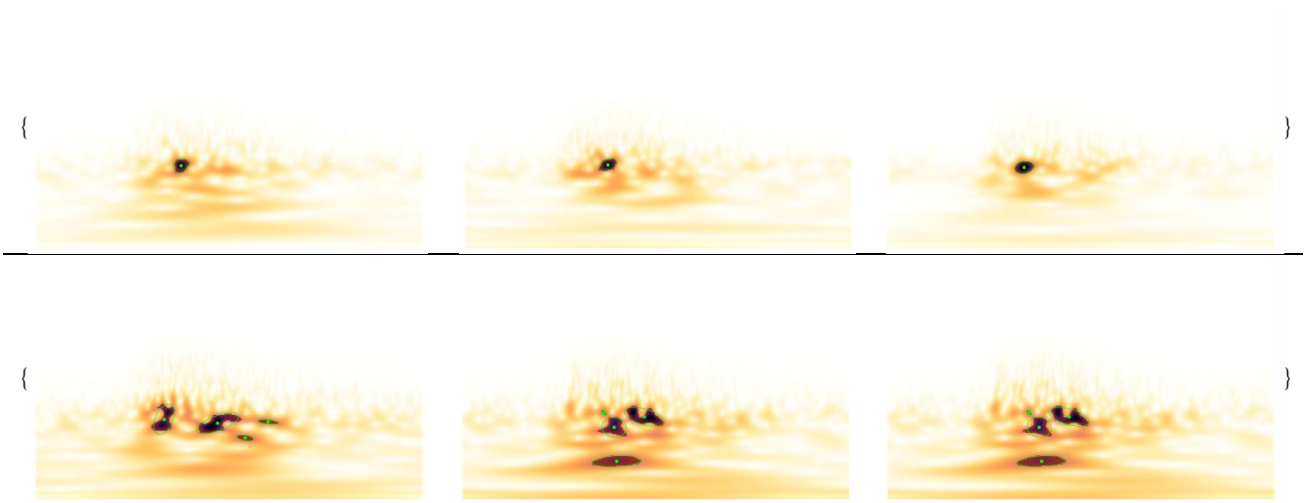


Figure 2. Number of fitted oriented ellipses (based on SURF) as a performance indicator, upper row: 3 ranked maximum and (lower row) 3 ranked minimum response.

The most general non-physical version of STS utilizes several pixels of small-resolution Wavelet Scalograms image for composition of R and R(G S_t) -> C_r mapping (step 4) based on a stochastic sensitivity analysis, returning pixels with state values that varies systematically according to the ranked small sample training sets, see Fig. 3.

The sensitive pixels are usually in clusters forming a line (indicating a dominant scale) and/or points (Fig. 3). Regardless of the attractiveness of emerging questions on physical connections of these clusters to the mechanical models (and dominant frequencies), such debates will not be detailed here due to the limited scope of the paper.

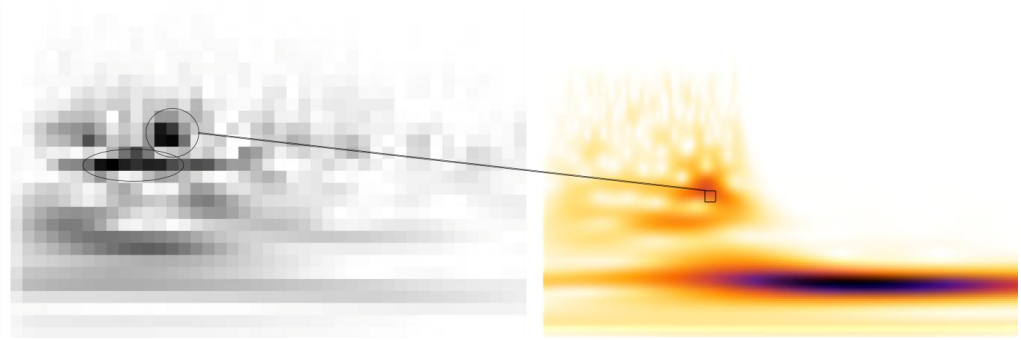


Figure 3. Left: Array of pixels (rescaled) according to their behaviour, darker the colour, more sensitive the pixel is to the ranked Gs, lighter colours indicates random or invariant behaviour; right: corresponding position of the sensitive pixel at the wavelet scalogram.

3. Acceleration and Structural Models

For validation of STS method four distinct combinations of two models (M1 and M2) and loadings (L1 and L2) are considered. The mechanical models represent a single degree of freedom (SDOF) damped linear oscillator (M1) and nonlinear seismically isolated SDOF on a friction pendulum system (M2) subjected to an earthquake loading $F(t) = -m\ a(t)$. Here $a(t)$ is the ground acceleration described as (L1) an amplitude modulated random process

$$a(t) = e(t) \cdot b(t) \tag{1}$$

where $e(t)$ is the amplitude modulating function given by

$$e(t) = 4[\exp(-0.25t) - \exp(-0.5t)] \qquad \text{for } t > 0 \tag{2}$$

and $b(t)$ denotes the stationary zero-mean Gaussian random process with power spectral density

$$S_{bb}(\omega) = S_0 \frac{4\zeta_g^2\omega_g^2\omega^2 + \omega_g^4}{\left(\omega_g^2 - \omega^2\right)^2 + 4\zeta_g^2\omega_g^2\omega^2} \tag{3}$$

and as (L2) an amplitude and frequency modulated random process whose objective is to reproduce the general frequency variation characteristics of the acceleration record from the 1964 Niigata earthquake (Shinozuka, 1991) described by the Bogdanhoff-Goldberg-Bernard (1961) envelope function

$$A(t) = a1\ t \exp(-a2\ t) \qquad \text{for } t > 0 \tag{4}$$

and Clough-Penzien acceleration spectrum with parameters S_0, ω_g and $\zeta_g = \zeta_f$ as functions of time:

$$S(\omega, t) = S_0(t) \left[\frac{1 + 4\zeta_g^2(t) \left[\frac{\omega}{\omega_g(t)}\right]^2}{\left\{1 - \left[\frac{\omega}{\omega_g(t)}\right]^2\right\}^2 + 4\zeta_g^2(t)\left[\frac{\omega}{\omega_g(t)}\right]^2} \right] \times \left[\frac{\left[\frac{\omega}{0.1\omega_g(t)}\right]^2}{\left\{1 - \left[\frac{\omega}{0.1\omega_g(t)}\right]^2\right\}^2 + 4\zeta_f^2(t)\left[\frac{\omega}{0.1\omega_g(t)}\right]^2} \right] \tag{5}$$

$$S_0(t) = \frac{\sigma^2}{\pi\,\omega_g(t)\left(2\zeta_g(t)+\frac{1}{2\zeta_g(t)}\right)} \tag{6}$$

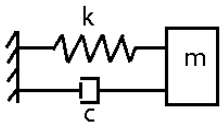
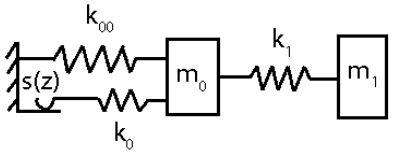
$$\omega_g(t) = \begin{cases} 15.56, & 0 \le t < 4.5 \\ 27.12\,(t-4.5)^3 - 40.68(t-4.5)^2 + 15.56, & 4.5 \le t < 5.5 \\ 2.0, & t \ge 5.5 \end{cases} \tag{7}$$

$$\zeta_g(t) = \begin{cases} 0.64, & 0 \le t < 4.5 \\ 1.25\,(t-4.5)^3 - 1.875(t-4.5)^2 + 0.64, & 4.5 \le t < 5.5 \\ 0.015, & t \ge 5.5 \end{cases} \tag{8}$$

where parameters $a_1 = 0.68$, $a_2 = 0.25$ and $\sigma = 100$.

The nonlinear mechanical model M2 represents a building (SDOF) combined with a friction based seismic isolation (friction pendulum system) device that introduces another mechanical degree of freedom as well as an internal variable representing plastic slip *z*. The implementation was adopted from (Bucher, 2010) and will not be detailed in this paper. The structural data for both M1 and M2 are provided in Table I, random realizations of L1 and L2 and response characteristics are depicted at Fig. 2.

Table I. Mechanical models and structural data

M1			M2		
	*m*	400 kg		m_0	6080 kg
	k	80000 KN/m		m_1	79770 kg
	c	120		k_0	42372 KN/m
				k_{00}	2629 KN/m
				k_1	62500 kN/m

Critical response criterion was formulated either as absolute values of top displacement of mass most distant from the application of seismic load or as given percentile of the mean-square values of the displacements. The former criterion led to better identification performance and therefore was adopted.

4. Identification Results

Development and testing of the STS on multiple scales and process-model scenarios showed that it is difficult, perhaps impossible, to formulate a general identification rule of physical interpretability, a fact that corresponds with the structural dynamics paradigm. One of such attempts led to the formulation of R incorporating the image keypoints as a way of quantifying the energy scatter in the loading process. Therefore, soft computing techniques were deployed in search for general black-box type method. The presented state of STS was tested on large number of clusters composed from a total of 4.2×10^4 realizations of Kt and Ni process in combination with various mechanical models. The stochastic simulations revealed the existence of R for every tested process-model scenario. Results presented in Fig. 5 were chosen to demonstrate the variability of performance and do not represent the best nor worst analyzed process-model instances.

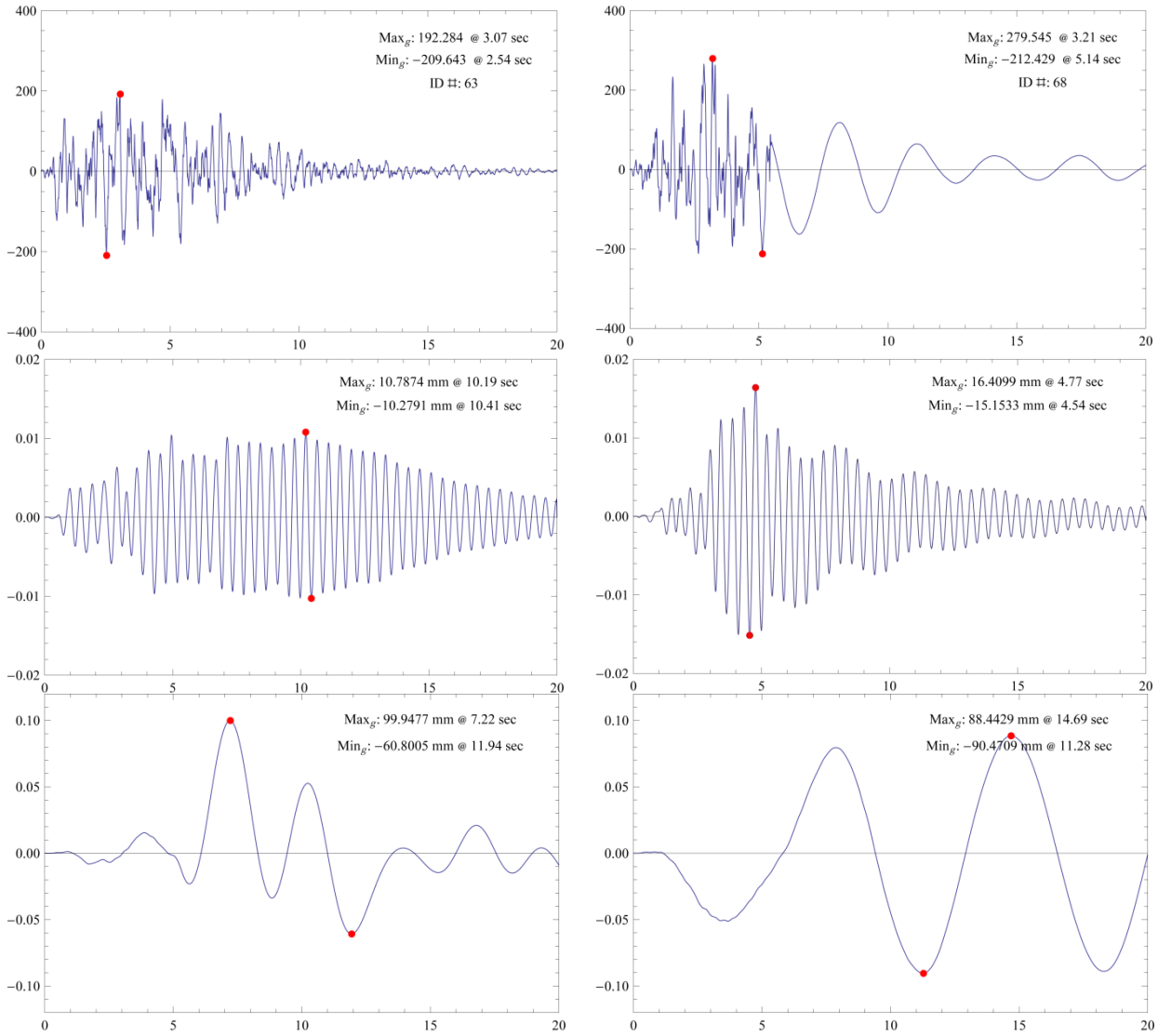


Figure 4. Example of realizations: left column top down: L1 process, M1 and M2 response to L1; right column top down: L2 process, SDOF response to L2, FPS response to L2 (note the abrupt change of frequency content at 5.5 sec); time at horizontal axes, acceleration/displacement on vertical axes.

The performance index was defined according to the following integral

$$P = \int_0^1 PDF_{min}\,(x).\,PDF_{max}(x)\,\mathrm{d}x \tag{9}$$

where $PDF_{min/max}$ states for the probability distribution function fitted to the ranked minimum/maximum set, growing isolation of these functions indicates better performance (see Fig. 6). The integration range corresponds to the admissible value of the *G* pixels.

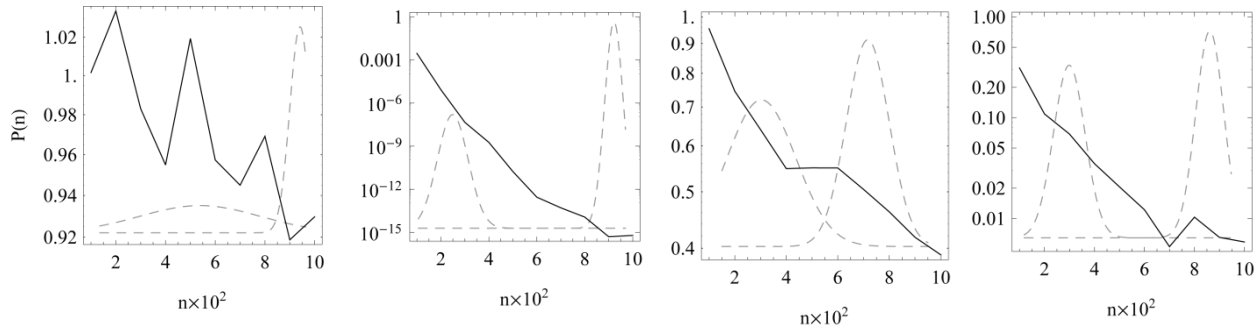


Figure 5. Performance index P as a function of sample size *n*, left to right: KtM1, KtM2, NiM1, NiM2; dashed line represents the Normal distribution PDFs fitted to 21 joined min and max sets for sample size $n = 2.1 \times 10^4$; Note the PDFs spacing effect on performance index.

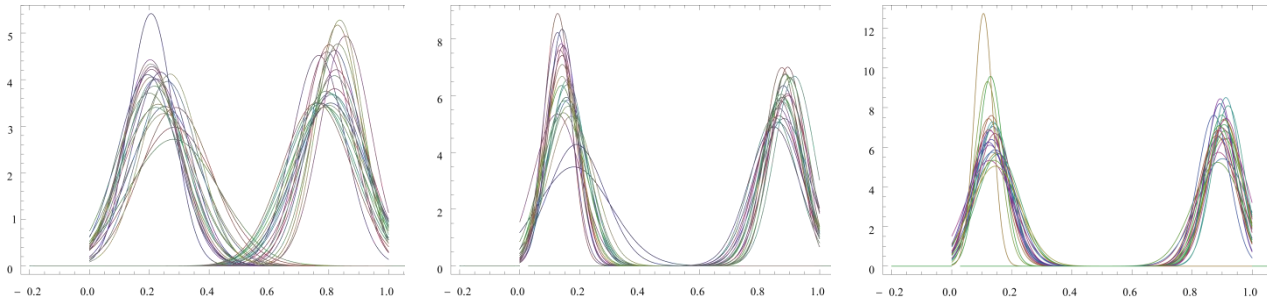


Figure 6. Ranked sets PDF {min, max} Isolation growth with increasing sample size $n = \{100,500,1000\}$; KtM2 realizations.

5. Importance Sampling

Following a successful formulation and validation of R according to the proposed STS, the importance sampling strategy is based on applying R to the full (original) set of realizations of stochastic processes and sorting the functional values of this product. Finally, the first *n* realizations corresponding to the ranked set are determined as critical input for numerical models. The determination of *n* depends on the required Importance Sampling confidence, e.g. in the presented case study (KtM2 model-process scenario) $n = 10$, i.e. 1% of the full set (1000), see figures 7 and 8.

The importance sampling test scenario, as described above, proved to be a consistent measure for reducing the 1000 sample set to a smaller set while maintaining the same critical response characteristics. The STS utilized 100 sample training set (10%) and the consequent importance sampling required additional 10 analyses (1%), therefore reducing the computational task by 89%. The additional 1% ensured that the important sample (most critical response) was captured by over 91% (within 21 test runs). Note the effect of emergent 2nd branch STS artefact from C_r distribution plot according to ranked R product. The inverse of the same plot (fig. 7) does not exhibit such effect, representing the amount of unaccounted information by STS. This is partly due to (i) incorporating only one sensitive point and (ii) ambiguous C_r -> R_p identifier based RGB channels. The performance of STS could be enhanced by including multiple

sensitive points with cross-correlations (i) and by modifying the R products by labels (e.g. random binary sequencing) or by enhancing the colour depth to ensure uniqueness (ii).

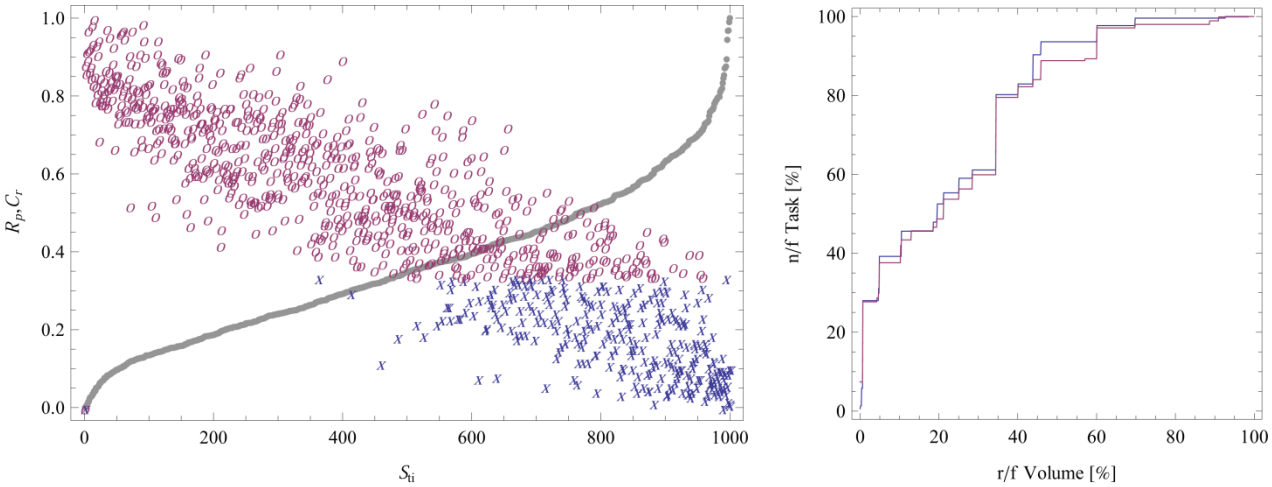


Figure 7. Left: Inverse property of ranked critical response (C_r points) and the R product (both Rescaled to (0,1) vs. sample size 1000, S_{ti}); here for illustration n = 10 and corresponding critical input markers "x", others "o". Right: Percentage of necessary/full computational task as a function of C_r ranked maxima (required/full volume) for 2 colour channels (R**GB**).

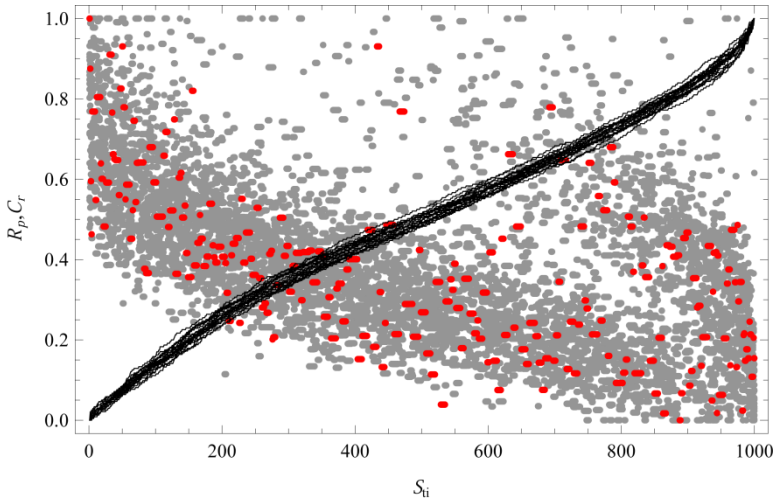


Figure 8. Rescaled distribution of C_r points (gray cloud) according to ranked R product (black line) from individual realizations S_{ti}; 21 repeated runs; particular realization in red points; note the emergent 2nd branch STS artefact.

The effect of unaccounted information does not only exhibit itself via the 2nd branch, but clearly also by the inability to always capture the single C_r maximum, as one might observe on the comparison plot at Fig. 9. Here the goal was to determine the probability of exceeding a critical displacement threshold u_{lim} at different sample size scales and compare it against reference pure Monte Carlo values. In terms of accuracy the maximum reached deviation between the MC reference and SST value was 15%, however, in terms of

computational efficiency the STS based importance sampling utilized only 4.6% of the MC computational cost, i.e. 0.6% for feature extraction and the remaining 4% for running the *n* realizations corresponding to the P_r ranked sets. This result indicates that there is significant potential in the application of the STS approach to the estimation of first passage probabilities. Nevertheless, the accuracy in its present form is not comparable to established simulation techniques. Due to its substantial computational advantage, however, the present approach will be suitable especially for reliability-based design optimization in which the reliability analysis has to be repeated frequently.

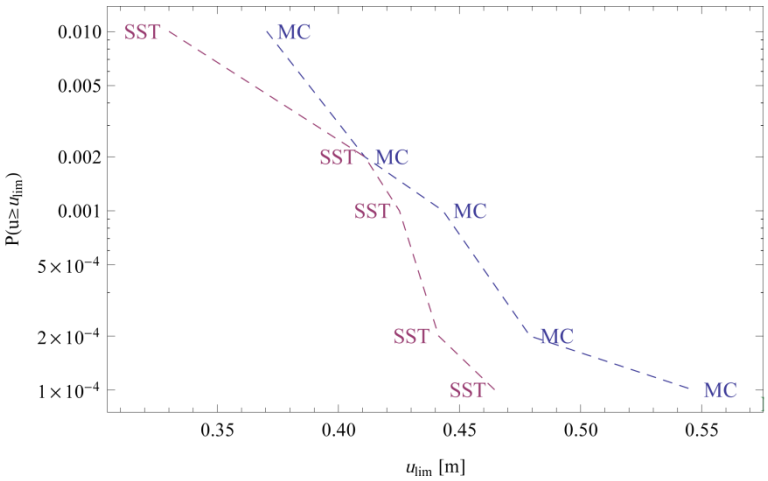


Figure 9. Determination of probability of exceeding a critical displacement threshold u_{lim} at different sample size scales: Comparison of pure Monte Carlo method (100% computational costs) and STS based importance sampling at 4.6% of computational cost.

6. Discussion and Conclusion

A novel **S**mall **T**raining **S**et strategy proposed by the author enables for identification of critical stochastic oscillatory processes with respect to given mechanical model. Such process is understood here as an environmental load acting on a structural system. From a design point of view, it is essential to understand what particular realization of such process has the critical impact on the structure. Traditionally, it is understood that each individual dynamical system has a very unique response to various stochastic loads. Therefore, for Monte-Carlo-based structural reliability considerations, all realizations of the stochastic load must be executed individually, making the task computationally unfeasible for realistic failure probabilities, since no sampling technique capable of reducing such task is available up to current date.

Motivated by the latter statement, an importance sampling strategy is formulated such that it reduces the size of the computational task without sacrificing any of the properties of fully probabilistic approach. As demonstrated on the numerical examples, the identification is feasible with varying performance according to the type of process-model scenario. As one may observe at Fig. 5, there is no relationship between the complexity of the process or model and the performance index.

Positively tested for both stationary and non-stationary processes, linear and non-linear mechanical models, an important implication is that the proposed STS strategy moves the fully probabilistic approach within the context of dynamical systems one step closer to the engineering practitioners, motivated by the ever-growing demand for performance-based design. Besides from the engineering community, STS may be a useful technique in the context of environmental sciences, such as water resources, solving analogous problems, e.g. realistic critical precipitation scenarios.

Further research will focus on possible extensions and improvements regarding the accuracy of the first passage probabilities as well as the treatment of more complex engineering models including structural dynamics and hydrology.

Acknowledgements

Author would like acknowledge financial support from the Austrian Science Funds (FWF) as part of the Vienna Doctoral Programme on Water Resource Systems (DK-plus W1219-N22) and the hospitality of the Department of Civil Engineering and Engineering Mechanics at Columbia University, NY, where part of the research was conducted, especially Prof. Deodatis for inspiring discussions.

References

Ang, A. H.-S. and Tang W.H. *Probability Concepts in Engineering Planning and design*, New York: John Wiley & Sons, Inc., 1990.

Bay, H., Ess, A. Tuytelaars, T., and Van Gool, L., "SURF: Speeded Up Robust Features", *Computer Vision and Image Understanding* (CVIU), Vol. 110, No. 3, 2008, 346-359.

Barbato, M. and Conte, J. P. Structural Reliability Applications of Nonstationary Spectral Characteristics, *Journal of Engineering Mechanics*, ASCE, May 2011, 371-382.

Bucher, Ch. Optimal probabilistic design of friction-based seismic isolation devices, *Safety, Reliability and Risk of Structures, Infrastructures and Engineering Systems* – Furuta, Frangopol & Shinozuka (eds), © 2010 Taylor & Francis Group, London, ISBN 978-0-415-47557-0.

Haldar, A. and Mahadevan, S. *Reliability assessment using finite element analysis*. New York: John Wiley & Sons, Inc. 2000.

Melchers, R.E. *Structural reliability analysis and prediction*. 2nd edn. Chichester: John Wiley & Sons, Inc. 2001.

Macke, M. and Bucher, Ch. Importance sampling for randomly excited dynamical systems, *Journal of Sound and Vibrations* 268, 2003, 269-290.

Moustafa, A. Damage-Based Design Earthquake Loads for Single-Degree-Of-Freedom Inelastic Structures, *Journal of Structural Engineering*, ASCE, March 2011, 456-467.

Novak, D., Strauss, A., Bergmeister, K., Hoffman, S. and Pukl, R. Advanced life-cycle analysis of existing concrete bridges, *Journal of Materials in Civil Engineering*, 20(1), 2008, 9-19.

Podrouzek, J. and Strauss, A. Probabilistic assessment of concrete structures in the context of model uncertainties. *Computer and Concrete* (In Print), 2011.

Priestley, M. B. Evolutionary Spectra and Non-Stationary Processes. *Journal of the Royal Statistical Society*. Series B (Methodological), Vol. 27, No. 2 (1965), 204-237.

Shinozuka, M. and Deodatis, G. Simulation of Stochastic Processes by Spectral Representation. *Appl. Mech. Rev*. 44, 191 (1991), DOI:10.1115/1.3119501.

Takewaki, I. Probabilistic Critical Excitation Method for Earthquake Energy Input Rate, *Journal of Engineering Mechanics*, ASCE, September 2006, 990-1000.

Wolfram Mathematica Documentation Center, ref/WaveletScalogram, 2011.

Simultaneous loads in structural design

T. Poutanen
Tampere University of Technology, tuomo.poutanen@tut.fi

Abstract: The current codes include three load combination methods: Permanent loads are always combined dependently. A permanent and a variable load are combined dependently or independently and two variable loads semi-dependently or dependently.

However, if the loads are simultaneous, i.e. the loads are active at the same time with one on the other, they are combined dependently. If the loads are not active at the same time, the distributions can be altered in a way the loads are simultaneous i.e. loads are always combined dependently in the structural design.

The dependent load combination results in higher safety factors, γ_G, γ_Q, γ_M, and combination factors, ψ_0, than the ones obtained semi-dependently and independently.

The dependent load combination is reliable, it is simple and it requires little calculation work.

Keywords: load combination, code, design

1. Introduction

The load combination is one of the key issues of the structural design and the design codes, (EN 1990, 2002; ISO 2394, 1996). A uniform theory of the load combination is missing. The dominant hypothesis is that the loads are combined independently if the loads are independent and dependently if the loads are dependent. However, the permanent loads are independent, but combined always in current codes dependently. The permanent load and the variable load are often considered independent and combined sometimes independently and sometimes dependently but these loads are dependent during the normal service time, 50 years, and must therefore be combined dependently. The one-year loads are independent but simultaneous and must be combined dependently i.e. by accumulation, too. The variable loads are combined semi-dependently but these loads should be combined dependently after the distributions are altered in a way the loads are simultaneous. This paper explains that loads are always combined dependently in the structural design which results in higher safety, γ_G, γ_Q, γ_M, and combination factors, ψ_0, than the ones in the current codes.

1.1. SYMBOLS

Symbols in this paper are mainly the same as used in the eurocode:

G Permanent load
Q Variable load
γ Safety factor
μ Mean
σ Deviation

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

β Reliability index, $P_f = \Phi(-\beta)$, Φ is standardised cumulative normal function
P_f Failure probability, in the eurocode $P_f = 1/15400$, $\beta = 3.826$
V Coefficient of variation $V = \sigma / \mu$
dp Design point value, the cumulative distribution value at the design point
d_s Duration of part time load, e.g. snow load
d_w Duration of full time load, e.g. wind or imposed load

1.2. TERMS

Some essential terms in this paper are:

Simultaneous – semi-simultaneous – non-simultaneous loads

Two loads are simultaneous during a reference time, if the loads are active during this time with one on the other. These loads may be dependent or independent but combined always dependently without combination or reduction factors which would result in a load vanish. However, the combination of two variable loads may include a combination factor induced by a distribution conversion.

The loads are non-simultaneous if the loads are not active at the reference time and semi-simultaneous if not simultaneous or non-simultaneous.

Permanent loads with each other and a permanent load and a variable load are always simultaneous.

Variable loads are normally semi-simultaneous but the distributions can be altered in a way these loads are simultaneous.

Dependent – semi-dependent – independent loads

Two loads are dependent if the loads at the same fractiles occur at the same time.

If one load of the two loads to be combined is a variable load, the loads may be equally dependent – semi-dependent – independent at the same time. The reason is that the load distribution of the variable load is defined to be the maximum load during one year. When time increases, several distributions and loads become available for the combination. At the infinite time there are infinite combination options and all dependent options (dependent – semi-dependent – independent) are equally possible.

Two individual simultaneous loads are independent but when time increases and/or number of loads increase, the loads become dependent.

When the distributions of the loads are fixed to the active time of both loads with one on the other and when time or number of loads increase while one load is selected the other load becomes automatically defined, too, as the loads occur at the same fractile at the same time. Due to this relation, the loads are dependent. A permanent and a variable load are dependent during one year only at low fractiles, 0.02 or less, i.e. the loads are semi-dependent and virtually independent, but these loads become more dependent when time increases, e.g. they are dependent up to fractile ca 0.98 in 50 years i.e. these loads are virtually fully dependent during the normal service time of structures, 50 years.

Two variable loads are similarly dependent, too, when the distributions are altered in a way both

loads are active at the same time with one on the other.

The dependence of the simultaneous loads may be explicit and physical i.e. the actual loads occur at the same time, e.g. the permanent and the variable load during 50 years.

The dependence of the simultaneous loads may also be implicit i.e. the individual loads are independent but a group of loads include at least one dependent load pair and the group acts dependently, e.g. imposed and permanent loads of a multi storey house.

The semi-dependent loads are normally dependent at low fractiles.

Two loads are independent if the loads do not occur at the same time at the same fractile.

Dependent – semi-dependent – independent load combination

Two loads are combined dependently by adding up the distributions by fractiles i.e. a load X with an item x_i in fractile i and a load Y with an item y_i in fractile i is combined dependently to obtain the combination load XY with an item xy_i in fractile i by adding up x_i and y_i, i.e. $xy_i = x_i + y_i$ (Poutanen, 2011). If the Monte Carlo simulation is used to combine the loads, in the dependent combination one seed number is used. If the convolution equation is used to combine the loads, the deviation of the combination load is fixed in a way the combination distribution crosses the crossing point of the partial distributions (Poutanen, 2011). In the dependent combination, the action of a new load in the combination is independent of other loads in the load combination.

The semi-dependent combination is an imprecise abstraction. Several semi-dependent combination methods exist, e.g. Turkstra's method where one load has the maximum deterministic value corresponding to the target reliability and the other load has a random value. The semi-dependent combination should lie between the dependent and independent combination. This is normally true at least at high fractiles.

In the independent combination, the loads are combined randomly e.g. by using the convolution equation or by using the Ferry Borges – Castanheta's method or by using the Monte Carlo simulation and two seed numbers. In the independent combination, the action of the new combination load depends on the earlier loads of the combination.

The current terms dependent and independent combination are misleading as in the independent combination the partial loads are independent but the combination load is dependent of the partial loads and in the dependent combination vice versa. Therefore it would be clearer to use terms random and accumulation summation.

The rule of the maximum load combination

A basic rule of the structural design is that the loads must be combined to obtain the maximum load. According to this rule, all loads should be combined dependently as the dependent combination results in the highest load. However, this rule is currently applied only to load combination alternatives with equal occurrence probability. Therefore this rule is not always applied as the independent or the semi-dependent combinations are considered more probable.

The permanent load and the variable load may be combined dependently, semi-dependently or independently. A new finding is that all these combination options are virtually equally possible during 50 years and therefore the maximum load combination rule must be applied i.e. these loads must be combined dependently. A further argument for the dependent combination is that even the

independent one year loads must be combined by accumulation. Variable loads are dependent, too if the distributions apply to the same time, and these loads are combined dependently.
Two permanent loads are independent but combined dependently.

1.3. ASSUMPTIONS, LIMITATIONS

The assumptions of the eurocode (EN 1990, 2002) are applied except for the load combination.

The variability and error induced by the code uncertainty, robustness, design, execution, use, degradation and wear out are excluded here.

2. Load combination in current codes

Loads are combined in the current codes in three ways:

- Permanent loads are always combined dependently.
- A permanent and a variable load are combined in the failure state sometimes dependently and sometimes independently but in the serviceability state always dependently.
- Two variable loads are always combined semi-dependently if these loads are the first and the second load in the load combination, but always dependently if the loads are third, fourth etc. load in the load combination.

The eurocode (EN 1990, 2002) includes three options to combine the permanent and the variable load, 6.10, 6.10a,b and 6.10a,mod. The first one is dependent and the others are independent. The Finnish eurocode is based on 6.10a,mod.

3. Load vanish

In the dependent load combination, no load vanishes. The loads are added up as such without any reductions or combination factors which would result in a load vanish.

In the semi-dependent and in the independent load combination a part of the load disappears in the combination. When two variable loads are combined semi-dependently a combination factor $\psi_0 \approx 0.6 \ldots 0.8$ is applied, which results in a load vanish of ca 0…20 %.

When a permanent and a variable load is combined independently, a load vanish of ca 0…10 % occurs, which is realized in the material safety factor γ_M.

We may deduce the independent load combinations wrong due to the load vanish: Assume a material (or a structure) has the survival probability S and the resistance 1 for the permanent load G alone and the variable load Q load alone. Now, if the material is loaded by $0.5G$ and $0.5Q$ and the loads are combined independently, the material has the resistance of ca 1.1[1] and if combined dependently, the resistance is 1. The independent combination is not viable. It is impossible that the effect of one load decreases if the other

[1] In the eurocode more precisely 1.0646, $V_{G,normal} = 0.09147$, $V_{Q,gumbel} = 0.4$, $\beta = 3.826$.

load is present as there is no link between the loads. The semi-dependent load combination is wrong for the same reason.

4. Simultaneous loads are combined dependently

Next examples are presented which show that simultaneous loads are combined dependently. The target survival probability is 0.98[2], $\gamma_G = 1$, $\gamma_Q = 1$ and the material is ideal with no variation $V_M = 0$, $\gamma_M = 1$.

4.1. COMBINATION OF PERMANENT LOADS

Permanent loads are simultaneous with each other. Permanent loads are independent, but these loads are combined dependently in all codes. This is correct as the independent and the semi-dependent load combinations are unrealistic: Assume a multi-storey house with the total design permanent load 1 on n floors. Now, if the number of floors n increases and the total permanent load remains and the total design load must remain, but the independent and the semi-dependent combination result in a decreased load.

4.2. COMBINATION OF PERMANENT AND VARIABLE LOADS

Structural design codes include a permanent load distribution G which defines the probability for the load not to reach the design point value. Accordingly, the variable load distribution Q defines an analogous probability for the variable load *during one year*. These distributions are independent and if combined, the combination distribution applies the random combination of the loads. The loads are combined in the structural design definitely, i.e. by accumulation and dependently.

A further reason for the dependent combination is that the permanent and the variable loads are dependent during the normal service time of structures: The probability of the variable load not to reach the design point value in one year is a low probability P_{f1} (0.02 in the eurocode). When time increases this probability increases, e.g. in t years it is $1 - (1 - P_{f1})^t$. We find that each fractile value of the Q distribution is associated to a fixed time. When the time is long, G and Q are fully dependent as all fractile values of G distribution occur at the same time as the corresponding fractile values of Q distribution. *Due to this relation, G and Q must be combined dependently, i.e. by the accumulation summation.*

4.3. COMBINATION OF VARIABLE LOADS

Two variable loads are combined almost analogously to the combination of the permanent and the variable load i.e. dependently when the distributions are first converted to the same time and simultaneous. In this combination, assumptions must be made about the basic characteristics of the variable loads. In my article (Poutanen, 2012) I assume that two kinds of variable loads exist: full time and part time loads, Figure 1. Each variable load has its characteristic duration d_w and d_s.

This variable load combination model is approximate: It is assumed that the load has a constant value through its duration i.e. the gradual increase and decrease of the load in the beginning and at the end is

[2] In this example the characteristic permanent load is the 0.98-value for the permanent load, in the eurocode it is the 0.5-value.

ignored. Further, appropriate load duration data, d_s and d_w, is lacking. However, this model allows us to make applicable conclusions by setting a range in the load durations.

When the wind load and the snow load is combined according to this model, the wind load distribution is converted to the time of the snow load and thereafter the loads are added up as such without any reductions, combination factors etc. In this combination, the snow load has no combination factor $\psi_0 = 1$, as the snow load distribution defined to one year is equal as the distribution defined to the winter. The wind load has a combination factor ψ_0 which is caused by the distribution conversion. Individual wind and snow loads are independent but combined dependently. These loads become dependent when time and number of loads increase. Ca 50 snow-wind-load pairs are virtually fully dependent during 50 years.

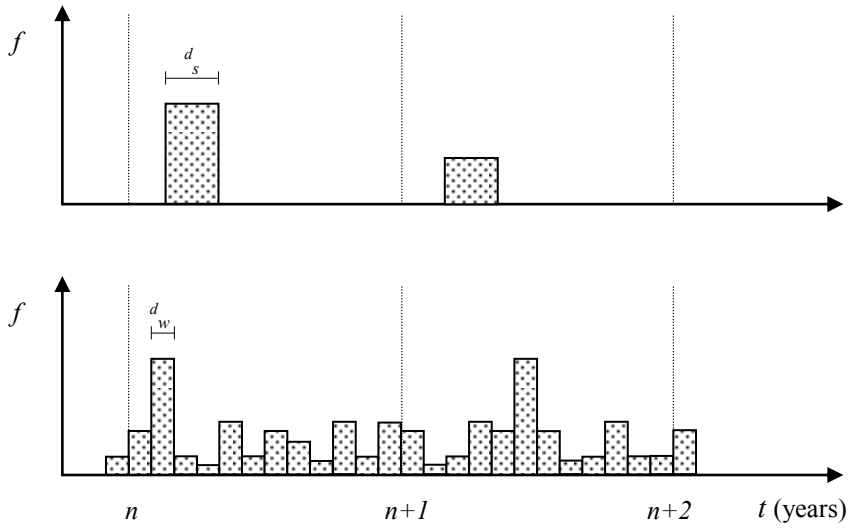


Figure 1. The variable load model, the part time load, upper Figure and full time load, lower Figure. The loads *f* are gumbel distributed and have a constant value within periods d_s and d_w.

We can change a distribution assigned to one year to *n* years by using an equation (EN 1990, 2002):

$$\Phi(\beta_n) = (\Phi(\beta_1))^n \tag{1}$$

where

β_1 reliability during one year, survival probability and the cumulative distribution value at the design point, in the eurocode $\beta_1 = 4.7$

β_n reliability at time *n*

n time (in years)

In the eurocode the variable load distribution $FG(x, \mu, \sigma)$ is gumbel and the target 50-year reliability value β_{50} is 3.826, where the parameters, μ, σ, can be solved $\mu = 0.2613$, $\sigma = 0.1045$. When the distribution is changed to time *n*, in years, the distribution must be multiplied by factor k_{Qn}, which can be solved from equation:

$$FG\left(1, \mu \cdot k_{Qn}, \sigma \cdot k_{Qn}\right) = \Phi\left(\beta_{50}\right)^{\frac{n}{50}} \quad (2)$$

Table 1 includes k_{Qn} – values for certain times n calculated from equation 2. The k_{Qn} – value is the load combination factor ψ_0, which is applied in the loads. Depending on the character of the loads the combination factor is applied in a fixed load, or in the lesser load.

Table 1. k_{Qn} –values calculated for certain times n.

Time n (load duration)	k_{Qn}
0.01 second	0.32
0.1 second	0.34
1 second	0.37
1 minute	0.41
10 minutes	0.45
1 hour	0.49
10 hours	0.53
1 day	0.56
1 week	0.61
2 weeks	0.63
1 month	0.66
2 months	0.68
3 months	0.70
6 months	0.73
1 year	1

According to this load combination model, the snow load never has a combination factor, $\psi_0 = 1$. A load or a sum of loads, combined to the snow load has a theoretical combination factor $\psi_0 = 0.63...0.73$ if the snow load lasts for 2 weeks...6 months. The code factor should be little higher to take into account the uncertainty of the model, ca $\psi_0 = 0.8$. If we assume that the live load and the imposed load lasts for 10 minutes...10 hours, these loads combined to each other have a theoretical combination factor $\psi_0 = 0.45...0.53$ and a code factor ca $\psi_0 = 0.6$ assigned in the lesser load.

When imposed loads are combined to each other a combination factor is not applied $\psi_0 = 1$ regardless of the duration of these loads as the imposed loads are proportions of the total imposed load in a house.

5. Conclusions

The simultaneous loads are always combined dependently in the structural design. If the loads are not simultaneous, the load distributions can be converted in a way the loads are simultaneous, i.e. the loads are always combined dependently.

Acknowledgements

When writing this paper, I have got valuable comments from Professor Keijo Ruohonen, Tampere University of Technology.

References

EN 1990:2002 Eurocode – Basis of structural design, European Committee for standardization, Bruxelles

Probabilistic Model Code, Part 1, Basis of Design, http://www.jcss.ethz.ch

Gulvanessian H., Holicky M., *Eurocodes: using reliability analysis to combine action effects*, Proceedings of the Institution of Civil Engineers, Structures and Buildings 158, pp 243-253, 2005

ISO 2394, *General principles on reliability for structures*, International Organization for Standardization, Geneva, Switzerland, 1996

Poutanen T., *Safety factors and design codes*, Joint IABSE – fib Conference, Dubrovnik, May 3-5, 2010

Poutanen T., *Calculation of partial safety factors*, Applications of Statistics and Probability in Civil Engineering – Faber, Köhler & Nishijima (eds), Taylor & Francis Group, London, 2011

Poutanen T., *Improved design code*, WCTE2012, Auckland, 2012

Poutanen T., *Combination factors in structural design*, article manuscript 2012

An Algorithm for Formal Safety Verification of Complex Heterogeneous Systems

Stefan Ratschan
Institute of Computer Science, Czech Academy of Science, stefan.ratschan@cs.cas.cz

Abstract. Modern technical systems are heterogeneous in the sense that they tightly integrate computational elements into physical surroundings. Computational elements usually require discrete, and physical systems continuous modeling. In this paper, we present an modeling formalism and safety verification algorithm for such heterogeneous systems.

1. Introduction

Modern technical systems more and more consist of a tight integration of computational devices into physical surroundings. For example, in modern cars, a large part of the development cost goes into software and digital electronics. Moreover, the complexity of such systems is growing rapidly. Hence it is of utmost importance to come up with formalisms for modeling, and algorithms for analyzing such systems.

The notion of a hybrid dynamical system is a current approach for modeling computation in physical surroundings (Lunze and Lamnabhi-Lagarrigue, 2009). Such systems integrate ordinary differential equations with finite state machines, based on a state space that is the Cartesian product of a subset of $\mathbb{R}^n$ and a set of finitely many states. Uncertainty is usually included by also allowing differential inequalities, or by allowing uncertain parameters in the differential equations. However, finite state machines do not suffice for modeling software of the complexity occurring in modern technical systems.

In our work, we will present an extension of the hybrid system model to systems that are parametric in k data types, with k an arbitrary, but fixed, positive integer. Those data types are generic in the sense that they can be chosen arbitrarily as long as they fulfill certain conditions that are met by the most widely-used data types such as integers, arrays, and lists. The state space of the new model is formed by the Cartesian product of a subset of $\mathbb{R}^n$ and the used data types. Again, the dynamics of the continuous part of the states space is given by ordinary differential equations (or inequalities).

Moreover, we provide an algorithm for the formal safety verification of such systems (i.e., the automatic verification that the system state always stays in a certain set of states considered to be safe) based on certain operations that the basic data types are required to provide. The algorithm is an extension of our earlier algorithm for hybrid systems verification (Dzetkulič and Ratschan, 2011).

This work was supported by MŠMT project number OC10048 and long-term financing of the Institute of Computer Science (RVO 67985807).

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

2. Problem Definition

Let us assume k (not necessarily distinct) data types $D_1, \ldots, D_k$. For each of those data types we assume certain functions and relations. For example for a certain $i \in \{1, \ldots, k\}$, D_i might be the set of integers with the functions addition and multiplication, and relations $=$ and $\leq$. Another examples is the set of lists of integers with the operations `nil` (for describing the empty list), `first` (for taking the first element of a list), `rest` (for taking the result of removing the first element of the list), and `cons` (for constructing a new list from an integer, and an old list). In the case of classical hybrid systems, there is just one data type, consisting of finitely many (but a potentially huge number) of so-called *modes*.

For modeling the physical surroundings we use the n-dimensional real space $\mathbb{R}^n$, with functions such as addition and multiplication, and relations such as equality $=$ and inequality $\leq$.

Now we assume a language $\mathcal{L}$ of constraints whose semantics is built on top of the semantics of the functions and relations of the data types $D_1, \ldots, D_k$, and of the real numbers. For example, having just one data type, the integers, $\mathcal{L}$ might consist of conjunctions of linear equalities and inequalities. Having both integers and lists over integers, we might allow expressions such as

$$x' = x + 1 \wedge l' = \texttt{cons}(x', l).$$

For us, the specific form of the language will not be important, but it will be essential to have certain constraint solving algorithms on them. We will introduce the specification of those algorithms in Section 3 and describe concrete possibilities for implementing them in Section 4.

Now, in order to describe the behavior of software within physical systems, we will introduce dynamical systems over those data types. The state space Φ will be given by $D_1 \times \ldots \times D_k \times C$, where $C \subseteq \mathbb{R}^n$.

DEFINITION 1. *A* system H *is a tuple* $(\texttt{Flow}, \texttt{Jump}, \texttt{Init}, \texttt{Unsafe})$*, where* $\texttt{Flow} \subseteq \Phi \times \mathbb{R}^n$*,* $\texttt{Jump} \subseteq \Phi \times \Phi$*,* $\texttt{Init} \subseteq \Phi$*, and* $\texttt{Unsafe} \subseteq \Phi$*.*

Informally speaking, the set `Init` specifies the initial states of a system and `Unsafe` the set of unsafe states that should not be reachable from an initial state. The relation `Flow` specifies the possible continuous behavior of the system by relating states with corresponding derivatives, and `Jump` specifies the possible discontinuous behavior by relating each state to a successor state.

We can describe those sets using the language $\mathcal{L}$. For example, using the constraint above to describe the set `Jump`—assuming that unprimed variables denote the current state and primed variables the successor state—will result in a system that creates a list of successive integers.

Another example, is a system with state space $\{\texttt{on}, \texttt{off}\} \times \mathbb{R}^2$, where the set `Flow` could be described by a constraint of the form

$$[\text{mode} = \texttt{on} \wedge \dot{x} = x + y \wedge \dot{y} = x - y] \vee [\text{mode} = \texttt{off} \wedge \dot{x} = x + y \wedge \dot{y} = x - 2y].$$

Here we view mode as a variable ranging over $\{\texttt{on}, \texttt{off}\}$, and $x, \dot{x}, y, \dot{y}$ as variables ranging over $\mathbb{R}$. Note that the dot in $\dot{x}$ is just used as a way of defining a new variable distinct from x—it is not yet connected to any form of derivation. It will be connected to derivation only now, in the following definition of system behavior:

DEFINITION 2. *For a certain discrete state $s \in D_1 \times \ldots \times D_k$, a* flow *of length $l \geq 0$ in s is a function $r : [0, l] \to \Phi$ such that the projection of r to its continuous part $C \subseteq \mathbb{R}^n$ is differentiable and for all $t \in [0, l]$, the projection of r to its discrete part $D_1 \times \ldots \times D_k$ is s. A* trajectory *of a system H is a sequence of flows $r_0, \ldots, r_p$ of lengths $l_0, \ldots, l_p$ such that for all $i \in \{0, \ldots, p\}$,*

1. *if $i > 0$ then $(r_{i-1}(l_{i-1}), r_i(0)) \in$ `Jump`, and*
2. *if $l_i > 0$ then $(r_i(t), \dot{r}_i(t)) \in$ `Flow`, for all $t \in [0, l_i]$, where $\dot{r}_i$ is the derivative of the projection of r_i to its continuous component.*

A (concrete) error trajectory *of a system H is a trajectory $r_0, \ldots, r_p$ of H such that $r_0(0) \in$ `Init` and $r_p(l) \in$ `Unsafe`, where l is the length of r_p. H is* safe *if it does not have an error trajectory.*

In the rest of the paper we will assume an arbitrary, but fixed system H. We will denote the set of its error trajectories by $\mathcal{E}$. In this paper, we study the problem of safety verification. This means that we want to check whether a given system has an error trajectory, that is, whether the set $\mathcal{E}$ is empty.

3. Safety Verification

One method for safety verification (that is used in so-called "bounded model checking"(Biere et al., 2003; Fränzle et al., 2007)) in the discrete time case is, to take the set of initial states, and compute the set of states reachable in one step, two steps, etc. and to check whether the result intersects the set of unsafe states. This has the drawback that it verifies safety of a given system over a bounded number of steps (at least for infinite state systems, and without additional techniques). Since the number of steps realistic systems can take is often huge, it is often more useful to design methods that check safety over unbounded time.

The straight-forward approach to verify safety over an unbounded number of steps is, to check whether the union of reachable states for subsequent time steps reaches a fix-point (in other words, further times steps do not result in further reachable states). For finite-state systems this is the main topic of the field of unbounded model checking (Clarke et al., 1999). For infinite state (but discrete time) systems in the form of computer programs, this is studied by abstract interpretation (Nielson et al., 1999). Such approaches make it necessary to first choose a representation for sets of system states for which a fix-point check can be easily done, and then to over-approximate the reachable states of the system using that representation.

This technique is also the basis of the first tools for safety verification of hybrid systems (Henzinger et al., 1997). However, for systems with non-trivial continuous evolution, this strategy has one severe drawback: For hybrid systems with non-trivial continuous dynamics even bounded time reach set computation necessarily involves over-approximation. A-priori it is not clear how precisely the reachable sets have to be computed to prove a given safety property. Hence, it may be advantageous to first compute approximate information using loose over-approximation, and then incrementally refine this.

Such techniques are popular in finite state and program verification under the name of "counter-example guided abstraction refinement" (CEGAR) (Clarke et al., 2003b) that has also been tried for hybrid systems (Alur et al., 2006; Clarke et al., 2003a). However, for systems with a partially continuous state space, this easily results in a behavior where the computed approximate information radically grows in size without representing enough information necessary for proving the safety property at hand.

In some earlier work (Dzetkulič and Ratschan, 2011), we presented a technique for avoiding this behavior for hybrid systems, and here we extend the technique to complex heterogeneous systems of the type described above.

Our approach is be based on an incremental refinement of a covering of the systems state space Φ by connected sets that we will call *regions*. We will form the regions in such a way that no two regions will overlap (i.e., regions are allowed to intersect, but only on their boundaries of the continuous part of the state space). The method is independent of the class of regions used. For example, in the special case of hybrid systems with a state space $M \times \mathbb{R}^n$, where M is finite, the regions can be formed by pairs consisting of an element of M and a Cartesian product of closed intervals (i.e., a *box*). But other classes of regions (e.g., based on polyhedra) are equally conceivable.

DEFINITION 3. *An* abstraction *is a graph whose vertices (which we will also call* abstract states*) are formed by regions that may be labeled with labels* `Init` *or* `Unsafe`*. We call the edges of an abstraction* abstract transitions*.*

This is the basic form of abstraction. However, an abstraction might be extended with much more information about the concrete system. For example, in our instantiation of this approach to the hybrid systems case (Dzetkulič and Ratschan, 2011), we store additional information on where trajectories might leave the regions.

For example, for a state space $\Phi = \mathbb{R}^2$, and the regions delimited by the black lines in Figure 1, a

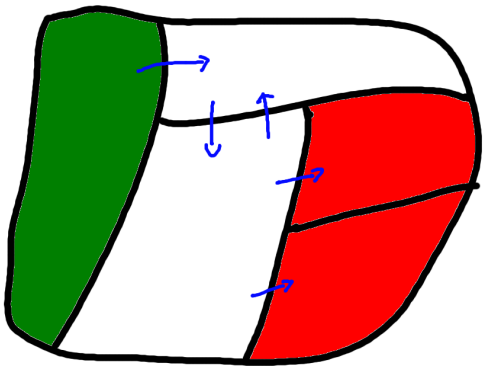


Figure 1. Abstraction Example

corresponding abstraction might be the graph consisting of those regions as nodes, the vertices given by the arrows (colored in blue), and with the region on the left-hand side (colored in green) marked as initial, and two regions on the right-hand side (colored in red) marked as unsafe. Such an abstraction represents the set of all trajectories that start in a region marked as initial (i.e., in a region colored in green in the figure), follows the edges of the graph (i.e.., the blue arrows in the figure), and ends in a region marked as unsafe (i.e., in a region colored in green in the figure). We will formalize this now.

We call a sequence of abstract states $a_1, \ldots, a_l$ an *abstract trajectory*. If all abstract states and all transitions between successive abstract states in an abstract trajectory belong to an abstraction $\mathcal{A}$, we call it an

$\mathcal{A}$-abstract trajectory and we denote it by $a_1 \to \ldots \to a_n$. An ($\mathcal{A}$-)abstract trajectory represents the set of concrete trajectories that begin in the region of a_1, move from one abstract state region to the next only if there is a corresponding concrete transition, and end in the region of a_n. We denote this set by $[\![a_1, \ldots, a_n]\!]$ for a given abstract trajectory or $[\![a_1 \to \ldots \to a_n]\!]$ for some $\mathcal{A}$-abstract trajectory.

This can be formalized as follows: We define a *splitting* of a flow l to be a sequence of flows $s_1, \ldots, s_r$ such that for all $i \in \{1, \ldots, r\}$, for all $t \in [0, \mathit{length}(s_i)]$, $s_i(t) = l(\sum_{j \in \{1,\ldots,i-1\}} t + \mathit{length}(s_j))$. A *trajectory splitting* is a concatenation of splittings of its individual contained flows. $[\![a_1, \ldots, a_n]\!]$ then is the set of all concrete trajectories $r_1, \ldots, r_p$ that have a trajectory splitting $q_1, \ldots q_n$, such that for all $i \in \{1, \ldots, n\}$, for all $t \in [0, \mathit{length}(q_i)]$, $q_i(t) \in a_i$.

An *$\mathcal{A}$-abstract error trajectory* is an $\mathcal{A}$-abstract trajectory $a_1 \to \ldots \to a_n$ such that in $\mathcal{A}$, a_1 is labelled initial, and a_n is labelled unsafe.

An abstraction $\mathcal{A}$ represents the set of all concrete trajectories $[\![a_1 \to \ldots \to a_n]\!]$ for abstract error trajectories $a_1 \to \ldots \to a_n$ in the abstraction $\mathcal{A}$. We denote this set by $[\![\mathcal{A}]\!]$.

The intuition is that, during abstraction refinement, the abstraction stays an over-approximation of the set of error trajectories $\mathcal{E}$ of a given system. We say that an abstraction $\mathcal{A}^*$ is a *refinement* of an abstraction $\mathcal{A}$ iff

- the abstraction $\mathcal{A}^*$ represents less trajectories than $\mathcal{A}$, that is, $[\![\mathcal{A}^*]\!] \subseteq [\![\mathcal{A}]\!]$, and
- the abstraction $\mathcal{A}^*$ does not lose error trajectories that are present in $\mathcal{A}$, that is $[\![\mathcal{A}^*]\!] \supseteq [\![\mathcal{A}]\!] \cap \mathcal{E}$.

Now we will come up with an algorithm that will incrementally improve an abstraction by refining it, *without* increasing the number of abstract states in the abstraction. Note that, in particular, $\mathcal{A}$ is a refinement of $\mathcal{A}$ itself, but in practice we will try to remove as many trajectories from the abstraction as possible.

Given abstract states a and a', we will assume a procedure $\mathit{Init}(a)$ that computes an over-approximation of the set of points in a that are initial (i.e., an element of `Init`), and a procedure $\mathit{Reach}(a, a')$ that computes an over-approximation of the set of points in a' reachable from a according to the system dynamics (here we do not assume any time bound, implementations of those procedures that compute reachability over bounded time would only require slight modifications of our algorithms). Our method is independent of the concrete technique used to compute those procedures. Still, in Section 4 we will discuss in detail how this can be implemented in practice. We assume that smaller inputs improve the precision of these operations, that is:

- $a_1 \subseteq a_2$ implies $\mathit{Init}(a_1) \subseteq \mathit{Init}(a_2)$
- $a_1 \subseteq a_2$ and $a_1' \subseteq a_2'$ implies $\mathit{Reach}(a_1, a_1') \subseteq \mathit{Reach}(a_2, a_2')$

Furthermore, we assume that these procedures exploit information about empty inputs, that is:

- $a = \emptyset$ implies $\mathit{Init}(a) = \emptyset$
- $a = \emptyset$ implies $\mathit{Reach}(a, a') = \emptyset$
- $a' = \emptyset$ implies $\mathit{Reach}(a, a') = \emptyset$

In the following, we require the existence of operations $\sqsubseteq$ and $\uplus$ on regions, with the following properties.

- $\sqsubseteq$ such that if $a^* \sqsubseteq a$, then for all $n \in \mathbb{N}$, for all $i \in \{1 \ldots n\}$ and for all regions $b_1 \ldots b_{i-1}, b_{i+1} \ldots b_n$ we have that $[\![b_1, \ldots, b_{i-1}, a^*, b_{i+1}, \ldots, b_n]\!] \subseteq [\![b_1, \ldots, b_{i-1}, a, b_{i+1}, \ldots, b_n]\!]$ i.e., less concrete trajectories follow a given abstract trajectory after replacing an abstract state by smaller one wrt. $\sqsubseteq$ operation.

– $\uplus$ s.t. for all regions $a_1, a_2, b : a_1 \sqsubseteq b \wedge a_2 \sqsubseteq b$ implies $a_1 \uplus_b a_2 \sqsubseteq b$, $a_1 \sqsubseteq a_1 \uplus_b a_2$ and $a_2 \sqsubseteq a_1 \uplus_b a_2$.

Since in our case abstract states represent sets, this can be ensured by the following:

– $\uplus$ s.t. for all $a_1, a_2 \subseteq b$: $a_1 \cup a_2 \ \subseteq \ a_1 \uplus_b a_2$ and $a_1 \uplus_b a_2 \ \subseteq \ b$

– $\sqsubseteq$ s.t. $a_1 \sqsubseteq a_2$ iff $a_1 \subseteq a_2$

This is our natural interpretation of $\uplus$ and $\sqsubseteq$. However, different choices are possible, as long as they fulfill the above properties: For certain representations of regions it might be convenient to use a weaker form of $\sqsubseteq$ efficiency reasons. Also, when encoding more information into abstract states (Dzetkulič and Ratschan, 2011), different interpretations of those symbols are often convenient.

In the instantiation of the method with boxes, $a_1 \uplus_b a_2$ is the smallest box that includes both argument boxes a_1 and a_2, but does not exceed b (i.e., box union intersected with bounding box), and $\sqsubseteq$ is the subset operation on boxes. Note that for $a_1, a_2 \subseteq \ b$ defining $a_1 \sqsubseteq a_2$ iff $a_1 \uplus_b a_2 = a_2$ fulfills the above property.

The following algorithm (which we will call *pruning algorithm*) computes a refinement of a given abstraction $\mathcal{A}$. The intuition is to remove parts from the regions forming the abstraction for which we can prove that they cannot lie on an error trajectory.

$\mathcal{A}^* \leftarrow$ copy of $\mathcal{A}$, all regions set to $\emptyset$, no initial labels, no edges
// from now on, for every abstract state a of $\mathcal{A}$,
// we denote by a^* the corresponding abstract state of $\mathcal{A}^*$
for all $a \in \mathcal{A}$, a is initial
    $a^* \leftarrow \mathit{Init}(a)$
    if $a^* \neq \emptyset$ **then**
        mark a^* as initial
while there is a pair of abstract states (a_1, a_2) in $\mathcal{A}$ with
    $a_1 \rightarrow a_2$, s.t. $\mathit{Reach}(a_1^*, a_2) \not\sqsubseteq a_2^*$ or ($a_1^* \not\rightarrow a_2^*$ and $\mathit{Reach}(a_1^*, a_2) \neq \emptyset$) **do**
        if $a_1^* \not\rightarrow a_2^*$ in $\mathcal{A}^*$ **then** introduce an edge $a_1^* \rightarrow a_2^*$ into $\mathcal{A}^*$
        if $\mathit{Reach}(a_1^*, a_2) \not\sqsubseteq a_2^*$ **then** $a_2^* \leftarrow (a_2^* \uplus_{a_2} \mathit{Reach}(a_1^*, a_2))$
return $\mathcal{A}^*$

Algorithms of such a type are known in the literature under them name "chaotic iteration" or "worklist algorithms" (Cousot and Cousot, 1977; Bourdoncle, 1993; Nielson et al., 1999; Apt, 1999; Apt, 2000).

Like similar algorithms in abstract interpretation, this algorithm computes unbounded reachability based on a fixpoint argument. However, unlike those algorithms, it exploits and refines the knowledge already available in the abstraction $\mathcal{A}$. In contrast to CEGAR approaches, the algorithm does *not* increase the size (i.e., the number of nodes) of the abstraction. Still it deduces some interesting information:

THEOREM 1. *The result of the pruning algorithm is a refinement of the input abstraction $\mathcal{A}$.*

Since the algorithm uses knowledge about the given system only through the operations Init and Reach, the correctness proof for the hybrid systems case (Dzetkulič and Ratschan, 2011) also applies here. For a similar approach in a completely discrete context see the notion of abstraction slicing (Brückner et al., 2008).

Note, that it is a-priori not clear, that the pruning algorithm terminates. However, termination can be ensured, for example, by using a representation for which, for given regions a and b, there is not infinite chain $a \sqsubseteq a_1 \sqsubseteq a_2 \sqsubseteq \ldots \sqsubseteq b$.

In the case where the discrete part of the state space is finite, and the regions describing the continuous part are formed by boxes, this holds if the endpoints of the corresponding intervals are formed by floating-point numbers, and the continuous state space is compact. Strategies for ensuring termination of such fixpoint computations are widely studied in the abstract interpretation community under the term "widening".

As already mentioned, the pruning algorithm tries to deduce information about a given system without increasing the size of the abstraction. In cases, where it can deduce no more information, we have to fall back to some increase of the size of the abstraction (cf. to a similar approach in constraint programming where one falls back to exponential-time splitting, when polynomial-time deduction does not succeed any more).

A simple method for doing this is a *Split* operation that chooses an abstract state and splits it into two, copying all the involved edges and introducing edges between the two new states. All the labels and abstract transitions to other abstract states are copied as well. Moreover, two new abstract transitions that connect the original abstract state with its copy are added. The region assigned to the abstract state is equally split among two abstract states. Such a refinement decreases the amount of over-approximation in subsequent calls to the pruning algorithm due to the properties of $Reach$ and $Init$. For example, Henzinger et. al. (Henzinger et al., 1998) use such a splitting step to reduce the over-approximation of the continuous system dynamics by differential inclusions of the form $\dot{x} \in A$, where A is a polyhedron. It is possible to use much more sophisticated splitting strategies, for example, a splitting step that removes as specific abstract error trajectory (i.e., one CEGAR step) (Alur et al., 2006; Clarke et al., 2003a) instead.

It is clear that the pruning algorithm can also be applied backward in time (i.e., removing parts of the abstraction not leading to an unsafe state) (Henzinger and Ho, 1995; Frehse et al., 2006). We will denote the resulting algorithm by $Prune^{-}(\mathcal{A})$.

Now we have to following overall algorithm for computing increasingly fine abstractions:

initialize $\mathcal{A}$ with an arbitrary abstraction such that
  $[\![A]\!]$ contains all error trajectories of the input system
while there is an $\mathcal{A}$-abstract error trajectory
  $\mathcal{A} \leftarrow Prune(\mathcal{A})$
  $\mathcal{A} \leftarrow Prune^{-}(\mathcal{A})$
  $\mathcal{A} \leftarrow Split(\mathcal{A})$
return "safe"

The most simple way to initialize the abstraction $\mathcal{A}$ in this algorithm is to use the trivial abstraction containing just one vertex for every mode marked as `Init` and `Unsafe`, containing a transition to all other vertices and itself, and a region containing the whole state space of the input system.

Since neither pruning nor splitting removes an error trajectory, the absence of an $\mathcal{A}$-abstract error trajectory at the termination of the while loop implies the absence of an error trajectory of the original system. Hence, in such a case, the algorithm correctly returns the information that the input system was safe. In cases where the input does have an error trajectory, this algorithm does not terminate. However, in such cases, the

algorithm maintains an abstraction that, at any time, can be used by other algorithms (Ratschan and Smaus, 2009) for searching for this error trajectory.

Note that forward pruning may enable further backward pruning and vice versa, hence the algorithm may be extended in such a way that forward and backward pruning are done in a loop until no further improvement occurs. If either forward or backward pruning is dropped from the algorithm, it will incrementally compute a tighter and tighter over-approximation of the (forward/backward) reach set.

The improvements to this algorithm introduced in an earlier paper (Dzetkulič and Ratschan, 2011) for the special case of classical hybrid systems can all easily be adapted to the more general case discussed in this paper.

4. Computation of Reachability Information

For applying the techniques in the previous sections to a concrete system one needs to

- choose a class of regions that will be used for representing subsets of the state space,
- instantiate the operations $\sqsubseteq$ and $\uplus$ with concrete algorithms, and
- provide algorithms for computing the reachability operations $Reach$ and $Init$.

Here we make the observation that for this, techniques from computational logic can be used. First we assume that (as in all examples above) the language $\mathcal{L}$ is based on first-order predicate logic. Moreover we assume that also the regions of the abstraction are formed by predicate logical formulas representing the set of all values that satisfy a given formula. For example, formulas of the form $\bigwedge_{i\in\{1,\ldots,n\}} \underline{a}_i \leq x_i \wedge x_i \leq \overline{a}_i$ represent hyper-rectangles (boxes). A concrete implementation may, of course, use a more optimized representation, but the usage of a formula representation in this section helps us to gain more insight into the nature of the problem.

Then, the region operations $\sqsubseteq$ and $\uplus$ can be implemented by (a sound approximation of) logical implication ($a \sqsubseteq b$ is such that $a \sqsubseteq b$ implies $a \Rightarrow b$) and (a conservative approximation of) disjunction ($a_1 \uplus_b a_2$ is such that $(a_1 \vee a_2) \wedge b$ implies $a_1 \uplus_b a_2$). For example, when using boxes, $a \uplus b$ might be the box hull (i.e., the smallest box containing both arguments) which clearly fulfills the above requirement.

Now we turn to the reachability operations $Reach$ and $Init$. We will write them as first-order predicate logical formulas. The operation $Init(a)$ must be such that the formula

$$a(x) \wedge \texttt{Init}(x)$$

implies $Init(a)(x)$.

For analyzing reachability we provide two separate logical formulas for reachability through jumps $Reach_J(a, a')$ and reachability through flows $Reach_F(a, a')$ which will result in reachability $Reach(a, a')$ being

$$Reach_J(a, a') \vee Reach_F(a, a').$$

The first part, $Reach_J(a, a')$ must be such that

$$\exists x \,.\, a(x) \wedge \texttt{Jump}(x, x') \wedge a'(x')$$

implies $Reach_J(a, a')(x')$, and $Reach_F(a, a')$ must be such that

$$\exists x \exists t \; . \; 0 \leq t \wedge t \leq c \wedge a(x) \wedge T_a(x, x', t) \wedge a'(x')$$

implies $Reach_F(a, a')(x')$, where $T_a(x, x', t)$ models the fact that there is a continuous flow from x to x' in a taking time t (we will later show how to model this as a logical formula), and c is an arbitrary positive real constant, or ∞, in which case the constraint $t \leq c$ can be dropped.

Now, one could just take the above formulas as the implementation of the operations themselves, in which case the implications above are implemented as equivalences. This is, in fact, the approach taken by bounded model checking of finite state systems, where there are techniques for representing and checking satisfiability of *huge* formulas in propositional logic using so-called SAT solvers (Biere et al., 2009). For discrete time systems with other data types/domains this can be extended to so-called satisfiability modulo theory (SMT) solvers (De Moura and Bjørner, 2011; Fränzle et al., 2007).

However, in the unbounded time case this has the disadvantage that when concatenating reachability over several steps (in our case, several applications of the $Reach(a, a')$ operator), more and more quantifiers accumulate, resulting in high-dimensional formulas, on which—in the unbounded time case—a fixpoint check (in our case based on $\sqsubseteq$) has to be done.

Another approach would be, to use the above formulas, but to eliminate the quantifiers in each application by quantifier elimination algorithms (Harrison, 2009), that is, algorithms that that compute equivalent, but quantifier-free formulas. In fact, this is precisely the approach taken in the finite state/propositional case, where practically efficient algorithms based on binary decision diagrams (BDDs) (Drechsler and Becker, 1998) exist. Also, early algorithms for hybrid systems verification took this approach (for very simple continuous dynamics). However, as soon as we leave the purely propositional case, even for quite simple individual theories, quantifier elimination is often not possible (e.g., non-linear integer arithmetic), or highly complex (e.g., linear integer arithmetic).

Hence, it makes sense, to use some form of over-approximation here, not implementing the above implications as equivalences any more. In program verification, the design of such regions with corresponding algorithms is the subject of *abstract domain design* (Filé et al., 1996), and—when using techniques from logic—*logical interpretation* (Tiwari and Gulwani, 2007; Gulwani and Tiwari, 2006). However, those abstract domains cannot directly be applied to systems with continuous time dynamics.

In our instantiation of the method for hybrid systems (Dzetkulič and Ratschan, 2011) we use interval constraint propagation (Benhamou and Granvilliers, 2006) (we also have a generalization of this technique available (Ratschan, 2006)). We also have investigated an alternative method based on an over-approximation of Fourier-Motzkin elimination (Dzetkulič and Ratschan, 2009).

We will now analyze how to handle continuous dynamics in this context. Above we used the expression $T_a(x, x', t)$ to model the fact that there is a continuous flow from x to x' in a taking time t. This could be directly written down in second order predicate logic (i.e., where variable and quantifier are allowed to range over functions and such functions can model system trajectories), however this would bring in additional difficulties for algorithmic analysis.

But, even if we need second order logic to model continuous reachability, we can at least *approximate* continuous reachability in first order logic. Here one get rid of second order quantifiers as follows: For each $v \in \{1, \ldots, n\}$, we can do a Taylor expansion at x of the projection of the trajectory to its v-th variable.

From this we get

$$\exists y, d \,.\, a(y) \wedge \texttt{Flow}^{k+1}(x, x^{(1)}) \wedge \ldots \wedge \texttt{Flow}(x, x^{(k)}) \wedge \texttt{Flow}^{k+1}(y, d) \wedge x' = x + \ldots + \frac{x^{(k)}}{k!} t^k + \frac{d|_v}{(k+1)!} t^{k+1},$$

where the notation $d|_v$ denote the projection of d to its v-th variable, and $\texttt{Flow}^i$ denotes a constraint that assigns to a point y its i-th derivative. If, as in the original Definition 1, only first-order derivatives are available, the formula can be applied only for $k = 0$ (i.e., the case corresponding to the mean-value theorem). The expression $T_a(x, x', t)$ can now be replaced by a conjunction of the above formula over all $v \in \{1, \ldots, n\}$. One can think of many variations of this approach, for example, by applying Taylor expansion backward in time.

5. Conclusion

In this paper we introduced a framework for formal safety verification of systems with both continuous and discrete dynamics, where the discrete part of the state space may include data structures such as lists and arrays. The framework includes an algorithm for safety verification of hybrid systems as an instantiation (Dzetkulič and Ratschan, 2011; Ratschan and She,). Computational experiments with that instantiation confirm the usefulness of the approach.

The remaining challenge is to instantiate the framework to cases with more interesting data structures, to design corresponding reachability analysis algorithms both for the discrete and continuous cases, and their combinations.

References

Alur, R., T. Dang, and F. Ivančić: 2006, 'Predicate abstraction for reachability analysis of hybrid systems'. *Trans. on Embedded Computing Sys.* **5**(1), 152–199.

Apt, K. R.: 1999, 'The Essence of Constraint Propagation'. *Theoretical Computer Science* **221**(1–2), 179–210.

Apt, K. R.: 2000, 'The Role of Commutativity in Constraint Propagation Algorithms'. *ACM Transactions on Programming Languages and Systems* **22**(6), 1002–1036.

Benhamou, F. and L. Granvilliers: 2006, 'Continuous and Interval Constraints'. In: F. Rossi, P. van Beek, and T. Walsh (eds.): *Handbook of Constraint Programming*. Amsterdam: Elsevier, Chapt. 16, pp. 571–603.

Biere, A., A. Cimatti, E. M. Clarke, O. Strichman, and Y. Zhu: 2003, 'Bounded Model Checking'. *Advances in Computers* **58**, 117 – 148.

Biere, A., M. J. H. Heule, H. van Maaren, and T. Walsh (eds.): 2009, *Handbook of Satisfiability*, Vol. 185 of *Frontiers in Artificial Intelligence and Applications*. IOS Press.

Bourdoncle, F.: 1993, 'Efficient chaotic iteration strategies with widenings'. In: D. Bjørner, M. Broy, and I. Pottosin (eds.): *Formal Methods in Programming and Their Applications*, Vol. 735 of *LNCS*. Springer, pp. 128–141.

Brückner, I., K. Dräger, B. Finkbeiner, and H. Wehrheim: 2008, 'Slicing Abstractions'. *Fundamenta Informaticae* **89**(4), 369–392.

Clarke, E., A. Fehnker, Z. Han, B. Krogh, J. Ouaknine, O. Stursberg, and M. Theobald: 2003a, 'Abstraction and Counterexample-Guided Refinement in Model Checking of Hybrid Systems'. *Int. J. of Foundations of Comp. Sc.* **14**(4), 583–604.

Clarke, E., O. Grumberg, S. Jha, Y. Lu, and H. Veith: 2003b, 'Counterexample-Guided Abstraction Refinement for Symbolic Model Checking'. *Journal of the ACM* **50**(5), 752–794.

Clarke, E. M., O. Grumberg, and D. A. Peled: 1999, *Model Checking*. MIT Press.

Cousot, P. and R. Cousot: 1977, 'Automatic synthesis of optimal invariant assertions: Mathematical foundations'. In: *Proceedings of the 1977 symposium on Artificial intelligence and programming languages*. pp. 1–12.

De Moura, L. and N. Bjørner: 2011, 'Satisfiability modulo theories: introduction and applications'. *Commun. ACM* **54**, 69–77.

Drechsler, R. and B. Becker: 1998, *Binary Decision Diagrams*. Springer.

Dzetkulič, T. and S. Ratschan: 2009, 'How to Capture Hybrid Systems Evolution Into Slices of Parallel Hyperplanes'. In: *ADHS'09: 3rd IFAC Conference on Analysis and Design of Hybrid Systems*. pp. 274–279.

Dzetkulič, T. and S. Ratschan: 2011, 'Incremental Computation of Succinct Abstractions For Hybrid Systems'. In: *FORMATS 2011*, Vol. 6919 of *LNCS*. pp. 271–285, Springer, Heidelberg (2011).

Filé, G., R. Giacobazzi, and F. Ranzato: 1996, 'A Unifying View of Abstract Domain Design'. *ACM Comput. Surv.* **28**, 333–336.

Fränzle, M., C. Herde, S. Ratschan, T. Schubert, and T. Teige: 2007, 'Efficient Solving of Large Non-linear Arithmetic Constraint Systems with Complex Boolean Structure'. *JSAT—Journal on Satisfiability, Boolean Modeling and Computation, Special Issue on SAT/CP Integration* **1**, 209–236.

Frehse, G., B. H. Krogh, and R. A. Rutenbar: 2006, 'Verifying Analog Oscillator Circuits Using Forward/Backward Abstraction Refinement'. In: *DATE 2006: Design, Automation and Test in Europe*.

Gulwani, S. and A. Tiwari: 2006, 'Combining abstract interpreters'. In: *Proceedings of the 2006 ACM SIGPLAN conference on Programming language design and implementation*. New York, NY, USA, pp. 376–386, ACM.

Harrison, J.: 2009, *Handbook of Practical Logic and Automated Reasoning*. Cambridge University Press.

Henzinger, T. A. and P.-H. Ho: 1995, 'A note on abstract interpretation strategies for hybrid automata'. In: *Hybrid Systems II*, Vol. 999 of *LNCS*. pp. 252–264.

Henzinger, T. A., P.-H. Ho, and H. Wong-Toi: 1997, 'HYTECH: a model checker for hybrid systems'. *International Journal on Software Tools for Technology Transfer (STTT)* **1**, 110–122.

Henzinger, T. A., P.-H. Ho, and H. Wong-Toi: 1998, 'Algorithmic Analysis of Nonlinear Hybrid Systems'. *IEEE Transactions on Automatic Control* **43**, 540–554.

Lunze, J. and F. Lamnabhi-Lagarrigue (eds.): 2009, *Handbook of Hybrid Systems Control*. Cambridge University Press.

Nielson, F., H. R. Nielson, and C. Hankin: 1999, *Principles of Program Analysis*. Springer.

Ratschan, S.: 2006, 'Efficient Solving of Quantified Inequality Constraints over the Real Numbers'. *ACM Transactions on Computational Logic* **7**(4), 723–748.

Ratschan, S. and Z. She, 'HSOLVER'. `http://hsolver.sourceforge.net`. Software package.

Ratschan, S. and J.-G. Smaus: 2009, 'Finding Errors of Hybrid Systems by Optimising an Abstraction-Based Quality Estimate'. In: C. Dubois (ed.): *Tests and Proofs*, Vol. 5668 of *LNCS*. pp. 153–168, Springer.

Tiwari, A. and S. Gulwani: 2007, 'Logical Interpretation: Static Program Analysis Using Theorem Proving'. In: F. Pfenning (ed.): *Automated Deduction CADE-21*, Vol. 4603 of *LNCS*. Springer, pp. 147–166.

Stochastic cracking of brittle matrix composites with heterogeneous reinforcement

Rostislav Rypl and Rostislav Chudoba
Institute of Structural Concrete, RWTH Aachen University, Germany

Miroslav Vořechovský
Institute of Structural Mechanics, Brno University of Technology, Czech Republic

Abstract. A model capable of capturing the effect of irregularly structured reinforcement in a brittle matrix is presented. It introduces a homogenization procedure for the state fields at the microscale in the vicinity of a crack bridge. In combination with a mesoscale model for matrix cracking the homogenized state fields are used for explicit calculation of the strain hardening-response of the composite. The model has been formulated for brittle matrix composites with reinforcement that exhibits random properties, e.g. due to random fiber orientation or because of an irregular penetration profile of the matrix into multifilament yarns. In the present paper we use multifilament yarns applied in textile reinforced concrete (TRC) to demonstrate the capabilities of the model. It is used for parametric studies to detect some qualitative and quantitative dependencies between the micromechanical material parameters and strain-hardening response of the composite.

Keywords: stochastic cracking, multifilament yarn, statistics, composite strength, micromechanics, filament

1. Introduction

Combining brittle matrix with fibrous reinforcement leads to quasi-ductile composite behavior with a high bearing capacity. When loaded in tension brittle matrix composites exhibit multiple cracks developing in the matrix perpendicularly to the loading direction (Li and Wu, 1992; Fantilli *et al.*, 2009). This process is accompanied with significant stress redistributions both between and within the constituents of the composite. The qualitative and quantitative characteristics of composites depend on the mechanical and geometrical properties of the components and their interface.

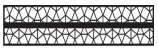
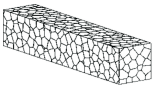
In order to study the response of the material structure subject to general loading conditions in 3D several models explicitly representing the geometrical distribution of fibers have been introduced using the finite element method (Radtke *et al.*, 2010) or lattice models (Bolander and Saito, 1997; Leite *et al.*, 2004). For purely tensile loading, models applying simplifying assumptions about the geometrical layout of the composite with respect to the tensile loding direction have been formulated with the goal to describe the tensile strain-hardening behavior. These models reflect the fragmentation process of composites starting from an elastic range, over a gradual evolution of matrix cracks with reinforcement strain localization up to a saturated crack density and/or ultimate failure of the weakest crack bridge.

The strain-hardening response of composites with elastic-brittle matrix and elastic reinforcement or similarly elastic-brittle reinforcement and elastic matrix can be described in closed form. In their classical work

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

Table I. State of the art overview

scheme	transverse	longitudinal	bond law	math	ref.
	assumed homogeneous	periodic CS, const. length	constant	1D, analytical	(Aveston *et al.*, 1971; Aveston and Kelly, 1973)
	assumed homogeneous	exact CS length distribution	constant	1D, analytical	(Curtin, 1992; Ahn and Curtin, 1996)
	averaged micromechanical	simplified	arbitrary analytical	1D, analytical	(Li *et al.*, 1991; Chudoba *et al.*, 2006a)
	partly homogeneous	explicit cracks	arbitrary	1D (2D), numerical	(Konrad and Chudoba, 2009; Azzam and Richter, 2011)
	explicit micromechanical	explicit cracks	arbitrary	2D, 3D, numerical	(Bolander and Saito, 1997; Radtke *et al.*, 2010)

Aveston, Cooper and Kelly (Aveston *et al.*, 1971) formulated the explicit relation between stress and strain of the composites with constant matrix strength under the assumption of aligned continuous reinforcement with an ideally plastic bond to the matrix. Later, they extended the model for elastic-plastic bond and included the effect of random fiber orientation by correspondingly reducing the number of bridging fibers compared to the aligned mode (Aveston and Kelly, 1973). An energy release rate approach taking into account the elastic stretching of matrix and fibers, matrix cracks propagation and fiber debonding has been presented by (Budiansky *et al.*, 1986).

(Cho *et al.*, 1992) studied ceramic composites with aligned fibers and provided analytical formulas for the composite stress-strain diagram assuming elastic material properties and a more sophisticated bond law. Using the stress criterion for debonding, the formulation of Cho et al. delivers a set of closed form solutions describing the composite behavior and crack spacing distributions for three different ratios of matrix and debonding strength. Moreover, a numerical study was performed for random matrix strength following the two parameter Weibull distribution furnishing stress-strain diagrams and crack spacing distribution. A remarkable method for arriving at the exact crack spacing distribution was developed by (Curtin, 1991) for a single filament embedded in a large failure strain matrix with randomized filament strength and constant frictional bond. The results show good agreement with an extensive Monte Carlo simulation performed earlier by (Netravali *et al.*, 1989; Henstenburg and Phoenix, 1989).

Curtin later applied his theory to composites with multiple matrix cracks (Curtin, 1992) and found a connection between the matrix flaw distribution and the crack spacing of a composite loaded in tension. Since the initial matrix flaws have to propagate through the cross-section while consuming energy a lower threshold is introduced for the stress-at-first-crack distribution. Having fitted two independent parameters from a composite tensile test, this model is able to predict the composite behavior and estimate the frictional bond. Another estimation of the bond stress based on experiments was performed earlier by (Marshall and Evans, 1985) offering three different methods for this purpose. An inherently statistical evaluation of the

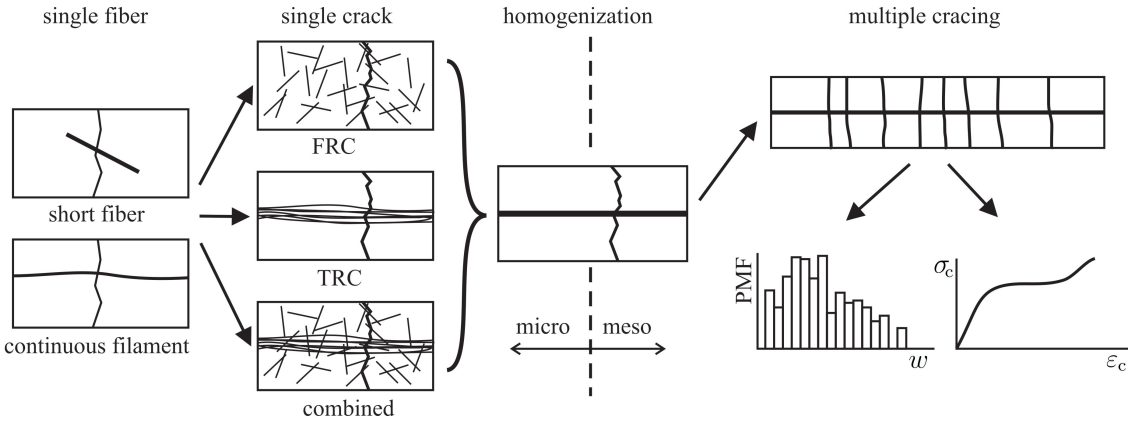


Figure 1. Scheme of the proposed model

stress-strain relationship and hysteretic behavior was performed by (Ahn and Curtin, 1996) and further simplified to closed forms by (Curtin *et al.*, 1998).

For composites with interface that can be described by one or a few bond parameters, e.g. for concrete reinforced with steel rebars, the aforementioned models can provide a realistic prediction. However, composites consisting of a large number of short or continuous fibers exhibit irregularities of the material structure (e.g. due random orientation of the fibers or due to an incomplete penetration of yarns by the matrix). The resulting variations in the stiffness and bond properties lead to a highly inhomogeneous microscopic bond stress fields. Such a field cannot be uniquely captured by a constant shear stress within the frictional bond model. This fact makes a more detailed resolution of the local stress and strain fields in the debonding zones inevitable. Models resolving the local fields in the vicinity in the crack bridge have been constructed using statistical averaging techniques (Li *et al.*, 1991; Chudoba *et al.*, 2006a; Kabele, 2003). With a higher computational effort, also finite element method has been used for local representation of the heterogeneous matrix-reinforcement bond structure (Konrad and Chudoba, 2009; Azzam and Richter, 2011; Nour *et al.*, 2011).

Table I summarizes the mentioned modeling approaches with a schematic picture of the assumed material representation. The models are classified according to the level of material resolution distinguished in transverse and longitudinal directions, kind of crack representation, applied bond law and dimensionality of the underlying mathematical formulation. In this paper a refined model for the simulation of strain-hardening response of a composite with heterogeneous structure of reinforcement is formulated. The effective stress –

Figure 2. TRC specimen reinforced with carbon fabrics after tensile test: localized failure crack (left), crack pattern (right)

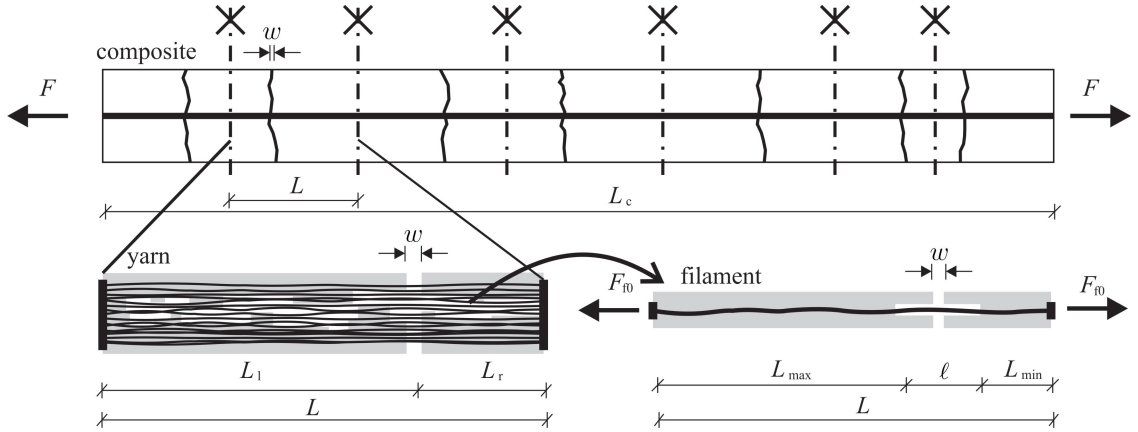


Figure 3. Application of the proposed model to textile reinforced concrete (from the top): homogenized composite model; statistical crack bridge model, microfilament double sided pullout model

strain relation is obtained using a multiscale homogenization procedure with separate integration loops at the micro- and mesoscale. At the microscale, the random structure of the reinforcement is reflected in the model of a representative crack bridge. At the mescoscale, the homogenization is performed over a representative series of emerging cracks sequentially introduced at positions where the matrix tensile stress reaches the level of the matrix strength.

The model components realizing the described homogenization procedure at the micro and mesoscales are depicted in Fig. 1. The hierarchical structure of the model opens up the possibility to include formulations of the crack bridge behavior for various types of reinforcement structure (e.g. short fibers, multifilament yarns or steel rebars) and their combinations. Let us also note, that the present modeling framework does not impose any limitations on the type of the bond law governing the interaction between fibers and matrix.

In order to make the explanation of the implementation of the model illustrative the formulation is provided for textile reinforced concrete consisting of a fine grained, brittle cementitious matrix reinforced by continuous multifilament yarns, such as AR-glass, carbon or aramid rovings (Fig. 2). The smallest scale considered deals with a single filament that bridges a matrix crack. The formulation of a single filament bridging a crack is provided in Sec. 2. A large number of such filaments form the reinforcing yarn which is assumed to be a multiple of the average filament response (Sec. 3). The evaluation of the homogenized strain field within a specimen with multiple cracks at a given level of stress is provided in Sec. 4. Results of computational examples showing some micromechanical dependencies on the global composite response are presented in Sec. 5 and concluding remarks summarize the capabilities and limitations of the model in Sec. 6.

2. Filament crack bridge model

At this level a single filament from a multifilament yarn embedded in matrix is observed and shall be represented by a parametric micromechanical model. For efficiency reasons, we assume symmetry at the half distance between adjacent cracks so that the filament is modeled only between two such symmetry

points crossing a single crack and its boundary conditions are fixed (Fig. 3). As the loading is increased, the crack width grows, both matrix and filament are stretched and debonding takes place at the filament-matrix interface. Filaments are assumed to have constant geometrical and physical properties over the length. However, the properties vary for individual filaments because they can, in general, have different physical, geometrical and bond properties. These assumptions together with linear elastic behavior of the matrix result in the following formulation for the filament response in terms of bridging force vs. crack width relationship:

$$F_{\mathrm{f0}} = F_{\mathrm{f0}}(w, A_{\mathrm{m}}, E_{\mathrm{m}}, L_{\mathrm{l}}, L_{\mathrm{r}}, \boldsymbol{\theta}_{\mathrm{f}}), \quad \boldsymbol{\theta}_{\mathrm{f}} = \{A_{\mathrm{f}}, E_{\mathrm{f}}, \tau, \ell, \theta, \xi, p\} \tag{1}$$

where the control variable w is the crack width, $A_{\mathrm{m/f}}$ and $E_{\mathrm{m/f}}$ the matrix/filament cross-sectional area and modulus of elasticity, respectively, $L_{\mathrm{l/r}}$ are distances from the crack to the boundaries at the left/right hand side, τ stands for the friction acting at the matrix-filament interface, ℓ denotes the bond free length of the filament, θ is the filament waviness in terms of additional strain (delayed activation), ξ is the filament breaking strain and p is the filament perimeter. Variables summarized in $\boldsymbol{\theta}_{\mathrm{f}}$ are the filament properties which are later in Sec. 3 eventually considered as random. For simplicity only the control variables w, x will be explicitly indicated further in the text.

Three stages (Fig. 4) of the filament crack bridge response have to be distinguished in the explicit notation:

(A) First, debonding (the bond law assumed here is a frictional resistance with constant magnitude τ) takes place at both sides of the crack and propagates towards the (fixed) boundaries

$$F_{\mathrm{f0(A)}}(w) = \begin{cases} \frac{1}{2\eta^2}\left(\sqrt{c_{\mathrm{A}}^2 + 4w_\theta K_{\mathrm{f}}\eta^2 T} - c_{\mathrm{A}}\right), & w_\theta \geq 0 \\ 0, & w_\theta < 0 \end{cases} \tag{2}$$

$$c_{\mathrm{A}} = LT - \eta(L_{\min} + L_{\max})T, \tag{3}$$

ℓ_θ and w_θ include the effect of the filament waviness in the following way: $\ell_\theta = \ell(1+\theta); w_\theta = w - \theta\ell$. $T = \tau p$ denotes the shear flow per unit length of a filament with the perimeter p, $K_{\mathrm{f/m}}$ is the filament/matrix tensile stiffness defined as $A_{\mathrm{f/m}}E_{\mathrm{f/m}}$ and η stands for the matrix/composite stiffness ratio $K_{\mathrm{m}}/(K_{\mathrm{f}} + K_{\mathrm{m}})$. $L_{\mathrm{min/max}}$ is the shorter/longer bonded length at the left or right hand side from the crack and is defined as $\min/\max\{L_{\mathrm{l}} - \ell/2, L_{\mathrm{r}} - \ell/2\}$ and L is the total filament length in the crack bridge (see Fig. 3).

(B) As soon as the debonding reaches the closer boundary, i.e. the bond is activated along the whole primarily bonded length $L_{\min}$ (Fig. 3) the model formally changes from a crack bridge to a pullout with free fiber length ℓ_{e} equal to $2L_{\min} + \ell$, bonded length L_{b} defined as $L_{\max} - L_{\min}$ and a force offset P_{A} accumulated in stage A due to the frictional bond along the debonded interface:

$$F_{\mathrm{f0(B)}}(w) = \frac{1}{\eta^2}\left(\sqrt{c_{\mathrm{B}}^2 + 2(w_\theta - w_{\theta,\mathrm{A}})K_{\mathrm{f}}\eta^2 T} - c_{\mathrm{B}}\right) + P_{\mathrm{A}} \tag{4}$$

$$c_{\mathrm{B}} = LT - \eta(L_{\max} - L_{\min})T \tag{5}$$

with $w_{\theta,\mathrm{A}}/P_{\mathrm{A}}$ the crack width/force at the transition between stage A and B.

(C) After the filament has been fully debonded along the whole length L, the model responses linearly-elastic to further loading with tensile stiffness K_{f}:

$$F_{\mathrm{f0(C)}}(w) = \frac{K_{\mathrm{f}}(w_\theta - w_{\theta,\mathrm{B}})}{L} + P_{\mathrm{B}}. \tag{6}$$

where $w_{\theta,\mathrm{B}}/P_{\mathrm{B}}$ in analogy to Eq. (4) stand for the crack width/force at the transition between stage B and C.

Putting Eqns. (2, 4, 6) together yields the formula for the bridging force F_{f0}:

$$F_{\mathrm{f0}}(w) = \begin{cases} P_{\mathrm{f0(A)}} & : 0 \leq F_{\mathrm{f0}} < P_{\mathrm{A}} \\ P_{\mathrm{f0(B)}} & : P_{\mathrm{A}} \leq F_{\mathrm{f0}} < P_{\mathrm{B}} \\ P_{\mathrm{f0(C)}} & : P_{\mathrm{B}} \leq F_{\mathrm{f0}} \end{cases} \tag{7}$$

Filament can break anytime during the loading which causes an immediate drop of the bridging force to zero. The remaining force carried by a broken filament being pulled out of the matrix is assumed to have minor contribution compared to intact filaments and is therefore neglected. This can be written using the Heaviside step function $H(x)$ defined as:

$$H(x) = \begin{cases} 0 & : x < 0 \\ 1 & : x > 0 \end{cases} \tag{8}$$

resulting in:

$$F_{\mathrm{f0}}(w) = F_{\mathrm{f0}} \cdot H(F_{\mathrm{f0}} - A_{\mathrm{f}} E_{\mathrm{f}} \xi) \tag{9}$$

Eqn. (7) delivers a base for the evaluation of strains in the filament $\varepsilon_{\mathrm{f}}(x)$ and matrix along the longitudinal axis x (Fig. 4). Highest values of filament strain occur at the crack position and with growing distance from the crack linearly descend with slope equal to the shear flow per length value T provided that the filament has a bond to the matrix. If there is a part of the filament with no contact to the matrix, the strain is constant along the region (see lower diagrams in Fig. 4). However, there is a lower bound $\varepsilon_{\mathrm{ff}}$ for the filament strain which equals the far field strain of the compact composite where no debonding takes place:

$$\varepsilon_{\mathrm{ff}}(w) = \frac{F_{\mathrm{f0}}}{K_{\mathrm{m}} + K_{\mathrm{f}}}. \tag{10}$$

The strain along a filament can be expressed as:

$$\varepsilon_{\mathrm{f}}(w, x) = \begin{cases} F_{\mathrm{f0}}/K_{\mathrm{f}} & : \text{free length } \ell_{\mathrm{e}} \\ \left[F_{\mathrm{f0}} - T\left(|x| - \ell_\theta/2\right)\right]/K_{\mathrm{f}} & : \text{debonded part } a \\ \varepsilon_{\mathrm{ff}} & : \text{bonded part } L_{\mathrm{PO}} - a \end{cases} \tag{11}$$

where the variable x is the position at the longitudinal axis with origin at the crack.

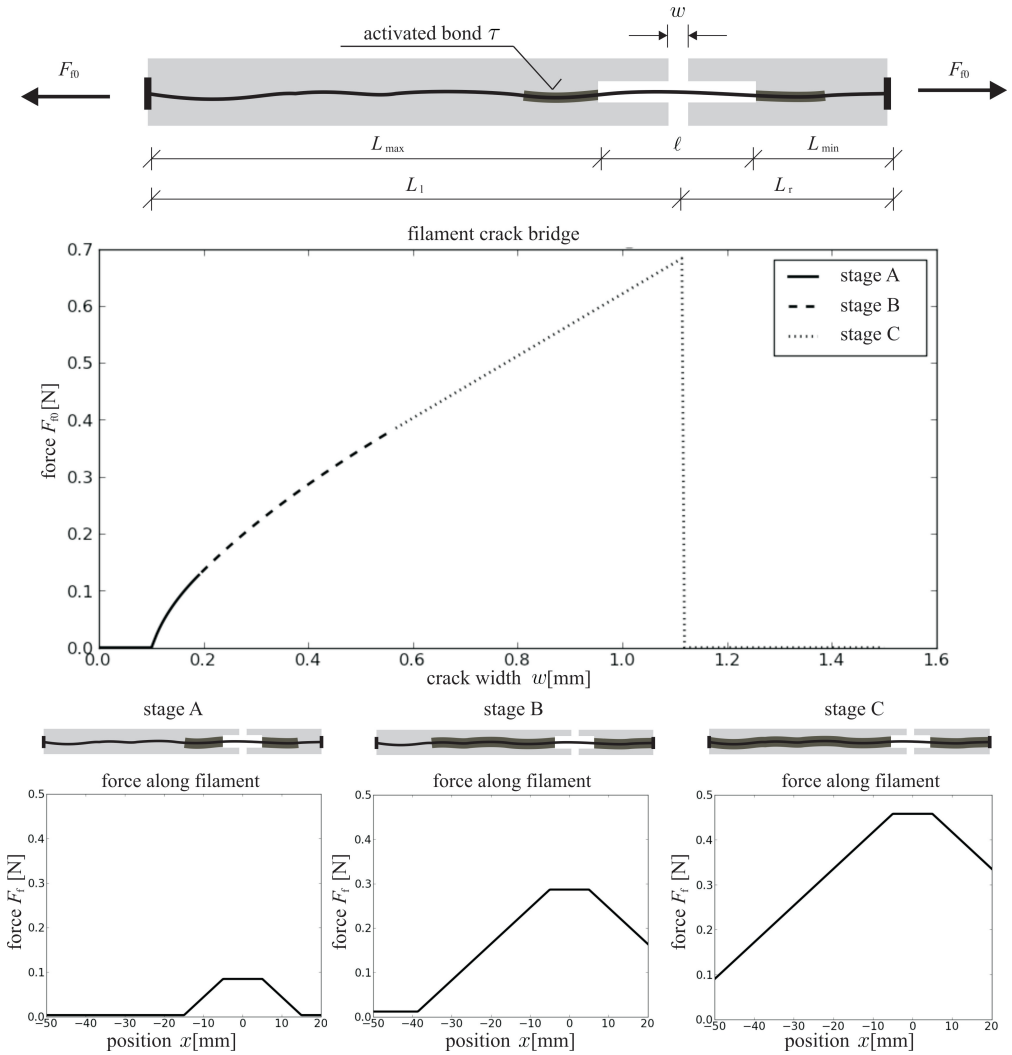


Figure 4. Filament crack bridge - force vs crack width with 3 distinguished phases (upper diagram); force in filament along the longitudinal axis for debonding stages A, B and C (lower diagrams). Parameters: $A_m = 29.4 \cdot 10^{-3}$ [mm^2], $E_m = 30 \cdot 10^3$ [MPa], $L_l = 50$ [mm], $L_r = 20$ [mm], $A_f = 5.31 \cdot 10^{-4}$ [mm^2], $E_f = 72 \cdot 10^3$ [MPa], $\tau = 0.1$ [N/mm^2], $\ell = 10$ [mm], $\theta = 0.01$ [-], $\xi = 0.0179$ [-], $p = 85.0 \cdot 10^{-3}$ [mm]. The three profiles are depicted for crack widths $w = 0.15$, 0.4 and 0.7 mm (from left to right).

3. Statistical crack bridge model

Sec. 2 creates a basis for the yarn crack bridge model. Since yarns consist of several hundreds or thousands of filaments, it would be very inefficient to simulate every single filament and sum their contributions. Therefore, the yarn is assumed to be represented by the average filament multiplied by the total number of filaments:

$$F_{y0}(w) = N_f \cdot \mu_{f0} = N_f \cdot E\left[F_{f0}\right] \tag{12}$$

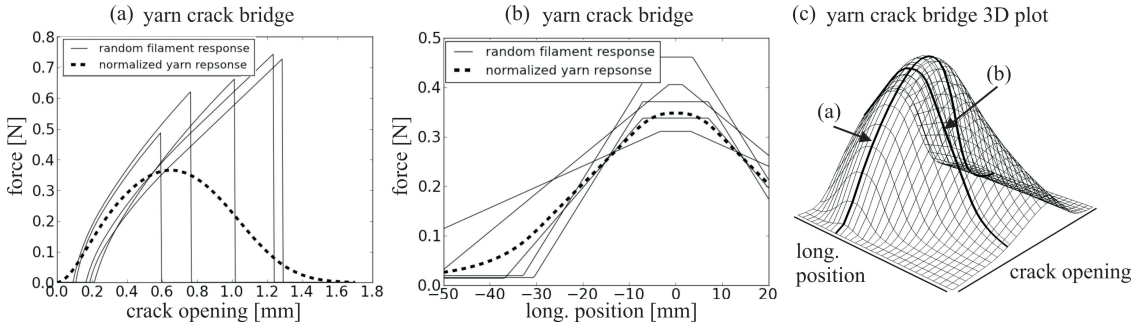


Figure 5. Yarn crack bridge - normalized yarn force $F_{\mathrm{y}0}/N_{\mathrm{f}}$ vs crack opening and 5 random filament realizations (a); normalized yarn force $F_{\mathrm{y}}/N_{\mathrm{f}}$ along longitudinal axis and 5 random filament realizations (b); 3D plot of the diagrams (c). Parameters for the yarn response: $A_{\mathrm{m}} = 50.0\,[\mathrm{mm}^2]$, $E_{\mathrm{m}} = 30 \cdot 10^3$ [MPa], $L_{\mathrm{l}} = 50$ [mm], $L_{\mathrm{r}} = 20$ [mm], $A_{\mathrm{f}} = 5.31 \cdot 10^{-4}\,[\mathrm{mm}^2]$, $E_{\mathrm{f}} = 72 \cdot 10^3$ [MPa], $\tau =$ uniform distribution $(\min = 0.05,\ \max = 0.20)$ [N/mm^2], $\ell =$ uniform distribution $(\min = 2.0,\ \max = 17.0)$ [mm], $\theta = 0.01$ [-], $\xi =$ Weibull distribution $(\mathrm{shape} = 5.0,\ \mathrm{scale} = 0.0179$ [-], $p = 85.0 \cdot 10^{-3}$ [mm], $N_{\mathrm{f}} = 1700$ [-]. The profiles in (b) are depicted for the crack width $w = 0.5$ mm.

for the force vs crack width and

$$F_{\mathrm{y}}(w, x) = N_{\mathrm{f}} \cdot \mu_{\mathrm{f}} = N_{\mathrm{f}} \cdot E\left[F_{\mathrm{f}}\right] \tag{13}$$

for the force along the yarn. The average or expected value of the filament crack bridge response $\mu_{\mathrm{f}0}$ and μ_{f} multiplied by the total number of filaments N_{f} can be alternatively written as follows:

$$F_{\mathrm{y}0}(w) = E\left[F_{\mathrm{f}0}(A_{\mathrm{f}} = A_{\mathrm{y}}, p = N_{\mathrm{f}} \cdot p)\right] \tag{14}$$

and

$$F_{\mathrm{y}}(w, x) = E\left[F_{\mathrm{f}}(A_{\mathrm{f}} = A_{\mathrm{y}}, p = N_{\mathrm{f}} \cdot p)\right] \tag{15}$$

respectively. This approach was used earlier e.g. for modeling fiber bundles with random fiber properties in (Phoenix, 1979; Phoenix and Taylor, 1973; Chudoba *et al.*, 2006b) and in the early works (Daniels, 1945; Coleman, 1958). Assumed that the filaments are statistically and mechanically independent it delivers an asymptotic result (for an infinite number of filaments) which is in this case justified by the large number of filaments forming a yarn. The average filament response (Fig. 5) is evaluated as stated in the cited literature in the following way:

$$\mu_{\mathrm{f}0}(w) = \int_{\boldsymbol{\theta}_{\mathrm{f}}} F_{\mathrm{f}0} \cdot f(\boldsymbol{\theta}_{\mathrm{f}})\,\mathrm{d}\boldsymbol{\theta}_{\mathrm{f}} \tag{16}$$

for the force resisting the crack opening and

$$\mu_{\mathrm{f}}(w, x) = \int_{\boldsymbol{\theta}_{\mathrm{f}}} F_{\mathrm{f}} \cdot f(\boldsymbol{\theta}_{\mathrm{f}})\,\mathrm{d}\boldsymbol{\theta}_{\mathrm{f}} \tag{17}$$

for the force along the composite longitudinal axis, with $f(\boldsymbol{\theta}_{\mathrm{f}})$ denoting the joint probability density function (PDF) of the random variables from the vector $\boldsymbol{\theta}_{\mathrm{f}}$. The average filament responses $\mu_{\mathrm{f}0}$ and μ_{f} are depicted in Fig. 5.

4. Composite model with stochastic cracking

With equations defined in Sec. 3 all tools needed for describing the stress state in the composite around a crack are provided. However, composite materials are designed to fail only after multiple transverse cracks have formed. To model this behavior, two additional variables have to be introduced: the load level F for new cracks to form and positions x_c of the cracks. F and x_c are revealed by satisfying the following condition:

$$\sigma_\mathrm{m}(F, x) \geq \sigma_\mathrm{mu}(x) \tag{18}$$

where σ_mu is a static autocorrelated random field (Vorechovsky, 2008). We define the local matrix strength as having the distribution of minimum extremes according to Weibull

$$W(\sigma_\mathrm{mu}) = 1 - \exp\left(-\left\langle\frac{\sigma_\mathrm{mu}}{\sigma_0}\right\rangle^m\right) \tag{19}$$

with m and σ_0 standing for the shape and scale parameter, respectively, and being cross correlated by the following definition

$$R(\mathrm{d}x, l_\rho) = \exp\left(-\frac{\mathrm{d}x}{l_\rho}\right)^2 \tag{20}$$

where l_ρ is the autocorrelation length and $\mathrm{d}x$ the distance between two points in the random field. The random field mimics the natural fluctuations of the local strength of reinforced matrix and automatically ensures the random distribution of the first cracks along the specimen (Fig. 6). A realization of this random field σ_mu is at a given load compared with the stress state in matrix σ_m and cracks are formed at the load level F and position x_c where the two functions first overlap if F is monotonically increased (Fig. 7).

4.1. Computation of σ_m

At small tensile loads at the beginning of the loading process, the strains in both matrix and reinforcement are assumed to be constant along the longitudinal axis x and described by:

$$\varepsilon_\mathrm{ff}(F) = \frac{F}{K_\mathrm{f} + K_\mathrm{m}} \tag{21}$$

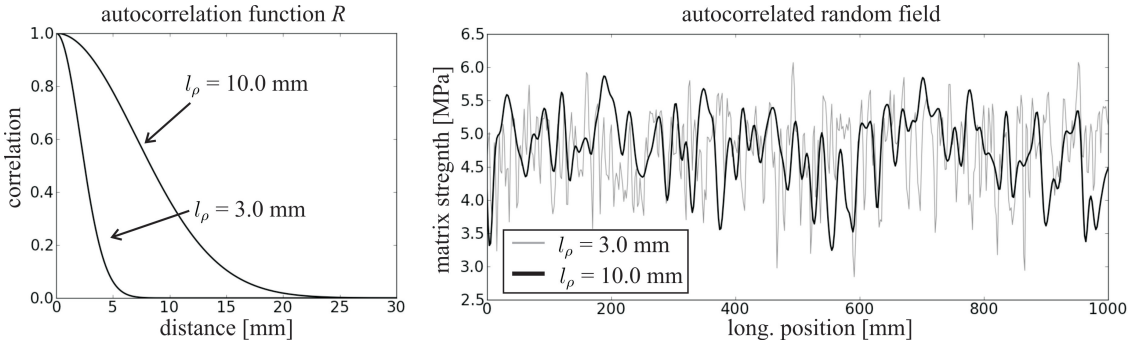


Figure 6. Autocorrelation function for $l_\rho = 3.0$ and 10.0 mm (left); corresponding realizations of a Weibull (shape = 10.0, scale = 5.0, location = 0.0) random field (right)

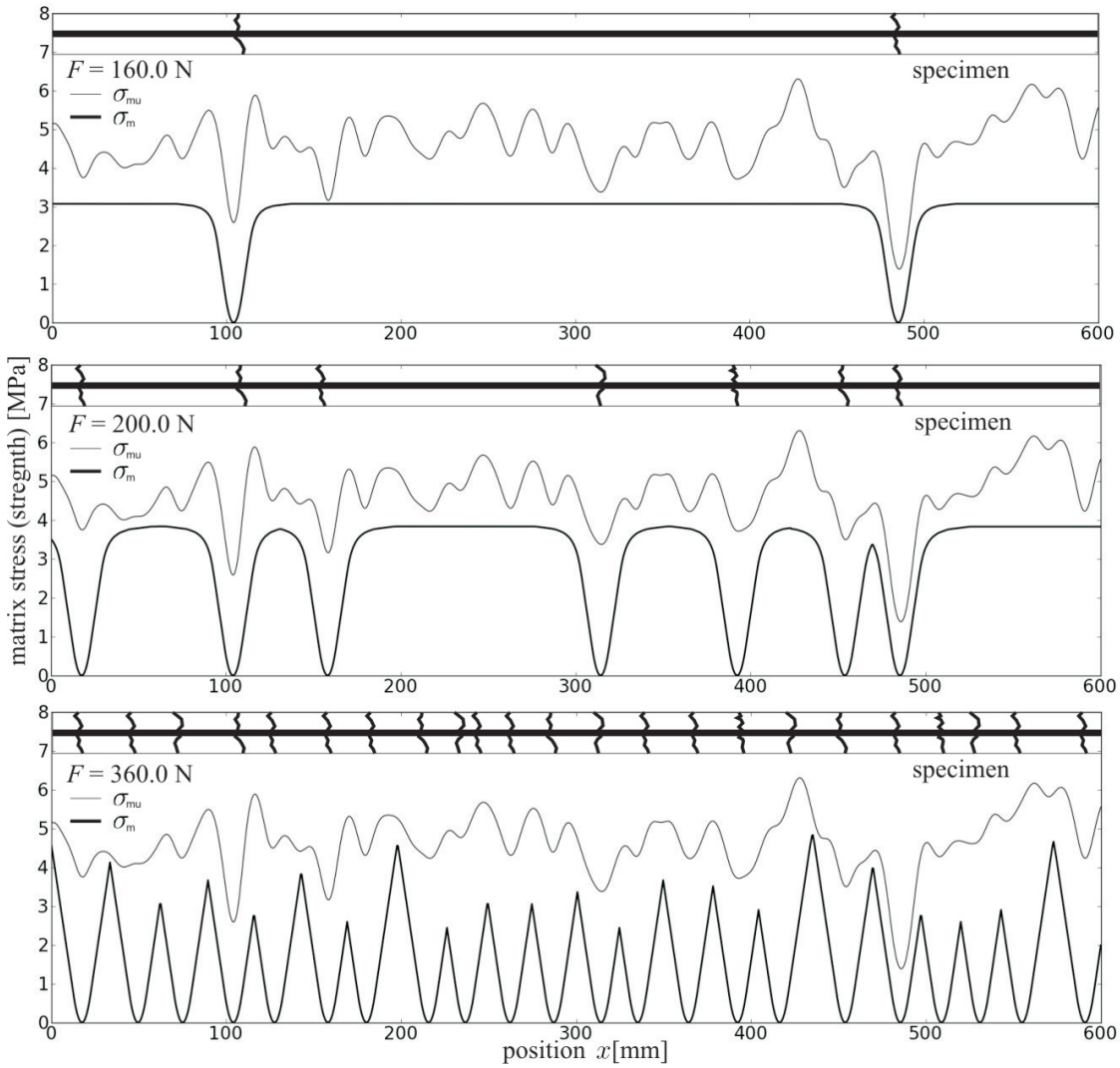


Figure 7. Matrix stress profiles at various loading stages

As soon as the tensile strength of the matrix is reached at some place, a matrix crack forms, the forces are redistributed and the reinforcement strain localizes at the position of the matrix crack (Fig. 7). Analogically, strain in the matrix drops to zero and is built up with growing distance from the crack. Cohesive forces between the newly created matrix surfaces are ignored here. The current matrix stress, given the yarn force by Eq. (15), is evaluated as:

$$\sigma_{\mathrm{m}}(F,x) = \frac{F - F_{\mathrm{y}}}{A_{\mathrm{m}}} \tag{22}$$

where the half distances to the neighboring cracks L_{l} and L_{r}, respectively, have to be taken into account and substituted into F_{y} in Eq. (15). Since F_{y} is controlled by crack opening, the value of crack opening

corresponding to the applied load F has to be evaluated first. This can be done by inverting Eq. (16) so that $\mu_{\mathrm{f}0}(w)$ becomes $w(\mu_{\mathrm{f}0})$ where $\mu_{\mathrm{f}0}$ is substituted by the applied load F. In this way, F_{y} in Eq. (15) is redefined as a function of F.

4.2. AVERAGE COMPOSITE STRAIN

The strain profile along the yarn has to be evaluated at every load step as it is qualitatively and quantitatively affected by both the load level and the crack positions. The overall (average) strain of the composite loaded in tension is evaluated by integrating the yarn strain, which is the parallel coupling of yarn strain profiles within the individual crack bridges $\varepsilon_{\mathrm{y}} = F_{\mathrm{y}}/(E_{\mathrm{f}}A_{\mathrm{y}})$, along the whole specimen - delivering the total displacement - and dividing it by the composite specimen length L_{c}:

$$\varepsilon_{\mathrm{c}}(F) = \frac{1}{L_{\mathrm{c}}} \int_0^{L_{\mathrm{c}}} \varepsilon_{\mathrm{y}}(F, x)\mathrm{d}x \tag{23}$$

5. Computational examples

To demonstrate the influence of randomized parameters we observe the composite stress – strain diagram Eq. (23) and the crack width distribution. Parameters shall be varied 'one at the time' to point out their particular contribution to the global response.

5.1. RANDOM MATRIX STRENGTH

Defining the matrix strength by a constant value results in a horizontal line in the stress – strain diagram (ACK (Aveston *et al.*, 1971)) at the composite stress level, which corresponds to the ultimate matrix stress. If the local matrix strength fluctuates around this value, cracks develop at earlier load stages (Fig. 8). In fact, the load at first crack is distributed as the minimum extreme value of the random matrix strength. Furthermore, the composite response is also affected by the predominant correlation of the matrix strength. If the fine grains are the main source of strength correlation, the autocorrelation length l_ρ is rather small. In contrary, a nearly homogeneous matrix with strength fluctuations caused predominantly by outer sources (e.g. casting process, geometrical inaccuracies) can be expected to have a large l_ρ. A case in between these extremes is e.g. the cross-section strength reduction caused by fine shrinkage cracks. Here, the l_ρ of the matrix strength is in the order of a few millimeters. In Fig. 9, the filament and bond parameters are fixed and l_ρ of the matrix strength is varied to represent the three cases mentioned. In the range of approx. $1.0 < \delta_0/l_\rho < 10.0$, where

$$\delta_0 = \frac{(K_{\mathrm{m}} + K_{\mathrm{f}})\sigma_0}{E_{\mathrm{m}}T} \tag{24}$$

is the shielded length and σ_0 is the scale parameter of the matrix strength distribution, the autocorrelation length l_ρ has a significant influence on the degree of tension stiffening - the difference in strains of a composite saturated with crack and of the reinforcement only. It is proportional to the stress remaining in the matrix in the saturated state (gray shaded area in Fig. 9) which on the other hand is proportional to the distance of minimums of the matrix strength For high δ_0/l_ρ ratios, the minimums of matrix strength along

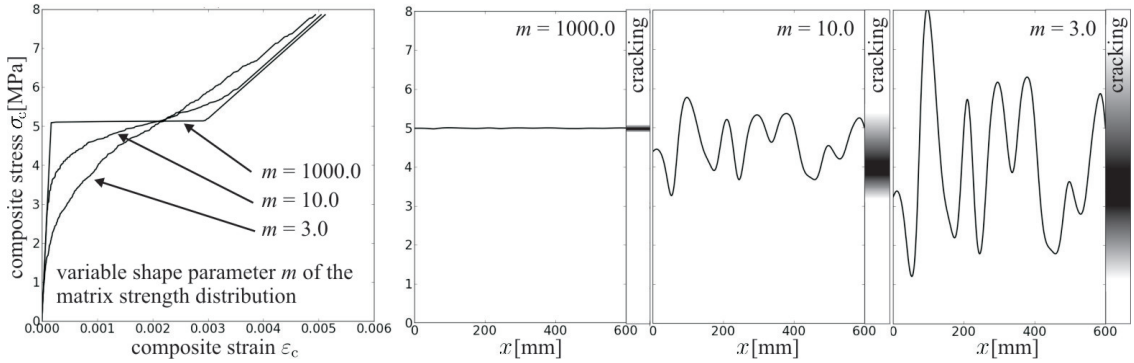


Figure 8. Variable shape parameter m of the matrix strength distribution. Stress – strain diagrams (left), matrix strength along x for variable m (right). Parameters: $A_\mathrm{m} = 50.0$ [mm^2], $E_\mathrm{m} = 30 \cdot 10^3$ [MPa], $L_\mathrm{c} = 3000$ [mm], $A_\mathrm{f} = 5.31 \cdot 10^{-4}$ [mm^2], $E_\mathrm{f} = 72 \cdot 10^3$ [MPa], $\tau = 0.1$ [N/mm^2], $\ell = 0.0$ [mm], $\theta = 0.0$ [-], $\xi = \infty$ [-], $p = 85.0 \cdot 10^{-3}$ [mm], $N_\mathrm{f} = 1700$ [-], $\sigma_\mathrm{mu} =$ autocorrelated field with Weibull distribution (shape = 3.0, 10.0, 1000.0, scale = 5.0, l_ρ = 30.0 [mm])

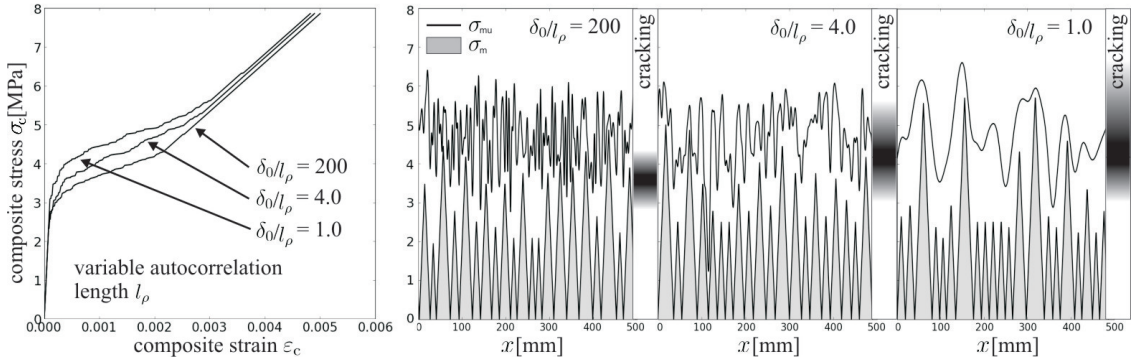


Figure 9. Variable autocorrelation length l_ρ of the matrix strength distribution. Stress – strain diagrams (left), matrix strength σ_mu along x for variable l_ρ and matrix stress $\sigma_\mathrm{m}(x)$ in saturated state (right). Parameters: $A_\mathrm{m} = 50.0$ [mm^2], $E_\mathrm{m} = 30 \cdot 10^3$ [MPa], $L_\mathrm{c} = 3000$ [mm], $A_\mathrm{f} = 5.31 \cdot 10^{-4}$ [mm^2], $E_\mathrm{f} = 72 \cdot 10^3$ [MPa], $\tau = 0.1$ [N/mm^2], $\ell = 0.0$ [mm], $\theta = 0.0$ [-], $\xi = \infty$ [-], $p = 85.0 \cdot 10^{-3}$ [mm], $N_\mathrm{f} = 1700$ [-], $\sigma_\mathrm{mu} =$ autocorrelated field with Weibull distribution (shape = 7.0, scale = 5.0, δ_0/l_ρ = 200.0, 4.0, 1.0)

x are close, so that the matrix cracking takes place mainly at lower stresses and the composite achieves the saturated state at relatively low loading stages. The crack spacing is dense and consequently there is not much stress remaining in the matrix in the saturated state. Low δ_0/l_ρ ratios result in a wider load range of matrix cracking, the saturated state is reached at higher load stages, the crack spacing is larger and the amount of stress stored in the matrix is higher.

The two examples demonstrate the feasibility of the implementation and its capability to reflect the effect of microstructural parameters on the strain-hardening response of the composite. Systematic parametric studies are currently being elaborated and will be presented in the following papers and during the conference presentation.

6. Concluding remarks

The paper describes a multiscale approach for modeling of the fragmentation process of composites with heterogeneous reinforcement. Components of the model were presented separately according to their respective scales. We have presented some results in Sec. 5, which reflect the sensitivity of the global composite behavior on the matrix strength distribution.

We remark that the matrix strength is represented by a single realization of a random field. Consequently, the results are single realizations of a function of random variables. However, the single realizations can be considered as fairly close to the expected values as the length of the modeled specimen L_{c} gets large compared to the autocorrelation length and therefore the variability, according to the central limit theorem, diminishes with the rate $\approx l_\rho / L_{\mathrm{c}}$. We kept this principle in mind when evaluating the presented results.

Summary of assumptions imposed in the model:

- loading results in uniaxial stress in the composite
- cracks are planar and perpendicular to the loading direction
- crack opening is uniform across the composite cross section

7. Acknowledgements

The work has been supported by the German Science Foundation under project number CH276/2-1 and SFB532/T05. Additionally, the work of the first and third author has been supported by the Czech Science Foundation under reg. no. P105/10/J028 and GD 103/09/H085. This support is gratefully acknowledged.

References

B. K. Ahn and W. A. Curtin. Strain and hysteresis by stochastic matrix cracking in ceramic matrix composites. *Journal of the Mechanics and Physics of Solids*, 45 (2):177 – 209, 1996.

J. Aveston and A. Kelly. Theory of multiple fracture of fibrous composites. *J. Mat. Sci. 8*, pages 411–461, 1973.

J. Aveston, G. Cooper, and A. Kelly. Single and multiple fracture, the properties of fibre composites. *Proc. Conf. National Physical Laboratories, IPC Science and Technology Press Ltd. London*, pages 15–24, 1971.

A. Azzam and M. Richter. Investigation of stress transfer behavior in textile reinforced concrete with application to reinforcement overlapping and development lengths. In Ortlepp R. Curbach, M., editor, *6th Colloquium on Textile Reinforced Structures (CTRS6)*, pages 103–116, 2011.

J.E. Bolander and S. Saito. Discrete modeling of short-fiber reinforcement in cementitious composites. *Advanced Cement Based Materials*, 6:76–86, 1997.

B. Budiansky, J.W. Hutchinson, and A.G. Evans. Matrix fracture in fiber-reinforced ceramics. *J. Mech. Phys. Solids*, 34(2):167–189, 1986.

Ch. Cho, J.W. Holmes, and J.R. Barber. Distribution of matrix cracks in a uniaxial ceramic composite. *J. Am. Ceram. Soc.*, 2(75):316–324, 1992.

R. Chudoba, M. Vorechovsky, J. Jerabek, and M. Konrad. Trc-specimens modeled as a chain of cracks. In *ECF16 16th European Conf. of Fracture*, 2006.

R. Chudoba, M. Vorechovsky, and M. Konrad. Stochastic modeling of multifilament yarns. I. Random properties within the cross-section and size effect. *International Journal of Solids and Structures*, 43(3-4):413–434, 2006.

B. D. Coleman. On the strength of classical fibres and fibre bundles. *Journal of Mechanics and Physics of Solids*, 7:60–70, 1958.

W.A. Curtin, B.K. Ahn, and N. Takeda. Modeling brittle and tough stress-strain behavior in unidirectional ceramic matrix composites. *Acta materialia*, 46(10):3409–3420, 1998.

W.A. Curtin. Exact theory of fiber fragmentation in a single-filament composite. *Journal of Materials Science*, 26:5239–5253, 1991.

W.A. Curtin. Multiple matrix cracking in brittle matrix composite. *Acta metall. mater.*, 41(5):1369–1377, 1992.

H. E. Daniels. The statistical theory of the strength of bundles of threads. i. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 183(995):405–435, 1945.

A.P. Fantilli, H. Mihashi, and P. Vallini. Multiple cracking and strain hardening in fiber reinforced concrete under uniaxial tension. *Cement and Concrete Research*, 39:1217–1229, 2009.

R.B. Henstenburg and S.L. Phoenix. Interfacial shear strength studies using the single-filament-composite test. part ii: A probability model and monte carlo simulation. *Polymer Composites*, 10(5):389–408, 1989.

P. Kabele. New developments in analytical modeling of mechanical behavior of ecc. *Journal of Advanced Concrete Technology*, 1(3):253–264, 2003.

M. Konrad and R. Chudoba. Tensile behavior of cementitious composite reinforced with epoxy impregnated multifilament yarns. *International Journal for Multiscale Engineering*, 7(2):115–133, 2009.

J.P.B. Leite, V. Slowik, and H. Mihashi. Computer simulation of fracture processes of concrete using mesolevel models of lattice structures. *Cement and Concrete Research*, 34:1025–1033, 2004.

V.C. Li and H.C. Wu. Conditions for pseudo strain-hardening in fiber reinforced brittle matrix composites. *Appl. Mech. Rev.*, 45(8):390–398, 1992.

V. C. Li, Y. Wang, and S. Backer. A micromechanical model of tension-softening and bridging toughening of short random fiber reinforced brittle matrix composites. *Journal of the Mechanics and Physics of Solids*, 39(5):607–625, 1991. fiber orientation considered by probabillity distribution function.

D.B. Marshall and A.G. Evans. Failure mechanisms in ceramic-fiber ceramic-matrix composites. *J. Am. Ceram. Soc.*, 68(5), 1985.

A.N. Netravali, R.B. Henstenburg, S.L. Phoenix, and P. Schwartz. Interfacial shear strength studies using the single-filament-composite test. i: Experiments on graphite fibers in epoxy. *Polymer Composites*, 10(4):226–241, 1989.

A. Nour, B. Massicotte, R. Montaignac, and J.P. Charron. Derivation of a crack opening deflection relationship for fibre reinforced concrete panels using a stochastic model: Application for predicting the flexural behaviour of round panels. *Cement and Concrete Research*, 41:964–974, 2011.

S.L. Phoenix and H.M. Taylor. The asymptotic strength distribution of a general fiber bundle. *Advances In Applied Probability*, 5:200 – 216, 1973.

S. L. Phoenix. Statistical theory for the strength of twisted fiber bundles with applications to yarns and cables. *Textile Research Journal*, 49:407–423, 1979.

F.K.F. Radtke, A. Simone, and L.J. Sluys. A computational model for failure analysis of fibre reinforced concrete with discrete treatment of fibres. *Engineering Fracture Mechanics*, 77:597–620, 2010.

M. Vorechovsky. Simulation of simply cross correlated random fields by series expansion methods. *Structural safety*, (30):337363, 2008.

A Multi-objective Optimization Approach with a View to Robustness Improvement

Aleksandra Serafinska, Wolfgang Graf, Michael Kaliske
Institute for Structural Analysis, Technische Universität Dresden, D-01062 Dresden,
aleksandra.serafinska@tu-dresden.de

Abstract. In this contribution, multi-objective optimization methods are applied together with uncertainty quantification approaches in order to provide a concept for a robust structural design. The concept enhances the utilization of numerical simulation methods (Finite Element analysis) and as such can be useful for the computer-aided engineering. In this study, the application of the approach for a design of tires is shown. The proposed methodology enables the consideration of fragmentary or dubious information within the design process, which leads to the introduction of fuzzy variables into the optimization task. The application of fuzzy set theory is motivated by the epistemic character of available uncertain data. The proposed concept enables the optimization of multiple objectives and simultaneously the uncertainty reduction in the optimization results, which leads to the robustness improvement. In order to increase the numerical efficiency of the proposed design approach, a response surface approximation based on artificial neural networks is applied.

Keywords: optimization, numerical simulations, uncertainty

1. Introduction

The intensive development of multi-objective optimization methods in the past decades, as well as the coupling of these methods with numerical simulation approaches, e.g. Finite Element method (FEM), enables currently the solution of complex design problems in numerous engineering fields. The main feature of these design tasks is that the considered multiple objectives are predominantly in conflict with each other. This means, that generally no ideal solution exists.

There are two well established optimization approaches, which yield a solution for these kind of problems. The first approach encloses the formulation of an aggregate objective function (AOF), which combines all considered objectives through the application of the weighted sum method. Within AOF, each single objective function is preserved with a weighting factor, which is chosen subjectively by the decision maker in order to express the preference of this objective. Thereby, the choice of weighting factors, which are gathered into a preference vector, strongly affects the optimization result.

In order to avoid the insertion of subjective decisions into the optimization process, some objective approaches are proposed. They enable the identification of a well distributed set of trade-off solutions (Pareto-optimal set), instead of finding one suboptimal solution (Pareto, 1971). The solutions in Pareto-optimal set P^* are satisfying the criterion of Pareto optimality. That means, that in the set of solutions P, the Pareto-optimal set P^* contains solutions, which are not dominated by any member of set P.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

After the completion of optimization, the decision maker can choose one solution from the Pareto-optimal set P^*, according to defined preferences. The main advantage of this approach, compared to the AOF method is that the expression of preferences occurs in the post-optimization step. A detailed description of the Pareto-optimality concept as well as the dominance concept is presented in Section 3.1.

The optimization approaches shortly described above are deterministic. Thereby, in the engineering design tasks, uncertainty has to be taken into account. Uncertainty is present in different forms: e.g. geometry parameters of structural parts can be regarded as uncertain, as well as the properties of materials utilized for the components. In some cases also the loads applied to the designed structure can be considered as uncertain. Thereby, the sources of uncertainty can be various, e.g. the uncertainty in geometry or material properties is caused by unstable production conditions of structural elements. The source of uncertainty in loading is vague information or variability, e.g. in the case of tire design, considered here, the vertical load applied to the tire changes in dependency on the car weight, which is different for diverse car models.

According to the uncertainty sources, it is distinguished between three characteristics of uncertainty: variability, imprecision and incompleteness. In this study, the focus is set on the incompleteness as in many engineering tasks we have to handle with vague information, leading to assumptions and expert evaluations. A suitable model for describing this kind of uncertainty is the uncertainty model fuzziness. In this paper, the application of fuzzy variables to the mentioned multi-objective optimization concepts will be studied.

If uncertainty is considered in the optimization process, a quantification of robustness can be accomplished subsequently, which is shown within the proposed optimization approach.

2. Modelling uncertain quantities

In order to properly consider the uncertainty within the design task, a suitable uncertainty model should be chosen, dependent on the type of available information. Commonly, the probabilistic uncertainty models are utilized (Benjamin and Cornell, 1970), which employ the random variables for the description of non-deterministic parameters. The application of probabilistic models is preconditioned by the availability of extensive statistical information. If this prerequisites are not met, other uncertainty models should be taken into account, especially enabling the consideration of subjective information.

These models account for the Bayesian methods (Bernardo and Smith, 1994) or approaches based on the fuzzy set theory. The uncertainty model fuzziness (Dubois and Prade, 1997) and (Zadeh, 1965) employs the fuzzy set theory for modelling the vague, incomplete or subjective information. Within fuzzy sets, the gradual membership of elements to the set is defined, which enable a subjective weightening of information inside the set. Alternatives for modelling with fuzzy sets represent the convex modelling (Elishakoff, 1995) and interval mathematics (Alefeld and Herzberger, 1983). Though, the last concepts are based on the assumption of crisp membership of the elements to the set and, therefore, provide limited modelling capabilities in comparison to fuzzy sets.

A generalized uncertainty model, enabling accounting for objective and subjective information simultaneously – fuzzy-randomness is described in (Kwakernaak, 1978) and (Möller and Beer, 2004). Fuzzy-randomness can be utilized for modelling of imprecise probabilities. Within this model, fuzziness and randomness might be considered as special cases. Another approach, which allows the consideration of uncertainty – the chaos theory (Kapitaniak, 2000) makes an attempt to describe the unpredictable behaviour of dynamical systems.

From the uncertainty models, mentioned above, the uncertainty model fuzziness is chosen for further investigation and applied within the multi-objective optimization approach. The choice of fuzziness is motivated by the capability of the model to describe vague, incomplete or subjective information, which is predominant in engineering design tasks.

The formulation of a fuzzy set refers to the definition of a crisp set. The membership to a crisp set $A \subseteq X$, where $X = \mathbb{R}^n$ can be defined in a binary way, an element either belongs to the set or not. Thereby, the membership $\mu_{\tilde{A}}$ to a fuzzy set $\tilde{A}$ is defined gradually, see Fig. 1. If $\mu_{\tilde{A}}$ takes values within the interval [0,1] and, at least, once the value 1 is achieved, than such a set is called a normalized fuzzy set $\tilde{A}$ or a fuzzy number on X

$$\mu_{\tilde{A}} : X \longrightarrow [0, \infty). \tag{1}$$

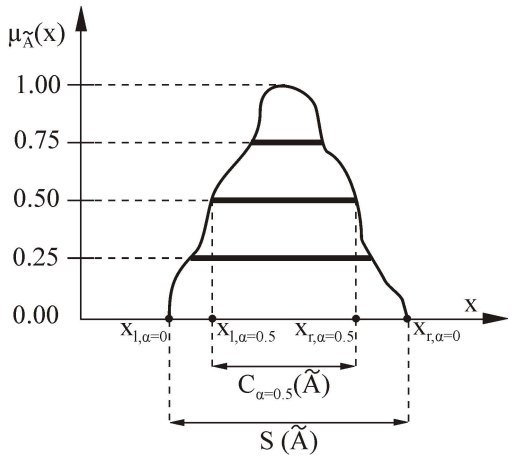


Figure 1. Fuzzy variable.

A fuzzy set is defined by its support $S(\tilde{A})$ and the membership function $\mu_{\tilde{A}}$. According to Fig. 1, $S(\tilde{A})$ is a crisp set, which contains elements

$$S(\tilde{A}) = \{x \in X, \mu_{\tilde{A}}(x) > 0\}. \tag{2}$$

In the optimization approach considered here, the fuzzy quantities are defined as normalized fuzzy sets. Thereby, the convexity of fuzzy sets is presumed. Convexity can be stated, if for every $x_1, x_2 \in X$ and $\lambda \in [0, 1]$

$$\mu_{\tilde{A}}(\lambda x_2 + (1 - \lambda)x_1) \geqslant min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2)). \tag{3}$$

The numerical treatment of a fuzzy quantity $\tilde{A}$ occurs by means of the discretization of $\tilde{A}$ by numerous crisp sets $C_\alpha(\tilde{A})$ – so-called α-level sets

$$C_\alpha(\tilde{A}) = \{x \in X : \mu_{\tilde{A}} \geqslant \alpha\}, \tag{4}$$

$$\tilde{A} = (C_\alpha(\tilde{A}))_{\alpha \in (0,1]}. \tag{5}$$

Fuzzy quantities regarded in this study are n-dimensional. They are enclosed in the set of all fuzzy quantities $\mathcal{F}(\mathbb{R}^n)$.

3. Multi-objective optimization

3.1. Deterministic multi-objective optimization approaches

In the deterministic multi-objective optimization (MOO) task, we consider design variables, described in terms of design vectors $\underline{x}_d = (x_{d1}, x_{d2}, ..., x_{dn})$, which are defined in the design space $X_d = \mathbb{R}^n$. The design vectors $\underline{x}_d$ are mapped by means of the evaluation function $f : X_d \rightarrow Z$ onto the objective vectors $\underline{z}(z_1, z_2, ..., z_k)$ in the objective space $Z = \mathbb{R}^k$. Due to the fact, that the evaluation function f is vector-valued, the objective space is k-dimensional. The multi-objective optimization task with objective functions $f_i,\ i = 1, ..., k$, subjected to equality constraints $h(\underline{x}_d)$ and inequality constraints $g(\underline{x}_d)$ is formulated as

$$min\{f(\underline{x}_d) \mid h(\underline{x}_d) = 0,\ g(\underline{x}_d) \leq 0\}\ ,\ f(\underline{x}_d) = [f_1(\underline{x}_d), ..., f_k(\underline{x}_d)]^T. \tag{6}$$

Thereby, it should be pointed out, that if $k = 1$, a single objective optimization problem is to solve, or accordingly a multi-objective optimization problem reduced by means of an aggregate function to a single objective problem. If $k = 1$, a direct comparison of one-dimensional objective vectors is accomplished within the optimization process. For $k > 1$, the k-dimensional objective vectors shall be compared, which can be carried out only through the utilization of the dominance concept.

Dominance: an objective vector $\underline{z}^* = f(\underline{x}_d^*)$ dominates another objective vector $\underline{z}' = f(\underline{x}_d')$ if no component of $\underline{z}^*$ is greater than the corresponding component of $\underline{z}'$ and at least one component is smaller

$$\forall i \in \{1, ..., k\} : f_i(\underline{x}_d^*) \leq f_i(\underline{x}_d') \wedge \exists i \in \{1, ..., k\} : f_i(\underline{x}_d^*) < f_i(\underline{x}_d'). \tag{7}$$

The dominance is formulated as $\underline{z}^* \succ \underline{z}'$.

Pareto-optimal set: the goal of the multi-objective optimization approach is to find a set of solutions, which are not dominated with respect to each other (non-dominated set). According to the definition provided in Section 1, if P is the entire design space X_d, than the non-dominated set P^* (or X_d^*) is a Pareto-optimal set. The visualization of the Pareto-optimal set X_d^* in the objective space is the Pareto-front $Z^* = f(X_d^*) \subseteq Z$.

As mentioned in Section 1, the state of the art in the multi-objective optimization methods are approaches that either evaluate the aggregate objective function and refer to single objective optimization methods or approximate the Pareto-optimal set in different manners.

According to (Deb, 2002a), within the available multi-objective optimization methods it is distinguished between classical methods and evolutionary algorithms. The classical methods transform the multi-objective optimization problem into a single-objective optimization task by the application of different user-specified techniques. In this group, the weighted sum method, enabling the formulation of an aggregate function or the ϵ-constraint method, converting all objective functions, except of one, into constraints can be identified. These methods yield, after completing the optimization, one sub-optimal objective vector. An aggregate function, created, using the weighted sum method is formulated

$$f_{obj}(\underline{x}_d) = \sum_{i=1}^{k} w_i f_i(\underline{x}_d). \tag{8}$$

Thereby, w_i defines a user-specified weight vector. Some approaches are proposed, which presume, that through the utilization of the weighted sum method and an appropriate choice of the weight vector, a Pareto-optimal solution can be identified. A purposeful manipulation of the weight vector components in multiple

optimization runs, could provide a whole Pareto-optimal set. Though, the evaluation of the optimization procedure numerous times, required for the identification of the whole Pareto-optimal set, is quite inefficient. Therefore, evolutionary algorithms, which enable providing the Pareto-optimal set in one optimization run, are commonly employed for the solution of multi-objective optimization problems.

The multi-objective evolutionary algorithms (MOEAs) provide not optimal trade-offs but an approximation of the Pareto-set. In general, MOEAs are expected to fulfil two tasks: guide the search through the Pareto-set and keep a diverse set of non-dominated solutions. The first goal is achieved by assigning the fitness to the population, based on the non-dominated sorting method, while the second goal by including the density information into the selection process. In 1990s several methods, e.g. the Nondominated Sorting Genetic Algorithm (NSGA) (Srinivas and Deb, 1995) or Multi-objective Genetic Algorithm (MOGA) (Fonseca and Fleming, 1993) were proposed, which were able to identify multiple diverse Pareto solutions. In further developments of MOEA, elitism was introduced in order to obtain a better convergence. Elitism enables the prevention of non-dominated solutions from being lost. Among the developed methods, which include elitism, three main approaches should be mentioned: NSGA-II (Deb, 2002b), the Strength-Pareto Evolutionary Algorithm (SPEA-2) and the Pareto-Archieved Evolution Strategy (PAES) (Knowles and Corne, 1999).

In the presented contribution, the application of fuzzy quantities within the weighted sum method (aggregate function) and to methods, providing the trade-off solutions will be studied.

3.2. MULTI-OBJECTIVE OPTIMIZATION WITH CONSIDERATION OF UNCERTAIN QUANTITIES

3.2.1. *Problem formulation*

In the literature, several methods were proposed which evaluate uncertainty within the multi-objective optimization. Thereby, referring to (Das et al., 2009), different sources of uncertainty are taken into account, e.g. noisy data, objective function evaluation errors or user indecision concerning the prioritization of objective functions. Generally, within the available approaches it is distinguished between methods, which evaluate aleatory uncertainty and utilize probabilistic concepts and methods, which account for epistemic uncertainty and employ concepts based on the fuzzy set theory. Recently, ideas and algorithms for the simultaneous consideration of different types of uncertainty and different uncertainty models (polymorphic uncertainty) within the optimization task are developed.

The main difference between deterministic multi-objective optimization approaches and approaches considering uncertainty, is expressed within the formulation of dominance. In the probabilistic concepts, the dominance of vector $\underline{u}$ over the vector $\underline{v}$ is expressed by the probability of dominance $pr(\underline{u} \succ \underline{v})$. Exemplary, such approaches model the random error, present by the evaluation of objective functions. Among the epistemic procedures, the idea proposed by (Farina and Amato, 2004) for the introduction of a fuzzy measure for the comparison of two non-dominated solutions, is worth mentioning. The comparison is accomplished through the evaluation of the number of objectives, in which one solution dominates another one. Further developments enabled the comparison of two solutions, which do not have to be Pareto-optimal – like in the previous approach – by the application of the concept of a fuzzy dominance.

In this study, instead of employing the concept of fuzzy dominance, ideas for obtaining and evaluation of a Fuzzy-Pareto-Front will be discussed. A Fuzzy-Pareto-Front contains fuzzy objective vectors, gained from the optimization with the application of fuzzy variables. The method for obtaining fuzzy results will be first introduced within the multi-objective optimization, using the weighted sum method for the creation

of the aggregate function. Subsequently, a method for evaluating fuzzy objective vectors within the Fuzzy-Pareto-Front will be discussed.

3.2.2. *Aggregate objective function evaluating fuzzy variables*

In the proposed methodology, a distinction is made within the definition of input quantities for the optimization task, between uncertain a-priori parameters $\underline{\tilde{p}}_a \in \mathcal{F}(\mathbb{R}^{np})$ and design variables $\underline{x}_d \in \mathbb{R}^n$. Uncertain a-priori parameters $\underline{\tilde{p}}_a$ are quantities, influencing the optimization task, which can not be modelled as crisp quantities due to information deficits. Therefore, they are considered as fuzzy numbers. Design variables $\underline{x}_d$ are defined within the user-specified ranges and can be arbitrarily chosen during the optimization.

The scheme of the optimization with fuzzy quantities is presented in Figure 2. The method is a three level approach. Optimization establishes the first level, that is the outer loop of the approach. Within the optimization loop, the fuzzy analysis is performed. The numerical realization encloses the execution of fuzzy analysis for every design vector $\underline{x}_d$. In consequence, if k objective functions are considered, k fuzzy result quantities are obtained for each design vector. These fuzzy result quantities are gathered into k-dimensional fuzzy objective vector $\underline{\tilde{z}} \in \mathcal{F}(\mathbb{R}^k)$.

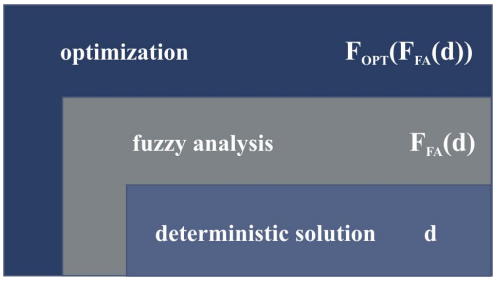


Figure 2. Optimization scheme.

Within the fuzzy analysis, the deterministic solution, that is a FE-solution or a response surface is evaluated numerous times. The fuzzy analysis uses the α-level optimization approach for the computation of membership functions of all fuzzy result quantities. For the α-level optimization, the modified evolution strategy is utilized, which was proposed in (Möller, Graf and Beer, 2000). In order to be able to evaluate the fuzzy result quantities, obtained for several designs, an aggregate objective function, enclosing the information reducing measures $\mathcal{M}_j$ is formulated

$$f_{obj} : \mathbb{R}^n \times \mathcal{F}(\mathbb{R}^{np}) \rightarrow \mathbb{R} \tag{9}$$

$$(\underline{x}_d, \underline{\tilde{p}}_a) \mapsto \sum_{i=1}^{k} \sum_{j=1}^{u} w_{ij} \, \mathcal{M}_j(f_i(\underline{x}_d, \underline{\tilde{p}}_a)). \tag{10}$$

In Eq. (10), w_{ij} defines the weighting factors, which enable the prioritization of chosen components (objectives). The information reducing measures $\mathcal{M}_j$ map the fuzzy result quantities onto real numbers and allow their quantification. As $\mathcal{M}_j$, uncertainty measures for fuzzy variables can be applied, e.g. the area (zeroth moment) of a fuzzy variable, variance or the Shannon's entropy. The uncertainty measure $\mathcal{M}_1$, based on the Shannon's entropy (Beer and Liebscher, 2008) and applied to the quantification of a one-dimensional

fuzzy result quantity $\tilde{z}$, is formulated as

$$\mathcal{M}_1 = H_U = -\int_{z=-\infty}^{z=+\infty} [\mu(z) \cdot ln\mu(z) + (1-\mu(z)) \cdot ln\,(1-\mu(z))]\, dz. \tag{11}$$

An uncertainty measure $\mathcal{M}_2$, evaluating the variance of a fuzzy variable $\tilde{z}$ is proposed in (Wu and Mendel, 2007)

$$\mathcal{M}_2 = V = \int_{z=-\infty}^{z=+\infty} (z-\overline{z})^2 \cdot \mu(z)dz \cdot \left(\int_{z=-\infty}^{z=+\infty} \mu(z)dz \right)^{-1}. \tag{12}$$

Another uncertainty measure $\mathcal{M}_3$, which assesses the area under the membership function $\mu(z)$ of a fuzzy variable is defined by

$$\mathcal{M}_3 = A = \int_{z=-\infty}^{z=+\infty} \mu(z)dz. \tag{13}$$

For the optimization approach, described in this study two information reducing measures are employed. The first measure is the uncertainty measure $\mathcal{M}_3$. The second information reducing measure $\mathcal{M}_4$ assesses the position of the first element of the fuzzy quantity support $S(\tilde{z})$ and can be regarded as a performance measure. Through the minimization of the smallest element of the support $S(\tilde{z})$, the minimization of the fuzzy result quantity $\tilde{z}$ is achieved.

In Fig. 3, the application of the information reducing measures for the comparison of two fuzzy result quantities $\tilde{z}_1$ and $\tilde{z}_2$, obtained for two different designs – A and B – is shown. Once the criteria for the comparison of uncertain quantities are formulated and included into the aggregate objective function, the optimization can succeed. The application of the method for single-objective optimization tasks is shown in (Pannier, 2011).

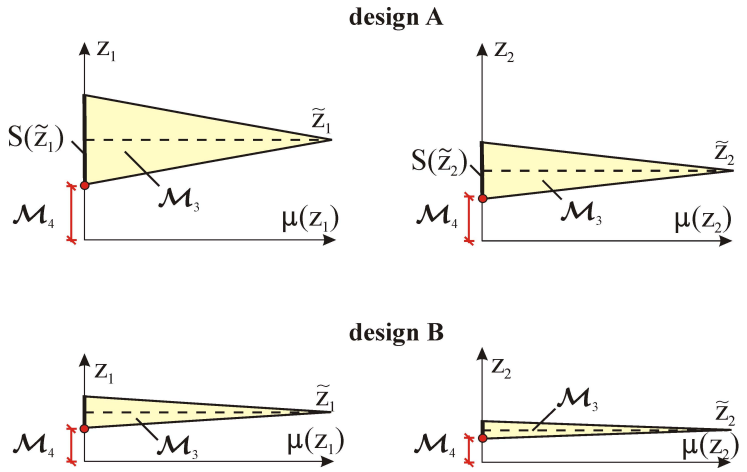


Figure 3. Comparison of fuzzy result quantities for design A and B.

3.2.3. *Fuzzy-Pareto solutions*

In Subsection 3.2.2., a method for handling fuzzy results within the optimization is presented, which focused on bringing together the fuzzy results from several objectives into one aggregate objective function, enclosing the information reducing measures $\mathcal{M}_j$. Another crucial issue, when regarding the objective space with numerous k-dimensional fuzzy objective vectors, is to find the Fuzzy-Pareto-Front. Fuzzy-Pareto-Front is the image of the Fuzzy-Pareto set in the objective space. In the context of the optimization approach with fuzzy variables, described above, the search for the best design should be coupled with the identification of the set of fuzzy result quantities, which are not dominated by any other fuzzy quantity from the set of all fuzzy quantities $\mathcal{F}(\mathbb{R}^k)$.

The objective space, containing k-dimensional fuzzy result vectors $\tilde{\underline{z}}^1 - \tilde{\underline{z}}^5$, obtained for corresponding design vectors $\underline{x}_{d1} - \underline{x}_{d5}$ is shown in Fig. 4 (here k = 2). Thereby, fuzzy result vectors $\tilde{\underline{z}}^1 - \tilde{\underline{z}}^3$ are not dominated by any other fuzzy result vector in the objective space.

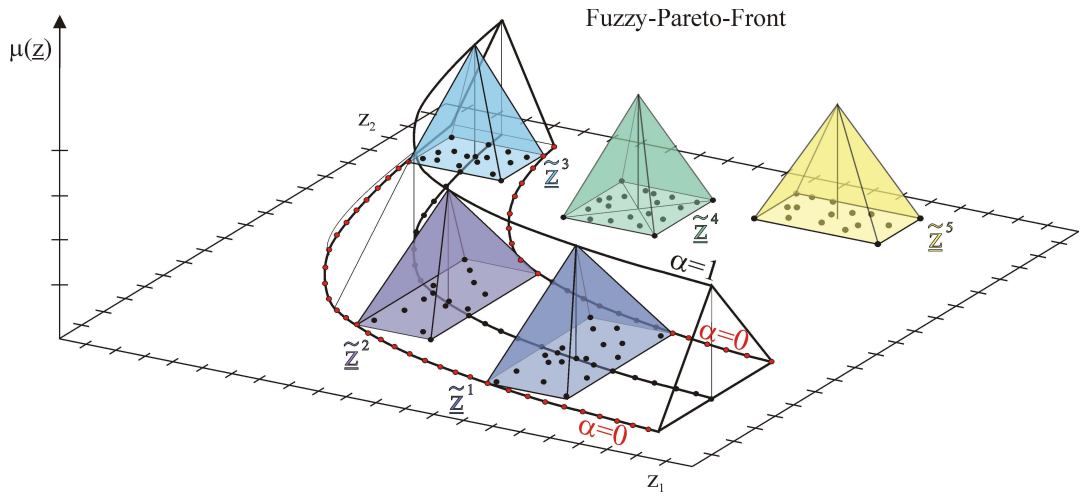


Figure 4. Fuzzy-Pareto-Front.

The k-dimensional fuzzy objective vector $\tilde{\underline{z}}^*$ is formulated

$$\tilde{\underline{z}}^* = \{(\underline{z}^* = (z_1, ..., z_k),\ \underline{\mu}^* = (\mu_1^*, ..., \mu_k^*)) \mid z_i \in \mathbb{R}^k\}. \tag{14}$$

The support $S(\tilde{\underline{z}}^*)$ is defined by

$$S(\tilde{\underline{z}}^*) = \{(\underline{z}^* = (z_1, ..., z_k),\ \underline{\mu}^* = (\mu_1^*, ..., \mu_k^*))\ \forall\ i \in [1, k],\ \mu_i^*(z_i) > 0\}. \tag{15}$$

The identification of the set of non-dominated k-dimensional fuzzy objective vectors prerequires the formulation of non-dominance criteria for fuzzy variables. A fuzzy objective vector $\tilde{\underline{z}}^*$ is non-dominated, if in the set of all fuzzy objective vectors $\mathcal{F}(\mathbb{R}^k)$, there exists no other fuzzy objective vector $\tilde{\underline{z}}'$, so that every element $\underline{z}' \in S(\tilde{\underline{z}}')$ dominates all elements $\underline{z}^* \in S(\tilde{\underline{z}}^*)$.

The proposed concept enables the check of dominance in the postcomputation step. Though, an approach is required, which would allow a dominance check for fuzzy quantities during the multi-objective optimization. In this way, the information concerning the dominance can influence the selection and variation step

in the evolutionary optimization algorithm. First attempts are made to extend the non-dominated sorting procedure, applied within NSGA II for the consideration of the domination criteria for fuzzy variables.

3.3. Coupling of multi-objective optimization approaches with FE solution

According to Fig. 2, if FE simulation is applied as deterministic solution d, the evaluation function f formulated in Eq. (6) as well as the aggregate function f_{obj} defined in Eq. (10), depend on the solution of a mechanical system, especially depend on displacements $\varphi(\underline{x}_d, \underline{\tilde{p}}_a)$ gained from the FE analysis. $\varphi(\underline{x}_d, \underline{\tilde{p}}_a)$ are obtained from the evaluation of the nonlinear equation of motion. In the case of a stationary rolling body (tire analysis), the equation of motion is formulated as

$$(\underline{K} - \underline{W}) \cdot \Delta\underline{\varphi} = \underline{f} - \underline{f}_{\sigma} + \underline{f}_{T}. \tag{16}$$

In Eq. (16), $\underline{K}$ denotes the tangential stiffness matrix, $\underline{W}$ the Arbitrary Lagrangian Eulerian inertia matrix, $\underline{f}$ nodal forces of external loads, $\underline{f}_{\sigma}$ nodal forces resulting from internal stress state and $\underline{f}_{T}$ nodal forces due to inertia. The dependency of the aggregate function f_{obj} on displacements $\varphi(\underline{x}_d, \underline{\tilde{p}}_a)$ is defined as

$$f_{obj} : \mathbb{R}^n \times \mathcal{F}(\mathbb{R}^{np}) \rightarrow \mathbb{R}, \tag{17}$$

$$(\underline{x}_d, \underline{\tilde{p}}_a) \mapsto \sum_{i=1}^{k} \sum_{j=1}^{u} w_{ij}\ \mathcal{M}_j(f_i(\varphi(\underline{x}_d, \underline{\tilde{p}}_a))). \tag{18}$$

3.4. Implicit robustness quantification within the multi-objective optimization

Due to the consideration of one uncertainty measure – $\mathcal{M}_3$ – within the aggregate objective function f_{obj}, the minimization of f_{obj} will automatically cause the reduction of uncertainty in the fuzzy results. This concept refers to the robustness measure $R_{l,k}^{[p]}$, proposed in (Beer and Liebscher, 2008) and in (Graf et al., 2010), which quantifies the ratio of the uncertainty of input quantities versus the uncertainty of result quantities for each design. $R_{l,k}^{[p]}$ adapted for the application within the multi-objective optimization is defined

$$R_{l,k}^{[p]} = \frac{\sum\limits_{p=1}^{np} w_p \left(\mathcal{M}_j \left(\underline{\tilde{p}}_a \right) \right)}{\sum\limits_{q=1}^{l} w_q \sum\limits_{i=1}^{k} w_i \left(\mathcal{M}_j \left(f_i \left(\underline{x}_d, \underline{\tilde{p}}_a \right) \right) \right)}. \tag{19}$$

In Eq. (19), $R_{l,k}^{[p]}$ denotes the robustness measure of $[p]$-th structural design under consideration of l load cases and k objective functions. $\mathcal{M}_j$ indicates the $[j]$-th uncertainty measure (here $j = 3$) and w_p, w_q, w_i the weighting factors. The configuration of uncertain a-priori parameters does not change during the optimization task. Therefore, the numerator of the fraction in Eq. (19) remains constant and can be neglected. Only the expression in the denominator of the fraction is considered within the aggregate objective function f_{obj}.

The proposed design approach enables beside of optimization of numerous objectives also the uncertainty reduction within the fuzzy results, which contributes to the robustness improvement. The application of the developed method to the structural tire design is shown by the way of an example.

4. Example

The goal of this study is the optimization and the robustness improvement of a passenger car tire 195/60 R15. Generally, within the tire structure the tread layer, made of rubber, several reinforcement layers – capplies and belts – as well as the bead, consisting of numerous steel cords can be identified, see Fig. 5.

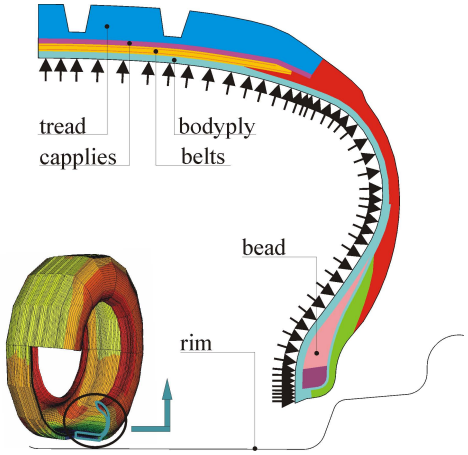


Figure 5. Tire cross-section.

Through the modifications of these structural parts, two objective functions f_1 and f_2 will be improved. The first objective function f_1 aims at providing regular tire wear, whereas f_2 at improving the fatigue performance. Regular tire wear is obtained, if a uniform contact pressure distribution in the tire-road contact zone occurs. Therefore, a ratio of the contact pressure in the shoulder region versus the contact pressure in the central region of the tire cross-section will be optimized. Within f_2, a fatigue criterion based on the evaluation of the strain energy density at the critical area – the belt edge – is applied.

The optimization of mentioned objectives is accomplished by the consideration of three design variables – the belt angle x_{d1}, the thickness of the tread layer x_{d2} and the number of capplies x_{d3} as well as three uncertain a-priori parameters – the tire inner pressure $\tilde{p}_{a1}$, the fiber spacing in bodyply $\tilde{p}_{a2}$ and the stiffness of the tread compound $\tilde{p}_{a3}$. The following ranges for the design variables are specified: x_{d1}: $\langle 18°; 30° \rangle$, x_{d2}: $\langle -1.5; 1.5 \rangle$mm and x_{d3}: $\langle 0; 2 \rangle$. The uncertain a-priori parameters are defined as fuzzy numbers – $\tilde{p}_{a1}$: $\langle 0.23; 0.25; 0.27 \rangle$N/mm^2, $\tilde{p}_{a2}$: $\langle 1.17; 1.304; 1.44 \rangle$mm and $\tilde{p}_{a3}$: $\langle 0.875; 0.976; 1.075 \rangle$N/mm^2.

In the first analysis step, the design of experiments (DOE) is performed. For each of the sampling points a 3D Finite Element tire model is evaluated in steady state rolling situation. The steady state rolling analysis is executed in terms of the Arbitrary Lagrangian Eulerian approach (Kaliske et al., 2003) and (Nackenhorst, 2004). Due to the high computational cost of the rolling tire analysis, the FE solution is substituted by a neural network based response surface approximation. The training of the feedforward neural networks occurs for the sampling points, evaluated within DOE. The obtained response surface can be applied as the deterministic solution within the optimization approach, according to the scheme in Fig. 2.

After the execution of the coupled approach of optimization and fuzzy analysis, which uses the aggregate objective function, one design vector is identified as the optimal solution. For this design, the FE analysis is performed in the 'post processing' step in order to validate the neural network outputs. The results, obtained

for the optimal solution are compared with results gained for reference designs. In Fig. 6, the contact pressure distributions in the optimal design – Fig. 6a), 6b) – and in the reference design – Fig. 6c), 6d) – are depicted. For each design, distributions obtained for the most advantageous – Fig. 6a), 6c) – and most disadvantageous – Fig. 6b), 6d) – combination of a-priori parameters are shown. These most advantageous and most disadvantageous combinations are derived from the definition of the contact pressure ratio as a fuzzy quantity, obtained for the regarded design.

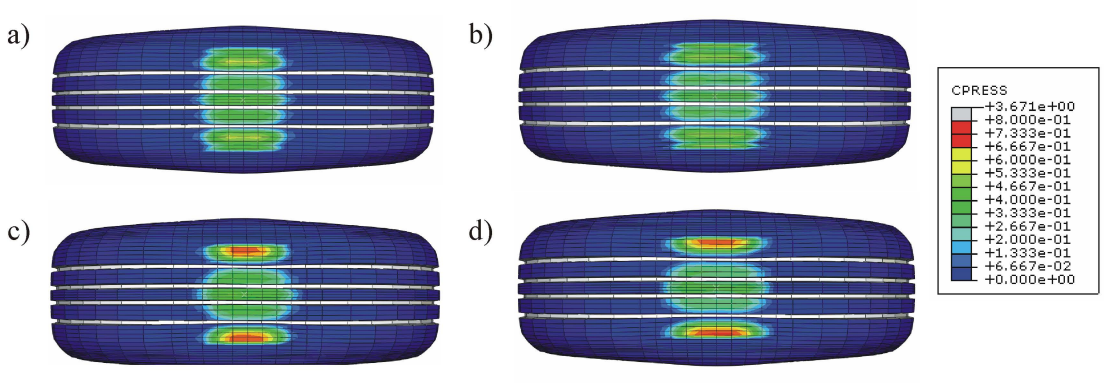


Figure 6. Contact pressure distribution for the optimal design: a) best case, b) worst case and reference design: c) best case and d) worst case.

It can be stated, that for the optimal design a uniform contact pressure distribution is obtained in the best case as well as in the worst case. This fact is confirmed by the according contact pressure ratios $p_{coeff} = 1.23$ and $p_{coeff} = 1.27$, see Fig. 7a). Therefore, the occurrence of a regular wear is expected for the optimal design. Within the reference design, a non-uniform contact pressure distribution occurs, resulting in $p_{coeff} = 1.68$ (best case) and $p_{coeff} = 1.97$ (worst case). Thus, not only the minimization of the contact pressure ratio but also the uncertainty reduction (small support) is achieved for the optimal design.

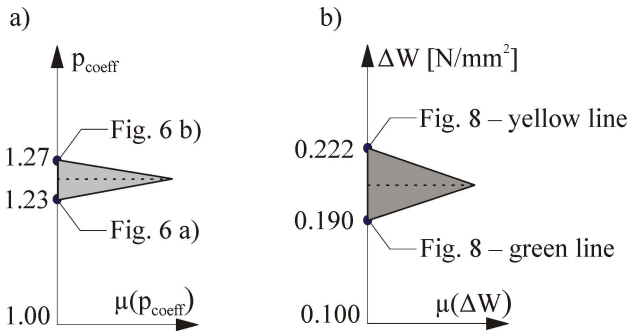


Figure 7. Contact pressure ratio and strain energy density amplitude obtained for the optimal design.

The second design task presumed the minimization of the strain energy density delta, which leads to the improvement of the tire resistance to fatigue. In Fig. 8, the strain energy density evaluated over the circumference of a tire is shown. According to the strain energy density as a fuzzy result quantity, four curves are shown – signifying the worse and best case for the optimal design and for the reference design as

well. The fuzzy result quantity, obtained for the optimal design is shown in Fig. 7b). From the comparison of the curves in Fig. 8, it can be stated that, peaks of curves obtained for the optimal design lie below the peaks of curves obtained for the reference design. The improvement is confirmed also by the strain energy density amplitudes: for the optimal design $\Delta W = 0.190\text{N/mm}^2$ (best case) and $\Delta W = 0.222\text{N/mm}^2$ (worst case), whereas for the reference design $\Delta W = 0.309\text{N/mm}^2$ and analogically $\Delta W = 0.321\text{N/mm}^2$.

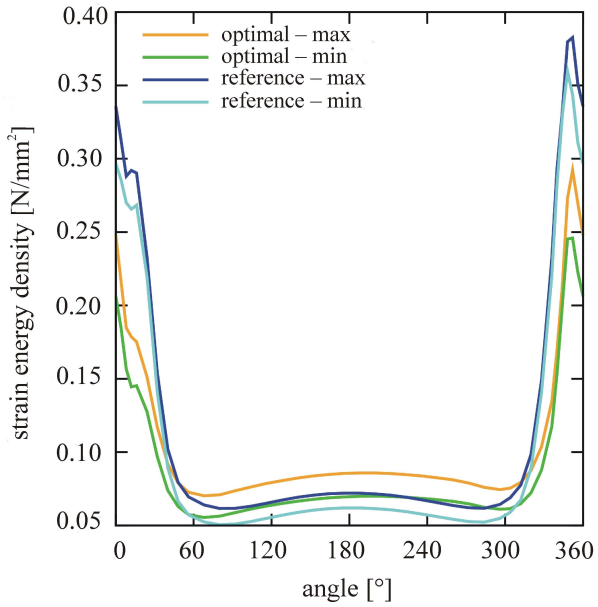


Figure 8. Strain energy density versus the circumference of a tire.

In the example, the capability of the approach to optimize the objectives and to reduce the uncertainty of fuzzy results is shown.

5. Conclusions

In this contribution, concepts for the consideration of fuzzy variables within the multi-objective optimization approaches are discussed. Fuzziness is chosen as an appropriate model for the description of parameters, which can not be defined as crisp quantities due to limited data and information deficits. Therefore, procedures are proposed, which enable handling of these uncertain a-priori parameters next to the design variables within the optimization task. The consideration of fuzzy variables within an aggregate objective function, formulated by means of the weighted sum method is studied. The application of the approach to the optimization and robustness improvement of a passenger car tire is shown. Additionally, concepts for the identification of the set of non-dominated fuzzy quantities are proposed in the context of Pareto-optimality. The developed approaches can be coupled with a FE simulation or a response surface approximation and therefore are suitable for the solution of engineering design tasks.

Acknowledgements

The authors gratefully acknowledge the support of the German Research Foundation (DFG) within the project KA 1163/13-1 "Design of durable tires under the consideration of data uncertainty".

References

Alefeld, G. and J. Herzberger. *Introduction to Interval Computations*. Academic Press, New York, 1983.

Beer, M. and M. Liebscher. Designing Robust Structures – A Nonlinear Simulation Based Approach. *Computers and Structures*, 86:1102–1122, 2008.

Benjamin, J. R. and C. A. Cornell. *Probability, Statistics and Decision for Civil Engineers*. McGraw-Hill, 1970.

Bernardo, J. M. and A. F. M. Smith. *Bayesian Theory*. Wiley, Chichester, 1994.

Das, S., Chowdhury, S. R., Panigrahi, B. K., Pattnaik, S. and S. Das. Multi-objective Optimization with Uncertainty: Probabilistic and Fuzzy Approaches. *Proceedings of the World Congress on Nature and Biologically Inspired Computing*, Coimbatore, 2009.

Deb, K. *Multi-objective Optimization using Evolutionary Algorithms*. Wiley, Chichester, 2002.

Deb, K. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *Evolutionary computation*, 6:182–197, 2002.

Dubois, D. and H. Prade. The Three Semantics of Fuzzy Sets. *Fuzzy Sets and Systems*, 90:141–150, 1997.

Elishakoff, I. Essay on Uncertainties in Elastic and Viscoelastic Structures: form A. M. Freudenthal's Criticisms to Modern Convex Modeling. *Computers and Structures*, 56:871–895, 1995.

Farina, M. and P. Amato. A Fuzzy Definition of "Optimality" for Many-criteria Optimization Problems. *IEEE Transactions on Systems, Man. and Cybernetics – Part A: Systems and Humans*, 34:315–326, 2004.

Fonseca, C. M. and P. J. Fleming. Genetic Algorithms for Multi-objective Optimization: Formulation, Discussion and Generalization. In Forrest, S., editor, *Proceedings of the Fifth International Conference on Genetic Algorithms*, Urbana-Champaign, 1993.

Graf, W., Sickert, J.-U., Pannier, S. and M. Kaliske. In Beer, M., Muhanna, R. L., Muellen, R. L. Robust Design with Uncertain Data and Response Surface Approximation. *Proceedings of the International Workshop on Reliable Engineering Computing*, Singapore, 2010.

Kaliske, M., Zheng, D., Andre, M. and C. Bertram. Efficient Steady-State Simulations up to High Speed with Dissipative Tire Characteristics. *Vehicle Systems Dynamics Journal Supplement*, 40:175–194, 2003.

Kapitaniak, T. *Chaos for Engineers: Theory, Applications and Control*. Springer, Berlin, 2000.

Knowles, J. and D. Corne. The Pareto Archived Evolution Strategy: A New Baseline Algorithm for Multiobjective Optimisation. *Proceedings of the 1999 Congress on Evolutionary Computation*, Washington, 1999.

Kwakernaak, H. Fuzzy random variables – I. Definitions and Theorems. *Information Sciences*, 15:1–29, 1978.

Möller, B. and M. Beer. *Fuzzy Randomness – Uncertainty in Civil Engineering and Computational Mechanics*. Springer, Berlin, 2004.

Möller, B., Graf, W. and M. Beer. Fuzzy Structural Analysis Using Alpha-level Optimization. *Computational Mechanics*, 26:547–565, 2000.

Nackenhorst, U. The ALE-formulation of Bodies in Rolling Contact – Theoretical Foundations and Finite Element Approach. *Computer Methods and Applied Mechanics in Engineering*, 193:4299–4322, 2004.

Pannier, S. *Effizienter Numerischer Entwurf von Strukturen und Prozessen bei Unschärfe*. Diss., Technische Universität Dresden, 2011.

Pareto, V. *Manual of Political Economy*. Macmillan, 1971.

Srinivas, N. and K. Deb. Multi-objective Function Optimization Using Non-dominated Sorting Genetic Algorithms. *Evolutionary Computation*, 2:221–248, 1995.

Wu, D. and J. M. Mendel. Uncertainty Measures for Interval Type-2 Fuzzy Sets. *Information Science*, 177:5378–5393, 2007.

Zadeh, L. Fuzzy Sets. *Information and Control*, 8:338–353, 1965.

Dynamic Response of Beams to Interval Load

Iwona Skalna
Department of Applied Computer Science, AGH University of Science and Technology, Krakow, Poland
skalna@agh.edu.pl

M.V. Rama Rao
Department of Civil Engineering, Vasavi College of Engineering, Hyderabad, India
dr.mvrr@gmail.com

Andrzej Pownuk
Department of Mathematical Sciences, University of Texas at El Paso, 500 West University Avenue El Paso, TX 79968, USA
andrzej@pownuk.com

Abstract. Parameters of mathematical models are most often represented by real numbers, while in practice it is impossible or at least very difficult to get reliable information about their exact values. Hence, it is unreasonable to take point data for that may lead to incorrect results, which is not welcome especially when inaccuracy cannot be neglected. Depending on available information, one can use different ways of modelling of uncertainty. Interval computing plays an important role in this field, because very often the only available information are lower and upper bounds on a physical quantity. This paper focuses on a transient dynamic analysis of a beam with uncertain parameters. Finite difference and finite element methods are used to solve partial differential equation which represents the model for the motion of a straight elastic beam. In order to compute the time-history response of the beam under uncertainty, interval dynamic beam equations are solved using Search method, Gradient method, Taylor method, adaptive Taylor method, direct optimisation and Direct method for solving parametric interval linear systems. The applicability, i.e. effectiveness and accuracy, of those methods is illustrated through solution of beams with interval value of modulus of elasticity and mass density and subjected to interval dynamic loading.

Keywords: Euler-Bernoulli beam, Dynamic response, Interval arithmetic, Search method, Gradient method, Taylor method, adaptive Taylor method, Direct method, Direct optimisation.

1. Introduction

Airplane wings, high-rise buildings and suspension bridges are just some of the mechanical and structural examples where vibration analysis of beams is essential for the safe design. Safety issues are the greatest concern of structural engineering as the design and construction of secure and safe structures can prevent disasters like the collapse of Tacoma Narrow Bridge November 7, 1940, just few month after it was finished. This was probably the most dramatic failure in bridge engineering history. Safety studies in structural engineering are supposed to prevent failure during the lifetime of a structure.

Constantly increasing computational capabilities allow for detailed numerical models of structural systems. However, those models are built, inter alia, on a number of model parameters subject to uncertainty.

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

The use of models that include the uncertainty, which is central to reliability/risk analysis of engineering systems, is of great importance for a design engineer.

Uncertainty of structural parameters is mainly due to the scarcity or lack of data which may be resulted from manufacturing/construction tolerances or caused by progressive deterioration of concrete and corrosion of steel. In engineering applications, uncertainty also exists in determining external loads. To make a decision based on an inexact data say some parameter $\tilde{p}$, a measurement error $\Delta p = |\tilde{p} - p|$ must known at least. Very often, the only available information about the error is its upper bound $\Delta p \leqslant \Delta$. In this case, once the measurement $\tilde{p}$ is obtained, one can conclude that the possible values $\tilde{p} + \Delta p$ form an interval $\boldsymbol{p} = [\tilde{p} - \Delta, \tilde{p} + \Delta]$ which is guaranteed to contain the exact value p of the parameter. Once interval quantities are introduced, they must be handled appropriately to obtain the result which is guaranteed to contain the exact solution.

Though interval arithmetic was introduced by Moore (Moore, 1966) already in 1966, the application of interval concepts to structural analysis is more recent. Some important advances on reliability-based design and modelling of uncertainty when data is limited were made during last years. Structural analysis using interval variables has been used by several researchers to incorporate uncertainty into structural analysis ((Köylüoglu et al., 1995), (Nakagiri and Yoshikawa, 1996), (Rao and Sawyer, 1995), (Rao and Berke, 1997), (Rao and Chen, 1998), (Mullen and Muhanna, 2001), (Neumaier and Pownuk, 2007), (Skalna, Pownuk and Rama Rao, 2008)).

In this paper, the problem of vibrations of an Euler-Bernoulli beam with interval material properties subjected to interval load is considered. Two different approaches are employed to obtain beam deflection in time. In the first approach, the Euler-Bernoulli equation governing the behaviour of the beam is descretized in space and time. The beam bending in the respective time step is obtained by solving a system of equations with coefficient depending on interval parameters. Several methods are used for this purpose. Search method, Gradient method ((Skalna, Pownuk and Rama Rao, 2008)), Taylor method and adaptive Taylor method (Pownuk, 2011) utilise the fact that in many structural engineering problems relation between the solution and uncertain parameters is monotone. In such a case, the extreme values of a solution are attained at respective endpoints of given intervals. Monotonicity can be verified by using Taylor series or an interval method (Hansen, 1992). Methods exploiting monotonicity tests are useful for solving large scale problems, but they may underestimate. When monotonicity is not assumed, the solution can be obtained using methods for solving parametric interval linear systems (Skalna, 2010)). Those methods give guaranteed enclosures, but their usage is limited e.g. by the amount of uncertainty. In the second approach, the Finite Element Method is a starting point for considerations. The Wilson-θ method and optimisation approach are used for the solution of the problem (Rama Rao, Pownuk and Vandewalle, 2010).

The paper has the following structure. In Sections 3 and 6, the considered problem is described in terms of the mathematical theory. Section 4 describes the discretization of the problem in time and space. Section 5 and 6.1 are devoted to the methods for solving interval linear systems obtained from the discretization of the Euler-Bernoulli equation. Numerical examples are given in Section 7. The paper ends with concluding remarks.

2. Interval uncertainty

If only very limited knowledge about the value of some structural parameter p_i is available, then this value can be conveniently described by an interval number in the following way:

$$p_i \in [\tilde{p}_i - \Delta p_i, \tilde{p}_i + \Delta p_i] = [\underline{p}_i, \overline{p}_i] = \boldsymbol{p}_i, \tag{1}$$

where $\tilde{p}_i$ can be considered as an approximation of the true value of p_i and Δp_i as an approximation error.

Now, if some output quantity y is related to parameters p by a known relation $y = f(p)$, then the calculation of the result, assuming p vary within $\boldsymbol{p}$, is numerically equivalent to finding the following solution set:

$$y_S = \{y \ : \ y = f(p), p \in \boldsymbol{p}\}. \tag{2}$$

The outcome of the interval analysis here is expressed as a set y_S of possible solutions as, in general, it cannot be described exactly by an interval or hypercube. The correct interpretation of this expression is that the set y_S contains all vectors y that are obtained from applying the function f on all possible vectors p within the interval vector $\boldsymbol{p}$.

An exact description of the solution set y_S is often extremely difficult to find. Therefore, usually an interval vector $\boldsymbol{x}^* \ni y_S$, called outer solution/enclosure, is computed instead and the goal is $\boldsymbol{x}^*$ to be as narrow as possible. The tightest interval vector containing y_S is called *hull solution* (or simply a *hull*). One can also calculate inner solution/approximation which is defined as an interval vector which is included in the hull. They are usually obtained using the "straightforward" interval arithmetic. However, this usually leads to large overestimation due to the so-called *dependency problem*. Keeping track of how intermediate results on input data may decrease excess with. This idea was successfully implemented in several approaches, e.g. affine arithmetic (Comba and Stolfi, 1993).

3. Forced vibration of a beam

Forced vibration of a beam is governed by Euler-Bernoulli equation (Ciarlet, 1997).

$$\frac{\partial^2}{\partial x^2}\left(EJ\frac{\partial^2 w}{\partial x^2}\right) = q - \rho A\frac{\partial^2 w}{\partial t^2}, \tag{3}$$

where E is the elastic modulus, J is the second area moment, A is the cross-sectional area, ρ is mass density of the material of the beam and q is an external load. The model (3) where the displacement w depends only on one-dimensional spatial variable x and time t is obtained upon the use of Hookes law and other simplifying assumptions. This model is a valid approximation for thin beams under small transverse deformations. As a good rule-of-thumb, 'small' is defined as deflections that are at least ten times smaller than beam thickness.

For an uniform beam (EJ is constant), Eq. (3) reduces to

$$EJ\frac{\partial^4 w}{\partial x^4} = q - \rho A\frac{\partial^2 w}{\partial t^2}. \tag{4}$$

Because vibration is an initial-boundary value problem, therefore both initial and boundary conditions are required to obtain a unique solution $w(x,t)$. Since the equation involves second order derivative with respect

to time and fourth derivative with respect to a space coordinate, thus four boundary and two initial conditions are necessary:

$$\begin{cases} w(0,t)=0 \\ w(L,t)=0 \\ \dfrac{\partial^2 w}{\partial x^2}(0,t)=0 \\ \dfrac{\partial^2 w}{\partial x^2}(L,t)=0 \end{cases}, \quad \begin{cases} w(x,0)=w_0(x) \\ v(x,0)=\frac{\partial w}{\partial t}(x,0)=v_0(x) \end{cases} \tag{5}$$

Endpoint displacements are equal to zero, which can be written as $w(0,t) = w(L,t) = 0$ for $t \in [0,T]$. Because bending moments at both endpoints are equal to zero, therefore $M(0,t) = EJ\dfrac{\partial^2 w}{\partial x^2}(0,t)$ and $M(L,t) = EJ\dfrac{\partial^2 w}{\partial x^2}(L,t)$, and consequently $\dfrac{\partial^2 w}{\partial x^2}(0,t) = \dfrac{\partial^2 w}{\partial x^2}(L,t) = 0$ for $t \in [0,T]$. For $t = 0$, both displacement and velocity are equal to zero and thus $w_0(x) = 0$, $v_0(x) = 0$ for $x \in [0,L]$.

4. Implicit Finite Difference Discretization

In this paper implicit Finite Difference Method has been applied to the problem of dynamic beam vibrations (Ciarlet, 1990). Discretization of Eq. (4) is performed at point $(i, j+1)$:

$$\left(EJ\frac{\partial^4 w}{\partial x^4}\right)_{i,j+1} = q_{i,j+1} - \left(\rho A\frac{\partial^2 w}{\partial t^2}\right)_{i,j+1} \tag{6}$$

which leads to the finite difference equation:

$$\begin{aligned} &E_{i,j+1}J_{i,j+1}\frac{w_{i+2,j+1}-4w_{i+1,j+1}+6w_{i,j+1}-4w_{i-1,j+1}+w_{i-2,j+1}}{\Delta x^4}+\frac{\rho_{i,j+1}A_{i,j+1}}{\Delta t^2}w_{i,j+1}= \\ &=q_{i,j+1}-\rho_{i,j+1}A_{i,j+1}\frac{2w_{i,j}-w_{i,j-1}}{\Delta t^2} \end{aligned} \tag{7}$$

Similarly, it is possible to discretize initial and boundary conditions. Finally, one obtains:

$$\begin{cases} w_{0,j+1}=0 \\ w_{0,j+1}-2w_{1,j+1}+w_{2,j+1}=0 \\ E_{i,j+1}J_{i,j+1}\dfrac{w_{i+2,j+1}-4w_{i+1,j+1}+6w_{i,j+1}-4w_{i-1,j+1}+w_{i-2,j+1}}{\Delta x^4}+\dfrac{\rho_{i,j+1}A_{i,j+1}}{\Delta t^2}w_{i,j+1}= \\ =q_{i,j+1}-\rho_{i,j+1}A_{i,j+1}\dfrac{2w_{i,j}-w_{i,j-1}}{\Delta t^2} \\ w_{n-2,j+1}-2w_{n-1,j+1}+w_{n,j+1}=0 \\ w_{n,j+1}=0 \\ w_{i,0}=w_i^* \\ w_{i,1}=w_{i,0}+v_i^*\Delta t. \end{cases} \tag{8}$$

It is important to note that $w_{i,j+1} = w(p_{i,j+1})$ where $p_{i,j+1} = (E_{i,j+1}, \rho_{i,j+1}, q_{i,j+1})$. Discretization reduces the problem of computing the dynamic response of a beam to the problem of solving a sequence of

parametric linear systems. Assuming the uncertainty of the parameters, an sequence of parametric interval linear systems must be solved. In order to increase the accuracy of the FDM, finite difference scheme of order higher than 3 has been applied for the time step.

$$\left(\frac{\partial^2 w}{\partial t^2}\right)_{i,j} \approx \frac{2w_{i,j} - 5w_{i,j-1} + 4w_{i,j-2} - w_{i,j-3}}{\Delta t^2} \tag{9}$$

5. Methods for solving parametric linear systems

Apart from the diversity caused by the nature of the numerical problem at hand, a clear distinction can be made between fundamental approaches for tackling the interval uncertainty. The interval arithmetic strategy approaches the exact hypercubic circumscription of the interval result from outside. It is based on the calculation of guaranteed outer bounds. The global optimisation approach on the other hand calculates an inner approximation. The interval arithmetic based methods proves to be computationally less expensive than the approximate method, it very often results in a huge overestimation of the actual interval result, due to the dependency problem. On the other hand, optimisation based approaches, though computationally expensive and time-consuming, provide an acceptable solution for practical engineering problems.

5.1. Interval solution as a function of uncertain parameters

Each interval solution is in fact a function of some specific combinations of the parameters:

$$\underline{w}_{i,j} = w_{i,j}(p_{i,j}^{\min}), \qquad \overline{w}_{i,j} = w_{i,j}(p_{i,j}^{\max}). \tag{10}$$

In the continuous case, one can write

$$\begin{aligned} \underline{w}(x,t) &= w(x,t,p^{\min}(x,t)), \\ \overline{w}(x,t) &= w(x,t,p^{\max}(x,t)). \end{aligned} \tag{11}$$

In some situations, the interval solution depends only on one combination of parameters for some domain $D_\alpha \subseteq [0,L] \times [0,T]$

$$\underline{w}(x,t) = w(x,t,p_\alpha^{\min}), \qquad \overline{w}(x,t) = w(x,t,p_\alpha^{\max}). \tag{12}$$

In such cases it is possible to calculate the interval solution exactly by using finite number of combinations of the parameters $p_1^{\min}$, $p_1^{\max}$,...,$p_\alpha^{\min}$, $p_\alpha^{\max}$,...,$p_q^{\min}$, $p_q^{\max}$ where $D_1 \cup ... \cup D_q = [0,L] \times [0,T]$.

5.1.1. *Search method*

In order to find the interval solution the Search method is applied. The method relies on solving parametric linear systems of equations corresponding to the specific combinations of the parameters. That is, each interval parameter p_i is replaced by the set of discrete points $p_{i1}, \ldots, p_{ik}$:

$$\boldsymbol{p}_i \approx \{p_{i1}, \ldots, p_{ik}\}. \tag{13}$$

A multidimensional interval $\boldsymbol{p} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_m]$ is approximated by the discrete set of points:

$$\boldsymbol{p} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_m] \approx \{(p_{1,i_1}, \ldots, p_{m,i_k}) : 0 \leqslant i_1, \ldots, i_k \leqslant k\} = \mathcal{P}_{m,k}. \tag{14}$$

Number of elements in the set (14) equals k^m where m is the number of interval parameters and k is the number of intermediate points in each interval p_i (for $k = 2$ the method reduces to the endpoints combination method). Then, an interval solution can be calculated in the following way:

$$\underline{w}_{i,j} \approx \underline{w}_{i,j}^{search} = \min\{w_{i,j}(p_{1,i_1}, \ldots, p_{m,i_k}) : (p_{1,i_1}, \ldots, p_{m,i_k}) \in \mathcal{P}_{m,k}\}, \tag{15}$$

$$\overline{w}_{i,j} \approx \overline{w}_{i,j}^{search} = \max\{w_{i,j}(p_{1,i_1}, \ldots, p_{m,i_k}) : (p_{1,i_1}, \ldots, p_{m,i_k}) \in \mathcal{P}_{m,k}\}. \tag{16}$$

The search method allows as well finding approximate values of $p_{i,j}^{\min,search}, p_{i,j}^{\max,search}$.

$$\underline{w}_{i,j} \approx \underline{w}_{i,j}^{search} = w_{i,j}(p_{i,j}^{\min,search}), \overline{w}_{i,j} \approx \overline{w}_{i,j}^{search} = w_{i,j}(p_{i,j}^{\max,search}) \tag{17}$$

According to numerical experiments

$$p_{i,j}^{\min} \approx p_{i,j}^{\min,search}, p_{i,j}^{\max} \approx p_{i,j}^{\max,search} \tag{18}$$

which means that the Search method can find approximate or exact values of $p_{i,j}^{\min}$ and $p_{i,j}^{\max}$.

5.1.2. *Gradient method*

The value of $p^{\min}(x,t)$ and $p^{\max}(x,t)$ can be found as well by solving respectively the following minimisation and maximisation problems

$$\begin{aligned}
&p^{\min}(x,t) = \arg\min_p w(x,t,p),\\
&p^{\max}(x,t) = \arg\max_p w(x,t,p),\\
&\text{s.t.} \begin{cases}
EJ\dfrac{\partial^4 w}{\partial x^4} = q - \rho A\dfrac{\partial^2 w}{\partial t^2}\\
w(0,t) = 0\\
w(L,t) = 0\\
\dfrac{\partial^2 w}{\partial x^2}(0,t) = 0\\
\dfrac{\partial^2 w}{\partial x^2}(L,t) = 0\\
w(x,0) = w_0(x)\\
v(x,0) = \dfrac{\partial w}{\partial t}(x,0) = v_0(x)\\
p \in \boldsymbol{p}
\end{cases}
\end{aligned} \tag{19}$$

Solutions $w_{i,j}$ are functions of uncertain parameters $w_{i,j} = w_{i,j}(p)$. If the function $w_{i,j} = w_{i,j}(p)$ is monotone, then $p_{i,j}^{\min}$ and $p_{i,j}^{\max}$ can be calculated as:

$$p_{i,j,k}^{\min,gradient} = \underline{p}_k, \;\; if \;\; \frac{\partial w_{i,j}}{\partial p_k} < 0 \;\; else \;\; p_{i,j,k}^{\min,gradient} = \overline{p}_k, \tag{20}$$

$$p_{i,j,k}^{\max,gradient} = \overline{p}_k, \;\; if \;\; \frac{\partial w_{i,j}}{\partial p_k} \geq 0 \;\; else \;\; p_{i,j,k}^{\max,gradient} = \underline{p}_k, \tag{21}$$

where $\frac{\partial w_{i,j}}{\partial p_k}$ are partial derivatives with respect to all uncertain parameters. Thus obtained combinations of endpoints can be utilised for calculation of upper and lower bounds of the solution

$$\underline{w}_{i,j} \approx \underline{w}_{i,j}^{gradient} = w_{i,j}(p_{i,j}^{\min,gradient}), \qquad \overline{w}_{i,j} \approx \overline{w}_{i,j}^{gradient} = w_{i,j}(p_{i,j}^{\max,gradient}). \tag{22}$$

5.1.3. *Taylor method*

The interval solution can be also calculated by using first order Taylor model:

$$w_{i,j}(p) \approx T_{i,j}^{(1)}(p) = w_{i,j}(p_0) + \sum_k \frac{\partial w_{i,j}}{\partial p_k}(p_0) \cdot (p_k - p_0). \tag{23}$$

In this approach the interval solution can be calculated in the following way

$$\underline{w}_{i,j} \approx \underline{w}_{i,j}^{Taylor} = w_{i,j}(p_0) - \sum_k \left| \frac{\partial w_{i,j}}{\partial p_k}(p_0) \right| \cdot \Delta p_k, \tag{24}$$

$$\overline{w}_{i,j} \approx \overline{w}_{i,j}^{Taylor} = w_{i,j}(p_0) + \sum_k \left| \frac{\partial w_{i,j}}{\partial p_k}(p_0) \right| \cdot \Delta p_k, \tag{25}$$

The result of the Taylor method can be calculated as well by using endpoint combinations and Taylor polynomial

$$\underline{w}_{i,j}^{Taylor} = T_{i,j}^{(1)}(p_{i,j}^{\min,gradient}), \quad \overline{w}_{i,j}^{Taylor} = T_{i,j}^{(1)}(p_{i,j}^{\max,gradient}). \tag{26}$$

5.1.4. *Adaptive Taylor approximation*

It is possible to increase the accuracy of the Taylor method results by using *adaptive approximation* (Pownuk, 2011). It is necessary to calculate all different combinations of parameters $L_1 = \{p^{*,1}, ..., p^{*,n_1}\}$ in the sets $p_{i,j}^{\min,gradient}$ and $p_{i,j}^{\max,gradient}$. For each combination $p^{(*,k)}$ from the list L_1 it is necessary to find a point solution $w^{(k)} = w(p^{*,k})$.

$$\underline{w}_{i,j}^{(1)} \approx min\{w_{i,j}^{(1)}, ..., w_{i,j}^{(n_1)}\}, \quad \overline{w}_{i,j}^{(1)} \approx max\{w_{i,j}^{(1)}, ..., w_{i,j}^{(n_1)}\} \tag{27}$$

For $\underline{w}_{i,j}^{(1)}$ and $\overline{w}_{i,j}^{(1)}$ it is necessary to calculate new values of $p_{i,j}^{\min,gradient,1}$ and $p_{i,j}^{\max,gradient,1}$. In the sets $p_{i,j}^{\min,gradient,1}$ and $p_{i,j}^{\max,gradient,1}$ it is necessary to find new combinations of parameters and add to the list L. New list will be denoted as L_2 and calculate new values of upper and lower bound from the formula (27). Calculations will be stopped if no new combinations of parameters will be found in the next iteration i.e. $L_i = L_{i+1}$.

5.2. DIRECT METHOD

To verify the results obtained using the approximate methods described in the previous sections, the direct method (DM) (Skalna, 2010) for solving parametric interval linear systems is applied to the problem.

5.2.1. *Iterative approach*
The system (8) can be written in matrix form as:

$$Kw^{j+1} = Q^j + M_2 w^j + M_1 w^{j-1},\ (j = 2, \ldots, n_t) \tag{28}$$

with

$$Q^j = \begin{bmatrix} 0 \\ 0 \\ q_{2,j+1} \\ q_{3,j+1} \\ \vdots \\ q_{n-2,j+1} \\ 0 \\ 0 \end{bmatrix},\ M_1 = \mu_2 \begin{bmatrix} 0 & 0 & & & & & 0 & 0 \\ 0 & 0 & & & & & 0 & 0 \\ & & -1 & & & & & \\ & & & \ddots & & & & \\ & & & & -1 & & & \\ 0 & 0 & & & & 0 & & \\ 0 & 0 & & & & & 0 & \end{bmatrix},\ M_2 = \mu_2 \begin{bmatrix} 0 & 0 & & & & & 0 & 0 \\ 0 & 0 & & & & & 0 & 0 \\ & & 2 & & & & & \\ & & & \ddots & & & & \\ & & & & 2 & & & \\ 0 & 0 & & & & 0 & & \\ 0 & 0 & & & & & 0 & \end{bmatrix}, \tag{29}$$

and

$$K = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & ... & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 0 & ... & 0 & 0 & 0 & 0 & 0 \\ \mu_1 & -4\mu_1 & 6\mu_1+\mu_2 & -4\mu_1 & \mu_1 & 0 & 0 & ... & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu_1 & -4\mu_1 & 6\mu_1+\mu_2 & -4\mu_1 & \mu_1 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \mu_1 & -4\mu_1 & 6\mu_1+\mu_2 & -4\mu_1 & \mu_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{30}$$

where $\mu_1 = \frac{EJ}{\Delta x^4}$, $\mu_2 = \frac{\rho A}{\Delta t^2}$. As can be seen, the coefficients of the system (28) are polynomial functions of the parameters E, ρ and q. Thus, the Direct Method can be applied. The interval result obtained in a given iteration step enters into the right-hand of the system to be solved in the next step. The iterative approach is quite efficient, however the overestimation of the result grows in successive iteration steps. This is mainly due to the so-called *dependency problem*. Namely, starting from the second iteration, the dependency in the right hand vector is lost since the solution enters the right-hand as an interval vector and not the affine one. Therefore, in what follows a non-iterative approach is considered.

5.2.2. *Non-iterative formulation*
Consider the linear system (28). Taking into account boundary conditions, one obtains:

$$\begin{cases} w^0 = 0 \\ w^1 = 0 \\ -M_1 w^0 - M_2 w^1 + K w^2 = Q^1 \\ \ldots \\ -M_1 w^{n_t-2} - M_2 w^{n_t-1} + K w^{n_t} = Q^{n_t-1} \end{cases}, \tag{31}$$

where $w^j = [w_{0,j}, w_{1,j}, \ldots, w_{n_x,j}]$. Finally, the following parametric linear system is obtained:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 \\ -M_1 & -M_2 & K & 0 & \ldots & 0 \\ 0 & -M_1 & -M_2 & K & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & -M_1 & -M_2 & K \end{bmatrix} \begin{bmatrix} w^0 \\ w^1 \\ w^2 \\ w^3 \\ \vdots \\ w^{n_t} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ Q^1 \\ Q^2 \\ \vdots \\ Q^{n_t-1} \end{bmatrix} \tag{32}$$

The non-iterative approach is free from the accumulation error problem described above, but suffers from the efficiency problem as a very large system (of size $(n_x + 1)n_t$) must be solved.

6. Wilson-θ method

Consider a discrete structural system with multi-degree of freedom (MDOF) described by equation

$$M\ddot{w} + C\dot{w} + Kw = F(t) \tag{33}$$

The damping matrix C is defined as

$$C = \alpha_0 M + \alpha_1 K, \tag{34}$$

and the coefficients α_0 and α_1 are computed by considering damping ratios ξ_1 and ξ_2 in the first two modes of vibration (with corresponding frequencies ω_1 and ω_2) as follows:

$$\begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \frac{1}{\omega_1} & \omega_1 \\ \frac{1}{\omega_2} & \omega_2 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \tag{35}$$

Wilson-θ method is used for the solution of the transient dynamic problem. This method is an implicit integration method and involves computation of dynamic response of a MDOF system by adopting a step by step integration process in the time domain. The Wilson-θ method assumes a linear variation of acceleration over the time interval $[t, t + \theta\delta t]$, where $\theta \geq 1.0$ and δt is a small time step. It has been shown by Wilson that the method becomes unconditionally stable for $\theta \geq 1.38$.

6.1. Optimisation Approach

Uncertainty is considered in the values of Young's modulus and mass density of steel and load. The solution to the resulting interval MDOF system is obtained by an optimisation procedure. This is done by utilising the $fmincon$ function from the optimisation toolbox of MATLAB (The Mathworks, 2011) which seeks to find the minimum of a constrained non-linear multivariate function. The $fmincon$ function finds a constrained minimum of a function $f(x)$ of several variables by solving a problem of the form:

$$\{x\} = fmincon(objfun, \{x_0\}, [A], [B], [A_{eq}], \{b_{eq}\}, \{lb\}, \{ub\}) \tag{36}$$

subject to the inequality constraints

$$Ax \leqslant b, \tag{37}$$

equality constraints

$$A_{eq}x_0 = b_{eq} \tag{38}$$

and bounds

$$lb \leq x \leq ub \tag{39}$$

with x_0 being the starting point for search. The last condition (39) defines a set of lower and upper bounds on the design variables $\{x\}$, so that a solution is found in the range $lb \leq x \leq ub$. Wilson-θ method is extended to compute the interval displacement response as a function of time by formulating it as a MATLAB function $objfun$ yielding a single output describing the deterministic transient vertical displacement of a given node of the structure. The displacement response is optimised and the bounds for the displacement response are obtained at each time step δt for $0 \leq t \leq t_{max}$. The normalised uncertainties associated with mass and stiffness and load terms are represented by normalised interval parameters $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ and $\boldsymbol{p}_3$ respectively. These upper and lower bounds of these interval parameters form the vertices of an uncertainty hypercube $\boldsymbol{p}^I$. Any point (p_1, p_2, p_3) inside this bounds $\{lb\}$ and $\{ub\}$ described in equation (39) are defined as

$$lb = \begin{bmatrix} \underline{p}_1 \\ \underline{p}_2 \\ \underline{p}_3 \end{bmatrix} \text{ and } ub = \begin{bmatrix} \overline{p}_1 \\ \overline{p}_2 \\ \overline{p}_3 \end{bmatrix} \tag{40}$$

Equation (33) is recast in interval parametric form as

$$\boldsymbol{p}_1\, M\ddot{w} + Cw + \boldsymbol{p}_2 Kw = \boldsymbol{p}_3 F(t) \tag{41}$$

where parameters $\boldsymbol{p}_i$ are defined as

$$\boldsymbol{p}_i = [\underline{p}_i, \overline{p}_i],\ (i = 1, 2, 3). \tag{42}$$

The objective function can be computed at any point p defined by coordinates (p_1, p_2, p_3) within the hypercube $\boldsymbol{p}^I$ that forms the search domain. Thus, using the procedure described above, the deterministic algorithm is translated to an interval algorithm using the global optimisation based approach. In this approach, the lower and upper bounds of interval displacement $\boldsymbol{w}_n$ at a given node n is determined, taking into account that the uncertain parameters p can vary within their intervals $\boldsymbol{p}^I$. This interval $\boldsymbol{w}_n$ of this displacement is determined by a minimisation and a maximisation over the uncertainty interval $\boldsymbol{p}^I$.

$$\boldsymbol{w}_n = \left[\min_{p \in \boldsymbol{p}^I}(\boldsymbol{w}_n), \max_{p \in \boldsymbol{p}^I}(\boldsymbol{w}_n)\right] \tag{43}$$

This is done by computing the displacement $\{w(x,t)\}$ at any location x along the span of the beam at a given time t, using the following deterministic matrix equation,by implementing Wilson-θ approach. :

$$p_1 M\ddot{w} + Cw + p_2 Kw = p_3 F(t). \tag{44}$$

The time history of the displacement response is obtained by computing the minimum and maximum values of the response at each time step. To compute the displacement at a certain time, it has to be computed at all earlier time steps too. However, in order to reduce the computational cost of the optimisations using the local optimisation algorithm $fmincon$, all function evaluations of the objective function are stored in a database

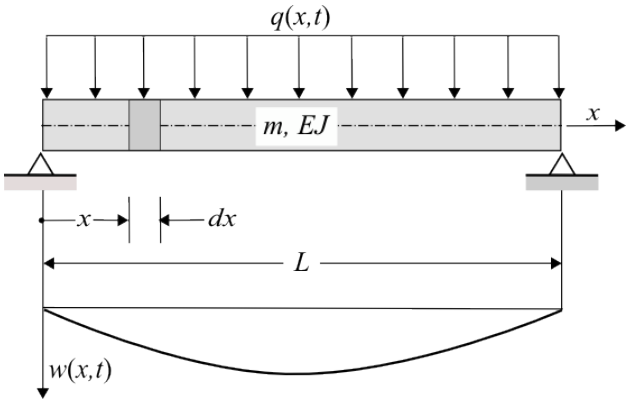


Figure 1. Geometry of the uniform beam with symmetrical load.

to enable the optimiser to reuse them for future optimisations without performing the same finite element analysis again. This database is also used to start optimisations from the point with the best function value found so far the lowest function value for minimisation and the highest function value for maximisation. All optimisations are performed from the highest to the lowest membership level.

7. Numerical experiments

In order to show the interval solutions obtained using the methods described in the paper, an example of the dynamically loaded beam with the load uniformly distributed over the entire span will be considered. Different cases of the amount of uncertainty are investigated.

Example 1. Consider Euler Bernoulli beam shown in Figure 1 with uniform load of 2.5kN applied for a short time of 0.009s. The beam has a span $L = 4$m, area of cross section $A = 0.01m^2$, second moment of area $J = 8.333 \times 10^{-6}\ m^4$ and Young's modulus $E = 200$GPa. It is assumed that mass density is uncertain $\pm 0.5\%$ and the load is uncertain $\pm 20\%$. This gives 2 interval parameters $p = (p_1, p_2) = (\rho, q)$.

The following discretization is applied $n_x = 20$, $(\Delta x = L/n_x)$, number of time steps is equal to $n_t = 100$ and time step is $\Delta t = 0.0015$s. The load is applied for 0.009s.

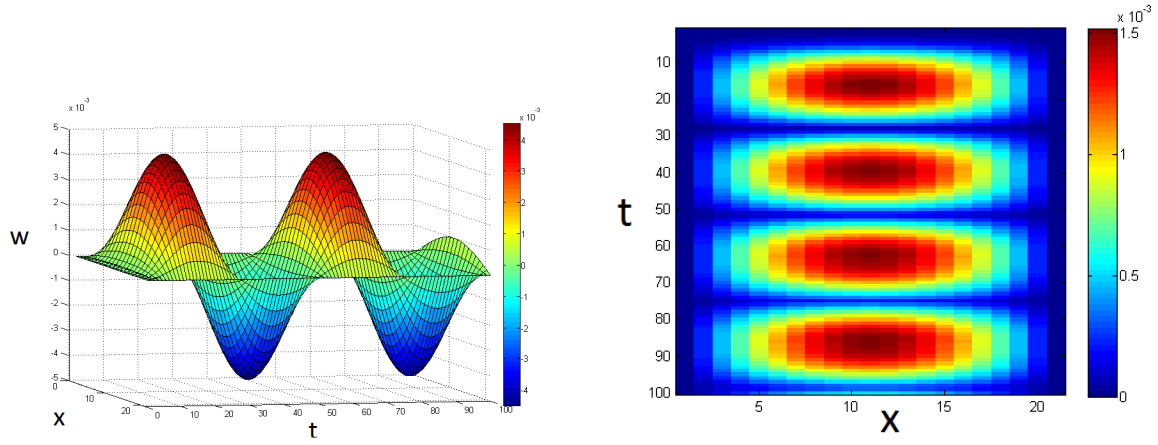
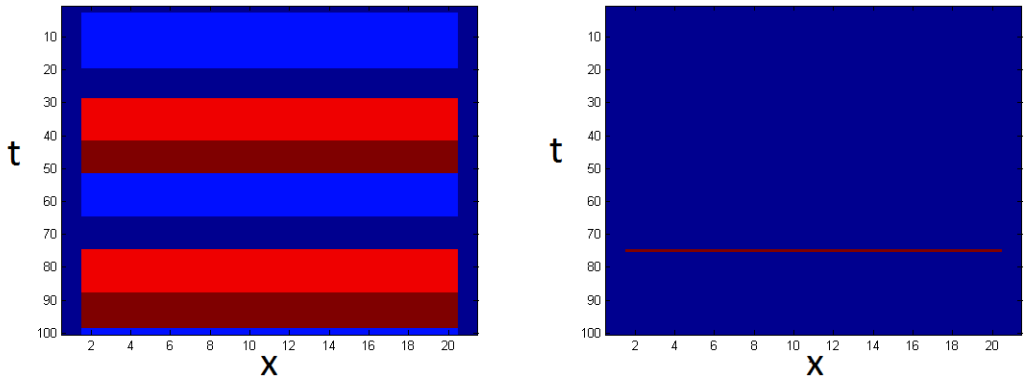


Figure 2. Search method result for $m = 2$, $k = 7$: (a) the lower and upper bound of the interval solution, (b) the difference between upper and lower bound, $\overline{w}^{search} - \underline{w}^{search}$.



(a) (b)

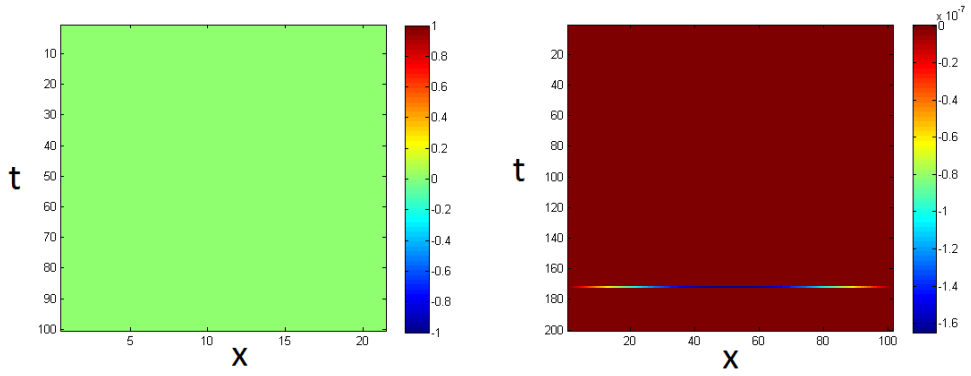
Figure 3. Combinations of parameters which correspond to $p^{\min}$: (a) Search Method, (b) comparison of the Search Method and the Gradient Method. The colours which represent each particular combination for the Search Method and the Gradient Method are different

From the Fig. 3 it is possible to see that the parameters which are calculated by using the Search Method and the Gradient Method are very similar. Only for one time step combinations of parameters were predicted incorrectly by the Gradient Method.

$$p_{i,j}^{\min,gradient} \approx p_{i,j}^{\min,search}. \tag{45}$$

The interval solution depends mostly on the endpoints of the intervals. In order to show this, $p_{i,j}^{\min,search}$, $p_{i,j}^{\max,search}$ is calculated with 2, 5 and 7 intermediate points in the given intervals. For $n_x = 20$, $n_t = 100$ the results are identical:

$$\underline{w}_{i,j}^{search,2} = \underline{w}_{i,j}^{search,5} = \underline{w}_{i,j}^{search,7}. \tag{46}$$



(a) (b)

Figure 4. Difference between the results of the search method $\underline{w}_{i,j}^{search,7} - \underline{w}_{i,j}^{search,2}$ for $k = 2$ and $k = 5$ for different discretization of the problem. $n_x = 20$, $n_t = 100$ (a), $n_x = 100$, $n_t = 200$ (b).

Of course, it is possible to find examples in which there are some difference between the solution for $k = 2$ and $k > 2$. However, according to numerical experiments for the equation which is discussed in this paper $\underline{w}_{i,j}^{search,2} \approx \underline{w}_{i,j}^{search,k}$ and $\overline{w}_{i,j}^{search,2} \approx \overline{w}_{i,j}^{search,k}$ where $k > 2$.

Figure 5 compares the results of the Direct Method and the Search Method. The solution for point data (solid black line) is presented as well.

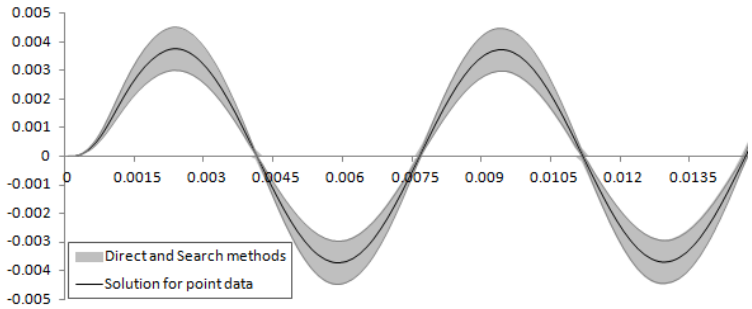


Figure 5. Vertical displacement of the midspan; comparison of the results of Direct method and Search method for the case: $E = 200$[GPa], $\rho = 7850$[kg/m^2]$\pm 0.5\%$, $q = 2.5$[kN]$\pm 20\%$, $n_x = 20$, $n_t = 100$, $\Delta t = 0.0015$.

As can be seen, the results of Direct method and Search method coincide. This proves the quality of the results of both methods.

Example 2. A beam similar to the one used in Example 1 is considered for analysis once again. The beam is acted upon by a load of 5kN/m uniformly distributed over the whole span suddenly for a duration of 0.4 seconds. Five percent damping is considered to be present. The transient dynamic response of the beam is computed using the procedure outlined in section 6.1 and results are presented . Figure 6 shows the time history plot of vertical displacement of the beam at mid-span corresponding to the case with $\boldsymbol{p}_1 = \boldsymbol{p}_2 = [0.95, 1.05]$and $\boldsymbol{p}_3 = [1.0, 1.0]$. This corresponds to deterministic load and interval values of stiffness and

mass matrices. Figure 7 shows the time history plot of vertical displacement corresponding to the case $\boldsymbol{p}_1 = \boldsymbol{p}_2 = [1.0, 1.0]$ and $\boldsymbol{p}_3 = [0.8, 1.2]$. Figure 6 clearly depicts the shifting of peaks of response and increase of uncertainty of response as time progresses as uncertainty in mass and stiffness causes a large uncertainty in the eigenfrequencies of the structure. However, no such shifting of peaks is observed in Figure 7 because mass and stiffness properties are deterministic and eigenfrequencies remain deterministic even as time progresses.

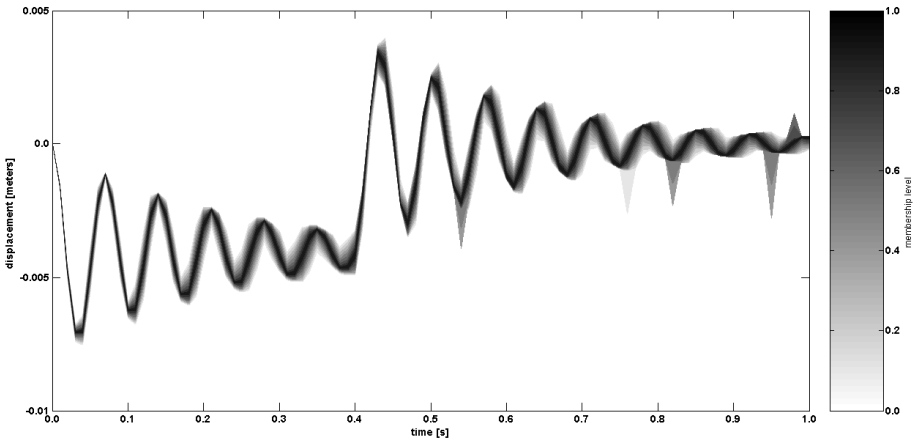


Figure 6. Vertical displacement of the midspan with $\boldsymbol{p}_1 = \boldsymbol{p}_2 = [0.95, 1.05]$, $\mathbf{p}_3 = [1.0, 1.0]$ and 5 percent damping.

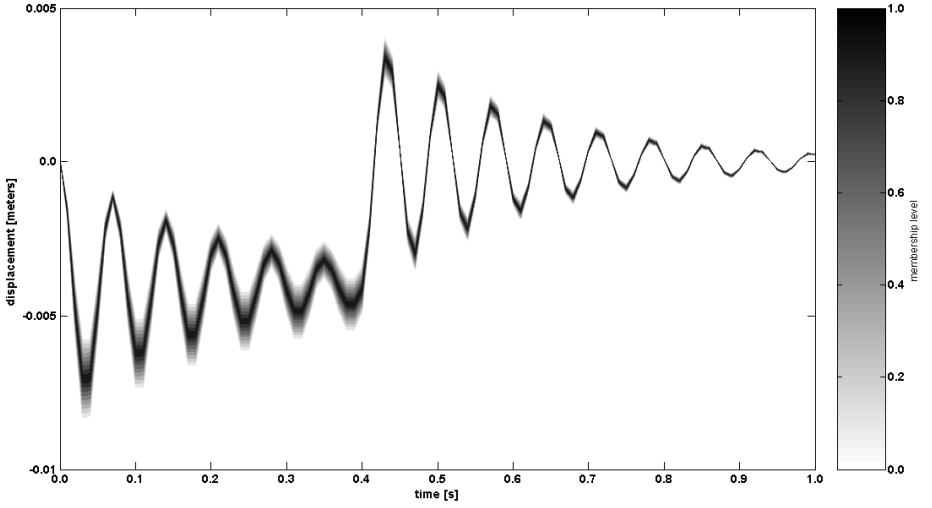


Figure 7. Vertical displacement of midspan with $\boldsymbol{p}_1 = \boldsymbol{p}_2 = [1.0, 1.0]$, $\boldsymbol{p}_3 = [0.8, 1.2]$ and 5 percent damping

8. Conclusions

Several methods for solving beam vibrations problem under interval uncertainty were considered. Based on illustrative examples, it can be stated that in the case of symmetrically loaded beam, the interval solution depends on very few combinations of uncertain parameters. In the example, the interval solution depends only on 4 combinations of parameters. This means that in practice it is possible to find four point solutions in order to compute the exact values of the interval solution. Appropriate combinations of parameters can be predicted by using the gradient of the solution. Moreover, it is possible to increase the accuracy of the calculations by using adaptive approximation (Pownuk, 2011) which will be a topic of future research. In more complex cases, as seen in Example 1, it is possible to find large areas in which the interval solution depends only on specific combinations of parameters. It is possible to use this information in order to improve accuracy of the interval solution. There are also situations in which the interval solution depends on infinite number of combinations of the parameters. According to numerical results, in considered example, the solution depends only on the endpoints of the parameters. In such situations it is possible to calculate the exact solution by using the gradient method (Pownuk, 2004). In this case it is possible to calculate the interval solution approximately using Taylor method, which is especially useful for narrow intervals. The guaranteed solution can be obtained using the Direct method for solving parametric interval linear systems, however the method requires some improvement do deal with large scale problems. Direct formulation of the iterative problems can eliminate the wrapping effects from the interval calculations. Optimisation approach is time consuming but produces acceptable results even with large intervals of input parameters.

References

Ciarlet P.G. Finite Difference Methods: Solution of Equations in R^n Elsevier, 1990.

Ciarlet P.G. Mathematical Elasticity, Volume II. Theory of Plates, North Holland, 1997.

Hansen, E.R. Global Optimization Using Interval Analysis Marcel Dekker, New York, 1992, ISBN 0824786963.

Köylüoglu, U. and Cakmak, S. and Ahmet, N. and Soren, R.K. Interval Algebra to Deal with Pattern Loading and Structural Uncertainty. *Journal of Engineering Mechanics*, 121(11): 1149-1157, 1995.

The Mathworks Optimization toolbox(R2011b) users guide. "http://www.mathworks.in/help/toolbox/optim/ug/bqar5w2.html"

Moore, R.E. Interval Analysis. Englewood Cliffs, N. J. Prentice-Hall, Inc., 1966.

Muhanna, R. L. and Mullen, R. L. Uncertainty in Mechanics ProblemsInterval-Based Approach *J. Engrg. Mech.* 127(6), 557-566, 2001.

Nakagiri, S. and Yoshikawa, N. Finite Element Interval Estimation by Convex Model. In *Proceedings of 7th ASCE EMD/STD Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability*, WPI, MA, 79 August 79, 1996.

Neumaier, A. and Pownuk, A. Linear Systems with Large Uncertainties, with Applications to Truss Structures. *Reliable Computing*, 13(2), 149-172, 2007.

Pownuk, A. Solution of the Interval Equations of Dynamics by Using Adaptive Approximation. In *Proceedings of NAFIPS 2011*, El Paso Texas, March 2011.

Pownuk A., Numerical solutions of fuzzy partial differential equation and its application in computational mechanics, FuzzyPartial Differential Equations and Relational Equations: Reservoir Characterization and Modeling (M. Nikravesh, L. Zadeh and V. Korotkikh, eds.), Studies in Fuzziness and Soft Computing, Physica-Verlag, 2004, pp.308-347

Rama Rao, M.V. and Pownuk, A. and Vandewalle, S. *A fuzzy finite element approach to computing the uncertain dynamic response of a simply-supported thin rectangular plate subjected to an impact load.* International Congress on Computational and Applied Mathematics(ICCAM-2010), Leuven, Belgium, July 2010.

Rao, S.S. and Sawyer, P. Fuzzy finite element approach for analysis of imprecisely defined systems. *AIAA J.*, 33(12), 2364-2370, 1995.

Rao, S.S. and Berke, L. Analysis of uncertain structural systems using interval analysis. *AIAA J.*, 35(4), 727-735, 1997.

Rao, S.S. and Chen, Li. Numerical solution of fuzzy linear equations in engineering analysis. *International Journal of Numerical Methods in Engineering.*, 43, 391-408, 1998.

Skalna, I. and Pownuk, A. On using global optimization method for approximating interval hull solution of parametric linear systems. In *Proceedings of NSF workshop on Reliable Engineering Computing(REC 2008)*, Savannah, Georgia, USA, February 2008.

Skalna, I. Direct method for solving parametric interval linear systems with non-affine dependencies. *Lecture Notes in Computer Science 6068*, 485-494, 2010.

Skalna, I. and Rama Rao, M.V. and Pownuk, A. Systems of fuzzy equations in structural mechanics. *Journal of Computational and Applied Mathematics*, 218(1) :149-156, 2008.

Comba J.L.D. and Stolfi J. Affine arithmetic and its applications to computer graphics. Proc. SIBGRAPI'93 VI Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens (Recife, BR), 9-18, 1993.

Terror, Security, and Money: Balancing the Risks, Benefits, and Costs of Critical Infrastructure Protection

Mark G Stewart
Australian Research Council Professorial Fellow
Centre for Infrastructure Performance and Reliability
The University of Newcastle, Australia
email: mark.stewart@newcastle.edu.au

John Mueller
Mershon Center for International Security Studies, Ohio State University
Cato Senior Fellow, Cato Institute, Washington, DC
email: bbbb@osu.edu

Abstract: The loading and response of structures to explosive blast loading is subject to uncertainty and variability. This uncertainty can be caused by variability of dimensions and material properties, model errors, environment, etc. Limit state and LRFD design codes for reinforced concrete and steel have been derived from probabilistic and structural reliability methods to ensure that new and existing structures satisfy an acceptable level of risk. These techniques can be applied to the area of structural response of structures subject to explosive blast loading. The use of decision theory to determine acceptability of risk is crucial to prioritise protective measures for built infrastructure. Government spending on homeland security will reach $141.6 billion worldwide in 2009 and is projected to reach $300 billion by 2016. The question is, is this or other expenditure necessary? Clearly, scientific rigour is needed when assessing the effectiveness and the need for protective measures to ensure that their benefits exceed the cost. The paper will assess terrorist threats to buildings and airport infrastructure and the cost-effectiveness of protective and counter-terrorism measures. Structural reliability and probabilistic methods are used to assess risk reduction due to protective measures. The key innovation is incorporating uncertainty modelling in the decision analysis, which in this case will maximise net benefit. This analysis will then consider threat likelihood, cost of security measures, risk reduction and expected losses to compare the costs and benefits of security measures to decide which security measures are cost-effective, and those which are not.

Keywords: risk, reliability, terrorism, security, cost-benefit analysis, infrastructure, aviation

1. Introduction

Terrorist threats against civilian and military infrastructure, particularly buildings, bridges, pipelines and aviation infrastructure, seem to be increasing, as evidenced by recent terrorist attacks including Manchester and London city centres (1992, 1993 and 1996), U.S. Embassy in Kenya (1998), Pentagon and World Trade Center (2001), night clubs and restaurants in Bali (2002, 2005), Marriott Hotel in Jakarta (2003), Australian Embassy in Indonesia (2004), and 'near misses' such as the recent Christmas Day Northwest Airlines

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

aircraft suicide bombing attempt (2009). The preferred method of attack is Improvised Explosive Devices (IEDs), often through suicide tactics, against buildings and transport infrastructure, see Figure 1.

Figure 1. VBIED Damage to Building in Jakarta (2004) and Bridge in Iraq (2009).

Securing airports and aircraft has been a high priority of governments world-wide after the 9/11 attacks. Several terrorist plots have recently been foiled, which if successful, would have killed many hundreds of people. The U.S. Transportation Security Administration (TSA) has arrayed '21 Layers of Security' to 'strengthen security through a layered approach'. This is similar to counter-terrorism (CT) strategies worldwide. Assessing the effectiveness and reliability of aviation CT measures is important to understanding their strengths and weaknesses, and assessing the need for additional security measures.

There are considerable uncertainties associated with threat scenarios, system response, effectiveness of CT measures and expected damage. Since IEDs are typically 'home made' and placed under imperfect conditions, then the probability of a successful detonation can be highly uncertain, as evidenced in recent failed attempts to blow up U.S. airliners. These uncertainties will affect damage risk predictions and the utility of subsequent decisions. Characterising these uncertainties using stochastic (probabilistic) methods is a logical step, which will lead to estimates of system reliability and risk. Only very few probabilistic and reliability analyses have been carried out for infrastructure systems subject to explosive blast loading (e.g. Twisdale 1994, Low and Hao 2001, 2002, Eamon 2007, Hao et al 2010). This is in contrast to the approach that has been used very widely and successfully for other man-made and natural hazards (e.g. Stewart and Melchers 1997). Risk and reliability analyses will allow comparisons to be made between the relative effectiveness of security measures, weapon selection, delivery method or other mitigation measures.

To compare costs and benefits requires the quantification of threat probability, risk reduction, losses, and security costs. This is a challenging task, but necessary for any risk assessment, and the quantification of security risks is recently being addressed (e.g. Stewart et al. 2006, Stewart and Netherton 2008, Netherton and Stewart 2009, Dillon et al. 2009, Cox 2009, Stewart and Mueller 2008a, 2008b, 2011), as well as recent life-cycle and cost-benefit analyses for infrastructure protective measures (Willis and LaTourette 2008, von Winterfeldt and O'Sullivan 2006, Stewart 2008, 2010a,b, 2011). Much of this work can be categorized as 'probabilistic terrorism risk assessment'.

The cumulative increase in expenditures on U.S. domestic homeland security over the decade since 9/11 exceeds one trillion dollars (Mueller and Stewart 2011a,b). Up to 45% of this expenditure is devoted to protecting critical infrastructure and key resources. Yet there is little evidence that such expenditures have

been efficient. Clearly, for efficient decision-support to occur there is a need to quantify security risks and assess their level of acceptability and cost-effectiveness. A significant challenge is balancing the costs and benefits of counter-terrorism measures when the threat scenarios are highly transient and considerable risk averseness displayed by decision makers. For security and public policy purposes a quantification of security risks is essential for risk acceptability and robust decision-making.

It was understandable, in the years immediately following the terrorist attacks of September 11, 2001 that there was a tendency to spend in haste on homeland security. For example, annual security costs for the U.S. airline industry have increased to over $8 billion (DHS 2011), yet little scientific rigour has been applied to assess the effectiveness of this expenditure as evidenced by a statement from the U.S. Department of Homeland Security that 'We really don't know a whole lot about the overall costs and benefits of homeland security.' (Anderson 2006). These concerns are equally valid for Australia, Canada and Europe. There is a need to examine homeland security expenditures in a careful and systematic way, applying the kind of system and reliability modelling approaches that are routinely applied to other hazards. This type of rigour, where security and public policy decisions are assessed on technical, social and economic considerations of risk acceptability, is much needed to ensure that public funds are expended on measures that maximise public safety.

Terrorism may be viewed as a 'new hazard', that although different in nature from other hazards, requires systems and reliability approaches similar to those adopted to other hazards to assess risk and safety. The paper will review recent research conducted at The University of Newcastle, including:

1. Stochastic modelling of blast loads
2. Stochastic modelling of structural response
3. Systems and Reliability analysis
4. Risk-based decision theory

This is a multi-faceted approach to probabilistic terrorism risk assessment that deals with existing and new (hardened) infrastructure. A capability to predict the likelihood and extent of damage and casualty levels has many potential uses; including:

1. infrastructure and security policy, as a decision support tool to mitigate damage
2. contingency planning and emergency response simulations
3. collateral damage estimation (CDE) for military planners
4. forensics to back-calculate charge weights.

A review of probabilistic risk assessments are given for specific example applications: (i) IED design and initiation, and predicting variability of time-pressure load history on infrastructure, (ii) reinforced concrete structural systems, (iii) airports subject to terrorist attack, and (iv) buildings subject to a terrorist Vehicle Borne Improvised Explosive Device (VBIED). The illustrative examples in this paper, where possible, use actual or representative threat, consequence and cost data. However, some hypothetical data is used (particularly when dealing with terrorist threats in Section 5) as the intention of the examples is to show the methodology of various risk acceptance criteria and not to make any definitive conclusions about a specific item of infrastructure.

For additional and wider-ranging assessments of the issues raised and the approaches used, including risk and cost-benefit assessments of buildings, bridges and aviation systems (air marshals, full-body scanners, etc.), see John Mueller and Mark G. Stewart, *Terror, Security, and Money: Balancing the Risks, Benefits, and Costs of Homeland Security*, Oxford University Press, 2011.

2. Risk-Based Decision Support Framework

An advantage of a probabilistic risk assessment is that it can include a risk-cost-benefit analysis that considers tradeoffs between risks and costs. An appropriate decision analysis compares the marginal costs of CT protective measures with the marginal benefits in terms of fatalities and damages averted. The decision problem is to maximise the net benefit (equal to benefits minus the cost) or net present value:

$$E_b = E\left(C_B\right) + \sum_T \sum_H \sum_L Pr(T)\, Pr\left(H|T\right) Pr\left(L|H\right) \bullet L \bullet \Delta R - C_{security} \tag{1}$$

where $E(C_B)$ is the expected benefit from the security measure not directly related to mitigating terrorist threats (e.g. increased consumer confidence, reduction in crime), $Pr(T)$ is the annual threat probability per item of infrastructure, $Pr(H|T)$ is the conditional probability of a hazard (successful initiation/detonation of an IED, or other initiating event leading to damage and loss of life) given occurrence of the threat, $Pr(L|H)$ is the conditional probability of a loss given occurrence of the hazard, L is the loss or consequence (i.e., damage costs, number of people exposed to the hazard), ΔR is the reduction in risk due to CT measures, and $C_{security}$ is the extra cost of CT protective measures including opportunity costs. The product $Pr(L|H)L$ refers to the expected loss given the occurrence of the hazard. The summation signs in Eqn. (1) refer to the number of possible threat scenarios, hazard levels and losses. A protective measure is viewed as cost-effective or efficient if the net benefit exceeds zero (OBPR 2010). There are many risk acceptance criteria and these depend on the type of risk being quantified (life safety, economic, environmental, social), the preferences of the interested parties and the decision maker, and the quality of the information available. Risk acceptance criteria based on annual fatality risk or failure probability may also be used (e.g. Stewart 2010a,b, 2011).

Terrorism is a frightening threat that affects our willingness to accept risk, a willingness that is influenced by psychological, social, cultural, and institutional processes. Moreover, events involving high consequences can cause losses to an individual that they cannot bear, such as bankruptcy or the loss of life. On the other hand, governments, large corporations, and other self-insured institutions can absorb such losses more readily and so governments and their regulatory agencies normally exhibit risk-neutral attitudes in their decision-making (e.g. Sunstein 2002, Ellingwood 2006). This is confirmed by the U.S. Office of Management and Budget (OMB) which requires cost-benefit analyses to use expected values (an unbiased estimate), and where possible, to use probability distributions of benefits, costs, and net benefits (OMB 1992). However, Eqn. (1) can be generalised for expected utility incorporating risk aversion (e.g. Stewart et al. 2011). The issue of risk aversion is an important one as this seems to dominate CT and other decisions (Jordaan 2005, Mueller 2006), but also arises from uncertainty of CT effectiveness (and threats).

Equation (1) can be generalised for any time period, discounting of future costs and more detailed time-dependent cost and damage consequences. Fatality risks can be computed as the product $Pr(T)Pr(H|T)Pr(L|T)$ which can be compared with appropriate societal risk acceptance criteria (Stewart and Melchers 1997). Security cost data are available from the literature and security practitioners. This is not so for losses, although indicative values for damages due to terrorist attacks in the UK, US and elsewhere are available from the literature (Mueller and Stewart 2011a).

It is very difficult to estimate the threat probability $Pr(T)$. Progress in quantifying $Pr(T)$ will need contributions from security analysts and other academic disciplines. If information about $Pr(T)$ is believed to be too unreliable, then the decision analysis can be used to calculate the minimum (threshold) threat probability for CT protective measures to be cost-effective (i.e., a break-even approach). It is then the

prerogative of the decision-maker, based on expert advice about the anticipated threat probability, to decide whether or not a CT protective measure is cost-effective. Moreover, a decision analysis based on scenario analysis where threat probability is decoupled from Eqn. (1) provides an alternative decision-making criteria based on expected costs. The challenging aspect of risk-based decision theory is predicting values of $Pr(H|T)$, $Pr(L|H)$ and ΔR. This information may be inferred from expert opinions, scenario analysis, and statistical analysis of prior performance data, as well as system and reliability modelling. Since there is uncertainty associated with such predictions, the use of probability distributions to describe mean, variance and distribution type is recommended. However, it is recognised that data or models are often incomplete for such low probability − high consequence events, and so a sensitivity analysis should always be conducted to assess the robustness of results to parameter and modelling uncertainty.

We recognise that Eqn. (1) is an overly simplification, however, it is a useful starting point for further discussion and perhaps for more detailed and complex analysis of how to manage the often conflicting societal preferences associated with assessments of risk, cost, and benefits. Clearly, risk and cost-benefit considerations should not be the sole criterion for public decision making. Nonetheless, they provide important insights into how security measures may (or may not) perform, their effect on risk reduction, and their cost-effectiveness. They can reveal wasteful expenditures and allow limited funds to be directed to where the most benefit can be attained. More important, if risk and cost-benefit advice is to be ignored, the onus is on public officials to explain why this is so, and the trade-offs and cuts to other programs that will inevitably ensue.

3. Probabilistic Blast Load Modelling

3.1. RELIABILITY OF IMPROVISED EXPLOSIVE DEVICES (IED)

Unlike conventional military hardware, the reliability of IEDs cannot be calculated through standard philosophies such as those identified at MIL-HDBK-217 (Department of Defense 1995). Much of this is because IEDs have not been designed, manufactured and utilised in accordance with standard systems engineering practices by competent personnel, nor necessarily have they been developed by personnel familiar with operations or with military training.

The threat of IED attack, and hence development of a probabilistic risk assessment, can be treated through a systems model, using an alternate paradigm to conventional munitions reliability. The components that make up the IED can be assessed as per traditional reliability methodologies, however, the effects of design, environment, manufacturing and operational considerations need to be independently considered and overlaid as performance shaping functions (PSFs) that introduce additional variability in traditional reliability functions.

A reliability function can then be used to identify what could be considered the reliability for an IED design and manufacture – that is, the reliability of the IED due to the selection of components, their format and the intended operating environment. A baseline reliability function adapted from Wolstenholme (1999) is employed to develop the baseline reliability of the IED (R) where the IED is modelled as a series system of n components:

$$R = \prod_{c=1}^{n} [\alpha_c - \lambda_s t_s] \tag{2}$$

where λ_s is the IED component storage failure rate, t_s is the time the IED component was in storage, α_c is the reliability of each IED component, and n is the number of components.

This paper uses several typical IED configurations of differing design complexities – simple (pipe bomb), medium (mobile phone initiated VBIED) and complex (improvised mortar). An example calculation for a medium complexity device, a mobile phone initiated VBIED (noting that most components are not disclosed for security reasons), derived from representative Operational Level Reliabilities for munitions systems data from Australia, U.K. and the U.S., and representative mobile phone data, to inform component reliabilities, is

$$R = 0.9994 \times 0.999 \times 0.98 \times 0.97 \times 0.97 \times 0.999 = 0.920 \tag{3}$$

Table I provides a summary of baseline IED reliabilities derived from conventional munitions' representative component reliability data for common IED designs (Grant and Stewart 2011). The baseline reliability assumes there are no errors in connecting components, and assumes statistical independence of component reliabilities. Hence, R reflects the reliability of an IED designed and manufactured to military specifications and standards.

Table I. Typical IED Baseline Reliability Estimates for Device Complexity

Device Complexity	Representative IED Design	Baseline Reliability R
Simple	Pipe Bomb	0.931
Medium	Mobile Phone initiated VBIED	0.920
Complex	Improvised Mortar	0.910

The probability of IED initiation is $Pr(H|T)$ where H is IED initiation(hazard) and T is the threat, is

$$Pr(H|T) = \prod_{i=1}^{K} PSF_i \bullet R \tag{4}$$

where PSF_i is the performance shaping function for attribute i. Typical PSFs might include design quality, manufacture quality, education, training and experience, organisational culture, stress, etc.

One open source database from which data is available to quantify the PSFs, the Global Terror Database (GTD), is collated by the National Consortium for the Study of Terrorism and Responses to Terrorism (START) at the University of Maryland. Terrorist incidents were filtered based on Weapon Type and date (1998 to 2008). The dataset was re-characterised based on categorisation of device operation and device complexity – Unknown (insufficient incident information to make a categorisation); Simple (consisting of roadside bombs, hand-thrown devices and those containing conventional munitions as a warhead); Medium (car bombs, remotely-fused IEDs and use of homemade explosive); and Complex (devices such as homemade rockets, mortars and projectiles or IEDs with complex triggers). The limitations associated with the GTD constrained the fidelity of our model, however, we have been able to consider a PSF pertaining to device complexity based on Region and Organisational Culture, see Table II.

Table II shows significant variability in PSFs between organisational types and regions. One significant limitation of using the GTD as a dataset is that it has significant potential for bias related to open-source reporting, this is thought to be the reason why the results at Table II imply that IED initiation rates for Criminal, Terrorist and Insurgent Organisations equal that of their conventional equivalents used by Western militaries (i.e. $PSF = 1$). Despite this, particularly taking the data for Western incidents where reporting is more likely to be reflective of the actual incident population, we can identify that the lowest levels of performance were observed for individuals, as would be expected for conventional engineering and manufacturing activities since the diversity within teams means that they are better equipped to design and manufacture IEDs than individuals. It is also notable that the PSFs that were identified are similar to the critical factors that have been identified as impacting the performance of personnel and equipment for other industries/professions involving processes, skill and stress.

For more details, including probabilistic estimates of loss (damage, casualties) due to IED initiation, see Grant and Stewart (2011).

Table II. PSFs for IEDs in Regions of Interest

Organisational Culture	Device Complexity	Global	Western	Middle East & North Africa
Individual	Simple	0.588	0.537	0.614
	Medium	0.695	0.521	-
	Complex	-	-	-
Criminal	Simple	→1	0.986	1
	Medium	0.972	0.956	1
	Complex	0.550	-	-
Terrorist Organisation	Simple	0.981	0.855	0.990
	Medium	0.980	0.928	0.953
	Complex	0.905	0.761	1
Insurgent Organisation	Simple	→1	NA	1
	Medium	→1	NA	1
	Complex	→1	NA	1

3.2. TIME-PRESSURE LOAD HISTORY OF EXPLOSIVES

The variability in blast loading can be traced to:

(a) Parameter uncertainty,
(b) Inherent variability − natural, intrinsic, irreducible uncertainty of a situation, and
(c) Model error − measure of accuracy of predictive model.

In all cases the variabilities can be represented as one or more random variables described by their mean, COV (coefficient of variation) and probability distribution function. The probabilistic blast load model considers parameter uncertainties for (Netherton and Stewart 2010):

(a) User factor for mass of explosive (W_{user}),
(b) Net equivalent quantity (NEQ) of an explosive in terms of a mass of TNT (W_{NEQ})
(c) The range (R) and Angle of Incidence (AOI), and
(d) Air temperature (T_a) and pressure (P_a).

Probabilistic models for model error and inherent variability were obtained from field data of repeatable tests. The polynomial curves from the explosive blast loading model proposed by Kingery and Bulmash (1984) have been incorporated into widely used and well respected blast load design models, such as ConWep (1991), TM5-1300 (1990) and LS-DYNA. Given such wide acceptance, the polynomials of Kingery and Bulmash (1984) are used for predicting blast load values. The time-pressure history is idealised by an equivalent triangular pressure pulse.

The variability of blast load will be influenced by the type of explosive used, its manufacturer, its placement, etc. One explosive of significant interest to counter-terrorism personnel is "home-made" Ammonium Nitrate Fuel Oil (ANFO) delivered by a VBIED. The statistical parameters describing the variability of input parameters and model error (accuracy) are given in Table III, for a VBIED that uses ANFO as the explosive. For more details of the probabilistic blast load model see Netherton and Stewart (2010), which also includes a blast scenario for weapon delivery of a 500 lb Mark-82 GP bomb (89 kg Tritonal) using GBU-38 JDAM (GPS) guidance control.

Table III. Statistical Parameters for Blast Loading Model (Netherton and Stewart 2010).

Parameter	Mean	COV	Distribution
Energetic Output:			
User factor	1.00	0.102	Normal
NEQ factor	Mode = 0.82	0.359	Triangular
Detonation Location:			
VBIED Location	x = 0	σ = 3.06 m	Normal
	y = R	σ = 1.53 m	Normal
	z = 0	σ = 0 m	Deterministic
Ambient Air Temperature (^{o}C)	21.9 ^{o}C	0.356	Normal
Ambient Air Pressure (hPa)	1015.0 hPa	0.014	Uniform
Model Error:			
Peak reflected pressure (P_r)	1.032	0.069	Normal
Peak reflected impulse (I_r):			
0.59 m/kg$^{1/3}$ $\le Z <$ 6.0 m/kg$^{1/3}$	0.991	$0.178 - 0.0236Z$	Normal
6.0 m/kg$^{1/3}$ $\le Z <$ 40.0 m/kg$^{1/3}$	0.991	0.036	Normal
Time of positive phase duration (t_d):			
0.59 m/kg$^{1/3}$ $\le Z <$ 6.0 m/kg$^{1/3}$	$0.43 + 0.596\log_{10}Z$	$C_0 + C_1Z + C_2Z^2 + C_3Z^3$	Normal
6.0 m/kg$^{1/3}$ $\le Z <$ 9.0 m/kg$^{1/3}$	$0.43 + 0.596\log_{10}Z$	0.046	
9.0 m/kg$^{1/3}$ $\le Z <$ 40.0 m/kg$^{1/3}$	1.00	0.046	Normal

Note: $C_0 = 0.6267$, $C_1 = -0.3510$, $C_2 = 0.0713$, $C_3 = -0.0048$, Z is scaled distance (m/kg$^{1/3}$)

The blast scenario considered herein is a small van-sized VBIED comprising 116 kg of "home-made" ANFO. The explosive for this scenario detonates on or very near to the ground. It is thus considered a hemispherical charge detonating against a reflecting surface. The blast load is from a single uninterrupted emanation of the shock-wave and that reflections from other structures or surfaces are not considered. The probability distribution of peak reflected pressure (P_r), impulse (I_r), and the time of a blast-waves first positive phase duration (t_d) are the outcomes of the probabilistic analysis − see Figure 2 for W = 116 kg ANFO and stand-off R = 50 m. Figure 2 also shows the TM5-1300 (or ConWep) design values. Note that

the design value based on the TM5-1300 approach includes a 'safety factor' where explosive mass (W) is increased by 20%. It is observed that the variability of blast load parameters is considerable, with COVs of 0.15 to over 1.0. These are significant variabilities, and roughly equivalent to the observed variability for earthquake loadings which has the highest variability of all natural hazards. It is observed that the probability that the explosive load exceeds the TM5-1300 design value is 28%, 4% and 19% for P_r, I_r and t_d, respectively. More research is needed that calculates the probability of exceedance for a wider range of blast scenarios before any definitive conclusions can be made about the conservatism (or not) of ConWep, TM5-1300 and other design tools for explosive blast loading.

4. Probabilistic Modelling of Structural Response and Reliability Analysis

The probability of the hazard for infrastructure conditional on the occurrence of a specific threat is

$$Pr\left(H|T\right) = Pr\left[G(\mathbf{X}) \le 0\right] \tag{5}$$

where $G(\mathbf{X})$ is the limit state function (of structural response) and $\mathbf{X}$ is the vector of all relevant variables. $G(\mathbf{X}) = 0$ defines the boundary between the 'unsafe' and 'safe' domains. The limit state functions can be expressed in terms of structural damage, safety hazards and casualties. The exposure of people to blast effects is highly dependent on site location, building layout, occupancy rates, etc. and so the effect of low and high exposures will be considered, both deterministically and probabilistically. As a structure ages the effect of deterioration and other time-dependent processes may lead to higher values of $Pr(H|T)$.

Computer software Blast-RF (Blast Risk for Facades) that calculates $Pr(H|T)$ for damage, safety level and casualties for glazing systems is currently under development and intended as freeware in the near future. Details are available elsewhere (Stewart and Netherton 2008, Netherton and Stewart 2009).

The discussion to follow will focus instead on the structural capacity and reliability of RC columns subject to explosive blast loading. The RC column is representative of a ground floor central column of a two storey RC frame building (Shi et al. 2008). The RC column is H = 4.6 m high and is of rectangular cross-section (see Figure 3). Table IV shows the design (nominal) material and dimensional properties of the RC column. The finite element model used herein is identical to that developed by Shi et al. (2008) using explicit FEM software LS-DYNA.

Since RC columns are designed to support an axial load, then the damage criterion is based in axial load-carrying capacity. The damage index (D) is defined as (Shi et al. 2008):

$$D = 1 - \frac{P_{residual}}{P_{design}} \tag{6}$$

where $P_{residual}$ is the residual axial load-carrying capacity of the damaged column, and P_{design} is the maximum axial load-carrying capacity of the undamaged column. Shi et al. (2008) define four damage limit states based on the damage index D:

1. D = 0−0.2 low damage
2. D = 0.2−0.5 medium damage
3. D = 0.5−0.8 high damage
4. D = 0.8−1.0 collapse

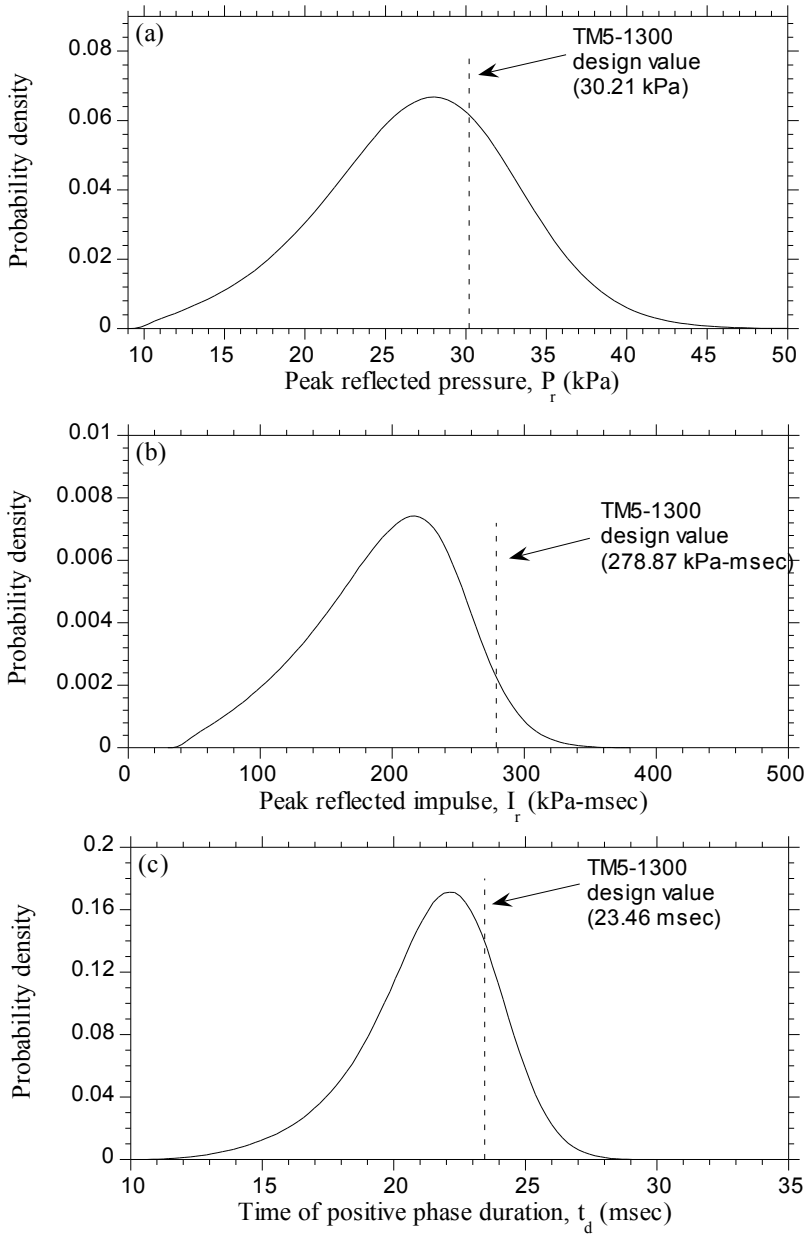


Figure 2. Probability Distributions of Blast Load Parameters and Comparison with TM5-1300 Design Values (adapted from Netherton and Stewart 2010).

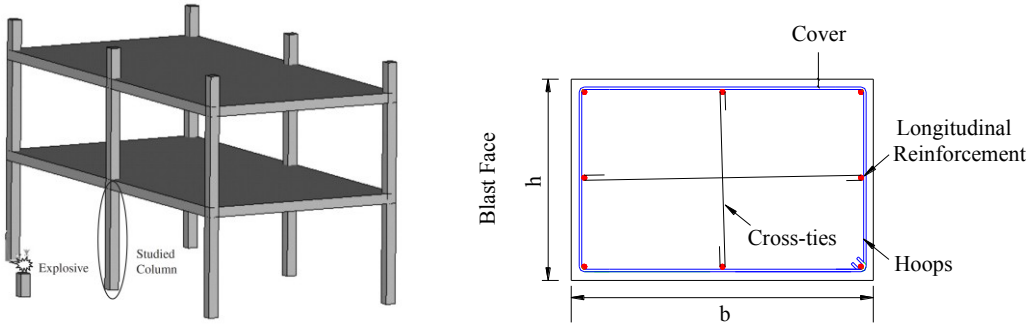


Figure 3. Location and Cross-section of RC Column.

Table IV. Material and Dimensional Properties for RC Column.

Parameter	Design Value
Column width (h)	400 mm
Column depth (b)	600 mm
Hoops/Cross ties spacing (s)	200 mm
Longitudinal reinforcement	8 × 20 mm diameter
Yield strength of longitudinal steel (F_y)	413.7 MPa (Grade 60)
Hoops/Cross ties	10 mm @ 200 mm spacing
Yield strength of hoops and cross-ties	275.8 MPa (Grade 40)
Cover	25 mm
Concrete Compressive Strength (F'_c)	42 MPa

Monte-Carlo simulation (MCS) is used for reliability estimation of the RC column. The probability of damage states conditional on threat T is $Pr(H|T)$:

$$Pr\left(\text{low damage}|T\right)=\frac{n\left[D<0.2\right]}{N} \qquad Pr\left(\text{medium damage}|T\right)=\frac{n\left[0.2\le D\le 0.5\right]}{N}$$
$$Pr\left(\text{high damage}|T\right)=\frac{n\left[0.5<D\le 0.8\right]}{N} \qquad Pr\left(\text{collapse}|T\right)=\frac{n\left[D>0.8\right]}{N} \tag{7}$$

where $n[]$ is the number of realisations when D matches the damage criterion, and N is the number of simulation runs.

The blast scenario considered is a $W = 100$ kg ANFO VBIED detonated from $R = 2.5$ m to $R = 20$ m from the front face of the RC column. The probabilistic load model described in Section 3.2 is used herein, where statistical parameters are given by Table III. The statistical parameters for cover, concrete compressive strength and yield strength of reinforcement are given in Table V. These statistics are representative of new RC columns constructed in the United States. Due to high computational demand associated with LS-DYNA, $N = 100$ simulation runs were used to generate distributions of load-carrying

capacity, damage index and probabilities of damage and collapse.

Table V. Statistical Parameters for RC Column (adapted from Stewart et al. 2011).

Parameter	Mean	COV	Distribution
Cover (mm)	C_{nom} + 6.4 + 0.004h	σ = 24.9 mm	Normal[a]
Yield Strength (MPa)	$1.145F_y$	0.05	Normal[b]
Concrete Compressive Strength	F'_c + 7.5 MPa	σ = 6 MPa	Lognormal

Note: [a] truncated at stirrup diameter (10 mm), [b] truncated at zero.

Results show that the COV of load-carry capacity of the undamaged (P_{design}) and damaged ($P_{residual}$) columns when R = 10 m are 0.13 and 0.32, respectively. Clearly, there is increased variability for a damaged structural element. Blast Reliability Curves (BRCs) are shown in Figure 4. The 90% confidence bounds are also shown − more simulation runs would reduce the 90% confidence intervals, but those shown in Figure 4 are sufficient to infer the BRCs. As expected, the probability of collapse reduces as stand-off (R) increases, and when R exceeds 15 m the probability of collapse is negligible. On the other hand, even though the risk of collapse is less than 10% when R = 10 m, there still remains a very high likelihood of low or medium damage. The BRCs provide a useful metric for assessing safety and damage risks. For more details see Stewart et al. (2011).

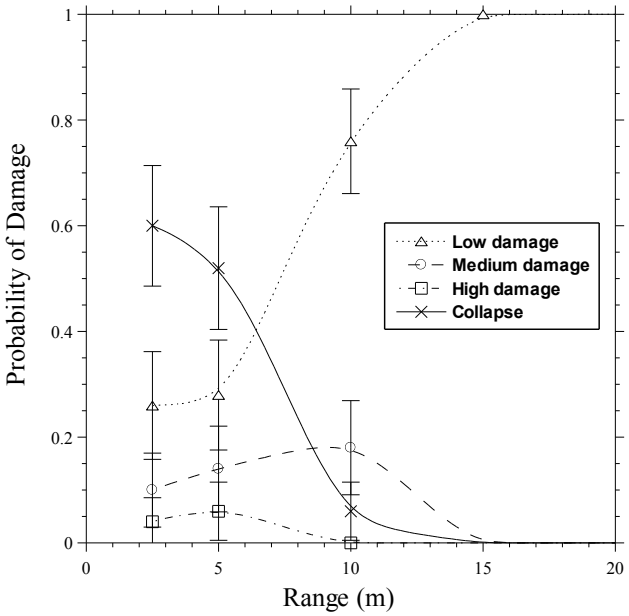


Figure 4. Blast Reliability Curves (BRC) for RC Column.

5. Cost Benefit Assessment of Infrastructure Protection

To illustrate the benefits of probabilistic terrorism risk assessment an airport terminal and institutional building subject to a terrorist Vehicle Borne Improvised Explosive Device (VBIED) are considered. The illustrative example will show under what combination of security costs, risk reduction, fatality and damage costs, and attack probability the protective measures would be cost-effective.

5.1. AIRPORTS

Although there may be special reasons to protect airplanes, it is not at all clear that there are any special reasons to protect airports. Compared with many other places of congregation, people are more dispersed in airports, and therefore, a terrorist attack is likely to kill far fewer than if, for example, a crowded stadium is targeted. The 2011 suicide bombing of the arrivals area of Moscow's Domodedovo airport, which killed 36 and injured 15 others, shows that airports are not unattractive targets, but in the previous year, suicide bombers targeted the Moscow metro, killing 25, and the year before that, derailed the Moscow to St. Petersburg high-speed train, killing 27.

In addition, airports sprawl and are only two or three stories high, and therefore damage to a portion is not likely to be nearly as significant as damage to a taller or more compact structure. Moreover, if a bomb does go off at an airport, the consequences would probably be comparatively easy to deal with: passengers could readily be routed around the damaged area, for example, and the impact on the essential function of the airport would be comparatively modest.

In the 10 year period 1998−2007 there were ten (2 fatalities) and nine (29 fatalities) attacks on airports in Europe and Asia-Pacific, respectively. The annual fatality risk is approximately 2×10^{-10} and 6.5×10^{-9} for Europe and Asia-Pacific, respectively. These are very low risks, and are considered "acceptable" based on a fair degree of agreement about acceptable risk (Stewart and Melchers 1997). However, terrorism is a hazard where risk acceptability might not be a matter of fatality risks due to the significant direct and indirect economic consequences of a terrorist attack. For example, losses inflicted by the terrorist attack that has been by far the most destructive in history, that of September 11, 2001 approached $200 billion (Mueller and Stewart 2011a).

The threat considered herein is a bombing of an airport terminal. A small IED might kill say five people, no structural damage, and minimal disruption to flight schedules - we value this attack at $50 million based on the value of a single life (VSL) is $6.5 million (Robinson 2010) plus other costs. On the other hand, a larger VBIED might kill 100 people ($65 million), severe structural damage to part of a terminal building ($100 million), and flight disruptions and relocation of check-in counters, etc. might total several billion dollars as a plausible upper bound. Security and protective measures to mitigate IED or VBIED attacks might include extra security personnel, vehicle entry screening for explosives, bollards, parking restrictions, etc. To be conservative, we assume that the increased cost of security is $C_{security}$ = $2 million per year for each airport terminal. For Sydney Airport, this would be equivalent to an 8% increase in their security budget. Opportunity costs associated with some security measures might be considerable, such as parking restriction near the terminal might deter passengers, or extra security screening will delay passengers. We do not consider such opportunity costs in this analysis.

Equation (1) can be simplified by assuming that $Pr(H|T) = Pr(L|H) = 1$, and so a break-even analysis to calculate how many attacks would have to take place to justify the expenditure gives

$$p_{attack-min} = \frac{C_{security}}{L \bullet \Delta R} \tag{8}$$

Table VI arrays the annual attack probabilities ($p_{attack\text{-}min}$) required at a minimum for enhanced security expenditures on protecting an airport terminal to be cost-effective.

This break-even analysis shows that protective measures that reduce risk by an impressive 75% and that successfully protect against an attack that would otherwise inflict $50 million in damage would be cost-effective only if the annual probability of a successful terrorist attack without them exceeds 0.05 or one in 20 per terminal per year. If we assume a $2.5 billion attack, and risk is reduced by 75 percent, the minimum attack probability per year required for airport protective measures to be considered cost-effective reduces to 0.001 per terminal per year. There have been five bomb attacks in the ten year period 1998−2007 in the Asia-Pacific region. If we assume there are 500 airport terminals in the Asia-Pacific region, then the attack probability is 0.001 per terminal per year. In this case, security and protective measures that cost $2 million per year would only be cost-effective if they reduce risk by 75% and prevent losses of $2.5 billion. For lower losses, or risk reductions, such security and protective measures would only be cost-effective if the attack probability greatly exceeded 0.001 per terminal per year.

Table VI. The number of otherwise successful attacks per year in which enhanced airport security would have to be solely responsible for deterring, foiling, or protection against in order for its enhanced yearly security budget of $2 million to be cost-effective, at various levels of loss and risk reduction − that is, for the security benefit of the expenditures to equal their costs.

Risk Reduction Caused by Enhanced Airport Security Expenditure (ΔR)	Losses from a Successful Terrorist Attack (C_{loss})				
	$50 million	$100 million	$500 million	$1 billion	$2.5 billion
5 percent	0.80	0.40	0.080	0.040	0.0160
10 percent	0.40	0.20	0.040	0.020	0.0080
25 percent	0.16	0.08	0.016	0.008	0.0032
50 percent	0.08	0.04	0.008	0.004	0.0016
75 percent	0.05	0.03	0.005	0.003	0.0011
100 percent	0.04	0.02	0.004	0.002	0.0008

5.2. BUILDINGS

A typical multi-storey building for which occupancy and loss data are available is an academic building located at the U.S. Naval Postgraduate School in Monterey, California (Lakamp and McCarthy 2003). In this case, measures to protect the building from VBIED and other explosive blast loads include strengthening perimeter columns and walls, blast-resistant glazing and other improvements to structurally harden the building.

Damage and loss parameters are considered as random variables that explicitly consider aleatory and epistemic uncertainties. Three threat scenarios are assumed as $i = 1$: low, $i = 2$: medium and $i = 3$: high terrorist threats, and two types of loss attributes $j = 1$: direct physical damage and $j = 2$: fatalities. The net benefit from eqn. (1) is re-written for this example as

$$E_b = \sum_{i=1}^{3}\sum_{j=1}^{2} p_{attack}\, Pr(T_i|attack) Pr(H_i|T_i) Pr(L_j|H_i) \bullet L_j \bullet \Delta R_i - C_{security} \quad (9)$$

where p_{attack} is the annual attack probability, $Pr(T_i|attack)$ is the relative threat probability given an attack, L_1 is the cost of direct physical damage (building replacement, damage to contents), L_2 is the number of people exposed to the hazard (building occupants), and ΔR_i is the percentage reduction in risk due to CT protective measures for the i^{th} threat. We assume that $C_B = 0$ and $Pr(H_i|T_i) = 1$.

A low threat may be a VBIED with low explosive weight or large stand-off, whereas medium or high threats would involve, for example, larger VBIED explosive weights and reduced stand-off. It is assumed that $Pr(T_i|attack)$ reduces as the threat level increases due to reduced likelihood of conducting such an attack undetected as the size of vehicle increases or as the vehicle moves closer to the target building, see Table VII. Stewart (2011) has shown that the probability of building occupant fatality given a terrorist attack $Pr(L_2|H_i)$ varies from 0.0003 to 0.45 and so $Pr(L_2|H_i)$ is assumed relatively low for low and medium threats, and is unlikely to reach above 0.5 even for a high threat. This example does not consider the risk and safety of people outside the building (such as pedestrians).

Although a small VBIED can cause low casualties, the effect on physical damages can be much higher as although a VBIED may not totally destroy a building, it will often need to be demolished and replaced, hence the probability of physical damage is high even for a medium threat. As there is uncertainty about these threat and loss probabilities then they are treated as random variables and Table VII shows their assumed statistical parameters and probability distributions. Note that a coefficient of variation (COV) of 0.25 represents a 95% confidence interval of approximately ± 50% about the mean value.

Table VII. Probabilistic Models for Hypothetical Threats and Losses (Stewart 2010b).

	Relative Threat Probability $Pr(T_i\|attack)$	Probability of Physical Damage $Pr(L_1\|H_i)$			Probability of Fatalities $Pr(L_2\|H_i)$		
Threat		mean	COV	Distribution	mean	COV	Distribution
$i = 1$ Low	0.6	0.25	0.1	Lognormal	0.1	0.25	Lognormal
$i = 2$ Medium	0.3	0.80	0.1	Lognormal	0.25	0.25	Lognormal
$i = 3$ High	0.1	1.0	-	-	0.5	0.25	Lognormal

Note: probability distributions censored at 0.0 and 1.0

Significant strengthening of a building is likely to reduce damage and fatality levels to near zero for low threat events, however, even a significantly strengthened structure can experience damage and casualties if the threat is high. It follows that risk reduction will reduce, perhaps marginally, as the size of the threat increases. Risk reductions are also modelled as a random variables, see Table VIII, where it is assumed that the risk reduction is accurate to ± 10%.

Table VIII. Probabilistic Models for Hypothetical Risk Reduction (Stewart 2010b).

	Risk Reduction ΔR_i		
Threat	mean	COV	Distribution
$i = 1$ Low	90%	0.064	Uniform [80−100]
$i = 2$ Medium	65%	0.089	Uniform [55−75]
$i = 3$ High	50%	0.115	Uniform [40−60]

The cost of physical damages is approximately L_1 = \$35 million − this includes replacement value of the building, value of contents, and demolition costs. There is more certainty about damage losses so L_1 is modelled as a normal distribution with mean = \$35 million and COV = 0.05. The academic building is sizeable, with offices and teaching space, and peak usage comprising 319 building occupants (Lakamp and McCarthy 2003). To maximise the impact of a terrorist attack, an attack would most likely occur at a time of high building occupancy, so it is assumed herein that the number of occupants (L_2) is modelled as a normal distribution with mean = 250 people and COV = 0.17 so that there is a 10% probability than occupancy will be higher than 319 occupants in the event of a terrorist attack. The value of a single life (VSL) is \$6.5 million (Robinson 2010), hence, mean L_2 = \$1.6 billion.

A literature review by Stewart (2011) found that the minimum cost of protective measures ($C_{security}$) needed for substantial risk reduction for an existing building is at least 10% of building costs. If we assume that the budget time period for providing protective measures to the building is five years, then if the 10% increase in costs is annualised over five years with a discount rate of 3% then this equates to a present value cost of $C_{security} \approx$ \$450,000 pa.

The net benefit is calculated from Eqn. (10) using Monte-Carlo simulation analysis for a range of attack probabilities. Figure 5 shows the simulation histogram of net benefit for three attack probabilities: $p_{attack} = 10^{-2}$, 10^{-3} and 10^{-4} /building/year. As there is random variability with many of the input parameters then net benefit is variable as shown in Figure 5. With reference to Figure 5 it is clear that if $p_{attack} = 10^{-2}$ per building per year then there is near 100% confidence that the net benefit is positive so near 100% sure that the protective measures are cost-effective. On the other hand, if $p_{attack} = 10^{-4}$/building/year then there is near 100% certainty that protective measures are not cost-effective. If $p_{attack} = 10^{-3}$/building/year then Figure 5 shows that there is only a 35% probability that protective measures are cost-effective (i.e., $Pr(E_b) > 0$). Figure 6 shows another way to present results and this shows the mean and lower and upper bounds (5th and 95th percentiles) of net benefit for various attack probabilities. The threshold threat probability is 5.6×10^{-4}/building/year so if an attack probability exceeds this threshold (or break-even) value then the protective measure is likely to be cost-effective. Note that Ellingwood (2006) suggests that the minimum attack probability be at least 10^{-4}/building/year for high density occupancies, key governmental and international institutions, monumental or iconic buildings or other critical facilities with a specific threat. It should be noted that although the probability of a terrorist attack may be high, the probability that any particular item of infrastructure will be attacked is very low. If the annual attack probability is 10^{-4}/building/year then the protective costs outweigh the benefits ($E_b < 0$) and so protective measures would not be cost-effective. Clearly, due to the uncertainties inherent in such an analysis, a sensitivity analysis is recommended, see Stewart (2010a) for further details and analysis.

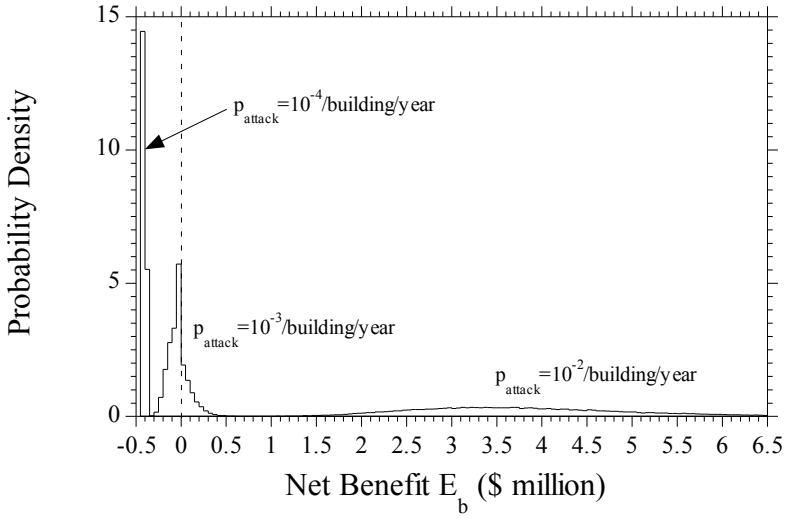


Figure 5. Histograms of Annual Net Benefit (E_b) for Institutional Building, for Attack Probabilities of 10^{-2}, 10^{-3} and 10^{-4} per year.

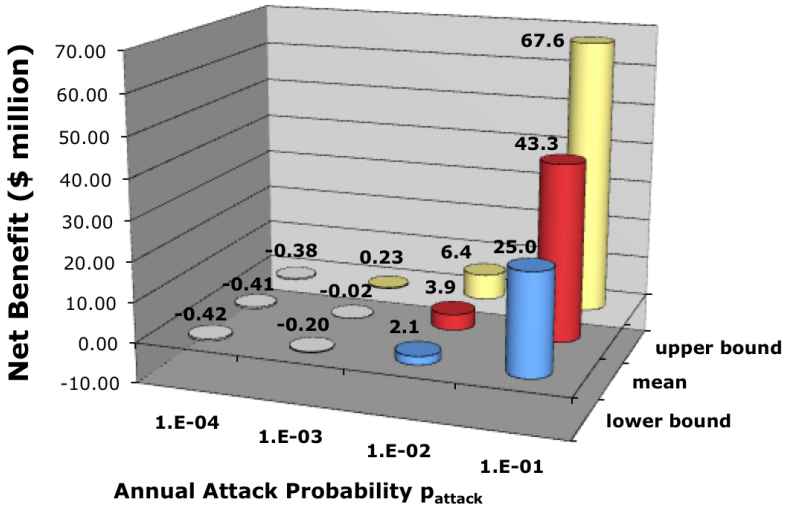


Figure 6. Annual Net Benefit (E_b) for Institutional Building.

6. Risk Transfer

An important consideration in critical infrastructure protection is the displacement effect, a transfer of risk. Terrorists can choose, and change, their targets, depending on local and immediate circumstances. This

process, of course, does not hold in the case of natural disasters: a tornado bearing down on Kansas does not decide to divert to Oklahoma if it finds Kansans too well protected. In contrast, if the protection of one target merely causes the terrorist to seek out another from among the near-infinite set at hand, it is not clear how society has gained by expending effort and treasure to protect the first. The people who were saved in the first locale are gainers, of course, but their grief is simply transferred to others.

For example, there is a program to protect bridges in the United States, and a list of something like 200 of the most important bridges has been drawn up. There seems to be no evidence terrorists have any particular desire to blow up a bridge, due in part, perhaps, to the facts that it is an exceedingly difficult task under the best of circumstances and that the number of casualties is likely to be much lower than for many other targets.

The apparent hope of the protectors in this case is that, after security is improved for all these targets, any terrorists who happen to have bridges on their hit list will become disillusioned. If so, however, they might become inclined to move on to the 201st bridge or, more likely perhaps, to another kind of bridge: the highway overpass, of which there are some 600,000 in the United States. If the terrorists' attention is drawn, further, to any one of a wide array of multiple overpass bridge networks, they might be inclined to destroy one of those. The financial and human consequence, not to mention the devastating traffic inconvenience, that could result from such an explosion might well surpass the destructive consequences of one directed at one of those 200 bridges. The issue, then, is: how has society been benefited by the protection of the bridges?

The 2011 suicide bombing at Moscow's Domodedovo airport took place in the arrivals area, well away from the passenger security screening. Accordingly, any risk reduction passengers gained by being in the secure zone of the airport was simply transferred to those outside, as the attackers targeted a place of public assembly for which there are few countermeasures.

Or there is the case of the installation of sensors to measure chemical, biological, or radiological levels in New York. Presumably, any terrorists clever enough to engineer the relevant weapons are likely to be able to learn where the sensors have been put in place, and there is no gain to society if they simply choose to move to Newark or Washington or Columbus. However, this elemental consideration does not appear to have been part of the decision process.

7. Assessing 'Critical' Infrastructure

There is no doubt that a terrorist attack on many infrastructure elements could cause considerable damage and significant loss of life. However, while such targets as buildings, bridges, highways, pipelines, mass transit, water supplies, and communications may be essential to the economy and well-being of society, damage to one or even several of these, with few exceptions, will not be "critical" to the economy, or to the state.

In part, this is because infrastructure designers and operators place much effort on systems modelling to ensure that a failure of one node will not keep the network from operating, even if at reduced efficiency. This is done routinely: for example, it is necessary to close many bridges from time to time for maintenance or repair, and therefore traffic is redirected so that the network is not interrupted. Other failures routinely planned for include traffic accidents, severe weather, earthquakes, and equipment malfunctions. In other

words, as a matter of course, infrastructure is designed with built-in redundancies and backup systems to ensure resilience in the event of anticipated or unexpected hazards.

There is also a displacement effect, a transfer of risk. Terrorists can choose, and change, their targets, depending on local and immediate circumstances. If the protection of one target merely causes the terrorist to seek out another from among the near-infinite set at hand, it is not clear how society has gained by expending effort and treasure to protect the first.

Relying on standard evaluative measures accepted for decades by analysts, governments, regulators, and risk managers, efforts to protect people and structures from the effects of a terrorist attack are unlikely in general to be cost-effective because of the multiplicity of targets, the ability of terrorists to shift targets as needed, the capacity in many cases to quickly rebuild, the exceedingly low likelihood of an attack on a specific target, the limited capability of most terrorist groups, and the difficulty of predicting which targets are most appealing to them. If the terrorists' goal is to kill people, lucrative targets are essentially everywhere. If their goal is to destroy property, protection measures may be able to deter, inconvenience, or complicate, but only to the point where the terrorists seek something comparable among a vast—or even effectively infinite—array of potential unprotected targets.

Our cost-benefit assessment suggests, then, that many individual items of infrastructure, including airports and buildings, require no protective measures unless, perhaps, there is a very specific threat to them.

Finally, we are not arguing that much of homeland security spending is wasteful because we believe there will be no more terrorist attacks. Like crime and vandalism, terrorism will always be a feature of life, and a condition of zero vulnerability is impossible to achieve. However, future attacks might not be as devastating as 9/11, as evidenced by the attacks on Western targets in the ten years since 9/11 that, although tragic, each have claimed victims numbering in the tens to a few hundred. The frequency and severity of terrorist attacks are low, very low in fact, which makes the benefits of enhanced counterterrorism expenditures challenging to justify by any rational and accepted standard of cost-benefit analysis.

8. Conclusions

Since there is uncertainty associated with terrorist threats, structural and system response, effectiveness of counter-terrorism and protective measures, and their ability to inflict damage, then there is a need for probabilistic approaches to assessing and mitigating terrorism risks. The paper reviews probabilistic risk assessments for (i) IED design and detonation, and predicting variability of time-pressure load history on infrastructure, (ii) reinforced concrete structural systems, (iii) airport protection, and (iv) buildings subject to a terrorist Vehicle Borne Improvised Explosive Device (VBIED). The illustrative examples highlighted the recent research, and identified research challenges to be faced in the future. It was found that attack probabilities have to be very high for security and protective measures for buildings and airports to be cost-effective.

Acknowledgements

The support of the Australian Research Council is gratefully acknowledged. The assistance of PhD students Michael Netherton, Yufeng Shi, and Matt Grant is greatly appreciated. Professor Mueller appreciates the financial support of a Distinguished Scholar Award at Ohio State University.

References

Anderson, T. (2006), Terror May Be at Bay at Port: Shipping Hubs Too Vulnerable, *Daily News of Los Angeles*, 18 May.

CONWEP (1991), *Conventional Weapons Effects Program*. Prepared by D.W. Hyde, US Waterways Experimental Station, Vicksburg.

Cox, L.A. (2009), Improving Risk-Based Decision-Making for Terrorism Applications, *Risk Analysis*, 29(3): 336-341.

Department of Defense (1995), *MIL-HDBK-217F Reliability Prediction of Electronic Equipment Notice 2*, Washington, D.C.

Dillon, R.L., Liebe, R. and Bestafka, T. (2009), Risk-based Decision Making for Terrorism Applications, *Risk Analysis*, 29(3): 321-335.

DHS (2011), FY2011 Budget in Brief, Department of Homeland Security, Washington, D.C.

Eamon, E. (2007), Reliability of concrete masonry unit walls subjected to explosive loads. *Journal of Structural Engineering* ASCE, 133(7):935–44.

Ellingwood, B.R. (2006), Mitigating Risk from Abnormal Loads and Progressive Collapse, *Journal of Performance of Constructed Facilities*, 20(4): 315-323.

Grant, M. and Stewart, M.G. (2011), System and Reliability Modelling of Improvised Explosive Devices, *PARARI 2011 - 10th Australian Explosive Ordnance Symposium*, Brisbane, 8-9 November, 2011.

Hao, H., Stewart, M.G., Li, Z-X. and Shi, Y. (2010), RC Column Failure Probabilities to Blast Loads, *International Journal of Protective Structures*. 1(4):571-591.

Jordaan, I. (2005), *Decisions Under Uncertainty: Probabilistic Analysis for Engineering Decisions*, Cambridge University Press.

Kingery, C.N. and Bulmash, G. (1984), *Airblast Parameters From TNT Spherical Air Burst and Hemispherical Surface Burst*. Technical Report ARBRL-TR-02555. US Army Armament Research and Development Centre, Maryland, USA.

Lakamp, D.J. and McCarthy, G.H. (2003), *A Cost-Benefit Analysis of Security at the Naval Postgraduate School*, MBA Professional Report, Naval Postgraduate School, Monterey, California.

Low, H.Y. and Hao, H. (2001), Reliability analysis of reinforced concrete slabs under explosive loading, *Structural Safety*, 23(2):157–178.

Low, H.Y. and Hao, H. (2002), Reliability analysis of direct shear and flexural failure modes of RC slabs under explosive loading, *Engineering Structures*, 24:189–198.

Mueller, J. (2006), *Overblown: How Politicians and the Terrorism Industry Inflate National Security Threats, and Why We Believe Them*, Free Press, New York.

Mueller, J. and Stewart, M.G. (2011a), *Terror, Security, and Money: Balancing the Risks, Benefits, and Costs of Homeland Security,* New York and Oxford, UK: Oxford University Press, September 2011.

Mueller, J. and Stewart, M.G. (2011b), The Price is Not Right: The U.S. spends too much money to fight terrorism, *Playboy*, 58(10), 149-150.

Netherton, M.D. and Stewart, M.G. (2009), The Effects of Explosive Blast Load Variability on Safety Hazard and Damage Risks for Monolithic Window Glazing, *International Journal of Impact Engineering*, 36(12): 1346-1354.

Netherton, M.D. and Stewart, M.G. (2010), Blast Load Variability and Accuracy of Blast Load Prediction Models, *International Journal of Protective Structures*. 1(4):543-570.

OMB (1992), *Guidelines and Discount Rates for Benefit-Cost Analysis of Federal Programs (Revised)*, Circular No. A-94, October 29, 1992, Office of Management and Budget, Washington, DC.

OBPR (2010), *Best Practice Regulation Handbook*, Office of Best Practice Regulation, Australian Government, Canberra, June 2010.

Robinson, L.A., Hammitt, J.K., Aldy, J.E., Krupnick, A. and Baxter, J. (2010), Valuing the Risk of Death from Terrorist Attacks, *Journal of Homeland Security and Emergency Management,* 7(1).

Shi, Y., Hao, H. and Li, Z-X. (2008), Numerical Derivation of Pressure-Impulse Diagrams for Prediction of RC Column Damage to Blast Loads, *International Journal of Impact Engineering,* 35(11): 1213-1227.

Stewart, M.G. and Melchers, R.E. (1997), *Probabilistic Risk Assessment of Engineering Systems*, Chapman & Hall, London.

Stewart, M.G., Netherton, M.D. and Rosowsky, D.V. (2006), Terrorism Risks and Blast Damage to Built Infrastructure, *Natural Hazards Review*, ASCE, 7(3):114-122.

Stewart, M.G. (2008), Cost-Effectiveness of Risk Mitigation Strategies for Protection of Buildings against Terrorist Attack, *Journal of Performance of Constructed Facilities*, ASCE, 22(2):115-120.

Stewart, M.G. and Netherton, M.D. (2008), Security Risks And Probabilistic Risk Assessment of Glazing Subject to Explosive Blast Loading, *Reliability Engineering and System Safety*, 93(4):627-638.

Stewart, M.G. and Mueller, J. (2008a), A Risk and Cost-Benefit and Assessment of U.S. Aviation Security Measures, *Journal of Transportation Security*, 1(3): 143-159.

Stewart, M.G. and Mueller, J. (2008b), A Cost-Benefit and Risk Assessment of Australian Aviation Security Measures, *Security Challenges,* 4(3): 45-61.

Stewart, M.G. (2010a), Acceptable Risk Criteria for Infrastructure Protection, *International Journal of Protective Structures*, 1(1):23-39.

Stewart, M.G. (2010b), Risk-Informed Decision Support for Assessing the Costs and Benefits of Counter-Terrorism Protective Measures for Infrastructure, *International Journal of Critical Infrastructure Protection*, 3(1):29-40.

Stewart, M.G. (2011), Life Safety Risks and Optimisation of Protective Measures against Terrorist Threats to Infrastructure, *Structure and Infrastructure Engineering*, 7(6): 431-440.

Stewart, M.G., Ellingwood, B.R. and Mueller, J. (2011), Homeland Security: A Case Study in Risk Aversion for Public Decision-Making, *International Journal of Risk Assessment and Management*, 15(5/6): 367-386.

Stewart, M.G. and Mueller, J. (2011), Cost-Benefit Analysis of Advanced Imaging Technology Fully Body Scanners for Airline Passenger Security Screening, *Journal of Homeland Security and Emergency Management*, 8(1): Article 30.

Stewart, M.G., Shi, Y. and Zhi, X. (2011), Structural Reliability Analysis of Reinforced Concrete Columns Subject to Explosive Blast Loading, *9th International Conference on Shock & Impact Loads on Structures*, Y. Sonoda and T.S. Lok (eds), CI-Premier, 91-102.

Sunstein. C.R. (2002), *The Cost-Benefit State: The Future of Regulatory Protection*, ABA Publishing, American Bar Association, Chicago.

TM5-1300 (1990), Design of Structures to Resist the Effects of Accidental Explosions, US Department of the Army Technical Manual TM5-1300, USA.

Twisdale, L.A., Sues, R.H. and Lavelle, F.M. (1994), Reliability-based design methods for protective structures. *Structural Safety*,15(1–2):17–33.

von Winterfeldt, D. and O'Sullivan, T.M. (2006), Should WE Protect Commercial Airplanes Against Surface-to-Air Missile Attacks by Terrorists?, *Decision Analysis*, 3(2): 63-75.

Willis, H. and LaTourette, T. (2008), Using Probabilistic Terrorism Risk-Modeling for Regulatory Benefit-Cost Analysis: Application to the Western Hemisphere Travel Initiative in the Land Environment*, Risk Analysis* 28(2):325-339.

Wolstenholme, L.C. (1999), *Reliability Modelling – A Statistical Approach*, Chapman & Hall/CRC, U.K.

Gamma Process –Life cycle analysis of the Neumarkt Bridge, IT

Alfred Strauss, Hirut Abebe Demissie
Department of Civil Engineering and Natural Hazards, University of Natural Resources and Life Science, 1190 Vienna, Austria, alfred.strauss@boku.ac.at, hirut.abebe@boku.ac.at

Abstract: In the maintenance management of infrastructures, stochastic processes as for example the Gamma process are gaining importance in methods for performance prediction. Gamma process approaches can be based on classical information gained from visual inspections, as for example crack formation, bending, but also the development of near surface strains (stresses), and enable the development of forecasting models as an effective decision making basis for the optimization of inspection intervals and maintenance measures. In this contribution, which is based on the material properties during and after the demolition of pre-stressed concrete bridges, we example the possibility of the Gamma process approach (in relation to visual inspections) to capture the internal mechanical changes, caused for example by pre-stressing steel corrosion processes. A pronounced correlation between the gamma process approach and the internal mechanical properties of structure are bases for (a) a quantitatively well ascertained remaining service life, (b) optimization of inspection periods, (c) identification of critical structural components for the overall condition and consequently (d) cost-efficient maintenance.

Keywords: Gamma process, Lifecycle analysis, aging of concrete structures

1. Introduction

Numerous highly developed industrial countries have pronouncedly different infrastructure management systems, which are there to ensure the safety levels demanded by the standards of their national economies. Every engineering structure is discrete and unique. The result is a deluge of data sets generated by these infrastructure management systems, which in turn leads to an overload of information that complicates the decision making process for the respective owners instead of facilitating it.

What is even more relevant, in the last few decades a wide range of modern monitoring technologies and numerical methods have been developed, ranging from novel sensor monitoring systems to routine video-imaging techniques. These developments make it possible to reform established infrastructure inspection methods and to bring them in line with a wide range of high-quality, site specific data based on physical and technical realities. Consequently these developments serve also to improve and adapt the inspection routines resulting in time-efficient and systematic optimization of those maintenance and management systems.

The present paper presents an approach for effective adaptive maintenance management employing additional information gained from new methods of inspection and analysis, as for example the simulation of deterioration process.

Generally speaking, the safety levels of all concrete structures decreases continuously as a result of aging processes caused by environment-induced mechanical and chemical loads. Today many countries prescribe certain specifications and norms (Schmidt and Sondermann, 2006; 2007; 2011) in order to control

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

the progressive development of these processes and to plan efficient maintenance measures. These specifications and norms standardize the regular examination of structures with adequately defined inspection intervals and clear specifications of the scope of those inspections. In line with the current practice, most examination programs based on a visual evaluation of structural components and accessories are performed by qualified personnel, whose qualification admittedly corresponds to the subjective international grades (Schmidt, 2008). As a general rule, it can be said that they deliver only conditioned exploration about the actual bearing capacity and the serviceability of the structure. Further estimates of the remaining service life of a structure are usually based exclusively on experience, depending on the assessment of the concrete surface condition.

In Austria an inspection routine which is subdivided into inspection, continuing monitoring and examination is the norm. These performance examinations can be carried out either at definite time intervals or after the occurrence of specific circumstances and events. For the general evaluation of the structure, a separate examination will be assigned. Despite all the advantages of these historically developed inspection routines, the following flaws of the system are listed below to raise the awareness:

- The quality of the data gained from inspections may be (partly) faulty or unsuitable for a systematic filling.
- Only in rare cases are inspection data suitable for a quantitative analysis.
- Alternative inspection methods for lifecycle assessment need to be included.
- Quantitative information regarding the efficiency of various inspection techniques is not available in most cases.
- The transfer of the examination results into the optimization of maintenance and repair activities is at best be carried out by modeling random variables.
- The number of structure specific inspection data, crucial for a comprehensive evaluation is very limited.

This particular study focuses on the specific degradation process that was observed in a damaged bridge and simulated by the application of a Cellular Automata approach compared with the gamma process approach. Specifically, the gamma process approach examined structural behavior, like crack formation, bending, and surface strain (stress development), which can be captured by traditional inspection and/or monitoring method. By employing material analysis after the demolition of the examined Neumarkt-bridge in South Tyrol, we were able to examine the degradation process due to corrosion and carbonation of prestressing in very deep detail.

Consequently, this contribution introduces the possibility of the gamma process approach (in relation to visual inspections) as a means of capturing the internal mechanical changes, for example due to prestressing steel corrosion processes. A pronounced correlation between the gamma process approach and the internal mechanical properties of a structure provides the basis for (a) a quantitatively well ascertained remaining service life, (b) the optimization of inspection periods, (c) the identification of critical structural components for the overall condition and consequently (d) cost- efficient maintenance.

2. Gamma Process Approach

Degradation and aging processes of a structure can be described by non-negative and continuous functions. These functions can be characterized by non-negative increments with independent path and variable uncer-

tainty. Furthermore the designated period of time until the occurrence of the observation of this undesirable incidence is mostly related to a considerable uncertainty and dependent on structural behavior.

Stochastic aging processes of mechanical components are frequently determined via the gamma process which is suitably applied in the analysis of civil engineering structures. Frangopol et al & Noortwijk recommended this approach for the assessment of deterioration processes and analyzed its suitability for engineering structures.

2.1. CHARACHTERISTICS OF A STOCHASTIC PROCESS

The reliability of assessment methods for engineering structures such as bridges, are usually based on assumptions of incomplete information on for instance, the material properties, the quality of the construction and the relevant loads. As a result of this insufficient information, the life time distributions which are derived from a very low failure rate are often contradictory and cannot be formulated accurately. Consequently, in structural engineering in general time dependent and highly uncertain properties/processes, such as an average deterioration per unit time are often considered as random quantities. For this, the suitably applicable process is the class of Markov processes as a class of stochastic processes which represents independent increments. Markov processes enable discrete and continuous processing (Noorwijk, 2009). For example Brownian motion with drift, Poisson Levy and Gamma process differ. The discrete stochastic model is, generally speaking, not suitable for determining the deterioration process in the field of engineering. Consequently, continuous gamma processes, as analyzed by Pandey et al are much better suited to this end. VanNoorwijk proved suitability of gamma processes illustrating the continuous stochastic process by which the temporal damage accumulation can be represented by small independent increments. In particular the essential positive increments are determined by a gamma distribution with identical scale parameter and a time dependent shape function. Consequently with this type of process, deteriorations such as wear, fatigue, creep, cracking corrosion etc. can be determined. Furthermore the gamma process offers the advantage of providing a relatively simple mathematical description.

In the gamma process modeling, we observe at the first step a random variable X with a Gamma distribution, which is characterized by the shape parameter $\alpha > 0$and the scale parameter $\beta > 0$.

From the above we note that sin $\theta = (x + y)z$ or:

$$Ga\left(x|\alpha,\beta\right) = \frac{\beta^{\alpha}}{\Gamma\left(\alpha\right)} \cdot x^{\alpha-1} \cdot \exp\left(-\beta \cdot x\right), \tag{1}$$

where

$$\Gamma\left(\alpha\right) = \int_{z=0}^{\infty} z^{\alpha-1} \cdot e^{-z} \cdot \mathrm{d}z \tag{2}$$

is the Gamma function for $\alpha > 0$

2.2. MODELING OF GAMMA PROCESSES

Gamma process distributions *Ga* are for different time variables are independent of each other. As a result it is possible to obtain conditional distribution of variable X only on the basis of current observations which are well suited to represent the degradation process of standard structures. Such types of deterioration pre-

diction models of a given structure based on the brief time period of observation, should consider the current status and the past events that preceded the current state. In the prediction models, past deterioration profiles can be used to acquaint the engineer with relevant information. In other words, it should be possible to integrate historical deterioration profiles. However, their knowledge should not be a prerequisite.

The parameter $\alpha(t)$ is a non-decreasing, right continuous with left limits, real valued function for $t \geq 0$, with $\alpha(0) = 0$. The gamma process with shape function $\alpha(t) > 0$ and scale parameter $\beta > 0$ is a continuous-time stochastic process $\{X(t), t \geq 0\}$ with the following properties:

$$\mathrm{P}\left(X(0)=0\right)=1, \tag{3}$$

$$X(\eta)-X(t)=Ga\left(\alpha(\eta)-\alpha(t),\beta\right) \quad \text{for} \quad \eta>t\geq 0, \tag{4}$$

where $X(t)$ is characterized by independent increments. Thus the corresponding probability distribution function of $X(t)$, with the time variable t, in accordance with the characteristics of the gamma process is defined as follows:

$$f_{X(t)}(x)=Ga\left(x,\alpha(t),\beta\right). \tag{5}$$

The corresponding expected value:

$$E\left(X(t)\right)=\frac{\alpha(t)}{\beta} \tag{6}$$

and variance

$$Var\left(X(t)\right)=\frac{\alpha(t)}{\beta^2}. \tag{7}$$

Finally the time variable coefficient of variation is obtained from the ratio of the standard deviation and mean value as follows:

$$COV\left(X(t)\right)=\frac{\sqrt{Var\left(X(t)\right)}}{E\left(X(t)\right)}=\frac{1}{\sqrt{\alpha(t)}}. \tag{8}$$

2.3. GAMMA PROCESS MODELING OF DETERIORATION PROCESSES

The following power law formulation is often suitable for the shape function of degradation gamma processes

$$\alpha(t)=c\cdot t^{b}. \tag{9}$$

This is a standard formulation, which has a linear shape ($b = 1$) for the corrosion of concrete reinforcement, a parabolic ($b = 2$) for the sulphate attack, and a square root ($b = 0.5$) for the diffusion-controlled aging according to Elingwood and Mori. The deterioration rate $X(t)$ at the time t, with $t \geq 0$ can simply be described by the shape function $\alpha(t) = ct^b$ and the scale parameter β. In engineering, in most cases the shape of

the expected deterioration b is known and can be taken as a constant, as pointed out by VanNoortwijk et al. However the random rate of degradation C and the scale parameter β are unknown and need to be determined by seeking experts' advice or applying statistical methods. The three most suited statistical methods are the Maximum Likelihood Method, the Method of Moments and Bayesian Statistics (VanNoortwijk, 2009). The determination of the population parameter by means of statistical sample moments is the simplest approach and provides in general a very good result for the first approximation. Provided that the main parameters, the expected value and the variance of the cumulative deterioration at time t are known, the non-stationary gamma process can be transformed to a stationary gamma process. This can be achieved by means of a monotonic transformation from the real time to the operational time as follows:

$$z(t) = t^b \cdot t(z) = z^{\frac{1}{b}}. \tag{10}$$

Thus the expected value is given as:

$$E\left(X\left(t(z)\right)\right) = \frac{c \cdot z}{\beta}, \tag{11}$$

and the equation of the variance as follows:

$$Var\left(X(t)\right) = \frac{c \cdot z}{\beta^2}. \tag{12}$$

Likewise, the transformation of the inspection times can be carried out, $z_i = t_i^b$, for $i = 1,\ldots,n$. The inspection interval between two inspection times is given as:

$$w_i = t_i^b - t_{i-1}^b \tag{13}$$

and

$$\gamma_i = X_i - X_{i-1} \tag{14}$$

as proposed by Van Noortwijk. The degradation increments γ_i have an approximate gamma distribution with a shape factor *cwi* and a scale parameter β for all $i = 1,2, \ldots, n$. The Method of Moments recommended by Van Noortwijk to estimate the parameters $\hat{c}$ and $\hat{\beta}$ is given by the following formulation:

$$\frac{\hat{c}}{\hat{\beta}} = \frac{\sum_{i=1}^{n}\gamma_i}{\sum_{i=1}^{n} w_i} = \frac{x_n}{t_n^b} \tag{15}$$

and with

$$\frac{\hat{c}}{\hat{\beta}} = \overline{\gamma} \cdot \frac{x_n}{\hat{\beta}} \cdot \left(1 - \frac{\sum_{i=1}^{n} {w_i}^2}{\left[\sum_{i=1}^{n} w_i\right]^2}\right) = \sum_{i=1}^{n}\left(\gamma_i - \overline{\gamma} \cdot w_i\right)^2. \tag{16}$$

The Method of Moments leads to a simplified formulation of parameter estimation and can be applied for the first estimation of the solutions of the probability equations. The intervals w_i can be the shorter or

longer periods between the main inspections This can be of particular interest for the optimized selection of inspection techniques and inspection intervals.

3. Case Study: Neumarkt Bridge

Gamma processes, as it is mentioned in the previous sections, are very suitable for the characterization and capture of information from the visual inspections as well as from monitoring systems. There is substantial interest to use this information for an efficient analysis and assessment of the mechanical changes in the structure, wherever no real solution statements can be provided.

In the present case study, a precast element bridge in South Tyrol was evaluated following RVS guideline. As part of the study we examined concrete, prestressed and reinforcement steels for signs of corrosion, general degradation processes and their correlations, before and during the demolition of the bridge. The changes within the mechanical systems, as for example the changes in cross section of the reinforcement and/or presstressing, were monitored during the last decades of the structure using Cellular Automata analysis time dependent reliability analysis and for comparison, nonlinear model. The structural responses provided by the nonlinear model analysis were compared with the result the gamma process analysis.

3.1. GEOMERTY OF THE NEUMARKT BRIDGE

The Neumarkt Bridge, a three-span bridge constructed from precast elements, crosses the A22, the Italian section of the Brenner highway, between the provincial towns of Neumarkt and Auer in South Tyrol. It exhibits features typical for the region's bridge design. The four V-shaped precast elements mounted side by side carry a thin concrete slab. In Figure 1 the most important dimensions of the bridge are illustrated with regard to elevation and transverse section. The main span of the bridge is 27.0 m long, and the outer spans are 9.14 m each. Each V-shaped girder of the main span was mounted with joints in the longitudinal axis as well as in the transverse axes. The girders were secured against side swaying by a 0.14 m strong concrete slab. As a result, the structure is identified as orthotropic with main parts of V-shaped girders. The bridge has two traffic lanes, each 3.75 m of wide and two sidewalks with a width of 1.0 m each.

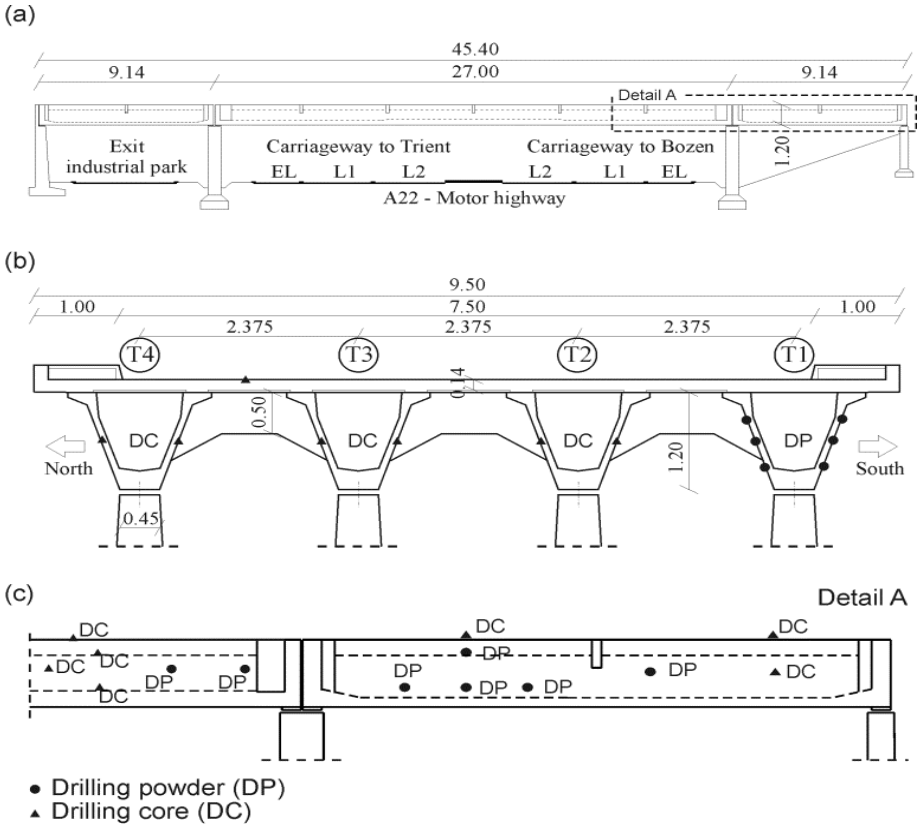


Figure 1. Geometry of the Neumarkt Bridge in South Tyrol: (a) longitudinal section, (b) section of the V-shaped girders of precast elements, and (c) sampling plan

Table I. Time variable statistical description values of pitting corrosion induced by prestressing

t_P [1)] [Year]	Corrosion depth [mm] and/or A/A_o		Remaining cross sectional area of one of the lower prestressing wire, $A_r(t)$ [mm²]			
			PDF	Mean	Hrs	COV
5 (3.5) [2)]	0.35	0.95	LN	126.54	5.07	0.040
10 (7.0)	0.70		LN	125.98	5.07	0.040
15 (10.0)	1.05	0.80	LN	125.06	5.07	0.041
20 (13.5)	1.39		LN	123.83	5.06	0.041
25 (17.0)	1.74		LN	122.25	5.06	0.041
30 (20.5)	2.09	0.60	LN	120.35	5.06	0.042
35 (24.0)	2.44		LN	118.14	5.05	0.043
40 (27.5)	2.78		LN	115.74	5.04	0.044
45 (31.0)	3.13		LN	112.96	5.03	0.045
50 (34.5)	3.50		LN	109.75	5.02	0.046

[1)] propagation time after initiation of corrosion
[2)] propagation time due to corrosion in 3 layers

The condition evaluation of the Neumarkt Bridge, taking into consideration the structural degradation was divided into the discrete steps of data acquisition, analysis simulation of chloride contamination and estimation of time dependent steel corrosion. In addition, the computation of the structural response, taking into account the expected steel section reduction and finally the estimation of the resulting safety level and/or prediction of remaining service life were performed.

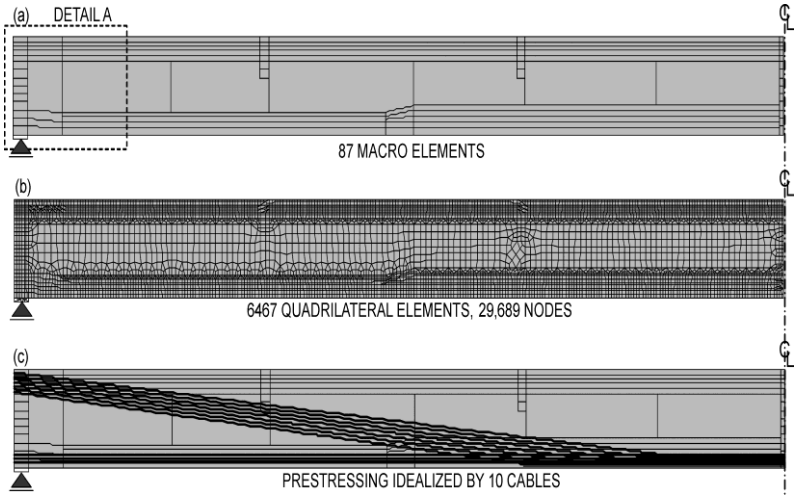


Figure 2. Finite element model to evaluate the SLS and ULS in the center of the beam for a cross section reduction of the prestressing wire in the center line (CL), (a) structuring of makro-elements, (b) FEM mesh generation, und (c) tendon layout of V- girders.

Table II. Limit states of the nonlinear model analysis

Limit state	Action S	Load Comb.	Barrier R	Unit
Material associated limit states				
G_1	Concrete comp. stress σ_c	QP	$0.45\,f_{ck}$	MPa
G_2	Concrete comp. stress σ_c	C	$0.60\,f_{ck}$	MPa
G_3	Mild steel stress σ_a	C	$0.80\,f_{yk}$	MPa
G_4	Mild steel stress σ_a	C	$1.00\,f_{yk}$	MPa
G_5	Pre-stressing steel stress σ_p	C	$0.75\,f_{pk}$	MPa
Deformation associated limit states				
G_6	Vertical deflection u	QP	$l/500$	mm
G_7	Vertical deflection u	FC	$l/250$	mm
G_8	Crack width w	QC	0.2	mm
G_9	Crack width w	C	0.3	mm

QP = Quasi permanent combination
C = Characteristic combination
FC = Frequent combination
f_{ck} = Characteristic concrete compressive strength
f_{yk} = Characteristic yield strength (reinforcement steel)
f_{pk} = Characteristic yield strength (prestressing steel)
l = Span width

The actual implementation of the steps outlined above was performed using the software package SARA, a program that allows the simulation of time dependent chloride ingression (CATES) and the FREET-D program (FREET-D, Teply et al., 2006) which was used to describe the degradation processes as for instance steel corrosion and carbonization induced by controlled inputs. The evaluation of the structural response for the degraded structure was performed via the nonlinear finite element software environment ATENA (Cervenka et al., 2011) on the basis of the fracture mechanical methods, where the generation of inputs and the evaluation of the limit state equation are accomplished by FREET (Novak et al., 2008).

The whole process predicting the chloride concentration up to reliability level for the discrete time t_i was carried out several times. Table I shows the corrosion progress from the time of corrosion occurrence in the lower prestressing position achieved by Cellular Automata Simulation. Accordingly a nonlinear reliability analysis was conducted for the limit states indicated in Table II in accordance with the FEM-Model shown in Figure 2.

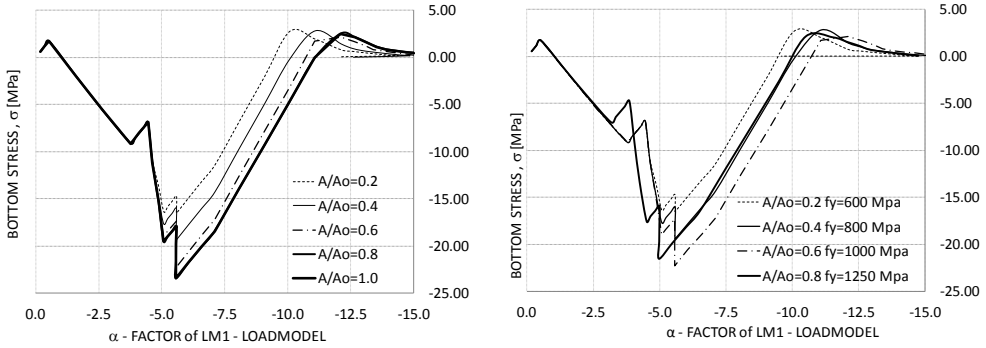


Figure 3. The course of the lower concrete stress- load model for the prestressing steel section lose A/A_o in the center of the beam (a) without and (b) with the consideration of the yield strength f_y reduction

The nonlinear probabilistic FEM analysis delivers on one hand the continuous process of the structural response with regard to the load application for different degree of deterioration, as shown for example in Figure 3 where the process of monitored concrete stress on the V-girder lower side in the center of the beam is illustrated and in Figure 4 where the process of the monitored bending in the center of the beam is shown. On the other hand the statistical characteristics of the structural response and also the probability of failure with regard to the defined limit states for the current and also for the future conditions (see Table III) need to be taken into consideration. These statistical structural responses can be captured by means of visual inspection and/or monitoring system and are consequently a link between the gamma process based description of the deterioration process and the assessment of the structural mechanical changes.

3.2. GAMMA PROCESS BASED CONDITION ASSESSMENT

For the lifetime condition assessment and the illustration of the time dependent structural deterioration, the evolution or progression of deterioration over time is modeled by gamma processes. In the following structural life time response modeling in relation to deflection is conducted for 80% of the LM1 load model. Throughout the analysis the independent deterioration increments are characterized by gamma distribution function with different shape and scale parameters. As a result, the deterioration profile at different ages of the structure was defined and visualized. The gamma process computation was conducted for predictions

that were based on inspections at the ages of 30 and 50. It may be worth mentioning that any estimation of parameters during the early years is best established by experts, as the Method of Moments provides unreliable results in the early lifetime of a structure. Table IV illustrates the gamma process prediction starting with the age of 30.

Table III. Statistical characteristics of the structural response and the corresponding safety levels for a) serviceability limit states and b) ultimate bearing capacity evaluated for 60%, 83% of the LM1 load model

Characteristica[1)]	Time t_P[4)] [years]	R Mean	R COV	S(60% LM1) Mean	S(60% LM1) COV	S(60% LM1) β[8)]	S(83% LM1) Mean	S(83% LM1) COV	S(83% LM1) β[8)]
Deflection,	0	0.11[5)]		− 0.07	0.04	+	− 0.09	0.04	5.1
u_z [m][2)]	30	0.11[5)]		− 0.07	0.04	+	− 0.10	0.05	2.0
	50	0.11[5)]		− 0.08	0.04	8.7	− 0.12	0.05	−
Crack	0	0.20[6)]		0.07	0.17	+	0.06	0.11	+
width	30	0.20[6)]		0.05	0.20	+	0.07	0.14	+
w [mm] [2)]	50	0.20[6)]		0.01	0.11	+	0.02	0.08	+
Concrete	0	18.00[7)]		− 14.12	0.01	+	− 19.02	0.01	−
stress[2)]	30	18.00[7)]		− 15.34	0.01	+	− 20.25	0.01	−
σ_{co}[MPa]	50	18.00[7)]		− 17.10	0.01	7.5	− 21.83	0.01	−
Bearing	0	48.30	0.08	20.00	7.3		30.00	4.7	
Capacity	30	45.10	0.01	20.00	5.6		30.00	3.3	
[load step]	50	39.00	0.06	20.00	8.1		30.00	3.8	

[1)] all variables normal distribution
[2)] serviceability limit state (SLS)
[3)] load level interpolated for 100% of LM1
[4)] time after corrosion initiation
[5)] $_{z,limit} = l/250$ according to [38], 7.4.1
[6)] $w_{,limit} = 0.2$ mm according to [38], 7.3.1
[7)] $\sigma_{,limit} = 0.6\, f_{ck}$ according to [38], 7.2
[8)] $\beta > 10$ is indicated by "+", $\beta > 1$ by "– "

Table IV. Gamma process prediction of statistical characteristics of structural response; evaluated for 83% of the LM1 load model

Time t_P[4)] [years]	S (83% LM1) Mean	COV	β	C	$\alpha(t)$
Deflection,	0	0.05	0.04	-	-	
	30*	**0.07**	**0.04**	**13.12**	**0.044**	**1.312**
u_z	50	0.09	0.04	**34.22**	**0.082**	**4.107**
[m][2)]	*35*	*0.08*	*0.08*	*1.764*	*0.006*	*0.206*
	40	*0.09*	*0.07*	*3.086*	*0.010*	*0.412*
	50	*0.12*	*0.06*	*5.967*	*0.020*	*0.995*
	60	*0.14*	*0.06*	*8.573*	*0.029*	*1.715*
	70	*0.16*	*0.05*	*10.87*	*0.036*	*2.538*
	80	*0.21*	*0.05*	*15.32*	*0.051*	*4.595*
	100	*0.23*	*0.05*	*17.49*	*0.058*	*5.830*

4. Conclusion

Within the current research project, three different testing methods were used to determine the material parameters of interest in experiment. For the assessment of the material characteristics of concrete, the compressive strength fc, the tensile strength fct and the fracture energy Gf were considered. Small deviations between the Vienna test results and the results of tests performed at TU Brno arose. Several of the stochastic concrete parameters, which were characterized by the Brno team, were modeled numerically based on data from three-point bending tests. In particular, the comparison of the fracture energy obtained from the three-point bending test with the result of the wedge splitting test revealed only minor divergences. These results allow the conclusion that all three test methods are reliable, comparable which each other, and provide consistent stochastic concrete methods. In addition to the verification of the test methods, the evaluation of the influence of concrete additives on the stochastic models were of interest.

Acknowledgements

This research was performed with the financial support of AUTOBRENNERO, ASFINAG and BMVIT as part of the research project "Monitoring und Analyse von integralen Brückenbauwerke" and with the financial support of FFG during the EUREKA Eurostar research project "Risk Lifetime Assessment of Concrete Structures".

The opinions and results published in this article are those of the authors and do not necessarily reflect those of the supporting organisations.

References

Cervenka, V., L. Jendele and J. Červenka. 2011. *ATENA Program documentation.* Cervenka Consulting, Czech Republic, 2011.

Ellingwood, B.R. and Y. Mori. Probabilistic methods for condition assessment and life prediction of concrete structures in nuclear power plants. *In: Nuclear Engineering and Design.* Vol. 142, Issues 2-3, pp.155-166, 1993.

Frangopol, D.M., Maatren J. Kallen and M. van Noortwijk. Probabilistic models for life-cycle performance of deteriorating structures (review). *In Structural Engineering Materials.* John Wiley & Sons: Vol. 6, pp, 197-212. 2004.

Novak, D., M. Vrechovsky and R. Rusina. *Feasibility Reliability Engineering Tool, FReET Program Documentation,* Cervenka Consulting, Czech Republic 2008.

Ohadi, A., T. Micic. Stochastic process deterioration modeling for adaptive inspections. *Applications of statistics and Probability in Civil engineering – Faber, Köhler & Nishijima,* Taylor & Francis group, London, pp. 1085-1091. 2011.

Pandey, M.D., X.X. Yausn and J.M. van Noortwijk. The influence of temporal uncertainty of deterioration on life-cycle management of structures. *In: Structure and Infrastructure Engineering.* Vol, 5(2); pp. 145-156, 2009

Schmidt, H. *Modelle der Baustoffwissenschaft. Hamburg.* Dissertation, 2008

Schmidt, H. and F.Sondermann. Baustoffe der Natur. *In: Karlstädter Bauseminar,* Karlstadt: self-published, 2006.

Schmidt, H. and F. Sondermann. Natur der Baustoffe. *In: Tölping, E. (ed.).* Internationale Baustoffwissenschaft: Hannover. 2007.

Schmidt, H. and F. Sondermann. *Buch der Baustoffe.* Part 7: Minerale. Bielefeld: Bauverlag, 2010.

Schmidt, H. and F. Sondermann. Zu den Grundlagen der Baustoffe. *In Baustoff-Journal 12.* S.10-20, 2011.

Teplý, B., M. Chromá, D. Matesová, P. Rovnaník, and V. Veselý. *Feasibility Reliability Engineering Tool for degradation effects assessment FReEt-D.* Cervenka Consulting, Czech Republic, 2003.

van Noortwijk J.M, J.A.M van Weide, M.J. Kallen and M.D. Pandey. Gamma process and peaks - over threshold distributions of time dependent reliability. *In Engineering and System Safety.* 92, pp. 1651-1658, 2007.

van Noortwijk, J.M. A survey of the application of gamma processes in maintenance. *In Reliability Engineering and Safety.* Vol. 94, pp. 2-21, 2009.

Durability Assessment of Large Surfaces Using Standard Reliability Methods

M. Sykora and M. Holicky
Czech Technical University in Prague, Klokner Institute,
miroslav.sykora@klok.cvut.cz, milan.holicky@klok.cvut.cz

Abstract: It is well recognized that initiation limit states defined in (ISO 13823, 2008) may be of uttermost importance for serviceability as well as ultimate limit states of civil engineering structures. However, practical applications of durability assessments may be difficult as basic variables influencing structural durability are often random quantities with a considerable spatial variability that should be considered as random fields. Application of common discretisation techniques may be rather cumbersome and require a considerable amount of input data. A simplified probabilistic model for spatial variation is thus proposed to allow for durability analysis of large surfaces using efficient reliability methods such as FORM/SORM. The technique is applied in the example of carbonation of concrete where spatial variation of the carbonation depth and concrete cover is considered. It appears that the failure probability increases with the size of surface exposed to unfavourable environmental influences. Optimisation study further indicates that the total costs primarily depend on the thickness of concrete cover, design service life, and the surface area exposed to the deterioration. However, the optimum concrete cover and optimum reliability index seem to be almost independent of the size of the surface area.

Keywords: durability, random fields, discretization, FORM.

1. Introduction

Durability is becoming an important issue of structural design. General principles on the probabilistic approach to verification of structural durability are provided in the new international standard (ISO 13823, 2008). The document is based on the fundamental principles provided in (ISO 2394, 1998), (ISO 19338, 2003) and (EN 1990, 2002). Materials of other international organisations such as CEB, fib, RILEM and findings in scientific publications have also been taken into account.

(Holicky, 2011) indicates that due to limited experience with the operational use of (ISO 13823, 2008), additional studies focused primarily on models of material deterioration, acceptance criteria, and theoretical models of basic variables are required. Difficulties in practical applications may arise particularly when basic variables influencing structural durability have a considerable spatial variability (e.g. for large surfaces concentrations of unfavourable agents or diffusion properties of construction materials). In probabilistic analyses the spatial variability is normally described by random fields. Application of common discretisation techniques, see e.g. (Allaix et al., 2009), may be rather cumbersome and may require a considerable amount of input data.

In the present study a simplified probabilistic model for spatial variation is thus proposed to allow for durability analysis of large surfaces using efficient reliability methods such as FORM/SORM. The

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

technique is applied in the example of carbonation of concrete where spatial variation of the carbonation depth and concrete cover is considered.

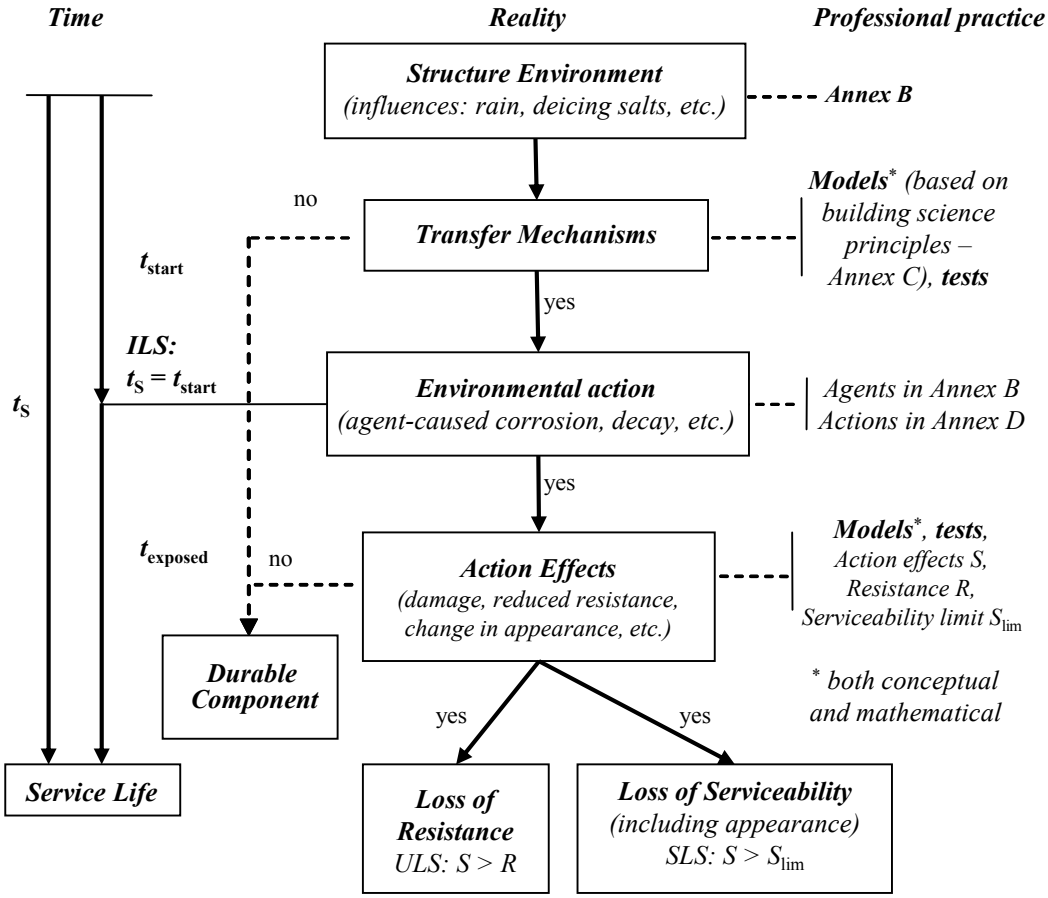


Figure 1. Limit state method for durability (accepted from (ISO 13823, 2008)).

2. Concept of Limit States

(ISO 13823, 2008) formulates the principles of limit state methods for durability. The key steps of deterioration modelling and reliability verification using the concepts of limit states are indicated in Figure 1. It provides a very general scheme that may be modified considering actual conditions of an investigated structure. It should be noted that Figure 1 is a result of many discussions and amendments made during the development of (ISO 13823, 2008).

The three vertical strands in Figure 1 indicate a time axis (on the left), reality (in the middle) and professional practice (on the right). The time axis is split into two parts by the point denoted as the Initiation Limit State (ILS). It corresponds to the point in time when environmental actions have turning point (for example the beginning of reinforcement corrosion or decays of construction materials).

The environmental effects may in general be combined with the action effects (the middle part of Figure 1). Resulting effects may then lead to the loss of resistance (bearing capacity) or to the loss of serviceability (excessive cracking or deformations). These limit states - ULS and SLS - are indicated in the lower part of Figure 1. However, an important question of load combination rules is not covered in (ISO 13823, 2008).

3. Verification of the Service Life

The fundamental durability requirement is represented by a simple condition that the predicted service life t_{SP} should be greater than the design service life t_D with a sufficient degree of reliability. Difficulties are obviously linked to the term „sufficient reliability". It is well recognised that the service life t_S is dependent on a number of basic variables and is consequently a random variable having a considerable scatter. The document (ISO 13823, 2008) thus provides a probabilistic formulation of this criterion:

$$\mathrm{P}[t_S < t_D] < P_{target} \tag{1}$$

where P_{target} denotes the target probability of the service life t_S being less than the design service life t_D. Commonly the design service life t_D is a deterministic quantity (for example 50 or 100 years) specified in advance.

4. Verification of the Limit States

The probabilistic formulation of the limit state conditions is similar to a case of the service life. For an arbitrary point in time $t \leq t_D$ the following condition should be satisfied:

$$P_f(t) = \mathrm{P}[Z(t) < 0] = \mathrm{P}[R(t) - S(t) < 0] < P_{target} \tag{2}$$

where $P_f(\cdot)$ denotes the failure probability; $Z(\cdot)$ = reliability margin; $R(\cdot)$ = resistance; and $S(\cdot)$ = action effect. The basic probabilistic condition for the serviceability can be written analogously as:

$$P_f(t) = \mathrm{P}[Z(t) < 0] = \mathrm{P}[S_{lim} - S(t) < 0] < P_{target} \tag{3}$$

where S_{lim} is the limit value of the serviceability indicator (for example of the crack width or deflection). The initiation limit state may be verified in accordance with Eqs. 2 or 3 depending on particular conditions.

5. Assessment of the Service Life

The probabilistic assessment of the predicted service life t_{SP} is schematically shown in Figure 2 adopted from (ISO 13823, 2008). Figure 2 describes monotonously varying action effects $S(t)$ and resistances $R(t)$. The horizontal axis denotes the time t and the vertical axis in the upper part denotes the resistance $R(t)$ and action effect $S(t)$, in the lower part the probability $P_f(t)$. Probability distributions of the variables $R(t)$ and $S(t)$ are indicated by probability density functions.

Obviously the failure probability $P_f(t)$ is an increasing function of time t. The predicted service life t_{SP} follows from the relationship:

$$P_f(t_{SP}) = P_{target} \tag{4}$$

However, there are no recommendations concerning the target probability P_{target} provided in (ISO 13823, 2008) and this open question may cause difficulties in the effective use of the document.

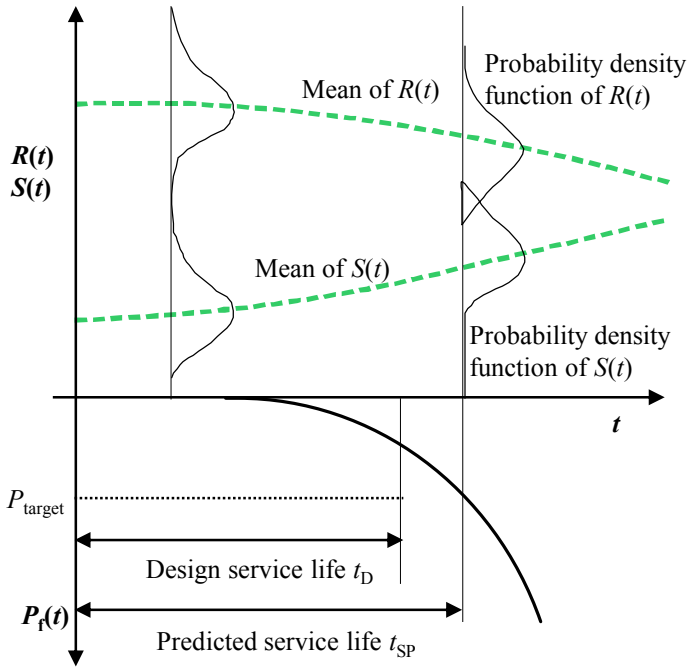


Figure 2. Probabilistic assessment of the service life.

6. Target Reliability Level

Target reliability level, indicated by the target probability P_{target} or reliability index β_{target}, depends in general on the definition of the service life time, whether the critical durability requirement concerns the ultimate limit state, serviceability limit state or initiation limit state and what are the consequences of their infringement (Holicky, 2011). Table I provides indicative intervals for the target reliability.

Table I. Indicative values of the target probability and reliability index.

Limit state	P_{target}	β_{target}
Ultimate limit state - ULS	$\sim 10^{-4}$	~ 3.7
Serviceability limit state - SLS	0.01 to 0.1	1.3 to 2.3
Initiation limit state - ILS	0.05 to 0.2	0.8 to 1.6

The target probability P_{target} and reliability index β_{target} given in Table I represent indicative values only. They are based on the target values recommended in (ISO 2394, 1998) and (EN 1990, 2002). It should be mentioned that (ISO 2394, 1998) indicates an additional dependence of the target values on relative costs of safety measures (required to increase the reliability level). This aspect should be also considered when specifying target reliability level for durability requirements. Specification of the appropriate reliability level remains, therefore, one of the most important open questions.

7. Simplified Model for Spatial Variability

(Faber and Rostam, 2001) suggested that a large surface exposed to deterioration effects should be analysed as an assembly of elementary surfaces rather than a whole structure. Probabilistic characteristics of the variables influencing the deterioration should then include also the spatial variability of the variables among elementary surfaces. For instance several studies focused on reinforced concrete structures, see e.g. (Vu and Stewart, 2005) or (Stewart and Mullard, 2007), reveal that the elementary surface may be a square with the side length varying from 1 to 3 meters. For steel structures the size of an elementary surface may correspond to a size of inspected areas (e.g. 3 m), (Straub, 2004).

The present study is based on the following assumptions:

- The basic (random) variables influencing a given limit state can be divided into random fields $\mathbf{W}(x,y)$ (e.g. some material properties) and variables attaining a single value for the whole structure $\mathbf{X}(t)$ (e.g. some environmental influences).
- Random fields $\mathbf{W}(x,y)$ are homogeneous and can be approximated by N elementary surfaces of the same size. Values of the random fields in each elementary surface, $W_{i,j}$ for $i = 1..n_W$ (number of the random fields) and $j = 1..N$ (number of elementary surfaces), are independent, identically distributed variables (having an appropriate probability distribution based on available data). To simplify the notation, vector of the values of the random fields in an elementary surface j is hereafter denoted as $\mathbf{W} = W_i$ for $i = 1..n_W$.
- Some of the variables $\mathbf{X}(t)$ can be time-dependent. Then they are either monotonously decreasing (when favourably influencing durability) or monotonously increasing (when unfavourable). Consequently the failure probability is monotonously increasing with time.

This simplified model for spatial variation is assumed to yield conservative results compared to standard techniques such as discretization at the centre of gravity or discretization by spatial mean proposed by (VanMarcke, 1983). The failure probability at the elementary surface p_f can be obtained from the following relationship:

$$p_f[\mathbf{W},\mathbf{X}(t)] = P\{Z[\mathbf{W},\mathbf{X}(t)] < 0\} \tag{5}$$

The failure probability at a whole surface can be written as:

$$P_f(t) = P\{n_{deg}[\mathbf{W},\mathbf{X}(t)] / N \geq \alpha_{lim}\} = E_{\mathbf{X}(t)}\{P[n_{deg}(\mathbf{W}|\mathbf{x}(t)) / N \geq \alpha_{lim}]\} \tag{6}$$

where $n_{deg}(\cdot)$ denotes the number of elementary surfaces for which $[Z(\cdot) < 0]$; α_{lim} the limiting value of the deterioration level $\alpha = n_{deg} / N$; $E(\cdot)$ expectation operator; and $\mathbf{x}(t)$ values of the variables $\mathbf{X}(t)$.

Given $\mathbf{x}(t)$, values of the reliability margin in each elementary surface are statistically independent. The probability of occurrence of ν "failed" elementary surfaces out of N is thus given by the binomial distribution, see (Faber and Rostam, 2001), (Malioka et al., 2011) and (Sýkora and Holický, 2011):

$$P\{n_{deg}(\mathbf{W}|\mathbf{x}(t)) = \nu\} = f_{binom}\{\nu, N, p_f[\mathbf{W}|\mathbf{x}(t)]\} \quad (7)$$

where $f_{binom}(\cdot)$ is the probability density function of a binomial distribution. Note that the number $n_{deg} = \nu$ actually represents the probability of ν successes out of N independent trials with the probability of success p_f. The failure probability (6) can then be modified as:

$$P_f(t) = E_{\mathbf{X}(t)}\{1 - F_{binom}[N\alpha_{lim}, N, p_f(\mathbf{W}|\mathbf{x}(t))]\} \quad (8)$$

where $F_{binom}(\cdot)$ is the cumulative distribution function of the binomial distribution.

The use of F_{binom} significantly decreases computational demands since the assessment of spatial variability simply reduces to evaluation of the cumulative distribution function of the binomial distribution. Note that the binomial distribution may be approximated by a normal distribution for, say, $N > 50$. The expectation in Eq. 8 can be carried out by the FORM/SORM methods, see e.g. (Wen, 1990).

The limiting value α_{lim} should be specified by an owner, preferably using cost optimisation and previous experience. As an example (Fitch et al., 1995) suggested $\alpha_{lim} = 0.12$ for corrosion-induced cracking of reinforced concrete bridges while $\alpha_{lim} = 0.2$ was considered in a general study by (Faber and Rostam, 2001).

It is emphasized that the proposed model of spatial variability may be oversimplified when the random fields need to be associated with different areas for which their values can be considered as independent. In such a case it would be necessary to modify Eqs. 5, 7 and 8. However, it is foreseen that the proposed approximation can be applied in a number of practical cases.

8. Numerical Example

8.1. Deterioration Model

The initiation limit state can be well illustrated by the carbonation of concrete. The limit state may be defined as a simple requirement that the carbonation depth $S(t)$ (action effect) is less than the concrete cover R (resistance). Note that it may be more suitable to define the failure considering an indicator that can be verified by visually (such as crack width).

A large, vertical concrete surface is investigated. Concrete cover R and inverse carbonation resistance under natural carbonation conditions $R_{NAC,0}^{-1}$ are assumed to be spatially variable. The size of an element is assumed to be 0.5 m in accordance with (Vu and Stewart, 2002) and (Malioka, 2009). The variables $\mathbf{X}$ and the model uncertainty of action effect K_S are time-independent. Notation and probabilistic models of the basic variables are given in Table II.

Given values $\mathbf{x}$ and k_S, the failure probability at an elementary surface is determined as follows:

$$p_f(\mu_R, t|k_S, \mathbf{x}) = P[R(\mu_R) - k_S S(R_{NAC,0}^{-1}, t|k_S, \mathbf{x}) < 0] \quad (9)$$

where μ_R denotes the mean of the concrete cover (nominal value – study parameter).

The point-in-space carbonation depth is described in accordance with (fib, 2006):

$$S\left(R_{\mathrm{NAC},0}^{-1}, t \middle| rh_{\mathrm{real}}, c_{\mathrm{s}}, b_{\mathrm{w}}\right) = \sqrt{t}\sqrt{2\left(\frac{1-rh_{\mathrm{real}}{}^{5}}{1-rh_{\mathrm{ref}}{}^{5}}\right)^{2.5}\left(\frac{t_{\mathrm{c}}}{7}\right)^{b_{\mathrm{c}}} R_{\mathrm{NAC},0}^{-1} c_{\mathrm{s}}}\left(\frac{0.0767}{t}\right)^{\frac{(p_{\mathrm{SR}} tow)^{b_{\mathrm{w}}}}{2}} \tag{10}$$

where t is time in years. Note that in Eqs. 9 and 10, the values of the random fields R and $R_{\mathrm{NAC},0}{}^{-1}$ are denoted by capital letters while values of the random variables and deterministic quantities are denoted by small letters.

Table II. Probabilistic models of the basic variables.

Type	Variable / random field	Symbol	Distribution	Unit	μ_X	V_X	Ref.
Random fields	Concrete cover	R	Beta (lower bound = 0, upper bound ≈ $3\mu_R$)	mm	μ_R	0.35	(Holický and Holická, 2006)
	Inverse carbon. resistance under natural carbonation	$R_{\mathrm{NAC},0}{}^{-1}$	Gamma	[(mm^2/year) / (kg/m^3)]	2×10^4	0.5	(fib, 2006)
Random variables	Relative humidity	RH_{real}	Beta	-	0.71	0.18	nearest weath. station
	CO_2 concentration	C_{s}	normal	kg/m^3	8.2×10^{-4}	0.12	(fib, 2006)
	Regression coefficient	B_{w}	normal	-	0.45	0.37	(fib, 2006)
	Model uncertainty	K_S	LN	-	1	0.1	-
Determ. variables	Refer. relative humidity	rh_{ref}	-	-	0.65	-	(fib, 2006)
	Curing period	t_{c}	-	day	5	-	-
	Regression coefficient	b_{c}	-	-	-0.57	-	(fib, 2006)
	Prob. driving rain	p_{SR}	-	-	0.4	-	nearest weath.
	Time of wetness	tow	-	-	0.27	-	station

The model for relative humidity is based on daily mean values. Probability of driving rain for vertical surface (facing to the west here) is determined from the distribution of wind directions during rain events. Time of wetness is assessed from the average number of days per year for which daily precipitation total exceeds 2.5 mm (in this case 100 days per year). Contrary to the recommendations of (fib, 2006), the regression coefficient b_{c} is considered here to be deterministic since numerical experience indicates that its variability is negligible.

The considered model for the carbonation depth has been calibrated against extensive measurements on cooling towers (unprotected external concrete) described by (Holický and Holická, 2006). Basically, the presented model leads to a similar mean value and somewhat lower coefficient of variation and skewness as compared to the measurements.

8.2. RELIABILITY ANALYSIS

The limiting deterioration level $\alpha_{lim} = 0.15$ is considered. The failure probability $P_f(t)$ is obtained from Eq. 8 by integration over K_S and **X**. In Figure 3 the failure probability $P_f(t)$ is indicated for $\mu_R = 25$ mm and $N = 1$, 20, and 100. In addition results based on a hypothetical assumption according to which correlation amongst the elementary surfaces is neglected and all the random variables are spatially variable, are plotted for the number of elementary surfaces $N = 100$.

It appears that the failure probability depends significantly on the number of elementary surfaces. The assumed model predicts significantly lower failure probabilities for $N = 1$ than for $N = 10$ or 100. Similarly as concluded by (Stewart, 2004), it follows that the spatial variability should be appropriately considered particularly when analysing large surfaces.

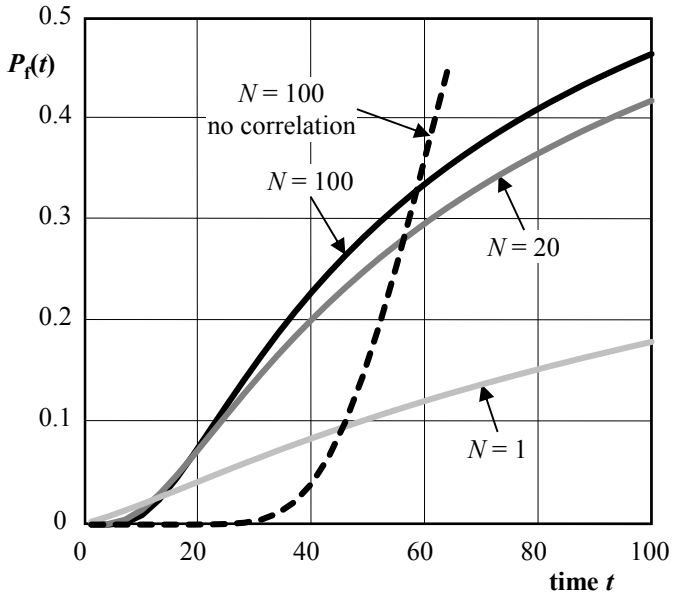


Figure 3. Variation of failure probability $P_f(t)$ with time t for $\mu_R = 25$ mm and $\alpha_{lim} = 0.15$.

Further Figure 3 shows that misleading results may be obtained when the correlation amongst the elementary surfaces is neglected. In this case, the failure probability is very low for $t < 30$ years and then significantly increases.

Figure 3 can be used to assess the service life t_{SP} defined by Eq. 4 for a specified target probability P_{target}, the mean of concrete cover μ_R and number of elementary surfaces N. If for example $P_{target} = 0.15$, then the mean $\mu_R = 25$ mm corresponds to $t_{SP} \approx 30$ years for $N = 20$ and 100, but for $N = 1$, $t_{SP} \approx 80$ years is estimated. Obviously, the service life t_{SP} appears to be significantly dependent on the number of elementary surfaces and on the target probability.

9. Probabilistic Optimisation

Methods of probabilistic analysis may be effectively used for the specification of the target reliability level and durability assessment, (Holicky, 2009) and (Holicky, 2011). The total costs of execution and repair of the structure due to failure (infringement of the initiation limit state) can be expressed as a function of the mean μ_R (decisive parameter):

$$C_{\text{tot}} = C_0 + C_1 \mu_R + \text{E}[C_{\text{f}}] \tag{11}$$

where C_0 denotes the initial costs independent of μ_R; C_1 the cost of a unit of μ_R; and $\text{E}[C_{\text{f}}]$ = expected expenses related to the durability failure given by:

$$\text{E}[C_{\text{f}}] = C_{\text{f}} \int_t \pi_{\text{f}}(\tau) \text{d}\tau \tag{12}$$

where C_{f} denotes a present value of the expected expenses related to the durability failure; q annual discount rate (around 0.03); and $\pi_{\text{f}}(\cdot)$ the discounted conditional failure rate given by:

$$\pi_{\text{f}}(t) = P_{\text{f}}(t)' \,/\, \{[1 - P_{\text{f}}(t)](1 + p)^t\} \tag{13}$$

where $P_{\text{f}}(t)'$ is the time derivative of the failure probability given in Eq. 8. Standardised total cost is considered as:

$$\kappa_{\text{tot}} = [C_{\text{tot}} - C_0] \,/\, C_1 = \mu_R + C_{\text{f}} \,/\, C_1 \int_t \pi_{\text{f}}(\tau) \text{d}\tau \tag{14}$$

The optimum mean $\mu_{R,\text{opt}}$ may be determined from:

$$\partial \kappa_{\text{tot}} \,/\, \partial \mu_R = 0 \tag{15}$$

Note that within the realistic domain of μ_R from 20 to 70 mm, Eq. 15 may not have a practical solution and the minimum of the total costs may not be attained.

Considering the above described initiation limit state, the standardised total costs κ_{tot} given by Eq. 14 are shown in Figure 4 assuming the design life time t = 40 years (typical for cooling towers), q = 0.03 and N = 100. In addition variation of the failure probability P_{f} with μ_R is also indicated. It appears that the optimum mean $\mu_{R,\text{opt}}$ considerably increases with increasing cost ratio $C_{\text{f}} \,/\, C_1$. More specifically, it follows that:

- For $C_{\text{f}} \,/\, C_1 = 10$ (“small” failure consequences or “high” unit costs), the optimum $\mu_{R,\text{opt}}$ is not attained in the practical range of μ_R.
- For $C_{\text{f}} \,/\, C_1 = 100$ (“medium” failure consequences and “medium” unit costs), the optimum mean is $\mu_{R,\text{opt}} \approx 29$ mm ($\beta_{\text{opt}}(\mu_{R,\text{opt}} = 29 \text{ mm}) \approx 1.1$).
- For $C_{\text{f}} \,/\, C_1 = 1\,000$ (“high” failure consequences or “small” unit costs), then the optimum mean increases up to $\mu_{R,\text{opt}} \approx 43.5$ mm ($\beta_{\text{opt}}(\mu_{R,\text{opt}} = 43.5 \text{ mm}) \approx 2.4$).

Additional parameter study reveals that the optimum mean concrete cover $\mu_{R,\text{opt}}$ is nearly independent of the number of elementary surfaces N.

Variation of the optimum (target) reliability index β_{opt} (based on $\mu_{R,\text{opt}}$) with the number of elementary surfaces N is shown in Figure 5 for the design life time $t = 40$ years, $q = 0.03$ and $C_\text{f} / C_1 = 100$ and 1 000. It follows that β_{opt} insignificantly increases with increasing N. In the first approximation the values $\beta_{\text{opt}} \approx 1.1$ ($C_\text{f} / C_1 = 100$) and $\beta_{\text{opt}} \approx 2.4$ ($C_\text{f} / C_1 = 1\,000$) may be considered.

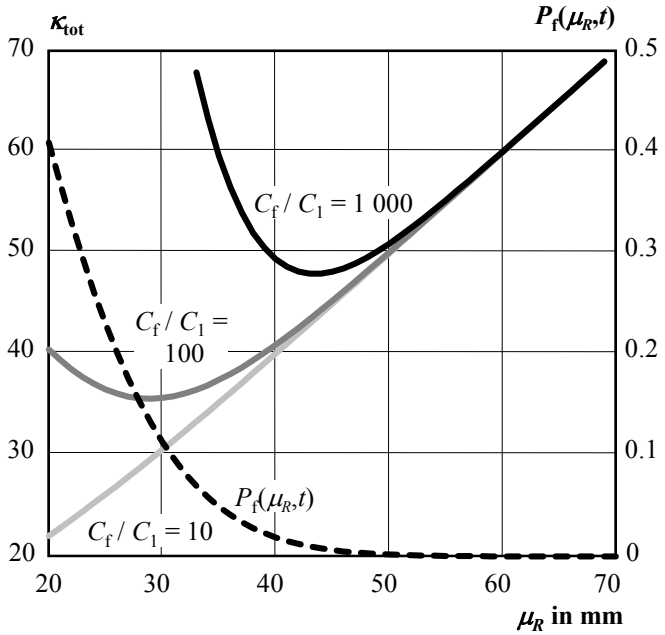


Figure 4. Variation of the total standardised costs κ_{tot} and failure probability P_f with the mean concrete cover μ_R for $q = 0.03$, $t = 50$ years and $N = 100$.

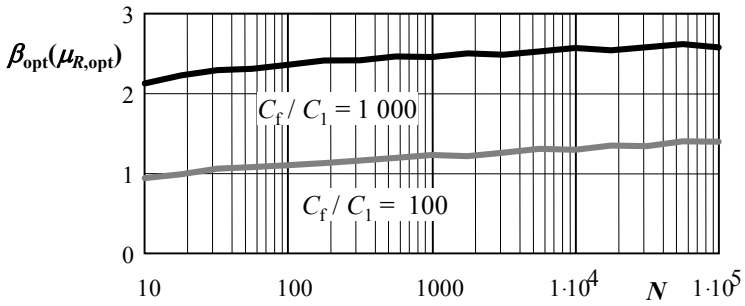


Figure 5. Variation of the optimum reliability index β_{opt} with the number of elementary surfaces N for $q = 0.03$, $t = 40$ years, and $C_\text{f} / C_1 = 100$ and 1 000.

10. Concluding Remarks

Structural durability is becoming an important part of structural design of buildings and other civil engineering works. It may be significantly affected the spatial variability particularly for large surfaces. Simplified model for deterioration of large surfaces, proposed here, seems to require less input data and significantly lower computational demands compared to random field techniques. It is foreseen that this model can be effectively used for optimisation studies when structural durability need to be assessed for various decision parameters.

Numerical example, focused on the carbonation of concrete, reveals that the failure probability increases with the size of surface exposed to unfavourable environmental influences. Optimisation study indicates that the total costs particularly depend on the thickness of the concrete cover, design service life, and the size of a surface area exposed to the deterioration. However, the optimum concrete cover and optimum reliability index seem to be almost independent of the size of the surface area. As a first approximation the optimum concrete cover of 30 mm and optimum reliability index of 1.1 may be recommended for the required design life of 40 years, discount rate 0.03 and the cost ratio $C_f / C_1 = 100$.

Further experimental data and appropriate models for the carbonation process, related model uncertainties and initial and failure costs are needed. Further research should be focused on the comparison of standard random field approaches with the proposed simplified model.

Acknowledgements

The study has been conducted in the framework of the research projects P105/12/G059 and P105/12/0589 supported by the Czech Science Foundation.

References

Allaix, D.L., Carbone, V.I. and Mancini, G. Recent advances in the modelling of space-variant structural properties. In H. Furuta, D.M. Frangopol & M. Shinozuka, editors. *Proc. ICOSSAR 2009,* Osaka, 13-17 September 2009. CRC Press/Balkema, Leiden, 2009.

EN 1990. *Eurocode - Basis of structural design.* CEN, Brussels, 2002.

Faber, M.H. and Rostam, S. Durability and Service life of Concrete Structures – the Owners Perspective. Anonymous *Proc. Int. IABSE Conf. on Safety, Risk, and Reliability - Trends in Engineering,* Malta, 21-23 March 2001. IABSE, 2001.

fib. *Model Code for Service Life Design.* fib, Lausanne, 2006.

Fitch, M.G., Weyers, R.E. and Johnson, S.D. Determination of end of functional service life for concrete bridge decks. *Transp. Res. Rec.,* (1490):60-66, 1995.

Holický, M. and Holická, N. Probabilistic optimisation of concrete cover exposed to carbonation. In K. Kovler, editor. *Proc. ConcreteLife'06,* Tallin, Estonia, RILEM Publications SARL, 2006.

Holicky, M. Forensic study of structures for durability. *Safety and Security Engineering III,* 1083-12, 2009.

Holicky, M. Forensic engineering of structures with durability problems. *Proceedings of the Institution of Civil Engineers-Structures and Buildings,* 164(1):19-25, 2011.

ISO 13823. *General principles on the design of structures for durability.* ISO, Geneve, Switzerland, 2008.

ISO 19338. *Performance and assessment requirements for design on structural concrete.* ISO, Geneve, Switzerland, 2003.

ISO 2394. *General principles on reliability for structures.* ISO, Geneve, Switzerland, 1998.

Malioka, V. *Condition indicators for the assessment of local and spatial deterioration of concrete structures (PhD thesis).* ETH, Zurich, 2009.

Malioka, V., Matsch, S. and Faber, M.H. Acceptance criteria for the spatial variability of the concrete cover depth in a Swiss tunnel. In M.H. Faber, J. Köhler & K. Nishijima, editors. *Proc. ICASP11,* ETH Zurich, 1-4 August 2011. CRC Press/Balkema, Leiden, 2011.

Stewart, M.G. Spatial variability of pitting corrosion and its influence on structural fragility and reliability of RC beams in flexure. *Structural Safety,* 26(4):453-470, 2004.

Stewart, M.G. and Mullard, J.A. Spatial time-dependent reliability analysis of corrosion damage and the timing of first repair for RC structures. *Engineering Structures,* 29(7):1457-1464, 2007.

Straub, D. *Generic approaches to risk based inspection planning for steel structures.* ETH, Zurich, 2004.

Sýkora, M. and Holický, M. Deterioration Model for Large Reinforced Concrete Structures. In C. Berenguer, A. Grall & C. Guedes Soares, editors. *Proc. ESREL 2011,* Troyes, France, 18 - 22 September 2011. CRC Press/Balkema, Leiden, 2011.

VanMarcke, E. *Random Fields: Analysis and Synthesis*. MIT Press, Cambridge MA, 1983.

Vu, K.A.T. and Stewart, M.G. Spatial Variability of Structural Deterioration and Service Life Prediction of Reinforced Concrete Bridges. Anonymous *Proc. 1st International Conference on Bridge Maintenance, Safety and Management,* Barcelona, 14-17 July 2002. IABMAS, 2002.

Vu, K.A.T. and Stewart, M.G. Predicting the Likelihood and Extent of Reinforced Concrete Corrosion-Induced Cracking. *Journal of Structural Engineering,* 131(11):1681-1689, 2005.

Wen, Y.K. *Structural load modeling and combination for performance and safety evaluation.* Elsevier, Amsterdam, 1990.

Probabilistic modelling of concrete structures degradation

B. Teplý, P. Rovnaníková, P. Rovnaník, D. Vořechovská
Faculty of Civil Engineering, Brno University of Technology, Brno, Czech Republic,
teply.b@fce.vutbr.cz, rovnanikova.p@fce.vutbr.cz, rovnanik.p@fce.vutbr.cz, vorechovska.d@fce.vutbr.cz

Abstract: Sulfate attack is a set of complex and overlapping chemical and physical processes caused by reactions of numerous cement components with sulfates originating from external or internal sources. These processes can lead to eventual deterioration of concrete composition and properties; in addition, these reactions depend on the environmental exposure, including access of moisture and temperature changes.
In the context of performance-based approaches, sustainability consideration and whole life costing, suitable models, which can reliably predict the behavior of hydrated cement systems of concrete structures subjected to sulfate attack during the service life, are needed. A consistent approach to the durability assessment of concrete structures is recommended, i.e. fully probabilistic durability design, which necessarily requires the utilization of stochastic approaches.
The present work is restricted to models for the description of concrete under the external sulfate attack. The primary sources of external sulfates are natural sulfates in soils or dissolved in ground water, sulfates originating from the atmospheric pollution or from the decomposition and oxidation of proteins in waste water. Several models for concrete corrosion rate are presented, compared and some parametric analyses are shown.

Keywords: sulfate attack, concrete, modeling, stochastic approach

1. Degradation Modeling

In the context of performance-based approaches for the design or assessment of concrete structures, time is the decisive variable and the durability issues are pronounced. Also, the reliability aspects are important as they is attribute to service life, maintenance, inspection, repair and the life-cycle cost – see also ISO 13823 (2008) and the fib Model Code (2010).

Evidently, durability and reliability are often crucial structural performance characteristics and the reliability level for relevant limit states has to be analysed frequently utilising mathematical models for degradation prognosis. When doing this it can be useful to have a range of models available for the structural task in question. The engineer can then select a suitable one with respect to the type of relevant limit state, although frequently, for pragmatic reasons, the model choice is based on the availability of model data and effective software.

Modelling of degradation processes may be based on models of different levels of sophistication:

a) macro-level;
b) simplified models, probabilistic approach;
c) micro-level.

The *level a* is the most simple, often being called a "deemed-to-satisfy" set of rules (mostly according to current codes), and does not allow for the design/assessment of a specified service life with a specified

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

reliability level. The *level b* comprises simple models (often semi-empirical) verified by comparisons with results obtained from testing under experimental and real-life conditions; the variables are treated as random quantities, so the outputs are also capable of expressing statistical and probabilistic quantities (service life assessment). This is the level dealt with in the present work. The *level c* is the most refined one, where the models are complex and are developed making use of basic physical laws and often also the constitutive laws of mechanics, thus leading to the problem of needing to solve partial differential equations. This level of sophistication is too high for everyday design practice. Note that *levels b* and *c* may be viewed as performance-based design types.

Many variables applied in the assessment of deteriorating concrete structures show random spatial variability. In contrast, the majority of published analyses deal with 1D representation, which enables the investigation of a "point in space" or "hotspot". Because of this, only temporal variability is taken into account. However, numerous proposals for approaches which also facilitate the analysis of the spatial characteristics of deterioration processes have recently appeared – e.g. Darmavan & Stewart (2003), Straub (2011). Frequently, random fields in 2D space are used, often simulated by means of random variables generated for a chosen mesh in stochastic finite element analysis. The requirement for data concerning the correlation structure in space creates a challenge in real-life cases; therefore, monitoring/testing can be employed. A more appropriate and economical decision about the service life consequences can then be based on defining the limit states for a certain proportion of the structure, and not merely for an isolated hotspot. The spatial variability can also be captured by using e.g. cellular automata technique – for an example of this applied to chloride ingress see Podroužek & Teplý (2008).

2. Tool for practical applications

Any decision process dealing with degradation prognosis needs a suitable set of models and efficient tools with stochastic capabilities. In this context, and with regard to the *level b* mentioned above, a software package called FReET-D (see e.g. Teplý et al. (2007)) has been developed. The utilization of stochastic approaches - a combination of analytical models and simulation techniques - was involved in the creation of this specialized software for assessing newly-designed as well as existing concrete structures.

FReET-D is a programme associated with the multipurpose probabilistic software for statistical, sensitivity and reliability analysis of engineering problems, FReET (Feasible Reliability Engineering Tool), which is based on efficient reliability techniques, see (Novák et al., 2003) and (Vořechovský and Novák, 2003). FReET can be utilized in two modes: as a stand-alone multipurpose programme for any user-defined problem, and as a module integrated with ATENA software, which is produced by Červenka Consulting, Prague; this integration has been developed within the SARA project (Bergmeister et al., 2007). FReET-D provides:

(i) modelling of degradation phenomena in concrete structures, statistical and sensitivity analyses; Bayes updating;
(ii) assessment of service life;
(iii) assessment of reliability measures;
(iv) risk assessment.

For the purposes of options (ii) and (iii) the user may create different simple limit conditions of the following types:

$$P_f = P(A \geq B) < P_d \tag{1}$$

or

$$P_f = P(t_S \leq t_D) \leq P_d \tag{2}$$

where A is the effect of the analysed action, B is the barrier and P_d is the target probability value. Generally, both A and B are time dependent and hence P_f or reliability index β are time dependent too. The time t_S is a predicted time value corresponding to the limit given by eq. (1), i.e. the predicted service life, t_D is design service life.

The FReET-D module has been developed by implementing a number of degradation models for reinforced concrete structures. Degradation models are time dependent mathematical functions that show the average increase of cumulative degradation over time. These models are parameterized with several material, structural and environmental parameters which are considered to be random variables. For all models, the factor ψ (the general multiplier) of model uncertainty is provided to compensate for the possible inexactness or incompleteness of results.

The main criteria in selecting the degradation model for each specific use are:

- the type of relevant degradation mechanism, the definition of the appropriate limit state and the given exposure conditions;
- the availability of statistical data or the testing method for the input variables of each model;
- the accuracy of the model when using the available data in relation to the required accuracy/strategy level.

The list of models currently implemented in FReET-D is specified in Table I (mainly 1D models). The implementation of additional models is still in progress. The original literature sources for all models that are predominantly deterministic are referenced in the FReET-D manuals.

The individual models listed in Table I can be used to construct and analyze different limit states according to Eq. (1) or (2). Input parameters are defined as random variables which are described by their probability density functions (PDF) and statistical parameters or can also be described by user-defined raw data. A mutual statistical dependence between input variables can be prescribed and is controlled by a simulated annealing method (Vořechovský & Novák, 2009).

Some models may be highly input-demanding; in order to simplify the handling of inputs their statistical sensitivity analysis is provided by means of Spearman rank-order correlation coefficients so the user may easily gain measurements of the relative effect of each basic variable.

The present paper briefly mentions “old” features of the software tool FReET-D described previously in Teplý et al. (2007) and concentrates in more detail on some recently attached models for different degradation effects, namely sulphuric and general acid attack.

According to the user-defined type of analysis FReET-D provides the following type of outputs:

- after performing the statistical analysis (via the Monte Carlo method or Latin Hypercube sampling method), the statistical moments of output variables are shown in a numerical and graphical way; also, the values of the sensitivity coefficients for individual inputs are provided; Bayes updating is performed when additional data are inputted;
- reliability analysis provides the probability of failure value or reliability index relevant to a user-designed limit condition. For this purpose the FORM technique may also optionally be utilized;
- the best-fitted PDF may be automatically found for the output quantity.

Table I. Models implemented in FReET-D

Degradation mechanism	Model notation	Output	Note
Carbonation	Carb1a Carb1b	Carbonation depth at time *t*; time to depassivation	Concretes from Portland cement; model b differs by RH function
	Carb2a Carb2b		Ditto; simplified model
	Carb3		Concretes from Portland cement; influence of temperature
	Carb4a Carb4b		Concretes from blended cements; model b differs by RH function
	Carb5a Carb5b		Concretes from blended cements; model b is for HVFA concretes
	Carb6		Concretes from blended cements; type of cement considered
	Carb7		
	Carb8		Concretes from blended cements; *fib*-Model Code 2010 model
	Carb9		Concrete from Portland cement with a lime-cement mortar coating
Chloride ingress	Chlor1a Chlor1b	Depth of chlorination at time *t*; time to depassivation	Model b provides calculation of the saturation concentration of Cl^- via an analytical formula
	Chlor2a Chlor2b	Concentration of chlorides at depth *x* and time *t*	Model b provides calculation of the diffusion coefficient via an experimentally derived formula
	Chlor3a Chlor3b		*fib*-Model Code 2010 model; model b provides calculation of surface Cl^- conc. for specific conditions via an analytical formula
Reinforcement corrosion	Corr1	Net rebar diameter at time *t*	Uniform type of corrosion
	Corr2	Pit depth at time *t*	Pitting type of corrosion
	Corr3	Net cross sectional area of rebar at time *t*	Pitting type of corrosion
	Corr4	Time to cracking due to corrosion	Crack initiation on the steel-concrete interface; uniform corrosion
	Corr5	Crack width due to corrosion at time *t*	Crack width on concrete surface; uniform corrosion
	Corr6	Time to cracking due to corrosion	Crack initiation on the steel-concrete interface; uniform corrosion
	Corr7	strength and ductility of corroded steel	Yield stress and ultimate stress - apparent values; strain at ultimate strength
	Scc1a Scc1b	Stress intensity factor at the pit tip at time *t*	Prestressed reinforcement, pitting corrosion; fracture mechanics approach
Sulphate attack	Sulf	Rate of concrete corrosion	Corrosion of concrete sewer pipes
Acid attack	Acid	Depth of concrete corrosion	i) User defined concentration of the acid ii) User defined pH of mineral acid solution iii) Buffering media consideration

3. Description of models recently added to FReET-D

3.1. SULPHIDE ATTACK ON CONCRETE

In sewer collection systems there is a corrosion problem involving the destruction of concrete pipes or structures by acid produced from hydrogen sulphide gas – ASCE Manual (2007). A certain kind of bacteria causes this type of corrosion, which is termed *microbiologically induced corrosion* (MIC). The bacteria responsible include sulphur-oxidizing bacteria (Thiobacillus), which oxidize sulphide into sulphuric acid.

The degradation of concrete sewer pipes due to sulphuric acid attack can be assessed by the frequently cited model by Pomeroy and Parkhurst (1977) for the deterioration rate c of concrete [mm/year]:

$$c = 11.5k\Phi A^{-1}\psi \tag{3}$$

where k = the factor representing the proportion of acid reacting (only an estimated value and ranging from 1.0 when the acid formation is slow to 0.3 when it is formed rapidly), Φ = the flux of H_2S (or sulphide release [g $H_2S/(m^2$ hr)]), A = acid-consumption capability, i.e. the alkalinity of the concrete, expressed as the proportion of equivalent calcium carbonate [g $CaCO_3$/g concrete]. For granitic aggregate concretes A ranges from 0.17 to 0.24, while for calcareous aggregates it ranges from 0.9 to 1.1, the equivalent A value for mortar-lined pipes being 0.4; ψ = the coefficient of model uncertainty (optional).

Flux Φ can be expressed according to the ASCE Manual (2007) and Tee et al. (2011) as

$$\Phi = 0.7(su)^{3/8} jS_{\lim}(b/P) \tag{4}$$

where j is the pH-dependent factor for the proportion of H_2S, s is the slope of the pipeline, u is the (m/sec), S_{lim} is the limiting value of sulphide concentration or dissolved sulphide concentration, P is the wetted perimeter P of the pipe wall and b is the surface width of the stream. Note that the variables s and u are in reality strongly dependent on each other (the greater the slope, the greater the velocity of the stream); this fact can be specified by appropriate statistical correlation of these two input variables while using FReET-D.

Starting from the example shown in Tee et al. (2011) a parametric study has been performed using FReET-D. The following data were selected for a sewer pipe of outer diameter = 2.286 m and wall thickness = 0.216 m, flowing half full:

Table II. Input data for an analysed example of sulphate attack

Notation	Unit	Distribution	Mean	COV
S_{lim}	mg/l	LN(2 par)	3.12	0.10
s	%	LN(2 par)	0.001 (up to 0.003)	0.10
u	m/sec	LN(2 par);	0.60	0.10
A	-	N	0.20	0.05
j	-	deterministic	0.28	
P	m	deterministic	2.91	
b	m	LN(2 par)	1.854	0.10
ψ	-	deterministic	1.0	

Note: (i) the imposed Spearman correlation coefficient of variables s and u equals 0.9
(ii) from pH = 7.4 it follows the value of constant j − Pomeroy and Parkhurst (1977)

The variable *s* was set as a parameter in the range from 0.0010 to 0.0030 to show the effect of possible scatter due to the non-precise embedding of pipes. It is depicted in Fig. 1.

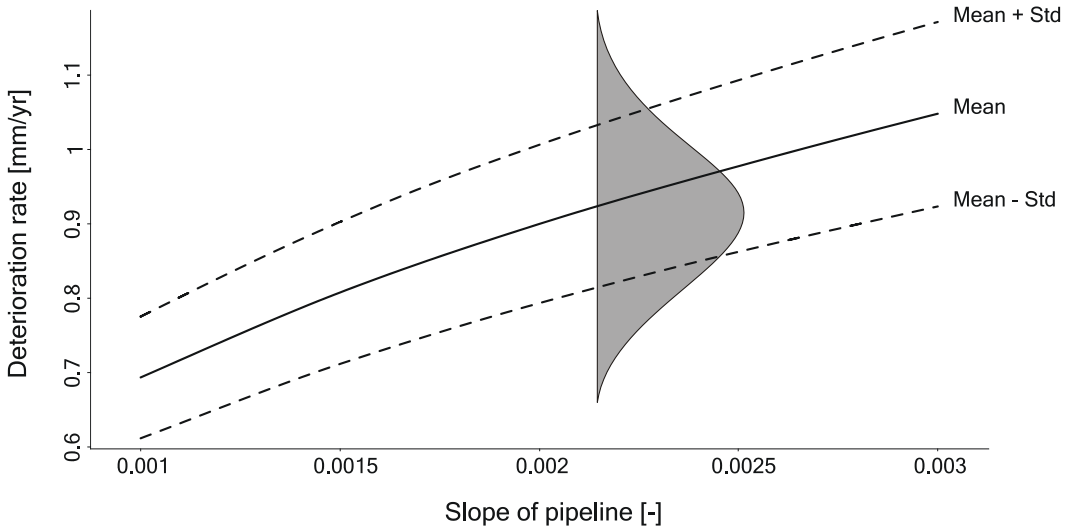


Figure 1. Deterioration rate of a sewer pipeline affected by sulphate attack as a function of the slope of the pipeline with depicted standard deviations resulting from statistical analysis.

3.2. ACID ATTACK

According to the fib Model Code (2010) the degree of concrete degradation caused by acid attack is defined by the corrosion depth *d* with respect to the original surface. It comprises the depth of material removed by abrasion and/or crystallization pressure and the depth of corroded material remaining on the concrete surface. The estimation of the time before a given depth of corrosion is reached is an important decision criterion for the designer.

Supposing the loss of surface material is negligible and the strength of the acid is assumed to be constant, the corrosion depth *d* [m] may then be estimated from the "square root" law

$$d = k_c \sqrt{ct}\ \psi \tag{5}$$

where c = the concentration of acid in [mol/L] and can be given as a (i) direct input quantity, or assessed either (ii) by Eq. (6) or (iii) by Eq. (7); t = contact time in [s], and ψ is the model uncertainty coefficient (optional). The constant k_c mediates the effect of concrete composition on the corrosion process and includes the effect of cement content and type, additions, w/c ratio and aggregate solubility. According to the fib Model Code (2010) no prediction formula for this constant may as yet be given − it should be determined by appropriate experiments, see e.g. Beddoe and Schmidt (2009a, b). As an example of utilizing this source the following k_c values for concrete fabricated from CEM I with a mass of 400 kg/m^3 were gained: w/c = 0.4 → k_c = 10.64; for w/c = 0.50 k_c = 12.46 and for w/c = 0.60 k_c = 14.10.

For mineral acids c [mol/L] is given by the proton concentration of the acid as calculated from its pH

$$c = 10^{-pH} \tag{6}$$

In the case of buffering media it is necessary to know the pH and the total content c_{tot} of acid and acid anions (e.g. acetate and acetic acid), dissolved CO_2 or ammonium:

$$c = \frac{10^{-pH} c_{tot}}{(10^{-pH} + K_s)} \tag{7}$$

where K_s = the dissociation constant in [mol/L] and c_{tot} = the total content of acid and conjugate base, dissolved CO_2 or ammonium in [mol/L].

For an illustrative example the following conditions and input data were used to calculate the corrosion depth. Concrete with a cement content of 400 kg/m^3 and w/c = 0.5, the variable k_c LN(12.46; 1.25), and the concentration of the acid was set to LN(0.1; 0.02) mol/L; it affected the concrete for a period of 10 years. Statistical analysis gives a mean corrosion depth value of 12.4 mm with COV = 0.14 and Lognormal (3par) as the best fitted PDF.

Acknowledgements

This paper has been prepared with the financial support of the Technologic Agency of the Czech Republic, project No. TA 01011019, and the Czech Science Foundation grant No. P105/10/1156 "COMOCOS - Complex modeling of concrete structures".

References

ASCE Manuals and Reports of Engineering Practice, No. 60, Gravity Sewers, American Society of Civil Engineers, New York, USA, 2007.

Beddoe, R.E. and Schmidt, K. Acid attack on concrete – effect of concrete composition. Part 1. *Cement International*, 7(3), 88-94, 2009.

Beddoe, R.E. and Schmidt, K. Acid attack on concrete – effect of concrete composition. Part 2. *Cement International*, 7(4), 86-93, 2009.

Bergmeister, K., Strauss, A., Pukl, R. and Novák, D. Structural analysis and reliability assessment, SARA Part I: Life-cycle reliability assessment, virtual simulation for health monitoring and damage identification of bridges. In *Proc. of 3rd International Conference on Structural Health Monitoring of Intelligent Infrastructure* (SHMII-3 2007), Vancouver, Canada, 2007.

Darmawan, M.S., Stewart, M.G. Spatial variability of pitting corrosion and its effect on the reliability of prestressing wires. In Der Kiureghian, Madanat & Pestana (eds.), *Application of Statistics and Probability in Civil Engineering*, 541-548. Rotterdam: Millpress, 2003.

EN 1990. Basis of Structural Design, European Standard, 2002.

EN 1992-1-1 Design of concrete structures – Part 1.1: General rules and rules for buildings, European Standard, 2003.

fib Bulletin No. 34, Service Life Design, 2006.

fib Model Code 2010, Final draft, Lausanne, 2010.

ISO 13823. General Principles in the Design of Structures for Durability, 2008.

Novák, D., Vořechovský, M. and Rusina, R. Small-sample Probabilistic Assessment – software FREET. In *Proceedings of the 9th International Conference on Applications of Statistics and Probability in Civil Engineering ICASP 9*, Millpress Rotterdam, San Francisco, USA, 91–96, 2003.

Podroužek, J., Teplý, B. Modelling of chloride transport in concrete by cellular automata. *Engineering Mechanics* 15(3): 1-10, 2008.

Pomeroy, R.D. and Parkhurst, J.D. The forecasting of sulfide build-up rates in sewers. *Progr. Water Technol.* 9 (3), 621-628, 1977.

Straub, D. Reliability updating with inspection and monitoring data in deteriorating reinforced concrete slabs. In Faber, Köhler & Nishijima (eds.), *Application of Statistics and Probability in Civil Engineering*, 2309-2316. London: Taylor & Francis Group, 2011.

Tee, K.F., Li, Ch.Q. and Mahmoodian, M.: Prediction of time-variant probability of failure for concrete sewer pipes. In *Proc. of XII DBMC*, Porto, Portugal, Vol. I, 447-454, 2011.

Teplý, B., Matesová, D., Chromá, M., Rovnaník, P. Stochastic degradation models for durability limit state evaluation: SARA – part VI. In *Proc. of 3rd International Conference on Structural Health Monitoring of Intelligent Infrastructure*, Vancouver, British Columbia, Canada, p. 187 (Abstract), full length paper on CD-ROM, 2007.

Vořechovský, M., Novák, D. Statistical correlation in stratified sampling. In *Proceedings of the 9th Int. Conf. on Applications of Statistics and Probability in Civil Engineering – ICASP 9*, San Francisco, USA, Rotterdam Millpress, 119-124, 2003.

Vořechovský, M., Novák, D. Correlation control in small-sample Monte Carlo type simulations I: A simulated annealing approach. *Probabilistic Engineering Mechanics* 24: 452-462, 2009.

Entropy Approach to Assessing Residual Life of Deteriorating Systems

S.A. Timashev[1)], A.V. Bushinskaya[2)]
[1)] *Science and Engineering Center "Reliability and Safety of Large Systems and Machines" Ural Branch, Russian Academy of Sciences, Ekaterinburg 620049, Russia, timashevs@cox.net*
[2)] *Science and Engineering Center "Reliability and Safety of Large Systems and Machines" Ural Branch, Russian Academy of Sciences, Ekaterinburg 620049, Russia, bushinskaya@gmail.com*

Abstract: The paper describes a novelty approach to estimating the entropy generated by the growth of multiple defects present in a distributed system. Examples of practical application of this methodology to assessing the residual life of pipeline systems (PS) affected by a multiple active degradation processes are presented.

The process of PS degradation (decrease of failure pressure of the defects due to the increase of their sizes and subsequent decrease of their residual strength) is considered as a non-homogeneous pure death Markov process (NPDMP) of the continuous time and discrete states type. Failure pressure is calculated, using one of the internationally recognized pipeline design codes: B13G, B31Gmod, DNV, Battelle and Shell-92.

The possible range of failure pressure (FP) change of defects is divided into M non-overlapping intervals. Thus, the structure of FP is a discrete finite set of states I_i (i = 1, 2,.. M). The probability of this state is the value of $P_i(t) = P\{P_f(t) \in I_i\}$, where $P_f(t)$ is the FP of defect.

The probability $P_i(t)$ is a measure of how definite is the occurrence of event the $P_f(t) \in I_i$. The measure of uncertainty of state I_i is [– ln $P_i(t)$], which is called *partial* entropy and characterizes this state only.

In the associative-structured approach the entropy associated with degradation of the residual strength of a defect is $H_d = \sum P_j(t)$ [– ln $P_j(t)$] (j = 1, 2, …, M).

Since entropy is a measure of uncertainty, it has the greatest value when the system states have an equiprobable distribution, i.e., when all the probabilities $P_i(t)$ are equal, and the uncertainty is maximal. The moment of time at which entropy is maximal can serve as an analog of the *conditional remaining life* (*warning* time of failure) of the defect or system. This approach has a potential of performing the role of early diagnostics of pipeline failure.

Keywords: entropy, residual strength, probability, Markov process, pipeline systems.

1. Description of entropy generated by the pipeline defects degradation described by a NPDMP

Consider a pipeline system (PS) with defects. Its residual strength (burst or fracture pressure FP) can be in a finite set of states: I_{ij} (i - number of defects, j - number of states) with probabilities P_{ij}.

Entropy as a function of time, generated by the degradation of pipeline cross section with a defect, which failure is of the rupture type, is calculated by equation:

$$H_d(t) = -\sum_{j=1}^{M} P_j(t) \ln\left[P_j(t)\right] \tag{1}$$

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

Where $P_i(t)$ is the probability that I_j $(j = M,..,1)$ at time t the FP $P_f(t)$ of a defect is in the j-th state. Assuming that the defects are independent of each other, their FPs will also be independent. Therefore, according to the property of entropy for independent random variables (Ventzel, 1969), the entropy of a pipeline as a system with independent defects can be calculated by formula:

$$H_P(t) = \sum_{i=1}^{n} H_{d_i}(t) = -\sum_{i=1}^{n}\sum_{j=1}^{M} P_{ij}(t)\ln\left[P_{ij}(t)\right] \quad (2)$$

Where n is the number of defects; $P_{ij}(t)$ is the probability that at time t the FP of i-th defect is in the j-th state I_j $(j = M,..,1)$.

To estimate the probabilities $P_j(t)$ non-homogeneous pure death Markov process (NPDMP) is used, which is described by systems of differential equations and do not depend on the nature of objects and their physical properties. In this sense the Markov processes are universal and are widely used in various fields of science and technology: nuclear physics, biology, astronomy, queuing theory, reliability theory, etc. (Ventzel, 1969; Gnedenko, 1988; Timashev, 1982; Feller, 1984; Bolotin, 1988; Gnedenko, Belyaev and Solovyev, 1965; Bogdanoff and Kozin, 1985).

A PS with many actively growing defects is a physical *distributed* system, which transits from one state to another. The degradation of the PS (measured as monotonous deterioration of its failure pressure) is considered as a non-homogeneous pure death Markov process (NPDMP) (Timashev and Bushinskaya, 2009; Bushinskaya, 2010).

Consider the cross section of pipeline with a defect. The burst pressure of a performing pipeline defective cross section at some fixed time t is a random variable (RV) $P_f(t) \geq P_{op}$, where Pop is the operating pressure in the pipeline. The burst pressure $P_f(t)$ can be assessed using one of the five internationally recognized pipeline design codes: B13G (ANSI/ASME B31G, 1991), B31Gmod (Kiefner and Vieth, 1989), DNV (DNV-RP-F101, 2004), Battelle (Stephens and Leis, 2000) and Shell-92 (Ritchie and Last, 1995).

Divide the possible range of change of the burst pressure of a pipeline defective cross section $\left(P_{op}; P_f(0)\right]$ into M-1 non-overlapping intervals I_i $(i = 1,..,M)$. Here $P_f(0)$ is the burst pressure of defect at the initial moment of time $t = 0$.

Take the last interval (conditional failure state), which includes the lowest values of burst pressure, as equal to (0; P_{op}].

The burst pressure of the defective cross section can only monotonically decrease over time, i.e., at random moments of time transit from i-th state only to the $(i + 1)$-th state, where the state is one of the intervals I_i $(i = 1,..,M)$. The system of differential equations (SDE) that describes the process of transition the burst pressure of a defect from one state to another has the form:

$$\begin{cases} \dfrac{dP_1(t)}{dt} = -\mu(t)P_1(t) \\ \dfrac{dP_i(t)}{dt} = \mu(t)P_{i-1}(t) - \mu(t)P_i(t), \ (i = 2,..,M-1) \\ \dfrac{dP_M(t)}{dt} = \mu(t)P_{M-1}(t) \end{cases} \tag{3}$$

where $P_i(t)$ is the probability that the RV $P_f(t)$ is in the i-th state at the moment of time t, $\mu_i(t)$ is intensity of transition from the i - th state to the $(i + 1)$-th state.

System (3) describes the non-homogeneous Markov process of pure death which is characterized by discrete number of states and continuous time. It is obvious that at the initial moment of time $t = 0$ RV $P_f(0) \in I_1$. Hence, the initial conditions for the SDE (3) have the form:

$$P_1(0) = 1, \ P_i(0) = 0, \ (i = 2,..,M) \tag{4}$$

Solving the SDE (3) by the method of variation of the constant, with initial conditions (4), obtain:

$$\begin{cases} P_1(t) = \exp\{-\rho(t)\} \\ P_i(t) = \dfrac{\rho^{i-1}(t)}{(i-1)!} \cdot \exp\{-\rho(t)\}, \ i = 2,..,M-1 \\ P_M(t) = 1 - \left[\exp\{-\rho(t)\} + \sum\limits_{i=2}^{M-1} \dfrac{\rho^{i-1}(t)}{(i-1)!} \cdot \exp\{-\rho(t)\} \right] \end{cases} \tag{5}$$

where $\rho(t)$ is calculated by formula:

$$\rho(t) = \int_0^t \mu(\tau)d\tau = -\int_0^t \frac{P_f'(\tau)}{\Delta I}d\tau = \frac{P_f(t) - P_f(0)}{\Delta I} \tag{6}$$

Since entropy is a measure of uncertainty, it is obvious that it is equal to zero when one of the probabilities $P_{ij}(t)$ is equal to unity (and all other probabilities are equal to zero), i.e., when the information is completely predictable and does not carry anything new. This situation occurs at the initial time $t = 0$ (see initial conditions (4)) and at the time when $P_M(t) = 1$ (i.e., when with probability 100% the FP of a defect is in the final, failure state).

The entropy takes the maximum value for the equiprobable distribution, where all probabilities $P_i(t)$ are equal, and therefore, the uncertainty is greatest. After reaching its maximum value the entropy begins monotonically decrease to zero, until the moment when $P_M(t) = 1$. Hence, the entropy maximum can serve as a measure of pipeline safety and is subject to further analysis.

The functions $P_i(t)$ are continuous and differentiable for any time t: $d(t) < wt$, wt is the pipe wall thickness, $d(t)$ is the defect depth at time t. Therefore, the maximum entropy can be found by solving the following equation with respect to time t:

$$\frac{dH_d(t)}{dt} = -\sum_{i=1}^{M}\left[\frac{dP_i(t)}{dt}\ln\left[P_i(t)\right] + \frac{dP_i(t)}{dt}\right] = 0 \tag{7}$$

2. Entropy Analysis of a Pipeline with Defects. Numerical examples

Evaluate the entropy, produced by the degradation of the pipeline cross section with a defect which fails by rupture. The pipeline parameters are presented in Table I.

Table I. Initial data. Pipeline parameters

Parameters	Symbol	Value
External pipeline diameter, mm	*D*	325.0
Pipeline wall thickness, mm	*wt*	9.0
Specified minimum yield strength of pipe material , MPa	*SMYS*	245.0
Minimum tensile strength of the pipe material, MPa	*UTS*	410.0
Design operating pressure, MPa	P_{op}	6.4

Consider that the last pipeline inspection discovered six characteristic defects with parameters given in Table II.

Table II. Initial data. Defects parameters

#	Defect depth *d*, mm	Defect length *l*, mm
1	10%wt	100
2	20%wt	
3	30%wt	
4	40%wt	
5	50%wt	
6	60%wt	

Assume that the growth rate (GR) of the defects depths a_d is equal to 0.5 mm/year and the growth rate GR of defects lengths a_l is 5 mm/year.

The stochastic process of degradation of FP of pipeline defect, which fails by rupture, is described by an NPDMP, using one of the five internationally recognized pipeline design codes: B31G, B31Gmod, DNV, Battelle and Shell92. Using these codes is acceptable for any time t: $d(t) < wt$, $d(t)$ is the defect depth at time t. Assuming that the defect sizes depend linearly from time t, define the limit time t_{ld} at which pipeline integrity is violated and a pinhole is formed ($d(t) = 0.9999wt \approx wt$):

$$t_{ld} = \frac{0.9999wt - d_0}{a_d} \tag{8}$$

Where a_d is the defect depth GR; d_0 is the defect depth at initial time $t = 0$.

The obtained graphs of the change of *local* entropy (i.e., entropy related to degradation of one of the pipeline defective cross sections) for codes B31G, B31Gmod and DNV are shown in Fig. 1−3. Fig. 4 shows entropy evolution of the most dangerous defect, with a 60%*wt* depth, using all the above five codes.

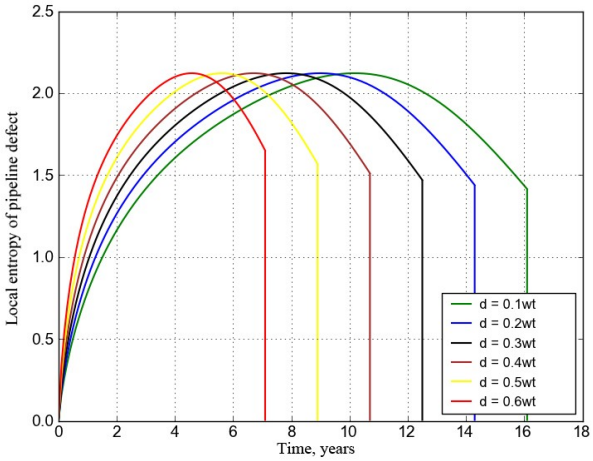


Figure 1. Entropy change of six defects over time (as per B31G code)

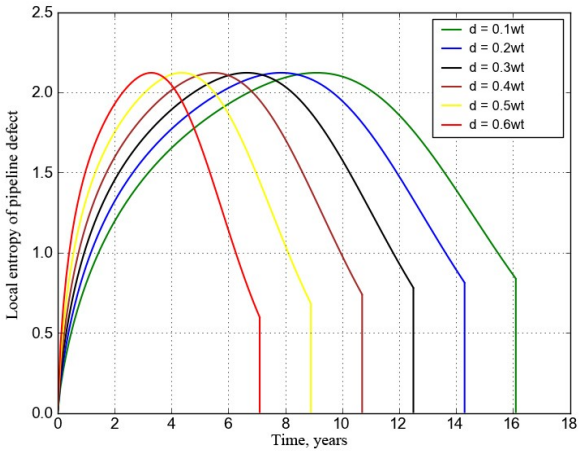


Figure 2. Entropy change of six defects over time (as per B31Gmod code)

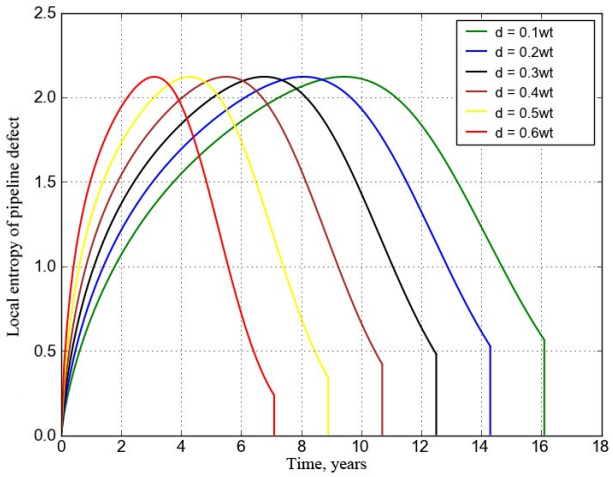


Figure 3. Entropy change of six defects over time (as per DNV code)

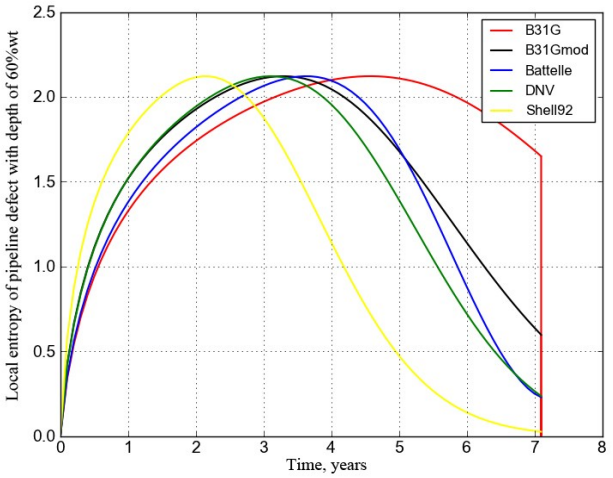


Figure 4. Entropy change of the most dangerous defect (with a 60%wt depth) over time (as per all five codes)

In Fig. 1−4 solid vertical lines, which descend from the ends of the entropy function, cut off on the x-axis the time t_{ld} , at which $d(t_{ld}) = 0.9999wt$.

According to Fig. 1−4, the entropy values for all defects have the same maximum values, independent from the GR and size of the defects parameters, indicating to entropy invariance, but these maximum values are reached at different moments of times.

Now calculate using eqn. (2) the entropy for the whole pipeline, considering it as the system with six identified above defects. The obtained results are shown in Fig. V.

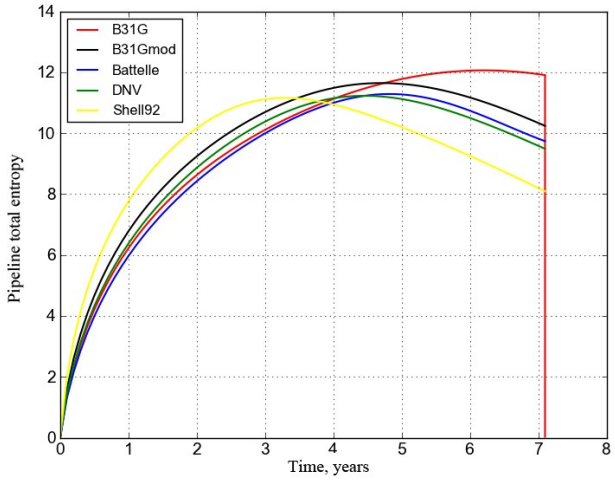


Figure 5. Entropy of the whole pipeline as a system with defects (as per all five codes)

According to Fig. 5, for the pipeline as a system the invariance of maximum entropy is not conserved. In other words, this maximum depends on the used code. This may be explained by the fact that the entropy of the whole pipeline is calculated as the sum of entropies of all defects and that the entropy maximum of each defect is reached at different times.

In Table III the times are presented, when the entropies of the whole pipeline and the most dangerous defect reach their maximums, calculated by eqn. (3) using corresponding codes.

Table III. The times when the entropies of the most dangerous defect (depth 60%*wt*) and the pipeline as a system reach their corresponding maximums (years)

Code	Time when entropy of the pipeline reaches its maximum, yrs	Time when entropy of the most dangerous defect reaches its maximum, yrs
B31G	6.2	4.6
B31Gmod	4.7	3.3
Battelle	4.8	3.6
DNV	4.5	3.1
Shell92	3.3	2.1

For all codes the entropy maximum of the whole pipeline is larger than entropy maximum of the most dangerous defect (see Table III). This fact can be easily explained, as the entropy of the whole pipeline is defined as the sum of entropies of all the defects present.

In Table IV the time distance from reaching the entropy maximum to the time at which the FP is equal to the operating pressure FP = OP (limiting condition, i.e. pipeline failure) is presented.

Table IV. Time (years) required for the occurrence of the limiting state (FP = OP), from the time the defect reaches its maximum entropy

Defect depth	B31G	B31Gmod	Battelle	DNV
0.1*wt*	5.7	4.7	1.5	1.6
0.2*wt*	5.3	4.2	1.2	1.4
0.3*wt*	4.7	3.8	0.9	1.0
0.4*wt*	4.2	3.3	0.6	0.7
0.5*wt*	3.7	2.7	0.3	0.3
0.6*wt*	3.1	2.2	0.0	0.1

According to Table IV, for all considered codes, the limiting state for the rupture type defects occurs later in time than the maximum of their entropy. With the defect depth growing, this time rapidly reduces: for B31G code from 5.7 to 3.1 years, for the B31Gmod code from 4.7 to 2.2 years, and for the Battelle and DNV codes to 1.5 (1.6) to 0.0 (0.1), respectively. Thus, the time of maximum entropy of the defect can be considered as an analog of its critical (warning) condition (Timashev and Bushinskaya, 2012) and used for early diagnostics of the defect condition and for optimizing the plan of its repair.

Now investigate the behavior of the entropy of a defect as a function of time, depending on the defects GR. Consider also the times of occurrence of the critical and limiting conditions of a defect on the entropy graph with respect to the time when the defect entropy reaches its maximum.

Consider one of the defects of the above pipeline, with following parameters: depth = 10%*wt*, length = 100 mm, defect depth GR = 1, 2 and 3 mm / year and defect length GR = 5 mm / year.

To construct the stochastic process of the degradation of BP of pipeline rupture type defect on the basis of NPDMP, use the two most commonly utilized codes: the B31Gmod code (most conservative) and the DNV code (least conservative).

The obtained graphs of the evolution of the local entropy (related to defect degradation) are shown in Table V and in Figs. 8 and 9.

Table V. The entropy maximum time and limit time t_{ld}

Defect depth GR	Limit time t_{ld}, yrs	Entropy maximum time, years		Entropy maximum value	The defect depth at time of entropy reaching its maximum, %*wt*	
		B31Gmod	DNV		B31Gmod	DNV
1 mm/year	8.09	4.65	4.02		61.71	54.57
2 mm/year	4.05	2.42	2.10	1.57	63.78	56.74
3 mm/year	2.69	1.64	1.43		64.60	57.58

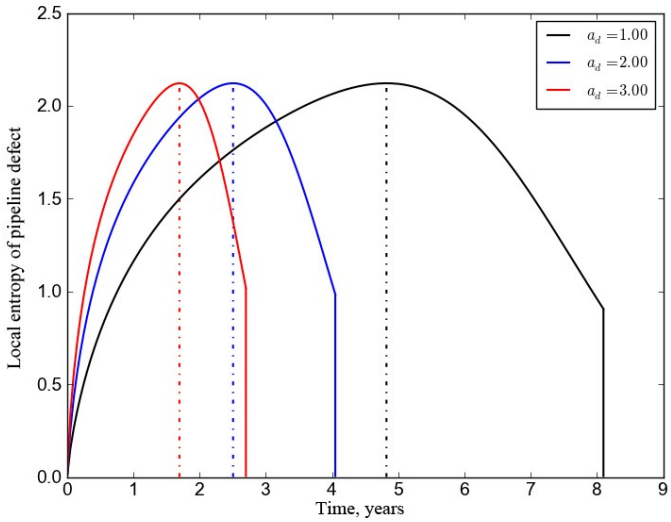


Figure 6. Defect entropy change in time, depending on the defect depth GR (as per B31Gmod code)

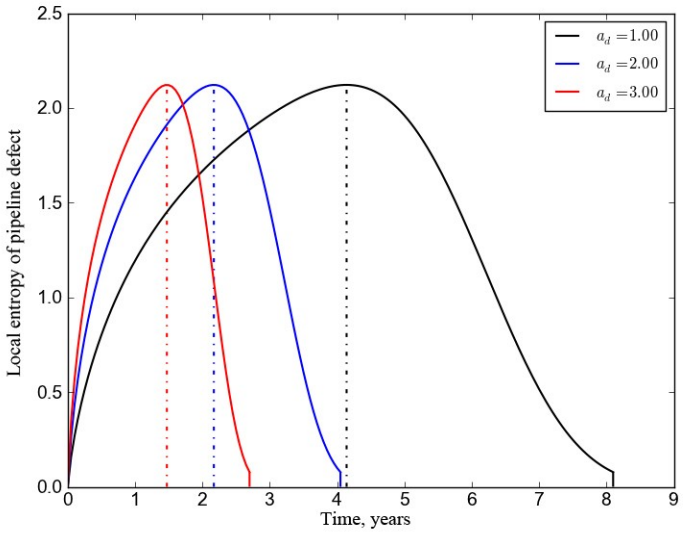


Figure 7. Defect entropy change in time, depending on the defect depth GR (as per DNV code)

According to Fig. 6, 7 and Table V the entropy maximum of a defect takes the same value, invariant to the used codes and defect depth GR. Given the results presented in Fig. 1−3, following conclusion can be made:

The maximum entropy of a defect takes the same value, independent from the used design code, defect depth and its GR. This fact confirms fundamental nature of entropy generated by deterioration of mechanical systems.

Now estimate:

- the moments of time t_1, t_2, t_3 at which the defect depth will be equal to 60, 70, and 80% of pipe wall thickness, respectively;
- the moment of time t_Q at which the probability of failure (the probability that the RV FP is in the last state) is equal to the maximum allowable value of POF = 10^{-5};
- the moment of time t_s at which the maximum safe operating pressure (MSOP) $P_s(t)$ of a defect will be equal to the operating pressure P_{op}: $P_s(t_s) = K \cdot P_f(t_s) = P_{op}$, where K is the strength safety factor;
- the moment of time t_f at which the FP of defect $P_f(t)$ is equal to the operating pressure: $P_f(t_f) = P_{op}$.

The obtained results are presented in Table VI and in Figs. 10.a, b, 11.a, b.

Table VI. Time (years) to occurrence of the critical and limit conditions of the defect

Grow rate of defect depth	t_1	t_2	t_3	B31Gmod code			DNV code		
				t_Q	t_s	t_f	t_Q	t_s	t_f
1 mm/yr	4.50	5.40	6.30	1.61	5.92	7.14	1.50	3.86	5.83
2 mm/yr	2.25	2.70	3.14	0.83	3.06	3.66	0.78	2.02	3.01
3 mm/yr	1.50	1.80	2.10	0.56	2.07	2.46	0.53	1.37	2.03

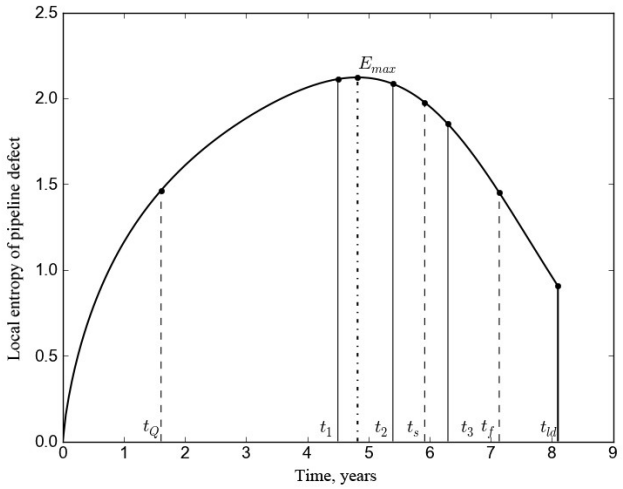


Figure 8.a.Defect entropy (as per B31Gmod code), the critical and limit moments of time (GR = 1 mm/yr)

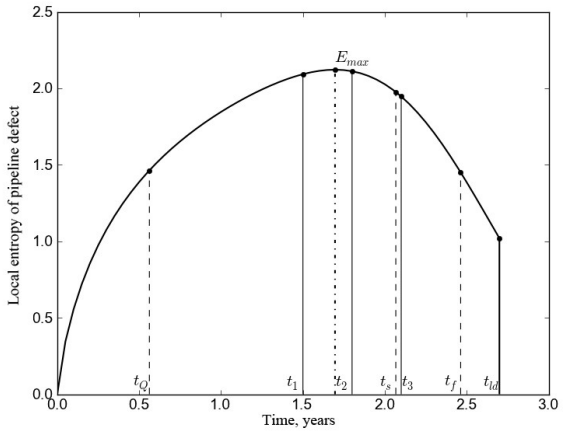


Figure 8.b. Defect entropy (as per B31Gmod code), the critical and limit moments of time (GR = 3 mm/yr)

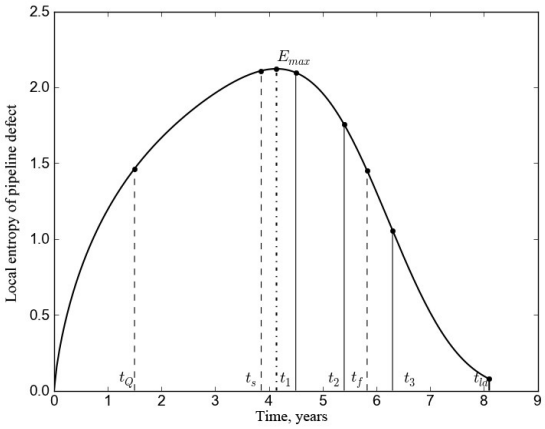


Figure 9.a. Defect entropy (as per DNV code), the critical and limit moments of time (GR =1 mm/yr)

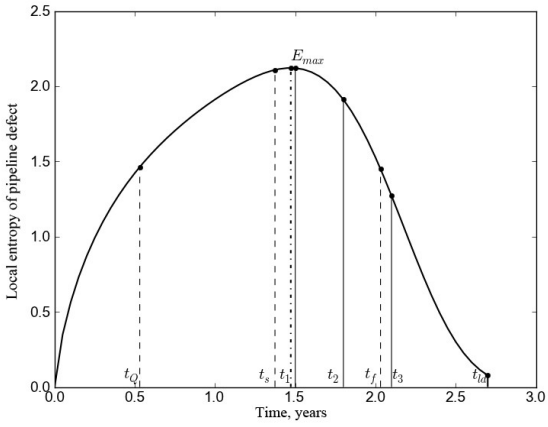


Figure 9.b. Defect entropy (as per DNV code), the critical and limit moments of time (GR = 3 mm/yr)

In Fig. 8.a,b, 9a,b, the dashed lines (except at the points of maximum entropy) correspond to the rupture type failure, the solid lines − to the leak type failure of defects.

According to Fig. 11.a,b, *in the case of using the DNV code for the defect under consideration, the entropy maximum, depending on the corrosion rates, is reached later than the critical condition MSOP = OP occurs*.

Moreover, in this case the limit condition of the rupture type defect occurs before the critical condition of the leak type, that is $t_f < t_3$; in the case of using the B31Gmod code the opposite is true (i.e., $t_3 < t_f$).

The difference (in years) between the time the defect entropy reaches its maximum and the time the defect reaches its critical and limit states are shown in Table VII.

Table VII. The difference (in years) between the time the defect entropy reaches its maximum and the time the defect reaches its critical and limit states

	according to B31Gmod code						according to DNV code					
Defect depth GR	t_Q	t_1	t_2	t_3	t_s	t_f	t_Q	t_1	t_2	t_3	t_s	t_f
1 mm/year	**4.26**	**0.15**	-0.75	-1.65	-1.27	-2.49	**3.64**	-0.48	-1.38	-2.28	**0.16**	-1.81
2 mm/year	**2.22**	**0.17**	-0.28	-0.72	-0.64	-1.24	**1.91**	-0.15	-0.60	-1.04	**0.08**	-0.91
3 mm/year	**1.50**	**0.14**	-0.16	-0.46	-0.43	-0.82	**1.30**	-0.07	-0.37	-0.67	**0.06**	-0.60

According to Table VII, when using the B31Gmod code, the time t_1 occurs earlier than the time when entropy reaches its maximum; when using DNV code, this condition is true for time t_s. According to the DNV code the critical state for a rupture type defect occurs earlier than the critical state for a leak type defect. When using the B31G code the opposite holds true.

Proceed to estimate intensity of the entropy change in time. Calculate the average value of the entropy function in intervals $I_1 = [0; t_{E\max}]$, $I_2 = [t_{E\max}; t_{ld}]$, $I_3 = [0; t_{ld}]$, by equations:

$$\text{for } I_1: \frac{\int_0^{t_{E\max}} f(t)\,dt}{t_{E\max}}; \qquad \text{for } I_2: \frac{\int_{t_{E\max}}^{t_{ld}} f(t)\,dt}{t_{ld} - t_{E\max}}; \qquad \text{for } I_3: \frac{\int_0^{t_{ld}} f(t)\,dt}{t_{ld}}, \tag{9}$$

where $t_{E\max}$ is the point of entropy maximum, $f(t)$ is the entropy function of a degrading defect. The obtained results are shown in Table VIII.

According to Table VIII, the average values of the entropy function for the same period of time do not depend on the defect depth GR and have nearly the same value, which means that they are practically invariant.

Table VIII. Average values of the entropy function of a degrading defect

Defect depth GR	For the B31Gmod code, in intervals			For the DNV code, in intervals		
	I_1	I_2	I_3	I_1	I_2	I_3
1 mm/yr	1.11	1.33	1.21	1.09	1.03	1.06
2 mm/yr	1.12	1.35	1.21	1.09	1.05	1.07
3 mm/yr	1.12	1.36	1.21	1.09	1.06	1.07

Now consider a numerical example of the entropy evolution for the whole pipeline as a system with defects, depending on adopted repair policies. Choose two policies of repair. According to the first policy, defect repair is performed either after the defect depth will be greater than 80%*wt*, or when the MSOP is equal to the OP. According to the second policy, defect is not repaired until pinhole formation (defect depth is equal to pipe wall thickness), or until the pipeline ruptures (FP = OP), depending on what type of failure occurs first. It is obvious that after repair the defect is nonexistent and its entropy is equal to zero.

Suppose there are nine defects in the pipeline, which parameters are given in Table IX.

Table IX. Initial parameters of defects

Defect #	Defect depth *d*, mm	Defect length *l*, mm
1	10%*wt*	
2	15%*wt*	
3	20%*wt*	
4	25%*wt*	
5	30%*wt*	100
6	35%*wt*	
7	40%*wt*	
8	45%*wt*	
9	50%*wt*	

Assume that the defect depth GR = 0.1 mm/year, the defects lengths GR = 5 mm/year, and the defects are independent of each other. Calculate by formula (2) entropy of the whole pipeline, taking into account two policies of defects repair. Consider the process of entropy variation of the whole pipeline up to the moment of time at which the depth of the first defect (10%*wt*) will be equal to the pipe wall thickness. In this case, the entropy of the whole pipeline takes the form shown in Figs. 10 and 11.

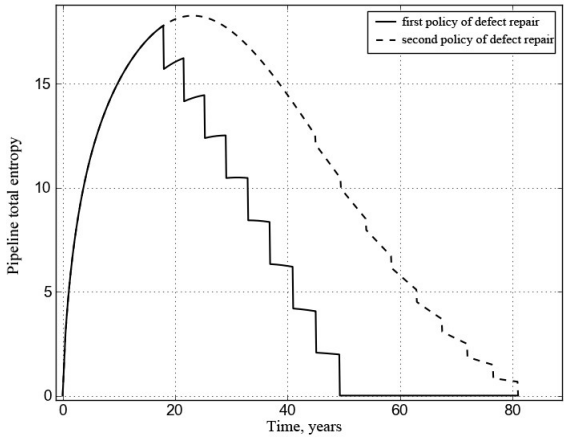


Figure 10. Evolution of the pipeline total entropy in time, depending on the adopted policy of defects repair (as per B31Gmod code)

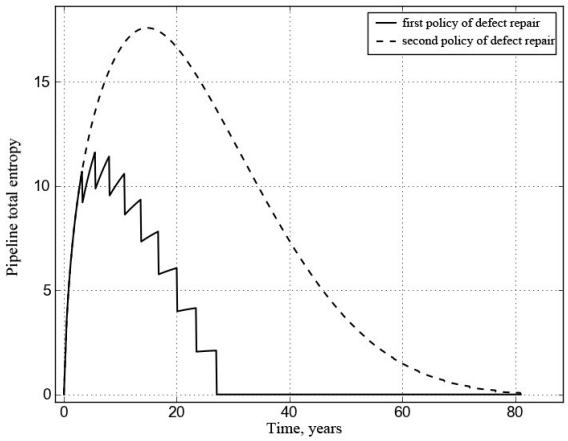


Figure 11. Evolution of the pipeline total entropy in time, depending on the adopted policy of defects repair (as per DNV code)

Now analyze the partially-total entropy (PTE) of the whole pipeline, considering for each defect only the descending branch of its entropy function, which originates right after the point where entropy reaches its maximum (hence, called *partial* entropy). Since each defect has a different time of reaching its entropy maximum, consider the process of PTE change for the whole pipeline after the time the deepest defect reaches its entropy maximum. This time is the earliest time when a pipeline defect reaches its entropy maximum. From this time on, one can get a complete picture of the pipeline PTE change with time. In our example, the deepest defect is 55% wt deep. According to the B31Gmod code the entropy maximum of this defect is reached in 15.21 years, all other defects reach their entropy maximums later. For moments of time $t_1 > 15.21$ yrs the pipeline entropy was calculated using formula (2), as the sum of the entropy of all defects for which the entropy maximum is reached at time $t_j \geq t_1$. The obtained partially-total entropy of the whole pipeline is shown in Figs. 12 and Fig. 13.

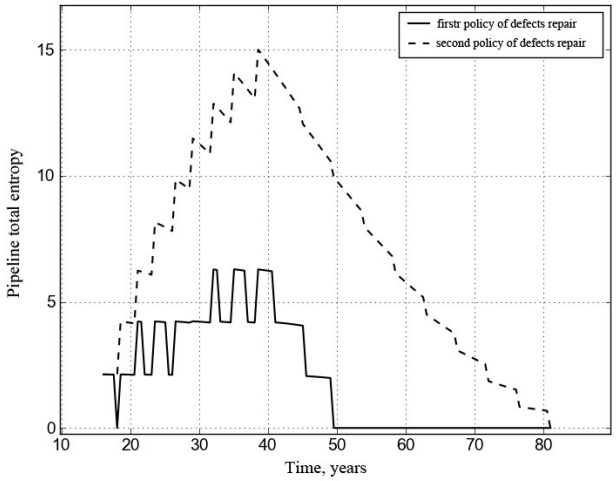


Figure 12. Evolution of pipeline system PTE in time depending on the policy of defects repair (as per B31Gmod code)

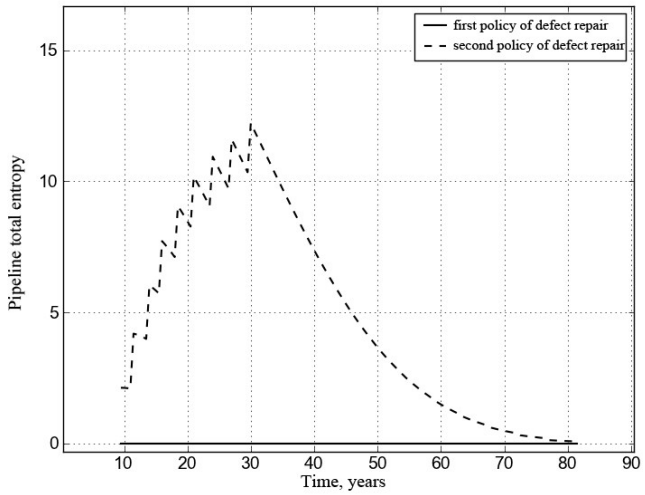


Figure 13. Evolution of pipeline system PTE in time depending on the policy of defects repair (as per DNV code)

The up and down jumps of the pipeline entropy function (see Fig.12) are caused by the following. The jumps down are due to the fact that at some moments of time, certain defects were already repaired. The jumps up are due to the fact that at some moments of time new defects came into consideration.

Such scheme for constructing the pipeline PTE leads to the curve given in Fig. 13. It has a maximum which is reached at the time when the shallowest defect (in our case, defect with a 10%wt depth) reaches its maximal entropy. After that, the PTE monotonically decreases, as the shallowest defect is the last defect providing a partial input into the pipeline PTE (all other defects are only repaired, and their entropies by that destroyed).

With the DNV code (Fig. 13) the pipeline PTE is everywhere equal to zero, since for all the defects the adopted policy of defects repair criterion (MSOP = OP) is reached *before* the deepest defect reaches its entropy maximum. In other words, when starting to visualize the evolution of pipeline entropy according to the adopted scheme all defects are already being repaired.

Conclusions

1. A methodology has been developed for estimating the entropy generated by degradation of pipeline defects of the rupture and leak type, as well as for the pipeline as a system with defects, apparently for the first time. Ratios have been defined between the different physical and probabilistic conditions of pipeline systems and the possibility of their early diagnostics in terms of information entropy was revealed.
2. According to conducted study, value of the defect entropy maximum is invariant with respect to the used design codes, defect depth and its GR, all of which confirms its fundamental nature.
3. This entropy measure could be useful for early diagnostics of pipeline systems condition, as well as for planning and optimizing their predictive maintenance (Timashev and Bushinskaya, 2012).

References

ANSI/ASME B31G-1991. Manual for Determining the Remaining Strength of Corroded Pipelines. National Standard of USA. – New York: ASME. – 1991. – 140 p.

Bogdanoff J. and Kozin F. Probabilistic models of cumulative damage. John Wiley & Sons NY, 1985, 344 p.

Bolotin V.V. Predicting life of machines and structures. – Moscow: Mashinostroenie, 1988. – 240 p. In Russian.

Bushinskaya, A.V. Description of Pipeline Strength Degradation as a Pure Death Markov Process and its Implementation in IMP // Proceedings of IPC Conference. Calgary, Canada. – September 2010. – Paper #IPC2008-31275

DNV-RP-F101. Corroded pipelines. Recommended Practice. – Norway: Det Norske Veritas, 2004. – 42 p.

Feller W. An Introduction to Probability Theory and Its Applications. Vol.1, 2. – Moscow: Mir, 1984. – 1266 p. In Russian.

Gnedenko B.V. Course on Probability Theory Manual.-6-th ed. Moscow: Nauka, 1988 – 448 p. In Russian.

Gnedenko B.V., Belyaev J.K. and Soloviev A.D. Mathematical methods in reliability theory. The main characteristics of reliability and statistical analysis. – Moscow: Nauka, 1965. – 524 p. In Russian.

Kiefner J.F. and Vieth P.H. A Modified Criterion for Evaluating the Remaining Strength of Corroded Pipe // AGA Pipeline Research Committee. Report PR 3-805. – 1989. – 78 p.

Ritchie D. and Last S. Burst Criteria of Corroded Pipelines - Defect Acceptance Criteria // Proceedings of the EPRG/PRC 10th Biennial Joint Technical Meeting on Line Pipe Research. – Cambridge, UK. – 1995. – Paper

Timashev S.A. and Bushinskaya A.V. Predictive Maintenance of Pipelines With Different Types of Defects Based on ILI Data Analysis, Pipeline Pigging and Integrity Management International Conference, PPIM, Houston, TX, USA, February 2012, Paper # 17.

Timashev S.A. and Bushinskaya A.V. Description of degradation process of pipelines with defects on the basis of pure death Markov model // Proceedings of III All Russia Conference «Safety of Critical Infrastructures and Territories», Ural Branch, Russian Academy of Sciences, Ekaterinburg, December, 2009., P. 338-339, *In Russian*

Timashev S.A. Reliability of Large Mechanical Systems. – M.: Nauka, 1982 . – 183 p. In Russian

Ventzel E.S. Theory of Probability. – M.: Nauka, 1969. – 576 p. In Russian

Assessment of the Reliability Level Embedded in International Codes for Pipeline Design

S.A. Timashev[1)], A.V. Bushinskaya[2)]
[1)] *Science and Engineering Center "Reliability and Safety of Large Systems and Machines" Ural Branch, Russian Academy of Sciences, Ekaterinburg 620049, Russia, timashevs@cox.net*
[2)] *Science and Engineering Center "Reliability and Safety of Large Systems and Machines" Ural Branch, Russian Academy of Sciences, Ekaterinburg 620049, Russia, bushinskaya@gmail.com*

Abstract: The paper analyses the actual reliability level which was empirically embedded in international codes for pipeline design [B31G, B31Gmod, Shell92, DNV, and Battelle (PCORRC)], using a real pipeline section as an example. A brief review of the above design codes and description of the plastic fracture criterion, which is the basis for these design codes are presented. The considered pipeline section contains longitudinally oriented surface defects of the corrosion/erosion type. Estimates of residual strength of the pipeline were carried out. For two characteristic defects the influence of randomness of pipeline geometry (diameter and pipe wall thickness), pipe material physical properties (ultimate tensile strength and specified minimum yield strength), load (operating pressure) and defect geometry (depth and length) on probability of failure (POF) is analyzed. This approach can be applied when studying the inherent reliability of other pipelines with defects.

Assessment of the reliability/probability of failure of pipeline with defects was performed using the Gram-Charlier-Edgeworth (GCE) method. This method is an assessment of the probability that the limit state function (*LSF*) of pipeline defect is positive at a given moment of time t. $LSF(t) = P_f(t) - P_{op}$, where $P_f(t)$ is the failure pressure, which is estimated by any of the above design code; P_{op} is the operating pressure.

The GCE method allows estimating the probability of failure/reliability of the pipeline defect and considering the stochastic character of pipeline geometry parameters, physical properties of pipe material, parameters of pipeline defects and operating pressure, treating them as random variables. Recommendations for choosing probability distributions for these random variables and calculating their statistical parameters are also presented. Extensive calculations permitted discovering the reliability levels which are actually present in the analyzed international pipeline design codes.

Keywords: reliability, probability of failure POF, design code, Gram-Charlier-Edgeworth series, pipeline systems.

1. Overview of the Most Common Pipeline Design International Codes

Currently, the most internationally recognized methodology for assessing the residual strength of longitudinally oriented surface (external or internal) corrosion defects are codes of the American Society of Mechanical Engineers (ASME). The initial code, ASME B31G, was adopted as the US national standard. (ANSI/ASME B31G, 1991) and, in simplified form, as the national standard of Canada (CAN3-Z183-M86, 1986). Subsequently, the modification of this standard was developed, which is called B31Gmod (Kiefner

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

and Vieth, 1989). In addition to the codes B31G and B31Gmod widely used methodologies are: DNV (DNV-RP-F101, 2004), Battelle (Stephens and Leis, 2000) and Shell-92 (Ritchie and Last, 1995).

All the above practical methodologies are based both on theory and extensive experiment conducted on real scale pipes. Their essence is in the calculated estimate of the burst (failure) pressure of the defective cross section of the pipeline. All the codes considered in this paper are based on the following formula which permits calculating the pipe burst pressure:

$$P_f(t) = \frac{2wt \cdot \sigma_f}{D} \frac{\left(1 - \frac{d(t)}{wt}\right)}{\left(1 - \frac{d(t)}{wt \cdot M(t)}\right)} \tag{1}$$

Where wt is the pipe wall thickness; D is the pipeline diameter; σ_f is the flow stress; $d(t)$ is the defect depth at the moment of t; M is the Folias factor.

The expression (1) for the evaluation of failure pressures for each of the codes (B31G, B31Gmod, Shell92, or DNV) is different, depending on the expressions used for the Folias factor, flow stress, and geometric shapes approximating the form of the defect.

According to (Alkazraji, 2008), all methods are classified as the SMYS-based and the UTS-based, where SMYS and UTS are specified minimum yield strength and ultimate tensile strength of the pipe material, respectively. UTS-based methods use the ultimate tensile strength of pipe material to define destruction of the pipeline defective cross section.

The described above codes can be applied only to a single cross section of the pipeline containing a longitudinally oriented, flat bottom surface defect of the corrosion/ erosion type.

Below formulas are given for calculating the failure pressure $P_f(t)$ of a single defective cross section of a pipeline for each of the codes.

The B31G (ANSI/ASME B31G code, 1991):

$$P_f(t) = \frac{2wt \cdot 1.1SMYS}{D} \cdot \frac{\left(1 - \frac{2}{3}\frac{d(t)}{wt}\right)}{\left(1 - \frac{2}{3}\frac{d(t)}{wt \cdot M_{B31G}(t)}\right)} \tag{2}$$

Where Folias factor is calculated by formula:

$$M_{B31G}(t) = \sqrt{1 + 0.893\frac{l^2(t)}{D \cdot wt}} \tag{3}$$

The flow stress $\sigma_s = 1.1SMYS$.

Equation (2) applies to defects with $l \le 1.12 \cdot 4\sqrt{D \cdot wt}$. For longer defects ($l > 4.48\sqrt{D \cdot wt}$) assessment of burst pressure is calculated by formula:

$$P_f(t) = \frac{2wt \cdot 1.1SMYS}{D} \cdot \left(1 - \frac{d(t)}{wt}\right) \tag{4}$$

The methodology is applicable only to pipes, which material class is below the X56 API 5L standard [5] (i.e., SMYS and UTS is less than, respectively, 386 and 489 MPa). The defects depth must be within the (10% - 80%) range of pipe wall thickness.

The B31Gmod code (Kiefner and Vieth, 1989):

$$P_f(t) = \frac{2wt(SMYS + 68.95MIIa)}{D} \cdot \frac{\left(1 - 0.85\frac{d(t)}{wt}\right)}{\left(1 - 0.85\frac{d(t)}{wt \cdot M_{B31G\,\mathrm{mod}}(t)}\right)} \tag{5}$$

This criterion uses a more accurate expression for the Folias factor than the code B31G and is less conservative (Alkazraji, 2008).

The Folias factor in this case is given by:

$$M_{B31G\,\mathrm{mod}}(t) = \begin{cases} \sqrt{1 + \frac{0.6275 \cdot l^2(t)}{D \cdot wt} - \frac{0.003375 \cdot l^4(t)}{(D \cdot wt)^2}}, & l(t) \le 7.07\sqrt{D \cdot wt} \\ \frac{0.032 \cdot l^2(t)}{D \cdot wt} + 3.3, & l(t) > 7.07\sqrt{D \cdot wt} \end{cases} \tag{6}$$

The flow stress is $\sigma_s = SMYS + 68.95MIIa$ $(10ksi)$.

The applicability of this methodology is similar to the applicability of the B31G code, with one difference: the defects depth must be within the (10% - 85%) range of pipe wall thickness.

The DNV code (DNV-RP-F101, 2004):

$$P_f(t) = \frac{2wt \cdot UTS}{D - wt} \cdot \frac{\left(1 - \frac{d(t)}{wt}\right)}{\left(1 - \frac{d(t)}{wt \cdot M_{DNV}(t)}\right)} \tag{7}$$

The Folias factor is calculated by formula:

$$M_{DNV}(t) = \sqrt{1 + 0.31\frac{l^2(t)}{D \cdot wt}} \tag{8}$$

The flow stress is equal to *UTS*. The methodology can be applied only to defects which depth is less than 85% of pipe wall thickness.

The Shell92 code (Ritchie and Last, 1995):

$$P_f(t) = \frac{2wt \cdot 0.9UTS}{D} \cdot \frac{\left(1 - \frac{d(t)}{wt}\right)}{\left(1 - \frac{d(t)}{wt \cdot M_{Shell92}(t)}\right)} \tag{9}$$

The Folias factor:

$$M_{Shell92}(t) = \sqrt{1 + 0.805 \frac{l^2(t)}{D \cdot wt}} \tag{10}$$

The flow stress in this code is equal to $0.9UTS$. This methodology should be applied only to defects which depth is less than 85% of pipe wall thickness.

The PCORRC (Battelle) code (Stephens and Leis, 2000):

$$P_f(t) = \frac{2wt \cdot UTS}{D} \cdot \left(1 - \frac{d(t)}{wt} M_{Battelle}(t)\right) \tag{11}$$

Where

$$M_{Battelle}(t) = 1 - \exp\left[-0.157 \frac{l(t)}{\sqrt{\frac{D}{2}(wt - d(t))}}\right] \tag{12}$$

This methodology can be used only for pipelines, which are operated at temperatures exceeding the temperature of pipe material ductile-brittle transition, and for pipe material with the impact energy of Charpy 61J and above.

2. Method of estimating the probability of failure/reliability of defective cross sections of pipeline system

2.1 The Gram-Charlier-Edgeworth Method

The probability of failure / reliability of pipeline defective cross section is estimated, using the probabilistic approach based on the Gram-Charlier-Edgeworth series (GCE) (Poluyan, Bushinskaya, Malyukova and Timashev, 2009; Timashev, Malyukova, Poluyan and Bushinskaya, 2008). This method allows assessing the probability of failure / reliability of a defective cross section of a pipeline, while taking into account the probabilistic nature of pipeline and defects geometry parameters, material properties and operating pressure.

The following main uncertainties of a specific cross section of a pipeline are considered as random variables (RVs): defect depth $d(t)$ and length $l(t)$; pipe wall thickness wt and diameter D; $SMYS$ and UTS of pipe material; and design operating pressure P_{op}.

The GSE method is an assessment of the probability that the limit state function (LSF) of a defective pipeline cross section has a positive value at an arbitrary moment of time t (i.e., assessment of defect reliability). The LSF is the difference between the value of the function which determines the burst pressure and the value of the design operating pressure:

$$Y(t) = P_f(t) - P_{op} \tag{13}$$

Where $P_f(t)$ is the burst (failure) pressure, assessed by any of the above codes.

The essence of this method is the approximation of the unknown probability density function PDF $f(y)$ of the RV $Y(t)$ by an orthogonal polynomial, which is a partial sum of the Gram-Charlier-Edgeworth series. According to this method, the reliability of the pipeline defective cross section is calculated by formula:

$$P\left(Y\left(t\right)>0\right)=1-F\left(0\right)=1-\left[\Phi\left(z\right)-\frac{S}{3!}\varphi^{(2)}\left(z\right)+\frac{E}{4!}\varphi^{(3)}\left(z\right)+\frac{10S^2}{6!}\varphi^{(4)}\left(z\right)\right]\Big/\sigma_Y \tag{14}$$

Where $z=(0-\mu_Y)/\sigma_Y$.

Then the probability of failure (POF) of the pipeline defective cross section can be calculated using following formula:

$$F\left(0\right)=\left[\Phi\left(z\right)-\frac{S}{3!}\varphi^{(2)}\left(z\right)+\frac{E}{4!}\varphi^{(3)}\left(z\right)+\frac{10S^2}{6!}\varphi^{(4)}\left(z\right)\right]\Big/\sigma_Y \tag{15}$$

Where $\varphi\left(y\right)$ is the PDF of the standard normal distribution, exponents *n* of this function indicate its *n*-th derivative; *y* is the normalized value of $y=\left(y-\mu_Y\right)/\sigma_Y$; μ_Y is the mean of the RV *Y*(*t*); σ_Y is the standard deviation of RV *Y*(*t*); *S* is the asymmetry coefficient of RV *Y*(*t*):

$$S=\frac{\mu_3\left[Y\right]}{\sigma_Y^3} \tag{16}$$

E is the curtosis of *Y*(*t*):

$$E=\left(\frac{\mu_4\left[Y\right]}{\sigma_Y^4}-3\right) \tag{17}$$

In eqns. (17) and (18) $\mu_3\left[Y\right],\mu_4\left[Y\right]$ are the third and fourth moments of the RV *Y*(*t*), respectively.

Denote as x_i its *i*-th RV in the LSF: $D=x_1$, $wt=x_2$, $d\left(t\right)=x_3$, $l\left(t\right)=x_4$, $SMYS=x_5$ or $UTS=x_5$, $P_{op}=x_6$. With these notations, the mean and variance of RV $Y(t)$ are calculated by eqns.:

$$\mu_Y=Y\left(t,m_{x_1},m_{x_2},..,m_{x_6}\right)+\frac{1}{2}\sum_{i=1}^{6}\left[\left(\frac{\partial^2 Y}{\partial x_i^2}\right)_{m_{x_1},..,m_{x_6}}\cdot D_{x_i}\right] \tag{18}$$

$$D_Y=\sum_{i=1}^{6}\left[\left(\frac{\partial Y}{\partial x_i}\right)^2_{m_{x_1},..,m_{x_6}}\cdot D_{x_i}\right]+\frac{1}{4}\sum_{i=1}^{6}\left[\left(\frac{\partial^2 Y}{\partial x_i^2}\right)^2_{m_{x_1},..,m_{x_6}}\cdot\left(\mu_4\left[x_i\right]-D_{x_i}^2\right)\right] \tag{19}$$

Where m_{x_i} is the mean of the *i*-th RV included in the LSF; D_{x_i} is the variance of the *i*-th RV included in the LSF, $\mu_4\left[x_i\right]$ is the fourth moment of the RV x_i.

The second terms in equations (18) and (19) are amendments for the nonlinearity of the LSF. The third and fourth central moments of RV $Y(t)$ are calculated by eqns.:

$$\mu_3\left[Y\right]=\sum_{i=1}^{6}\left[\left(\frac{\partial Y}{\partial x_i}\right)^3\Bigg|_{m_{x_1},..,m_{x_6}}\cdot\mu_3\left[x_i\right]\right] \tag{20}$$

$$\mu_4[Y] = \sum_{i=1}^{6}\left[\left(\frac{\partial Y}{\partial x_i}\right)^4\Bigg|_{m_{x_1},..,m_{x_6}} \cdot \mu_4[x_i]\right] + 6\sum_{\substack{i,k=1 \\ i<k}}^{6}\left[\left(\frac{\partial Y}{\partial x_i}\right)^2\Bigg|_{m_{x_1},..,m_{x_6}} \cdot \left(\frac{\partial Y}{\partial x_k}\right)^2\Bigg|_{m_{x_1},..,m_{x_6}} D_{x_i} D_{x_k}\right] \quad (21)$$

Where $\mu_3[x_i]$ is the third moment of RV x_i.

Note that the calculation by formula (14) or (15) is always carried out for the zero value of *F*. In a particular calculation only the statistical characteristics (mean, variance and moments) of LSF will vary. Thus, for a less dangerous defect, the LSF mean will be will be farther away from zero, due to which the POF will decrease. In the opposite case, for a more dangerous defect, the burst pressure will decrease and the difference between the burst and operating pressure will decrease, hence, the POF will increase.

2.2 Recommendations for Choosing the Probability Characteristics of the Pipeline Systems Parameters

For practical use of the GCE method probabilistic characteristics of the initial data have to be known: the geometric parameters of pipeline and defects, load parameters and mechanical properties of pipe steel.The probabilistic characteristics of these parameters are recommended in (C-FER Project No. L128, 2005).

2.2.1 *Probabilistic characteristics of the geometric parameters of the pipeline*

The probabilistic characteristics of the relationship between the actual and the nominal value of the pipeline geometric parameters are presented in Table I.

Table I. The probabilistic characteristics of the relationship between the actual and the nominal value of the pipeline geometric parameters

Parameter	Distribution law	Mean	Coefficient of variation
Diameter	Deterministic [1]	1.00	0.00
	Normal	1.00	0.06 %
Pipe wall thickness	Normal [2]	1.00	0.25/*wt*
	Normal [3]	1.10	3.30 %
	Normal	1.01	1.00 %

Different metallurgical plants can produce pipes with different levels of variability in their geometric parameters. It is shown (Jiao, Sotberg and Igland, 1995) that the standard deviation (SD) can be calculated on the assumption that the width of the interval of deviation from the nominal value is three sigma on either side of the nominal value, i.e., the geometrical parameters of the pipeline are normally distributed.

It should be noted that the use of a normal distribution for pipeline geometry parameters is not in contradiction with their physical properties (strictly positive values), as the scatter of a RV around the mean

[1] Based on the pipes with diameter from 16 to 56 inches. RVs with a COV < 0.1% could be considered as deterministic parameters.
[2] For welded pipes with wall thickness in the range between 15 and 37 mm. In eqn. of COV *wt* is the nominal pipe wall thickness, i.e., the standard deviation = 0.25 mm.
[3] For seamless pipe, based on tolerances for wall thickness.

value is such, that the occurrence of negative values is *practically impossible*. For example, according to Table I, the mean of pipe wall thickness (*wt*) is equal to 110% of its nominal value and the COV is equal to 3.3% (the largest value in Table I). Hence, the standard deviation (SD) is 3.63% of *wt*. With this value of SD, the negative values of *wt* are outside the 30.3SD range. Using properties of the normal distribution, the probability of hitting a RV outside six SD is equal to $1.97 \cdot 10^{-9}$ (0.002 cases per million), for seven SD such probability is equal to $1.28 \cdot 10^{-12}$, and for 30.3SD this probability is infinitesimal ($5.23 \cdot 10^{-202}$).

2.2.2 *Probabilistic characteristics of the operating pressure*

The relationship between operating pressure (OP) and the design pressure (DP) at an arbitrary time can be described by a beta distribution with mean 0.865 and coefficient of variation (COV) 0.084. In this case, the operating pressures have as its lower limit 60% and as its upper limit 110% of the design pressure. This distribution is based on the assumption that the maximum operating pressure in the pipeline is equal to the design pressure.

In the pipeline section located immediately after the pumping or compressor stations, the ratio between the maximum annual operating and design pressure has a Gumbel distribution with mean between 1.03 and 1.07 and COV of 1 to 2%. This distribution provides a conservative estimate of the annual maximum operating pressure for those pipeline sections, which are farther downstream from the pumping/compressor stations.

2.2.3 *Probabilistic characteristics of the pipe material mechanical properties*

Probabilistic characteristics of the relationship between the actual yield and ultimate strength and SMYS and UTS of the pipe material are presented in Table II.

Table II. Probabilistic characteristics of pipe material properties

Parameter	Disribution law	Mean	COV, %
SMYS	Normal [4]	1.11	3.40
	Normal [4]	1.08	3.30
	Normal or Lognormal [5]	1.10	3.50
	Lognormal	1.08	4.00
UTS	Normal [4]	1.12	3.00
	Normal [4]	1.12	3.50

Probabilistic characteristics of the relationship between the actual yield (ultimate) stress and the SMYS (UTS) of the pipe material is based on data from plant tests. In general, the distribution does not depend on the steel grade (C-FER Project No. L128, 2005). It is shown (Jiao, Sotberg and Igland, 1997), that the COV may vary, depending on the quality of the pipe steel from different plants. So, as far as possible, results of plants tests should be used.

[4] Based on multiple tests of samples of different manufacturing plants (760 test samples of steel grade X60 (API 5L standard, 2004) and 2753 test samples of steel X65 (API 5L standard, 2004)).
[5] Based on plants data (pipe steel grade X60 and X70 (API 5L standard, 2004)).

Use of the normal distribution with parameters from Table II for the properties of pipe material also totally excludes negative values of RVs, as in the case with the pipe and defects geometry parameters..

3. Analysis of the reliability level actually embedded in the international codes for pipeline design

To assess the actual level of reliability embedded in the international codes [B31G, B31Gmod, Shell92, DNV and PCORRC], consider first the problem of evaluating reliability of an ideal pipe (without defects).

The failure pressure for the ideal pipe is given by formula:

$$P_f = \frac{2wt \cdot \sigma_f}{D} \tag{22}$$

The basic equation (22) for each code varies depends on which expression for the flow stress is utilized. Since all the above codes are either SMYS-based or UTS-based, in eqn (22) the UTS or SMYS is used, instead of the flow stress:

$$P_f = \frac{2wt \cdot SMYF}{D} \quad \text{or} \quad P_f = \frac{2wt \cdot UTS}{D} \tag{23}$$

The reliability / probability of failure of the ideal pipe were assessed using the GCE method. The analysis was performed on a real pipeline segment, which parameters are defined in Table III.

Table III. Initial data. Nominal values of pipeline parameters

Parameter	Symbol	Value
External pipeline diameter, mm	D	325.00
Pipe wall thickness, mm	wt	9.00
Specified minimum yield strength, MPa	$SMYS$	245.00
Ultimate tensile strength, MPa	UTS	410.00
Design operating pressure, MPa	P_{op}	6.40

Probabilistic characteristics of the pipeline geometric parameters were calculated on the assumption (Jiao, Sotberg and Igland, 1995) that these parameters are normally distributed.

The maximum deviation for the outside diameter is ± 2 mm, for pipe wall thickness – +0.6 and -0.8 mm. Assume that the pipeline parameters are normally distributed. Then, using the three-sigma rule and the tolerances as specified by Table III, it is possible to calculate their mean and standard deviation. The mean of pipeline diameter is equal to 325 mm and standard deviation SD = 2/3 mm; the mean of pipe wall thickness is equal to (9.6+8.2)/2 = 8.9 mm and the SD = (9.6-8.2)/6= 0.23 mm. For SMYS and UTS the distribution law and parameters were chosen according to Table II.

Initial data for further analysis are presented in Table III and IV.

Table IV. Initial data. Pipeline parameters, considered as random variables

Parameter	Symbol	Distribution law	Mean, mm	SD, mm
External pipeline diameter, mm	D	Normal	325.00	0.67
Pipe wall thickness, mm	wt	Normal	8.90	0.23
Specified minimum yield strength, MPa	$SMYS$	Normal	269.50	9.43
Ultimate tensile strength, Mpa	UTS	Normal	459.00	16.07
Design operating pressure, MPa	P_{op}	Normal	6.40	0.70

It is easy to show that for the distribution parameters of Table IV the probability of negative values of pipeline parameters is equal to zero. Indeed, for the pipe wall thickness its negative values are outside the 38.7 SD (8.90/0.23 ≈ 38.69) range, for the pipeline diameter – outside the 485.1 SD (325.00/0.67 ≈ 485.07) range, and for operating pressure – outside the 9.1 SD (6.40/0.70 ≈ 9.14) range. In all these cases the probability of occurrence of negative values is infinitesimal. For the pipe material properties the possibility of negative values is similarly infinitesimal.

In the deterministic setting, or considering any pipeline parameter deterministic when using the GCE method, the nominal values of pipeline parameters as defined by the Table III should be used.

The LSF for eqns (23) take the form:

$$Y_1 = \frac{2wt \cdot SMYS}{D} - P_{op} \quad and \quad Y_2 = \frac{2wt \cdot UTS}{D} - P_{op} \tag{24}$$

Introduce following notation: $D = x_1$, $wt = x_2$, $SMYS(UTS) = x_3$, $P_{op} = x_4$. Applying the GCE method and performing all necessary calculations as prescribed by formula (15), the POFs of the ideal pipe, calculated using the basic eqns. (23), are equal to:

$$F_1(0) = P\left(Y_1(t) < 0\right) = 7.39 \cdot 10^{-19}$$

$$F_2(0) = P\left(Y_2(t) < 0\right) = 1.55 \cdot 10^{-47}$$

These are the hidden levels of pipeline reliability embedded in the basic methodologies (23) of international codes for pipeline design.

Analyze now the sensitivity of the POF for the defect-free pipeline to the randomness of parameters in the LSFs Y_1, Y_2. The obtained results are shown in Table V.

Table V. The POF of geometrically ideal pipelines, calculated by the GCE method, based on methodologies (23), using SMYS or UTS

Formula (23) based on	All pipeline parameters are RV	All parameters of PL are RV, except			
		Diameter	Pipe wall thickness	SMYS or UTS	Operating pressure
SMYS	$7.39 \cdot 10^{-19}$	$7.10 \cdot 10^{-19}$	$4.38 \cdot 10^{-22}$	$6.71 \cdot 10^{-26}$	$8.74 \cdot 10^{-39}$
UTS	$1.55 \cdot 10^{-47}$	$1.31 \cdot 10^{-47}$	$1.44 \cdot 10^{-62}$	$9.28 \cdot 10^{-86}$	$5.63 \cdot 10^{-66}$

According to Table V, as expected, the highest POF of the ideal pipe belongs to the case when all parameters were considered as RVs. POF is most sensitive to operating pressure, when SMYS is used in eqn.(23), and to the ultimate strength, when UTS is used in eqn. (23), since in these cases the POF is minimal. Further down, in order of decreasing sensitivity of POF, are SMYS and pipe wall thickness, in the case of using SMYS in eqn. (23), and the operating pressure and pipe wall thickness in the case of using UTS in eqn. (23). In both cases, POF is not sensitive to the pipeline diameter.

Now analyze the sensitivity of POF of an ideal pipe to the randomness of pipeline parameters, when using international codes B31G, B31Gmod, Shell92, DNV и Battelle. For an ideal pipe these codes differ only by the coefficient at SMYS or UTS, which either are more than or equal to one. The obtained results are shown in Table VI.

Table VI. The POF of a geometrically ideal pipeline, calculated by the GCE method, using international codes for pipeline design

Code	All pipeline parameters are RV	All parameters of PL are RV except			
		Diameter	Pipe wall thickness	SMYS or UTS	Operating pressure
B31G	$2.39\cdot10^{-23}$	$2.27\cdot10^{-23}$	$6.16\cdot10^{-28}$	$1.30\cdot10^{-33}$	$3.80\cdot10^{-44}$
B31Gmod	$1.34\cdot10^{-34}$	$1.20\cdot10^{-34}$	$2.13\cdot10^{-44}$	$1.41\cdot10^{-46}$	$1.76\cdot10^{-66}$
Battelle	$1.83\cdot10^{-48}$	$1.53\cdot10^{-48}$	$5.15\cdot10^{-65}$	$1.49\cdot10^{-87}$	$6.61\cdot10^{-66}$
DNV	$1.88\cdot10^{-41}$	$1.64\cdot10^{-41}$	$2.14\cdot10^{-53}$	$7.05\cdot10^{-71}$	$3.52\cdot10^{-61}$
Shell92	$1.55\cdot10^{-47}$	$1.31\cdot10^{-47}$	$1.44\cdot10^{-62}$	$9.28\cdot10^{-86}$	$5.63\cdot10^{-66}$

According to Table VI, as in the previous example, for all codes the POF of an ideal pipeline has the maximal value, when all the pipeline parameters are considered as RVs. The POF of an ideal pipe is most sensitive to pressure, for SMYS-based codes (B31G, B31Gmod), and to the ultimate strength, for UTS-based codes (Battelle, DNV, Shell92), since in these cases the POF is minimal. Further down, in order of decreasing sensitivity of the POF, are SMYS and pipe wall thickness, for the SMYS-based codes, and operating pressure and pipe wall thickness, for the UTS-bases codes. In both cases, POF is not sensitive to the pipeline diameter.

4. Analysis of sensitivity of pipeline defects POF to the randomness of its parameters

Now consider the same pipeline with some typical (representative) dangerous defects. Assume that as a result of a pipeline inspection two most dangerous characteristic defects of the metal loss type were found, with depths equal to 20% wt (1.8 mm) and 42% wt (3.78 mm) and lengths - 246 mm and 70 mm, respectively. Distribution parameters of sizes of these defects are shown in Table VII.

Table VII. Initial data. Distribution parameters of sizes of most dangerous defects

Parameter	Symbol	Distribution law	Mean, mm	SD, mm
Parameters of defect #1				
Depth	d_1	Normal	2.25	0.92
Length	l_1	Normal	246.00	15.63
Parameters of defect #2				
Depth	d_2	Normal	5.62	0.99
Length	l_2	Normal	70.00	15.63

Calculate the POF of these defects, varying the number of pipeline and defects parameters, which are considered as RVs. This allows tracing the contribution of each parameter of the pipeline system to the POF of the pipeline defective cross section.

Calculate now the failure pressure for these two defects, using the nominal values of the defect sizes, as listed in Table III. The obtained results are shown in Table VIII and Fig. 1.

Table VIII. The burst pressure of most dangerous defects

Code	Burst pressure, MPa, for	
	Defect #1	Defect #2
B31G	11.20	11.83
B31Gmod	14.57	12.98
Battelle	18.80	17.40
DNV	19.29	17.78
Shell92	16.30	12.95

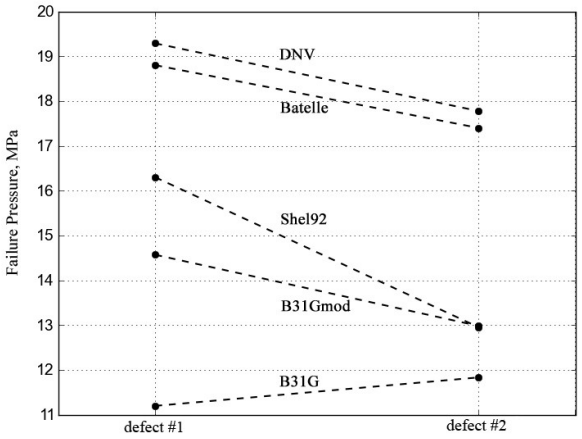


Figure 1. Burst pressure of defects under consideration

According to Fig. 1 and Table VIII, the most conservative code is the B31G, and, according to this code, the longer defect (#1) is more dangerous than the deeper. Further, in descending order of conservatism, follow codes B31Gmod, Shell92, Battelle and DNV.

Now calculate POF of the defects under consideration, varying the number of parameters of the defects and the pipeline, which are considered as RVs. The obtained results are presented in Table IX and Figs. 2 and 3.

Table IX. The POF of the defects under consideration, depending on the number of parameters of the pipeline and the defects, considered to be RVs

Code	Probability of failure	
	defect #1	defect #2
Case #1: all parameters are RVs, except diameter		
B31G	$3.55 \cdot 10^{-3}$	$2.06 \cdot 10^{-6}$
B31Gmod	$4.23 \cdot 10^{-8}$	$4.40 \cdot 10^{-5}$
Battelle	$1.46 \cdot 10^{-10}$	$1.88 \cdot 10^{-8}$
DNV	$1.81 \cdot 10^{-10}$	$3.53 \cdot 10^{-6}$
Shell92	$1.31 \cdot 10^{-7}$	$1.59 \cdot 10^{-3}$
Case #2: all parameters are RVs, except pipe wall thickness		
B31G	$9.71 \cdot 10^{-4}$	$5.22 \cdot 10^{-7}$
B31Gmod	$1.23 \cdot 10^{-8}$	$2.05 \cdot 10^{-5}$
Battelle	$2.29 \cdot 10^{-11}$	$2.62 \cdot 10^{-9}$
DNV	$2.92 \cdot 10^{-11}$	$1.31 \cdot 10^{-6}$
Shell92	$5.25 \cdot 10^{-8}$	$1.21 \cdot 10^{-3}$
Case #3: all parameters are RVs, except defect length and depth		
B31G	$3.53 \cdot 10^{-3}$	$1.64 \cdot 10^{-11}$
B31Gmod	$1.49 \cdot 10^{-21}$	$1.38 \cdot 10^{-13}$
Battelle	$3.48 \cdot 10^{-32}$	$3.74 \cdot 10^{-24}$
DNV	$5.07 \cdot 10^{-33}$	$2.44 \cdot 10^{-22}$
Shell92	$2.14 \cdot 10^{-25}$	$2.29 \cdot 10^{-11}$
Case #4: all parameters are RVs, except SMYS or UTS		
B31G	$3.54 \cdot 10^{-3}$	$2.49 \cdot 10^{-10}$
B31Gmod	$1.68 \cdot 10^{-8}$	$7.43 \cdot 10^{-7}$
Battelle	$1.83 \cdot 10^{-11}$	$1.23 \cdot 10^{-8}$
DNV	$2.53 \cdot 10^{-11}$	$3.35 \cdot 10^{-6}$
Shell92	$4.86 \cdot 10^{-8}$	$1.88 \cdot 10^{-3}$
Case #5: all parameters are RVs, except operating pressure		
B31G	$3.53 \cdot 10^{-3}$	$7.70 \cdot 10^{-8}$
B31Gmod	$2.76 \cdot 10^{-9}$	$1.55 \cdot 10^{-5}$
Battelle	$2.16 \cdot 10^{-11}$	$4.68 \cdot 10^{-9}$
DNV	$3.25 \cdot 10^{-11}$	$2.12 \cdot 10^{-6}$
Shell92	$3.75 \cdot 10^{-8}$	$1.32 \cdot 10^{-3}$
Case #6: all parameters are RVs		
B31G	$3.54 \cdot 10^{-3}$	$2.06 \cdot 10^{-6}$
B31Gmod	$4.24 \cdot 10^{-8}$	$4.41 \cdot 10^{-5}$
Battelle	$1.47 \cdot 10^{-10}$	$1.89 \cdot 10^{-8}$
DNV	$1.82 \cdot 10^{-10}$	$3.53 \cdot 10^{-6}$
Shell92	$1.31 \cdot 10^{-7}$	$1.59 \cdot 10^{-3}$

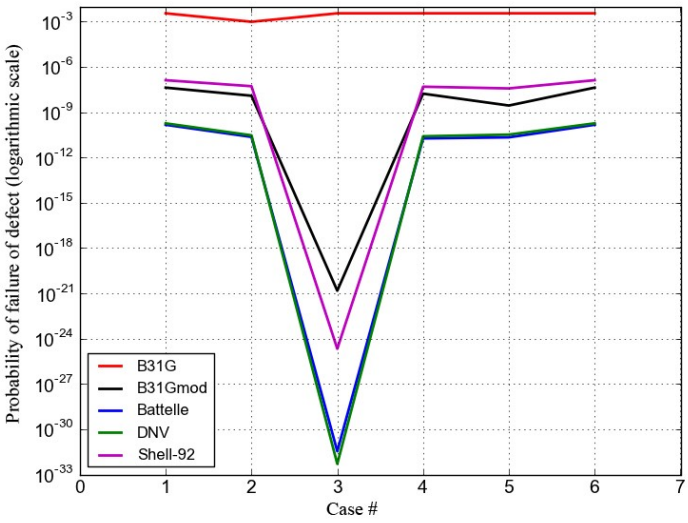


Figure 2. The POF of defect #1, depending on used codes and the number of considered design cases (logarithmic scale)

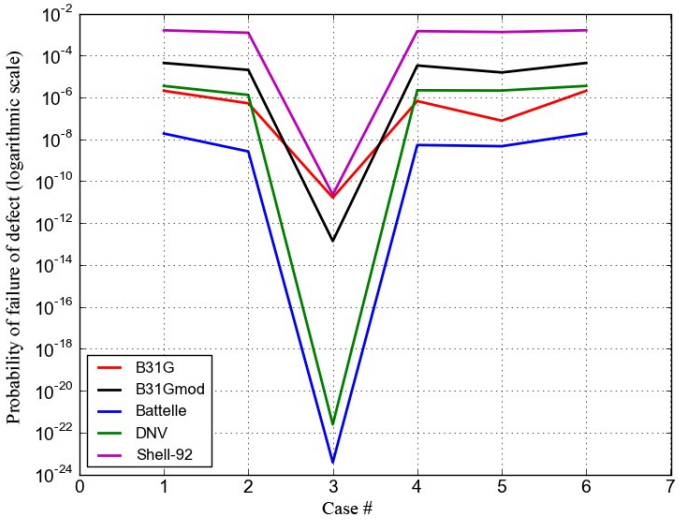


Figure 3. The POF of defect #2, depending on used codes and the number of considered design cases (logarithmic scale)

According to Fig. 2 and Table IX for the first, longer defect following conclusions could be made:

- The B31G code is the most conservative. Next follow codes Shell92, B31Gmod. The codes DNV and Battelle are least conservative.
- With respect to POF, the B31G code is absolutely not sensitive to any of the considered cases except for the second, when the pipe wall thickness was assumed to be a deterministic value (DV). This can be

explained by the fact that the formula for calculating longer defects does not depend on the Folias factor.
- Assessments of the POF, obtained using codes Battelle and DNV are almost identical.
- Methods of the Battelle, DNV, and Shell92 codes are equally sensitive to randomness of the pipe wall thickness and operating pressure, because in cases #2 and #5 the assessments of the POF, obtained by using these codes, are practically identical.

The results shown in the Fig. 3 and Table IX for the second, deeper defect lead to the following conclusions:

- The Shell92 code is the most conservative. Next follow B31Gmod and DNV.
- Codes B31Gmod, Battelle, DNV, and Shell92 are equally sensitive to the randomness of pipe wall thickness and operating pressure, due to the fact that in cases #2 and #5 the values of POF are almost identical.
- The Shell92 code is sensitive only to randomness of defect parameters (case #3), because in other cases the assessments of POF are the same.

Summarizing the overall results of the conducted research, for a given pipeline with only two characteristic defects permits following generalized, but conditional (due to limited scope of analysis) conclusions:

- All standards, except B31G, were found to be very sensitive to case #3, in which the defects parameters were considered as DV. In this case the assessments of POF are minimal for all the used codes.
- All codes are not sensitive to the randomness of the pipe diameter. This is shown by the fact that the assessment of POF for the case #1 coincides with the assessment for case #6, where all parameters are considered as RVs.
- With respect to POF, all codes are most sensitive to the randomness of the defect depth and length. The difference between the POF for the case when the defect parameters are considered as DVs and the case when they are RVs is very significant. When the parameters of defects are DV, the POF is minimal, even when considering deep and long defects.
- For all codes the second most sensitive parameter of the POF is pipe wall thickness. All codes are less sensitive to the randomness of SMYS / UTS and the operating pressure.

It should be noted that quantitatively the sensitivity of the POF will change with the change of the stochastic characteristics of parameters in consideration, but the ranking of cases in Table IX will remain the same.

The results of the above analysis demonstrate that it is not prudent to consider the most important pipeline parameters (pipe wall thickness, SMYS and UTS), load (pressure) and the defects parameters (depth and length) as deterministic values. Using deterministic approach is fraught with inconsistent estimates of pipeline reliability and the degree of danger of its defects.

References

Alkazraji D. A Quick Guide to Pipeline Engineering. – Cambridge, England, Series editor: Clifford Matthews, Woodhead Publishing Limited and Matthews Engineering Training Ltd, 2008. – 157 p.

ANSI/ASME B31G-1991. Manual for Determining the Remaining Strength of Corroded Pipelines. National Standard of USA. – New York: ASME. – 1991. – 140 p.

API 5L. Specification for Line Pipe. Exploration and Production Department. 43rd edition, Washington: American Petroleum Institute, DC, 2004

CAN3-Z183-M86. Oil Pipeline System. National Standard of Canada. – Ontario: CSA, 178 Rexdale Blvd., September, 1986. – 185 p.

C-FER Project No. L128. Reliability-based Design and Assessment of Onshore Natural Gas Transmission Pipelines - A Proposed Standard, 2005

DNV-RP-F101. Corroded pipelines. Recommended Practice. – Norway: Det Norske Veritas, 2004. – 42 p.

Jiao G., Sotberg T. and Igland R The Superb Project: Linepipe Statistical Properties and Implications in Design of Offshore Pipelines // OMAE – Pipeline Technology, ASME. – 1997. – Vol. V. – P. 45-56.

Jiao G., Sotberg T. and Igland R. Report № STF70 F95212 – SUPERB 2M Statistical Data – Basic Uncertainty Measures for Reliability Analysis of Offshore Pipelines // Submarine Pipelines. –June 1995. – Superb Project № 700411

Kiefner J.F. and Vieth P.H. A Modified Criterion for Evaluating the Remaining Strength of Corroded Pipe // AGA Pipeline Research Committee. Report PR 3-805. – 1989. – 78 p.

Matsushima I. Carbon Steel-Corrosion by Soils. In Uhlig's Handbook of Corrosion - Second Edition edited by R.W. Revie. – NY: John Wiley & Sons, Inc., 2000. – 1302 p.

Poluyan L.V., Bushinskaya A.V., Malyukova M.G. and Timashev S.A. On-line Method of Reliability Analysis of Pipelines with Growing Defects // ICOSSAR. – Japan. – 2009. Paper #ICOSSAR09-0155.

Ritchie D. and Last S. Burst Criteria of Corroded Pipelines - Defect Acceptance Criteria // Proceedings of the EPRG/PRC 10th Biennial Joint Technical Meeting on Line Pipe Research. – Cambridge, UK. – 1995. – Paper

Sheikh A.K., Boah J.K., Hansen D.A. Statistical Modelling of Pitting Corrosion and Pipeline Reliability [Text] // Corrosion. – 1999. – V. 46. – №3. – P. 190–197

Stephens D.R. and Leis B.N. Development of an Alternative Criterion for Residual Strength of Corrosion Defects in Moderate to High-Toughness Pipe // Proceedings of the Third International Pipeline Conference. – Calgary, Canada, American Society of Mechanical Engineers. – 2000. – P. 781-792.

Timashev S.A., Malyukova M.G., Poluyan L.V. and Bushinskaya A.V. Internet – Oriented Method Of Reliability Analysis Of On-Shore Pipelines With Growing Defects // Proceedings of IPC Conference. – 2008. – Paper № IPC2008-64545

Bounding the dependence measures for spatial uncertainties

Wim Verhaeghe, Wim Desmet and Dirk Vandepitte
Dept. of Mechanical Engineering, KU Leuven, B3001 Heverlee, Belgium,
wim.verhaeghe@mech.kuleuven.be

Isaac Elishakoff
Dept. of Ocean and Mechanical Engineering, Florida Atlantic University, Boca Raton, FL 33431, USA

David Moens
Dept. of Appl. Engineering, Lessius Hogeschool, K.U.Leuven Association, B2860 St-Katelijne-Waver, Belgium

Abstract. The analysis and design of mechanical engineering systems requires to take into account the influence of uncertainties on the system's performance. Depending on the available amount of information, the designer or analyst can choose from a wide variety of methods in the probabilistic (see e.g. (Schuëller, 2001)) or non-probabilistic (see e.g. (Elishakoff and Ohsaki, 2010), (Moens and Hanss, 2011)) approaches to describe the uncertainties. However, the selection of a suitable uncertainty model for the different uncertainties most often is not enough to yield satisfactory information on the reliability or bounds of the system's performance. A crucial piece of information appears to be the dependence of the uncertain variables. This is especially the case for uncertainties with a spatial character, e.g. material properties or distributed loads.

The study gives an overview of the existing probabilistic and non-probabilistic methods to represent this kind of dependencies. In the probabilistic setting the concepts of the covariance function associated with a random field, a copula and several correlation measures are treated. In the non-probabilistic setting the concepts of interval fields, convex modeling and interactive fuzzy numbers are reviewed. Of special interest is the ability to bound these dependence measures. For the case of a spatial uncertainty, this generally comes down to specifying the maximum distance between points that are influencing each other. Points further away from each other than this distance are considered practically independent. For points closer to each other than this distance the interaction may be described, introducing a notion of perfect dependence. Finally, of utmost importance is to study the effect of the bounds on the dependence on the uncertainty in the system's performance.

Keywords: Random field, Interval field, Finite Element analysis

1. Introduction

A description of dependence can take many faces. First of all, the word itself has different meanings. Mosteller and Tukey (1977) emphasize: "We must be clearer about the abused word *dependence* and its relatives." (Drouet Mari and Kotz, 2001) On the one hand, dependence may mean perfect dependence, i.e. if one knows the value of x, then one knows exactly the value of y. On the other hand, dependence can be more flexible, i.e. when we know x, we may know something more about y as opposed to the situation

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

when we know nothing about the value of x. This piece of information can be expressed in a probabilistic framework (section 2) as well as a non-probabilistic framework (section 3).

The modeling of a spatial uncertainty exemplifies the need for a proper description of dependence in uncertainty modeling. It is a given that spatially distributed model parameters for locations adjacent to each other assume uncertain but very similar values. The uncertain values for points further away from each other can be very dissimilar. The crucial piece of information is the notion of spatial closeness. This notion consists in actually two things. The first is being able to measure how close points are to each other in a numerical model. The second is to be able to compare this distance to a reference to assess whether higher or lower dependence is assumed between these two points. For the details of a distance measuring method in a finite element model, the reader may consult (De Mulder et al., 2012). Specifying a value for the reference distance depends on the actual uncertainty modeling framework in use. In a probabilistic framework, the concept of correlation length is widely used, whereas in the non-probabilistic world a related concept does not (yet) exist. Next issue is to take into account this dependence when propagating input uncertainty to output uncertainty: According to (Kurowicka and Cooke, 2006) "... an essential part of uncertainty analysis is the analysis of dependence. Indeed, if all uncertainties are independent, then their propagation is mathematically trivial (though perhaps computationally challenging). Sampling and propagating independent uncertainties can easily be trusted to the modellers temselves. However, when uncertainties are dependent, things become much more subtle, and we enter a domain for which the modellers' training has not prepared them." From a practical perspective, a tool is needed to translate the spatial dependence information given by an expert to a representative set of realisations of the uncertain model parameter. Section 4 sheds some light on this topic.

An important feature of the spatial dependence modeling is yet untouched. Although the more flexible notion of dependence (i.e. *not* the perfect dependence: if $x = a$ then $y = b$) and its related probabilistic and non-probabilistic descriptions allows a more versatile treatment of the dependence phenomenon, it becomes increasingly clear that the assignment of one single value to a dependence descriptor is still far from feasible. Either because the data set on which such a single value assignment could be made is too small or the data set simply does not exist and the dependence information is based on expert knowledge. To quote again (Kurowicka and Cooke, 2006) "Engineers and scientists sometimes cover their modesty with churlish acronyms designating the source of ungrounded assessments. 'Wags' (wild-ass guesses) and 'bogsats' (bunch of guys sitting around a table) are two examples found in published documentation." It is suggested in section 5 to put bounds on the dependence measures, instead of assigning a single value to them.

Finally, in section 6 a numerical example is given to show the effect of such bounds on a probabilistic and non-probabilistic dependence measure in the context of spatial uncertainty modeling.

2. Probabilistic dependence measures

In scientific literature notions and definitions of *independence* preceded the notion of *dependence* (the latter was just regarded as the negation of the former). Related to this seems the fact that it was easier to understand independence. "Saying that variables are *not* independent does not say much about their joint distribution. What is the nature of this dependence? How dependent are they? How can we measure the dependence? These questions must be addressed in building a dependence model." (Kurowicka and Cooke, 2006) The first probabilistic concept of dependence (correlation) emerged at the end of the 19^{th} century in the field of

social study and psychology. “The concept of correlation (and its modifications) introduced by F. Galton in 1885 dominated statistics during some 70 years of the 20-th century, practically serving as the only measure of dependence, often resulting in somewhat misleading conclusions.” (Drouet Mari and Kotz, 2001) On the other hand, the correlation has great merit. “By now, over a century later, contemporary scientists often take the correlation coefficient for granted. It is not appreciated that before Galton and Pearson, the only means to establish a relationship between variables was to deduce a causative connection. There was no way to discuss - let alone measure - the association between variables that lack a cause-effect relationship.” (Rodgers and Nicewander, 1988) In the following, an overview of several dependency measures is given. First, the scalar aggregate (global) measures of bivariate dependence *product moment correlation*, *rank correlation*, *Kendall's tau* and *relative entropy* are discussed. Next, the more thorough (local) measure of dependence *copula* is discussed.

2.1. PRODUCT MOMENT CORRELATION

A historic account of the product moment correlation, together with as much as 13 ways to look at it can be found in (Rodgers and Nicewander, 1988). The product moment correlation of random variables X,Y with finite expectations $E(X)$, $E(Y)$ and finite variances σ_X^2, σ_Y^2, is

$$\rho(X,Y) \equiv \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y} \tag{1}$$

Some properties of the product moment correlation are listed below:

- The product moment correlation depends on the marginal distribution functions F_X and F_Y.
- The product moment correlation is bounded by $-1 \leq \rho(X,Y) \leq 1$ and it's minimum and maximum (not necessarily -1 and 1, as in the example (Kurowicka and Cooke, 2006), pages 29-30) are attained for X and Y countermonotonic and comonotonic, respectively. We say that random variables X and Y are comonotonic if there is a strictly increasing function G such that $X = G(Y)$. X and Y are countermonotonic if X and $-Y$ are comonotonic.
- The product moment correlation is invariant under linear strictly increasing transformations of X or Y, but is not invariant under non-linear strictly increasing transformations.
- If X and Y are independent, then $\rho(X,Y) = 0$. The reverse is not generally true.

2.2. RANK CORRELATION

The rank correlation or Spearman correlation was introduced by Spearman in 1904. The rank correlation of random variables X, Y with cumulative distribution functions F_X and F_Y is

$$\rho_r(X,Y) \equiv \rho(F_X(X), F_Y(Y)) \tag{2}$$

Since for any X with a continuous invertible F_X, $F_X(X)$ is uniform on $\lfloor 0,1 \rceil$, the rank correlation is a correlation of random variables transformed to uniform random variables. This leads to the following properties:

- The rank correlation is independent of marginal distributions.
- The rank correlation is bounded by $-1 \leq \rho_r(X,Y) \leq 1$ and it's minimum (-1) and maximum (1) are attained for X and Y countermonotonic and comonotonic, respectively.
- The rank correlation is invariant under non-linear strictly increasing transformations.
- If X and Y are independent, then $\rho_r(X,Y) = 0$. The reverse is not generally true.

In (Kurowicka and Cooke, 2006) an efficient technique is presented to extract information on the rank correlation from experts. The technique is based on an indirect question: "Suppose Y were observed in a given case and its values were found to lie above the median value for Y; what is your probability that, in this same case, X would also lie above its median value?" Formally this comes down to assess

$$\pi_{\frac{1}{2}}(X,Y) \equiv P\left(F_X(X) > \frac{1}{2} | F_Y(Y) > \frac{1}{2}\right) \tag{3}$$

Based on the minimum information copula (see below), the probability assigned by the expert can be directly related to a rank correlation. It is obvious that a probability 0 would mean X and Y are anti-correlated, a value $\frac{1}{2}$ would suggest a rank correlation equal to 0 whereas a value 1 leads to completely rank-correlated X and Y.

2.3. KENDALL'S TAU

Let (X_1, Y_1) and (X_2, Y_2) be two independent pairs of random variables with joint distribution function F and marginal distributions F_X and F_Y. Kendall's rank correlation, also called Kendall's tau (1938) is given by

$$\tau \equiv P\left[(X_1 - X_2)(Y_1 - Y_2) \geq 0\right] - P\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right] \tag{4}$$

The following properties hold for Kendall's tau:

- Kendall's tau is independent of marginal distributions.
- Kendall's tau assumes values between -1 and 1.
- Kendall's tau is invariant under non-linear strictly increasing transformations.
- If X and Y are independent, then $\tau(X,Y) = 0$. The reverse is not generally true.

For a discussion of two correlation measures, namely *sup correlation* and *monotone correlation*, where a zero value implies independence, see (Devroye, 1986) pp.574-576.

2.4. RELATIVE ENTROPY

Consider the pair of random variables (X,Y) with $f(x,y)$ the joint density and $f_1(x)$ and $f_2(x)$ the marginal densities. Then the relative entropy is defined as:

$$\delta_{X,Y} \equiv \int\int f(x,y)\, log\left(\frac{f(x,y)}{f_1(x)f_2(y)}\right) dxdy \tag{5}$$

The entropy of the density $f(x, y)$ is compared with the maximum attainable entropy, namely when X and Y are independent. For independent X and Y, $\delta_{X,Y}$ is zero, and for maximal dependence, $\delta_{X,Y}$ approaches infinity. The concept of maximum entropy or equivalently minimum information will be reappearing in the next section on copulae.

2.5. COPULA

Copulae are tools for modeling dependence of several random variables. In particular a copula allows to seperate the effect of dependence from the effect of marginal distributions in a joint distribution. The term copula was first introduced by Sklar in 1959 (Schmidt, 2006). A copula C is defined as a function which is a cumulative distribution function with uniform marginals. Random variables X and Y are joined by copula C if their joint distribution can be written as

$$F_{XY}(x, y) = C(F_X(x), F_Y(y)) \tag{6}$$

For the bivariate case, a copula is the joint distribution of two random variables that are each converted to the uniform distribution by means of their respective marginal distribution functions. A copula is always contained in between the Fréchet-Hoeffding bounds C_L and C_U (see Figure 1). C_L represents the case when all of the probability mass is spread uniformily on the main diagonal $v = 1 - u$ and C_U is attained when all of the mass is on the main diagonal $v = u$.

$$C_L = max(u + v - 1, 0) \leq C(u, v) \leq C_U = min(u, v) \tag{7}$$

with $(u, v) \in \lfloor 0, 1 \rceil^2$. Next, two copulae from a sheer endless list of copulae are described (based on (Kurow-

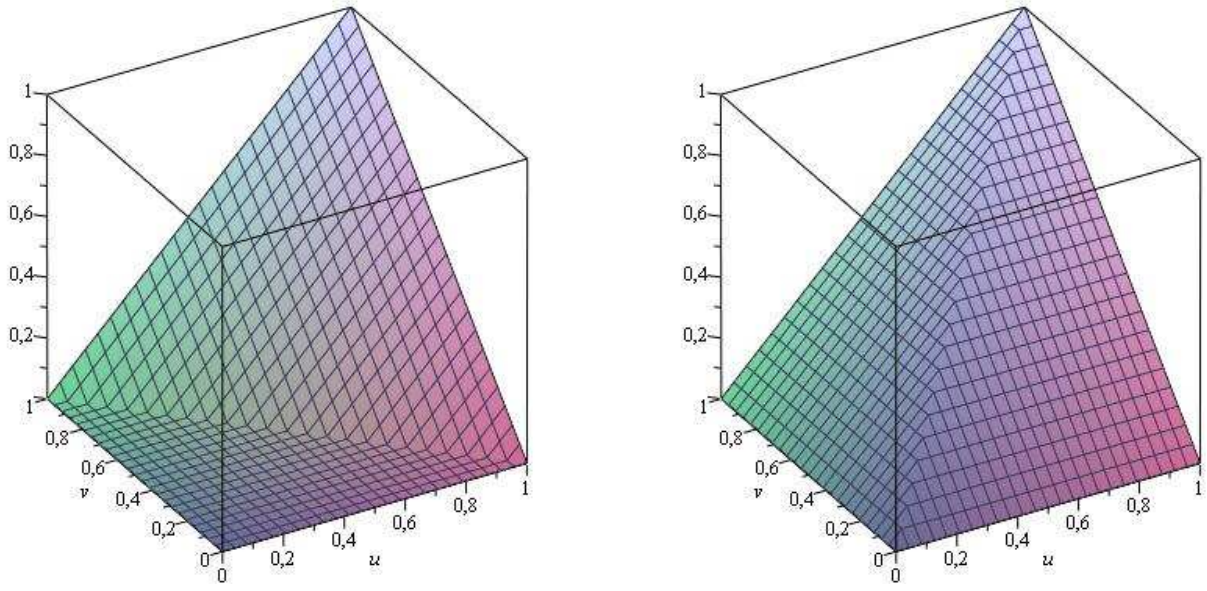


Figure 1. The lower and upper Frchet-Hoeffding bounds C_L and C_U

icka and Cooke, 2006)) as they are of particular interest in next section's discussion on non-probabilistic dependence measures.

2.5.1. *Diagonal band copula*
One natural generalization of the bounding copula C_U is the diagonal band copula introduced by Cooke and Waij in 1986. In contrast to C_U, for possitive correlation the mass is concentrated on the diagonal band with vertical bandwidth $b = 1 - a$. Mass is distributed uniformly on the inscribed rectangle and is uniform but 'twice as thick' in the triangular corners. A clear relationship exists between the product moment correlation and the parameter a of the diagonal band copula.

2.5.2. *Minimum information copula*
The minimum information copula was introduced and studied by Meeuwissen in 1993. The construction is based on the fact that for any $\rho_r \in \rfloor -1, 1 \lceil$ there is a unique bivariate joint distribution satisfying the following constraints:

- the marginal distributions are uniform on $I = [-\frac{1}{2}, \frac{1}{2}]$
- the rank correlation is $\rho_r \in \rfloor -1, 1 \lceil$
- the distribution has minimal information relative to uniform distribution (or maximum entropy as defined higher) among all distributions with rank correlation ρ_r.

The minimum information copula is attractive because it realizes a specified rank correlation by 'adding as little information as possible' to the product of the marginals. Its main disadvantage is that it does not have a closed functional form. All calculations with this copula must involve numerical approximations.

3. Non-Probabilistic dependence measures

All too often the following typical 'jump of thought' is made, it is first noted that "..., the dependence is obviously not deterministic but of a stochastic nature." (Drouet Mari and Kotz, 2001), and then a book all about probabilistic dependence follows. Present authors maintain however that if something is not deterministic it is not necessarily probabilisitic. The non-probabilistic methods cover a very large set of alternative uncertainty treatments, of which (together with probabilistic theory) an excellent overview is given in (Klir, 2006). We will limit our focus to crisp and so called graded possibilities. In crisp possibilities a clear distinction is made between members and non-members of a set, by assigning membership level one and zero, respectively. A typical example of a crisp set is an interval. In graded possibilities the full range of membership degrees between zero and one is available. A typical example of a graded set is a fuzzy number.

In this framework, the treatment of independence and noninteraction overshadows, to the best of our knowledge, study in the field of dependence descriptions and measures. We follow mainly the description by (Klir, 2006) on both crisp possibilities and graded possibilities.

3.1. CRISP POSSIBILITIES AND INFORMATION TRANSMISSION

For finite sets (generalization to infinite sets is given in (Klir, 2006)) of possible alternatives Hartley proposed in 1928 to measure the amount of uncertainty of such a set by

$$H(r_E) = \log_2 |E|, \quad r_E(x) = \begin{cases} 1 & \text{when } x \in E \\ 0 & \text{when } x \notin E \end{cases} \tag{8}$$

with $|E|$ the cardinality of the set E. Assume two sets X and Y and the set $R \subseteq X \times Y$ that describe possible alternatives in some situation of interest. Then following relations between the marginal, joint and conditional Hartley measures hold:

$$H(X|Y) = H(X \times Y) - H(Y) \tag{9}$$
$$H(Y|X) = H(X \times Y) - H(X) \tag{10}$$

If possible alternatives from X do not depend on selections from Y, and vice versa, then $R = X \times Y$ and the sets R_X and R_Y (the projections of R on X and Y, respectively) are called noninteractive:

$$H(X|Y) = H(X) \tag{11}$$
$$H(Y|X) = H(Y) \tag{12}$$
$$H(X \times Y) = H(X) + H(Y) \tag{13}$$

In the general case of unknown interactivity, these equalities become inequalities $\leq$. To indicate the strength of constraint between possible alternatives in sets X and Y, the information transmission is defined as

$$T_H(X, Y) = H(X) + H(Y) - H(X \times Y) \tag{14}$$

When the sets are noninteractive, $T_H(X, Y) = 0$; otherwise, it is positive. Its maximum value is obtained if $H(X|Y) = H(Y|X) = 0$, or in other words, if the value of X is specified, only one value for Y is possible. This indicator can be considered a measure of dependence in this paper.

3.2. GRADED POSSIBILITIES

In the framework of graded possibilities a value between zero and one is assigned to every singleton of a set X by the basic possibility function $r(x)$, with the requirement of possibilistic normalization to assign at least to one x the value 1. The possibility assigned to a subset A of X is determined by

$$Pos(A) = \max_{x \in A} \{r(x)\} \tag{15}$$

Based on a joint possibility function $r(x, y)$ defined over $X \times Y$ the marginal possibility functions are defined as

$$r_X(x) = \max_{y \in Y} \{r(x, y)\}, \quad \forall x \in X \tag{16}$$
$$r_Y(y) = \max_{x \in X} \{r(x, y)\}, \quad \forall y \in Y \tag{17}$$

If it is known that variables X and Y do not interact, then the joint possibility function r is defined by

$$r(x,y) = \min\{r_X(x), r_Y(y)\} \tag{18}$$

Among all joint possibility profiles that are consistent with the given marginal possibilities, this one based on the assumption of noninteractive marginals, is the largest one. The definition of conditional possibilities can go many ways (Fonck, 2006). We follow the definition by (Hisdal, 1978):

$$r_{X|Y} = \begin{cases} r(x,y) & \text{when } r_Y(y) > r(x,y) \\ 1 & \text{when } r_Y(y) = r(x,y) \end{cases} \tag{19}$$

$$r_{Y|X} = \begin{cases} r(x,y) & \text{when } r_X(x) > r(x,y) \\ 1 & \text{when } r_X(x) = r(x,y) \end{cases} \tag{20}$$

To define possibilistic independence based on this definition of conditional possibilities, one can again go many ways. The key is to compare $r_X(x)$ and $r_{X|Y}(x|y)$, this can be done in at least three ways as suggested in (de Campos and Huete, 1999). Possibilistic independence can be defined based on: equality $r_X(x) = r_{X|Y}(x|y)$ (not modifying information), inequality $r_X(x) \leq r_{X|Y}(x|y)$ (not gaining information) or similarity $r_X(x) \simeq r_{X|Y}(x|y)$ (obtaining similar information, but specification of the similarity measure is needed). Here, as in (Klir, 2006), the equality operator is adopted. With such a definition possibilistic independence implies possibilistic noninteraction, but not the other way around.

This definition of noninteraction or independence does not give us a measure of interaction or dependence. First attempts to come up with such a measure are apparently found in (Fuller and Majlender, 2004).

4. Spatial uncertainties and dependence measures

As mentioned in the introduction, the modeling of a spatial uncertainty calls for a dependence measure. In particular, one needs a dependence description in function of the distance between points.

4.1. Probabilistic spatial uncertainty and correlation length

In the probabilistic framework, the concept of a random field (Vanmarcke, 1993) is well developed. In its application the crucial element is the specification of the correlation structure. For homogeneous random fields, this correlation structure describes the value of the correlation as a function of the distance between two points. A crucial parameter in this function is the correlation length as made clear in the illustrative sensitivity study (Charmpis et al., 2007). The parameter largely dominates the discretization of a random field. For an overview of discretization methods applicable to finite element analysis, the reader is referred to the excellent report by (Sudret And Der Kiureghian, 2000). Three groups of discritization are identified: point discretization, average discretization and series expansion methods.

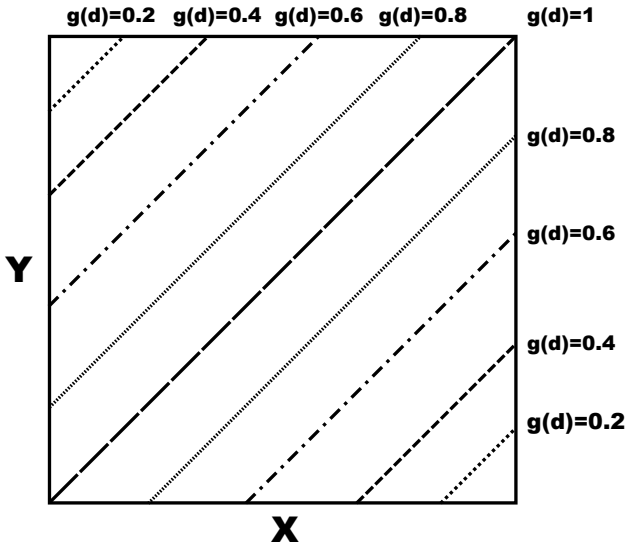


Figure 2. The possible values of $R \subseteq X \times Y$ as a function of $g(d)$. The limits on the possible values are given by identically dashed lines.

4.2. NON-PROBABILISTIC SPATIAL UNCERTAINTY AND DOMAIN OF INFLUENCE

Apparently a link between a crisp possibilistic measure of dependence and spatial uncertainty is not yet formulated in literature. Our suggestion consists in specifying a function $g(d) : d \to \lfloor -\infty, 1 \rceil$ with d the (non-negative) distance between two points in a model. To every value $g(d)$ corresponds a set $R \subseteq X \times Y$ of possible values. For $g(d) \leq 0$ the set $R = X \times Y$, for $g(d) = 1$ the set R reduces to the single line $X = Y$. In other words, the information transmission becomes maximal and the possible alternatives for X given Y reduce to one, if the distance between X and Y reduces to zero. All this leads to the following conditions on $g(d)$:

$$\begin{cases} g(0) = 1 \\ g(d_1) > g(d_2) \quad \text{for } d_1 < d_2 \end{cases} \tag{21}$$

A simple example of such a function $g(d)$ is

$$g(d) = 1 - \frac{d}{a} \tag{22}$$

with $a > 0$ a parameter specifying the domain of influence. If $d < a$ then $R \subset X \times Y$, for $d > a$ $R = X \times Y$.

For values $0 < g(d) < 1$ the domain of possible values R can take many shapes, our focus is restricted to two cases. The first is a shape similar to the diagonal copula discussed above. Figure 2 illustrates the concept, where the limits on the possible values are given by identically dashed lines. A clear link between this representation and bounding the spatial derivative $0 < |X'| \leq z$ of a model parameter (in addition to bounding its value $l \leq X \leq u$) can be established. Let two locations in a model seperated by a distance d be given index 1 and 2. Given a value for x_1, the value of x_2 is bounded by $x_1 - zd$ and $x_1 + zd$ as long as $x_1 - zd > l$ and $x_1 + zd < u$. These bounds can be directly related to the model for $g(d)$ in Eq. (22) with

$a = (l - u)/z$. An example where the bounds on the spatial derivative of an uncertain model parameter are used to describe the spatial dependence can be found in (Ben-Haim and Elishakoff, 1990).

The second shape is a nesting of ellipses for $0 < g(d) < 1$, that degenerate to a square for $g(d) = 0$ and degenerate to a line for $g(d) = 1$. The basis of a practical method to calculate the ellipses based on experimentally measured spatial data can be found in (Zhu et al., 1996).

5. Bounding dependence measures for spatial uncertainties

The actual value of the correlation length or the parameter a in Eq. (22) is seldomly known. For this reason, it is suggested to treat them as intervals. For a study of the influence of an interval correlation length on the series expansion of a random field, the reader is referred to (Verhaeghe et al., 2011). In the example below a similar analysis is performed for a point discretization of a random field. After the discretization is fixed, the interval on the correlation length, actually results in intervals on the correlations between all the discretization points.

For the non-probabilistic case, after fixing the discretization, the interval on a results in a set of domains of influence. Points that are not influencing eachother for small values of a, become important to each other with increasing a. Depending on the studied output, its interval can increase or decrease in size with increasing a.

6. Numerical example

In this section, the influence of bounds on a dependence measure for both a probabilistic and a non-probabilistic analysis is studied in the context of a numerical example.

6.1. Reliability of beam with random flexibility

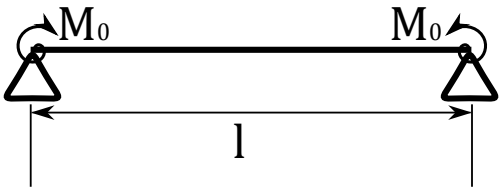


Figure 3. A simply supported beam of length l, loaded with constant moment M_0

A beam of length l is simply supported at its both ends and loaded with a constant moment M_0 (see Figure 3). The flexibility $h(x)$ is characterized by

$$h(x) = H_1\phi_1(x) + H_2\phi_2(x) \tag{23}$$

where $\phi_1(x)$, takes value 1 for the left half ($0 \leq x < 0.5l$) and has value 0 for the right half of the beam. Conversely, the function $\phi_2(x)$ takes value 1 for the right half of the beam and 0 for the left half. The amplitudes H_1 and H_2 are each a uniformly distributed random variable. As such, the flexibility is modelled

by a very coarse point discretisation (only the two points $x = 0.25l$ and $x = 0.75l$ are considered) of the random field for the flexibility.

The performance of the beam in this case is determined by the displacement difference between two points symmetrically located on both sides of its mid-point. The displacement is calculated using

$$\begin{aligned} W(x) &= M_0 \int_0^x \int_o^v h(u)dudv - M_0 \frac{x}{l} \int_0^l \int_0^v h(u)dudv \\ &= M_0 \left[H_1 \int_0^x \int_o^v \phi_1(u)dudv + H_2 \int_0^x \int_o^v \phi_2(u)dudv \right. \\ &\quad \left. - \frac{x}{l} \left(H_1 \int_0^l \int_0^v \phi_1(u)dudv + H_2 \int_0^l \int_0^v \phi_2(u)dudv \right) \right] \end{aligned} \tag{24}$$

The displacement difference between points x_1 and x_2 is thus found by

$$\Delta W_{x_1,x_2} = W(x_1) - W(x_2) \tag{25}$$

Defining $A(x) = \int_0^x \int_o^v \phi_1(u)dudv$ and $B(x) = \int_0^x \int_o^v \phi_2(u)dudv$, the reliability can be calculated as:

$$\begin{aligned} R &= P\left(|\Delta W_{x_1,x_2}| \leq \Delta W_{specified}\right) \\ &= P\left(|M_0 \left[H_1 \left(A(x_1) - \frac{x_1}{l} A(l) \right) + H_2 \left(B(x_1) - \frac{x_1}{l} B(l) \right) \right. \right. \\ &\quad \left. \left. -H_1 \left(A(x_2) - \frac{x_2}{l} A(l) \right) - H_2 \left(B(x_2) - \frac{x_2}{l} B(l) \right) \right] | \leq \Delta W_{specified} \right) \\ &= P\left(|M_0 \left[H_1 \left(A(x_1) - A(x_2) - \frac{x_1}{l} A(l) + \frac{x_2}{l} A(l) \right) \right. \right. \\ &\quad \left. \left. +H_2 \left(B(x_1) - B(x_2) - \frac{x_1}{l} B(l) + B(l) \right) \right] | \leq \Delta W_{specified} \right) \\ &= P\left(|M_0 \left[H_1 A^* + H_2 B^* \right] | \leq \Delta W_{specified} \right) \\ &= \int_0^{\overline{H_1}} \int_{\frac{-\Delta W_{specified}/M_0 - \xi_1 A^*}{B^*}}^{\frac{\Delta W_{specified}/M_0 - \xi_1 A^*}{B^*}} f_H(\xi_1, \xi_2) d\xi_2 d\xi_1 \\ &= \int_0^{\overline{H_2}} \int_{\frac{-\Delta W_{specified}/M_0 - \xi_2 B^*}{A^*}}^{\frac{\Delta W_{specified}/M_0 - \xi_2 B^*}{A^*}} f_H(\xi_1, \xi_2) d\xi_1 d\xi_2 \end{aligned} \tag{26}$$

with A^* and B^* the weights of the random variables due to the integrals of $\phi_1(x)$ and $\phi_2(x)$ respectively; $\overline{H_1}$ and $\overline{H_2}$ the upper bounds of the random variables.

Let us assume $l = 1$, $M_0 = 1$, both H_1 and H_2 uniform on $\lfloor 0.95, 1.05 \rceil$, $x_1 = 0.4$, $x_2 = 0.6$ and $\Delta W_{specified} = 0.001$. The joint density $f_H(\xi_1, \xi_2)$ is chosen equal to the diagonal band copula discussed above. The density is characterised by the parameter $b = 1 - a$. For $b = 1$, the copula represents independence. For $b = 0$, the copula represents perfect positive dependence. The density is described by

$$C_a(u_1, u_2) = \frac{1}{2(1-a)} \left(I_{[a-1,1-a]}(u - v) + I_{[0,1-a]}(u + v) + I_{[1+a,2]}(u + v) \right) \tag{27}$$

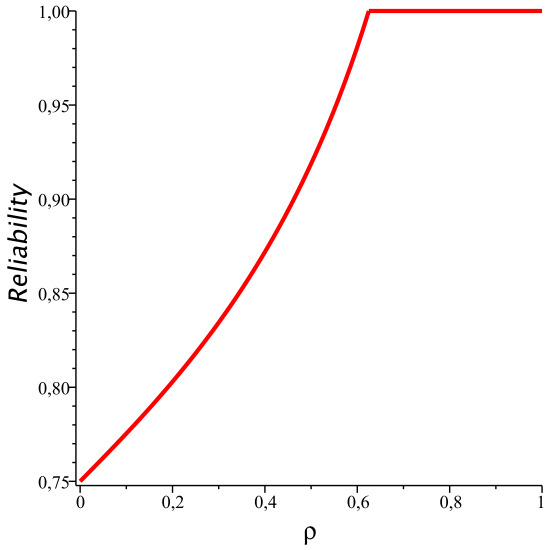


Figure 4. The reliability of the beam as a function of the correlation

With I_A, the indicator function of A. Parameter b is a function of the (rank) correlation as found in (Kurowicka and Cooke, 2006):

$$b = 1 - a = \frac{2}{3} - \frac{4}{3}\sin\left(\frac{1}{3}\arcsin\left[\frac{27}{16}\rho - \frac{11}{16}\right]\right) \tag{28}$$

Figure 4 shows the reliability for the correlation varying between 0 and 1. For a higher correlation coefficient the reliability tends to 1. If both variables H_1 and H_2 are completely independent, the reliability is only 0.75. If one is able to bound the correlation between the two variables, the corresponding bounds on the reliability can be found from the figure.

6.2. Bounds on the displacement difference of the beam with interval flexibility

The same beam as above is considered, but now H_1 and H_2 are intervals between 0.95 and 1.05. The quantity of interest is again the displacement difference between the same two points symmetrically located on both sides of the mid-point. The dependence between the two intervals is characterised by a joint set as illustrated in Figure 2, with $g(d)$ as in Eq. (22) with $d = 0.5l$, the distance between the two discretization points on the beam and $a \in \lfloor 0.5l, 2l \rceil$. This interval description for a results in a situation where one of the two discretization points is just on the boundary of the domain of influence of the other point when $a = 0.5l$ and $g(0.5l) = 0$. In the other extreme case, the other point resides on the $g(05l) = 0.75$ (with $a = 2$) limit in the domain of influence.

The upper bound for the absolute value of the displacement difference as a function of a is shown in Figure 5. The lower bound is always 0.

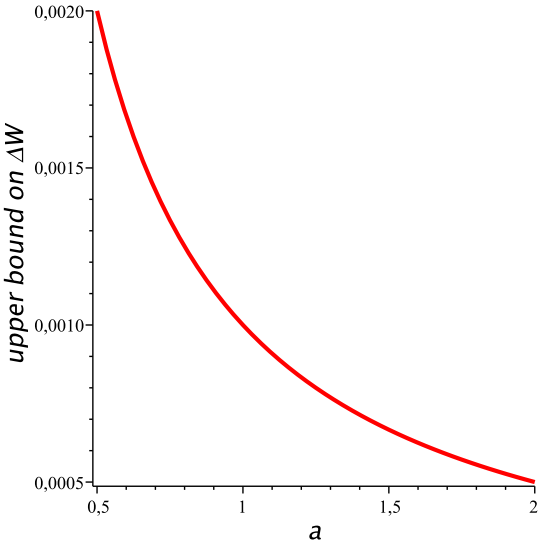


Figure 5. The upper bound on the displacement difference

7. Conclusion

The paper presents a review of dependence measures for both probabilistic and non-probabilistic descriptions of uncertainty. The link with numerically modeling a spatial uncertainty is established based on the functional relation between the dependence measure and a distance measure in a numerical model. The additional uncertainty related to this functional relationship is treated by representing the reference distance (i.e. correlation length or domain of influence) as an interval. The procedure is illustrated on a numerical example.

Acknowledgements

This study was performed during the research stay of W.V. at the Florida Atlantic University as a Fulbright scholar. W.V. would like to acknowledge the highly appreciated support of the IWT-Flanders, the Commission for Educational Exchange Between The United States of America, Belgium and Luxembourg (The Fulbright Fund and Vesuvius) and the Raymond Snoeys fund of the K.U.Leuven.

References

Ben-Haim, Y. and I. Elishakoff. *Convex Models of Uncertainty in Applied Mechanics* Elsevier Science, Amsterdam, 1990.

Charmpis, D., G.I. Schuëller and M. Pellissetti. The need for linking micromechanics of materials with stochastic finite elements: A challenge for materials science. *Computational Materials Science*, 41:27-37, 2007.

de Campos, L. and J. Huete. Independence concepts in possibility theory: Part I. *Fuzzy Sets and Systems*, 103:127-152, 1999.

De Mulder, W., D. Moens and D. Vandepitte. Modeling uncertainty in the context of finite element models with distance-based interpolation. In *Proceedings of the 1st International Symposium on Uncertainty Quantification and Stochastic Modeling, Maresias, Brazil*, 2012.

Devroye, L. *Non-Uniform Random Variate Generation*. Springer-Verlag, New York, 1986.

Drouet Mari, D. and S. Kotz. *Correlation and Dependence*. Imperial College Press, London, 2001.

Elishakoff, I. and M. Ohsaki. *Optimization and Anti-Optimization of Structures under Uncertainty*. Imperial College Press, London, 2010.

Fonck, P. A Comparative Study of Possibilistic Conditional Independence and Lack of Interaction. International Journal of Approximate Reasoning, 16:149-171, 1997.

Fuller, R. and P. Majlender. On interactive fuzzy numbers. *Fuzzy Sets and Systems*, 143:355-369, 2004.

Hisdal, E. Conditional possibilities, independence and noninteraction. *Fuzzy Sets and Systems*, 1:283-297, 1978.

Klir, G. *Uncertainty and Information, Foundations of Generalized Information Theory*. John Wiley & Sons, Hoboken, 2006.

Kurowicka, D. and R. Cooke. *Uncertainty Analysis with High Dimensional Dependence Modelling*. John Wiley & Sons, Chichester, 2006.

Moens, D. and M. Hanss. Non-probabilistic finite element analysis for parametric uncertainty treatment in applied mechanics: Recent advances. *Finite Elements in Analysis and Design*, 47:4-16, 2011.

Rodgers, J.L. and W.A. Nicewander. Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*, 42:59-66, 1988.

Schmidt, T. Coping with Copulas. *http://www.math.uni-leipzig.de/∼tschmidt/TSchmidt_Copulas.pdf*, Dec. 2006.

Schuëller, G.I. Computational Stochastic Mechanics - Recent Advances. *Computers & Structures*, 79:2225-2234, 2001.

Sudret, B. and A. Der Kiureghian. *Stochastic finite element methods and reliability: A state-of-the-art report*. Report No. UCB/SEMM-2000/08, Berkeley, 2000.

Vanmarcke, E. *Random fields: analysis and synthesis*. MIT Press, Cambridge, 1993.

Verhaeghe, W., W. Desmet, D. Vandepitte and D. Moens. Analysis of efficient interval field techniques for random field analysis with uncertain correlation length. In *Proceedings of the ASME 2011 International Mechanical Engineering Congress & Exposition, Denver, USA*, 2011.

Zhu, L.P., I. Elishakoff and J.H. Starnes, Jr. Derivation of Multi-Dimensional Ellipsoidal Convex Model for Experimental Data *Mathematical and Computer Modelling*, 24:103-114, 1996.

Interval quasi-Monte Carlo method for reliability assessment with imprecise probability

Hao Zhang
School of Civil Engineering, University of Sydney, NSW 2006, Australia.
Michael Beer
Centre for Engineering Sustainability, School of Engineering, University of Liverpool, Liverpool, UK

Abstract. Reliability analysis of structures is often performed on the basis of limited data. Under this circumstance, there are practical difficulties in identifying unique distributions as input for a probabilistic analysis. However, the selection of realistic probabilistic input is critical for the quality of the results of the reliability analysis. This problem can be addressed by using an entire set of plausible distribution functions rather than one single distribution for random variables based on limited data. An imprecise probability distribution can be modeled by a probability box, i.e., the bounds of the cumulative distribution function for the random variable. Sampling-based methods have been proposed to perform reliability analysis with probability boxes. Direct sampling of probability boxes requires a large number of samples. The computational cost can be very high as each simulation involves an interval analysis. This study proposes an interval quasi-Monte Carlo simulation methodology to efficiently compute the bounds of structure failure probabilities. The methodology is based on deterministic low-discrepancy sequences, which are distributed more regularly than the (pseudo) random points in direct Monte Carlo simulation. The efficiency and accuracy of the present method is illustrated using two examples. The reliability implications of different approaches for construction of probability boxes are also investigated through the example.

Keywords: epistemic uncertainty, imprecise probability, interval, low-discrepancy sequence, probability box, quasi-Monte Carlo, structural reliability.

1. Introduction

The building process of civil engineering structures and infrastructure is complicated by the various sources of uncertainties in structural resistance and loads, as well as in computational models. These uncertainties are treated as random variables when using established probabilistic methods for reliability analysis. Structural reliability is then measured by a probability of failure, denoted P_f (Melchers, 1999).

Although the mathematical formulation as well as the basic numerical techniques for the calculation of P_f appears to be straightforward, difficulties appear in practical applications. For one thing, the evaluation of the involved multi-dimensional integral can be challenging. Development of numerical techniques target at a high numerical efficiency and concern advanced concepts of Monte Carlo (MC) simulation such as subset simulation (Au and Beck, 2001) and line sampling (Koutsourelakis et al., 2004). Particular attention is currently paid to simulation schemes for high-dimensional problems. Moreover, there are practical difficulties in identifying the proper distribution for the random variables particularly to model their extremes

5th International Conference on Reliable Engineering Computing (REC 2012)
Edited by M. Vořechovský, V. Sadílek, S. Seitl, V. Veselý, R. L. Muhanna and R. L. Mullen

ISBN 978-80-214-4507-9. Published by Ing. Vladislav Pokorný − LITERA

("tails") which are of greatest concern in reliability assessment. Structural failure probabilities, generally very small, are sensitive to the choice of probability distributions (Ellingwood, 2001). However, available data on structural strength and loads are typically limited, and competing distributions often cannot be distinguished by standard statistical tests. When observational data are limited, the analyst may not be able to identify the type of the distribution of a random variable, or precise values for the distribution parameters, or there may be competing probabilistic models. The selection of a distribution for the probabilistic input is so generally realized based on ambiguous information and indications. This may lead to a wrong model choice and a strong overestimation of structural reliability resulting in critical decisions with severe consequences.

It is advisable to consider the distribution itself as uncertain. These statistical uncertainties are epistemic (knowledge-based) in nature (Der Kiureghian and Ditlevsen, 2009). Because of the distribution sensitivity problem, the failure probability calculated on the basis of small data samples is only a "notional" value and should not be interpreted as a relative frequency in an absolute sense (Ellingwood, 2001; Der Kiureghian and Ditlevsen, 2009). To overcome the distribution sensitivity problem, the development of first-generation probability-based structural design codes utilized a process of calibration to existing practice. The notional reliabilities associated with existing and new practices were computed in the same model universe of probability distributions and used as a means of comparison (Ellingwood et al., 1982). If the calculated reliabilities are notional and only used for the purpose of reliability comparison, then the uncertainty in the distributions is generally of minor importance and can be ignored. A reliability comparison in this sense represents a ranking of alternatives, whereby the uncertainties in the distributions associated with the alternatives are normally not significantly different in magnitude since they origin from the same source (same model universe of probability distributions). The effects of the uncertainties in the distributions, hence, cancel out one another on an ordinal (ranking) scale almost completely.

However, there are circumstances where epistemic uncertainties due to limited availability of data need to be included explicitly in reliability analysis and, further on in risk assessment. One such case is the risk-informed decision-making, in which the regulatory authorities often see a need to quantify the confidence in the results of the risk assessment, particularly if the event is rare but the consequence is severe. Another case is the performance-based approach to structural design. In this new paradigm of structural engineering, it is necessary to establish explicit reliability/risk terms to rationalize the selection of performance levels in structural design. Designers using innovative building materials and technologies are also concerned with computing realistic structural reliabilities because they cannot rely on past experience to calibrate the reliabilities. In all these circumstances a notional reliability measure without taking into account the epistemic uncertainty in the distributions is not very helpful.

Within a pure probabilistic framework, epistemic uncertainty can be handled with Bayesian approaches. Uncertain parameters of a probabilistic model can be described with prior distributions and updated by means of limited data. They can then be introduced formally, with the remaining (aleatory) uncertainties, in the reliability calculation (Der Kiureghian and Ditlevsen, 2009; Ellingwood and Kinali, 2009). In the case where competing probabilistic models exist, each model is considered separately with an assigned probability mass. A failure probability can be computed for each probabilistic model. The expectation of the failure probability can then be calculated as a characteristic result, and the frequency distribution (or variance) of the failure probability can be evaluated to separate the effects of aleatory and epistemic uncertainty. In practical applications, this requires a high numerical effort or statistical approximations.

Alternatively, an imprecisely known probability distribution can be modeled by a family of all candidate probability distributions which are compatible with available data. This is the idea of the theory of imprecise

probabilities (Walley, 1991). Dealing with a set of probability distributions is essentially different from a Bayesian approach. A practical way to represent the distribution family is to use a probability bounding approach by specifying the lower and upper bounds of the imprecise probability distribution. This corresponds to the use of an interval to represent an unknown but bounded number. Consequently, a unique failure probability cannot be determined. Instead, the failure probability is obtained as an interval whose width reflects the imprecision of the distribution model in the calculated reliability.

A popular uncertainty model using the probability bounding approach is the probability box (p-box for short) structure (Ferson et al., 2003). A p-box is closely related to other set-based uncertainty models such as random sets, Dempster-Shafer evidence theory and random intervals. In many cases, these uncertainty models can be converted into each other, and thus considered to be equivalent (Walley, 2000; Ferson et al., 2003; Joslyn and Ferson, 2004; Möller and Beer, 2008; Baudrit et al., 2008). Therefore, the methodology presented in this paper is also applicable to other set-based uncertainty models.

Within the reliability analysis context, simulation-based methods have been suggested to propagate p-boxes (Alvarez, 2006; Baraldi et al., 2008; Batarseh and Wang, 2008; Zhang et al., 2010a). Direct sampling of p-boxes requires a large number of samples to control the sampling uncertainty. The total computational cost can be very high as each simulation may involve an expensive range-finding problem. The issue of computational cost becomes more serious when the limit state function is only implicitly known through a computational model. Another concept follows the idea of global optimization to directly identify the bounds of probabilistic results (Möller and Beer, 2004). Although this concept is general and can be applied in association with variance-reduction methods (Zhang et al., 2010b), the computations are still numerically demanding. There is some urgency for further developments towards efficient methodologies for analysis with imprecise probabilities.

This study focuses on the reduction of sampling uncertainty with quasi-Monte Carlo technique. Quasi-MC method is typically used for multidimensional numerical integration problems. It performs in a similar manner as the Monte Carlo integration with the exception that samples used are a systematically constructed deterministic sequence instead of random samples. This deterministic sequence, known as low-discrepancy sequence, has the property that its points approximate the uniform distribution as closely as possible, so that a better sampling of the function can be achieved (Lemieux, 2009). As compared to MC method, Quasi-MC method can often achieve a desired tolerance of the result variance using much less simulation effort. In addition, the implementation of quasi-MC is as simple as replacing the pseudo-random numbers in MC method by low-discrepancy sequences.

Quasi-Monte Carlo has been traditionally used in computational finance and computational physics. Its applications in structural reliability analysis have been more recent (Nie and Ellingwood, 2004; Shinoda, 2007; Dai and Wang, 2009). In this paper the quasi-Monte Carlo technique is combined with the interval sampling to develop a simple, yet efficient sampling procedure for reliability analysis with p-boxes. Variance-type error estimates for the proposed method are calculated. Different approaches for constructing p-boxes are investigated and compared.

2. Construction of probability box

An imprecise distribution can be represented by the bounds on its cumulative distribution function (CDF). Let $F_X(x)$ denote the CDF for a random variable X. With an imprecise distribution, for any reference

number x it is assumed that a closed interval $[\underline{F}_{\mathrm{X}}(x),\ \overline{F}_{\mathrm{X}}(x)]$ can be found to bound the possible values of $F_{\mathrm{X}}(x)$, i.e., $\underline{F}_{\mathrm{X}}(x) \leq F_{\mathrm{X}}(x) \leq \overline{F}_{\mathrm{X}}(x)$. $\underline{F}_{\mathrm{X}}(x)$ and $\overline{F}_{\mathrm{X}}(x)$ thus form the envelopes of the probability family

$$\mathscr{P} = \{P | \forall x \in \mathbb{R}, \underline{F}_{\mathrm{X}}(x) \leq F_{\mathrm{X}}(x) \leq \overline{F}_{\mathrm{X}}(x)\}. \tag{1}$$

Such a pair of two CDF's $\underline{F}_{\mathrm{X}}(x)$ and $\overline{F}_{\mathrm{X}}(x)$ defines a "probability box". Detailed background can be found in (Ferson et al., 2003).

Various ways have been suggested to construct bounds on cumulative distribution functions. Depending on the amount of available information, the analyst may choose among the following options to construct a p-box.

2.1. KOLMOGOROV-SMIRNOV (K-S) CONFIDENCE LIMITS

K-S confidence limits on an empirical distribution have been proposed in (Ferson et al., 2003) to define bounds of the CDF as

$$\begin{aligned} \underline{F}_{\mathrm{X}}(x) &= \min(1, \max(0, F_n(x) - D_n^{\alpha})), \\ \overline{F}_{\mathrm{X}}(x) &= \min(1, \max(0, F_n(x) + D_n^{\alpha})), \end{aligned} \tag{2}$$

in which $F_n(x)$ is the empirical cumulative frequency function, and D_n^{α} is the K-S critical value at significance level α with a sample size of n.

2.2. CHEBYSHEV'S INEQUALITY

If only knowledge of the first two moments (μ and σ) for the random variable is available, Chebyshev's inequality is proposed in (Oberguggenberger and Fellin, 2008) to define a p-box as

$$\begin{aligned} \underline{F}_{\mathrm{X}}(x) &= \begin{cases} 0, & x < \mu - \sigma, \\ 1 - \frac{\sigma^2}{(x-\mu)^2}, & x \geq \mu + \sigma, \end{cases} \\ \overline{F}_{\mathrm{X}}(x) &= \begin{cases} \frac{\sigma^2}{(x-\mu)^2}, & x < \mu - \sigma, \\ 1, & x \geq \mu - \sigma. \end{cases} \end{aligned} \tag{3}$$

2.3. DISTRIBUTIONS WITH INTERVAL PARAMETERS

Probability families can also be defined by distributions with interval parameters as described in (Zhang et al., 2010a). This concept corresponds to the bunch parameter representation of fuzzy random variables described in (Möller and Beer, 2004). Confidence intervals on statistics (e.g., mean, variance) provide a natural way to define interval bounds of the distribution parameters. Based on the observational data, a confidence interval can be established using classical statistical approaches, such that the (unknown) distribution parameter is located within the interval at a specified level of confidence. Let Θ denote the (unknown) statistical parameter of the distribution, and its interval range is $[\underline{\theta},\ \overline{\theta}]$, thus the p-box can be constructed as

$$\begin{aligned} \underline{F}_{\mathrm{X}}(x) &= \min\{F_{\mathrm{X}}(x|\theta) : \underline{\theta} \leq \theta \leq \overline{\theta}\}, \\ \overline{F}_{\mathrm{X}}(x) &= \max\{F_{\mathrm{X}}(x|\theta) : \underline{\theta} \leq \theta \leq \overline{\theta}\}, \end{aligned} \tag{4}$$

in which $F_{\mathrm{X}}(x|\theta)$ is the conditional CDF, given that the parameter Θ takes the value of θ. A related concept for the quantification of fuzzy random variables can be found in (Beer, 2009).

2.4. Envelope of competing probabilistic models

When there are multiple candidate distribution models which cannot be distinguished by standard goodness-of-fit tests, the envelope of the candidate distributions may be used as the p-box (Fu et al., 2011), i.e.,

$$\begin{aligned} \underline{F}_{\mathrm{X}}(x) &= \min\{F_i(x),\ i = 1, \ldots\}, \\ \overline{F}_{\mathrm{X}}(x) &= \max\{F_i(x),\ i = 1, \ldots\}, \end{aligned} \tag{5}$$

where $F_i(x)$ denotes the ith candidate CDF.

2.5. Summary

One important observation about p-box is that although its construction still needs some subjective judgement, such as selecting an appropriate confidence level in the approach of K-S confidence limit, it generally requires less subjective information than the Bayesian approaches.

Among the above four approaches of deriving p-boxes, the K-S approach and Chebyshev's inequality are non-parametric and do not require a distribution assumption, in contrast to the third approach "distributions with interval parameters". The fourth approach "envelope of competing probabilistic models" may be parametric or non-parametric. As will be seen in the second example, the p-boxes obtained from these approaches may differ significantly from each other, leading to different results. This brings us to an important question: for a given set of small samples, how should we construct an appropriate p-box for reliability assessment? We will return to this question in our discussion of Example 2.

3. Direct interval Monte Carlo simulation

Reliability analysis with probability boxes can be performed using sampling-based methods (Alvarez, 2006; Baraldi et al., 2008; Batarseh and Wang, 2008; Zhang et al., 2010a). We start with the conventional MC simulation, in which the failure probability P_f is estimated by

$$P_f \approx \hat{P}_f = \frac{1}{n} \sum_{j=1}^{n} \mathbf{1}[g(\mathbf{x}_j) \leq 0], \tag{6}$$

where n is the total number of simulations conducted, and $\mathbf{x}_j$ represents the jth simulated vector of basic random variables in accordance to their joint probability density function $f_{\mathbf{X}}(\,)$. $g(\,)$ is the limit state function and failure occurs when $g(\mathbf{X}) \leq 0$. $\mathbf{1}[\ ]$ is the indicator function, having the value 1 if $[\ ]$ is "true" and the value 0 if $[\ ]$ is "false".

As a basis for our development we use the inverse transform method (Lemieux, 2009) for generation of random numbers:

$$\mathbf{x}_j = F_{\mathrm{X}}^{-1}(\mathbf{r}_j), \quad j = 1, 2, \ldots, n \tag{7}$$

with $F_{\mathrm{X}}(\,)$ being the target CDF of X, and $\mathbf{r}_1, \ldots \mathbf{r}_n$ a sample of independent and identically distributed (i.i.d.) standard uniform random variates.

Now consider the situation in which $F_{\mathrm{X}}(\,)$ is unknown but bounded by $\underline{F}_{\mathrm{X}}(\,)$ and $\overline{F}_{\mathrm{X}}(\,)$. Since only bounds of $F_{\mathrm{X}}(\,)$ are known, it is not possible to generate point samples but only interval samples. Let

$$\begin{aligned} \underline{\mathbf{x}}_j &= \overline{F}_X^{-1}(\mathbf{r}_j), \\ \overline{\mathbf{x}}_j &= \underline{F}_X^{-1}(\mathbf{r}_j), \quad j = 1, \ldots, n, \end{aligned} \tag{8}$$

it follows that $\underline{\mathbf{x}}_j \leq \mathbf{x}_j \leq \overline{\mathbf{x}}_j$. Note that the inverse of a CDF, $F_X^{-1}(\,)$, is monotonically decreasing. Eq. (8) suggests that for a given $\mathbf{r}_j$, the interval $[\underline{\mathbf{x}}_j, \overline{\mathbf{x}}_j]$ contains all possible simulated numbers from the probability family represented by the p-box. Define

$$\begin{aligned} \underline{g}_j(\mathbf{x}) &= \min\{g(\mathbf{X}) : \underline{\mathbf{x}}_j \leq \mathbf{X} \leq \overline{\mathbf{x}}_j\}, \\ \overline{g}_j(\mathbf{x}) &= \max\{g(\mathbf{X}) : \underline{\mathbf{x}}_j \leq \mathbf{X} \leq \overline{\mathbf{x}}_j\}, \end{aligned} \tag{9}$$

one has

$$\mathbf{1}[\overline{g}_j(\mathbf{x}) \leq 0] \leq \mathbf{1}[g(\mathbf{x}_j) \leq 0] \leq \mathbf{1}[\underline{g}_j(\mathbf{x}) \leq 0]. \tag{10}$$

Applying the inequality (10) to Eq. (6) yields

$$\frac{1}{n}\sum_{j=1}^{n} \mathbf{1}[\overline{g}_j(\mathbf{x}) \leq 0] \leq P_f \leq \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}[\underline{g}_j(\mathbf{x}) \leq 0]. \tag{11}$$

Therefore, a lower and an upper bound for P_f can be estimated as:

$$\begin{aligned} \underline{P}_f &= \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}[\overline{g}_j(\mathbf{x}) \leq 0], \\ \overline{P}_f &= \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}[\underline{g}_j(\mathbf{x}) \leq 0]. \end{aligned} \tag{12}$$

Because the randomly generated samples are intervals, this procedure is referred to as direct interval Monte Carlo simulation (Zhang et al., 2010a).

3.1. Variance of direct interval Monte Carlo

The accuracy of the reliability bounds as obtained from Eq. (12) can be measured by their variance. An estimator for the variance of $\underline{P}_f$, denoted $\widehat{\mathrm{Var}}(\underline{P}_f)$, is constructed by calculating its sample variance

$$\widehat{\mathrm{Var}}(\underline{P}_f) = \tfrac{1}{n(n-1)}\left(\textstyle\sum_{j=1}^{n} \mathbf{1}^2[\overline{g}_j \leq 0] - \underline{P}_f^2\right) = \tfrac{\underline{P}_f - \underline{P}_f^2}{n-1}. \tag{13}$$

Similarly, the variance of $\overline{P}_f$ can be estimated as

$$\widehat{\mathrm{Var}}(\overline{P}_f) = \tfrac{1}{n(n-1)}\left(\textstyle\sum_{j=1}^{n} \mathbf{1}^2[\underline{g}_j \leq 0] - \overline{P}_f^2\right) = \tfrac{\overline{P}_f - \overline{P}_f^2}{n-1}. \tag{14}$$

4. Interval quasi-Monte Carlo simulation

The starting point of our development is the difficulty to implement ideally uniform random variates in the simulation procedure as described in Section 3. The accuracy of Monte Carlo sampling relies on the "uniformity" (equidistribution) of the used uniform random numbers, i.e., the better the samples approximate the uniform distribution the more precise the Monte Carlo result is (Kalos and Whitlock, 1986). In practice, pseudo uniform random numbers which are produced by deterministic algorithms are used. Unfortunately, the pseudo uniform random numbers commonly have poor equidistribution properties, and accordingly, a good result is only achievable, if possible at all, with a large number of samples. Hence, the computational cost of interval Monte Carlo simulations can be very high as each simulation may involve an expensive range-finding problem (i.e., Eq. 9).

4.1. Low-discrepancy sequences

Consider a general multivariate integral

$$I(h) = \int_{C^s} h(\mathbf{u}) d\mathbf{u}, \tag{15}$$

where $\mathbf{u} = (u_1, \ldots, u_s)$ is an s-dimensional vector in the hypercube $C^s = [0,1)^s$, and $h : C^s \to \mathbb{R}$ is a real-valued function. Note that, in general, the integration problem characterizing the reliability assessment can always be reformulated as an equivalent one with the integration domain defined over the unit hypercube C^s (Lemieux, 2009). Given some sample points $P_n = \{\mathbf{u}_j, j = 1, \ldots, n\} \subset C^s$, an estimator Q_n for the integral $I(h)$ can be calculated as

$$I(h) \approx Q_n = \frac{1}{n} \sum_{j=1}^{n} h(\mathbf{u}_j). \tag{16}$$

With Monte Carlo method, the points $\mathbf{u}_j$ are i.i.d. samples from the uniform distribution in C^s. Alternatively, deterministic low-discrepancy sequences can be used as $\mathbf{u}_j$ to improve sampling efficiency. The theoretical background of this can be seen from the Koksma-Hlawka inequality (as discussed in Niederreiter, 1992), which gives the upper bound on the absolute error of Q_n:

$$|I(h) - Q_n| \leq V(h) D_n^{(s)} \tag{17}$$

where $V(h)$ is the bounded total variation of function h over C^s in the sense of Hardy and Krause (see Niederreiter, 1992). The term $D_n^{(s)}$ is defined as

$$D_n^{(s)} = \sup_{\mathbf{u} \in C^s} |F_n(\mathbf{u}) - F(\mathbf{u})|, \tag{18}$$

in which $\mathbf{u} = \{u_1, \ldots, u_s\} \in C^s$, $F(\mathbf{u})$ is the CDF of the uniform distribution in C^s, i.e.,

$$F(\mathbf{u}) = \prod_{i=1}^{s} u_i, \tag{19}$$

and $F_n(\mathbf{u})$ represents the empirical CDF induced by the sample points P_n, i.e.,

$$F_n(\mathbf{u}) = \frac{1}{n} \sum_{j=1}^{n} \mathbf{1}(\mathbf{u}_j \leq \mathbf{u}). \tag{20}$$

It may be observed that $D_n^{(s)}$ is actually the Kolmogorov-Smirnov statistics. In analytical number theory, $D_n^{(s)}$ is known as the *discrepancy* of the point set P_n. $D_n^{(s)}$ can be thought of as a measure of the uniformity of P_n. The inequality (17) suggests that the absolute error of the integration estimator is bounded and dominated by $D_n^{(s)}$ as $V(h)$ is a constant as far as h is given. Therefore, the quality of the estimator Q_n depends only on the discrepancy $D_n^{(s)}$, and the sample points with the smallest discrepancy (i.e., the so-called *low-discrepancy* sample) are the optimal sampling points in this sense. Since the low-discrepancy sequences are deterministic, simulation with low-discrepancy sequences is often referred to as quasi-Monte Carlo simulation.

It is possible to construct low-discrepancy sequence points for a given dimension (s) and number of samples (n) such that the deterministic error bound for Q_n is in $O(n^{-1}(\log(n))^s)$ (Fang and Wang, 1994). This is a substantial improvement of the direct Monte Carlo convergence rate of $O(n^{-1/2})$, particularly when the dimension s is not too large. The most commonly used low-discrepancy point sets and sequences include good lattice points (GLP), Halton sequence, Hammersley sequence, Sobol' sequence, Faure sequence, etc. Among them GLP is a type of lattice points, and the others are digital sequences. Dai and Wang (2009) have compared the efficiency of various digital sequences and GLP in the context of structural reliability assessment. This study uses GLP, Halton, and Faure sequences. The construction of these point sets is presented in Appendix. Details can be found, for example, in (Niederreiter, 1992) and (Lemieux, 2009).

The uniformity of a low-discrepancy sequence may be visualized by plotting its two-dimensional projection (Morokoff and Caflisch, 1994). Fig. 1 shows two-dimensional scatter plots for random sampling, GLP, Halton and Faure sequences. It is evident that the low-discrepancy points fill in C^2 more uniformly than the pure random samples.

4.2. Interval quasi-Monte Carlo sampling procedure

The proposed interval quasi-Monte Carlo method performs in the same manner as the direct interval Monte Carlo method, except that the random uniform variates are replaced by deterministic low-discrepancy sequences. With the interval quasi-Monte Carlo method, the bounds of failure probability are given by:

$$\begin{aligned} \underline{P}_f &= \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}[\overline{g}_j(\mathbf{u}_j) \le 0], \\ \overline{P}_f &= \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}[\underline{g}_j(\mathbf{u}_j) \le 0], \end{aligned} \tag{21}$$

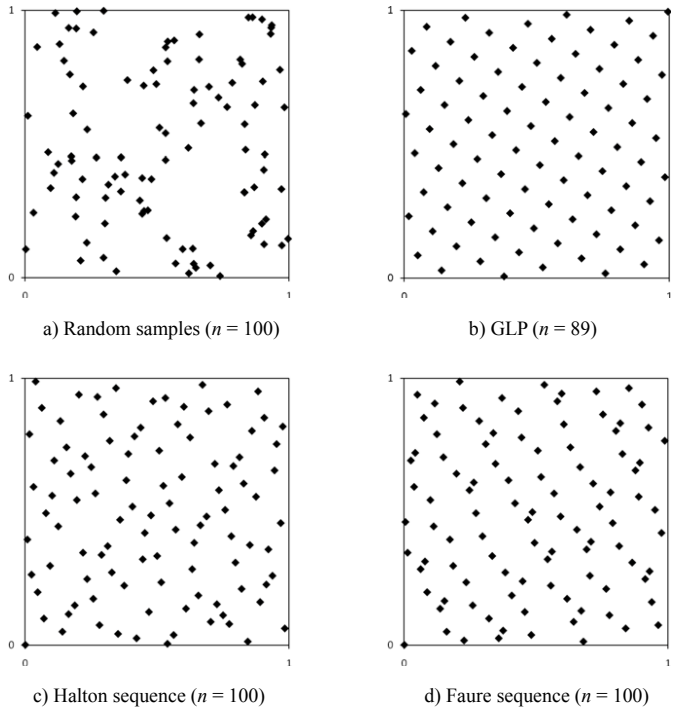
in which the sequence $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is a low-discrepancy point set, and

$$\begin{aligned} \underline{g}_j(\mathbf{u}_j) &= \min\{g(\mathbf{X}) : \overline{F}_X^{-1}(\mathbf{u}_j) \le \mathbf{X} \le \underline{F}_X^{-1}(\mathbf{u}_j)\}, \\ \overline{g}_j(\mathbf{u}_j) &= \max\{g(\mathbf{X}) : \overline{F}_X^{-1}(\mathbf{u}_j) \le \mathbf{X} \le \underline{F}_X^{-1}(\mathbf{u}_j)\}. \end{aligned} \tag{22}$$

The interval quasi-MC sampling procedure can be summarized as follows.

Step 1: Generate a low-discrepancy point set $P_n = \{\mathbf{u}_j, j = 1, \ldots, n\}$.

Step 2: Generate interval samples $[\overline{F}_X^{-1}(\mathbf{u}_j),\ \underline{F}_X^{-1}(\mathbf{u}_j)]$.



a) Random samples (n = 100) b) GLP (n = 89)

c) Halton sequence (n = 100) d) Faure sequence (n = 100)

Figure 1. 2D scatter plots of different point sets.

Step 3: For the jth simulation, compute the extrema of $g(\mathbf{X})$ for $\mathbf{X} \in [\overline{F}_X^{-1}(\mathbf{u}_j),\ \underline{F}_X^{-1}(\mathbf{u}_j)]$.

Step 4: Compute the lower and upper bounds for P_f by Eq. (21).

The proposed method requires the additional computational cost for constructing low-discrepancy sequences, as compared to the direct interval MC method. However, in a practical structural reliability analysis, the CPU time needed for the construction of low-discrepancy sequences is generally negligible in comparison with that of performing multiple simulations (Dai and Wang, 2009). Open-source libraries/routines for computer implementation of popular low-discrepancy point sets are available in programming environments, such as C++, Fortan, Matlab and others (Burkardt, 2011).

Step 3 in the proposed method involves the calculation of the range of the limit state function g when the inputs vary in certain closed intervals. The problem of finding the range of values of a function is solved on the basis of interval analysis (Moore, 1966). In many practical cases of structural reliability assessment, the limit state of interest (e.g., stress, displacement) is computed through finite element (FE) analysis, thus the limit state function is only implicitly known. In this case, an interval FE analysis is needed for the purpose of evaluating Eq. (9). An implementation of these techniques in the proposed method is straightforward. Theory and algorithms for linear elastic interval FE analysis are well developed. Reliable methods are available to compute the bounds of responses of realistic structures with reasonable accuracy when the structural stiffness and geometrical properties and loads vary in relatively narrow intervals. Reviews of interval FE analysis can be found in (Zhang, 2005) and (Moens and Vandepitte, 2005). It should be noted that, even

with the relatively efficient interval FE analysis, the computing cost is generally considerably higher than the corresponding deterministic FE analysis. It is crucial to limit the number of simulations needed.

At present, the proposed interval quasi-MC method can only handle independent p-boxes, but extensions to handle problems with dependencies are under development. Dependence information between p-boxes can be implemented via copulas as discussed in Ferson et al. (2004). Another option for consideration of dependencies in the p-box framework is the construction of multivariate models for associated random sets (Fetz and Oberguggenberger, 2010). Or one may consider the p-boxes as generated by a parametrized set of real-valued random variables and implement the dependencies between them in the traditional manner. This is associated with the concept proposed in (Möller et al., 2004) for fuzzy probabilities. In general, these concepts of imprecise probabilities provide even more flexibility in specifying dependencies compared to traditional probabilistic concepts. They can handle vaguely defined, i.e. imprecise, dependencies, such as a fuzzy correlation (Möller et al., 2006), and even cases in which only marginals are known but the copula unknown (Klir, 2006). In those cases, the indeterminacy, or imprecision, in the dependencies are translated into imprecision in CDF's, i.e. into p-boxes.

4.3. VARIANCE FOR INTERVAL QUASI-MONTE CARLO

Because standard low-discrepancy sequences are deterministic, a variance-type error estimate cannot be calculated directly for the interval quasi-MC using Eq. (13). This is one of the limitations of using low-discrepancy sequences. To address this issue, we randomize the low-discrepancy sequences to create a random sample of low-discrepancy sequences, each will be used to compute a reliability bound. Empirical variances of the obtained results can then be calculated to measure the quality of the interval quasi-MC.

A simple randomization method is to use a random shift (Cranley and Patterson, 1976). Consider an s-dimensional low-discrepancy point set $P_n = \{\mathbf{u}_j, j = 1, \ldots, n\}$. Let $\mathbf{v}$ be an s-dimensional uniform random vector. The randomized points $\tilde{\mathbf{u}}_j$ can be constructed as

$$\tilde{\mathbf{u}}_j = (\mathbf{u}_j + \mathbf{v}) \bmod 1, \quad \text{for } j = 1, \ldots, n \tag{23}$$

in which the modulo 1 operation is taken componentwise. Other more advanced randomization methods include random scrambling and permutations, see (Lemieux, 2009).

Using the random shift method, we can create m independent randomized copies of P_n, i.e., $\{\tilde{\mathbf{u}}_{j,1}, j = 1, \ldots, n\}, \ldots, \{\tilde{\mathbf{u}}_{j,m}, j = 1, \ldots, n\}$. For the lth copy of P_n, a lower and an upper bound of P_f are calculated as:

$$\begin{aligned} \underline{P}_{f,\,l} &= \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}[\overline{g}_j(\tilde{\mathbf{u}}_{j,\,l}) \le 0], \\ \overline{P}_{f,\,l} &= \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}[\underline{g}_j(\tilde{\mathbf{u}}_{j,\,l}) \le 0]. \end{aligned} \tag{24}$$

The variance of $\underline{P}_{f,\,l}$ can be estimated by its empirical variance

$$\widehat{\mathrm{Var}}(\underline{P}_{f,\,l}) = \frac{1}{m-1}\sum_{l=1}^{m}(\underline{P}_{f,\,l} - \underline{P}_f)^2, \tag{25}$$

where

$$\underline{P}_f = \frac{1}{m}\sum_{l=1}^{m} \underline{P}_{f,\,l}. \tag{26}$$

The empirical variance of the upper bound $\underline{P}_{f,\,l}$ can be calculated in a similar manner.

5. Examples

5.1. RIGID-PLASTIC PORTAL FRAME

A rigid-plastic portal frame, shown in Fig. 2, is subjected to a horizontal load H and a vertical load V. The example is adopted from (Melchers, 1999). The failure probability of the structure is mainly controlled by three limit states defined as follows:

$$\begin{aligned} g_1(\mathbf{X}) &= M_1 + 2M_3 + 2M_4 - H - V, \\ g_2(\mathbf{X}) &= M_2 + 2M_3 + M_4 - V, \\ g_3(\mathbf{X}) &= M_1 + M_2 + M_4 - H, \end{aligned}$$

in which M_1,...,M_4 are the plastic bending capacities at the joints. Since this is a series system, the system limit state function g is the minimum of the above, i.e.,$g = \min(g_1, g_2, g_3)$. The basic random variables are $\mathbf{X} = (M_1, \ldots, M_4, H, V)$. All the variables are normally distributed and mutually independent. Assume that because of limited knowledge, the mean values of X_i can only be estimated as intervals. Table I summarizes the statistics for the basic random variables.

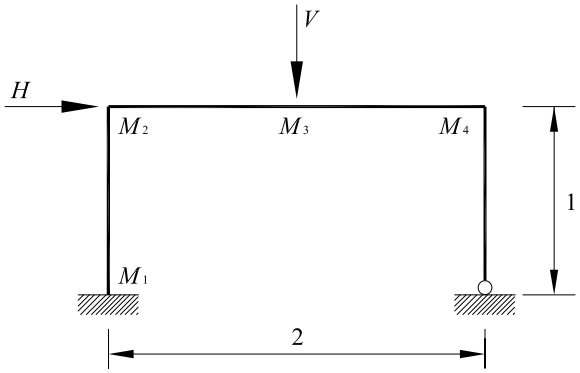


Figure 2. A rigid-plastic portal frame.

The proposed quasi-interval Monte Carlo method is applied using the GLP, Halton and Faure points with different numbers of samples. The random shift method (Eq. 23) was used to randomize the sample points. The first 10,000 points in the Halton and Faure sequences were omitted, as suggested in (Shinoda, 2007). The Halton points were also leaped with $L = 61$. As an independent check, direct interval Monte Carlo simulation was conducted. Table II compares the computed reliability bounds from the two methods with

Table I. Random variables for the rigid frame.

Variable	Mean	Standard dev.
M_i	[0.98,1.02]	0.15
H	[1.87,1.93]	0.45
V	[1.18,1.22]	0.3

different numbers of simulations. With the direct interval Monte Carlo, the failure probability is found to be [0.0118, 0.0287] after 10^5 simulations. It can be seen from Table II that the interval quasi-Monte Carlo methods with GLP, Halton and Faure sequences all give reasonable results after 3,000 simulations. The results were somewhat improved when the number of simulations was increased to 10,000.

To achieve a variance-type error estimate, each method was then repeated 30 times to obtain a sample of 30 results. On the basis of the 30 calculations, the sample variance for the estimated reliability bounds by each method were calculated and compared in the last two columns of Table II. For the direct interval Monte Carlo results, the estimates of variance calculated using Eqs. (13) and (14) are also presented to compare with the empirical results. It is evident that with the same number of simulations, the (randomized) interval quasi-Monte Carlo results have less variance than the direct interval Monte Carlo results. For instance, with 3000 simulations, the sample variance for $\underline{P}_f$ is 0.00224 for direct interval Monte Carlo; it reduces to around 0.001 for the present method. From Table II, it can also be seen that for the direct interval Monte Carlo, variance estimates obtained from Eqs. (13) and (14) agree well with the empirical results.

Table II. Reliability bounds and their empirical variances (Example 1: rigid frame).

Method-No Simul.	$\underline{P}_f \times 10^2$	$\overline{P}_f \times 10^2$	$\widehat{\mathrm{Var}}_{\underline{P}_f} \times 10^2$	$\widehat{\mathrm{Var}}_{\overline{P}_f} \times 10^2$
GLP-3001	1.13	2.87	0.105	0.140
GLP-5003	1.12	2.96	0.895	0.124
GLP-10007	1.19	2.86	0.051	0.045
Halton-3000	1.20	2.70	0.109	0.169
Halton-5000	1.18	3.06	0.083	0.132
Halton-10^4	1.19	2.94	0.043	0.075
Faure-3000	1.27	2.57	0.118	0.175
Faure-5000	1.28	2.78	0.104	0.161
Faure-10^4	1.21	2.83	0.076	0.123
interval MC-3000	1.33	3.03	0.224 (0.195) *	0.302 (0.302)
interval MC-10^4	1.11	2.84	0.112 (0.109)	0.182 (0.167)
interval MC-10^5	1.18	2.87	0.038 (0.034)	0.061 (0.053)

* values in parentheses are based on Eqs. (13) and (14).

5.2. TEN-BAR TRUSS UNDER WIND LOAD

The second example is a planar 10-bar truss subject to wind load, as shown in Fig. 3. The structure layout and member sizes are adopted from (Nie and Ellingwood, 2005). The response of the truss is assumed to be linear elastic. The limit state in question is the roof drift with a limit value of 17.78 mm. We are interested in finding the failure probability under the (annual maximum) wind load. The member cross-sectional areas and the wind load are taken to be random variables. Let A_i denote the cross-sectional area for the ith member. The basic random variables are $\mathbf{X} = (A_1, \ldots, A_{10}, W)$. Table III summarizes the statistics for the cross-sectional areas. All cross-sectional areas are assumed to be mutually statistically independent.

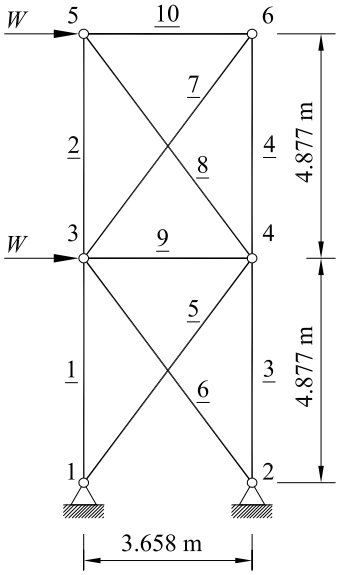


Figure 3. Ten-bar plane truss subject to wind load.

Table III. Statistics for the cross-sectional areas for the 10-bar truss (unit: cm^2).

Variable	Mean	COV	Distribution
$A_1 - A_4$	53.226	0.1	Normal
$A_5 - A_8$	45.677	0.1	Normal
A_9, A_{10}	9.548	0.1	Normal

Suppose that a 30-year record of the annual maximum wind speed at the location of the structure is available, which represents a sample of size 30 for the load, see Table IV. The sample mean and sample standard deviation for the wind load W are 112.99 kN and 35.51 kN, respectively.

Based on this rather restricted set of wind load data, the probability boxes for the wind load were constructed using the four approaches introduced in Section 2. Fig. 4 shows the p-boxes derived by the

Table IV. Samples of annual maximum wind load, W (kN).

Year	W	Year	W	Year	W
1	81.33	11	128.79	21	149.51
2	121.94	12	97.37	22	80.05
3	175.01	13	72.13	23	134.91
4	91.50	14	88.03	24	74.02
5	101.02	15	110.05	25	150.12
6	110.66	16	130.78	26	100.81
7	106.58	17	114.06	27	86.55
8	83.71	18	69.09	28	198.92
9	94.88	19	123.18	29	80.61
10	151.61	20	198.16	30	84.17

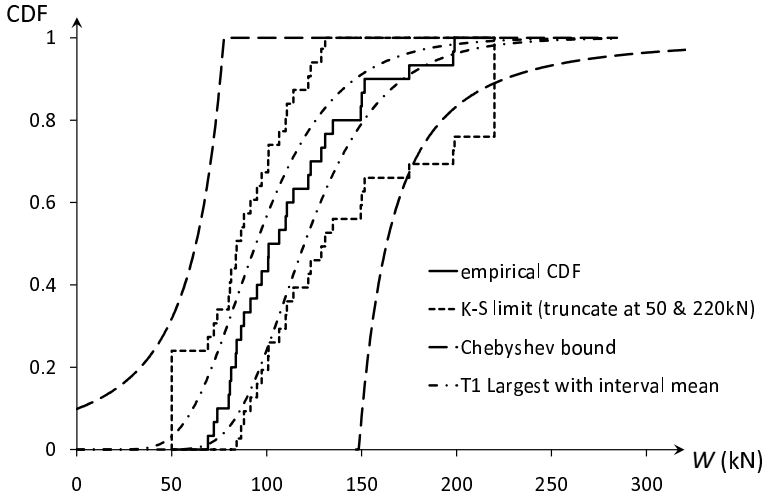


Figure 4. CDF bounds for the wind load W constructed using different methods.

Kolmogorov-Smirnov confidence limits (at 5% significance level) and Chebyshev's inequality. Since the K-S confidence limits only give the p-box for the range of the samples, it is necessary to assume a smallest and a largest possible value to truncate the K-S limits. In this example, the wind load samples have a range of 69.09 to 198.92 kN. The K-S limits are truncated at 50 and 220 kN. These two limit values were chosen based on the authors' engineering judgement.

Next, the confidence interval on the mean value was used to define the p-box. Assuming the population standard deviation σ_W is equal to the sample standard deviation, the 95% confidence interval on the population mean is approximately [100.28, 125.69] kN. Assume that based on experience, the (annual maximum)

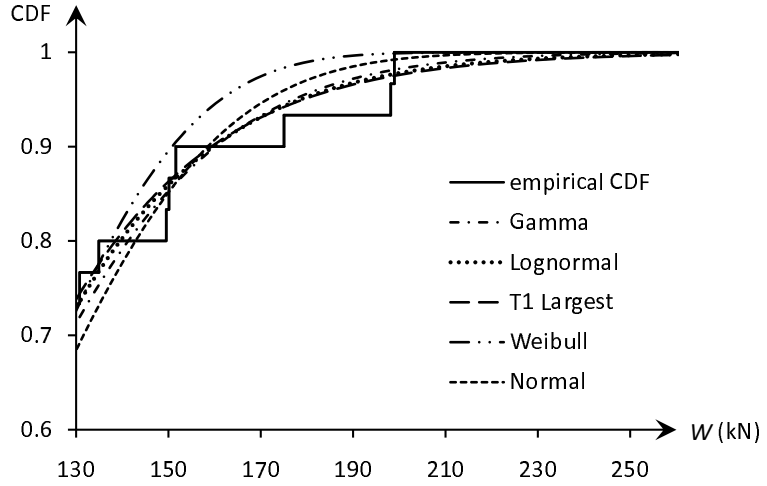


Figure 5. Candidate distributions for the wind load (upper tails).

wind load can be properly modeled by a Type 1 Largest (T1 Largest) extreme value distribution. The p-box for the T1 Largest with an interval mean [100.28, 125.69] kN and a standard deviation of 35.51 kN is illustrated in Fig. 4.

The wind load p-box can also be constructed using the envelope of a set of candidate distributions. Five candidate distributions were chosen to fit the wind load data, i.e., the T1 Largest, lognormal, Gamma, Normal and Weibull distributions. The Kolmogorov-Smirnov goodness-of-fit tests were performed to test the validity of these assumed distributions. It can be shown that all candidate distributions passed the K-S goodness-of-fit tests at the significance level of 5%. As the K-S statistics for the different distributions are relatively similar, their CDFs would hardly be distinguished if plotted in one figure. However, the difference between the five candidate distributions in their upper tails can be clearly seen from Fig. 5. As the structural reliability is dominated by the upper tail of the distribution of the wind load, it is expected that the failure probabilities obtained from different wind models will differ significantly.

The proposed interval quasi-Monte Carlo method is applied with the four different p-box models for the wind load. The results by GLP, Halton and Faure sequences are very similar. Only the latter is presented herein. The first 10,000 points in the Faure sequence were omitted. The next 5,000 points were then used for the simulation. The results were verified by direct interval Monte Carlo simulation with 10^5 samples. The two sets of the results are compared in Table V. It may be observed from Table V that the results from the present method agree well with those from direct interval Monte Carlo.

It is more interesting to note that the bounds of P_f for different wind load p-box models vary considerably. Using Chebyshev's inequality yields the widest P_f, i.e., $P_f = [0,\ 0.37]$. It is unlikely that the analyst will find such a wide P_f practically useful. This is because the wind load p-box based on Chebyshev's inequality is very wide, particularly in the upper tail as can be seen in Fig. 4. It is due to the fact that only the information of the first two moments about the random variable is used. Note that the p-box derived by Chebyshev' inequality (Eq. 3) is independent of the sample size given that the first two moments are known. Two sets of data, one with limited samples and a second with comprehensive samples, would lead to the

same p-box if the first two moments from the two sets of data were the same. This lack of distinction is undesirable. Moreover, with Chebyshev's inequality, the p-box will not necessarily become narrower when additional data are available. This does not agree with the general conception that epistemic uncertainty can be reduced when the quality of data is refined. Because of these limitations, it appears that Chebyshev's inequality is not suitable for the construction of p-box in structural reliability assessment.

With the wind load p-box constructed using the K-S confidence limits, P_f is found to vary in a wide range between 0 and 0.246. Unlike Chebyshev's inequality, the K-S confidence limits will become narrower as the number of samples increase. However, the K-S approach has a main drawback, i.e., the derived p-boxes have to be truncated at the tails, and the result may be influenced strongly by the values of the truncation, which are often chosen arbitrarily. In this example the K-S limits are truncated at 220 kN at the upper end, and $\overline{P}_f$ is found to be 0.246. If the tail is truncated at 198.92 kN, i.e. the maximum of the sampled wind loads, $\overline{P}_f$ will reduce from 0.246 to 0.0288. Similar observation is made in (Ferson et al., 2003). Because of the sensitivity of results to the truncation, it is not practical to use the Kolmogorov-Smirnov confidence limit to construct p-boxes for the purpose of reliability analysis.

Table V. Reliability bounds of the 10-bar truss.

P-box construction	quasi-interval MC (Faure, $n = 5000$)	interval MC ($n = 10^5$)
K-S limits	[0, 0.246]	[0, 0.246]
Chebyshev	[0, 0.368]	[0, 0.359]
P-box 3	$[1.16, 2.66] \times 10^{-2}$	$[1.1, 2.67] \times 10^{-2}$
P-box 4	$[0.06, 1.62] \times 10^{-2}$	$[0.052, 1.73] \times 10^{-2}$

P-box 3: T1 Largest with interval mean.

P-box 4: envelope of 5 candidate distributions.

If the effect of small sample size is considered through the confidence interval of the first-order statistics, the wind load is modeled by a Type 1 Largest with an interval mean of [100.28, 125.69] kN (95% confidence interval). Under this assumption, P_f varies between 0.0116 and 0.0266. This interval bound clearly demonstrates the effect of small sample size on the calculated structural reliability. It indicates that confidence intervals on distribution parameters is a reasonable way to define p-boxes, provided that the appropriate distribution form can be discerned.

In the case where the wind load p-box is constructed using the envelope of five candidate distributions, the variation of P_f is from 6×10^{-4} to 0.0162. The very small lower bound of P_f is contributed by the Weibull distribution. As evidenced from Fig. 5, the Weibull distribution is the least conservative one in the upper tail region. If the Weibull is discarded, the bounds of P_f will become [0.32, 1.62] $\times 10^{-2}$. These results highlight the sensitivity of the failure probability to the choice of the probabilistic model for the wind load. Although the candidate distributions all passed the K-S goodness-of-fit test and their K-S statistics are relatively similar, their tails can be different, leading to very different reliability results.

6. Conclusions

A new interval quasi-Monte Carlo method has been developed to perform structural reliability analysis with probability boxes. The methodology is based on deterministic low-discrepancy sequences, which are more uniformly distributed than pure random samples. Compared with the direct interval Monte Carlo method, the present method can effectively reduce the required number of simulations to achieve comparable accuracy. The proposed interval quasi-Monte Carlo method can also be randomized so that variance-type error estimates can be obtained. The implementation of the present method is relatively straightforward; it replaces the pseudo random numbers in interval Monte Carlo by the low-discrepancy sequences. It is a general technique that can be applied to a variety of uncertainty models, including probability boxes, random sets, random intervals, Dempster-Shafer evidence theory, fuzzy random variables, etc.

A frame and a truss have been analyzed to demonstrate the efficiency and accuracy of the proposed method. For both problems, the present method yields reasonably good results with at least one order-of-magnitude less computational effort than the direct interval Monte Carlo simulation.

For the truss example, the wind load p-box was constructed using four different approaches suggested in the literature, i.e., the Kolmogorov-Smirnov confidence limits, Chebyshev's inequality, confidence interval on the mean value, and the envelope of five candidate distributions which all pass K-S goodness-of-fit tests. The reliability results suggest that the K-S approach or Chebyshev's inequality do not lend themselves to the practical construction of probability boxes in the context of structural reliability analysis. The p-boxes based on these two approaches seem to be too conservative. It is unlikely that the analysts/decision-makers will find such conservative P_f practically useful. There are also practical difficulties in using the K-S confidence limits and Chebyshev's inequality to define p-boxes. The most reasonable method to construct p-boxes for the purpose of reliability assessment seems to be their construction based on confidence intervals of statistics.

Appendix. Low-discrepancy sequences

This section briefly introduces the constructions of the GLP, Halton, and Faure sequences. See (Niederreiter, 1992) and (Lemieux, 2009) for details.

GOOD LATTICE POINT

Let n be an integer ≥ 2, $\mathbf{a} = (a_1, \ldots, a_s)$ be an integer vector, and gcd(a_i, n)=1 for all i, where gcd = greatest common divisor. A point set $P_n = \{\mathbf{u}_j, j = 1, \ldots, n\}$ can be constructed by

$$\mathbf{u}_j = \mathrm{frac}(\frac{j-1}{n}\mathbf{a}) \quad \text{for } j = 1, \ldots, n \tag{27}$$

where $\mathrm{frac}(z)$ denotes the fractional part of z. Such a point set is known as a lattice point set with the generating vector **a**. Among all possible lattice point sets, the one that has the smallest discrepancy is called the GLP point set. It has been suggested that for a prime number n, the generating vector can be taken as (Korobov, 1959):

$$\mathbf{a} = (1, a, \ldots, a^{s-1}) \bmod n, \tag{28}$$

where a is an integer, $1 \leq a \leq n-1$, and $\gcd(a,n) = 1$. It was shown that an optimal value of a exists for given values of s and n such that the resulting good lattice point set has a discrepancy $D_n^{(s)} = O(n^{-1}(\log(n))^s)$. The method for determining the optimal generating vectors can be found in (Fang and Wang, 1994).

HALTON SEQUENCE

Let $b \geq 2$ be an integer, then any integer $n \geq 0$ has a unique digit expansion in base b as

$$n = \sum_{j=0}^{r} a_j(n) b^j = a_0(n) + a_1(n)b + \ldots + a_r(n)b^r, \tag{29}$$

where $a_j(n) \in \{0, 1, \ldots, b-1\}$ for all i, and $b^r \leq n < b^{r+1}$. Thus, r can be calculated as

$$r = \lfloor \log_b n \rfloor, \tag{30}$$

in which $\lfloor x \rfloor$ denotes the integral part of x. Define the radical-inverse function $\phi_b(n)$ in base b as

$$\begin{aligned} \phi_b(n) &= \sum_{j=0}^{r} a_j(n) b^{-j-1} \\ &= a_0(n)b^{-1} + a_1(n)b^{-2} + \ldots + a_r(n)b^{-r-1}. \end{aligned} \tag{31}$$

Note that for all integers $n \geq 0$, $\phi_b(n) \in [0, 1)$. With this definition of radical-inverse function, the nth vector of the s-dimensional Halton sequence is defined as

$$\mathbf{u}_n = (\phi_{b_1}(n), \phi_{b_2}(n), \ldots, \phi_{b_s}(n)), \; n = 1, 2, \ldots \tag{32}$$

with the bases $b_1, \ldots, b_s > 1$ being pairwise relatively prime. In practice, the first s primes are usually used as the bases. It has been shown that the Halton sequence formed by the first n points ($n > \max(b_1, \ldots, b_s)$) has the discrepancy $O(n^{-1}(\log(n))^s)$ (Halton, 1960).

FAURE SEQUENCE

The construction of Faure sequence appears to be like the Halton sequence, but it uses only one base for all dimensions. Given a dimension s, let b be a prime number such that $b \geq s$ and $b \geq 2$. The first dimension of the Faure sequence corresponding to n is given by

$$\phi_b^1(n) = \sum_{j=0}^{r} a_j^1(n) b^{-j-1}, \quad n = 1, 2, \ldots \tag{33}$$

in which the superscript 1 denotes the first dimension, and $a_j^1(n)$ equals to the expansion coefficient defined in Eq. (29), i.e.,

$$a_j^1(n) = a_j(n). \tag{34}$$

The elements of the higher dimensions of the Faure sequence are then constructed as

$$\phi_b^k(n) = \sum_{j=0}^{r} a_j^k(n) b^{-j-1}, \quad 2 \le k \le s,\ n = 1, 2, \ldots \tag{35}$$

and the coefficients $a_j^k(n)$ are calculated recursively from those of the lower dimensions:

$$a_j^k(n) = \sum_{i \ge j}^{r} C(i,j) a_i^{k-1}(n) \bmod b \tag{36}$$

in which $C(i,j) = i!/j!(i-j)!$ is the combinatorial function. Thus the nth vector of the sample points in the sequence is

$$\mathbf{u}_n = (\phi_b^1(n), \phi_b^2(n), \ldots, \phi_b^s(n)),\ n = 1, 2, \ldots \tag{37}$$

With digital sequences such as Halton and Faure sequences, it has been suggested to omit some initial sequence points as they often exhibit undesirable correlations among different dimensions. Another technique is to use "leaped" sequence points, i.e., to use only every Lth points in the sequence, where L is the leap (Kocis and Whiten, 1997).

References

Alvarez, D. A., 2006. On the calculation of the bounds of probability of events using infinite random sets. Int. J. Approximate Reasoning 43, 241–267.

Au, S. K., Beck, J. L., 2001. Estimation of small failure probabilities in high dimensions by subset simulation. Probab. Engrg. Mech. 16, 263–277.

Baraldi, P., Popescu, I. C., Zio, E., 2008. Predicting the time to failure of a randomly degrading component by a hybrid monte carlo and possibilistic method. In: IEEE Proc. International Conference on Prognostics and Health Management, 2008.

Batarseh, O. G., Wang, Y., 2008. Reliable simulation with input uncertainties using an interval-based approach. In: IEEE Proc. Winter Simulation Conference, 2008.

Baudrit, C., Dubois, D., Perrot, N., 2008. Representing parametric probabilistic models tainted with imprecision. Fuzzy Sets Syst. 159, 1913–1928.

Beer, M., 2009. Engineering quantification of inconsistent information. Int. J. Reliab. Saf. 3 (1–3), 174–200.

Burkardt, J., 2011. Quasirandom sequence generator: source codes. http://people.sc.fsu.edu/~jburkardt/cpp_src/cpp_src.html.

Cranley, R., Patterson, T. N. L., 1976. Randomization of number theoretic methods for multiple integration. SIAM J. Numer. Anal. 13 (6), 904–914.

Dai, H., Wang, W., 2009. Application of low-discrepancy sampling method in structural reliability analysis. Struct. Saf. 31, 55–64.

Der Kiureghian, A., Ditlevsen, O., 2009. Aleatory or epistemic? Does it matter? Struct. Saf. 31, 105–112.

Ellingwood, B. R., 2001. Acceptable risk bases for design of structures. Prog. Struct. Engng. Mater. 3, 170–179.

Ellingwood, B. R., Kinali, K., 2009. Quantifying and communicating uncertainty in seismic risk assessment. Struct. Saf. 31, 179–187.

Ellingwood, B. R., MacGregor, J. G., Galambos, T. V., Cornell, C. A., 1982. Probability based load criteria: load factors and load combinations. J. Struct. Div., ASCE 108 (5), 978–997.

Fang, K. T., Wang, Y., 1994. Number-theoretic methods in statistics. Chapman & Hall, London.

Ferson, S., Joslyn, C., Helton, J. C., Oberkampf, W. L., Sentz, K., 2004. Summary from the epistemic uncertainty workshop: consensus amid diversity. Reliability Engineering & System Safety 85, 355–369.

Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D. S., Sentz, K., 2003. Constructing probability boxes and Dempster-Shafer structures. Tech. Rep. SAND2002-4015, Sandia National Laboratories.

Fetz, T., Oberguggenberger, M., 2010. Multivariate models of uncertainty: A local random set approach. Struct. Saf. 32 (6), 417–424.

Fu, G., Butler, D., Khu, S.-T., Sun, S., 2011. Imprecise probabilistic evaluation of sewer flodding in urban drainage systems using random set theory. Water Resour. Res. 47, W02534.

Joslyn, C., Ferson, S., 2004. Approximate representations of random intervals for hybrid uncertainty quantification in engineering modeling. In: Hanson, K., Hemez, F. (Eds.), Proc. 4th International Conference on Sensitivity Analysis of Model Output (SAMO 2004). Santa Fe, New Mexico, U.S.A., pp. 453–469.

Kalos, M. H., Whitlock, P. A., 1986. Monte Carlo methods. Vol. 1: basics. John Wiley, New York.

Klir, G. J., 2006. Uncertainty and information: foundations of generalized information theory. Wiley-Interscience.

Koutsourelakis, P. S., Pradlwarter, H. J., Schuëller, G. I., 2004. Reliability of structures in high dimensions, part I: algorithms and applications. Probab. Engrg. Mech. 19, 409–417.

Lemieux, C., 2009. Monte Carlo and Quasi-Monte Carlo sampling. Springer, New York.

Melchers, R. E., 1999. Structural reliability analysis and prediction, 2nd Edition. John Wiley & Sons, West Sussex, England.

Moens, D., Vandepitte, D., 2005. A survey of non-probabilistic uncertainty treatment in finite element analysis. Comput. Methods Appl. Mech. Engrg. 194, 1527–1555.

Möller, B., Beer, M., 2004. Fuzzy-randomness-uncertainty in civil engineering and computational mechanics. Springer-Verlag.

Möller, B., Beer, M., 2008. Engineering computation under uncertainty-capabilities of non-traditional models. Comput. Struct. 86 (10), 1024–1041.

Möller, B., Beer, M., Graf, W., Sickert, J.-U., 2006. Time-dependent reliability of textile-strengthened rc structures under consideration of fuzzy randomness. Comp. Struc. 84, 585–603.

Möller, B., Graf, W., Nguyen, S. H., 2004. Modeling the life cycle of a structure using fuzzy processes. Compu.-Aided Civ. Infrastruct. Engrg. 19, 157–169.

Moore, R. E., 1966. Interval Analysis. Prentice-Hall, Inc., Englewood Cliffs, N. J.

Morokoff, W. J., Caflisch, R. E., 1994. Quasi-random sequences and their discrepancies. SIAM J. Sci. Comput. 15, 1251–1279.

Nie, J., Ellingwood, B. R., 2004. A new directional simulation method for system reliability. Part I: application of deterministic point sets. Probab. Engrg. Mech. 19 (4), 425–436.

Nie, J., Ellingwood, B. R., 2005. Finite element-based structural reliability assessment using efficient directional simulation. J. Engrg. Mech., ASCE 131 (3), 259–267.

Niederreiter, H., 1992. Random number generation and quasi-Monte Carlo methods. SIAM, Philadelphia.

Oberguggenberger, M., Fellin, W., 2008. Reliability bounds through random sets: non-parametric methods and geotechnical applications. Comput. Struct. 86 (10), 1093–1101.

Shinoda, M., 2007. Quasi-Monte Carlo simulation with low-discrepancy sequence for reinforced soil slopes. J. Struct. Engrg., ASCE 133 (4), 393–404.

Walley, P., 1991. Statistical reasoning with imprecise probabilities. Chapman and Hall, London.

Walley, P., 2000. Towards a unified theory of imprecise probability. Int. J. Approximate Reasoning 24, 125–148.

Zhang, H., 2005. Nondeterministic linear static finite element analysis: an interval approach. Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA, USA.

Zhang, H., Mullen, R. L., Muhanna, R. L., 2010a. Interval Monte Carlo methods for structural reliability. Struct. Saf. 32, 183–190.

Zhang, M., Beer, M., Quek, S. T., Choo, Y. S., 2010b. Comparison of uncertainty models in reliability analysis of offshore structures under marine corrosion. Struct. Saf. 32, 425–432.

Halton, J., 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer. Math. 2, 84–90.

Kocis, L., Whiten, W. J., 1997. Computational investigations of low-discrepancy sequences. ACM Trans. Math. Software 23 (2), 266–294.

Korobov, N. M., 1959. The approximate computation of multiple integrals. Doklady Akademii Nauk SSSR (In Russian) 124, 1207–1210.

Author Index