W. Fellin · H. Lessmann · M. Oberguggenberger · R. Vieider (Eds.)

Analyzing Uncertainty in Civil Engineering

Wolfgang Fellin · Heimo Lessmann
Michael Oberguggenberger · Robert Vieider (Eds.)

Analyzing Uncertainty in Civil Engineering

With 157 Figures and 23 Tables

Springer

Editors

a.o. Univ.-Prof. Dipl.-Ing. Dr. Wolfgang Fellin
Institut für Geotechnik und Tunnelbau
Universität Innsbruck
Technikerstr. 13
6020 Innsbruck
Austria

em. Univ.-Prof. Dipl.-Ing. Heimo Lessmann
Starkenbühel 304
6073 Sistrans
Austria

a.o. Univ.-Prof. Dr. Michael Oberguggenberger
Institut für Technische Mathematik,
Geometrie und Bauinformatik
Universität Innsbruck
Technikerstr. 13
6020 Innsbruck
Austria

Dipl.-Ing. Robert Vieider
Vieider Ingenieur GmbH
Rebschulweg 1/E
39052 Kaltern an der Weinstraße
Italy

ISBN 3-540-22246-4 **Springer Berlin Heidelberg New York**

Library of Congress Control Number: 2004112073

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in other ways, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable to prosecution under German Copyright Law.

Springer is a part of Springer Science+Business Media

springeronline.com

© Springer-Verlag Berlin Heidelberg 2005
Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Data conversion by the authors.
Final processing by PTP-Berlin Protago-TeX-Production GmbH, Germany
Cover-Design: medionet AG, Berlin
Printed on acid-free paper 62/3020Yu - 5 4 3 2 1 0

Preface

This volume addresses the issue of uncertainty in civil engineering from design to construction. Failures do occur in practice. Attributing them to a residual risk or a faulty execution of the project does not properly cover the range of causes. A closer scrutiny of the design, the engineering model, the data, the soil-structure-interaction and the model assumptions is required. Usually, the uncertainties in initial and boundary conditions as well as material parameters are abundant. Current engineering practice often leaves these issues aside, despite the fact that new scientific tools have been developed in the past decades that allow a rational description of uncertainties of all kinds, from model uncertainty to data uncertainty.

It is the aim of this volume to have a critical look at current engineering risk concepts in order to raise awareness of uncertainty in numerical computations, shortcomings of a strictly probabilistic safety concept, geotechnical models of failure mechanisms and their implications for construction management, execution, and the juristic question as to who has to take responsibility. In addition, a number of the new procedures for modelling uncertainty are explained.

Our central claim is that doubts and uncertainties must be openly addressed in the design process. This contrasts certain tendencies in the engineering community that, though incorporating uncertainties by one or the other way in the modelling process, claim to being able to control them.

In our view, it is beyond question that a mathematical/numerical formalization is needed to provide a proper understanding of the effects of the inherent uncertainties of a project. Available information from experience, in situ measurements, laboratory tests, previous projects and expert assessments should be taken into account. Combining this with the engineering model(s) - and a critical questioning of the underlying assumptions -, insight is generated into the possible behavior, pitfalls and risks that might be encountered at the construction site. In this way workable and comprehensible solutions are reached that can be communicated and provide the relevant information for all participants in a complex project.

This approach is the opposite of an algorithm that would provide single numbers pretending to characterize the risks of a project in an absolute way (like safety margins or failure probabilities). Such magic numbers do not exist. Instead of seducing the designing engineer into believing that risks are under

control, we emphasize that understanding the behavior of the engineering system is the central task and the key to responsible decisions in view of risks and imponderables.

The book is the result of a collaborate effort of mathematicians, engineers and construction managers who met regularly in a post graduate seminar at the University of Innsbruck during the past years. It contains contributions that shed light on the central theme outlined above from various perspectives and thus subsumes the state of discussion arrived at by the participants over those years. Except for three reprints of foundational papers, all contributions are new and have been written for the purpose of this collection.

The book starts with three papers on geotechnics. The first two articles by Fellin address the problem of assessment of soil parameters and the ambiguity of safety definition in geotechnics. The third paper by Oberguggenberger and Fellin demonstrates the high sensitivity of the failure probability on the choice of input distribution. This sets the stage for the theoretically oriented paper by Oberguggenberger providing a survey of available models of uncertainty and how they can be implemented in numerical computations. The mathematical foundations are complemented by the following paper of Fetz describing how the joint uncertainty in multi-parameter models can be incorporated. Next, Ostermann addresses the issue of sensitivity analysis and how it is performed numerically. This is followed by a reprint of a paper by Herle discussing the result of benchmark studies. Predictions of deformations obtained by different geotechnicians and numerical methods in the same problem are seen to deviate dramatically from each other. Lehar et al. present an ultimate load analysis of pile-supported buried pipelines, showing the extensive interplay between modelling, laboratory testing and numerical analysis which is necessary to arrive at a conclusive description of the performance of the pipes. The paper by Lessmann and Vieider turns to the implications of the geotechnical model uncertainty to construction management. It discusses the type of information the construction manager would need as well as the question of responsibility in face of large model uncertainties. The following paper by Oberguggenberger and Russo compares various uncertainty models (probability, fuzzy sets, stochastic processes) at the hand of the simple example of an elastically bedded beam, while the article by Oberguggenberger on queueing models ventures into a similar comparison of methods in a theme relevant for project planning. The book is completed by a reprint of a survey article showing how fuzzy sets can be used to describe uncertainty throughout civil engineering.

Innsbruck,
May 2004

Wolfgang Fellin
Heimo Lessmann
Michael Oberguggenberger
Robert Vieider

Contents

Assessment of characteristic shear strength parameters of soil and its implication in geotechnical design
Wolfgang Fellin 1
1 Characteristic values of soil parameters 1
2 Example 6
3 Influence on design 9
4 Conclusions 13
References 14

Ambiguity of safety definition in geotechnical models
Wolfgang Fellin 17
1 Slope stability of a vertical slope 17
2 Various safety definitions 19
3 Different geotechnical models 27
4 Sensitivity analysis 28
5 Conclusion 30
References 31

The fuzziness and sensitivity of failure probabilities
Michael Oberguggenberger, Wolfgang Fellin 33
1 Introduction 33
2 Probabilistic modeling 34
3 Sensitivity of failure probabilities: two examples 36
4 Robust alternatives 44
5 Conclusion 48
References 48

The mathematics of uncertainty: models, methods and interpretations
Michael Oberguggenberger 51
1 Introduction 51
2 Definitions 53

3 Semantics . . . 57
4 Axiomatics . . . 63
5 Numerics . . . 64
6 The multivariate case . . . 66
References . . . 67

Multi-parameter models: rules and computational methods for combining uncertainties
Thomas Fetz . . . 73
1 Introduction . . . 73
2 Random sets and sets of probability measures . . . 74
3 Numerical example . . . 77
4 Types of independence . . . 82
5 Sets of joint probability measures generated by random sets . . . 85
6 The different cases . . . 87
7 Numerical results for Examples 1, 2, 3, 4 and Conclusion . . . 96
References . . . 98

Sensitivity analysis
Alexander Ostermann . . . 101
1 Introduction . . . 101
2 Mathematical background . . . 102
3 Analytic vs. numerical differentiation . . . 104
4 Examples . . . 106
5 Conclusions . . . 114
References . . . 114

Difficulties related to numerical predictions of deformations
Ivo Herle . . . 115
1 Introduction . . . 115
2 Predictions vs. measurements . . . 116
3 Constitutive model . . . 118
4 Mathematical and numerical aspects . . . 123
5 Concluding remarks . . . 124
References . . . 125

FE ultimate load analyses of pile-supported pipelines - tackling uncertainty in a real design problem
Hermann Lehar, Gert Niederwanger, Günter Hofstetter . . . 129
1 Introduction . . . 129
2 Pilot study . . . 131
3 Laboratory tests . . . 133
4 Numerical model of the pile-supported pipeline . . . 147
5 On-site measurements . . . 155
6 Design . . . 158
7 Conclusions . . . 161

References 162

The Implications of Geotechnical Model Uncertainty for Construction Management
Heimo Lessmann, Robert Vieider 165
1 The "last" 165
2 The soil / building interaction 166
3 The computational model 172
4 Safety 174
5 The probabilistic approach 176
6 Information for the site engineer 178
7 Conclusion 178
References 181

Fuzzy, probabilistic and stochastic modelling of an elastically bedded beam
Michael Oberguggenberger, Francesco Russo 183
1 Introduction 183
2 The elastically bedded beam 184
3 Fuzzy and probabilistic modelling 185
4 Stochastic modelling 189
5 Summary and Conclusions 195
References 195

Queueing models with fuzzy data in construction management
Michael Oberguggenberger 197
1 Introduction 197
2 The fuzzy parameter probabilistic queueing model 199
3 Fuzzy service and return times 204
References 208

Fuzzy models in geotechnical engineering and construction management
Thomas Fetz, Johannes Jäger, David Köll, Günther Krenn, Heimo Lessmann, Michael Oberguggenberger, Rudolf F. Stark 211
1 Introduction 211
2 Fuzzy sets 213
3 An application of fuzzy set theory in geotechnical engineering 215
4 Fuzzy differential equations 225
5 Fuzzy data analysis in project planning 227
References 238

Authors 241

Assessment of characteristic shear strength parameters of soil and its implication in geotechnical design

Wolfgang Fellin

Institut für Geotechnik und Tunnelbau, Universität Innsbruck

Summary. The characteristic shear strength parameters of soil are obviously decisive for the geotechnical design. Characteristic parameters are defined as cautious estimates of the soil parameters affecting the limit state. It is shown how geotechnical engineers interpret this cautious estimate. Due to the inherent lack of data in geotechnical investigations there is always a certain degree of subjectivity in assessing the characteristic soil parameters. The range of characteristic shear parameters assigned to the same set of laboratory experiments by 90 geotechnical engineers has been used to design a spread foundation. The resulting geometrical dimensions are remarkably different. It is concluded that geotechnical calculations are rather estimates than exact predictions. Thus for intricate geotechnical projects a sensitivity analysis should be performed to find out critical scenarios. Furthermore a continuous appraisal of the soil properties during the construction process is indispensable.

1 Characteristic values of soil parameters

1.1 Definition

European geotechnical engineers proposed a definition of the characteristic value of soil or rock parameters given in EC 7:

> "The characteristic value of a soil or rock parameter shall be selected as a cautious estimate of the value affecting the occurrence of the limit state." [4, 2.4.3(5)]

Failure in soils is generally related with localisation of strains in shear bands. Therefore, simple geotechnical limit state analyses are based on assuming shear surfaces, e.g., the calculation of stability of slopes using a defined shear surface, see Fig. 2. Thus the value affecting the limit state is the shear strength of the soil. The shear strength in the failure surfaces is usually modelled by the Mohr-Coulomb failure criterion $\tau_f = c + \sigma \cdot \tan\varphi$, with the stress σ acting normal to the shear surface. The validity of this model will not be discussed here, it should only be mentioned that it is not applicable in all cases.

Assuming that the Mohr-Coulomb failure criterion is an appropriate model, the parameters whose distributions we have to analyse are the friction coefficient $\mu = \tan\varphi$ and the cohesion c.

In the limit state[1] the shear strength is mobilised over the whole length of the shear surface. Accounting for this is usually done in the way as EC 7 proposes:

> "The extent of the zone of ground governing the behaviour of a geotechnical structure at a limit state is usually much larger than the extent of the zone in a soil or rock test and consequently the governing parameter is often a mean value over a certain surface or volume of the ground. The characteristic value is a cautious estimate of this mean value. ..." [4, 2.4.3(6)]

1.2 Intuitive Model

A very instructive model to explain this idea was presented in [6]. We consider the base friction of n equally weighted blocks on a horizontal soil surface, see Fig. 1. The blocks are pushed by the horizontal force H. The total weight of the blocks is W. Each block has the weight of W/n and the friction coefficient μ_i.

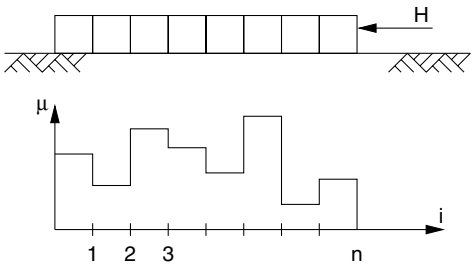


Fig. 1. Equally weighted blocks pushed by a horizontal force on a horizontal soil surface.

Each block i contributes to the resistance $\mu_i W/n$. For a constant pushing force H all blocks act together. Thus the total resisting force is

$$\sum_{i=1}^{n} \mu_i \frac{W}{n} = W \frac{1}{n} \sum_{i=1}^{n} \mu_i = W\overline{\mu} \; .$$

The limit state function for slip is therefore

$$g := \overline{\mu} W - H \; .$$

[1] Strictly spoken this is only true in critical state.

Slip occurs when $g \leq 0$. We see clearly, that in this example the mean value of the friction coefficient $\overline{\mu}$ is affecting the limit state.

Extending this idea to general geotechnical design raises at least three obvious difficulties:

Do all blocks act together? Consider that the blocks are pulled and not pushed. If the blocks are not glued together, the crucial value for the limit state is the friction coefficient of the right block and not the (spatial) mean value.

Equally weighted? The normal stress in the shear planes is usually not constant, e.g. in the case of a slope stability calculation in Fig. 2. A mean value is therefore only approximately valid for a specific stress range, i.e. depth.

Defined shear surface? The model of a defined shear surface is of course a very simple one. In reality the shear surface tends to find its way through the weakest zones. Therefore, the shape of the real shear surface in an inhomogeneous soil is different from that in the model. This increases the uncertainties, e.g., the bearing capacity is lower in inhomogeneous soil than in homogeneous soil [5].

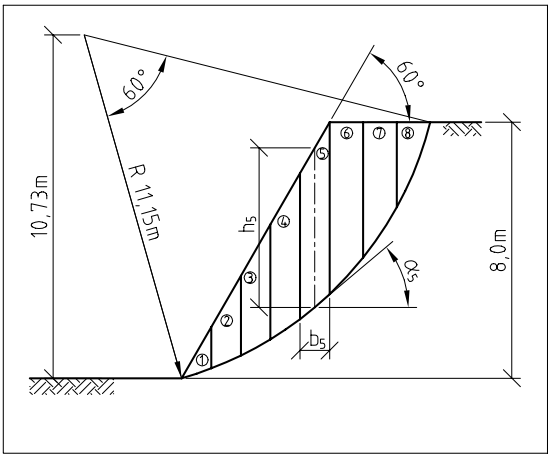


Fig. 2. Stability of slope.

Simple geotechnical limit state models use a predefined shear surface, e.g., in the calculation of slope stability a circular shear surface is used to find the circle with the minimum global safety, see Fig. 2. This implies that we can model certain regions of the soil as homogeneous materials with characteristic shear parameters. When using (spatial) mean values in these regions, one has to check carefully if the assumed failure mechanism in such simple models is really valid or if, e.g., a series of weak layers changes the presumed shear surface considerably, and therefore the mechanical behaviour of the whole

system, e.g. Fig. 3. In other words, mean values of the shear strength parameters are appropriate for being used for a soil region if the defined shear surface of the geotechnical model in this region is not significantly changed by inhomogeneities.

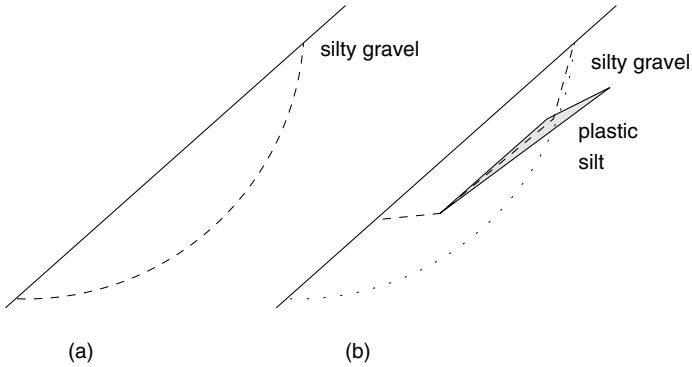


Fig. 3. Change of the shear surface, due to a weak layer:
(a) In homogeneous conditions the shear surface (dashed line) can be assumed as a circular slip curve.
(b) A weak layer attracts the shear surface.

1.3 Basic statistical methods

In geotechnical investigations the number of samples is mostly very small. Therefore, statistical methods cannot be used in a straightforward manner. Thus choosing a characteristic value requires subjective judgement, as we will see in the example below. However, knowing the statistical background gives a better understanding of what a *cautious estimate* could mean. Furthermore, using statistical methods as a basis of determining characteristic values structures the decision making process and makes it therefore clearer and, maybe, more exchangeable between different individuals.

In EC 7 statistical methods are explicitly allowed:

> "... If statistical methods are used, the characteristic value should be derived such that the calculated probability of a worse value governing the occurrence of a limit state is not grater than 5%." [4, 2.4.3(6)]

This is in agreement with EC 1:

> "Unless otherwise stated in ENVs 1992 to 1999, the characteristic values should be defined as the 5% fractile for strength parameters and as the mean value for stiffness parameters." [3, 5(2)]

Pure statistics

As discussed above, we assume in the following that the mean values of the shear parameters are decisive for the limit state. An estimate of the mean value of the population X (cohesion or friction coefficient) is the sample mean value of the n sample members

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \ . \tag{1}$$

We can further estimate a confidence interval for the mean value. The mean value of the population lies in this interval with a probability of (1-α). Thus that the probability of a mean value lower than the lower bound of the confidence interval is $\alpha/2$, because there are also values higher than the upper bound with a probability of $\alpha/2$. We can therefore define the characteristic value (5% fractile) as lower bound of the confidence interval using $\alpha/2 = 0.05$. This means a 90% confidence interval, compare [13, 15]. Using the Student's t distribution to estimate the confidence interval the characteristic value can be calculated

$$x_k = \bar{x} - t_{n-1,\frac{\alpha}{2}} \frac{s_x}{\sqrt{n}} \ , \tag{2}$$

with the sample variance s_x^2

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \ . \tag{3}$$

Values of the (1-$\alpha/2$)-quantile of the Student's t distribution with $n-1$ degrees of freedom $t_{n-1,\alpha/2}$ can be found in standard text books, e.g. [11].

Note that if certain weak zones trigger the failure, it first has to be checked whether the used simplified geotechnical model is able to capture this behaviour. There are cases in geotechnical problems in which the 5% fractile of the distribution of the soil parameters have to be used as characteristic values, as it is common for materials in structural engineering, see e.g. [1]. Remember the intuitive model in Sec. 1.2. If the blocks are pulled, one would use the 5% fractile of the distribution of the friction coefficient μ as characteristic value, and not the lower bound of the 90% confidence interval of the mean value of μ.

Incorporation of engineering experience

It is obvious that using only one to four specimens, as it is common in geotechnical engineering, a purely statistically determined characteristic value tends to be very low, due to the high uncertainty expressed by the wide confidence interval (large $t_{n-1,\alpha/2}$ and s_x).

There are several methods to incorporate a priori information about the soil — the so called engineering experience — into the determination of the characteristic value, e.g. [1, 12, 7], which are often based on Bayesian methods. They all bring additional subjectivity into the estimation of the characteristic value, and their rigorous (and sometimes complicated) formulas only pretend to be more accurate.

The simplest method is just to use the knowledge of the scatter of the properties of soils classified to be of the same soil type, e.g. from a laboratory database [1]. Some widely used values are summarised in Table 1.

Table 1. Coefficient of variation V for shear properties of soil

Soil property	V typical range		V mean [8] (recommended [12])
φ	0.06 – 0.14	[8]	0.1
	0.05 – 0.15	[12]	
	0.02 – 0.13	[2]	
c	0.3 – 0.5	[8, 12]	0.4

Assuming from experience that these variations can be applied in the case under investigation we "know" the coefficient of variation, and therefore (2) changes to (see [1])

$$x_k = \bar{x} - t_{\infty,\frac{\alpha}{2}} \frac{s_x}{\sqrt{n}} = \bar{x} - 1.645 \frac{s_x}{\sqrt{n}} \ , \tag{4}$$

with

$$s_x = \bar{x} V_x \ ,$$

and $\alpha/2 = 0.05$ as before to get the lower bound of the 90% confidence interval.

Note that adjusting the variation and/or the mean value of the sample is a very subjective decision, which must be justified by the responsible engineer. This requires usually a database [1] and local experience.

Moreover, when estimating the characteristic shear parameters, one should never try to be more accurate than the underlying mechanical model. The Mohr-Coulomb failure criterion is a very simplified model, e.g., the friction angle is actually pressure dependent and therefore not a material constant or a soil property.

2 Example

The results of classification and ring shear tests of glacial till of northern Germany are listed in Table 2 [13].

Table 2. Classification and ring shear tests, glacial till (marl), northern Germany [13]

specimen No .	fines %	sand %	gravel %	I_p	γ kN/m^3	φ °	c kN/m^2
1	55	45	0	0.25	20.7	23	55
2	28	67	5	0.17	19.7	27	15
3	47	52	1	0.20	21.3	29	16
4	40	60	0	0.17	21.9	33	13

Using (1) and (3) with $x_i = \tan\varphi_i$ we end up with the sample mean and standard deviation of the friction coefficient and angle

$$\overline{\tan\varphi} = 0.5344 \ , \ s_{\tan\varphi} = 0.0937$$
$$\overline{\varphi} = 28.1^\circ \ , \ s_\varphi = 5.4^\circ \ .$$

The same procedure with $x_i = c_i$ yields

$$\bar{c} = 24.8 \text{ kN/m}^2 \ , \ s_c = 20.2 \text{ kN/m}^2 \ .$$

The characteristic values taken as lower bounds of the 90% confidence intervals (2) with $t_{n-1,\alpha/2} = t_{3,5\%} = 2.353$ are

$$\varphi_k = 23^\circ$$
$$c_k = 1 \text{ kN/m}^2 \ .$$

Statistical information from the small sample ($n = 4$) gives rather low characteristic values due to the high uncertainty expressed by the large Student's t quantile and the large standard deviation.

A database for glacial till from the same geographical region (131 shear box tests) gives additional regional experience [1]

$$\overline{\tan\varphi} = 0.627 \ , \ V_{\tan\varphi} = 0.09$$
$$\bar{c} = 9 \text{ kN/m}^2 \ , \ V_c = 0.95 \ .$$

If we decide that these values of $V_{\tan\varphi}$ and V_c are the coefficients of variations of the population of our laboratory tests, we can use (4) to estimate the characteristic values

$$\varphi_k = 26^\circ$$
$$c_k = 4 \text{ kN/m}^2 \ .$$

All of a sudden we end up with two different sets of characteristic values. This seems to be an intrinsic problem in geotechnical engineering, due to the persistent lack of data. There is no purely mathematical justification which values are the best. The engineer has to decide which information is used, either the results of laboratory tests only, or additional information from a database or experience, see Fig. 4. Therefore (among other uncertainties) the level of safety in geotechnical problems will never be known (exactly).

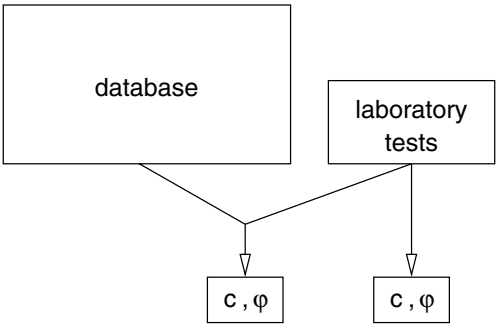


Fig. 4. Different information leads to different characteristic values.

2.1 Geotechnical engineers' opinions

In a survey the set of experimental data of Table 2 was given to 90 geotechnical engineers in Germany [13]. They were asked to determine the characteristic shear parameters for using them in a slope stability problem. The wide range of answers is illustrated in Fig. 5 and Fig. 6.

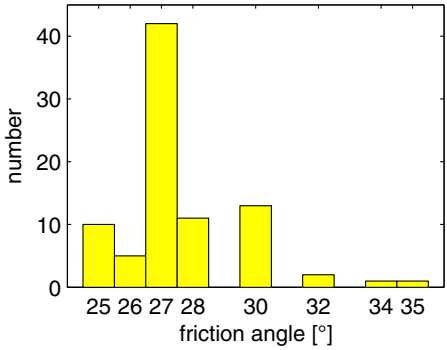


Fig. 5. Frequency of given characteristic friction angles [13]

The recommended values for the characteristic friction angle φ_k are in between 25° and 35°, where 27° was most frequently given. The range of the characteristic cohesion is much wider, from 0 to 27 kN/m^2, with 10 kN/m^2 as the most common value. The method of estimating these values remains unclear.

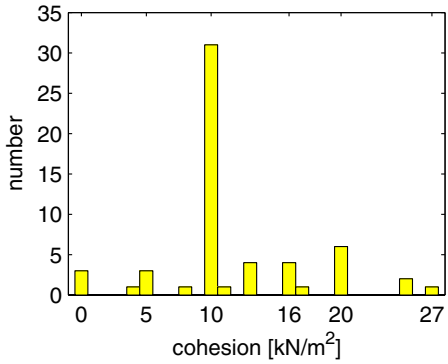


Fig. 6. Frequency of given characteristic cohesions[13]

3 Influence on design

The influence of the range of given characteristic values on geotechnical design is illustrated using the simple example of the bearing capacity of a square spread foundation, see Fig. 7.

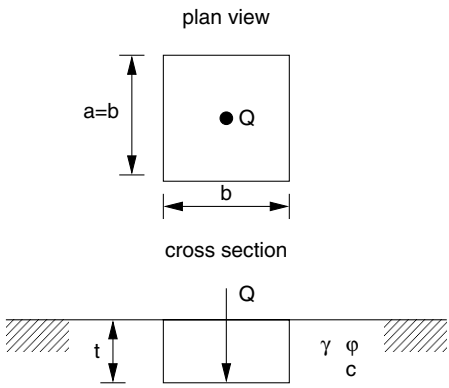


Fig. 7. Vertically and centrically loaded quadratic footing: $t = 1$ m, $Q_k = 1000$ kN, $\gamma_k = 20.9$ kN/m^3 (mean value from Table 2).

The bearing capacity is calculated according to [9]. The design value of the long term bearing capacity (drained conditions) is

$$Q_{f,d} = A'(\gamma'_{u,d} b' N_\gamma + \gamma'_{o,d} t N_q + c_d N_c) \ ,$$

with the bearing capacity factors for vertical load, horizontal surface and base plate

$$
\begin{aligned}
N_q &= s_q \frac{1+\sin\varphi_d}{1-\sin\varphi_d}\, e^{\pi\tan\varphi_d} \ , \\
N_\gamma &= s_\gamma (N_q - 1)\tan\varphi_d \ , \\
N_c &= s_c (N_q - 1)\varphi_d \ ,
\end{aligned}
$$

and the shape factors

$$
s_\gamma = 1 - 0.3\frac{b'}{l} \ , \ s_q = 1 + \frac{b'}{l}\sin\varphi_d \ , \ s_c = \frac{N_q s_q - 1}{N_q - 1} \ .
$$

In our simple example $b' = b$, $l = a$, $A' = b^2$ and the unit weight $\gamma'_{u,d} = \gamma'_{o,d} = \gamma_d$.

The design values of the soil parameters are ([9]; regular loading):

$$
\begin{aligned}
c_d &= \frac{c_k}{1.3} \ , \\
\varphi_d &= \arctan\left(\frac{\tan\varphi_k}{1.3}\right) \ , \\
\gamma_d &= \frac{\gamma_k}{1.0} \ .
\end{aligned}
$$

The design value of the load is according to [9]: $Q_d = Q_k \cdot 1.0$.

We search for the minimal dimensions $a = b$ of the footing, i.e. $a = b$ such that the design bearing capacity $Q_{f,d}$ is equal to the design load Q_d. This yields

$$
b^3 \gamma_d N_\gamma + b^2(\gamma_d t N_q + c_d N_c) - Q_d = 0 \ .
$$

The bearing capacity factors N_γ, N_q and N_c for a quadratic footing are functions of the design friction angle only. We assume the unit weight as constant, because its variability is comparably small. Thus the width b is only a function of the characteristic shear parameters

$$
b = f(\varphi_k, c_k) \ .
$$

Of course, the geotechnical engineers determined the values φ_k and c_k simultaneously. Unfortunately, the information about which φ_k corresponds to which c_k is not accessible from [13]. Therefore, we deal in the following with all possible combinations.[2]

The resulting variation of the width of the spread foundation is assembled in Figs. 8–10. The width ranges between $b = 0.85$ m und $b = 2.08$ m for all combinations of recommended characteristic values. For the most common values $\varphi_k = 27°$ and $c_k = 10$ kN/m^2 the footing dimension is $a = b = 1.50$ m.

[2] This is synonymous with assuming that the friction angle and the cohesion are not correlated, which is obviously not true. But in this qualitative study this simplification should be appropriate, as it gives an upper limit of the range of the footing dimensions.

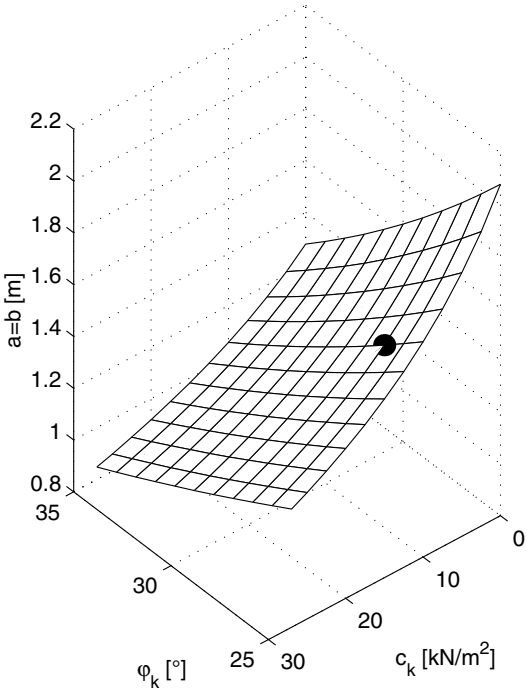


Fig. 8. Dimension of the footing (Fig. 7) for the recommended characteristic values (Fig. 5 and Fig. 6). The label • refers to the most common values $\varphi_k = 27°$ and $c_k = 10$ kN/m^2.

A calculation with the characteristic values statistically determined from the sample, $\varphi_k = 23°$ and $c_k = 1$ kN/m^2, gives even a result outside this range $b = 2.71$.

The relative frequency of the footing dimension can be calculated directly from the frequency distribution of the recommended characteristic values (Figs. 5 and 6) assuming again uncorrelated shear parameters, see Fig. 11. As above the most common value is $b = 1.5$ m, but also values of $b = 1.3$ m and $b = 1.6$ m would result with similar frequency. Extreme values of the dimensions are predicted very seldom.

Which footing dimension is *correct*? This answer can only be given in a model experiment. Assuming that all footings are loaded below the limit state, the actual safety level is drastically different. Note, that the area $a \cdot b$ of the largest footing is approximately six times greater than the area of the smallest footing. This would lead to very different construction costs.

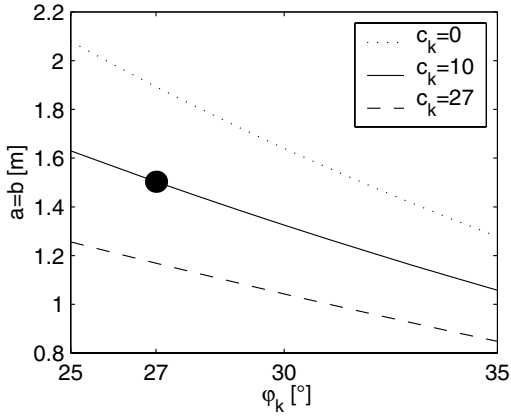


Fig. 9. Dimension of the footing (Fig. 7) for the recommended characteristic values (Fig. 5 and Fig. 6). The label • refers to the most common values $\varphi_k = 27°$ and $c_k = 10$ kN/m^2.

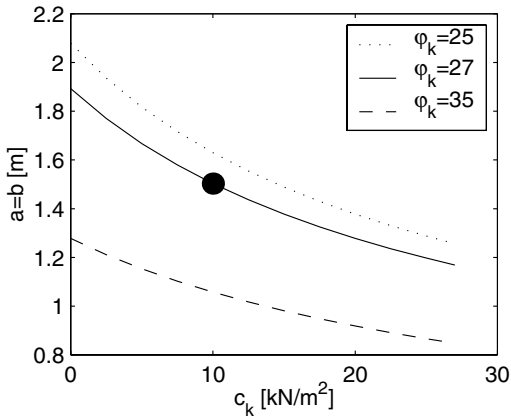


Fig. 10. Dimension of the footing (Fig. 7) for the recommended characteristic values (Fig. 5 and Fig. 6). The label • refers to the most common values $\varphi_k = 27°$ and $c_k = 10$ kN/m^2.

Is the world black and white?

Let's try to compare geotechnical engineering with the simple problem of lifting a black bowl with a crane operated by a myopic driver. The operator sees the bowl fuzzy, like geotechnical engineers that cannot exactly determine the soil parameters. The operator will open his gripper as wide as he estimates the diameter of the bowl. This can be compared with his estimation of the confidence interval. Depending on his experience with his myopic view, there

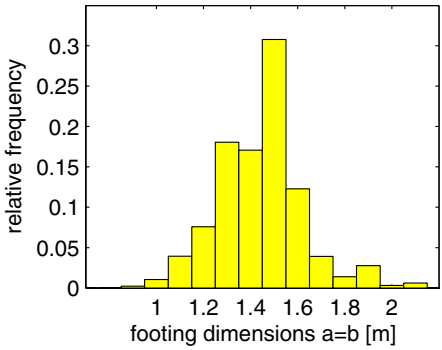


Fig. 11. Relative frequency of the the footing dimension.

is a certain possibility that gripping the bowl fails. This is something like a failure probability.

Defining a single characteristic value is comparable to wearing spectacles which cut off the light grey zones at the border of the operators' fuzzy view of the bowl and fill the rest in black. In such a way, the operator will have the illusion of a sharp (deterministic) view and tend to be sure that gripping the bowl would work. He has no chance to bring in his personal experience and is behaving like a "blind" robot.

Codes tend to offer engineers such spectacles. By fixing characteristic values and partial safety factors they cut away the intrinsic fuzziness of geotechnical design. This seems to make decisions easier but it pretends a not existing safety level.

4 Conclusions

Obviously the determination of the characteristic values as an input value for geotechnical calculations is decisive for the numerical result. The usual lack of data in geotechnical investigations leads to an ambiguous assessment of these input data. As seen in the example above this ambiguity leads to significantly different characteristic values given by different geotechnical engineers and therefore to considerable differences in geotechnical design.

The main point which has to be kept in mind is that the input of a geotechnical calculation — whatever method is used to determine the parameters — is a cautious *estimate*. Combined with other uncertainties, e.g. a crude mechanical model, the result of the calculation is therefore also an *estimate*. In 1936 Therzaghi stated:

> "In soil mechanics the accuracy of computed results never exceeds that of a crude estimate, and the principal function of the theory consists in teaching us what and how to observe in the field." [14]

One has to face the fact that regardless of the enormous developments in geotechnical theories and computational engineering since 1936, the variability of the soil will never vanish, and geotechnical engineers will always suffer from a lack of information. Thus even the result of the highest sophisticated numerical model will be more or less a crude estimate.

Therefore, an additional crucial information from a calculation is the *sensitivity* of the output with respect to the input. That provides a linearised estimation of the design variability at the considered design point due to a possible scatter of all input parameters, among them the characteristic soil properties. Input parameters with the highest output sensitivity are the crucial ones. Their values should be determined either with higher accuracy or more caution or both.

A better solution would be to take into account the uncertainties as fuzziness and end up with a fuzzy result, which provides the whole band width of the design variability due to scattering input parameters. To extract a deterministic dimension of the building from the fuzzy result is not easy either, but the responsible engineer will have much more information about the behaviour of the construction than from a single calculation.

In spite of all numerical predictions it is indispensable for sensitive geotechnical projects to appraise the soil properties and to observe the construction behaviour continuously during the construction process. A proper use of the observational method is a must [10].

References

[1] C. Bauduin. Ermittlung characteristischer Werte. In U. Smoltcyk, editor, *Grundbau-Taschenbuch*, volume 1, pages 15–47. Ernst & Sohn, 6 edition, 2001.
[2] J.M. Duncan. Factors of safety and reliability in geotechnical engineering. *Jornal of geotechnical and geoenvironmental engineering*, 126(4):307–316, 2000.
[3] Eurocode 1, Part 1. Basis of design and actions on structures, basis of design.
[4] Eurocode 7, Part 1. Geotechnical design – general rules.
[5] D.V. Griffiths and A. Fenton. Bearing capacity of spatially random soil: the undrained clay Prandtl problem revisited. *Geotechnique*, 51(4):351–359, 2001.
[6] G. Gudehus. Sicherheitsachweise für den Grundbau. *Geotechnik*, 10:4–34, 1987.
[7] J. Hanisch and W. Struck. Charakteristischer Wert einer Boden- oder Materialeigenschaft aus Stichprobenergebnissen und zusätzlicher Information. *Bautechnik*, 10:338–348, 1985.
[8] H.G. Locher. Anwendung probabilistischer Methoden in der Geotechnik. In *Mitteilungen der Schweizer Gesellschaft für Boden- und Felsmechanik*, volume 112, pages 31–36. 1985.
[9] ÖNORM B4435-2. Erd- und Grundbau, Flächengründungen, EUROCODE-nahe Berechnung der Tragfähigkeit.

[10] R.B. Peck. Advantages and limitations of the observational method in applied soil mechanics. *Géotechnique*, 19(2):171–187, 1969.

[11] E. Plate. *Statistik und angewandte Wahrscheinlichkeitslehre für Bauingenieure.* Ernst & Sohn, 1993. ISBN 3-433-01073-0.

[12] H.R. Schneider. Definition and determination of characteristic soil properties. In *Proceedings of the Fourteenth International Conference on Soil Mechanics and Foundation Engineering*, pages 2271–2274. Balkema, 1997.

[13] B. Schuppener. Die Festlegung charakteristischer Bodenkennwerte – Empfehlungen des Eurocodes 7 Teil 1 und die Ergebnisse einer Umfrage. *Geotechnik*, Sonderheft:32–35, 1999.

[14] K. Terzaghi. Relation between soil mechanics and foundation engineering. In *Proceedings of the International Conference on Soil Mechanics and Foundation Engineering*, volume 3, pages 13–18. Harvard University, Cambridge, Mass., 6 1936.

[15] A. Weißenbach, G. Gudehus, and B. Schuppener. Vorschläge zur Anwendung des Teilsicherheitskonzeptes in der Geotechnik. *Geotechnik*, Sonderheft:4–31, 1999.

Ambiguity of safety definition in geotechnical models

Wolfgang Fellin

Institut für Geotechnik und Tunnelbau, Universität Innsbruck

Summary. The design state of buildings and structures should be far from failure. To achieve this, engineers define a number called safety factor. Safety factor equal one defines the limit state. Safety factor greater than one identifies safe conditions. This paper shows with the help of a simple example that the value of the safety number depends on the used mechanical model and the definition of safety within this model. This means that the value of the safety is not an absolute measure of the distance to failure.

Recently safety has often been described by the failure probability. This failure probability can vary by orders of magnitude depending on the inevitable assumptions due to some typical unknowns in geotechnical engineering. It provides therefore just another qualitative indicator for failure.

In conclusion it is recommended to perform an additional sensitivity analysis to find the unfavourable variations and combinations of the input parameters. Therewith the worst case scenario can be found and a minimal (worst case) safety can be estimated.

1 Slope stability of a vertical slope

The simplest geotechnical model for a stability analysis of the vertical slope in Fig. 1 is to assume that failure occurs due to a concentration of the shear deformation in a planar shear band inclined with the angle ϑ, which is for the present unknown, see Fig 2.[1] The slope will fail if the shear strength τ_f is reached along the entire shear band. In a design situation we want to be far away from this state. Thus, only the so called mobilised shear strength $\tau_{f,m}$ is activated in the failure plane

$$\tau_{f,m} = c_m + \sigma \tan \varphi_m < \tau_f \ , \tag{1}$$

which is the usually used Mohr-Coulomb criterion for the shear strength.

[1] For the sake of simplicity we do not want to include any shrinking cleavage, which would drastically change the model and reduce the safety.

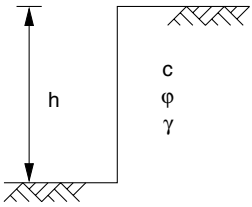


Fig. 1. Vertical slope: $h = 3$ m, unit weight $\gamma = 16$ kN/m^3.

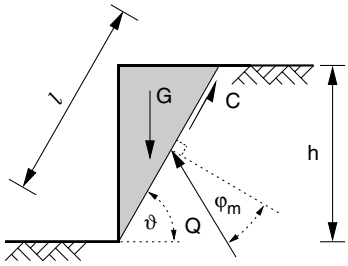


Fig. 2. Geotechnical model: planar failure surface.

In other words, we calculate an artificial equilibrium (limit) state using reduced shear parameters, namely the mobilised friction angle $\varphi_m < \varphi$ and the mobilised cohesion $c_m < c$.

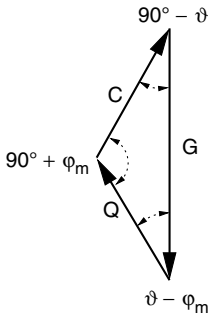


Fig. 3. Equilibrium state.

The wedge in Fig. 2 will not slide down, if the weight G (load) is in equilibrium with the cohesion force C and the friction force Q (resistance), see Fig. 3.

The weight of the wedge is

$$G = \frac{1}{2}\gamma h l \cos\vartheta \ . \tag{2}$$

The mobilised cohesion force is

$$C = c_m l. \tag{3}$$

The mobilised friction force Q acts with the angle φ_m to the geometric normal of the failure plane.

From Fig. 3 we can deduce

$$\frac{C}{\sin(\vartheta - \varphi_m)} = \frac{G}{\cos\varphi_m} \ . \tag{4}$$

Using (2) and (3) in (4) we end up with

$$\frac{c_m}{\gamma h} = \frac{1}{2}\frac{\sin(\vartheta - \varphi_m)\cos\vartheta}{\cos\varphi_m} \,. \tag{5}$$

The critical angle of the failure plane ϑ will be such that a maximal mobilised cohesion is needed to establish equilibrium for any given mobilised friction angle. Thus we can get from

$$\frac{\mathrm{d}\frac{c_m}{\gamma h}}{\mathrm{d}\vartheta} = 0 \tag{6}$$

the angle of the critical failure plane

$$\vartheta = \frac{\pi}{4} + \frac{\varphi_m}{2} = 45^\circ + \frac{\varphi_m}{2} \,. \tag{7}$$

We can also find the maximal mobilised friction angle for any given mobilised cohesion by reformulating (5) in terms of $\varphi_m(\vartheta)$ and setting $\mathrm{d}\varphi_m/\mathrm{d}\vartheta = 0$. This leads to the same angle of the failure plane (7).

Substituting ϑ in (5) using (7) provides the limit state function

$$g = 4c_m \tan\left(\frac{\pi}{4} + \frac{\varphi_m}{2}\right) - \gamma h \,, \tag{8}$$

where $g < 0$ indicates failure; $g = 0$ is called limit state.

2 Various safety definitions

2.1 Reducing shear parameters

The limit state $g = 0$ can be reached by reducing the shear parameters. Two safety factors, one for the cohesion and another for the friction angle, can be defined

$$\eta_c = \frac{c}{c_m} \,, \qquad \eta_\varphi = \frac{\tan\varphi}{\tan\varphi_m} \,. \tag{9}$$

Generally, they may have different values, i.e. the cohesion and the friction are reduced differently. We will subsequently use the following definitions

$$\eta_c = \frac{c}{c_m} \qquad \text{with} \qquad \varphi_m = \varphi \,, \tag{10}$$

$$\eta_\varphi = \frac{\tan\varphi}{\tan\varphi_m} \qquad \text{with} \qquad c_m = c \,. \tag{11}$$

By introducing (10) and (11) in the limit state function (8) and setting $g = 0$ we obtain

$$\eta_c = \frac{4c}{\gamma h}\tan\left(\frac{\pi}{4} + \frac{\varphi}{2}\right) \tag{12}$$

and

$$\eta_\varphi = \tan\varphi \cdot \tan\left[2\arctan\left(\frac{\gamma h}{4c}\right)\right] , \tag{13}$$

respectively.

A third safety factor can be defined by applying the same reduction to the cohesion and the friction, which is generally known as the FELLENIUS rule. We set

$$c_m = \frac{c}{\eta} \qquad \text{and} \qquad \tan\varphi_m = \frac{\tan\varphi}{\eta} \tag{14}$$

in the limit state function (8), and from $g = 0$ follows

$$\eta = \frac{2\sqrt{2c(\gamma h \tan\varphi + 2c)}}{\gamma h} . \tag{15}$$

2.2 Increasing loads

Another idea to reach the limit state is to increase the load with fixed shear parameters. In our example we can increase the gravity to obtain a modified unit weight

$$\gamma_m = \eta_\gamma \gamma . \tag{16}$$

Replacing γ by γ_m in the limit state function (8), setting $\varphi_m = \varphi$, $c_m = c$ and $g = 0$ yields

$$\eta_\gamma = \frac{4c}{\gamma h} \tan\left(\frac{\pi}{4} + \frac{\varphi}{2}\right) = \eta_c . \tag{17}$$

2.3 Disturbing forces compared with resistance forces

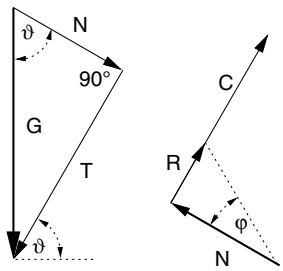


Fig. 4. Disturbing force T and resistance force $R + C$.

A safety factor can also be defined by dividing the shear resistance forces by the disturbing forces in the failure plane. This is according to Fig. 4

$$\eta_{FS} = \frac{C + R}{T} = \frac{cl + G\cos\vartheta\tan\varphi}{G\sin\vartheta} . \tag{18}$$

With $c_m = c/\eta_{FS}$, $\tan\varphi_m = \tan\varphi/\eta_{FS}$, (2) and (3) this can be rewritten

$$\frac{c_m}{\gamma h} = \frac{1}{2}(\cos\vartheta\sin\vartheta - \cos^2\vartheta\tan\varphi_m) = \frac{1}{2}\frac{\sin(\vartheta-\varphi_m)\cos\vartheta}{\cos\varphi_m} \,. \tag{19}$$

This is the same equation as (5) and therefore

$$\eta_{FS} = \eta \,. \tag{20}$$

2.4 Comparison of the different safety definitions

All aforementioned safety definitions give the same value for a slope in limit state: $\eta = \eta_c = \eta_\varphi = \eta_\gamma = 1$. If the shear parameters are higher than required, all safety factors are bigger than one but different! This is illustrated in Figs. 5 and 6.

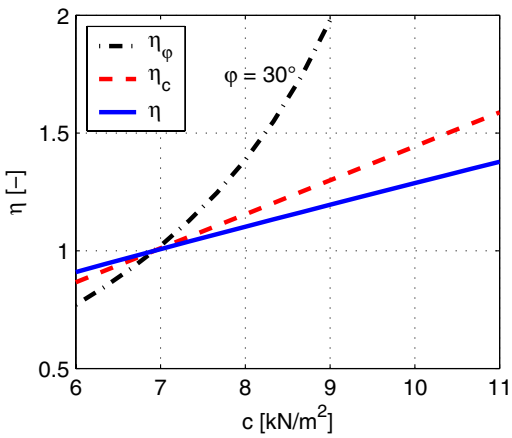


Fig. 5. Factors of safety for the slope in Fig. 1 for $\varphi = 30°$ and varying cohesion; $\eta < 1$ indicates failure.

In geotechnical calculations mostly the Fellenius definition η (14) and sometimes η_c are used; η_φ is rather unusual. One method for limit state analyses in finite element calculation is to increase the gravity until the structure fails, i.e. η_γ is used in such simulations.

The Fellenius definition gives in our example the lowest value of safety. This safety η is plotted in Fig. 7 as function of the cohesion and the friction angle.

2.5 Distance to failure?

Geotechnical engineers desire to know how far their structures are from collapse. This is a kind of distance to failure. Let us assume that the soil of the

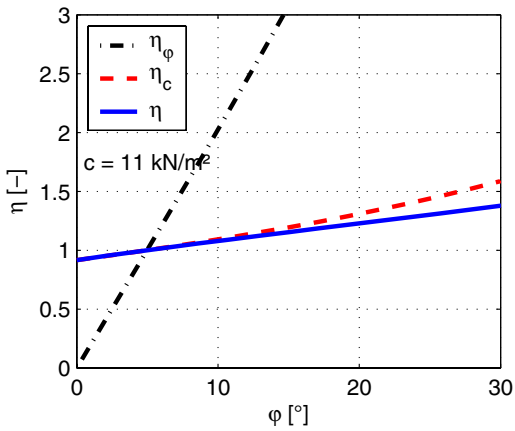


Fig. 6. Factors of safety for the slope in Fig. 1 for $c = 11$ kN/m^2 and varying friction angle; $\eta < 1$ indicates failure.

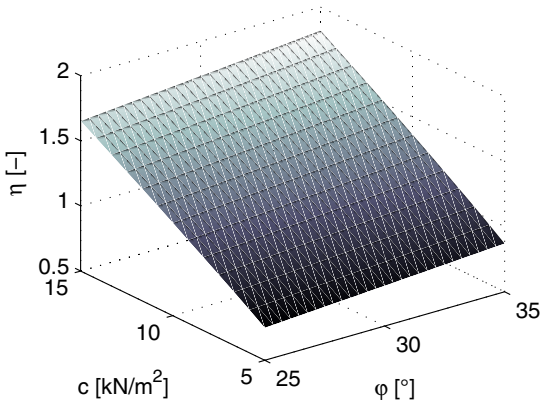


Fig. 7. Factor of safety due to Fellenius rule for the slope in Fig. 1.

slope in Fig. 1 has a cohesion of $c = 11$ kN/m^2 and a friction angle of $\varphi = 30°$. How far is this from limit state?

One measure of the distance is $(c - c_m)$ with $\varphi_m = \varphi$, which appears as vertical distance in Fig. 8. This distance is represented by the safety factor η_c: $c - c_m = c(1 - 1/\eta_c)$. Another measure is the distance $(\tan\varphi - \tan\varphi_m)$ with $c_m = c$, which is a horizontal distance in Fig. 8. An expression for this distance is η_φ. Using the Fellenius definition we measure a third distance, which is represented by η. Obviously all three distances – and therefore the safety factors – are different, although they are measures of the same state

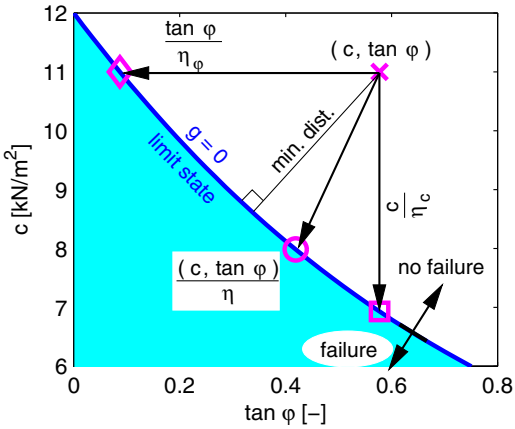


Fig. 8. Distance to limit state (8): $c = 11$ kN/m^2, $\varphi = 30°$; $\eta_c = 1.59$, $\eta_\varphi = 6.63$, $\eta = 1.38$

difference. Using the Euclidean metric the minimal distance can be found perpendicular to the curve defined by the limit state $g = 0$, which is lower than the distance obtained by η, see Fig. 8. The direction and the value of the minimal distance will change, if the coordinates are changed, e.g. if we plot g in a c–φ coordinate system, instead of a c–$\tan\varphi$ system. All these underlying conventions make the distance interpretation ambiguous.

One has to keep in mind that any safety definition is kind of arbitrary. Therefore values higher than one are no absolute measure of a distance to failure.

2.6 Design according to European codes

Based on the semi-probabilistic safety concept the new codes EC 7 [2] and DIN 1054-100 [1] are using partial safety factors for load and resistance. The characteristic values of load and resistance are increased and decreased to the design values by partial factors γ, respectively:

$$\gamma_d = \gamma_G \gamma_k \ , \tag{21}$$

$$c_d = \frac{c_k}{\gamma_c} \ , \tag{22}$$

$$\tan\varphi_d = \frac{\tan\varphi_k}{\gamma_\varphi} \ , \tag{23}$$

with the characteristic values of the shear parameters c_k and φ_k (resistance), the characteristic value of unit weight γ_k (load), the partial factors for the shear parameters γ_c and γ_φ, and the partial factor for persistent loads γ_G.

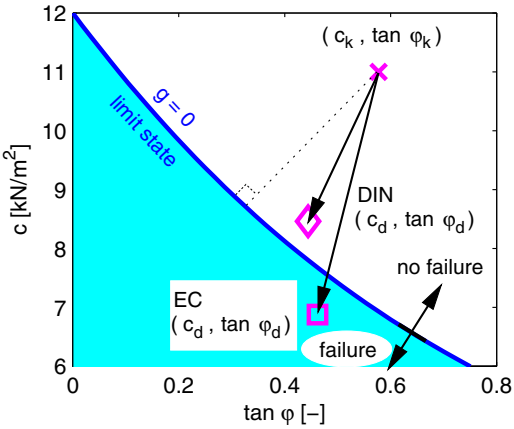


Fig. 9. Distance to limit state (8) measured with partial safety factors: $c_k = 11$ kN/m^2, $\varphi_k = 30°$, design values according to (21)-(23) and Table 1.

Table 1. Partial factors in EC 7 [2] and DIN 1054-100 [1].

	EC 7 case C	DIN 1054-100 case GZ 1C, load class LF 1
γ_G	1.0	1.0
γ_φ	1.25	1.3
γ_c	1.6	1.3

In a calculation using the design values the structure must not fail: $g > 0$.

The partial factors given in the two codes are different, see Table 9. This causes different design states, see Fig. 9. In our particular case a calculation based on EC 7 would predict unsafe conditions, whereas a calculation according to DIN 1054-100 would state safe conditions. Not only the value of the distance between characteristic and design state is different, also the direction of the distance is changing when using different codes. Again, the definition of safety is kind of arbitrary.

2.7 Probabilistic approach

The safeties η, η_c and η_φ generally do not measure the shortest distance between the actual state of the structure to the limit state $g = 0$, see Fig 10. In addition, such safety definitions do not account for the curvature of the limit state function, see Fig 11.

A more objective assessment of the risk of failure is desired. Soil varies from point to point, results of experiments scatter and there is generally a lack of site investigations. Provided that one accepts that such uncertainties

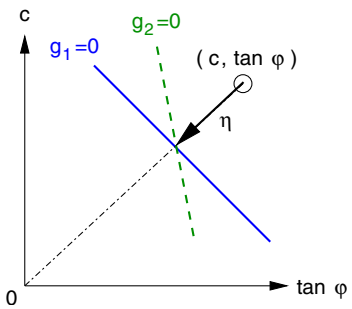


Fig. 10. Two situations with equal η (14) and different limit state functions g_1 and g_2: The distance to g_2 is shorter.

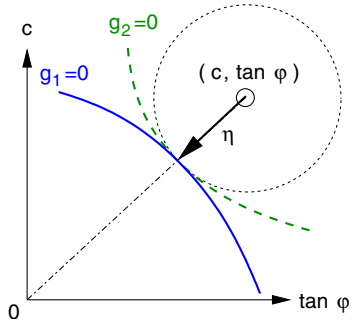


Fig. 11. Two situations with equal η (14) and different limit state functions g_1 and g_2: more combinations of c and φ exist with failure conditions in the case of g_2 (area below the limit state $g_2 = 0$ is larger than the area below $g_1 = 0$).

can be modelled with random fields, the probabilistic safety concept seems to offer such an objective measure. The idea sounds simple: If a unique and well defined limit state function exists, we just have to calculate the probability of $g < 0$, which is called failure probability p_f.

Let us try to estimate the failure probability of the vertical cut in Fig. 1. First we have to determine the probability distribution functions of the shear parameters and the unit weight. The type of these functions is usually unknown, and the choice of different types and parameter fittings has a strong influence on the resulting failure probability, especially for such small failure probabilities as required in the codes [6]. To stay simple, we assume that the unit weight does not scatter and choose normal distributions for the shear parameters, with the distribution parameters mean value μ and standard deviation σ.

The afore used characteristic values $c_k = 11$ kN/m^2 and $\varphi_k = 30°$ are cautious estimates of the mean values [3]. Thus, the mean value μ lies within an interval, of which the lower boundary is the characteristic value. Typical intervals may be

$$\mu_c = [11 \ldots 23] \text{ kN/m}^2 \tag{24}$$

$$\mu_\varphi = [30 \ldots 35] \ ^\circ \ . \tag{25}$$

If statistical data were available (24) and (25) would be chosen as confidence intervals.

We estimate the standard deviation using coefficients of variations of soil properties proposed in [5, 8]: $V_\varphi = 0.1$, $V_c = 0.4$; $\sigma = \mu V$. We further assume that $V_\varphi = V_{\tan\varphi}$.

Monte-Carlo simulations applied to the limit state function (8) with deterministic $\gamma = 16$ kN/m^3, normally distributed $c \sim \mathcal{N}(\mu_c, \sigma_c^2)$ and $\tan\varphi \sim \mathcal{N}(\mu_{\tan\varphi}, \sigma^2_{\tan\varphi})$ results in the solid line in Fig. 12. In such a simple calculation the shear parameters are assumed to be equal over the whole length l of the shear band, no spatial averaging of the shear parameters is taken into account. Therefore the resulting failure probability p_f is an upper bound.

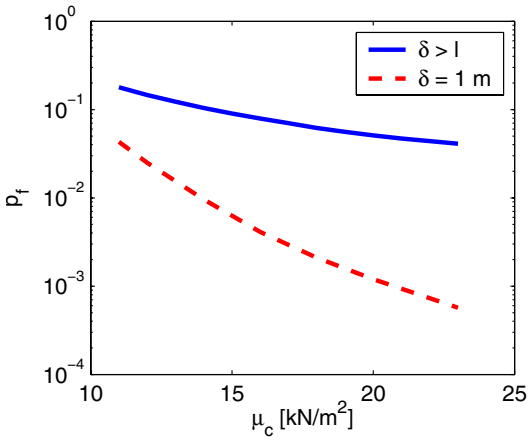


Fig. 12. Failure probability for the slope in Fig. 1 with correlation length $\delta > l$ and $\delta = 1$ m: $\mu_c = [11 \ldots 23]$, $\mu_\varphi = 30°$.

If we want to account for a spacial averaging, we have to model the shear parameters as random fields [7]. For this we have to introduce a correlation function. This is another unknown in our calculations. A list of possible choices of the correlation function is given in [7]. We want to use the simplest correlation function, where the correlation is one within the correlation length δ and zero outside. With this simple rectangular function the coefficient of variation of the averaged parameter in the shear band of length l is

$$V_\mu = V\sqrt{\frac{\delta}{l}} \qquad \text{with} \qquad \delta < l\,, \tag{26}$$

compare [4].

We set $\delta = 1$ m and calculate $l = h/\sin\vartheta$. This l would change for every realisation of $\tan\varphi$ due to (7). We use the same length l in every realisation by setting $\varphi_m = \mu_\varphi = 30°$ in (7). This error is small compared with other uncertainties. With $l = 3.46$ m the coefficients of variations are: $V_{\mu_\varphi} = 0.054$, $V_{\mu_c} = 0.21$.

Monte-Carlo simulations with $c \sim \mathcal{N}(\mu_c, \sigma^2_{\mu_c})$ and $\tan\varphi \sim \mathcal{N}(\mu_{\tan\varphi}, \sigma^2_{\mu_{\tan\varphi}})$ lead to the dashed curve in Fig. 12. Depending on the actual choice μ_c the

calculations with and without spacial averaging differs up to two orders of magnitude. In addition, different choices of the cohesion μ_c – all inside the confidence interval, and therefore with the same probability – also change p_f drastically.

At first glance the probabilistic approach looks like an objective way to measure the distance to failure. Assuming that soil parameters can be modelled in the probabilistic framework, and provided that the probability distribution functions of load and resistance, as well as the correlation functions of the shear parameters are known, the failure probability could be a good measure for the distance to failure. However, there are too many unknowns and crude estimations in such calculations, so that the failure probability is just another qualitative indicator for safety.

3 Different geotechnical models

Up to now we studied variations of safety definitions using the same mechanical model, i.e. we used the same limit state function g. But the slope stability can be analysed with different models, each of which has its certain degree of simplification. We compare four models:

- the afore presented planar shear surface;
- Bishop's method for slope stability calculation (circular shear surface);
- a wedge analysis with two rigid bodies, forces between the bodies due to fully mobilised friction and cohesion;
- a finite element calculation with reduction of the shear parameters down to collapse.

The limit state functions of these models are different in each case.

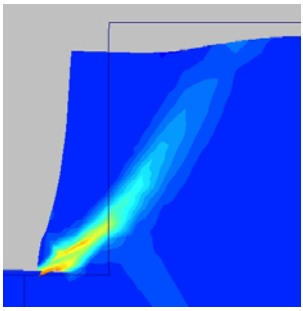


Fig. 13. Finite element calculation with Plaxis, contour plot of the total shear deformation, bright regions indicate large deformations.

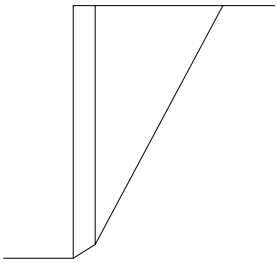


Fig. 14. Wedge analysis with GGU, shape of the rigid bodies for minimal safety η.

The standard slope stability analyses, Bishop's method and the wedge analysis were performed with the software GGU [2]. The finite element program Plaxis [3], which provides a shear parameter reduction according to Fellenius for a Mohr-Coulomb elasto-plastic constitutive model, was used with the parameters: $E = 20$ MN/m^2, $\nu = 0.3$, $c = [6 \ldots 11]$ kN/m^2, $\varphi = 30°$, $\psi = 10°$.

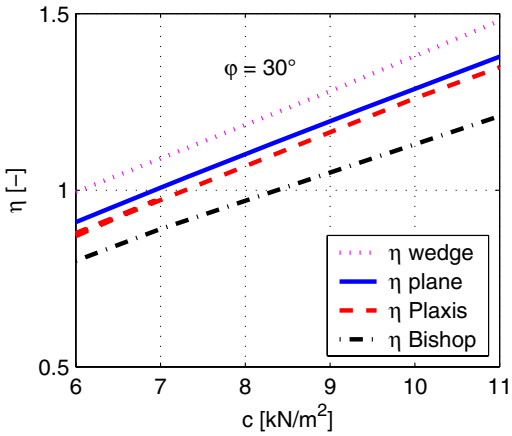


Fig. 15. Factor of safety η (14) for the slope in Fig. 1 calculated with different mechanical models: η wedge: wedge analysis with rigid bodies; η plane: planar shear surface; η Plaxis: finite element calculation, reduction of the shear parameters; η Bishop: Bishop's method, circular shear surface

The results are plotted in Fig. 15. They differ remarkably. From the collapse theorems of plasticity theory we know that the wedge analysis gives an upper bound for η, which is clearly represented in Fig. 15. For the values $c = 11$ kN/m^2 and $\varphi = 30°$ the obtained safeties are listed in Table 2. This additional uncertainty enlarges the fuzziness of the safety assessment, see Fig. 16.

4 Sensitivity analysis

As seen before, the safety factor is sensitive to the input parameters, compare Figs. 5, 6, 12 and 15. We restrict our attention to the study of the sensitivity of η with respect to the shear parameters. The variation of the safety due to variations of the shear parameters can be found by differentiation

[2] Civilserve GmbH: http://www.ggu-software.de
[3] http://www.plaxis.nl

Table 2. Factors of safety η (14) for the slope in Fig. 1 calculated with different mechanical models; $c = 11$ kN/m^2 and $\varphi = 30°$.

model / method	η
wedge analysis with rigid bodies	1.48
planar shear surface	1.38
finite element calculation	1.35
BISHOP's method	1.21

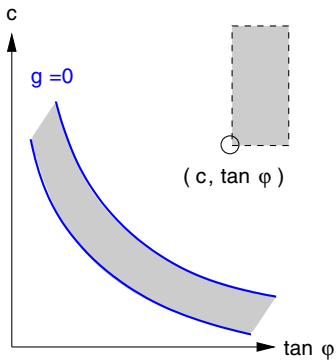


Fig. 16. Uncertainties in geotechnical safety assessment: Additional to the uncertainty of the mean values of the shear parameters (rectangular area bounded with dashed line with $(c, \tan\varphi)$ the lower bounds of the confidence intervals) there is a possible range of limit state function due to different mechanical models (area between solid curves).

$$\Delta\eta \approx \left.\frac{\partial\eta}{\partial c}\right|_{c,\varphi} \Delta c + \left.\frac{\partial\eta}{\partial\varphi}\right|_{c,\varphi} \Delta\varphi \,. \tag{27}$$

We bring this in dimensionless form

$$\frac{\Delta\eta}{\eta} \approx \underbrace{\left.\frac{\partial\eta}{\partial c}\right|_{c,\varphi} \frac{c}{\eta}}_{=:\,k_c} \frac{\Delta c}{c} + \underbrace{\left.\frac{\partial\eta}{\partial\varphi}\right|_{c,\varphi} \frac{\varphi}{\eta}}_{=:\,k_\varphi} \frac{\Delta\varphi}{\varphi} \,. \tag{28}$$

and define the dimensionless gradients k_c and k_φ. These gradients expose the most decisive parameters. The sign of k indicates if a cautious estimate of a parameter is gained by reducing (if $k > 0$) or increasing (if $k < 0$) the mean value. Furthermore one can easily estimate variations of the safety due to variations of the input.

Example:

We calculate the safety factor η for the slope in Fig. 1 with a planar failure surface. Using (15) with $c = 11$ kN/m^2 and $\varphi = 30°$ gives $\eta = 1.38$. The

partial derivatives of η are numerically evaluated at the values $c = 11$ kN/m^2 and $\varphi = 30°$

$$\left.\frac{\partial \eta}{\partial c}\right|_{c,\varphi} = \left.\frac{\partial \eta}{\partial c}\right|_{11,30} = 0.0904 \quad , \quad \left.\frac{\partial \eta}{\partial \varphi}\right|_{c,\varphi} = \left.\frac{\partial \eta}{\partial \varphi}\right|_{11,30} = 0.0155 \,. \tag{29}$$

With this we obtain the dimensionless gradients

$$k_c = \left.\frac{\partial \eta}{\partial c}\right|_{c,\varphi} \frac{c}{\eta} = 0.0904 \cdot \frac{11}{1.38} = 0.72 \tag{30}$$

$$k_\varphi = \left.\frac{\partial \eta}{\partial \varphi}\right|_{c,\varphi} \frac{\varphi}{\eta} = 0.0155 \cdot \frac{30}{1.38} = 0.34 \,. \tag{31}$$

From this we see immediately that the safety η is more sensitive to the cohesion than to the friction angle, compare Figs. 5 and 6. For a 10% variation of the cohesion we can estimate a $k_c \cdot 10\% = 7.2\%$ variation in safety: $\eta - \Delta\eta \approx 1.38 \cdot (1 - 7.2/100) = 1.28$.

5 Conclusion

The wish of an engineer to know (exactly) how far a structure is from collapse cannot be fulfilled, at least in geotechnical engineering. The calculated safety of a structure depends on the used mechanical model and the definition of safety within this model. The value of this safety is therefore no absolute measure for a distance to failure. When a certain value of safety is required, e.g. in codes, it should be stated for which model and safety definition it is valid.

In addition to the calculation of a single safety value a sensitivity analysis should be done. This helps to find the unfavourable variations and combinations of parameters. Input parameters with a strong influence on the safety should either be estimated very cautiously or investigated more intensively.

It is further recommended to do at least two calculations. First perform a standard calculation with the parameters determined as *cautious estimates*, which are in a statistical sense bounds of the 90% confidence intervals. In this calculation the safety should be greater than the required one $\eta > \eta_{\text{req}}$, or the design resistance should be larger than the design load $R_d > S_d$. In addition a second calculation, with the *worst parameters* one can think of, should be performed. In a statistical sense one can chose e.g. a bound of a 99% confidence interval for each parameter. For these worst parameters the safety should be at least larger than one $\eta_w > 1$, or the worst resistance should be larger than the worst load $R_w > S_w$ (no partial factors applied). Last but not least it should not be forgotten that not only the shear parameters scatter, there can also be a variation in the geometry of soil layers, the groundwater table and much more. A worst case scenario must also take account of such variations.

References

[1] DIN 1054-100. 3. Normvorlage der DIN 1054-100 „Sicherheitsnachweise im Erd- und Grundbau“, 1999.
[2] Eurocode 7, Part 1. Geotechnical design – general rules, 1994.
[3] W. Fellin. Assessment of characteristic shear strength parameters of soil and its implication in geotechnical design. In this volume.
[4] G. Gudehus. Sicherheitsachweise für den Grundbau. *Geotechnik*, 10:4–34, 1987.
[5] H.G. Locher. Anwendung probabilistischer Methoden in der Geotechnik. In *Mitteilungen der Schweizer Gesellschaft für Boden- und Felsmechanik*, volume 112, pages 31–36. 1985.
[6] M. Oberguggenberger and W. Fellin. The fuzziness and sensitivity of failure probabilities. In this volume.
[7] R. Rackwitz. Reviewing probabilistic soils modelling. *Computers and Geotechnics*, 26:199–223, 2000.
[8] H.R. Schneider. Definition and determination of characteristic soil properties. In *Proceedings of the Fourteenth International Conference on Soil Mechanics and Foundation Engineering*, pages 2271–2274. Balkema, 1997.

The fuzziness and sensitivity of failure probabilities

Michael Oberguggenberger[1] and Wolfgang Fellin[2]

[1] Institut für Technische Mathematik, Geometrie und Bauinformatik, Universität Innsbruck
[2] Institut für Geotechnik und Tunnelbau, Universität Innsbruck

Summary. In this article, we scrutinize basic issues concerning the interpretation of probability in the probabilistic safety concept. Using simple geotechnical design problems we demonstrate that the failure probability depends in an extremely sensitive way on the choice of the input distribution function. We conclude that the failure probability has no meaning as a frequency of failure. It may supply, however, a useful means for decision making under uncertainty. We suggest a number of alternatives, as interval probability, random and fuzzy sets, which serve the same purpose in a more robust way.

1 Introduction

This paper addresses the role of probabilistic modeling in geotechnical problems. There is a general awareness of the uncertainties in all questions of geotechnical engineering; by now, it has also become clear that the uncertainties themselves have to be modeled. Common practice is to use a probabilistic set-up to achieve this task. The probabilistic format should be seen as supporting the decision process under uncertainty, as formulated e. g. by Einstein [6]. It helps structuring the problem and aids in obtaining qualitative judgments.

The numerical values thus obtained, like failure probabilities and safety factors, play an important role in comparative and qualitative studies. The point we wish to make, however, is that these numerical values do not allow quantitative assertions about reality.

In particular, contrary to common language, the failure probability cannot be interpreted as a frequency of failure. This fact was already pointed out by Bolotin [2] as early as 1969. To support our claims, we start off with a general discussion of the probabilistic format and its underlying assumptions in the second section. In the third section we present two examples that dramatically exhibit the sensitivity of the failure probability to the choice of probabilistic model. We first fit different distributions (normal, lognormal, triangular) to the same input data obtained from laboratory experiments. All fits are verified as acceptable by means of standard statistical tests. This defines a probabilistic model of the data. Then we apply this model in safety assessments according to the procedures prescribed by

the European and Austrian codes. The first example concerns slope stability. There the failure probabilities are seen to vary by a factor of 100 and more when the type of fitted distribution is changed, depending on the safety class employed for the design. In the second example, we consider the bearing capacity of a centrically loaded square footing. In this case we demonstrate that the failure probability may vary by even higher orders of magnitude: they are seen to range between 10^{-11} and 10^{-3}, depending on the choice of fitted distribution. The required design dimensions of the footing are less sensitive to the input distribution, but still vary by a factor of 1.5 to 1.9. The sensitivity of the tail probabilities in engineering applications has been observed by other authors as well [7]. In the last section we discuss alternative concepts that are currently under investigation in reliability theory. We start with lower and upper failure probabilities obtained from interval valued probabilistic parameters, then demonstrate how random set theory can be used to assess the plausibility of the computed failure probabilities, and finally suggest the possible usefulness of fuzzy set theory in modeling expert estimates.

2 Probabilistic modeling

To set the stage, we briefly recall the format of the probabilistic safety concept. The vector R lumps together all variables describing the resistance of a structure, while S signifies the loads. An engineering model incorporating the structure as well as its geotechnical environment supplies the limit state function $g(R, S)$. Negative values of $g(R, S)$ correspond to failure. Modeling R and S as random variables, we can compute the failure probability

$$p_f = P\big(g(R, S) < 0\big)$$

provided the probability distributions of R and S and their parameters are known. However, the current codes employ critical values R_k and S_k (certain percentiles of R and S) and partial safety factors γ_R and γ_S, so that the designing engineer has to verify a relation of the type

$$R_k/\gamma_R \geq \gamma_S S_k. \tag{1}$$

In theory, the critical values and the partial safety factors are computed in such a way that (1) holds if and only if p_f attains a certain required value p_{fr}. For example, in the case of normally distributed resistance and loads and $g(R, S) = R - S$ one must choose

$$\gamma_R = \frac{1 - Q_p V_R}{1 - \beta \alpha_R V_R},\ \alpha_R = \frac{\sigma_R}{\sqrt{\sigma_R^2 + \sigma_S^2}},\ V_R = \frac{\sigma_R}{\mu_R}$$

where μ_R, σ_R are the mean and standard deviation of R, $p_{fr} = \Phi(-\beta)$, $Q_p = \Phi(1-p)$, R_k is the $(1-p) \cdot 100\%$ - percentile of R, and Φ the cumulative standard normal distribution. In practice, γ_R and γ_S are not computed but rather prescribed in the codes, where they entered as results of negotiations of the respective committees.

For a general description of the uncertainties involved in geotechnical modeling we refer e.g. to [6]. Here we wish to emphasize a few aspects that are important for interpreting the results of the probabilistic method.

Uncertainty of non-probabilistic input. At the foremost place, the soil model has to be mentioned. Is it a continuum model with an elasto-plastic, a hypoplastic

or another constitutive law, or is it a discrete model? Completely different sets of influence variables may arise in this way. Second, what is the failure mechanism? Is it assumed that failure occurs when averages of the parameters exceed certain values, or is failure assumed to be due to localized disturbances?

Uncertainty of probabilistic input. The probabilistic input in the model consists in specifying the types of probability distributions of the variables under consideration as well as the parameters of the distributions. We first point out that probability plays several roles here. It has to describe as diverse things as

- measurement errors,
- spatial data fluctuations,
- model deficiencies,
- lack of information.

While measurement errors can be considered as approximately random, there are strong arguments against the presumption of randomness of spatial data fluctuations. The fact that the soil *in situ* has a definite history contradicts the assumption of a random character. Major model deficiencies arise from the fact that parameters which vary spatially or actually depend on the state of the system are treated as constants in models that neglect precisely this variability. There is a notorious lack of data in geotechnics which often leads to insufficient information. One should also mention the often enormous deviation of parameter values obtained in comparative studies (experiments with the same soil specimen in different laboratories). Further, there are systematic errors in the choice of model; but the soil properties and reactions are interpreted through the filter of the chosen model. In addition, soil samples are necessarily disturbed when extracted, and thus the initial and boundary conditions applied in laboratory experiments may be different from the corresponding conditions in situ, or even unsuitable to reproduce them. In short, laboratory experiments introduce systematic (and at the same time unknown) errors in the determination of the parameter values. When applying a probabilistic formulation, the probabilistic model has to carry the burden of accounting for the data variability that comes from all these sources in an unknown mix.

Second, what *is* probability? In engineering modeling, at least three interpretations can be identified. There is *classical probability* which assigns probabilities from combinatorial considerations (for example, the assertion that the number k of successes in a sequence of n trials has a binomial distribution). Then there is the *frequentist interpretation* in which probability is an approximation to relative frequencies of outcomes in large samples. For the practical purpose of determining confidence intervals and performing statistical tests for the parameters of a single random variable "large" means $n \geq 20$. This sample size is rarely available in engineering practice, and this makes *subjective probability* attractive in engineering: in this interpretation, probability is just a subjective measure of confidence in the available information. Using prior knowledge and Bayes' rules, even the information of a sample of size $n = 1$ may be incorporated in the assessment of a certain parameter. It should be pointed out that *frequentist probability*, in its applications, has two facets: there is *individual* and *collective frequency*. For example, there were 75.000 large dams (more than 7.7 m high) in operation in the United States in 1999. In the preceding decade, 440 dam failures (loss of pool) of large dams had been recorded (data according to [1]; current statistics can be found in [11]). That translates into a collective frequency of failure of about 1/1700 per year. Does that mean that a

specific dam under consideration will fail once in the next 1700 years? Of course, for an assessment of the individual failure rate, individual data like local yearly precipitation averages etc. are needed. If these data are lacking, the collective frequency is a highly questionable estimate for the individual frequency. Often only a mixture of individual and collective data is available.

The necessity of estimating the individual risk from the collective risk is a well known problem for insurance premium calculations, especially in property and casualty insurance. It is known as the *insurance dilemma*. As stated, for example, by Bühlmann [4, Sect. 4.2], the "collective premium is essentially derived from statistical observation data", while "the [individual] risk premium will simply remain unknown in most cases".

Summarizing, we note that the failure probability in the engineering safety concept is an amalgam of classical, subjective and frequentist bits of information. That observation does not diminish its importance and usefulness for comparative studies of scenarios. However, the failure probabilities obtained in different engineering projects cannot be related to each other, as they depend on many individual choices that had to be made along the way. All the more, they do not have a meaning as an expected frequency of failure for an individual project. They are what Klingmüller and Bourgund have termed *operational probabilities* [10]. There is one more point which emphasizes these assertions: The probability distributions of the input data have to be chosen as part of the modeling procedure. It is common statistical practice to fit them by some likelihood principles and accept them if they pass a number of statistical tests (Chi-square test of fit, Kolmogorov-Smirnov-test). It is the aim of this paper to show that the choice of probability distribution, even among the standard types in use, has dramatic effects on the failure probability. Thus keeping the limit state model as well as all data fixed, while slightly changing the input distributions may (and usually does) completely change the value of p_f.

3 Sensitivity of failure probabilities: two examples

3.1 Soil parameters and fitting of probability distributions

For the purpose of demonstration, we choose a soil which was very well tested. The data stem from twenty direct shear tests [3] with disturbed samples of silt that had been performed with equal initial conditions. Note that this is much more than usually available on construction site. In each experiment, the water content of the specimen was near the liquid limit of the silt; the unit weights were nearly identical: $\gamma = 19.8$ kN/m^3. The shear velocity was constant and equal in each experiment. The cohesion resulted to $c = 0$ for all tests. The obtained friction angles are listed in Tab. 1.

The Eurocode EC 7, part 1, 2.4.3.(5) suggests the following definition of the characteristic soil parameters:

> *The characteristic value of a soil or rock parameter shall be selected as a cautious estimate of the value effecting the occurrence of the limit state.*

The value effecting the limit state is the shear strength of the soil $\tau_f = c + \sigma \cdot \tan\varphi$, with the normal stress σ. Therefore, the parameter whose distribution we analyze is the friction coefficient

$$\nu = \tan\varphi \ .$$

The mean value and the standard deviation of the sample are

$$\bar{\nu} = \frac{1}{n}\sum_{i=1}^{n}\nu_i = 0.474 \ , \qquad s_\nu = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\nu_i - \bar{\nu})^2} = 0.0452 \ .$$

Table 1. $n = 20$ direct shear tests

friction angle φ [°]									
25.6	25.5	24.0	26.0	24.1	24.0	28.5	25.3	23.4	26.5
23.2	25.0	22.0	24.0	24.9	30.0	27.0	24.4	24.3	29.5

There is some discussion in the literature, but no decisive conclusion, what distribution is appropriate for the friction coefficient. We consider some of those that have been proposed.

Normal distribution: The simplest and most common choice is the Gaussian normal distribution $\nu \sim \mathcal{N}(\mu_\nu, \sigma_\nu^2)$, with the estimations $\mu_\nu = \bar{\nu}$ and $\sigma_\nu^2 = s_\nu^2$. This distribution has the disadvantage that the friction coefficient can be negative, which is physically impossible. Though this happens only with low probability, a better choice may be the following lognormal distribution with two parameters, which is strictly non-negative. Nevertheless, the normal distribution is often used, and can be seen as providing a conservative (high) estimate of the failure probability.

Lognormal distribution with two parameters: In this case the natural logarithm of the friction coefficient is assumed to be normally distributed: $\ln\nu \sim \mathcal{N}(\mu_{\ln\nu}, \sigma_{\ln\nu}^2)$, with the estimations for the parameters

$$\mu_{\ln\nu} = \overline{\ln\nu} = \frac{1}{n}\sum_{i=1}^{n}\ln\nu_i \ , \qquad \sigma_{\ln\nu}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\ln\nu_i - \overline{\ln\nu})^2 \ .$$

This distribution is sometimes criticized to give too high probabilities for high friction coefficients.

When fitted to the data of Tab. 1, both distributions, normal and lognormal, pass the standard statistical tests (Kolmogorov-Smirnov and χ^2, see Tab. 3), but as we see in Fig. 1, they do not match the peak of the histogram and its asymmetry very well. A better statistical fit is given by the two following distributions.

Lognormal distribution with three parameters: We shift the lognormal distribution by the lower limit of the friction coefficient ν_0: $\ln(\nu-\nu_0) \sim \mathcal{N}(\mu_{\ln(\nu-\nu_0)}, \sigma_{\ln(\nu-\nu_0)}^2)$. The distribution parameters are chosen by maximizing the likelihood function of the given sample (maximum likelihood method). This distribution fits the histogram very well, see Fig. 2.

Triangular distribution: The triangular distribution is seldom used in engineering, but it is simple and matches the histogram also well (Fig. 2): $\nu \sim \mathcal{T}(l_\nu, m_\nu, u_\nu)$. The lower boundary l_ν, modal value m_ν and upper boundary u_ν were fitted with the maximum likelihood method.

Table 2. Parameters of the distributions

normal:	$\mu_\nu = 0.474$	$\sigma_\nu = 0.0452$	
two parameter lognormal:	$\mu_{\ln\nu} = -0.749$	$\sigma_{\ln\nu} = 0.0923$	
three parameter lognormal:	$\mu_{\ln(\nu-\nu_0)} = -2.192$	$\sigma_{\ln(\nu-\nu_0)} = 0.3601$	$\nu_0 = 0.355$
triangular:	$l_\nu = 0.394$	$m_\nu = 0.445$	$u_\nu = 0.598$

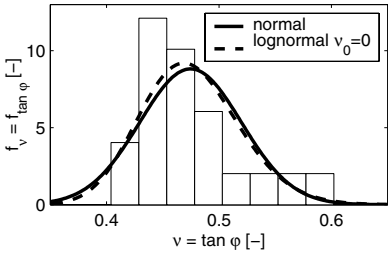


Fig. 1. Fitted normal and two parameter lognormal distribution

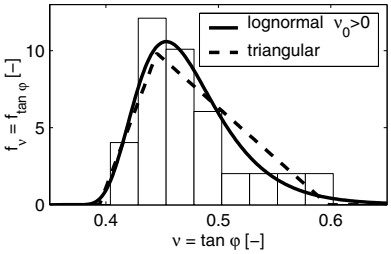


Fig. 2. Fitted three parameter lognormal and triangular distribution

As pointed out above, all four fitted distributions pass the Kolmogorov-Smirnov as well as the χ^2-test at the 95%-level (see Tab. 3). The fact that the goodness of fit is quite reasonable in all cases can also be observed visually in Fig. 3.

Table 3. Statistical tests of the distributions: computed values are below the 95%-percentile

	Kolmogorov-Smirnov		Chi-square	
	computed	95%-percentile	computed	95%-percentile
normal	0.162	0.294	5.12	5.99
two parameter lognormal	0.142	0.294	3.71	5.99
three parameter lognormal	0.113	0.294	1.99	3.84
triangular	0.154	0.294	3.48	3.84

Characteristic values: When statistical data are available (as in our case), it is commonly agreed upon [19, 20] that the lower boundary of a (1-α)-confidence interval for the mean value of $\nu = \tan\varphi$ should be taken as the "cautious estimate" of the friction coefficient $\nu_k = \tan\varphi_k$: $\nu_k = \bar{\nu} - t_{n-1,\alpha/2} s_\nu/\sqrt{n}$, where $t_{n-1,\alpha/2}$ is the (1-$\alpha/2$)-percentile of the Student t-distribution with $(n-1)$ degrees of freedom. Usually $\alpha = 0.1$ is used, that is a 90% confidence interval. Here we want to use a slightly more conservative estimate, namely $\alpha = 0.05$, which gives a 95% confidence

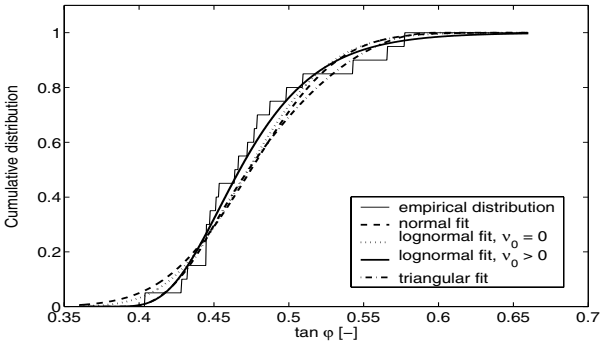


Fig. 3. Fits, cumulative distribution functions

interval. For $(1-\alpha/2)\cdot 100\% = 97.5\%$ and $n = 20$ we have $t_{n-1,\alpha/2} = 2.09$. Therefore, the characteristic friction coefficient is $\nu_k = \tan\varphi_k = 0.454$.

3.2 Spatial and local variation

We mention briefly the well-known geotechnical spatial versus local variation problem. On the one hand the strength of the material cannot be measured exactly at each point, and on the other hand the soil is usually spatially inhomogeneous. Both effects contribute to the variability of the data. For the sake of simplicity we assume in this example $\tan\varphi$ to be homogeneously distributed in space, but otherwise a random variable with statistical fluctuations.

Griffiths and Fenton [9] showed in the case of a bearing capacity problem in an undrained soil, that the spatial distribution of the undrained cohesion c_u leads to a reduced capacity factor N_c, because the failure plane tends to find its way through the weakest zones. The failure mechanism becomes asymmetric. The characteristic value of the cohesion effecting the occurrence of the limit state is therefore not the spatial mean value, it is somewhat less, depending on the practically unknown spatial correlation length.

The results in [9] let us presume that a calculation with spatially homogeneous soil but allowing statistical variation of the soil strength leads to an upper bound of the failure probability for small values of p_f (<0.05). Introducing a spatial variability of the soil parameters would bring additional unknown probability parameters into the calculation. It would not show new effects for this presentation, except a prolongation of our list of failure probabilities.

3.3 Example 1: slope stability

We are going to investigate what effect the choice of probability distribution (fitted to the same data) has on the failure probability. Our first example addresses the stability of a slope with angle β of inclination, excavated in the type of soil described in Section 3.1, see Fig. 4.

Angle of inclination: Failure of the slope occurs when

$$\tan\varphi - \tan\beta < 0.$$

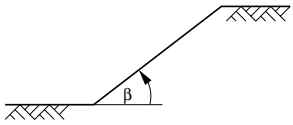


Fig. 4. Example of a slope

We choose the (deterministic) angle β as

$$\tan\beta = \tan\varphi_d = \frac{\tan\varphi_k}{\gamma_\varphi}$$

with various coefficients γ_φ as commonly used: 1.3 (persistent design situation); 1.2 (transient design situation) and 1.1 (accidental design situation). These are either the global safety factors according to Austrian code B 4433 [14] for safety class 3, or else the partial safety factors from Austrian code B 4435-2 [15].

Failure probability: In this simple example, the failure probability is given by $p_f = P(\tan\varphi < \tan\beta)$ - thus it is nothing but the cumulative distribution function of the friction coefficient $\nu = \tan\varphi$, evaluated at the tangent of the slope angle β. For example, in case of a normally distributed friction coefficient it is $p_f = \Phi\left(-(\mu_\nu - \tan\beta)/\sigma_\nu\right)$. The failure probabilities corresponding to the different types of fitted distributions are displayed in Tab. 4. In the standard case $\gamma_\varphi = 1.3$, the

Table 4. Failure probabilities p_f of slope

	p_f		
distribution	$\gamma_\varphi = 1.3$	$\gamma_\varphi = 1.2$	$\gamma_\varphi = 1.1$
normal	2.7×10^{-3}	1.6×10^{-2}	8.4×10^{-2}
2 parameter lognormal	4.9×10^{-4}	7.6×10^{-3}	6.9×10^{-2}
3 parameter lognormal	0	3.9×10^{-6}	3.0×10^{-2}
triangular	0	0	3.2×10^{-2}

failure probability p_f differs by one order of magnitude, depending on whether the friction coefficient is assumed to follow a normal or a two parameter lognormal distribution. For a three parameter lognormal and a triangular distribution, absolute safety ($p_f = 0$) is predicted. Only for a slope near its limit (failure) state (accidental situation: $\gamma_\varphi = 1.1$), p_f reacts less sensitively.

It is also interesting to compare the failure probabilities for different values of γ_φ when the type of distribution is fixed. In its traditional interpretation, the global safety factor of $\gamma_\varphi = 1.3$ means that the actual friction coefficient of the soil on site is 30% higher than the one required to guarantee stability. Intuitively, this means that there is a 30% safety margin for the slope stability. The failure probability changes from $\gamma_\varphi = 1.1$ to 1.3 by a factor of about 100 and more. This suggests a – possibly misleading – quantitatively much higher gain in safety as in the traditional interpretation using the global safety coefficient, for which the safety margin changes only from 10% to 30%.

3.4 Example 2: bearing capacity of a square footing

The second example concerns a centrically loaded square footing with the dimensions $a = b = 1.0$ m and $t = 0.5$ m in silt (Fig. 5). In this section, we will again demonstrate

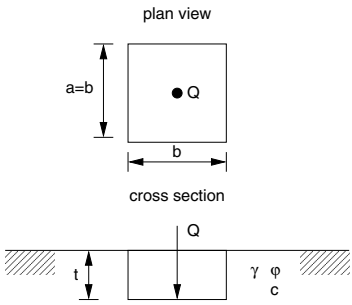


Fig. 5. Bearing capacity example

that the failure probability depends decisively on the chosen probability distribution for the friction coefficient, even more dramatically than in Section 3.2. But we will also show here that the design dimensions of the footing, when computed according to Austrian code B4435-2 [15], will vary significantly when the type of probability distribution is changed. Thus the sensitivity of the failure probability may have considerable economic effects in practice.

We use the design bearing capacity due to [15] as loading of the footing. The design values of the soil parameters are $\tan\varphi_d = \tan\varphi_k/\gamma_\varphi$, $\gamma_d = \gamma_k/\gamma_\gamma$, with $\gamma_\gamma = 1.0$ and $\gamma_\varphi = 1.3$. The design bearing capacity results to $Q_{f,d} = 101.93$ kN. The characteristic load is $S_k = Q_{f,d}/\gamma_S$ with $\gamma_S = 1.0$ according to [15]. We interpret this characteristic load as 95%-percentile of a normally distributed load with a coefficient of variation $V_S = 0.1$ according to EC 1, part 1, 4.2 (4): $S_k = \mu_S + k_N\sigma_S$, with $k_N = 1.645$. Thus the mean value and the standard deviation of the load is

$$\mu_S = \frac{Q_k}{1 + 1.645V_S} = 87.53 \text{ kN}\,, \qquad \sigma_S = V_S\mu_S = 8.75 \text{ kN}\,. \tag{2}$$

Numerical simulation: A conceptually simple method for computing the failure probability is Monte-Carlo simulation, with which we begin our elaboration. In this special example, the failure probabilities actually can be expressed analytically by a single integral; this will be presented in a later paragraph. The Monte-Carlo simulations were done with the aid of a FORTRAN-program using the random number generators for a uniform distribution $\mathcal{U}(l, u)$ (for high number of calls $N > 10^8$) and a normal distribution $\mathcal{N}(0, 1)$ published in [18]. The realizations of the friction coefficient for the various distributions were calculated as follows:

normal distribution: $\nu = \mu_\nu + \sigma_\nu \cdot \mathcal{N}(0, 1)$

two parameter lognormal distribution: $\nu = \exp\left(\mu_{\ln\nu} + \sigma_{\ln\nu} \cdot \mathcal{N}(0, 1)\right)$

three parameter lognormal distribution:

$$\nu = \exp\left(\mu_{\ln(\nu-\nu_0)} + \sigma_{\ln(\nu-\nu_0)} \cdot \mathcal{N}(0, 1)\right) + \nu_0$$

triangular distribution:

$$\nu = \begin{cases} l_\nu + \sqrt{\mathcal{U}(0,1)\cdot(u_\nu - l_\nu)\cdot(m_\nu - l_\nu)} & \text{if} \quad \mathcal{U}(0,1) \le \frac{m_\nu - l_\nu}{u_\nu - l_\nu} \\ u_\nu - \sqrt{(1-\mathcal{U}(0,1))\cdot(u_\nu - l_\nu)\cdot(u_\nu - m_\nu)} & \text{if} \quad \mathcal{U}(0,1) > \frac{m_\nu - l_\nu}{u_\nu - l_\nu} \end{cases}$$

A series of numerical simulations were carried out. In each simulation N independent realizations of the friction coefficient and the load

$$S = \mu_S + \sigma_S \cdot \mathcal{N}(0,1)$$

were computed. With each realization of the friction coefficient $\nu = \tan\varphi$ a realization of the resistance

$$R = Q_f = ab(\gamma b N_\gamma s_\gamma + \gamma t N_q s_q) \tag{3}$$

with

$$N_q = \tfrac{1+\sin\varphi}{1-\sin\varphi}\, e^{\pi\tan\varphi}\,,\ N_\gamma = (N_q - 1)\tan\varphi,$$
$$s_q = 1 + \tfrac{b}{a}\sin\varphi,\ s_\gamma = 1 - 0.3\tfrac{b}{a}$$

was calculated. In Fig. 6 we show histograms of $z = r - s$ for various distributions of the friction coefficient, each for a simulation with $N = 10^5$ number of realizations. The number of failure events $Z = R - S < 0$ was counted in H_f. The relative

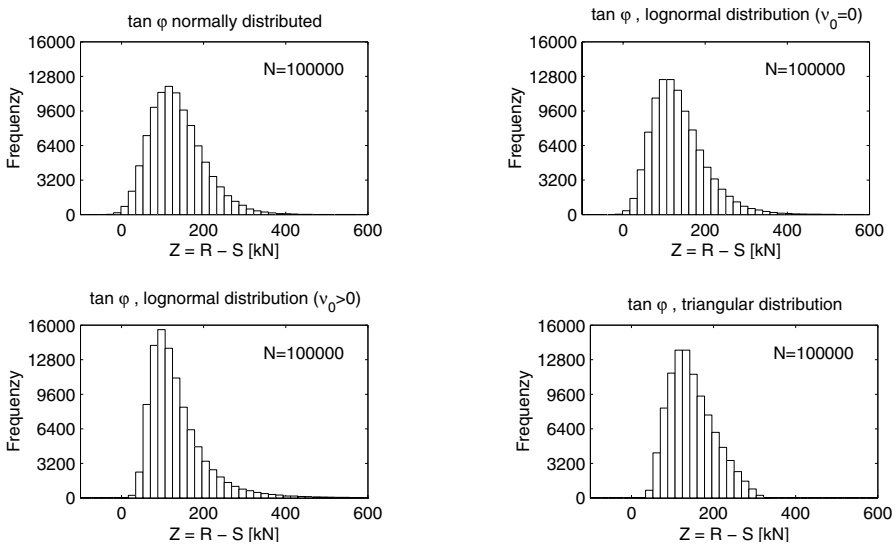


Fig. 6. Simulations with various distributions for $\tan\varphi$

frequency provides an estimate for the failure probability $p_f \approx H_f/N$. The approximative failure probabilities are listed in Tab. 5. The value of p_f varies in between 10^{-3} and 10^{-11} for different distributions of $\tan\varphi$. These variations of several orders of magnitude are confirmed by direct numerical integration below.

For curiosity we mention that transforming the data for $\tan\varphi$ in Table 1 into a histogram for R via (3) and fitting a normal distribution to R, we get an even larger failure probability of $p_f = 2 \times 10^{-2}$.

Table 5. Failure probabilities p_f in m simulations with N realizations

distribution	p_f	$N \times m$
normal	0.81×10^{-3}	$10^8 \times 4$
two parameter lognormal ($\nu_0 = 0$)	1.1×10^{-4}	$10^8 \times 4$
three parameter lognormal ($\nu_0 > 0$)	1.0×10^{-9}	$10^{11} \times 5$
triangular	1×10^{-11}	$10^{11} \times 9$

Design dimensions: If we design the dimension $a = b$ of the footing with the given failure probability $p_f = 10^{-6}$ (probability index $\beta = 4.7$, EC 1, Table A.2) we get the varying areas of the footing in Tab. 6. The minimal and the maximal footing area differ by a factor of 1.9. For a failure probability of $p_f = 10^{-4}$ this factor is 1.5.

Table 6. Dimensions of the footing for fixed failure probability $p_f = 10^{-6}$

distribution	$a = b$ [m]	difference [%]	$A = a \cdot b$ [m^2]	difference [%]
normal	1.26	35	1.59	85
two parameter lognormal ($\nu_0 = 0$)	1.11	19	1.23	43
three parameter lognormal ($\nu_0 > 0$)	0.95	2	0.90	5
triangular	0.93	–	0.86	–

Direct computation of failure probabilities: In this example, the failure probabilities can be analytically expressed as a single integral involving the distribution functions F_S of the load and the probability density f_ν of the friction coefficient. Indeed, we have that

$$\begin{aligned} p_f &= P\big(g(R,S) < 0\big) = P\big(Q_f(\nu) < S\big) \\ &= \iint\limits_{Q_f(\nu) < S} f_S(s) f_\nu(\nu)\, ds\, d\nu = \int\limits_{-\infty}^{\infty} \int\limits_{Q_f(\nu)}^{\infty} f_S(s)\, ds\; f_\nu(\nu)\, d\nu \\ &= \int\limits_{-\infty}^{\infty} \Big(1 - F_S\big(Q_f(\nu)\big)\Big)\; f_\nu(\nu)\, d\nu\,. \end{aligned}$$

In our case, we had assumed that the load was normally distributed according to $S \sim \mathcal{N}(\mu_S, \sigma_S^2)$. Thus

$$p_f = \int\limits_{-\infty}^{\infty} \left(1 - \Phi\left(\frac{Q_f(\nu) - \mu_S}{\sigma_S}\right)\right)\; f_\nu(\nu)\, d\nu\,. \tag{4}$$

The function $Q_f(\nu)$ is given by formula (3). The integral can be evaluated numerically. The values computed in this way are listed in Tab. 7 and coincide with the ones obtained by Monte-Carlo simulation (see Tab. 5).

Table 7. Failure probabilities for the footing

distribution	p_f
normal	0.811×10^{-3}
two parameter lognormal ($\nu_0 = 0$)	1.11×10^{-4}
three parameter lognormal ($\nu_0 > 0$)	1.03×10^{-9}
triangular	0.96×10^{-11}

4 Robust alternatives

4.1 Interval probabilities

As demonstrated in Sections 3.3 and 3.4, the failure probability is highly sensitive to changes in the input distribution parameters. A more robust measure as a basis for decisions appears desirable, in particular, a measure that would allow to account for model dependence and non-probabilistic uncertainties as well. Current trends in reliability and decision theory aim at relaxing some axioms of probability theory. This goes under the heading of "imprecise probability" and subsumes interval probability, convex sets of probability measures, random sets, plausibility and belief functions, fuzzy sets, and more (see [5] for a review as well as [13]). We exemplify three of these concepts in the case of the footing from Section 3.4. First, the most conservative estimate for the failure probability is obtained by admitting all probability distributions that could feasibly produce the data from Table 1. These could e.g. be modeled by the set of all probability distributions concentrated in the interval $[\underline{\varphi}, \overline{\varphi}] = [20^\circ, 32^\circ]$. Still assuming that the load S is normally distributed according to (2), it follows that the lower and upper failure probabilities are obtained by

$$\underline{p}_f = \inf\{P(Q_f(\varphi) < S) : \varphi \in [20^\circ, 32^\circ]\} = P(S > Q_f(\overline{\varphi})) = 0,$$
$$\overline{p}_f = \sup\{P\big(Q_f(\varphi) < S\big) : \varphi \in [20^\circ, 32^\circ]\} = P(S > Q_f(\underline{\varphi})) = 0.00235.$$

The correspondence of lower and upper probabilities with interval estimates is proved, e.g., in [8]. The computations are aided by the fact that the failure probabilities depend antitonically on the friction angle. Thus one simply has to compute Q_f using formula (3) and then evaluate the complementary distribution function of an $\mathcal{N}(\mu_s, \sigma_S^2)$-distribution with μ_s, σ_S from (2) at Q_f. Fig. 7 illuminates the situation. The numerical values are $Q_f(20^\circ) = 112.26$ kN which, after centering of the normal distribution, leads to $\overline{p}_f = 1 - \Phi(2.83) \approx 0.00235$, while $Q_f(32^\circ) = 543.11$ kN, leading to $\underline{p}_f = 1 - \Phi(52.07) \leq 10^{-589} \approx 0$. Less conservative estimates would be obtained by restricting the set of admitted input distributions.

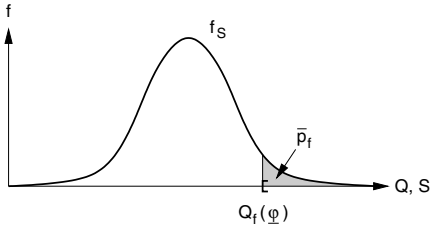


Fig. 7. Upper failure probability.

4.2 Random sets

The second idea is to get some valuation of the plausibility of the computed failure probability (cf. (Tab. 7) by means of random set theory. We demonstrate the concepts under two modeling assumptions: we assume that the load is normally distributed according to (2), and we postulate that the friction coefficient $\nu = \tan\varphi$ has a normal distribution $\mathcal{N}(\mu_\nu, \sigma_\nu^2)$. This is not mandatory, but leads to the convenient fact that the STUDENT t-distribution is applicable when computing confidence intervals, as will be needed below. The starting point of our analysis is the idea that the parameters of the probability distributions are random themselves. However, we will not describe them as probabilistic random variables (this would just lift the difficulties with the choice of type of distribution to the second level), but rather model their randomness in a more robust way as random sets. To be specific, we exemplify the approach by means of the parameter μ_ν. The point estimate inferred from the sample (Tab. 1) resulted in $\mu_\nu = \bar{\nu} = 0.474$ with the corresponding failure probability $p_f = 0.81 \times 10^{-3}$. What further evidence can be deduced from the sample for the plausibility of this estimate? The random variable

$$T = \frac{\bar{\nu} - \mu_\nu}{s_\nu / \sqrt{n}}$$

has a STUDENT t-distribution with $(n-1)$ degrees of freedom, where $s_\nu^2 = 0.0452^2$ is the sample variance and $n = 20$ the sample size. Denote by t_q the $q \cdot 100\%$-percentile of this t-distribution. Then the $q \cdot 100\%$-right sided confidence interval for μ_ν is the interval $[l_q, \infty)$ with $l_q = \bar{\nu} - t_q s_\nu / \sqrt{n}$. The confidence intervals are displayed in Fig. 8. We construct a random set whose plausibility function corresponds to the significance levels as follows (Tab. 8): We take 10 focal sets $A_i, i = 1, \ldots, 10$, each with basic probability weight $m_i = 0.1$, such that $A_1 = (-\infty, l_{0.9}], A_2 = (l_{0.9}, l_{0.8}], A_3 = (l_{0.8}, l_{0.7}], \ldots, A_{10} = (l_{0.1}, \infty)$. According to the concepts of random set theory [13], the plausibility η of the event $\mu_\nu \geq l$ is

$$\eta(\mu_\nu \geq l) = \sum_{[l,\infty) \cap A_i \neq \emptyset} m_i$$

and this equals q precisely when l belongs to the $q \cdot 100\%$-right sided confidence interval $[l_q, \infty)$ (see Fig. 9).

The friction coefficient ν was assumed to be normally distributed according to $\mathcal{N}(\mu_\nu, s_\nu^2)$. Formula (4) assigns to each value of μ_ν a failure probability $p_f = p_f(\mu_\nu)$ (observe that the probability density function f_ν in (4) depends on μ_ν). Similarly,

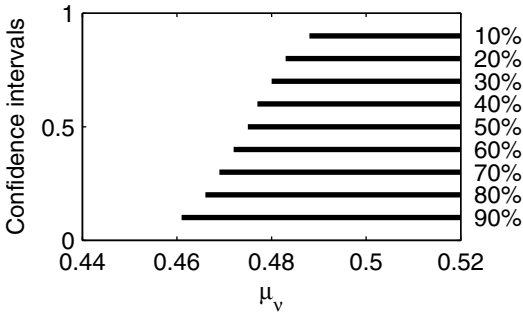


Fig. 8. Right sided confidence intervals for μ_ν

each focal set A_i is mapped into a corresponding interval $B_i = [\inf\{p_f(\mu_\nu) : \mu_\nu \in A_i\}, \sup\{p_f(\mu_\nu) : \mu_\nu \in A_i\})$ of failure probabilities. These constitute a random set of failure probabilities (Tab. 8), from which the plausibility of the event $p_f \leq p$ can be computed as follows:

$$\eta(p_f \leq p) = \sum_{[0,p]\cap B_i \neq \emptyset} m_i \, .$$

The results are shown in Fig. 10 and state, for example, that $p_f \leq 0.002$ has a plausibility degree of 0.9, while $p_f \leq 0.0005$ has a plausibility degree of only 0.3. Taking the sample mean $\bar{\nu}$ as a point estimate for μ_ν leads to the failure probability of 0.81×10^{-3} computed earlier. This estimate carries a degree of plausibility of 0.5 (dotted lines in Figs. 9 and 10).

Table 8. Random sets for μ_ν an for p_f

weight	focal sets A_i	focal sets B_i
0.1	$(-\infty, 0.461]$	$[0.00205, 1]$
0.1	$(0.461, 0.466]$	$[0.00150, 0.00205)$
0.1	$(0.466, 0.469]$	$[0.00119, 0.00150)$
0.1	$(0.469, 0.472]$	$[0.00098, 0.00119)$
0.1	$(0.472, 0.475]$	$[0.00081, 0.00098)$
0.1	$(0.475, 0.477]$	$[0.00067, 0.00081)$
0.1	$(0.477, 0.480]$	$[0.00055, 0.00067)$
0.1	$(0.480, 0.483]$	$[0.00043, 0.00055)$
0.1	$(0.483, 0.488]$	$[0.00030, 0.00043)$
0.1	$(0.488, \infty]$	$[0, 0.00030)$

The results in Figs. 9 and 10 may change quantitatively, but will be the same qualitatively if other types of distributions are assumed for the load S or the friction coefficient ν.

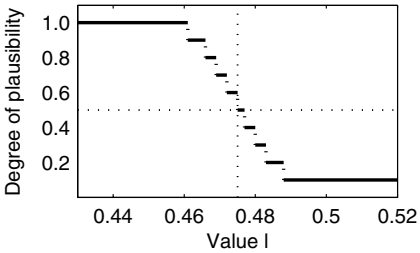


Fig. 9. Plausibility η of event $\mu_\nu \geq l$

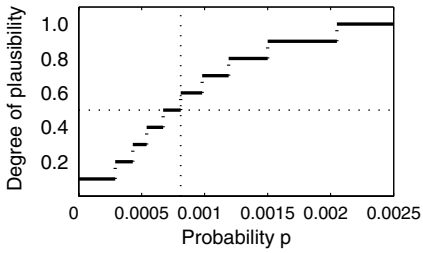


Fig. 10. Plausibility η of event $p_f \leq p$

4.3 Fuzzy sets

An even more robust non-probabilistic approach would employ fuzzy sets to describe the parameter uncertainties. The construction of a fuzzy set describing input variability may be based on an expert's risk assessment of possible ranges of the angle of internal friction, say. A detailed discussion of the interpretation and the construction of fuzzy sets can be found in [13] and the references therein. An example of a triangular fuzzy number is given in Figure 11. This says, e.g., that the

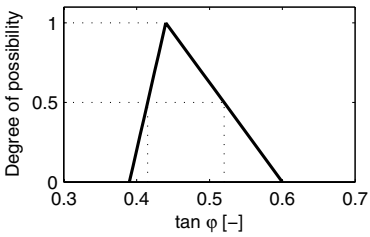


Fig. 11. Triangular fuzzy number for $\tan\varphi$

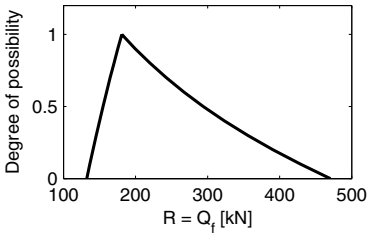


Fig. 12. Fuzzy number for bearing capacity R

expert had assigned the degree of possibility $\alpha = 0.5$ that $\nu = \tan\varphi$ is in the range of [0.415, 0.52] and possibility $\alpha = 1$ that it has the value 0.44. The computation of the fuzzy set describing the resistance is done according to the rules of fuzzy set theory (intervals of equal possibility α in input and output correspond to each other - this just means that Eq. (3) has to be evaluated on the level intervals of the fuzzy input data). The result is depicted in Figure 12 and allows, e.g., to estimate the degree of possibility that $R < S$ for any given value of S (even fuzzy S as well).

Leaving the interpretation of a fuzzy set as describing the degree of possibility of an outcome aside, one can view Figures 11 and 12 as showing interval computations, performed on each level α. This way a picture of the total sensitivity of the model - the variability of the output depending on the variability of the input - is obtained. Such an information is equally useful as a framework for decision making under uncertainty.

5 Conclusion

The computational outcome of a probabilistic model is the failure probability. It reacts highly sensitively to the choice of the input parameters and thus cannot be taken as an absolute measure of safety, in general. A reduction of the fuzzy input data to characteristic values, as suggested in the new semi-probabilistic safety concept, makes sense only for the simplest design problems, because the variability of geotechnical parameters is generally very high and largely unknown at the same time. The failure probability is a poor basis for the final assessment of the safety of a structure. When dealing with complex geotechnical problems, the engineer should give extensive consideration to modeling the data uncertainty and study how it propagates through the model. For this task, methods extending the probabilistic framework, like random sets or fuzzy sets, provide simple additional tools. The responsibility for interpreting the uncertainty of the result case by case rests with the designing engineer.

References

[1] G.B. Baecher and J.T. Christian. Risk analysis and the safety of dams. Keynote Address, 4th Cairo International Conference on Geotechnical Engineering, Cairo University, January 2000.

[2] V. V. Bolotin. *Statistical Methods in Structural Mechanics.* Holden-Day, Inc., San Francisco, 1969.

[3] E. Bösinger. Serienuntersuchungen zum Vergleich verschiedenener Rahmenschergeräte. Institut für Bodenmechanik und Felsmechanik der Universität Karlsruhe, personal communication, 1996.

[4] H. Bühlmann. *Mathematical Methods in Risk Theory.* Springer-Verlag, Berlin 1970.

[5] G. De Cooman, T. Fine, and T. Seidenfeld, editors. *ISIPTA'01, Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications.* Shaker Publishing BV, Maastricht, 2001.

[6] H.H. Einstein. Quantifying uncertain engineering geologic information. *Felsbau*, 19(5):72–84, 2001.

[7] I. Elishakoff. What may go wrong with probabilistic methods? In I. Elishakoff, editor, *Whys and Hows in Uncertainty Modelling: Probability, Fuzziness and Anti-Optimization*, volume 388 of *CISM Courses and Lectures*, pages 265–283. Springer, Wien, New York, 1999.

[8] Th. Fetz and M. Oberguggenberger. Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering and Systems Safety,* 85:73–87, 2004.

[9] D.V. Griffiths and A. Fenton. Bearing capacity of spatially random soil: the undrained clay Prandtl problem revisited. *Geotechnique*, 51(4):351–359, 2001.

[10] O. Klingmüller and U. Bourgund. *Sicherheit und Risiko im Konstruktiven Ingenieurbau.* Verlag Vieweg, Braunschweig/Wiesbaden, 1992.

[11] M. W. McCann. National Performance of Dams Program. Stanford University, http://npdp.stanford.edu/center.html.

[12] M. A. Meyer, J. M. Booker. *Eliciting and Analyzing Expert Judgement.* SIAM, Philadelphia PA, 2001.

[13] M. Oberguggenberger. The mathematics of uncertainty: models, methods and interpretations. *In this volume*.
[14] ÖNORM B4433. Erd- und Grundbau, Böschungsbruchberechnung, 1987.
[15] ÖNORM B4435-2. Erd- und Grundbau, Flächengründungen, EUROCODE-nahe Berechnung der Tragfähigkeit, 1999.
[16] ÖNORM ENV 1991-1. Eurocode 1: Grundlagen der Tragwerksplanung und Einwirkung auf Tragwerke, Teil 1: Grundlagen der Tragwerksplanung, 1996.
[17] ÖNORM ENV 1997-1. Eurocode 7: Entwurf, Berechnung und Bemessung in der Geotechnik, Teil 1: Allgemeine Regeln, 1996.
[18] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes*. Cambridge University Press, 2nd edition, 1992.
[19] B. Schuppener. Die Festlegung charakteristischer Bodenkennwerte – Empfehlungen des Eurocodes 7 Teil 1 und die Ergebnisse einer Umfrage. *Geotechnik*, Sonderheft:32–35, 1999.
[20] A. Weißenbach, G. Gudehus, and B. Schuppener. Vorschläge zur Anwendung des Teilsicherheitskonzeptes in der Geotechnik. *Geotechnik*, Sonderheft:4–31, 1999.

The mathematics of uncertainty: models, methods and interpretations

Michael Oberguggenberger

Institut für Technische Mathematik, Geometrie und Bauinformatik, Universität Innsbruck

Summary. This article discusses various mathematical theories that have been put forth as tools for modelling uncertainty, among them probability, interval arithmetic, random sets, and fuzzy sets. After recalling the definitions, we stress their interpretations (semantics), axioms, interrelations as well as numerical procedures and demonstrate how the concepts are applied in practice.

1 Introduction

An adequate understanding of the influence of input parameter variability on the output of engineering computations requires that the uncertainty itself is captured in mathematical terms. This section serves to present a number of mathematical approaches, focused around generalizations of probability theory, that are able to formalize the state of knowledge about parameter uncertainty.

Scientific modelling in engineering has to deal with three facets. First, there is reality (with soils, materials etc.). Second, there is the model of reality (formulated in mathematical terms and containing physical laws and constitutive equations). Third, correspondence rules (prescribing how to translate one into the other) are needed. It is important to note that theory is prior to observation [77, Sect. 30]. In this context this simply means that the physical model decides what are the state variables and what are the material constants, the *parameters* to be observed. Once this has been decided, the values of the parameters have to be determined from information extracted from the real world and will serve as input in the physical model. This input in turn is processed numerically and should deliver an output describing the behavior of the structure under investigation plus an assessment of the output uncertainty. Thus, models of the data uncertainty should reflect and incorporate the level of information available on the data and, second, must be able to propagate it through numerical computations and deliver an output whose uncertainty is formulated in the same terms. In addition, the uncertainty models need correspondence rules themselves, that is, well-defined semantics.

There is one more aspect to be taken into account, and that is how the information on the uncertainty of *different parameters* is combined – this refers to the

axiomatics of the underlying approach. For example, in a sum $A + B$ of two parameters A and B, will the joint uncertainty be the smallest interval containing all realizations $a+b$, or do we believe that extreme combinations of realizations $a+b$ are less *probable* than those near the *standard values*? The first choice implies adherence to the axiomatics of interval arithmetic, the second to the axiomatics of probability theory. As we will see, many more axiomatic systems exist in between these two extremal viewpoints.

Thus three aspects of the modelling of uncertainty are isolated:

- **Definition and axiomatics:** How is uncertainty described and what are the combination rules?
- **Numerics:** How is uncertainty propagated through the computational model?
- **Semantics:** What is the meaning of the results - what do they say about our conception of reality?

In this paper, we will first introduce a number of theories that are in use in engineering risk assessment, in a section entitled *definitions*. We will touch upon the corresponding axiomatics, viewed from the perspective of set functions, in the section on *axiomatics*, which will reveal a certain hierarchical order of the approaches we present. Big emphasis will be laid on a discussion of the different *semantics*, while the computational aspects will be briefly discussed in the section on *numerics*. Finally, we will comment on *multiple parameter* models, notions of independence, and on incorporating spatial or temporal parameter variability.

The need for a proper understanding of semantics is illustrated by the use and notorious public abuse of the concept of failure probability. For example, in an interview an expert on nuclear power plants claimed on Austrian public television [45, in 2001] that Western European nuclear power plants are 100 times safer than the Czech plant Temelin. Asked why, he answered that in a risk analysis, Czech scientists had computed a probability of 10^{-4} for a major nuclear accident in Temelin, while the Western European plants are designed so that such an accident may occur only once in 10^6 years. By nature, failure probabilities are clearly *subjective* probabilities which can only and actually do play a useful rule in studying scenarios. However, no absolute (that is, non-contextual) meaning can be attached to their numerical values, as already pointed out by the founders of the probabilistic safety concept in engineering [12, Chap. 2]. In the context of a particular risk analysis, given knowledge about the premises, the number 10^{-4} has a well-defined interpretation. On the other hand, the number 10^{-6} has the character of a normative computational design value in the European codes [48]. It is devoid of any real-world meaning, and in particular, can *never* have a rational - verifiable - interpretation as a frequency of failure. It is regrettable that in public statements about safety the different interpretations of probability theory are frequently mixed in an unjustified and possibly misleading way. For a similar, historical discussion of the evolvement of the numerical values of reliability assessments in the aerospace-, nuclear- and chemical process sectors we refer to [8].

Putting these delicate general questions aside (an excellent overview of the different interpretations of probability is [33]), we shall pursue the more modest goal of capturing *parameter variability* in this article. Prior to parameter variability, major aspects that have to be addressed in the analysis of uncertainty are: the choice of structural model, the adequate selection of state variables and constants, and the proper limit state function (failure mechanism). Some of these questions are

addressed in the contributions [27, 44, 53] in this volume. It seems probable that the only road to understanding these kinds of uncertainties is through case studies, studies of scenarios, simulations, and developing an appreciation of what the underlying structural theory says. A promising new approach to building models of the model uncertainty can be found in [54].

The question of parameter variability is intricate enough. It can be attributed to a number of causes. There are random fluctuations, lack of information, random measurement errors, but also systematic measurement errors (e.g. in soil mechanics, when the laboratory test destroys the history of a material with memory). There are fluctuations due to spatial inhomogeneity, errors made by assigning parameter status to state variables; and variability arising from the fact that parameters have to carry the burden of model insufficiency. The available information on data uncertainty may range from frequency distributions obtained from large samples, values from small samples or single measurements, interval bounds, to experts' point estimates and educated guesses from experience.

Probabilistic models of these kinds of parameter uncertainty may require the engineer to supply more stringent information than he has at his hands. In such a case it may be preferable to formalize uncertainty in one of the alternative ways that have been developed in the past decades, that may be better adapted to the type of available data and that do not require unwarranted assumptions (thereby avoiding the *law of decreasing credibility*, according to which the credibility of inference decreases with the strength of assumptions maintained [59]). We shall now present a spectrum of approaches ranging from interval analysis, probability theory, random sets to fuzzy sets and possibility theory. All these theories come with appropriate correspondence rules (interpretations) and can be correctly and consistently entered in numerical computations.

Additional theoretical background and references to further approaches not discussed here can be found at the end of Section 2.

2 Definitions

In this section we shall focus on describing the main theories of uncertainty in the univariate case of a single parameter. The following convention will be in use throughout the paper: parameters will be denoted by upper case letters, e.g. A, while corresponding lower case letters, as a, will be reserved for their realizations. The description of the semantics is deferred to the next section, except for the straightforward first two cases.

a. Deterministic values: The simplest approach is what in engineering terminology usually is called *deterministic description*, that is, the parameter A is described by a single value a, see Figure 1.

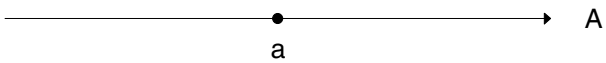


Fig. 1. Deterministic value a of parameter A

The semantics is simply that a is an expert estimate (or educated guess). Clearly, possible variations are not modelled in this approach, though their influence can be assessed to some extent by a classical sensitivity analysis. This will be elaborated a bit more in the section on *numerics*; a detailed exposition is in [75].

b. Intervals: The next level in modelling uncertainty is interval analysis. The uncertainty of the input A is described by an interval $[a_L, a_R]$, signifying bounds in terms of a worst/best case assumption. This way, the total variability is captured, but no detailed information on the uncertainty is provided (Figure 2).

Fig. 2. Interval bounds for parameter A

c. Probability: The most informative, but also most stringent description of the uncertainty of a parameter A is by means of *probability*. If the probability distribution is given by a density $p_\lambda(a)$, the probability that the realizations of the parameter A lie in a set S is

$$P(A \in S) = \int_S p_\lambda(a)\,\mathrm{d}a\,,$$

see Figure 3. The notation p_λ indicates that, usually, the probability distributions

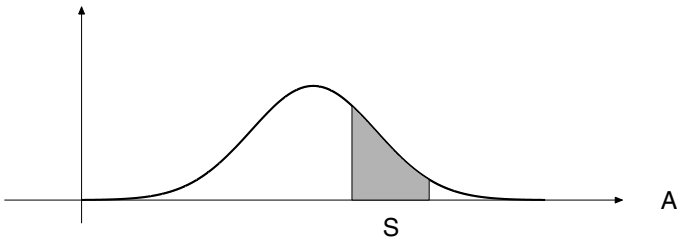


Fig. 3. Probability distribution

arise as members of a class of distributions which in turn are parametrized by parameters λ. For example, the class of Gaussian normal distributions $\mathcal{N}(\mu, \sigma^2)$ is given by the Gaussian densities $p_\lambda(a)$ with parameters $\lambda = (\mu, \sigma)$. Thus the complete specification of a probability distribution requires determination of the type it belongs to as well as the values of its parameters.

d. Sets of probability measures: A central idea in relaxing the precision inherent in a probabilistic model is to replace the single measure by a set of probability measures, say a family

$$\mathcal{M} = \{p_\lambda : \lambda \in \Lambda\}$$

where the parameter $\lambda \in \Lambda$ specifies each participating single measure. A set of probability measures defines lower and upper probabilities according to the rules

$$\underline{P}(A \in S) = \inf\{P(A \in S) : P \in \mathcal{M}\}, \quad \overline{P}(A \in S) = \sup\{P(A \in S) : P \in \mathcal{M}\}.$$

The idea of sets of probability measures is inherent in Bayesian statistics, where the distribution parameters λ are considered to be random variables themselves. In the frequentist interpretation, there is the notion of sample statistics with its confidence intervals for certain sample parameters. Convex sets of probability measures play an important role in Choquet's theory of capacities [14], see [46]. A new axiomatic approach of interval probability has been introduced by [94]. We will say more about the role of sets of probability measures in the section on semantics.

As most of the models of uncertainty, the two subsequent theories - random and fuzzy sets - can be seen as special prescriptions for obtaining *sets of probability measures*.

e. Random sets: A *random set*, sometimes also referred to as a *Dempster-Shafer structure*, is given by finitely many 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. 4. The general case of an infinite number of focal sets can be treated as well; the concept of a set-valued random variable being the defining notion. For this general case we refer e.g. to [36].

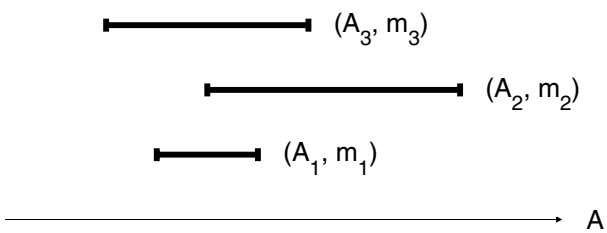


Fig. 4. A random set

In the Dempster-Shafer approach [16, 85], 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

$$\gamma(S) = \sum_{A_i \subset S} m(A_i)\,, \quad \eta(S) = \sum_{A_i \cap S \neq \emptyset} m(A_i).$$

A random set can also be interpreted as a prescription for a set of probability distributions. Denote by $\mathcal{M}(A_i)$ the totality of all probability measures supported by A_i, that is, a probability measure P on the underlying set X belongs to $\mathcal{M}(A_i)$ iff $P(A_i) = 1$. The set of probability measures induced by the given random set is

$$\mathcal{M} = \{P : P = \sum m(A_i) P_i, P_i \in \mathcal{M}(A_i)\}.$$

One can show [30, 32] that the corresponding lower and upper probabilities coincide with the degrees of belief and plausibility, respectively. That is, for any subset[1] S of

[1] When considering sets of probability measures a common framework is needed. Full mathematical precision can be achieved, for example, by requiring that X is a topological space, probability measure is interpreted as *regular Borel measure of mass one*, and subset means *Borel-measurable subset*, cf. [40].

X, it holds that

$$\underline{P}(S) = \gamma(S)\,, \quad \overline{P}(S) = \eta(S).$$

f. Fuzzy sets: Fuzzy sets can be viewed as ordered families of sets or as membership functions. It is simplest to describe the ideas by means of the special case of a fuzzy real number A. From the first point of view, A is a family of parametrized intervals. The parametrization is done in terms of levels α, $0 \leq \alpha \leq 1$. Each level α has a corresponding interval A^α so that $A^\beta \subset A^\alpha$ if $\alpha \leq \beta$. Thus the intervals are stacked and can be depicted by their left/right contour functions, see Figure 5. More generally, one could allow the A^α to be arbitrary, stacked subsets of a given set of objects under investigation (complex numbers, vectors, matrices, functions or the like).

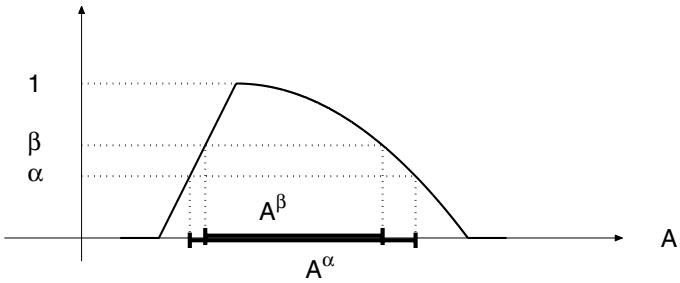


Fig. 5. Fuzzy set as family of intervals

In the second approach, the contour function is taken as the primary object, and a fuzzy set A (over the real numbers) is just a map from the real line to the interval $[0,1]$, assigning to each real number a a value $\pi_A(a) \in [0,1]$. This value may be interpreted as the membership degree to which a belongs to the fuzzy set A, or in the language of parameters, as the degree of possibility that the parameter A takes the value a, see Figure 6. In classical set theory, the membership degree is either 0 or 1; fuzzy set theory allows gradual membership as well. The intervals from the first interpretation are now the *α-level sets* $A^\alpha = \{a : \pi_A(a) \geq \alpha\}$.

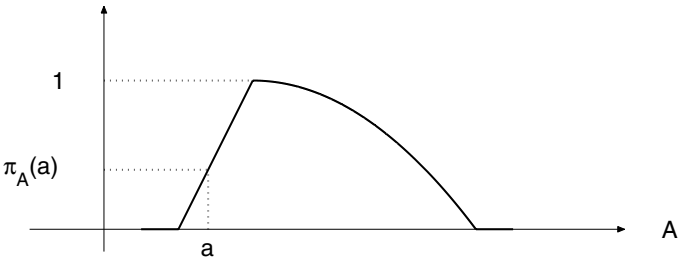


Fig. 6. Fuzzy set as degree of possibility

In analogy to the situation in probability theory, one can introduce a *possibility measure* on the underlying set, defining a degree of possibility for each subset by

$$\pi_A(S) = \sup\{\pi_A(a) : a \in S\},$$

giving the degree of possibility that the parameter A takes a value in S. The possibility measure is monotone, i.e., $\pi_A(S) \leq \pi_A(T)$ if $S \subset T$. Possibility measures are actually in one-to-one correspondence with fuzzy sets; given a possibility measure π, its evaluation on singletons defines the membership function of a fuzzy set: $\pi_A(a) = \pi(\{a\})$.

For additional theoretical background on the concepts discussed here we refer the interested reader to the following sources: interval analysis [64], fuzzy sets [6, 11, 20, 47, 49, 95, 96], possibility theory [20, 98], random sets [28, 37, 67, 85, 95], random sets and possibility theory in engineering [1, 3, 7, 9, 11, 29, 31, 35, 38, 39, 61, 72, 88, 89], probability in geotechnics [2, 23, 69, 78], comparative studies and further concepts [15, 25, 26, 49, 57, 63, 70, 71, 73, 83, 86, 87, 91, 92]; see also the unifying concept of *clouds*, recently proposed by [66].

3 Semantics

As outlined in the introduction, the interpretation of a theory is an essential ingredient for achieving an adequate translation from model into reality and back. Needless to say that the assertions made by a model become meaningful only in the context of the underlying semantics. Different semantics imply different meanings. One has to be aware of the interpretations used when comparing assertions made by different authors, all the more so as often the same vocabulary is employed for notions that differ in the various interpretations.

a. Probability: The interpretation of probability has been the subject of scientific dispute for centuries. We mentioned [33] as a recommendable reference in the Introduction. The most prevalent and important semantics in engineering practice are:

1. *Classical probability*, based on principles like the *principle of non-sufficient reason* would, in colloquial terms, determine the probability of an event S as the fraction of favorable cases among the possible cases.
2. *Frequentist probability*, based on the idea of random occurrence of an event in a sequence of independent trials, would approximate the probability of an event S by its relative frequency.
3. *Subjective probability* is meant to be a measure of personal confidence. It can be assessed by introspection and/or elicitation through experts.

In its applications, classical probability often takes the form of *combinatorial* probability. Aside from the trivial application in computing the chances in a lottery, it is often the means by which the standard probability distributions are derived, like the binomial or geometric distributions. Another example would be the exponential distribution for the survival time of a radioactive particle which is an immediate combinatorial consequence of the law of radioactive decay.

The central idea of frequentist probability is the sample with its statistical parameters. It is viewed or designed as a sequence of independent realizations of the random variable whose distribution parameters have to be determined - keeping the boundary conditions constant. The relative frequencies of the realizations of an event are taken as estimates for the probability of the event. The sample parameters like sample mean or sample variance correspond to moments of the random variable - expectation value and variance in this case. From there, the parameters of the distribution of the random variable can be estimated. This is one of the wide-spread procedures for fitting models based on the frequentist interpretation.

Going one step further and allowing the notion of samples of samples, criteria for decision making about the quality of a frequentist estimate are obtained: confidence intervals or tests of fit, and so on. From the viewpoint of the philosophy of science, the frequentist interpretation carries a number of problems, among them the question *whose* probability is realized in the sample (of the random variable, or of the experiment which was designed to measure it - a clear answer to this question has been given by [76]). A pragmatic approach with a cautious and critical attitude has proven to provide a successful basis for probabilistic models in many sciences, including physics.

A further issue of debate has been the fact that the decision aids mentioned above provide meaningful evaluations only if the sample size is *sufficiently large*, a condition which remains vague, and in civil engineering with sample sizes of 0, 1, 2 or 3 is often lacking. This is the point where subjective probability enters engineering. When such a switch of interpretation is undertaken, we believe that it is the responsibility of the engineer to put it in the open. Otherwise the meaning of probability in the final result is lost or at least obscured.

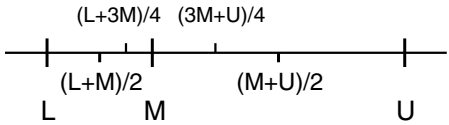
Turning to subjective probability, we first mention that schemes have been developed that allow to deduce it from decision theoretic principles, assuming *rational behavior* of the agent. This has been done to obtain operational ways of extracting the personal probability assessment of an agent quantitatively. One way promoted by [80] has been the notion of *indifference price*. To determine a probability of an event S, the decision maker is required to imagine a gamble which pays one monetary unit if event S occurs and zero otherwise. The decision maker surely would buy the gamble at the price of zero units, but surely not for more than one unit. Raising the lower bound for the price and lowering the upper bound should eventually lead to a price at which the decision maker is indifferent to buying the gamble or not. This indifference price is the probability $P(S)$ of the event. It is argued that the indifference price also equals the minimal price $\underline{P}(S)$ at which the decision maker is willing to sell the gamble, as well as the maximal buying price $\overline{P}(S)$.

This has been a point of critique, because real world persons do not behave strictly rational in this sense and often lack the information to decide about the minimal selling and maximal buying price, see e.g. the discussions in [18, 24]. Thus an interval $[\underline{P}(S), \overline{P}(S)]$ appears to be a more accurate description of a decision maker's information. This line of argument directly leads to *imprecise probability*, probability intervals, and lower and upper probabilities [15, 91, 94].

In practical engineering applications, *elicitation* of probabilities from *experts* is the paradigm for obtaining subjective probabilities quantitatively. We refer e.g. to [24, 60, 79]. For the purpose of demonstration, we exemplify an elicitation procedure for a continuous probability distribution, as suggested in [74], from which the numerical data are taken.

Example: To determine the distribution of a random variable S, may be the monetary value of the damage associated with a certain risk, experts are first asked to determine a lower bound L, an upper bound U and a modal value M. Next, they are required to assign probabilities to the following intervals:

(a) $[L, M]$,
(b) $[L, (L+M)/2]$,
(c) $[(M+U)/2, U]$,
(d) $[L, (L+3M)/4]$,
(e) $[(3M+U)/4, U]$.



For example, the following bounds might have been given:

$$L = 165, \ M = 190, \ U = 250,$$

together with the following elicited probabilities:

(a) $[165, 190] : P = 0.500$,
(b) $[165, 177.5] : P = 0.175$,
(c) $[220, 250] : P = 0.050$,
(d) $[165, 183.75] : P = 0.330$,
(e) $[205, 250] : P = 0.250$.

This implies the following probabilities for the non-overlapping subintervals:

bin	bin center z_j	probability p_j	height of bin
$[165, 177.5]$	171.250	0.175	0.0140
$[177.5, 183.75]$	180.625	0.155	0.0248
$[183.75, 190]$	186.875	0.170	0.0272
$[190, 205]$	197.500	0.250	0.0167
$[205, 220]$	212.500	0.200	0.0133
$[220, 250]$	235.000	0.050	0.0017

The next step consists in fitting a probability distribution. For the sake of simplicity, we demonstrate the procedure with a lognormal distribution $\mathcal{L}og\mathcal{N}(\mu, \sigma^2)$ which is easily fitted using the mean and variance of the logarithms of the bin centers: $\mu \approx \sum_j \log z_j \, p_j \approx 5.2609$, $\sigma^2 \approx \sum_j (\log z_j - \mu)^2 \, p_j$; $\sigma \approx 0.0853$. Displaying the histogram together with the fitted lognormal distribution would reveal a somewhat unsatisfactory fit near the modal value. Thus the elicitation procedure incorporates a second turn, *stepping back* and trying to get a distribution which more accurately reflects the experts' assessment of the probabilities. It turns out that a three parameter lognormal distribution (with origin shifted to the right by 150 units) shows a satisfactory representation of the original histogram. The result is depicted in Figure 7; for a more detailed discussion of the actual eliciting procedure we refer to [74].

Finally, we should not fail to mention the *Bayesian* approach to assessing probability distributions. From the Bayesian viewpoint, everything is a random variable, including the parameters, say Θ, of the distributions of the original variables, say

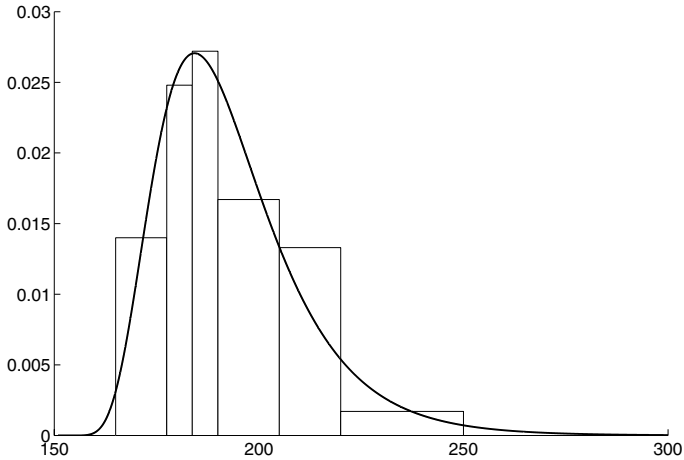


Fig. 7. Histogram and fit after step-back

X, to be assessed. The Bayesian approach has interpretations both in the frequentist as well as the subjective setting. In the civil engineering literature, it has been found useful for combining expert knowledge with sample data [19, 41, 78, 81]. The expert knowledge may be based on known frequencies or on subjective estimates and is encoded in the *a priori distribution* of the parameter θ. Sample data x (or again expert estimates obtained in situ) are then used to produce an *a posteriori distribution* of the distribution parameter θ according to Bayes' rule, loosely stated as

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} \,.$$

The data vector x may have length 1, or may consist in a large sample of size n, so that the Bayesian procedure may accommodate single estimates up to sample sizes satisfying frequentist requirements. The assessment of the distribution of the original variable X is completed by the *Bayes estimate* of its distribution parameter

$$\hat{\theta} = \int_{-\infty}^{\infty} \theta \, p(\theta|x) \, \mathrm{d}\theta \,.$$

Further powerful concepts start out from here, like odds ratios, high density regions, posterior predictive distributions, and so on. We refer e.g. to [8, 34, 42, 90] for further information on Bayesian methods.

b. Sets of probability measures: Sets of probability measures can arise both in a frequentist as well as in a subjectivist approach. In a frequentist setting, sets of probability measures arise as sets of fitted distributions: in fact, confidence regions for the distribution parameters imply that parametrized *families* of distributions are employed. As we have just seen above, sets of probability distributions are inherent in Bayesian statistics (each parameter value θ defines a distribution of the original variable). In order to avoid implausible determinations due to the choice of an a

priori distribution, families of a priori distributions have been employed, as well as fuzzy a priori distributions [90].

Further, referring to the decision theoretic foundation of probability, an interval $[\underline{P}(S), \overline{P}(S)]$ replacing the indifference price and signifying incomplete knowledge leads to interval valued probabilities, which again imply that a set of probabilities is considered.

c. Random sets: In a frequentist interpretation, this might correspond to a sample of size n of interval data for a parameter A, the probability weight being approximated by the relative frequency. The difference to a histogram then is that the focal sets A_i may overlap. In geotechnical engineering, for example, interval data arise as ranges of rock parameters associated with certain rock classes. These rock classes in turn may be the outcomes, obtained with a certain frequency from in situ measurements. This approach has been proposed and elaborated by [88, 89]. In a subjectivist interpretation, the focal sets A_i may be (possibly conflicting) estimates given by different experts and the weights might correspond to each expert's relative credibility.

Random sets have turned out to be useful for bracketing probability estimates given by different sources [52, 86] as well as for combining information of different type, due to the observation that every histogram, every interval and every fuzzy set can be viewed as a random set, without a need for artificial transformations [36, 51].

d. Fuzzy sets: In engineering and in risk analysis applications, probabilistic models have been criticized as requiring more input from the designing engineer or the decision maker than could be plausibly provided - or that would be reasonably required for a rough estimate. In particular, the requirement that probabilities have to add up to 1 causes the problem that probabilities of events change when additional events are taken into consideration. Further, probabilities have to be set up in a consistent way, e.g. satisfying the rule $p(S \cup T) = p(S) + p(T) - p(S \cap T)$ and thus do not admit incorporating conflicting information. From the current perspective, fuzzy set theory appears to provide a resolution of these difficulties in as much as it admits much more freedom in modelling. Further, fuzzy sets may be used to model vagueness and ambiguity.

The notion of *possibility* provides an interpretation for a fuzzy set and an operational method of constructing it, as we wish to argue here. We imagine that a fuzzy set describing the uncertainty of a real valued parameter has to be designed. The procedure proposed below is based on a scale $0 \leq \alpha \leq 1$. The linguistic meaning of the α-values is specified verbally by the designing engineer or the decision maker in advance, but then remains fixed during the whole modelling process (for example, $\alpha = 1$ signifies the standard value of the parameter, $\alpha = 2/3, 1/3, 0$ might indicate high, medium, and low *degree of possibility*).

One would start by specifying the standard value a_S of the parameter, in engineering terminology often referred to as the *deterministic approximation* and assign degree of possibility $\alpha = 1$ to it. Next, possible deviations of the parameter from the standard value are taken into account, corresponding to decreasing degree of possibility, until the minimal and maximal values, which are assumed with very small degree of possibility, are reached at level $\alpha = 0$, see Figure 8.

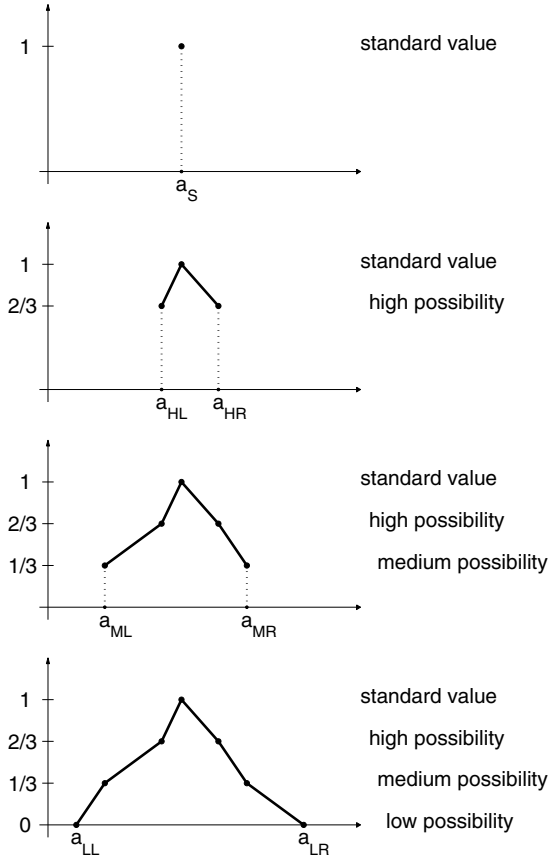


Fig. 8. Construction of a fuzzy set

To counter the criticism that the notion of *possibility* has no operational meaning, one might argue that it is intended to encode subjective risk assessments. Thereby, it is envisaged that the risks leading to parameter fluctuations at the corresponding possibility level are established in an analysis of scenarios. We refer to [55] where the procedure has been demonstrated in the case of network analysis of a civil engineering project. Similar procedures have been proposed by [4]. There is also a normative approach which uses pre-shaped, parametrized membership functions as well as the suggestion to use elicitation procedures; see [5, 20, 79, 99, 100] for further details.

An alternative way of establishing the semantics of *possibility* is to start from the notion of *potential surprise* and define possibility as its complementary notion [84] or as a transformed quantity [65] thereof. For a decision-theoretic foundation, see [21], for possibility as a transformation of probability, see [22].

Finally, the *contour function* of a random set is the function $a \to \overline{P}(\{a\}) = \eta(\{a\})$, assigning each singleton a its plausibility. It is simply obtained by adding the probability weights $m(A_i)$ of those focal sets A_i to which a belongs. Figure 9 shows a random set and the resulting contour function where weights have been chosen as $m(A_1) = 1/2, m(A_2) = 1/3, m(A_3) = 1/6$. The contour function may be viewed as the membership function defining a fuzzy set and thus in turn provides a further interpretation for fuzzy sets. In the case of nested focal sets, possibility equals plausibility in this interpretation.

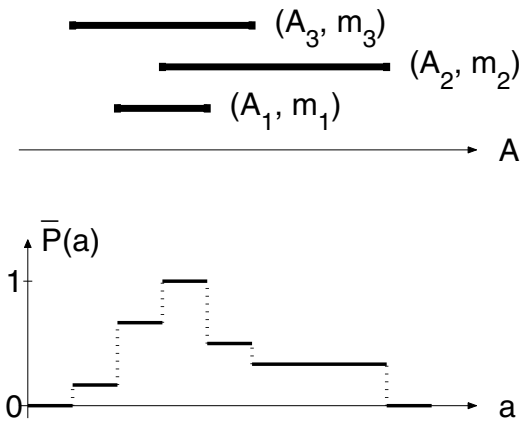


Fig. 9. Contour function of a random set

4 Axiomatics

Referring to probability theory, it has been emphasized by [77, Sect. 71] that a formal mathematical system of axioms and postulates is required in order to approach the problem of relations between the different interpretations of probability. All the more so, this applies to the problem of comparing the different theories of uncertainty. By now, all these theories can be based on appropriate axioms, the most well known being the Kolmogorov [50] axioms of probability (though criticized by [77]) and the Choquet [14] axioms of capacities.

We do not wish to give a detailed exposition of the axioms here, but just highlight a few of them, showing that they lead to different combination rules in the various models of uncertainty. This may serve as a means of distinguishing the models as well as a guiding criterion to decide which model should be selected for what purpose.

The axioms fix the algebraic properties of the corresponding set functions. For example, probability measures p define additive set functions, that is,

$$p(S \cup T) = p(S) + p(T) - p(S \cap T)$$

for sets S, T. This is not true of possibility measures π, which in turn satisfy

$$\pi(S \cup T) = \max\{\pi(S), \pi(T)\}\,.$$

Both probability measures and possibility measures are special cases of plausibility measures η which enjoy the more general property

$$\eta(S \cup T) \leq \eta(S) + \eta(T) - \eta(S \cap T)\,.$$

All these set functions are contained in the largest class of monotone set functions μ, characterized by the property

$$\mu(S \cup T) \geq \max\{\mu(S), \mu(T)\}\,.$$

Suitably completing these algebraic properties to full systems of axioms, it is actually possible to *characterize* all these type of measures (and many more). The following hierarchical relations obtain:

$$\begin{array}{ll} \text{probability} & \rightarrow \text{plausibility} \\ \text{possibility} & \rightarrow \text{plausibility} \\ \text{plausibility} & \rightarrow \text{monotonicity.} \end{array}$$

This means that every probability measure is a plausibility measure, and so on (thus plausibility theory is more general than probability theory). Probability measures and possibility measures are in no implication relation in either direction.

Alternatively, if one starts from the notion of random sets, probability would correspond to singletons as focal sets, possibility to nested focal sets, and plausibility to arbitrary focal sets. These short indications may suffice; a discussion of the complete axiomatic systems as well as much more elaborated hierarchies can be found in [17, 49, 92, 93].

5 Numerics

Most engineering models are input-output systems. Given certain input values (model parameters, initial conditions, dimensions, etc.) the model produces output values (displacements, stresses, costs, etc.). In other words, the model is a function F that assigns to the input data A certain output values $F(A)$. Both A and $F(A)$ may be multidimensional (for simplicity, we shall consider $F(A)$ as one-dimensional in the sequel - corresponding to a a single component of a multidimensional output). Often, the function F is a computer code, in which case the output is obtained as a numerical approximation. If the input data consist of a single, *deterministic* data set, then the model produces a uniquely determined output. If the input data fluctuate, so does the output. If the fluctuation of the input is described by one or the other theories of uncertainty discussed so far, the fluctuation of the output should be captured on the same terms. This is the issue of this section: how is data uncertainty propagated through an input-output system?

a. Deterministic values: If a is an expert estimate of some input parameter A, the output is just the value $F(a)$ of the function F at a. In this framework, the effects of the uncertainty of the input parameters can still be assessed by performing a *sensitivity analysis*. In its classical form, sensitivity analysis means the study of the derivatives of the function F at the fixed value a of interest, that is, the linear

approximation to the change in output when the input parameter a is changed to a nearby value b:

$$F(b) \approx F(a) + \frac{\mathrm{d}F}{\mathrm{d}a}(a)(b-a)$$

or in case the parameter $a = (a_1, \ldots, a_n)$ is n-dimensional,

$$F(b) \approx F(a) + \sum_{i=1}^{n} \frac{\partial F}{\partial a_i}(a)(b_i - a_i)$$

where the approximation error is of order $|b-a|^2$. The sign and size of the partial derivative $\frac{\partial F}{\partial a_i}(a)$ is an indicator of the influence of the i-th component of the parameter a on the output. For more details on classical sensitivity analysis and its numerical aspects we refer to [75].

In case $A = (A_1, \ldots, A_n)$ is a vector of n independent random variables and a is its expectation value, the variance σ^2 of $F(A)$ can be approximated by

$$\sigma^2 \approx \sum_{i=1}^{n} \left(\frac{\partial F}{\partial a_i}(a)\right)^2 \sigma_i^2$$

where σ_i^2 is the variance of A_i. Then $\frac{\partial F}{\partial a_i}(a)\sigma_i/\sigma$ is an indicator of the fractional contribution of A_i to the total standard deviation σ, and thus another indicator of the influence of the uncertainty of the i-th component on the output. For more details, we refer to [43]. When the input-output function F is not differentiable, one may still use fuzzy sets to assess the partial influence of the components of the input parameters (see below) or study the change of variance when all but one component are kept fixed (see e.g. [8]).

b. Intervals: If A is an interval, the functional evaluation $F(A)$ results in a set of values (an interval, if F is continuous and one-dimensional). In general, both A and $F(A)$ could be sets of arbitrary geometry. In interval arithmetic, one would bound these sets by the smallest multidimensional intervals (boxes) that contain them (see [64]). In any case this approach represents the full range of the possible output values without further fine structure.

c. Probability measures and random sets: Given a single probability measure and a (measurable [40]) map F, the output probabilities are determined through the induced image measure, that is, $P(F(A) \in S) = P(A \in F^{-1}(S))$. Though the distribution of the random variable $F(A)$ can be computed by this prescription in principle, this is practically impossible as soon as F attains a rather moderate complexity. *The* numerical method for approximating the output distribution by means of an artificially created sample is *Monte Carlo simulation*. We refer e.g. to [8] as well as the survey articles [43, 82].

In case the uncertainty of the input is modelled by a set $\mathcal{M}$ of probability measures, the map F induces a set of probability measures as well, namely the collection of all image measures, obtained from $\mathcal{M}$ under this map. The computation of lower and upper probabilities turns into an optimization problem.

If the input is described by a random set with focal sets $A_i, i = 1, \ldots, n$ and probability weights $m(A_i)$, the output is again a random set which consists of the focal sets $F(A_i), i = 1, \ldots, n$, supplied with the original weights $m(A_i)$. In case the sets $F(A_i)$ are intervals, their boundaries can be found by optimization. The determination of lower and upper probabilities is then a combinatorial task involving

the formulas for belief and plausibility. For further details on computational methods with random sets, see [30, 32].

d. Fuzzy sets: The question of propagating the uncertainty of a fuzzy input A through a function F needs some more explication. If the input is a fuzzy set with membership function $\pi_A(a)$, say, the output will also be a fuzzy set, described by a membership function $\pi_{F(A)}(b)$. According to the *Zadeh extension principle* [97], it is given by

$$\pi_{F(A)}(b) = \sup \left\{ \pi_A(a) : F(a) = b \right\} .$$

In case the input consists of a vector of parameters $A = (A_1, \ldots, A_m)$, the extension principle takes the form

$$\pi_{F(A)}(b) = \sup \left\{ \min \left(\pi_{A_1}(a_1), \ldots, \pi_{A_m}(a_m) \right) : F(a_1, \ldots, a_m) = b \right\} .$$

Note that this comes naturally from the possibility-theoretic interpretation: to compute the degree of possibility that $F(A)$ takes the value b, one has to look for all combinations $a_1, \ldots, a_m$ producing the value b; each single combination gets the smallest possibility among its participants, while b gets the supremum of all possibility degrees that can be obtained in this way. In case F is continuous and the α-level sets of $A_1, \ldots, A_m$ are compact ($0 < \alpha \leq 1$), this corresponds exactly to computing the range of the function F on each α-level set,

$$F(A)^\alpha = F(A^\alpha) , \quad \text{respectively,} \quad F(A_1, \ldots, A_m)^\alpha = F(A_1^\alpha, \ldots, A_m^\alpha) .$$

When the A_j^α are intervals, in addition, the set $F(A_1, \ldots, A_m)^\alpha$ is an interval as well. The computation of its boundaries is then a task of global optimization: finding the minimum and maximum value of F on the set A^α. In any case, the procedure is consistent: if the input data are structured as stacked intervals, so is the output.

Fuzzy sets can be used to assess the partial influence of each component A_i of the input. One may keep all components $A_j, j \neq i$, at a fixed value and model A_i by means of a triangular shaped fuzzy number, say. Then the distortion and spread of the output (as a fuzzy number), compared with the resulting fuzzy numbers when other components are allowed to vary, give a good impression of how (smaller and bigger) changes of A_i effect the output.

6 The multivariate case

In principle, the multivariate case, that is, the case when the input A has several components $A_1, \ldots, A_n$, has been covered by what has been said above - all applies to multidimensional intervals, random sets, fuzzy sets and multivariate distribution functions. However, the issue is how to model mutual dependence, correlation, interaction, influence of the different components. In addition, the task remains to model infinitely many components, as arising in spatial fields or temporal processes, when the parameters are functions of space and/or time.

There is a rather well established notion of independence in probability theory: two random variables are *independent* when their joint distribution function is the product of the individual (marginal) distribution functions. The situation is also clear in interval analysis: two parameters taking interval values are *non-interactive* when their joint behavior is described by the product of the two intervals (a rectangle), and *interactive* when their joint range is a proper subset of the product interval.

This idea extends to fuzzy sets: Two fuzzy sets are non-interactive, when all their α-level sets are rectangles.

Switching to random sets, the situation becomes more complicated. The concept of independence splits into a number of different concepts, all coinciding when the focal sets are singletons (i.e. when the random set is actually a probability). *Random set independence* is characterized by two properties: the joint focal sets are product intervals, and the joint weights are the products of the corresponding individual (marginal) weights. *Strong independence* is obtained when the underlying set of joint probability measures is required to consist of product measures only. Going deeper into the structure of the underlying set of probability measures, many more notions of independence can be considered, one of the more prominent being *epistemic independence*. For further details on this, see [30, 32].

What concerns spatial fields or temporal processes, there is the fully developed probabilistic theory of stochastic processes, see e.g. [58]. Processes involving random sets have been considered in [38]. A fully intrinsic theory of processes of fuzzy sets (with interactivity replacing dependence) has not yet been developed. A useful substitute has been the notion of a fuzzy random process, either obtained as a stochastic process with fuzzy parameters or as a process of fuzzy numbers whose α-level sets are made up by *imbedded* stochastic processes. For more references on both approaches, see [10, 13, 56], for engineering applications, see [62, 68].

For the sake of introduction, these indications may suffice. Research on stochastic processes constitutes a very active branch of probability theory and stochastic analysis. In the other approaches to uncertainty, the theory of processes has been developed to a lesser degree and constitutes an important topic for future research.

References

[1] H. Adeli, B. Möller (Eds.). *Special Issue: Fuzzy Modeling.* Vol. **14**(2) of *Computer Aided Civil and Infrastructure Engineering*, 1999.

[2] G. Auvinet, J. L. González. Three-dimensional reliability analysis of earth slopes. Computers and Geotechnics **26** (2000), 247–261.

[3] B. M. Ayyub (Ed.). *Uncertainty Modeling and Analysis in Civil Engineering.* CRC Press, Boca Raton 1998.

[4] H. Bandemer. *Ratschläge zum mathematischen Umgang mit Ungewißheit - Reasonable Computing.* Teubner, Leipzig 1997.

[5] H. Bandemer, S. Gottwald. *Einführung in Fuzzy-Methoden.* 4. Aufl., Akademie Verlag, Berlin 1993.

[6] H. Bandemer, W. Näther. *Fuzzy Data Analysis.* Kluwer Acad. Publ., Dordrecht 1992.

[7] A. Bárdossy, L. Duckstein. *Fuzzy Rule-Based Modeling with Applications to Geophysical, Biological and Engineering Systems.* CRC Press, Boca Raton 1995.

[8] T. Bedford, R. Cooke. *Probabilistic Risk Analysis.* Cambridge University Press, Cambridge 2001.

[9] A. Bernardini. Qualitative and quantitative measures in seismic damage assessment and forecasting of masonry buildings. In: A, Bernardini (Ed.). *Seismic Damage to Masonry Buildings.* A. A. Balkema, Rotterdam 1999, 169–177.

[10] C. Bertoluzza, M. Á. Gil, D. A. Ralescu (Eds.). *Statistical Modeling, Analysis and Management of Fuzzy Data.* Physica-Verlag, Heidelberg 2002.

[11] D. Blockley, I. Ferraris. Managing the quality of a project. In: *Seguridad en Ingeniería*, Academia Nacional de Ingeniería, Editorial Dunken, Buenos Aires 2000, 69–89.

[12] V. V. Bolotin. *Statistical Methods in Structural Mechanics.* Holden-Day, Inc., San Francisco, 1969.

[13] J. J. Buckley, E. Eslami, T. Feuring. *Fuzzy Mathematics in Economics and Engineering.* Physica-Verlag, Heidelberg 2002.

[14] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier* **5** (1953/54), 131–295.

[15] G. De Cooman, T. Fine, T. Seidenfeld (Eds.). *ISIPTA'01, Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications.* Shaker Publishing BV, Maastricht, 2001.

[16] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics* 38:325–339, 1967.

[17] D. Denneberg. *Non-additive Measure and Integral.* Kluwer Acad. Publ., Dordrecht 1994.

[18] R. Diedrich. *Entscheidungen bei Ungewißheit.* Physica-Verlag, Heidelberg 1999.

[19] O. Ditlevsen, N. J. Tarp-Johansen, H. Denver. Bayesian soil assessments combining prior with posterior censored samples. *Computers and Geotechnics*, **26** (2000), 187–198.

[20] D. Dubois, H. Prade. *Possibility Theory.* Plenum Press, New York, 1988.

[21] D. Dubois, H. Prade, R. Sabbadin. Decision-theoretic foundations of qualitative possibility theory. *European J. Operations Res.* **128** (2001), 459–478.

[22] D. Dubois, H. Prade, S. Sandri. On possibility/probability transformations. In: R. Lowen, M. Roubens (Eds.). *Fuzzy Logic: State of the Art.* Kluwer Academic Publishers, Dordrecht 1993, 103–112.

[23] H. H. Einstein. Quantifying uncertain engineering geologic information. *Felsbau*, **19**(5) (2001), 72 - 84.

[24] F. Eisenführ, M. Weber. *Rationales Entscheiden.* 3rd Ed., Springer-Verlag, Berlin 1999.

[25] I. Elishakoff (Ed.). *Whys and Hows in Uncertainty Modelling: Probability, Fuzziness and Anti-Optimization.* Springer-Verlag, Wien 1999.

[26] I. Elishakoff. What may go wrong with probabilistic methods? In I. Elishakoff (Ed.). *Whys and Hows in Uncertainty Modelling: Probability, Fuzziness and Anti-Optimization* Springer-Verlag, Wien 1999, 265–283.

[27] W. Fellin. Assessment of characteristic soil parameters and its implication in geotechnical design. *In this volume.*

[28] S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, K. Sentz. Constructing probability boxes and Dempster-Shafer structures. SAND REPORT SAND2002-4015, Sandia National Laboratories, Albuquerque, 2003.

[29] Th. Fetz. Sets of joint probability measures generated by weighted marginal focal sets. In: G. De Cooman, T. Fine, T. Seidenfeld (Eds.). *ISIPTA'01, Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications.* Shaker Publishing BV, Maastricht, 2001, 171–178.

[30] Th. Fetz. Multi-parameter models: rules and computational methods for combining uncertainties. *In this volume.*

[31] Th. Fetz, J. Jäger, D. Köll, G. Krenn, H. Lessmann, M. Oberguggenberger, R. Stark. Fuzzy models in geotechnical engineering and construction management. *In this volume.*

[32] Th. Fetz, M. Oberguggenberger. Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering and Systems Safety,* **85** (2004), 75–87.

[33] T. L. Fine. *Theories of Probability. An Examination of Foundations.* Academic Press, New York 1973.

[34] A. B. Gelman, J. S. Carlin, H. S. Stern, D. B. Rubin. *Bayesian Data Analysis.* Chapman & Hall/CRC, Boca Raton 1997.

[35] A. Gheorghe, J.-P. Krause, R. Mock. *Integration von Fuzzy-Logic für eine regionale Umweltrisiko-Analyse.* Hochschulverlag an der ETH, Zürich 1996.

[36] I. R. Goodman, H. T. Nguyen. Fuzziness and randomness. In: C. Bertoluzza, M. Á. Gil, D. A. Ralescu (Eds.), *Statistical Modeling Analysis and Management of Fuzzy Data.* Physica-Verlag, Heidelberg 2002, 3–21.

[37] J. Goutsias, R. P. S. Mahler, H. T. Nguyen (Eds.). *Random Sets. Theory and Applications.* Springer-Verlag, New York 1997.

[38] J. Hall, J. Lawry. Imprecise probabilities of engineering system failure from random and fuzzy set reliability analysis. In: G. De Cooman, T. Fine, T. Seidenfeld (Eds.). *ISIPTA'01, Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications.* Shaker Publishing BV, Maastricht, 2001, 195–204.

[39] J. Hall, E. Rubio, M. Anderson. Random sets of probability measures in slope hydrology and stability analysis. *Zeitschrift f. Angewandte Mathematik u. Mechanik*, to appear.

[40] P. R. Halmos. *Measure Theory.* Van Nostrand, New York 1950.

[41] J. Hanisch, W. Struck. Charakteristischer Wert einer Boden- oder Materialeigenschaft aus Stichprobenergebnissen und zusätzlicher Information. *Bautechnik* **10** (1985), 338–348.

[42] J. A. Hartigan. *Bayes Theory.* Springer-Verlag, New York 1983.

[43] J. C. Helton, F. J. Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. SAND REPORT SAND2001-0417, Sandia National Laboratories, Albuquerque, 2002.

[44] I. Herle. Difficulties related to numerical predictions of deformations. *In this volume.*

[45] H. Hirsch. Interview by H. Hochner. In: *ZIB3, Austrian Television*, May 5, 2001.

[46] P. J. Huber. Robust Statistics. John Wiley & Sons, New York 1981.

[47] A. Kandel. Fuzzy Mathematical Techniques with Applications. Addison-Wesley, Reading, 1986.

[48] O. Klingmüller, U. Bourgund. *Sicherheit und Risiko im Konstruktiven Ingenieurbau.* Verlag Vieweg, Braunschweig, 1992.

[49] G. J. Klir, M. J. Wiermann. *Uncertainty-Based Information. Elements of Generalized Information Theory.* Physica-Verlag, Heidelberg 1998.

[50] A. Kolmogoroff. *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Springer-Verlag, Berlin 1933.

[51] V. Kreinovich. Random sets unify, explain, and aid known uncertainty methods in expert systems. In: J. Goutsias, R. P. S. Mahler, H. T. Nguyen (Eds.). *Random Sets. Theory and Applications.* Springer-Verlag, New York 1997.

[52] E. Kriegler, H. Held. Global mean temperature projections for the 21st century using random sets. In: J.-M. Bernard, T. Seidenfeld, M. Zaffalon (Eds.). *ISIPTA'03. Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications.* Proceedings in Informatics 18, Carlton Scientific 2003.

[53] H. Lehar, G. Niederwanger, G. Hofstetter. FE ultimate load analyses of pile-supported pipelines - tackling uncertainty in a real design problem. *In this volume.*

[54] K. Lesny, A. Kisse, W. Richwien: Proof of foundation stability using a consistent failure model. In: R. Pöttler, H. Klapperich, H. F. Schweiger (Eds.). *Probabilistics in GeoTechnics: Technical and Economic Risk Estimation.* United Engineering Foundation, Verlag Glückauf, Essen, 2002, 95–103.

[55] H. Lessmann, J. Mühlögger, M. Oberguggenberger. Netzplantechnik mit unscharfen Methoden. *Bauingenieur* **69** (1994) 469–478.

[56] R.-J. Li, E. S. Lee. Analysis of fuzzy queues. *Computers Math. Appl.*, **17** (1989) 1143–1147.

[57] B. S. L. P. de Lima, N. F. F. Ebecken. A comparison of models for uncertainty analysis by the finite element method. *Finite Elements in Analysis and Design*, **34** (2000), 211-232.

[58] M. Loève. *Probability Theory II.* 4th Ed., Springer-Verlag, New York 1978.

[59] C. R. Manski. *Partial Identification of Probability Distributions.* Springer-Verlag, New York 2003.

[60] M. A. Meyer, J. M. Booker. *Eliciting and Analyzing Expert Judgement.* SIAM, Philadelphia PA, 2001.

[61] B. Möller. Fuzzy-Modellierung in der Baustatik. *Bauingenieur*, **72** (1997) 75–84.

[62] B. Möller. Fuzzy randomness - a contribution to imprecise probability. *Zeitschrift f. Angewandte Mathematik u. Mechanik*, to appear.

[63] P.-A. Monney. *A Mathematical Theory for Statistical Evidence.* Physica-Verlag, Heidelberg 2003.

[64] A. Neumaier. *Interval Methods for Systems of Equations.* Cambridge University Press, Cambridge 1990.

[65] A. Neumaier. Fuzzy modeling in terms of surprise. *Fuzzy Sets and Systems* **135**, 21–38 (2003).

[66] A. Neumaier. Clouds, fuzzy sets and probability intervals. *Reliable Computing,* **10** (2004), 249–272.

[67] H. T. Nguyen, T. Wang. Belief functions and random sets. In: J. Goutsias, R. P. S. Mahler, H. T. Nguyen (Eds.). *Random Sets. Theory and Applications.* Springer-Verlag, New York 1997.

[68] M. Oberguggenberger. Queueing models with fuzzy data in construction management. *In this volume.*

[69] M. Oberguggenberger, W. Fellin. The fuzziness and sensitivity of failure probabilities. *In this volume.*

[70] M. Oberguggenberger, F. Russo. Fuzzy, probabilistic and stochastic modelling of an elastically bedded beam. *In this volume.*

[71] M. Oberguggenberger, G. I. Schuëller, K. Marti (Eds.). *Special Issue: Fuzzy Sets, Imprecise Probability and Stochastics in Engineering. Zeitschrift f. Angewandte Mathematik u. Mechanik*, to appear.

[72] W. L. Oberkampf, J. C. Helton. Investigation of Evidence Theory for Engineering Applications. American Institute of Aeronautics and Astronautics, Reston VA, AIAA 2002 - 1569, Non-Deterministic Approaches Forum, Seattle, WA (2002).

[73] W. L. Oberkampf, J. C. Helton, K. Sentz. Mathematical Representation of Uncertainty. American Institute of Aeronautics and Astronautics, Reston VA, AIAA 2001 - 1645, Non-Deterministic Approaches Forum, Seattle, WA (2001).

[74] A. O'Hagan. Eliciting expert beliefs in substantial practical applications. *The Statistician* **47** (1998), 21–35.

[75] A. Ostermann. Sensitivity analysis. *In this volume.*

[76] K. Popper. The propensity interpretation of the calculus of probability, and the quantum theory. In: S. Körner ((Ed.), *Observation and Interpretation*, London 1957.

[77] K. Popper. *Logik der Forschung.* 10. Aufl., J. C. B. Mohr (Paul Siebeck), Tübingen 1994.

[78] R. Rackwitz. Reviewing probabilistic soils modelling. *Computers and Geotechnics* **26** (2000), 199–223

[79] T. J. Ross, J. M. Booker, W. J. Parkinson (Eds.). *Fuzzy Logic and Probability Applications - Bridging the Gap.* SIAM, Philadelphia PA, ASA, Alexandria VA, 2002.

[80] L. J. Savage. *The Foundations of Statistics.* Wiley, New York 1954.

[81] H. R. Schneider. Definition and determination of characteristic soil properties. In: *Proceedings of the Fourteenth International Conference on Soil Mechanics and Foundation Engineering, Hamburg.* A. A. Balkema, Rotterdam 1997, 2271–2274.

[82] G. I. Schuëller, H. J. Pradlwarter, P. S. Koutsourelakis. A critical appraisal of reliability estimation procedures for high dimensions. *Probabilistic Engineering Mechanics,* to appear.

[83] R. Seising (Ed.). *Fuzzy Theorie und Stochastik.* Vieweg, Braunschweig 1999.

[84] G. L. S. Shackle. *Decision - order and time in human affairs.* Cambridge University Press, Cambridge 1969.

[85] G. Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, Princeton, NJ 1976.

[86] F. Tonon. On the use of random set theory to bracket the results of Monte Carlo simulations. *Reliable Computing,* **10** (2004), 107–137.

[87] F. Tonon, A. Bernardini, I. Elishakoff. Hybrid Analysis of uncertainty: probability, fuzziness and anti-optimization *Chaos, Solitons and Fractals* **12** (2001), 1403–1414.

[88] F. Tonon, A. Bernardini, A. Mammino. Determination of parameters range in rock engineering by means of random set theory. *Reliabilty Engineering & System Safety* **70**, 241–261 (2000).

[89] F. Tonon, A. Bernardini, A. Mammino. Reliability analysis of rock mass response by means of random set theory. *Reliabilty Engineering & System Safety* **70**, 263–282 (2000).

[90] R. Viertl. *Einführung in die Stochastik.* Mit Elementen der Bayes-Statistik und der Analyse unscharfer Information. 3. Aufl., Springer-Verlag, Wien 2003.

[91] P. Walley. *Statistical Reasoning with Imprecise Probabilities.* Chapman & Hall, London 1991.

[92] P. Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning* **24** (2000), 125–148.
[93] Z. Wang, G. J. Klir. *Fuzzy Measure Theory.* Plenum Press, New York, 1992.
[94] K. Weichselberger. *Elementare Grundbegriffe einer allgemeinen Wahrscheinlichkeitsrechnung I.* Physica-Verlag, Heidelberg 2001.
[95] O. Wolkenhauer. *Data Engineering. Fuzzy Mathematics in Systems Theory and Data Analysis.* John Wiley & Sons, New York 2001.
[96] L. A. Zadeh. Fuzzy sets. *Information and Control*, **8** (1965) 338–353.
[97] L. A. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning, Part 1 - 3. *Information Sci.* **8** (1975), 199–250; 301–357; **9** (1975), 43–80.
[98] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* **1** (1978), 3–28.
[99] H.-J. Zimmermann. *Fuzzy Sets, Decision Making, and Expert Systems.* Kluwer Academic Publishers, Boston 1987.
[100] H.-J. Zimmermann. *Fuzzy Set Theory – and Its Applications.* 2nd Ed., Kluwer Academic Publishers, Boston 1991.

Multi-parameter models: rules and computational methods for combining uncertainties

Thomas Fetz

Institut für Technische Mathematik, Geometrie und Bauinformatik, Universität Innsbruck

Summary. This paper is devoted to the construction of sets of joint probability measures for the case that the marginal sets of probability measures are generated by random sets. Different conditions on the choice of the weights of the joint focal sets and on the probability measures on these sets lead to different types of independence such as strong independence, random set independence, fuzzy set independence and unknown interaction. As an application the upper probabilities of failure of a beam bedded on two springs are computed.

1 Introduction

Precise probability theory alone often does not suffice for modelling the uncertainties arising in civil engineering problems such as the reliability analysis of structures and much more in soil mechanics. The goal is to have practicable measures for the risk of failure in the case where the material properties are not or only partly given by precise values or probability measures. One should also have the possibility to assess subjective knowledge and expert estimates.

It is usually easy for the planning engineer to provide such information by using random sets (weighted focal sets) to model the fluctuations of the parameters involved. In most cases intervals are used for the focal sets which has the advantage that the computations can be performed by interval analysis. This leads, if the intervals are nested, to fuzzy numbers and possibility measures or, if not nested, to plausibility measures and evidence theory. Fuzzy numbers or possibility measures [18, 19] have been applied in civil engineering problems in [4, 7, 8, 12, 13, 14] and plausibility measures [18] in [10, 15, 16]. Using the more general plausibility measures has the advantage that one can mix e.g. fuzzy numbers with histograms or probability measures directly without transforming the probability measures into fuzzy numbers and neglecting information. For the engineer it is often more practicable to use the semantics of probability and interpret these measures as upper probabilities as done in

[5, 10, 15, 16] rather than dealing with the meaning of different new measures of risk.

In general the uncertainty of the parameters has to be propagated through a numerical model. Therefore, the joint uncertainty of the parameters is needed, but in many cases the random sets are given for each uncertain parameter separately. So first we have to model the joint uncertainty by means of the given marginal random sets. As in classical probability theory this procedure is not uniquely determined; even the notion of independence is not uniquely defined.

If the marginal focal sets are nested, we can generate a joint possibility measure and a joint plausibility measure. Unfortunately these two measures are not the same in general, which leads to ambiguities in interpreting both measures as upper probabilities. In this article we model the uncertainty by random sets (and by their special cases fuzzy sets and histograms) and we assume that the uncertain parameters are "independent".

The plan is as follows: First we give definitions of random sets, the corresponding measures of plausibility and possibility and the joint measures generated by marginal random sets. As a civil engineering example a beam bedded on two springs with uncertain spring constants given by random sets, fuzzy sets or histograms is considered. Then we introduce the different notions of independence and construct sets of joint probability measures by means of random sets. We list different conditions on choosing the weights of the joint focal sets and the probability measures on these sets. Depending on these conditions we get different sets of joint probability measures and different types of independence, respectively. We show that some of these cases lead to types of independence as described in [1] such as strong independence, random set independence and unknown interaction. For these types of independence the upper probabilities of failure of the beam will be computed.

2 Random sets and sets of probability measures

In this article we consider two uncertain values or parameters λ_1 and λ_2. The possible realizations ω_k of an uncertain parameter λ_k belong to a set Ω_k. The uncertainty of a parameter λ_k is always modelled by a random set.

2.1 Random Sets

A random set $(\mathcal{F}_k, m_k)$ consists of a finite class $\mathcal{F}_k = \{F_k^1, \ldots, F_k^{n_k}\}$ of focal sets and of a weight function

$$m_k : \mathcal{F}_k \longrightarrow [0,1] : F \mapsto m_k(F)$$

with $\sum_{i=1}^{|\mathcal{F}_k|} m_k(F_k^i) = 1$ where $|\mathcal{F}_k|$ is the number of focal sets. Then the plausibility measure of a set $A_k \in \Omega_k$ is defined by

$$\mathrm{Pl}_k(A_k) = \sum_{F_k^i \cap A_k \neq \varnothing} m_k(F_k^i) \tag{1}$$

and the belief measure by

$$\mathrm{Bel}_k(A_k) = \sum_{F_k^i \subseteq A_k} m_k(F_k^i). \tag{2}$$

Three important special cases:

1. The focal sets are singletons: Then the plausibility measure Pl_k is a classical probability measure defined by the weights m_k.
2. The random set is consonant: That means that the focal sets are nested, e.g. $F_k^1 \supseteq F_k^2 \supseteq \cdots \supseteq F_k^{|\mathcal{F}_k|}$. Then the contour $\omega_k \mapsto \mathrm{Pl}_k(\{\omega_k\})$ of the random set is the membership function $\mu_{\tilde{A}_k}$ of the corresponding fuzzy set $\tilde{A}_k$. On the other hand a fuzzy set can be represented by a consonant random set. If in addition the focal sets are intervals in $\mathbb{R}$, we get a fuzzy number.
3. The focal sets F_k^i are disjoint intervals $[a_i, b_i)$ in $\mathbb{R}$: Then the random set is a histogram with bins $[a_i, b_i)$ and heights $m_k([a_i, b_i))/(b_i - a_i)$.

For more information and references concerning random sets and fuzzy sets we refer to the article "Mathematics of uncertainty: models, methods and interpretations" in this volume.

2.2 Sets of probability measures generated by random sets

If the uncertainty of a parameter λ_k is modelled by a set of probability measures $\mathcal{M}_k$, then the upper probability $\overline{P}_k$ and the lower probability $\underline{P}_k$ are defined in the following way:

$$\overline{P}_k(A_k) = \sup\{P_k(A_k) : \ P_k \in \mathcal{M}_k\},$$
$$\underline{P}_k(A_k) = \inf\{P_k(A_k) : \ P_k \in \mathcal{M}_k\}.$$

Here we want to generate the set $\mathcal{M}_k$ of probability measures by a random set $(\mathcal{F}_k, m_k)$. Let $\mathcal{M}_k^i = \{P_k^i : P_k^i(F_k^i) = 1\}$ be the set of probability measures P_k^i "on" the corresponding focal set F_k^i. Then

$$\mathcal{M}_k = \left\{ P_k = \sum_{i=1}^{|\mathcal{F}_k|} m_k(F_k^i) P_k^i : \ P_k^i \in \mathcal{M}_k^i \right\} = \sum_{i=1}^{|\mathcal{F}_k|} m_k(F_k^i) \mathcal{M}_k^i$$

is the set of probability measures generated by $(\mathcal{F}_k, m_k)$.

The upper probability $\overline{P}_k(A_k)$ is obtained by solving the following optimization problem (c.f. [3]):

$$\overline{P}_k(A_k) = \max\{P_k(A_k) : P_k \in \mathcal{M}_k\} = \sum_{i=1}^{|\mathcal{F}_k|} m_k(F_k^i) P_k^{i*}(A_k)$$

with certain $P_k^{i*} \in \mathcal{M}_k^i$. Such an optimal P_k^{i*} can be chosen in the following way: $P_k^{i*} = \delta_{\omega_k^{i*}}$ with

$$\omega_k^{i*} \in \begin{cases} F_k^i \cap A_k, & \text{if } F_k^i \cap A_k \neq \varnothing, \\ F_k^i \text{ arbitrary}, & \text{if } F_k^i \cap A_k = \varnothing \end{cases}$$

where δ_{ω_k} is the Dirac measure at $\omega_k \in \Omega_k$. Then $P_k^{i*}(A_k) = 1$ for $F_k^i \cap A_k \neq \varnothing$ and 0 otherwise which leads to the same result as in formula (1). The lower probability $\underline{P}_k(A_k)$ is obtained by:

$$\underline{P}_k(A_k) = \min\{P_k(A_k) : P_k \in \mathcal{M}_k\} = \sum_{i=1}^{|\mathcal{F}_k|} m_k(F_k^i) \delta_{\omega_k^{i**}}(A_k)$$

with

$$\omega_k^{i**} \in \begin{cases} F_k^i \setminus A_k, & \text{if } F_k^i \not\subseteq A_k, \\ F_k^i \text{ arbitrary}, & \text{otherwise.} \end{cases}$$

Then $\delta_{\omega_k^{i**}}(A_k) = 1$ for $F_k^i \subseteq A_k$ and 0 otherwise and we get the same result as in formula (2).

If the random set is consonant, e.g. $F_k^1 \supseteq F_k^2 \supseteq \cdots \supseteq F_k^{|\mathcal{F}_k|}$, then the plausibility measure is also a possibility measure Pos_k with possibility density $\mu_k(\omega_k) = \mathrm{Pos}_k(\{\omega_k\})$ which is the membership function of the corresponding fuzzy number or the contour of the random set.

Thus we have for arbitrary random sets:

$$\overline{P}_k = \mathrm{Pl}_k, \quad \underline{P}_k = \mathrm{Bel}_k$$

and for consonant random sets:

$$\overline{P}_k = \mathrm{Pl}_k = \mathrm{Pos}_k, \quad \underline{P}_k = \mathrm{Bel}_k = \mathrm{Nec}_k$$

where Nec_k is the necessity measure ($\mathrm{Nec}_k(A_k) = 1 - \mathrm{Pos}_k(\Omega_k \setminus A_k)$).

2.3 Combining random sets

Let two random sets $(\mathcal{F}_1, m_1)$ and $(\mathcal{F}_2, m_2)$ be given. Then the joint random set $(\mathcal{F}, m)$ is defined by joint focal sets

$$F_1^i \times F_2^j \in \mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$$

and joint weights

$$m(F_1^i \times F_2^j) = m_1(F_1^i) m_2(F_2^j).$$

And for the joint plausibility Pl measure we have

$$\mathrm{Pl}(A) = \sum_{i,j:\ A \cap F_1^i \times F_2^j \neq \varnothing} m_1(F_1^i) m_2(F_2^j).$$

If the above random sets $(\mathcal{F}_1, m_1)$ and $(\mathcal{F}_2, m_2)$ are consonant, then the two random sets are also fuzzy sets $\tilde{A}_1$ and $\tilde{A}_2$. The membership function $\mu_{\tilde{A}}$ of the joint fuzzy set $\tilde{A}$ is defined by

$$\mu_{\tilde{A}}(\omega_1, \omega_2) = \min(\mu_{\tilde{A}_1}(\omega_1), \mu_{\tilde{A}_2}(\omega_2))$$

which is the possibility density of the joint possibility measure Pos. $\tilde{A}$ again defines a random set $(\mathcal{F}_f, m_f)$ with plausibility measure Pl_f.

Then in general we have

$$\mathrm{Pl} \neq \mathrm{Pos} = \mathrm{Pl}_f \text{ and } (\mathcal{F}, m) \neq (\mathcal{F}_f, m_f).$$

If we now reinterpret Pl and Pos $= \mathrm{Pl}_f$ as upper probabilities we get different results depending on the generation of the joint random sets which depends on the notion of independence.

3 Numerical example

As a numerical example we consider a beam of length 3 m supported on both ends and additionally bedded on two springs, see Fig 1. The values of the beam rigidity $EI = 10$ kNm^2 and of the equally distributed load $f(x) = 1$ kN/m are assumed to be deterministic. But the values of the two spring constants λ_1 and λ_2 are uncertain. We also assume that λ_1 and λ_2 are independent.

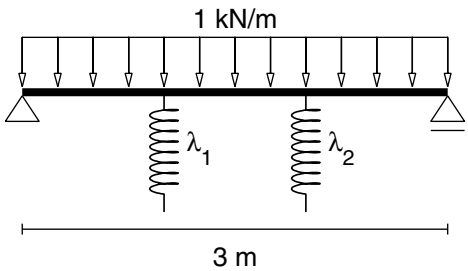


Fig. 1. Beam bedded on two springs.

We want to compute measures for the risk of failure of the beam. The criterion of failure is here

$$\max_{x\in[0,3]} |M(x)| \geq M_{\mathsf{yield}},$$

where $M(x)$ is the bending moment at a point $x \in [0,3]$ and M_{yield} the elastic limit moment. The bending moment M also depends on the two spring constants. We define a map

$$g : \Omega_1 \times \Omega_2 \longrightarrow \mathbb{R} : (\omega_1, \omega_2) \mapsto \max_{x\in[0,3]} |M(x, \omega_1, \omega_2)|,$$

which is the maximal bending moment of the beam depending on values ω_1 and ω_2 for the two spring constants. In Fig. 2 the function g is depicted as a contour plot for values $(\omega_1, \omega_2) \in [20, 40] \times [20, 40]$.

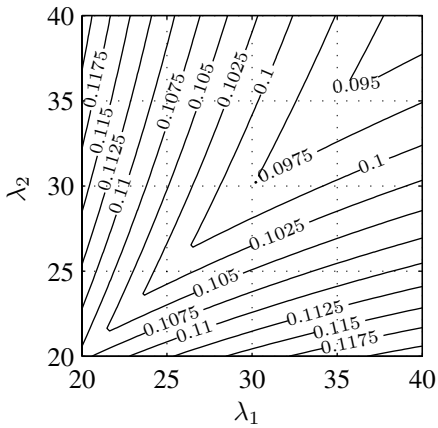


Fig. 2. Function g on $[20, 40] \times [20, 40]$.

In the following four examples the uncertainty about the values of the spring constants λ_1 and λ_2 is modelled by

- random intervals,
- fuzzy numbers and
- histograms.

3.1 Example 1

Uncertainty of λ_1:

The uncertainty of parameter λ_1 is modelled by a consonant random interval with focal sets

$$F_1^1 = [20, 40],\ F_1^2 = [30, 40],\ F_1^3 = \{30\},\ F_1^1 \supseteq F_1^2 \supseteq F_1^3$$

and weights

$$m_1(F_1^1) = 0.2,\ m_1(F_1^2) = 0.3,\ m_1(F_1^3) = 0.5.$$

In Fig. 3 the focal sets are plotted as a stack where the height is determined by the cumulative sum of the weights. Since the focals are nested the stack can be interpreted as a fuzzy number. Then the membership function is the contour of the stack of focals (light gray line).

Uncertainty of λ_2:

A consonant random set with focal sets

$$F_2^1 = [20, 40],\ F_2^2 = [28, 38],\ F_2^3 = \{31\},\ F_2^1 \supseteq F_2^2 \supseteq F_2^3,$$

and weights

$$m_2(F_2^1) = 0.3,\ m_2(F_2^2) = 0.3,\ m_2(F_2^3) = 0.4.$$

See Fig. 4.

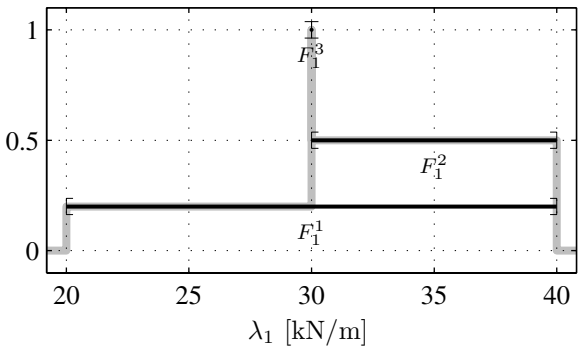


Fig. 3. Example 1: Uncertainty of λ_1.

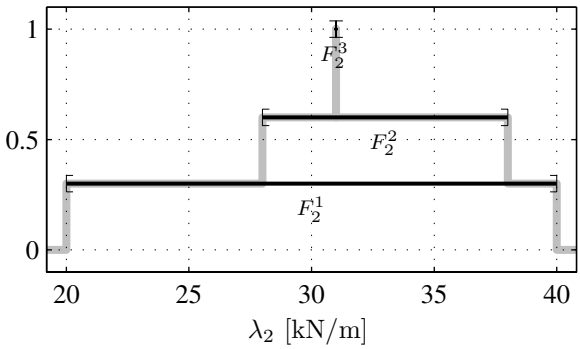


Fig. 4. Example 1: Uncertainty of λ_2.

3.2 Example 2

Uncertainty of λ_1:

The same consonant random interval as in example 1.

Uncertainty of λ_2:

The histogram depicted in Fig. 5. The focal sets of the corresponding random interval are

$$F_1^1 = [20, 25],\ F_1^2 = [25, 33],\ F_1^3 = [33, 40]$$

and the weights are

$$m_2(F_1^1) = 0.1,\ m_2(F_1^2) = 0.7,\ m_2(F_1^3) = 0.2$$

which are the areas of the three bins. See Fig. 5.

3.3 Example 3

Uncertainty of λ_1 and λ_2:

Again a consonant random interval, but with more focals to approximate a triangular fuzzy number. See Fig. 6.

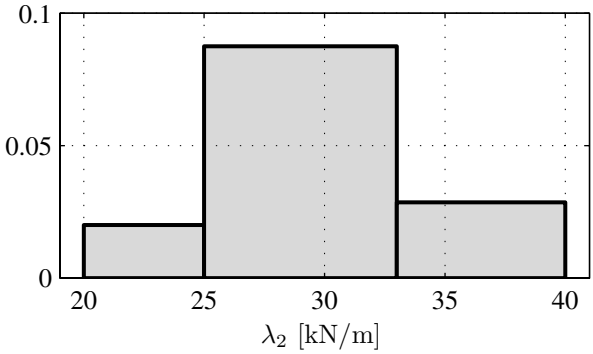


Fig. 5. Example 2: Uncertainty of λ_2.

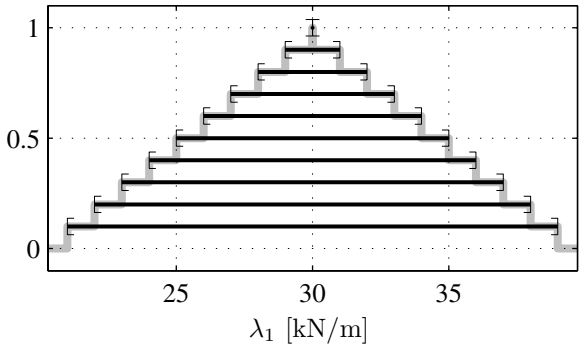


Fig. 6. Example 3: Uncertainty of λ_1 and λ_2.

3.4 Example 4

Uncertainty of $\boldsymbol{\lambda_1}$:

The same consonant random interval or triangular fuzzy number as in Example 3.

Uncertainty of $\boldsymbol{\lambda_2}$:

The histogram depicted in Fig. 7.

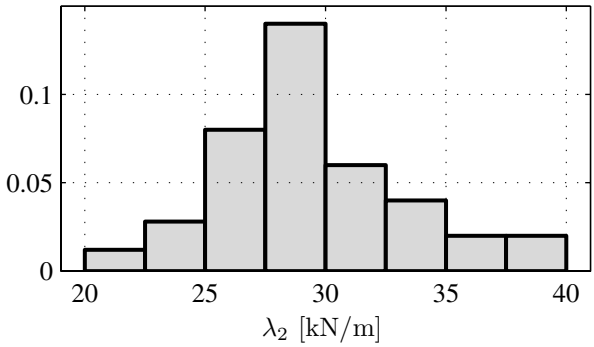


Fig. 7. Example 4: Uncertainty of λ_2.

4 Types of independence

Assume the two parameters λ_1, λ_2 are random variables defined by the probability measures P_1 and P_2 and the sets of possible values for the parameters are Ω_1 and Ω_2. If λ_1 and λ_2 are stochastically independent, then the joint probability measure is the product measure $P = P_1 \otimes P_2$. In this case the conditional probabilities satisfy

$$P(\cdot \times \Omega_2 |\ \Omega_1 \times \{\omega_2\}) = P_1 \text{ and}$$
$$P(\Omega_1 \times \cdot |\ \{\omega_1\} \times \Omega_2) = P_2$$

for all $\omega_1 \in \Omega_1$ with $P_1(\{\omega_1\}) > 0$ and for all $\omega_2 \in \Omega_2$ with $P_2(\{\omega_2\}) > 0$. That means:

If we learn the value ω_1 of parameter λ_1, our knowledge about the probability measure for parameter λ_2 does not change.

This will be our motivation for the notions of independence for sets of probability measures we subsequently discuss. For an introduction to independence for sets of probability measures see [1, 17].

Now assume that the uncertainty of the parameters λ_k is modelled by sets of probability measures $\mathcal{M}_k$, $k = 1, 2$. For our general introduction of the notions of independence the sets $\mathcal{M}_k$ will be arbitrary sets of probability measures. Later on we will consider $\mathcal{M}_k$ generated by ordinary sets, by random sets and fuzzy numbers.

4.1 Unknown interaction

We start with the case where nothing is known about dependence or independence of λ_1 and λ_2.

Definition 1 *The set $\mathcal{M}_{\mathrm{U}}$ is the set of all probability measures P for which the following conditions hold:*

$$P(\cdot \times \Omega_2) \in \mathcal{M}_1 \quad \textit{and} \quad P(\Omega_1 \times \cdot) \in \mathcal{M}_2. \tag{U}$$

$\mathcal{M}_{\mathrm{U}}$ is the set of *all* probability measures with marginals in $\mathcal{M}_1$ and $\mathcal{M}_2$. So $\mathcal{M}_{\mathrm{U}}$ is a good choice if it is not known what probability measure of $\mathcal{M}_1$ and $\mathcal{M}_2$ has to be combined and if the interaction or correlation between the probability measures is unknown. The subscript U indicates unknown interaction.

4.2 Epistemic independence

Definition 2 *Let Ω_1 and Ω_2 be finite sets. Then $\mathcal{M}_{\mathrm{E}}$ is the set of all probability measures which satisfy condition (U) and for which the conditional probabilities satisfy*

$$\begin{aligned} &P(\cdot \times \Omega_2 |\, \Omega_1 \times \{\omega_2\}) \in \mathcal{M}_1, \\ &P(\Omega_1 \times \cdot |\, \{\omega_1\} \times \Omega_2) \in \mathcal{M}_2 \end{aligned} \tag{E'}$$

for all $\omega_1 \in \Omega_1$ with marginal probability $P_1(\{\omega_1\}) > 0$ and $\omega_2 \in \Omega_2$ with $P_2(\{\omega_2\}) > 0$.

Using the formula of Bayes the conditions (E') can be rewritten in the following way: Let $P \in \mathcal{M}_{\mathrm{E}}$ and $P_1 \in \mathcal{M}_1$, $P_2 \in \mathcal{M}_2$, respectively, the marginal probability measures. Then P must satisfy

$$\begin{aligned} &P(\cdot \times \{\omega_2\}) = P_1^{|\omega_2} P_2(\{\omega_2\}), \\ &P(\{\omega_1\} \times \cdot) = P_1(\{\omega_1\}) P_2^{|\omega_1} \end{aligned} \tag{E"}$$

for all $(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$ with probability measures $P_1^{|\omega_2} \in \mathcal{M}_1$ and $P_2^{|\omega_1} \in \mathcal{M}_2$. In the case where $P_2(\{\omega_2\}) > 0$, $P_1^{|\omega_2}$ is the conditional probability $P(\cdot \times \Omega_2 | \Omega_1 \times \{\omega_2\})$. So the superscript $|\omega_2$ indicates conditioning with respect to $\{\omega_2\}$.

For singletons $\mathcal{M}_i = \{P_i\}$, $i = 1, 2$, the condition (E") leads to stochastic independence, because we always have $P_1^{|\omega_2} = P_1$ and $P_2^{|\omega_1} = P_2$ since there is only one probability measure in $\mathcal{M}_1$, $\mathcal{M}_2$, respectively. So epistemic independence is a sort of generalization of the notion of independence to sets of probability measures.

Epistemic independence means:

If we have learned the value ω_1 of λ_1, then the probability measure for the parameter λ_2 is again one of the probability measures in $\mathcal{M}_2$ (but in general not always the same for different ω_1), because the conditional probability measures have to be in $\mathcal{M}_2$, and vice versa.

Condition (E") can be generalized to infinite sets Ω_k in the following way: The set $\mathcal{M}_{\mathrm{E}}$ is the set of probability measures P for which the following holds:

There is a family $(P_1^{|\omega_2})_{\omega_2 \in \Omega_2}$ and $(P_2^{|\omega_1})_{\omega_1 \in \Omega_1}$ of probability measures with $P_1^{|\omega_2} \in \mathcal{M}_1$ and $P_2^{|\omega_1} \in \mathcal{M}_2$ such that for all P-measurable sets A

1. $\omega_1 \mapsto P_2^{|\omega_1}(A_{\omega_1})$ is P_1-measurable,
2. $\omega_2 \mapsto P_1^{|\omega_2}(A_{\omega_2})$ is P_2-measurable and
3. $P(A) = \int_{\Omega_1} P_2^{|\omega_1}(A_{\omega_1})\, \mathrm{d}P_1(\omega_1) = \int_{\Omega_2} P_1^{|\omega_2}(A_{\omega_2})\, \mathrm{d}P_2(\omega_2)$

holds where $A_{\omega_1} = \{\omega_2 : \ (\omega_1, \omega_2) \in A\}$ and similarly for A_{ω_2}.

4.3 Strong independence

Definition 3 *The set*

$$\mathcal{M}_{\mathrm{S}} = \{P = P_1 \otimes P_2 : \ P_1 \in \mathcal{M}_1, P_2 \in \mathcal{M}_2\}$$

is the set of all product measures which can be generated by probability measures in $\mathcal{M}_1$ *and* $\mathcal{M}_2$.

This is a stronger generalization of stochastic independence. We also get the set $\mathcal{M}_{\mathrm{S}}$ if we add the following conditions

$$\forall \omega_2 \in \Omega_2 : \ P_1^{|\omega_2} = P_1 \text{ and } \forall \omega_1 \in \Omega_1 : \ P_2^{|\omega_1} = P_2 \tag{S}$$

to the conditions for epistemic independence. This type of independence is called strong independence [1] or type-1 extension [17].

Strong independence is an appropriate choice, if the following assumptions are satisfied [1]:

1. The values of parameter λ_1 and λ_2 are random, each described by a unique but unknown probability distribution.
2. These probability distributions belong to the sets $\mathcal{M}_1$, $\mathcal{M}_2$, respectively.
3. The parameters λ_1 and λ_2 are stochastically independent.
4. It is not known which $P_1 \in \mathcal{M}_1$ is allowed to be combined with $P_2 \in \mathcal{M}_2$.

4.4 Relations between the above types of independence

Summarizing we recall that

- $\mathcal{M}_{\mathrm{U}}$ is the set of probability measures P subject to condition (U).
- $\mathcal{M}_{\mathrm{E}}$ is the set of probability measures P subject to conditions (U)+(E").
- $\mathcal{M}_{\mathrm{S}}$ is the set of probability measures P subject to conditions (U)+(E")+(S).

Since the conditions are successively added we get for the sets

$$\mathcal{M}_{\mathrm{U}} \supseteq \mathcal{M}_{\mathrm{E}} \supseteq \mathcal{M}_{\mathrm{S}}$$

and therefore for the lower and upper probabilities

$$\underline{P}_{\mathrm{U}} \leq \underline{P}_{\mathrm{E}} \leq \underline{P}_{\mathrm{S}} \ \leq \ \overline{P}_{\mathrm{S}} \leq \overline{P}_{\mathrm{E}} \leq \overline{P}_{\mathrm{U}}.$$

where the subscripts indicate the type of independence.

5 Sets of joint probability measures generated by random sets

5.1 Sets of joint probability measures generated by ordinary sets

Given a subset $F_k \subseteq \Omega_k$, we denote by $\mathcal{M}(F_k) = \{P : \ P(F_k) = 1\}$ the set of probability measures on F_k. Letting $\mathcal{M}_k = \mathcal{M}(F_k)$ we have that $\mathcal{M}_{\mathrm{U}} = \mathcal{M}(F_1 \times F_2)$ where $\mathcal{M}_{\mathrm{U}} := \mathcal{M}_{\mathrm{U}}(\mathcal{M}(F_1), \mathcal{M}(F_2))$ is the set of joint probability measures generated by the two sets $\mathcal{M}(F_1)$ and $\mathcal{M}(F_2)$ of probability measures according to condition (U). To get the upper and lower probability $\overline{P}_{\mathrm{U}}(A)$ and $\underline{P}_{\mathrm{U}}(A)$ with $A \subseteq \Omega = \Omega_1 \times \Omega_2$ we simply have to put a Dirac measure into $(F_1 \times F_2) \cap A$ for $\overline{P}_{\mathrm{U}}(A)$ or into $(F_1 \times F_2) \cap (\Omega \setminus A)$ for $\underline{P}_{\mathrm{U}}(A)$. Since a Dirac measure is a product measure conditions (E") and (S) are satisfied as well. So we get $\overline{P}_{\mathrm{S}}(A) = \overline{P}_{\mathrm{E}}(A) = \overline{P}_{\mathrm{U}}(A)$ and $\underline{P}_{\mathrm{U}}(A) = \underline{P}_{\mathrm{E}}(A) = \underline{P}_{\mathrm{S}}(A)$.

5.2 General formulation of the generation of sets $\mathcal{M}_k$ of joint probability measure for random sets

Let the random set $(\mathcal{F}_k, m_k)$ be given with focals $\mathcal{F}_k = \{F_k^1, \ldots, F_k^{n_k}\}$ and weight function m_k. The set of probability measures $\mathcal{M}_k$ is generated as follows:

$$\mathcal{M}_k := \mathcal{M}(\mathcal{F}_k, m_k) := \sum_{i=1}^{|\mathcal{F}_1|} m_k(F_k^i)\, \mathcal{M}(F_k^i) :=$$

$$= \left\{ P : \ P = \sum_{i=1}^{|\mathcal{F}_1|} m_k(F_k^i) P_k^i, \ P_k^i \in \mathcal{M}(F_k^i) \right\}.$$

We write a joint probability measure P in the following way:

$$P = \sum_{i=1}^{|\mathcal{F}_1|} \sum_{j=1}^{|\mathcal{F}_2|} m(F_1^i \times F_2^j)\, P^{ij}$$

where $F_1^i \times F_2^j$, $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$, are the joint focal sets and where $P^{ij} \in \mathcal{M}(F_1^i \times F_2^j)$ are probability measures on $F_1^i \times F_2^j$. We assume that there are no interactions between the marginal focal sets themselves. We have to choose the joint weights m, the probability measures P^{ij} and the type how the P^{ij} interact. This leads to three groups of conditions which will be indicated by a triple (A,B,C). The notation is explained in Fig. 8.

We emphasize that the choice of the Cartesian products $F_1^i \times F_2^j$ as joint focals is no restriction of generality. Joint focal sets $V \subseteq F_1^i \times F_2^j$ of arbitrary shape can be subsumed in our approach by restricting sets of joint probability measures on $F_1^i \times F_2^j$ to those whose support lies in V. Such subsets would describe specific types of dependence and thus will not enter our investigation of independence.

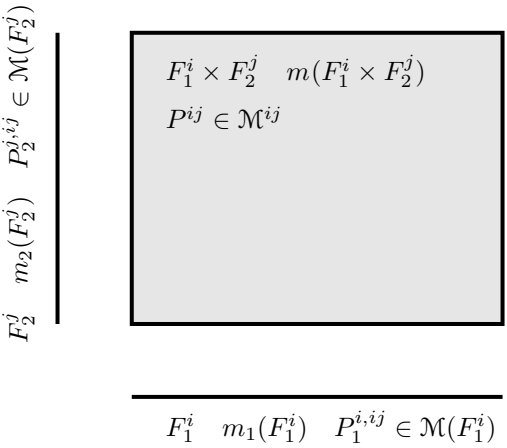


Fig. 8. Marginal focal sets F_1^i, F_2^j and joint focal set $F_1^i \times F_2^j$, marginal weights $m_1(F_1^i)$, $m_2(F_2^j)$ and joint weight $m(F_1^i \times F_2^j)$, joint probability measure P^{ij} and marginal probability measures $P_1^{i,ij}$, $P_2^{j,ij}$. The superscripts ij after the comma indicate that the marginal probability measure results from focal set $F_1^i \times F_2^j$.

5.3 The choice of the joint weights $m(F_1^i \times F_2^j)$

The weights m_1 and m_2 for the parameters λ_1 and λ_2 are probability measures on the sets of focal sets $\{F_1^1, \ldots, F_1^{n_1}\}$, $\{F_2^1, \ldots, F_2^{n_2}\}$ respectively. So if we want to choose the joint focal sets in a stochastically independent way, then $m = m_1 \otimes m_2$ which means $m(F_1^i \times F_2^j) = m_1(F_1^i)m_2(F_2^j)$ for all i, j. If we do not know how m_1 and m_2 interact, we choose unknown interaction.

Case (U−−):

Unknown interaction, m must satisfy condition (U). That means

$$m_1(F_1^i) = \sum_{j=1}^{|\mathcal{F}_2|} m(F_1^i \times F_2^j), \quad i = 1, \ldots, |\mathcal{F}_1|,$$
$$m_2(F_2^j) = \sum_{i=1}^{|\mathcal{F}_1|} m(F_1^i \times F_2^j), \quad j = 1, \ldots, |\mathcal{F}_2|.$$

Case (S−−):

Stochastic independence: $m(F_1^i \times F_2^j) = m_1(F_1^i)m_2(F_2^j)$.

Case (F−−):

m_1 and m_2 are correlated in such a way that the joint upper probability coincides with the joint possibility measure.

5.4 The choice of P^{ij}, $\mathcal{M}^{ij}$, respectively

$P^{ij} \in \mathcal{M}^{ij}$ is a probability measure on the joint focal set $F_1^i \times F_2^j$, $\mathcal{M}^{ij} \subseteq \mathcal{M}(F_1^i \times F_2^j)$. How a P^{ij} looks like depends on how $\mathcal{M}^{ij}$ is constructed from $\mathcal{M}(F_1^i)$ and $\mathcal{M}(F_2^j)$.

Case (−U−):

$\mathcal{M}^{ij} := \mathcal{M}_{\mathrm{U}}(\mathcal{M}(F_1^i), \mathcal{M}(F_2^j)) = \mathcal{M}(F_1^i \times F_2^j)$ which is the set of all joint probability measures generated by the sets $\mathcal{M}(F_1^i)$ and $\mathcal{M}(F_2^j)$ according to condition (U).

Case (−S−):

$\mathcal{M}^{ij} := \mathcal{M}_{\mathrm{S}}(\mathcal{M}(F_1^i), \mathcal{M}(F_2^j))$ which is the set generated according to strong independence.

5.5 The choice of interaction between the P^{ij}

Case (−−1):

Row- and columnwise equality conditions on the marginals of the probability measures on the joint focal sets:

$$P_1^i := P_1^{i,i1} = \cdots = P_i^{i,in_2},\ i = 1, \ldots, n_1,$$
$$P_2^j := P_2^{j,1j} = \cdots = P_i^{j,n_1 j},\ j = 1, \ldots, n_2.$$

The notation is explained in Fig. 8.

Case (−−0):

No interactions, this means that we can choose a $P^{ij} \in \mathcal{M}^{ij}$ on $F_1^i \times F_2^j$ irrespective of the probability measures chosen on other joint focal sets.

6 The different cases

Now we will discuss combinations of the above cases which lead to random set independence, unknown interaction, strong independence, fuzzy set independence and epistemic independence.

6.1 The cases (SU0) and (SS0), random set independence

Let $\mathcal{M}_{\mathrm{SU0}}$ be the set of probability measures generated according to case (SU0). A probability measure $P \in \mathcal{M}_{\mathrm{SU0}}$ can be written as

$$P = \sum_{i=1}^{|\mathcal{F}_1|} \sum_{j=1}^{|\mathcal{F}_2|} m(F_1^i \times F_2^j)\, P^{ij} = \sum_{i=1}^{|\mathcal{F}_1|} \sum_{j=1}^{|\mathcal{F}_2|} m_1(F_1^i) m_2(F_2^j)\, P^{ij}.$$

$\mathcal{M}_{\mathrm{SU0}}$ is the set of probability measures generated by the random set $(\mathcal{F}, m)$ with $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$ and weights $m(F_1^i \times F_2^j) := m_1(F_1^i) m_2(F_2^j)$ for all i, j. This type of independence is called random set independence [3]. For $\mathcal{M}_{\mathrm{SU0}}$ we have

$$\mathcal{M}_{\mathrm{R}} := \mathcal{M}_{\mathrm{SU0}} = \sum_{i=1}^{|\mathcal{F}_1|} \sum_{j=1}^{|\mathcal{F}_2|} m_1(F_1^i) m_2(F_2^j)\, \mathcal{M}(F_1^i \times F_2^j)$$

where R indicates random set independence.

The upper probability $\overline{P}_{\mathrm{R}} := \overline{P}_{\mathrm{SU0}}$ is the corresponding joint plausibility measure

$$\mathrm{Pl}(A) = \sum_{\substack{i,j:\\ F_1^i \times F_2^j \cap A \neq \varnothing}} m_1(F_1^i) m_2(F_2^j)$$

or the solution of the optimization problem:

Find $P^{ij}(A)$ such that

$$\sum_{i=1}^{|\mathcal{F}_1|} \sum_{j=1}^{|\mathcal{F}_2|} m_1(F_1^i) m_2(F_2^j)\, P^{ij}(A) = \max!$$

subject to $P^{ij} \in \mathcal{M}(F_1^i \times F_2^j)$. It is sufficient to search for appropriate Dirac measures for optimal P^{ij} which amounts to searching for points in $F_1^i \times F_2^j \cap A$. Since for Dirac measures condition $(-\mathrm{S}-)$ is satisfied we get $\overline{P}_{\mathrm{R}} = \mathrm{Pl} = \overline{P}_{\mathrm{SU0}} = \overline{P}_{\mathrm{SS0}}$. For the lower probability we get analogous results: $\underline{P}_{\mathrm{R}} = \mathrm{Bel} = \underline{P}_{\mathrm{SU0}} = \underline{P}_{\mathrm{SS0}}$.

Computational method to obtain $\overline{P}_{\mathrm{R}}$ for our examples:

Let $A = \{(\omega_1, \omega_2) :\ y = g(\omega_1, \omega_2) \geq M_{\mathsf{yield}}\}$. Then

$$\begin{aligned}
\overline{P}_{\mathrm{R}}(y \geq M_{\mathsf{yield}}) &= \sum_{\substack{i,j:\\ F_1^i \times F_2^j \cap A \neq \varnothing}} m_1(F_1^i) m_2(F_2^j) = \\
&= \sum_{\substack{i,j:\\ g(F_1^i \times F_2^j) \cap [M_{\mathsf{yield}}, \infty) \neq \varnothing}} m_1(F_1^i) m_2(F_2^j) = \\
&= \sum_{i,j:\ \overline{y}^{ij} \geq M_{\mathsf{yield}}} m_1(F_1^i) m_2(F_2^j)
\end{aligned}$$

with

$$\overline{y}^{ij} = \max_{(\omega_1,\omega_2)\in F_1^i\times F_2^j} g(\omega_1,\omega_2).$$

So for each joint focal set $F_1^i \times F_2^j$ an optimization problem has to be solved to get $\overline{y}^{ij}$.

Result for Example 1: The upper probability of failure for random set independence as a function of M_{yield} is depicted in Fig. 9.

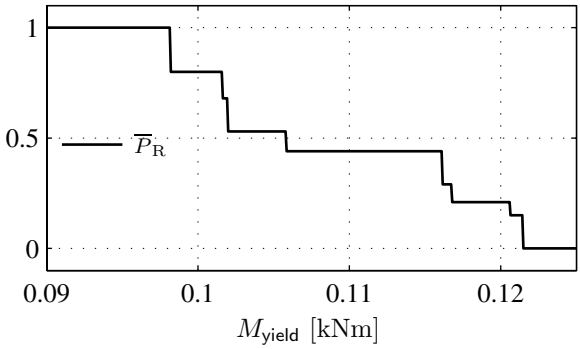


Fig. 9. Example 1: Upper probability $\overline{P}_{\mathrm{R}}(y \geq M_{\mathsf{yield}})$.

6.2 The cases (UU0) and (US0), unknown interaction

Let $\mathcal{M}_{\mathrm{UU0}}$ be the set of probability measures generated according to case (UU0). A computational method for $\overline{P}_{\mathrm{UU0}}$ is obtained in the following way:

$$\begin{aligned}\overline{P}_{\mathrm{UU0}}(A) &= \max\left\{P(A) : P \in \mathcal{M}_{\mathrm{UU0}}\right\} = \\ &= \sum_{i=1}^{|\mathcal{F}_1|}\sum_{j=1}^{|\mathcal{F}_2|} m^*(F_1^i \times F_2^j)\, P^{ij*}(A) = \sum_{\substack{i,j:\\ F_1^i\times F_2^j\cap A\neq\varnothing}} m^*(F_1^i \times F_2^j),\end{aligned} \tag{3}$$

where P^{ij*} are the same Dirac measures as for $\overline{P}_{\mathrm{R}}$. The weights m^* are obtained by solving the following linear optimization problem:

$$\sum_{\substack{i,j:\\ F_1^i\times F_2^j\cap A\neq\varnothing}} m(F_1^i \times F_2^j) = \max! \tag{4}$$

subject to

$$m_1(F_1^i) = \sum_{j=1}^{|\mathcal{F}_2|} m(F_1^i \times F_2^j), \quad i = 1, \ldots, |\mathcal{F}_1|, \tag{5}$$

$$m_2(F_2^j) = \sum_{i=1}^{|\mathcal{F}_1|} m(F_1^i \times F_2^j), \quad j = 1, \ldots, |\mathcal{F}_2|. \tag{6}$$

Minimization instead of maximization in (3) and (4) leads to the lower probability $\underline{P}_{\mathrm{UU0}}(A)$.

The set $\mathcal{M}_{\mathrm{UU0}}$ is just the set of probability measures which is generated by the least restrictive conditions on m and P^{ij}. It is proved in [5, 6] that

$$\mathcal{M}_{\mathrm{UU0}} = \mathcal{M}_{\mathrm{U}} := \big\{P : P(\cdot \times \Omega_2) \in \mathcal{M}(\mathcal{F}_1, m_1), P(\Omega_1 \times \cdot) \in \mathcal{M}(\mathcal{F}_2, m_2)\big\};$$

this is precisely the case of unknown interaction. By the same arguments as in the previous case we get $\overline{P}_{\mathrm{U}} = \overline{P}_{\mathrm{UU0}} = \overline{P}_{\mathrm{US0}}$ and $\underline{P}_{\mathrm{U}} = \underline{P}_{\mathrm{UU0}} = \underline{P}_{\mathrm{US0}}$.

Computational method to obtain $\overline{P}_{\mathrm{U}}(A)$:

The set A determines the linear objective function, which is in our examples:

$$\sum_{i,j:\ \overline{y}^{ij} \geq M_{\mathsf{yield}}} m(F_1^i \times F_2^j) = \max!$$

The linear conditions (5) and (6) are always the same and have to be generated only once.

Result for Example 1: The upper probability $\overline{P}_{\mathrm{U}}(y \geq M_{\mathsf{yield}})$ for unknown interaction is depicted in Fig. 10.

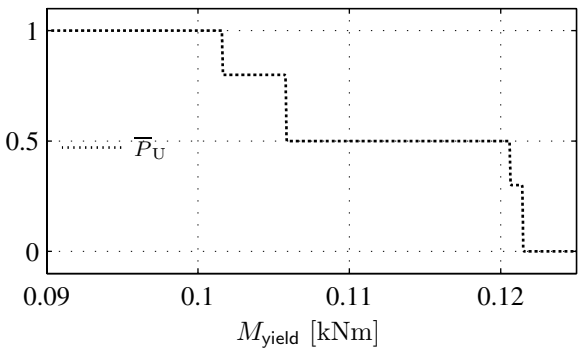


Fig. 10. Example 1: Upper probability $\overline{P}_{\mathrm{U}}(y \geq M_{\mathsf{yield}})$.

6.3 The case (SS1), strong independence

We write a probability measure $P_{\mathrm{SS1}} \in \mathcal{M}_{\mathrm{SS1}}$ in the following way:

$$\begin{aligned} P_{\mathrm{SS1}}(A) &= \sum_{i=1}^{|\mathcal{F}_1|} \sum_{j=1}^{|\mathcal{F}_2|} m(F_1^i \times F_2^j)\, P^{ij}(A) = \\ &= \sum_{i=1}^{|\mathcal{F}_1|} \sum_{j=1}^{|\mathcal{F}_2|} m_1(F_1^i) m_2(F_2^j)\, (P_1^i \otimes P_2^j)(A) = \\ &= \left(\sum_{i=1}^{|\mathcal{F}_1|} m_1(F_1^i)\, P_1^i \right) \otimes \left(\sum_{j=1}^{|\mathcal{F}_2|} m_2(F_2^j)\, P_2^j \right)(A) = \\ &= (P_1 \otimes P_2)(A) \end{aligned}$$

with $P_1 \in \mathcal{M}(\mathcal{F}_1, m_1)$ and $P_2 \in \mathcal{M}(\mathcal{F}_2, m_2)$. This leads to

$$\mathcal{M}_{\mathrm{SS1}} = \mathcal{M}_{\mathrm{S}} = \left\{ P_1 \otimes P_2 : P_1 \in \mathcal{M}(\mathcal{F}_1, m_1), P_2 \in \mathcal{M}(\mathcal{F}_2, m_2) \right\}$$

which is the case of strong independence.

Computational method:

We obtain a computational method for $\underline{P}_{\mathrm{S}}(A)$ and $\overline{P}_{\mathrm{S}}(A)$ from the following theorem.

Theorem 1 *The upper probability $\overline{P}_{\mathrm{S}}(A)$ is the solution of the following global optimization problem:*

$$\sum_{i=1}^{|\mathcal{F}_1|} \sum_{j=1}^{|\mathcal{F}_2|} m(F_1^i \times F_2^j)\, \chi_A(\omega_1^i, \omega_2^j) = \max!$$

subject to

$$\begin{aligned} \omega_1^i &\in F_1^i,\ i = 1, \ldots, |\mathcal{F}_1|, \\ \omega_2^j &\in F_2^j,\ j = 1, \ldots, |\mathcal{F}_2|, \end{aligned}$$

where χ_A is the indicator function of the set A. The lower probability $\underline{P}_{\mathrm{S}}(A)$ is obtained by minimization.

Proof: see [5, 9].

In general it is very hard to solve the above optimization problem because there may be many local maxima (or minima) and because the objective function is not continuous. Here we have adapted a branch and bound algorithm (c.f. [11]) to find the global optima.

For our example we have the set

$$A = \{(\omega_1, \omega_2) \in \Omega : y = g(\omega_1, \omega_2) \geq M_{\mathsf{yield}}\}$$

and therefore

$$\chi_A(\omega_1, \omega_2) = \begin{cases} 1, & y = g(\omega_1, \omega_2) \geq M_{\mathsf{yield}}, \\ 0, & \text{otherwise.} \end{cases}$$

Result for Example 1: The upper probability $\overline{P}_{\mathrm{S}}(y \geq M_{\mathsf{yield}})$ of failure for strong independence is depicted in Fig. 11 as a function of M_{yield}.

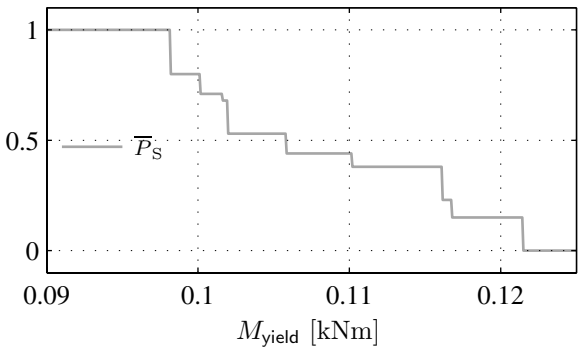


Fig. 11. Example 1: Upper probability $\overline{P}_{\mathrm{S}}(y \geq M_{\mathsf{yield}})$.

6.4 The cases (FU0) and (FS0), fuzzy set independence

Consonant random sets $(\mathcal{F}_k, m_k)$ can be interpreted as fuzzy sets $\tilde{A}_k$ or as possibility measures with density $\mu_{\tilde{A}_k}(\omega_k) = \sum_{i:\, \omega_k \in F_k^i} m_k(F_k^i)$. The joint possibility measure has the density

$$\mu_{\tilde{A}}(\omega_1, \omega_2) = \min(\mu_{\tilde{A}_1}(\omega_1), \mu_{\tilde{A}_2}(\omega_2))$$

which can again be transformed back into a joint random set.

We exemplify this for the consonant focal sets in Example 1: Recalling that the focal sets for parameter λ_1 and λ_2 are given as the intervals

$$F_1^1 = [20, 40],\ F_1^2 = [30, 40],\ F_1^3 = [30, 30],\ F_1^1 \supseteq F_1^2 \supseteq F_1^3$$

with weights $m_1 = (0.2, 0.3, 0.5)$ and for parameter λ_2

$$F_2^1 = [20, 40],\ F_2^2 = [28, 38],\ F_2^3 = [31, 31],\ F_2^1 \supseteq F_2^2 \supseteq F_2^3$$

with weights $m_2 = (0.3, 0.3, 0.4)$.

Summing up the weights leads to α-values at which the density function has a jump. These α-levels are for λ_1:

$$\alpha_1 = (\alpha_1^1, \alpha_1^2, \alpha_1^3) = (0.2, 0.5, 1)$$

and for λ_2:

$$\alpha_2 = (\alpha_2^1, \alpha_2^2, \alpha_2^3) = (0.3, 0.6, 1).$$

The "joint" α-values

$$\alpha = (\alpha_1^1, \alpha_2^1, \alpha_1^2, \alpha_2^2, \alpha_1^3) = (0.2, 0.3, 0.5, 0.6, 1)$$

are just the values at which the density of the joint possibility measure has a jump.

The corresponding random set $(\mathcal{F}, m)$ has focal sets $F^k = \tilde{A}_1^{\alpha_k} \times \tilde{A}_2^{\alpha_k}$, $\alpha_k \in \alpha$, with weights $m(F^k) = \alpha_k - \alpha_{k-1}$ $(\alpha_0 = 0)$, $k = 1, \ldots, 5$:

$$F^1 = \tilde{A}_1^{0.2} \times \tilde{A}_2^{0.2} = [20, 40] \times [20, 40], \quad F^2 = \tilde{A}_1^{0.3} \times \tilde{A}_2^{0.3} = [30, 40] \times [20, 40],$$
$$F^3 = \tilde{A}_1^{0.5} \times \tilde{A}_2^{0.5} = [30, 40] \times [28, 38], \quad F^4 = \tilde{A}_1^{0.6} \times \tilde{A}_2^{0.6} = [30, 30] \times [28, 38],$$
$$F^5 = \tilde{A}_1^{1} \times \tilde{A}_2^{1} = [30, 30] \times [31, 31],$$

with weights

$$m = (0.2, 0.1, 0.2, 0.1, 0.4).$$

The nested joint focal sets $F^1, \ldots, F^5$ form a subset of the family of all product sets $F_1^i \times F_2^j$, $i, j = 1, 2, 3$. If we take all of them and set the weights of the "unused" sets to 0, then we get a random set with the same focals as for random set independence, but with different weights. This is the case (FU0) which we call the case of fuzzy set independence (or possibilistic independence [2]) in analogy to random set independence, and we let $\mathcal{M}_{\mathrm{F}} = \mathcal{M}_{\mathrm{FU0}}$. By similar arguments as in case (SU0) we get $\overline{P}_{\mathrm{FU0}} = \mathrm{Pos} = \overline{P}_{\mathrm{FS0}}$ and $\underline{P}_{\mathrm{FU0}} = \mathrm{Nec} = \underline{P}_{\mathrm{FS0}}$.

Result for Example 1: The upper probability $\overline{P}_{\mathrm{F}}(y \geq M_{\mathsf{yield}})$ of failure for fuzzy set independence is depicted in Fig. 12.

6.5 $\mathcal{M}_{\mathrm{E}}$ is a subset of $\mathcal{M}_{\mathrm{R}}$ in the finite case

Theorem 2 *Let Ω_1 and Ω_2 be finite sets, $\mathcal{M}_{\mathrm{E}} = \mathcal{M}_{\mathrm{E}}(\mathcal{M}(\mathcal{F}_1, m_1), \mathcal{M}(\mathcal{F}_2, m_2))$ the set of probability measures generated by two random sets according to epistemic independence and $\mathcal{M}_{\mathrm{R}} = \mathcal{M}_{\mathrm{R}}(\mathcal{M}(\mathcal{F}_1, m_1), \mathcal{M}(\mathcal{F}_2, m_2))$ the set generated according to random set independence. Then $\mathcal{M}_{\mathrm{E}} \subseteq \mathcal{M}_{\mathrm{R}}$.*

Proof: see [6, 9].

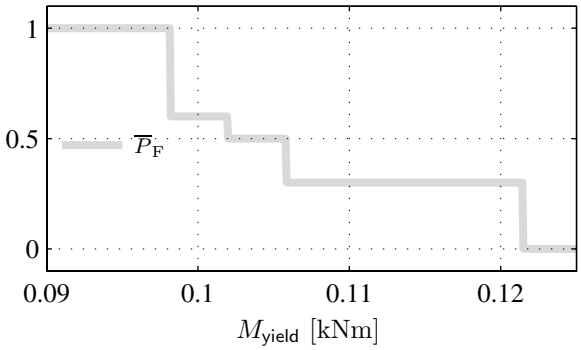


Fig. 12. Upper probability $\overline{P}_{\mathrm{F}}(y \geq M_{\mathsf{yield}})$ of failure.

6.6 Relations between the different cases

Fig. 13 depicts the relations between the sets of joint probability measures. For the upper probabilities and lower probabilities see Fig. 14 and 15. The results for epistemic independence have been proven only for finite sets Ω_1 and Ω_2 (indicated by the dash box).

$$
\begin{array}{ccc}
\boxed{\mathcal{M}_{\mathrm{F}} = \mathcal{M}_{\mathrm{FU0}}} & \supseteq & \mathcal{M}_{\mathrm{FS0}} \\
\cap\!| & & \cap\!| \\
\boxed{\mathcal{M}_{\mathrm{U}} = \mathcal{M}_{\mathrm{UU0}}} & \supseteq & \mathcal{M}_{\mathrm{US0}} \\
\cup\!| & & \cup\!| \\
\boxed{\mathcal{M}_{\mathrm{R}} = \mathcal{M}_{\mathrm{SU0}}} & \supseteq & \mathcal{M}_{\mathrm{SS0}} \\
\cup\!| & & \cup\!| \\
\boxed{\mathcal{M}_{\mathrm{E}}} & \supseteq & \boxed{\mathcal{M}_{\mathrm{S}} = \mathcal{M}_{\mathrm{SS1}}}
\end{array}
$$

Fig. 13. Relations between the different sets of joint probability measures.

6.7 Computational effort

To obtain $\overline{P}_{\mathrm{F}}$, $\overline{P}_{\mathrm{R}}$ and $\overline{P}_{\mathrm{U}}$ we have to solve a global optimization problem for each joint focal set because we need the images $g(F_1^i \times F_2^j)$ of the focals.

$$
\begin{array}{ccc}
\boxed{\overline{P}_{\mathrm{F}} = \mathrm{Pos} = \overline{P}_{\mathrm{FU0}}} & = & \overline{P}_{\mathrm{FS0}} \\
\wedge\!| & & \wedge\!| \\
\boxed{\overline{P}_{\mathrm{U}} = \overline{P}_{\mathrm{UU0}}} & = & \overline{P}_{\mathrm{US0}} \\
\vee\!| & & \vee\!| \\
\boxed{\overline{P}_{\mathrm{R}} = \mathrm{Pl} = \overline{P}_{\mathrm{SU0}}} & = & \overline{P}_{\mathrm{SS0}} \\
\vee\!| & & \vee\!| \\
\boxed{\overline{P}_{\mathrm{E}}} & \geq & \boxed{\overline{P}_{\mathrm{S}} = \overline{P}_{\mathrm{SS1}}}
\end{array}
$$

Fig. 14. Relations between the upper probabilities for the different cases.

$$
\begin{array}{ccc}
\boxed{\underline{P}_{\mathrm{F}} = \mathrm{Nec} = \underline{P}_{\mathrm{FU0}}} & = & \underline{P}_{\mathrm{FS0}} \\
\vee\!| & & \vee\!| \\
\boxed{\underline{P}_{\mathrm{U}} = \underline{P}_{\mathrm{UU0}}} & = & \underline{P}_{\mathrm{US0}} \\
\wedge\!| & & \wedge\!| \\
\boxed{\underline{P}_{\mathrm{R}} = \mathrm{Bel} = \underline{P}_{\mathrm{SU0}}} & = & \underline{P}_{\mathrm{SS0}} \\
\wedge\!| & & \wedge\!| \\
\boxed{\underline{P}_{\mathrm{E}}} & \leq & \boxed{\underline{P}_{\mathrm{S}} = \underline{P}_{\mathrm{SS1}}}
\end{array}
$$

Fig. 15. Relations between the lower probabilities for the different cases.

Therefore it is very important to use all properties of the function g such as monotonicity and convexity to minimize the effort. The computation of the images of the focals can be parallelized, that means that each computation of $g(F_1^i \times F_2^j)$ can be done on a separate CPU or on a separate computer (grid computing).

The effort for the computation of $\overline{P}_{\mathrm{F}}$, $\overline{P}_{\mathrm{R}}$ or $\overline{P}_{\mathrm{U}}$ depends on the number of joint focals generated. For $\overline{P}_{\mathrm{F}}$ it is very low but in general for $\overline{P}_{\mathrm{R}}$ and $\overline{P}_{\mathrm{U}}$ it is high. In addition to compute $\overline{P}_{\mathrm{U}}$ we have to solve a linear optimization problem to get the joint weights m. As already mentioned the computational effort

for $\overline{P}_{\mathrm{S}}$ is very high, but if the function g is monotonic, the result coincides with $\overline{P}_{\mathrm{R}}$, see [6].

7 Numerical results for Examples 1, 2, 3, 4 and Conclusion

First we show the results for Example 1 in one picture so that we can see the relations between the different types of independence. Since the results for different types of independence partly coincide, we have to recall the rule $\overline{P}_{\mathrm{S}} \leq \overline{P}_{\mathrm{R}} \leq \overline{P}_{\mathrm{U}}$. So if the curve for $\overline{P}_{\mathrm{S}}$ or $\overline{P}_{\mathrm{U}}$ partly disappears, then such a curve is "under" the graph of $\overline{P}_{\mathrm{R}}$, see Fig. 16. In this Figure we can also see that sometimes $\overline{P}_{\mathrm{R}}$ is greater than $\overline{P}_{\mathrm{F}}$ and sometimes smaller. So especially the choice between fuzzy set and random set independence should depend on which type of independence is appropriate for the current problem and not on the computational effort.

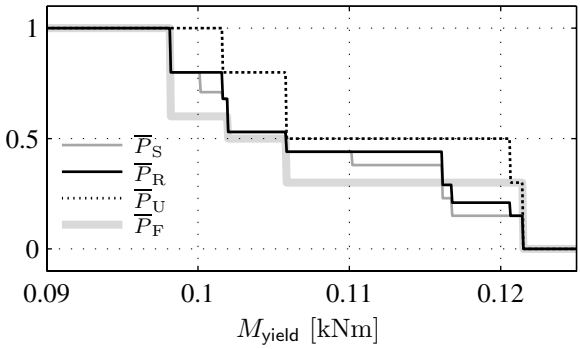


Fig. 16. Results for example 1.

In Fig. 17 the results $\overline{P}_{\mathrm{S}}$, $\overline{P}_{\mathrm{R}}$ and $\overline{P}_{\mathrm{U}}$ are depicted for Example 2. Of course there is no result for fuzzy set independence, because the random set of parameter λ_2 is not consonant.

In Fig. 18 the results $\overline{P}_{\mathrm{F}}$, $\overline{P}_{\mathrm{R}}$ and $\overline{P}_{\mathrm{U}}$ for two smoother fuzzy numbers are depicted. We can see that for $\overline{P}_{\mathrm{R}}$ the steps are very small, because there are a lot of joint focal sets (in contrary to fuzzy set independence). In Fig. 19 the results $\overline{P}_{\mathrm{R}}$ and $\overline{P}_{\mathrm{U}}$ for Example 4 are plotted. For Example 3 and 4, we did not compute $\overline{P}_{\mathrm{S}}$ because of the large number of joint focal sets, but it can be bounded by $\overline{P}_{\mathrm{R}}$.

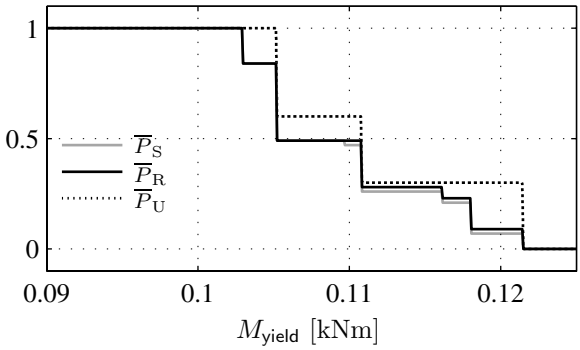


Fig. 17. Results for example 2.

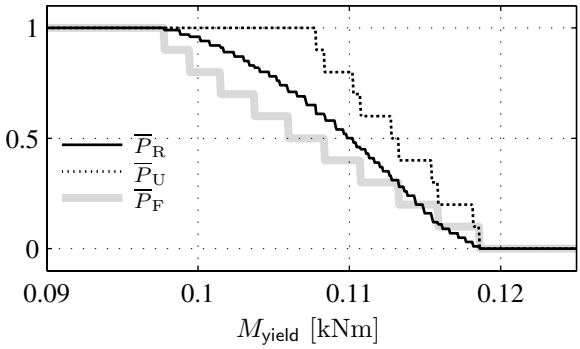


Fig. 18. Results for example 3.

Summarizing, we have considered models of parameter uncertainty using intervals, random sets, histograms and fuzzy sets. In our approach, we have shown that sets of probability measures provide a unifying interpretation for all these models and allow the propagation through numerical algorithms. One advantage of this interpretation is that it admits using the semantics of probability theory for interpreting the data. Another advantage is that it makes it possible to analyze and give a meaning to various concepts of dependence or independence of multivariate data, whose differences are not visible when one remains on the basic level of random sets or fuzzy sets.

In solving our examples, we have shown that the type of independence chosen crucially influences the lower and upper probabilities of the outcome

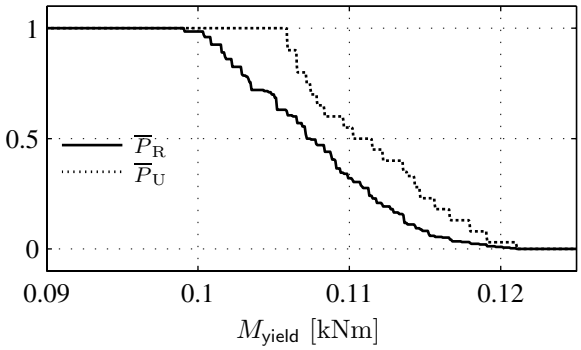


Fig. 19. Results for example 4.

in general. For computational purposes, it is not necessary to go down to the level of sets of probability measures. The computations of the upper probabilities can be reduced to optimization problems taking into account the special structures of the different types of independence considered here.

References

[1] I. Couso, S. Moral, and P. Walley. Examples of independence for imprecise probabilities. In G. de Cooman, G. Cozman, S. Moral, and P. Walley, editors, *Proceedings of the first international symposium on imprecise probabilities and their applications*, pages 121–130, Ghent, 1999. Universiteit Gent.

[2] G. De Cooman. Possibility theory III: possibilistic independence. *International Journal of General Systems*, 25:353–371, 1997.

[3] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.*, 38:325–339, 1967.

[4] Th. Fetz. Finite element method with fuzzy parameters. In Troch I. and Breitenecker F., editors, *Proceedings IMACS Symposium on Mathematical Modelling*, volume 11, pages 81–86, Vienna, 1997. ARGESIM Report.

[5] Th. Fetz. Sets of joint probability measures generated by weighted marginal focal sets. In *G. de Cooman, T. Fine, T. Seidenfeld (Eds.), ISIPTA'01, Proceedings of the Second Symposium on Imprecise Probabilities and Their Applications.*, pages 171–178, Maastricht, 2001. Shaker Publ. BV.

[6] Th. Fetz. Mengen von gemeinsamen Wahrscheinlichkeitsmaßen erzeugt von zufälligen Mengen, Dissertation, Universität Innsbruck, 2003.

[7] Th. Fetz, M. Hofmeister, G. Hunger, J. Jäger, H. Lessman, M. Oberguggenberger, A. Rieser, and R. F. Stark. Tunnelberechnung – Fuzzy? *Bauingenieur*, 72:33–40, 1997.

[8] Th. Fetz, J. Jäger, D. Köll, G. Krenn, H. Lessmann, M. Oberguggenberger, and R. Stark. Fuzzy models in geotechnical engineering and construction management. In this volume.

[9] Th. Fetz and M. Oberguggenberger. Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering and Systems Safety*, 85:73–87, 2004.

[10] Th. Fetz, M. Oberguggenberger, and S. Pittschmann. Applications of possibility and evidence theory in civil engineering. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 8(3):295–309, 2000.

[11] R. Hammer, M. Hocks, U. Kulisch, and D. Ratz. *C++ Toolbox for Verified Computing*. Springer, Berlin, 1995.

[12] B. Möller. Fuzzy-Modellierung in der Baustatik. *Bauingenieur*, 72:75–84, 1997.

[13] B. Möller, M. Beer, W. Graf, and A. Hoffmann. Possibility theory based safety assessment. *Computer-Aided Civil and Infrastructure Engineering*, 14:81–91, 1999.

[14] R. L. Muhanna and R. L. Mullen. Formulation of fuzzy finite-element methods for solid mechanics problems. *Computer Aided Civil and Infrastructure Engineering*, 14:107–117, 1999.

[15] F. Tonon and A. Bernardini. A random set approach to the optimization of uncertain structures. *Comput. Struct.*, 68(6):583–600, 1998.

[16] F. Tonon and A. Bernardini. Multiobjective optimization of uncertain structures through fuzzy set and random set theory. *Computer-Aided Civil and Infrastructure Engineering*, 14:119–140, 1999.

[17] P. Walley. *Statistical reasoning with imprecise probabilities.* Chapman and Hall, London, 1991.

[18] Z. Wang and G. J. Klir. *Fuzzy Measure Theory.* Plenum Press, New York, 1992.

[19] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.

Sensitivity analysis

Alexander Ostermann

Institut für Technische Mathematik, Geometrie und Bauinformatik, Universität Innsbruck

Summary. When studying a mathematical model it is not enough to compute individual solutions. It is equally important to determine systematically the influence of parameter variations on these solutions. This is particularly true in engineering applications where parameters of the underlying model are often only imprecisely known. The main task of sensitivity analysis is to identify critical parameter dependences. In this short article, we review the basic ideas of sensitivity analysis for deterministic models. We emphasize the use of internal numerical differentiation which is a reliable and robust method. A greater part of the paper is devoted to typical applications, illustrated by numerical examples.

1 Introduction

Sensitivity analysis is concerned with the propagation of uncertainties in mathematical models. Its main task is to assess the influence of parameters on the state of the system. The term parameter is used here in a broad sense. For example, if the underlying problem consists in calculating the shape of a rod under loading, then the length of the rod itself can be considered as a parameter and the purpose of sensitivity analysis could be to determine the influence of the length on the deformation of the rod.

In this paper we will model parameter dependent systems with the help of nonlinear functions. Within this framework, sensitivity can be defined as the Fréchet derivative of the function with respect to the parameters. It thus describes the variations of the output values for small parameter changes. Very often, however, the underlying function is not known explicitly, but only defined as the result of a possibly involved numerical calculation. This leads to some minor complications. In Sect. 2 we introduce the mathematical framework and recall the definition of differentiability of a function in several variables. We further discuss the relations between sensitivity and condition. The latter is a common concept in numerical analysis.

Since sensitivity is closely related to derivatives, we discuss in Sect. 3 some computational issues of differentiation. We emphasize here the concept

of numerical differentiation which goes well together with the fact that the modelling function is often known numerically only.

The main part of this paper is devoted to numerical examples. In Sect. 4 we will discuss five typical applications from engineering: systems of linear equations, eigenvalue problems, systems of nonlinear equations, ordinary differential equations, and elliptic boundary value problems. A straightforward method for computing sensitivities would be *external differentiation*. This approach relies on the approximation of derivatives by finite differences. It only requires a few evaluations of the underlying function. In practice, however, it is often preferable to solve the so-called sensitivity equations instead. The approach based on sensitivity equations is sometimes also called the method of *internal differentiation*. It gives much better results than external differentiation, if the underlying functions are evaluated with low accuracy only. Since the sensitivity equations have a structure that is similar to the underlying function, the computational costs of this method are comparable to external differentiation. For each of the five problems in Sect. 4, we will derive the corresponding sensitivity equations. Further, we will give some numerical examples in order to illustrate our theoretical considerations. Some final conclusions are given in Sect. 5.

For engineers, sensitivity analysis can mean much more than estimating derivatives. In a stochastic framework, for example, it might be interesting to model the parameters by probability distributions and to ask for the corresponding probabilities of the output values of the model. Answering this question requires much more computational effort than our above described methods. An illustrative example from electrical engineering can be found in [9, Sect. 2.7]. To obtain the desired probability distribution there, one has to solve the underlying problem 10 000 times. Such extensions to sensitivity analysis, however, will not be discussed here.

2 Mathematical background

Consider a system with m output values $\mathbf{u} = [u_1, \ldots, u_m]^\mathsf{T}$, depending on n input parameters $\mathbf{p} = [p_1, \ldots, p_n]^\mathsf{T}$. We model this system mathematically by a function

$$\mathbf{F} : D \subset \mathbb{R}^n \to \mathbb{R}^m : \ \mathbf{p} \mapsto \mathbf{u} = \mathbf{F}(\mathbf{p}), \tag{1}$$

defined on an open domain D. The properties of the model are thus fully determined by $\mathbf{F}$. A pitfall of the approach, however, is that $\mathbf{F}$ is *not* always known *explicitly*. In finite element calculations, for example, $\mathbf{F}(\mathbf{p})$ might be the calculated strains at m predefined points of a tunnel during excavation for assumed soil parameters $\mathbf{p}$. An evaluation of $\mathbf{F}$ then corresponds to a possibly involved computation on a computer, typically with low accuracy.

Since sensitivity analysis is concerned with the influence of parameters on the output of the system, we are led to investigate the behavior of $\mathbf{F}$ in a ball $\mathcal{B}_R(\mathbf{q})$ of radius R around the point $\mathbf{q} \in D$

$$\mathcal{B}_R(\mathbf{q}) = \{\mathbf{p} \in \mathbb{R}^n : \|\mathbf{p} - \mathbf{q}\| \leq R\}.$$

Here, $\|\cdot\|$ denotes an arbitrary norm on $\mathbb{R}^n$. After a possible reduction of R, we can always assume that this ball lies in the domain D of $\mathbf{F}$. Assume that $\mathbf{F}$ is Fréchet differentiable at $\mathbf{q}$, i.e. there exists a $m \times n$ matrix $\mathbf{J}$ such that for $\mathbf{p} \in \mathcal{B}_R(\mathbf{q}) \setminus \{\mathbf{q}\}$

$$\lim_{\mathbf{p}\to\mathbf{q}} \frac{\|\mathbf{F}(\mathbf{p}) - \mathbf{F}(\mathbf{q}) - \mathbf{J}\cdot(\mathbf{p}-\mathbf{q})\|}{\|\mathbf{p}-\mathbf{q}\|} = 0. \tag{2}$$

Obviously, the existence of all partial derivatives of $\mathbf{F}$ at $\mathbf{q}$ is a necessary condition for Fréchet differentiability. Therefore, the above matrix $\mathbf{J}$ is just the Jacobian matrix of $\mathbf{F}$ at $\mathbf{q}$

$$\mathbf{J} = \mathbf{F}'(\mathbf{q}) = \begin{bmatrix} \frac{\partial F_1}{\partial p_1}(\mathbf{q}) & \dots & \frac{\partial F_1}{\partial p_n}(\mathbf{q}) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial p_1}(\mathbf{q}) & \dots & \frac{\partial F_m}{\partial p_n}(\mathbf{q}) \end{bmatrix}.$$

We further recall that $\mathbf{F}$ is Fréchet differentiable at $\mathbf{q}$, if all partial derivatives of $\mathbf{F}$ exist in $\mathcal{B}_R(\mathbf{q})$ and are *continuous* at $\mathbf{q}$, see [7, Chap. 5.2].

2.1 Sensitivity

Let the behavior of the system be characterized by a Fréchet differentiable function $\mathbf{u} = \mathbf{F}(\mathbf{p})$. Its *sensitivity* with respect to the parameter p_j at the point $\mathbf{q}$ is *defined* as

$$\mathbf{z} = \frac{\partial \mathbf{u}}{\partial p_j} = \frac{\partial \mathbf{F}}{\partial p_j}\Big|_{\mathbf{p}=\mathbf{q}} \tag{3}$$

and is thus given by the jth column of the Jacobian matrix $\mathbf{F}'(\mathbf{q})$. It is sometimes convenient to collect these partial sensitivities in the $m \times n$ matrix

$$\mathbf{Z} = \mathbf{F}'(\mathbf{q}).$$

The sensitivity matrix $\mathbf{Z}$ describes the variation of the output values for small changes in the parameters, viz.

$$\mathbf{F}(\mathbf{p}) - \mathbf{F}(\mathbf{q}) \approx \mathbf{Z}\cdot(\mathbf{p}-\mathbf{q}) \tag{4}$$

for $\|\mathbf{p}-\mathbf{q}\|$ sufficiently small.

2.2 Condition numbers

The concept of condition is common in numerical analysis. It is a means to classify whether the solution of a problem depends sensitively on its input

data. Since this concept is very close to sensitivity analysis, we will shortly comment on it.

The problem $\mathbf{p} \mapsto \mathbf{F}(\mathbf{p})$ is called *well-posed* in $\mathcal{B}_R(\mathbf{q})$ if and only if there exists a constant $L(R) \geq 0$ such that for all $\mathbf{p} \in \mathcal{B}_R(\mathbf{q})$

$$\|\mathbf{F}(\mathbf{p}) - \mathbf{F}(\mathbf{q})\| \leq L(R) \cdot \|\mathbf{p} - \mathbf{q}\|.$$

Without loss of generality, the Lipschitz constant $L(R)$ may be chosen here minimal. It is called *condition number* of problem (1) in $\mathcal{B}_R(\mathbf{q})$ and describes the maximum amplification of parameter variations in $\mathcal{B}_R(\mathbf{q})$. Under appropriate regularity assumptions, it can be shown that

$$\lim_{R \to 0} L(R) = \|\mathbf{F}'(\mathbf{q})\|,$$

see [9, Chap. 2.6]. The operator norm of the sensitivity matrix is thus the limit of the corresponding condition numbers. This shows that condition is only a coarse means to estimate sensitivity.

3 Analytic vs. numerical differentiation

Since differentiation plays a prominent role in sensitivity analysis, we recall some algorithmic facts about it.

3.1 Analytic differentiation

If the function $\mathbf{F}$ is given explicitly, then its partial derivatives can be obtained in a straightforward way by applying the standard rules of differentiation. However, if $\mathbf{F}$ is made up of involved expressions, this could be quite a tedious and error-prone task. Alternatively, the derivatives can by obtained symbolically with the help of computer algebra systems like Maple or even with automatic differentiation tools, see [5]. We propose here still another possibility, namely numerical differentiation.

3.2 Numerical differentiation

Let φ denote a sufficiently smooth scalar function in one variable, defined in some neighborhood of $q \in \mathbb{R}$. We approximate its first derivative at q numerically by the one-sided difference scheme

$$\varphi'(q) \approx \frac{\varphi(q+\delta) - \varphi(q)}{\delta}. \tag{5}$$

Note that in practical calculations, the function φ is only evaluated up to a relative accuracy ACC. Let φ_* denote the numerical evaluation of φ, hence

$$\varphi_*(p) = (1+\varepsilon)\,\varphi(p) \qquad \text{with} \quad |\varepsilon| \le \mathrm{ACC}.$$

In engineering applications like finite element calculations, we typically have $\mathrm{ACC} \approx 10^{-3}$ only. On the other hand, whenever φ is explicitly known, evaluation in double precision IEEE arithmetic yields $\mathrm{ACC} \approx 10^{-16}$.

The accuracy of (5) obviously depends on δ. On the one hand, Taylor series expansion of the right-hand side of (5) shows that the discretization error is proportional to δ. On the other hand, since φ is only evaluated up to a relative accuracy ACC, the finite difference contains an additional error term proportional to ACC/δ, hence

$$\varphi'(q) = \frac{\varphi_*(q+\delta) - \varphi_*(q)}{\delta} + \mathcal{O}(\delta) + \mathcal{O}\left(\frac{\mathrm{ACC}}{\delta}\right). \tag{6}$$

Best results for (5) are expected if both error terms in (6) are of the same magnitude, viz. $\delta \approx \sqrt{\mathrm{ACC}}$. With this choice, the error of (5) is $\sqrt{\mathrm{ACC}}$, too.

More accurate results can be achieved with central differences

$$\varphi'(q) \approx \frac{\varphi(q+\delta) - \varphi(q-\delta)}{2\delta}. \tag{7}$$

An error analysis similar to the one above shows that

$$\varphi'(q) = \frac{\varphi_*(q+\delta) - \varphi_*(q-\delta)}{2\delta} + \mathcal{O}(\delta^2) + \mathcal{O}\left(\frac{\mathrm{ACC}}{\delta}\right).$$

Best results for (7) are expected if both of the above error terms are of the same magnitude. For the choice $\delta \approx \sqrt[3]{\mathrm{ACC}}$, the error of (7) is proportional to $\mathrm{ACC}^{2/3}$.

The partial derivatives of a vector-valued function $\mathbf{F}$ in several variables are approximated in a similar way. Let $j \in \{1, \ldots, n\}$ and let $\mathbf{e}_j = [\delta_{jk}]_{k=1}^n$ denote the standard basis vector, defined with the Kronecker symbol

$$\delta_{jk} = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{else.} \end{cases}$$

The jth column of the Jacobian matrix can be approximated by one-sided differences

$$\frac{\partial \mathbf{F}}{\partial p_j}(\mathbf{q}) = \begin{bmatrix} \dfrac{\partial F_1}{\partial p_j}(\mathbf{q}) \\ \cdots \\ \dfrac{\partial F_m}{\partial p_j}(\mathbf{q}) \end{bmatrix} \approx \frac{\mathbf{F}(\mathbf{q} + \delta \mathbf{e}_j) - \mathbf{F}(\mathbf{q})}{\delta}. \tag{8}$$

An approximation of the whole Jacobian requires $n+1$ evaluations of the function $\mathbf{F}$ at most. Taking central differences in (8) yields again a slightly higher accuracy at the price of more function evaluations.

4 Examples

In engineering applications, many interesting systems are, in general, not given in explicit form (1). Nevertheless, the above ideas can be applied to compute sensitivities. In this section, we will illustrate this with five typical examples: systems of linear equations, eigenvalue problems, systems of nonlinear equations, ordinary differential equations, and elliptic boundary value problems.

4.1 Systems of linear equations

Let our problem first take the form of a linear system of equations

$$\mathbf{A}(\mathbf{p})\,\mathbf{u} = \mathbf{b}(\mathbf{p}). \tag{9}$$

For simplicity, we assume that the matrix $\mathbf{A}$ and the vector $\mathbf{b}$ possess continuous partial derivatives in a neighborhood of $\mathbf{q}$. If $\mathbf{A}(\mathbf{q})$ is regular, i.e. $\det \mathbf{A}(\mathbf{q}) \neq 0$, then

$$\mathbf{u} = \mathbf{F}(\mathbf{p}) = \mathbf{A}(\mathbf{p})^{-1}\mathbf{b}(\mathbf{p}) \tag{10}$$

for all $\mathbf{p}$ in a neighborhood of $\mathbf{q}$. This proves that $\mathbf{F}$ and thus $\mathbf{u}$ are Fréchet differentiable with respect to $\mathbf{p}$ at $\mathbf{q}$. Applying the chain rule further shows that

$$\frac{\partial \mathbf{A}}{\partial p_j}(\mathbf{q})\,\mathbf{u} + \mathbf{A}(\mathbf{q})\frac{\partial \mathbf{u}}{\partial p_j} = \frac{\partial \mathbf{b}}{\partial p_j}(\mathbf{q})$$

and thus

$$\frac{\partial \mathbf{u}}{\partial p_j}(\mathbf{q}) = \mathbf{A}(\mathbf{q})^{-1}\left(\frac{\partial \mathbf{b}}{\partial p_j}(\mathbf{q}) - \frac{\partial \mathbf{A}}{\partial p_j}(\mathbf{q}) \cdot \mathbf{A}(\mathbf{q})^{-1}\mathbf{b}(\mathbf{q})\right). \tag{11}$$

Note that the numerical evaluation of this formula requires only *one* decomposition of $\mathbf{A}$, but it needs the partial derivatives of $\mathbf{A}$ and $\mathbf{b}$, or at least some numerical approximations thereof.

An alternative way would be to differentiate (10) numerically, i.e.

$$\frac{\partial \mathbf{u}}{\partial p_j}(\mathbf{q}) \approx \frac{1}{\delta}\Big(\mathbf{A}(\mathbf{p}+\delta\mathbf{e}_j)^{-1}\mathbf{b}(\mathbf{p}+\delta\mathbf{e}_j) - \mathbf{A}(\mathbf{p})^{-1}\mathbf{b}(\mathbf{p})\Big). \tag{12}$$

This approach *does not* need the partial derivatives of $\mathbf{A}$ and $\mathbf{b}$. On the other hand, it requires two decompositions of $\mathbf{A}$, one at $\mathbf{q}$ and one at $\mathbf{q}+\delta\mathbf{e}_j$.

To illustrate the above ideas, we consider

$$\mathbf{A} = \begin{bmatrix} 1+p & p^2 \\ 1-p & 2-p \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} p^2-1 \\ 2p \end{bmatrix}. \tag{13}$$

For this simple example, the exact solution $\mathbf{u} = [u_1, u_2]^\mathsf{T}$ of $\mathbf{A}\mathbf{u} = \mathbf{b}$ can be calculated easily and is given by

$$u_1 = -\frac{2 - p - 2p^2 + 3p^3}{2 + p - 2p^2 + p^3}, \qquad u_2 = \frac{1 + p + p^2 + p^3}{2 + p - 2p^2 + p^3}.$$

At $p = 1$, the solution and its sensitivity are thus

$$\mathbf{u}(1) = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \qquad \frac{\partial \mathbf{u}}{\partial p}(1) = \begin{bmatrix} -2 \\ 3 \end{bmatrix}.$$

We compare these exact values with numerically obtained approximations based on methods (11) and (12). We work in IEEE double precision arithmetic with ACC $\approx 10^{-16}$ and approximate the partial derivatives that are required in (11) numerically with one-sided differences. The errors of the corresponding methods as a function of the discretization parameter δ are given in Table 1. As expected, both methods yield best results for $\delta \approx 10^{-8}$.

δ	10^{-4}	10^{-6}	10^{-8}	10^{-10}	10^{-12}
errors of (11)	$5.000 \cdot 10^{-5}$	$4.998 \cdot 10^{-7}$	$2.191 \cdot 10^{-8}$	$2.983 \cdot 10^{-7}$	$3.205 \cdot 10^{-4}$
errors of (12)	$3.162 \cdot 10^{-4}$	$3.162 \cdot 10^{-6}$	$2.082 \cdot 10^{-8}$	$2.983 \cdot 10^{-7}$	$3.205 \cdot 10^{-4}$

Table 1. Numerical approximation of the sensitivity of problem (13) with respect to p. The errors of methods (11) and (12) are displayed for different values of δ, respectively. With one-sided differences, best results are obtained for $\delta \approx 10^{-8}$.

4.2 Eigenvalue problems

Let $\mathbf{A}(p) = \mathbf{A}^\mathsf{T}(p)$ be a symmetric matrix that depends smoothly on a parameter $p \in \mathbb{R}$ and consider the eigenvalue problem

$$\mathbf{A}\mathbf{u} = \lambda\,\mathbf{u}, \qquad \mathbf{u}^\mathsf{T}\mathbf{u} = 1. \tag{14}$$

Further, for fixed q, let $\mu = \lambda(q)$ be a simple eigenvalue of (14) with a corresponding eigenvector $\mathbf{v} = \mathbf{u}(q)$ of unit length. Differentiating (14) with respect to p at q gives the *sensitivity equations*

$$\mathbf{A}'(q)\,\mathbf{v} + \mathbf{A}(q)\,\mathbf{u}'(q) = \lambda'(q)\,\mathbf{v} + \mu\,\mathbf{u}'(q), \qquad \mathbf{v}^\mathsf{T}\mathbf{u}'(q) = 0. \tag{15}$$

Here, primes denote differentiation with respect to the argument p. Multiplying the first equation from the left-hand side with $\mathbf{v}^\mathsf{T}$ and using (14) shows that

$$\lambda'(q) = \mathbf{v}^\mathsf{T}\mathbf{A}'(q)\,\mathbf{v}. \tag{16a}$$

The sensitivity $\mathbf{u}'(q)$ is then given as the unique solution of the linear system

$$\begin{bmatrix} \mathbf{A}(q) - \mu\,\mathbf{I} \\ \mathbf{v}^\mathsf{T} \end{bmatrix} \mathbf{u}'(q) = \begin{bmatrix} \lambda'(q)\,\mathbf{v} - \mathbf{A}'(q)\,\mathbf{v} \\ 0 \end{bmatrix}. \tag{16b}$$

For the purpose of numerical illustration, we consider the matrix

$$\mathbf{A}(p) = \begin{bmatrix} 1-p & (1+p)^{-1} \\ (1+p)^{-1} & 1+p \end{bmatrix}.$$

At $p = 0$, it has the eigenvalues 0 and 2. The latter takes the form

$$\lambda(p) = \frac{1+p+\sqrt{1+p^2+2p^3+p^4}}{1+p}.$$

for p near 0. A straightforward calculation from this explicit representation gives the sensitivities

$$\lambda'(0) = -1, \qquad \mathbf{u}'(0) = \frac{\sqrt{2}}{4}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.3535533905932738... \\ 0.3535533905932738... \end{bmatrix}.$$

In general, however, an explicit representation of λ as a function of p is not available. Then (16) combined with a numerical differentiation of $\mathbf{A}$ is a good alternative to calculate the sensitivities. Choosing $\delta = 10^{-8}$ we obtain in IEEE arithmetic the following numerical approximations

$$\lambda'(0) \approx -\underline{0.9999999}883714141, \qquad \mathbf{u}'(0) \approx \begin{bmatrix} -\underline{0.3535533390}4071787 \\ \underline{0.3535533390}4071790 \end{bmatrix}.$$

Since we used one-sided differences, the accuracy is about 8 digits.

4.3 Systems of nonlinear equations

Let us next discuss the case when our problem takes the form of a system of nonlinear equations

$$\mathbf{G}(\mathbf{u}, \mathbf{p}) = \mathbf{0}. \tag{17}$$

For this we assume that $\mathbf{G}$ has continuous partial derivatives in a neighborhood of $(\mathbf{v}, \mathbf{q})$ and that $\mathbf{G}(\mathbf{v}, \mathbf{q}) = \mathbf{0}$. If

$$\det \frac{\partial \mathbf{G}}{\partial \mathbf{u}}(\mathbf{v}, \mathbf{q}) \neq 0,$$

then the implicit function theorem provides the existence of a function $\mathbf{u} = \mathbf{u}(\mathbf{p})$ which uniquely solves (17) in a neighborhood of $\mathbf{q}$ with $\mathbf{u}(\mathbf{q}) = \mathbf{v}$. Further, this function is differentiable with respect to $\mathbf{p}$ and

$$\frac{\partial \mathbf{u}}{\partial p_j}(\mathbf{q}) = -\left(\frac{\partial \mathbf{G}}{\partial \mathbf{u}}(\mathbf{v}, \mathbf{q})\right)^{-1} \frac{\partial \mathbf{G}}{\partial p_j}(\mathbf{v}, \mathbf{q}). \tag{18}$$

Note that the implicit function theorem actually holds under slightly weaker assumptions on $\mathbf{G}$, see [7, Chap. 5.2].

For an evaluation of (18) we can employ numerical differentiation to approximate the partial Jacobians

$$\frac{\partial \mathbf{G}}{\partial \mathbf{u}}(\mathbf{v}, \mathbf{q}) \qquad \text{and} \qquad \frac{\partial \mathbf{G}}{\partial p_j}(\mathbf{v}, \mathbf{q}),$$

and then solve a linear system of equations.

4.4 Ordinary differential equations

Suppose that we have to solve the following system of time dependent differential equations for $t > 0$

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}t} = \mathbf{f}(t, \mathbf{y}, \mathbf{p}), \qquad \mathbf{y}(0) = \mathbf{a}. \tag{19}$$

If the right-hand side $\mathbf{f}$ is Lipschitz continuous with respect to $\mathbf{y}$, the solution at time t is a function of the parameters and the initial data

$$\mathbf{y}(t) = \mathbf{F}(t, \mathbf{p}, \mathbf{a}).$$

In general, however, this function $\mathbf{F}$ is not known explicitly.

We are again interested in obtaining the sensitivities of the solution $\mathbf{y}(t)$ with respect to the parameters $\mathbf{p}$. Note that the sensitivity of the solution with respect to initial data can be studied in a similar way. One possibility is to consider $\mathbf{v}(t) = \mathbf{y}(t) - \mathbf{a}$ which transforms initial data into parameters. Applying this transformation to (19) results in the equivalent problem

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \mathbf{f}(t, \mathbf{v} + \mathbf{a}, \mathbf{p}) =: \mathbf{g}(t, \mathbf{v}, \mathbf{a}, \mathbf{p}), \qquad \mathbf{v}(0) = \mathbf{0}$$

with a parameter dependent right-hand side $\mathbf{g}$ and zero initial data.

We will explain two strategies for computing the sensitivity of $\mathbf{y}(t)$ with respect to the parameter p_j. *External differentiation* enjoys great popularity in engineering, but it may suffer from a loss of accuracy. Our method of choice is therefore *internal differentiation*, which is based on solving sensitivity equations. It gives more accurate results for a comparable amount of computational effort.

External differentiation

In view of Sect. 3.2, the sensitivity of $\mathbf{y}(t)$ with respect to p_j can be approximated by numerical differentiation

$$\mathbf{z}(t) = \frac{\partial \mathbf{y}(t)}{\partial p_j} \approx \frac{\mathbf{F}(t, \mathbf{p} + \delta \mathbf{e}_j, \mathbf{a}) - \mathbf{F}(t, \mathbf{p}, \mathbf{a})}{\delta}.$$

In practice, each evaluation of $\mathbf{F}$ corresponds to solving (19) numerically for a given tolerance TOL. Let $\mathbf{y}_n$ and $\widetilde{\mathbf{y}}_n$ denote the numerical solutions of (19) at time t with parameters $\mathbf{p}$ and $\widetilde{\mathbf{p}} = \mathbf{p} + \delta \mathbf{e}_j$, respectively, and let

$$\widehat{\mathbf{z}}_n = \frac{\widetilde{\mathbf{y}}_n - \mathbf{y}_n}{\delta}. \tag{20}$$

An analysis along the lines of Sect. 3.2 then shows that

$$\|\mathbf{z}(t) - \widehat{\mathbf{z}}_n\| = \mathcal{O}(\delta) + \mathcal{O}\left(\frac{\mathrm{TOL}}{\delta}\right). \tag{21}$$

For large tolerances, which are typical in engineering applications, the attainable accuracy of the numerical approximation $\widehat{\mathbf{z}}_n$ can thus be quite poor. This is a severe drawback of external differentiation.

Internal differentiation

If the partial derivatives of $\mathbf{f}$ with respect to $\mathbf{y}$ and $\mathbf{p}$ are continuous in a neighborhood of the exact solution, then the sensitivity

$$\mathbf{z}(t) = \frac{\partial \mathbf{y}(t)}{\partial p_j}, \quad j \in \{1, \ldots, n\}$$

is continuous and fulfils the so-called *variational equations*

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}t} = \frac{\partial \mathbf{f}}{\partial \mathbf{y}}(t, \mathbf{y}, \mathbf{p}) \cdot \mathbf{z} + \frac{\partial \mathbf{f}}{\partial p_j}(t, \mathbf{y}, \mathbf{p}), \qquad \mathbf{z}(0) = \mathbf{0}. \tag{22}$$

This well-known result goes back to Gronwall [4]. For a modern presentation, we refer to the monograph [6, Chap. I.14].

The variational equations are formally obtained by differentiating (19) with respect to the parameters. Note that the explicit calculation and implementation of the expressions appearing on the right-hand side of (22) can be a tedious task. We therefore strongly recommend to replace (22) by the following approximation that is obtained from (19) by *numerical differentiation*

$$\frac{\mathrm{d}\widehat{\mathbf{z}}}{\mathrm{d}t} = \frac{\mathbf{f}(t, \mathbf{y} + \delta\widehat{\mathbf{z}}, \mathbf{p} + \delta\mathbf{e}_j) - \mathbf{f}(t, \mathbf{y}, \mathbf{p})}{\delta}, \qquad \widehat{\mathbf{z}}(0) = \mathbf{0}. \tag{23}$$

The variational equations (23) together with the original problem (19) form a coupled system of differential equations which has to be solved simultaneously for $\mathbf{y}$ and $\widehat{\mathbf{z}}$. Let $\widehat{\mathbf{z}}_n$ denote the numerical solution of (23) at time t, obtained with a prescribed tolerance TOL. An analysis along the lines of Sect. 3.2 shows that

$$\|\mathbf{z}(t) - \widehat{\mathbf{z}}_n\| = \mathcal{O}(\mathrm{TOL}) + \mathcal{O}(\delta) + \mathcal{O}\left(\frac{\mathrm{ACC}}{\delta}\right). \tag{24}$$

Here, ACC denotes the accuracy of the function evaluations in (19) and (23).

Comparing (21) with (24) reveals the superiority of internal differentiation. Most accurate results are obtained for the choice

$$\delta \approx \sqrt{\mathrm{ACC}} \approx \mathrm{TOL}, \tag{25}$$

when the three error terms in (24) are of the same order of magnitude.

Applications

The sensitivity of the solution of a differential equations with respect to initial values or parameters is frequently needed in combination with Newton's method. For example, in finite element calculations with a nonlinear constitutive equation, a consistent tangent operator is required to achieve quadratic convergence in Newton's method. If the constitutive equation consists of a system of ordinary differential equations, this requires the sensitivity of the

solution with respect to parameters, namely the derivative of the computed stresses with respect to the prescribed strain increments. In [3], we have used internal differentiation for a successful implementation of a hypoplastic constitutive equation in a finite element code.

Further applications are encountered in computational dynamical systems. A typical problem there is the numerical computation of periodic orbits which requires the derivative of the solution with respect to the initial data, see [2].

4.5 Elliptic boundary value problems

When elliptic boundary value problems are formulated in the language of *functional analysis*, they take the form of linear or nonlinear equations. At first glance, the sensitivity analysis is therefore quite close to that of Sects. 4.1 and 4.3. The mathematical framework, however, is much more involved since the analysis takes now place in abstract function spaces. The technical details are well explained in the monograph [8] and need not be repeated here.

On the other hand, the applied engineer should not be deterred by the above mentioned abstract framework. Although the *mathematical analysis* cannot be done without it, numerical examples nevertheless can be handled to some extent without going into the intricacies of functional analysis.

To be more specific, we consider a nonlinear model for the buckling of an elastic rod, see [1, Chap. 1.3]. We describe the shape of the rod in problem adapted coordinates by a parameterized curve $[x(s), u(s)]^\mathsf{T}$ in the (x, u)-plane. As parameter s, we choose the arc length of the curve. One end of the rod is kept fixed at the origin whereas the other one is allowed to move along the x-axis, subject to a horizontal force $\mathbf{p} = [p, 0]^\mathsf{T}$ applied at this end, see Fig. 1.

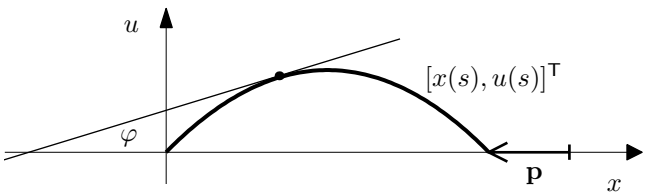


Fig. 1. The deformed rod $[x(s), u(s)]^\mathsf{T}$ under constant axial loading $\mathbf{p}$ from the right. The oriented angle between the positive x-axis and the tangent of the curve is denoted by φ.

Let φ denote the oriented angle between the positive x-axis and the tangent of the curve $[x(s), u(s)]^\mathsf{T}$. We then have the following relations

$$\frac{\mathrm{d}x}{\mathrm{d}s} = \cos\varphi(s), \quad \frac{\mathrm{d}u}{\mathrm{d}s} = \sin\varphi(s), \qquad x(0) = u(0) = 0. \tag{26}$$

In standard elasticity, the curvature $\frac{\mathrm{d}\varphi}{\mathrm{d}s}$ is assumed to be proportional to the moment $p{\cdot}u(s)$ of the force. Due to (26), the deformation of the rod is therefore described by the following integro-differential equation

$$\frac{\mathrm{d}\varphi}{\mathrm{d}s}(s) + \lambda \int_0^s \sin\varphi(\sigma)\,\mathrm{d}\sigma = 0, \tag{27}$$

where the constant λ is proportional to p. We differentiate (27) with respect to s to obtain a second-order nonlinear boundary value problem for φ

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}s^2}(s) + \lambda\,\sin\varphi(s) = 0, \qquad \frac{\mathrm{d}\varphi}{\mathrm{d}s}(0) = \frac{\mathrm{d}\varphi}{\mathrm{d}s}(1) = 0. \tag{28}$$

This is exactly the same equation that models the mathematical pendulum. For any positive integer k and $(k-1)^2\pi^2 < \lambda < k^2\pi^2$, problem (28) has k different solutions which are all locally unique, see [1]. Let

$$\psi(s,\lambda) = \frac{\partial\varphi}{\partial\lambda}(s,\lambda)$$

denote the sensitivity of φ with respect to λ. Then ψ satisfies the sensitivity equation

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}s^2}(s) + \lambda\cos\varphi(s)\cdot\psi(s) = -\sin\varphi(s), \qquad \frac{\mathrm{d}\psi}{\mathrm{d}s}(0) = \frac{\mathrm{d}\psi}{\mathrm{d}s}(1) = 0, \tag{29}$$

which is obtained from (28) by formal differentiation. For the numerical solution of (28) and (29), we use standard finite differences on an equidistant grid with 200 meshpoints. The resulting nonlinear system for (28) is solved by Newton's method. We set $\lambda = 72$ and look for the unique solution with two buckles. The linearization of (28) motivates to start the iteration with $\varphi(s) = \beta\cos(2\pi s)$. Taking $\beta = 2$, the method converges to machine precision after five steps. In order to obtain second-order approximations for the positions and their sensitivities, we use (26) and discretize the integrals

$$\begin{aligned} x(s) &= \int_0^s \cos\varphi(\sigma)\,\mathrm{d}\sigma, & \xi(s) &= \frac{\partial x(s)}{\partial\lambda} = -\int_0^s \sin\varphi(\sigma)\,\psi(\sigma)\,\mathrm{d}\sigma,\\ u(s) &= \int_0^s \sin\varphi(\sigma)\,\mathrm{d}\sigma, & z(s) &= \frac{\partial u(s)}{\partial\lambda} = \int_0^s \cos\varphi(\sigma)\,\psi(\sigma)\,\mathrm{d}\sigma \end{aligned}$$

with the trapezoidal rule. The numerical results are displayed in Figs. 2 and 3.

In this example, the information on the Jacobian matrix is needed in the Newton iterations for solving (28). The additional costs for solving the sensitivity equations (29) are therefore negligible. This situation is quite typical of nonlinear problems.

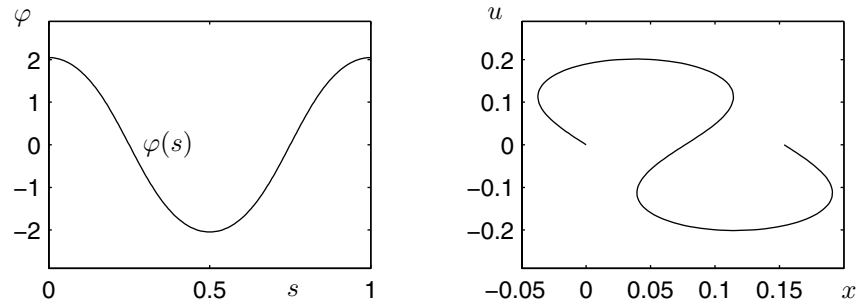


Fig. 2. The locally unique solution $\varphi(s)$ of (28) for $\lambda = 72$ (left) and the deformed rod $[x(s), u(s)]^{\mathsf{T}}$ with two buckles (right).

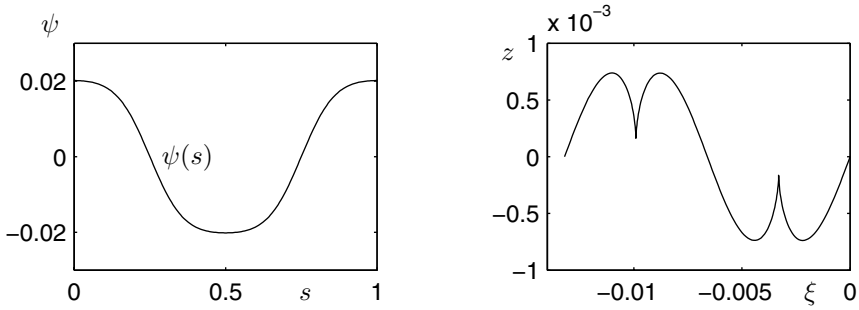


Fig. 3. The sensitivity $\psi(s)$ of the angle $\varphi(s)$ (left) and the sensitivity $[\xi(s), z(s)]^{\mathsf{T}}$ of the deformed rod $[x(s), u(s)]^{\mathsf{T}}$ (right) for $\lambda = 72$.

We close this section with a simple but yet important remark. We illustrate with a simple example how shape parameters can be transformed to parameters of the right-hand side function. For this, let

$$u_{xx}(x) = f\big(x, u(x), u_x(x)\big), \quad 0 < x < \ell, \qquad u(0) = u(\ell) = 0$$

be a nonlinear second-order boundary value problem on the interval $[0, \ell]$. Here subscripts denote corresponding differentiations. In order to determine the sensitivity of the solution u with respect to the length ℓ, one might consider the new coordinates

$$x = \ell\xi, \qquad \frac{\mathrm{d}^2}{\mathrm{d}x^2} = \frac{1}{\ell^2}\frac{\mathrm{d}^2}{\mathrm{d}\xi^2}, \qquad z(x) = w(\xi).$$

This transformation yields the equivalent boundary value problem

$$w_{\xi\xi}(\xi) = \ell^2 f\big(\ell\xi, w(\xi), \ell^{-1} w_\xi(\xi)\big) =: g\big(\xi, w(\xi), w_\xi(\xi), \ell\big), \quad 0 < \xi < 1$$

with homogeneous boundary conditions $w(0) = w(1) = 0$ on the constant interval $[0, 1]$. The transformed problem, however, has a right-hand side function g that depends on the parameter ℓ.

5 Conclusions

In this article, we discussed sensitivity analysis for deterministic mathematical models. For approximating the sensitivities of a particular solution, we emphasized the method of internal differentiation. This corresponds to differentiating the underlying model first analytically with respect to the parameters and then approximating the arising derivatives numerically. The resulting equations, which define the sensitivities, are linearizations of the underlying equations. In particular, they possess a structure similar to that of the underlying equation which in turn allows for an efficient implementation.

When dealing with nonlinear problems, some information on the Jacobian matrix is typically needed in Newton-type iterations. The additional costs for solving the sensitivity equations for a few parameters are then often negligible.

Acknowledgement. I am most grateful to Ernst Hairer who introduced me to sensitivity analysis during my postdoc in the late eighties at the University of Geneva. I will always remember his wonderful lectures on computational dynamical systems where he emphasized the advantages of internal differentiation.

References

[1] S.-N. Chow, J.H. Hale. *Methods of Bifurcation Theory.* Springer, New York, 1982.

[2] P. Deuflhard. Computation of periodic solutions of nonlinear ODEs. *BIT* 24 (1984) 456–466.

[3] W. Fellin, A. Ostermann. Consistent tangent operators for constitutive rate equations. *Int. J. Numer. Anal. Meth. Geomech.* 26 (2002) 1213–1233.

[4] T.H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Ann. Math.* 20 (1919) 292–296.

[5] A. Griewank and G.F. Corliss (eds). *Automatic Differentiation of Algorithms: Theory, Implementation, and Application.* SIAM, Philadelphia, 1991.

[6] E. Hairer, S.P. Nørsett, G. Wanner. *Solving Ordinary Differential Equations I. Nonstiff Problems.* Springer, Berlin, 1993[2].

[7] J.M. Ortega and W.C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, New York, 1970.

[8] L.G. Stanley, D.L. Stewart. *Design Sensitivity Analysis. Computational issues of Sensitivity Equation Methods.* SIAM, Philadelphia, 2002.

[9] Ch. Überhuber. *Computer-Numerik 1.* Springer, Berlin, 1995.

Difficulties related to numerical predictions of deformations

Ivo Herle

Institut für Geotechnik, Technische Universität Dresden

1 Introduction

Outputs of numerical calculations are rather impressive today. Contour plots of stresses and deformations contribute to the fast interpretation of results and create the impression of being perfect and reliable. Nevertheless, many of the end-users overlook the complicated process leading to such results. This process can be divided into several basic steps:

1. *Simplification of the reality*
   (choosing important variables, geometry, selection of substantial aspects of the problem and disregarding minor ones)
2. *Discretization*
   (space and time discretization — element size, time step, boundary conditions, construction details)
3. *Constitutive model*
   (framework for calculation of strains, selection of the appropriate material description, model for interfaces, calibration of the parameters, determination of the initial state)
4. *Mathematical and numerical aspects*
   (type of time integration, equation solver, iteration scheme, well-posedness).

All steps are equally important. If one step is not correctly implemented, the whole result may become wrong. All steps compose a chain which fails at the weakest link. The outlined steps are valid for any calculation. It is not possible to distinguish seriously between calculations of deformation and limit states. Even in problems related to deformation the material usually reaches the limit state of stress at several points! Moreover, we need deformations in order to calculate limit states since the subsoil and structure interact and we

Published originally in: S. Springman (ed.), *Constitutive and Centrifuge Modelling: Two Extremes,* Workshop in Monte Verita, Switzerland. A. A. Balkema Publishers, 2002, pp. 239–248.

cannot distinguish between action and reaction (although it is a wish of code designers). The necessity of an overall approach was already discussed e.g. in the Rankine lecture by Roscoe: "Influence of strains in soil mechanics" [22].

2 Predictions vs. measurements

The only way to evaluate numerical simulations is to compare them with measurements and observations. However, opportunities to compare measurements with true predictions (so-called class A predictions [15]) are very rare. They teach us that in most cases our predictions are rather far away from measured values.

It is useful briefly to review some competitions on predictions in geotechnical engineering, with emphasis on deformation calculations (all competitions were designed as plane strain problems):

1. *MIT trial embankment* [34]
   A normally consolidated soft clay layer beneath a trial embankment controlled the deformation behaviour. Laboratory experiments and field measurements for the first construction stage up to 12.2 m height were done prior to the prediction calculations. Predictions of deformations, pore pressures and maximum additional height of the embankment at subsequent rapid filling to failure were received from ten groups. There was a large scatter of the numerical results: e.g. additional height of fill at failure ranged between 2.4 m and 8.2 m (measured: 6.4 m), additional settlement at the ground surface in the centre of the embankment due to 1.8 m of fill — between 1.9 cm and 34.8 cm (measured: 1.7 cm), additional horizontal movement beneath the middle of the embankment slope due to 1.8 m of fill ranged between 0.4 cm and 21.8 cm (measured: 1.3 cm), etc. One of the best predictions, based on the Modified Cam Clay model [34], was very good with respect to pore pressures but still less accurate for deformations.
2. *Excavation in sand* [33]
   A 5 m deep excavation in a homogeneous sand layer above the groundwater level in Germany near Karlsruhe was supported by a sheet pile wall. Struts were installed at 1.5 m depth. An additional surcharge was placed at the ground surface after the excavation and finally the struts were loosened. In situ and laboratory soil investigations were performed prior to the excavation. Results from 43 predictions concerned horizontal displacements of the wall, vertical displacements at the ground surface, earth pressures on the wall and bending moments in the wall. The comparison with measured values was very disappointing. Especially worrying was that displacements were often predicted wrong in a qualitative sense (i.e. in the opposite direction).

3. *Excavation in clay* [16]
   A 7 m deep excavation in soft clay and peat with a high groundwater level was located in the Netherlands near Rotterdam. The purpose of this field test was to complement the field test in Karlsruhe with another type of soil. Again, predicted values were often far away from the measured ones.
4. *Tunnel and deep excavation* [24]
   Both benchmark calculations were intended to check only the numerical aspects and were not accompanied by any measurements. Geometry, constitutive model (Mohr-Coulomb), model parameters and boundary and initial conditions were specified. Still the results diverge in a considerable manner: In case of the tunnel excavation, most calculations predicted nearly the same surface settlement only for one step excavation. Applying two steps, the maximum surface settlement ranged from 3.3 cm to 5.8 cm (10 calculations). Even worse was the case of a deep excavation supported by a diaphragm wall with two rigid struts. Vertical displacement of the surface behind the wall ranged between 1.3 cm and 5.2 cm and the horizontal displacement of the head of wall oscillated between −1.3 and 0.5 cm (12 calculations).
5. *Excavation with a tied back diaphragm wall* [25]
   A geometrical specification of the 30 m deep excavation in Berlin sand supported by a diaphragm wall anchored at three levels was accompanied by a detailed specification of the simulation stages and basic laboratory experiments (index tests, oedometer and triaxial tests). Even after "filtering out" the most extreme and questionable results, the values of the maximum horizontal displacements ranged between 0.7 cm and 5.7 cm, and the surface settlements between 5.0 cm and −1.5 cm. Large scatter was obtained even for results from the same code using the same material model.
6. *Vertically loaded small model footings on sand* [10, 27, 30]
   Small model strip footings ($B = 1.0, 2.5, 5.0$ and 10 cm) were placed on the surface of a dense sand and loaded vertically. Load-displacement curves were calculated prior to the experiments. Due to small dimensions of the footings, a pronounced scale effect was involved. The overall results of the prediction competition [29] have never been published. However, a personal communication to the organizer[1] confirmed that the predicted load-displacement curves [10], based on the polar hypoplastic model, were the best ones. Although the scale effect and the strains at the peak of the load-displacement curve were reproduced very well, the maximum load was overpredicted by almost 40%.

The above overview of several prediction competitions shows an unsatisfactory state of the art. It seems that we are not able to predict reliably the behaviour even of the simplest geotechnical structures.

[1] Personal communication from Professor F. Tatsuoka at the IS-Nagoya, 1997.

Leaving aside the question as to why there are, in reality, only a few geotechnical structures which fail (probably due to very high values of safety), I will try to analyze in more detail the steps No. 3 and 4 of Section 1, which describes the process of the creation of a numerical model.

3 Constitutive model

From the point of view of a geotechnical engineer, the question of the constitutive model and of its calibration is the crucial one. A large number of soil models exist but still no general agreement on their quality (and how the quality can be defined [13]) and on the range of their applicability. The second criterion is probably more important since every model has some limitations which are unfortunately usually hidden by the designer of the model. Moreover, the complexity of advanced models often prevents other potential users from implementing and checking them[2]. This results in applying but very simple models in most practical cases.

It is not the aim of this contribution to evaluate the suitability of particular soil models in describing the soil behaviour. Alongside prediction competitions on boundary value problems (BVP), prediction competitions on soil behaviour in element tests have also been performed [6, 23]. Nevertheless, there are several topics closely related to constitutive models which become obvious first in calculations of BVP.

3.1 Calculation of strains

Constitutive models use strains instead of displacements in order to be scale independent. However, there is no unique definition of strain. In soil mechanics, we usually consider small strains which refer to a very small change of lengths and angles: in 1-D we write $\varepsilon = (l - l_0)/l_0$ with l_0 being the initial length and l the actual length. This approach is the simplest one, but mostly it cannot be justified. Using small strains, we refer to a reference (initial) geometrical configuration without taking into account geometrical effects. Although it may often seem that the overall deformations are small, local strains

[2] The question of checking the models is extremely important. There is a useful analogy with the open source software e.g. under the GNU license. This software is accessible to everyone who is interested and can read the source code. Consequently, a large number of people test the software and report the bugs as well as propose improvements. It is a public secret that many open source codes are today better and more reliable than commercial software which can be supported only by limited financial means on development and testing. Unfortunately, the situation in the development of constitutive models resembles more of a struggle between advertisement agencies. Users are merely convinced on superiority of each particular model even without understanding it and having a possibility to check it.

can reach high values in some cases, consider e.g. cavity expansion problem or shear zones. Another class of problems sensitive to the strain calculation refers to situations where the stiffness moduli are comparable with the stress level. We should always bear in mind that using highly non-linear constitutive models, which correspond better to the soil behaviour than the linear ones, even a small change in strains can yield a substantial change in stresses.

3.2 Strain decomposition

The vast majority of constitutive models assumes the strain decomposition into elastic and plastic parts thus implicitly assuming the existence of an elastic range. Elastic strains are recoverable (reversible) after so-called unloading. Usually linear elasticity is assumed. But it is not simple to detect an elastic range even in highly sophisticated laboratory experiments in the range of very small deformations [28].

One often argues that this decomposition is not so important for monotonic deformation processes where only one loading direction appears. However, if the strain decomposition is not important, so why do we do it? Do we want to forget the basic rule ,,*Keep the model as simple as possible (but not simpler)*"? Moreover, the argument mentioned above is not true. Practically all boundary value problems involve regions of simultaneous loading and other regions of unloading. Probably no BVP exist which include only one loading direction. A classical demonstration can be a tunnel excavation or a deep excavation [8, 31]. Both, loading and unloading is also present in coupled problems with the generation and the subsequent dissipation of pore water pressures [14].

The problem of the elastic range is closely related to the non-linearity of the soil behaviour. Although, according to experimental results, the elastic strain range is almost vanishing, many models presuppose quite a large domain of (linear) elastic strains (except for very sophisticated elasto-plastic models which, on the one hand can reproduce the soil behaviour in such a small elastic range, but on the other hand are extremely complex and remain a topic of the academic research of few specialists). Even rather advanced models can suffer from the effect of an unrealistic elastic range, see Figure 1 [7].

An example of failing to reproduce the observed deformation pattern by applying linear elasticity may be the class A prediction of displacements due to the excavation for the underground car park at the palace of Westminster [3]. Although the horizontal displacements at the surface behind the wall were in a tolerable range, the predicted vertical displacements had the opposite sign to the measured values. This resulted in the wrong prediction of the direction of rotation of the Big Ben Tower. An additional calculation using a bilinear model with ten times higher initial stiffness yielded a substantial improvement [26].

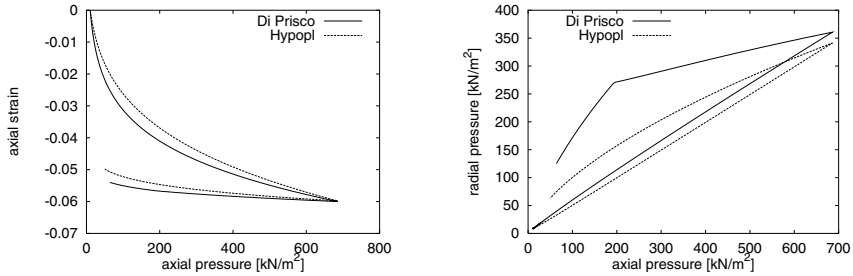


Figure 1. Elastic range of *di Prisco model* yields an unrealistic oedometer stress path during unloading (right)

3.3 Hypoplasticity

Let me give an example of a non-linear constitutive model without strain decomposition. There is a class of so-called hypoplastic models which abandon the strain decomposition and consequently do not use notions of yield and potential surfaces, flow rules etc. [12, 5]. The basic idea can be explained for a 1-D case. The stress rate follows from

$$\dot{\sigma} = E_1 \dot{\varepsilon} + E_2 |\dot{\varepsilon}| \tag{1}$$

where $E_1 > E_2 > 0$ are the stiffness moduli. Defining "loading" as $\dot{\varepsilon} < 0$, we have $|\dot{\varepsilon}| = -\dot{\varepsilon}$ and therefore

$$\dot{\sigma} = (E_1 - E_2)\, \dot{\varepsilon} \tag{2}$$

whereas for "unloading", $|\dot{\varepsilon}| = \dot{\varepsilon} > 0$ and

$$\dot{\sigma} = (E_1 + E_2)\, \dot{\varepsilon}\ . \tag{3}$$

Obviously, the stiffness for loading is smaller than for unloading, as usually observed.

Considering two stress components, σ and τ, and the corresponding strain components, ε and γ, a hypoplastic equation can be written as

$$\begin{Bmatrix} \dot{\sigma} \\ \dot{\tau} \end{Bmatrix} = \begin{bmatrix} L_{\sigma\sigma} & L_{\sigma\tau} \\ L_{\tau\sigma} & L_{\tau\tau} \end{bmatrix} \begin{Bmatrix} \dot{\varepsilon} \\ \dot{\gamma} \end{Bmatrix} + \begin{Bmatrix} N_\sigma \\ N_\tau \end{Bmatrix} \sqrt{\dot{\varepsilon}^2 + \dot{\gamma}^2} \tag{4}$$

with L_{ij} and N_i being components of the tangent stiffness matrix. A particular representation of this equation [9] may be

$$\dot{\sigma} = c_2 \sigma \left(\dot{\varepsilon} + c_1^2 \frac{\sigma \dot{\varepsilon} + \tau \dot{\gamma}}{\sigma^2}\, \sigma - c_1 \sqrt{\dot{\varepsilon}^2 + \dot{\gamma}^2} \right) \tag{5}$$

$$\dot{\tau} = c_2 \sigma \left(\dot{\gamma} + c_1^2 \frac{\sigma \dot{\varepsilon} + \tau \dot{\gamma}}{\sigma^2}\, \tau - 2 c_1 \frac{\tau}{\sigma} \sqrt{\dot{\varepsilon}^2 + \dot{\gamma}^2} \right) \tag{6}$$

with c_1 (=1/tanφ, φ is a friction angle) and c_2 being the model parameters. With such a very simple model one can obtain many features of the soil behaviour, including non-linearity, shear-volumetric coupling or stress-dependent stiffness.

In the general case, the hypoplastic model includes an objective stress rate tensor $\mathring{\mathbf{T}}$ instead of $\dot{\sigma}$ and the stretching tensor $\mathbf{D}$ instead of $\dot{\varepsilon}$. The modulus $|\dot{\varepsilon}|$ is replaced by $\|\mathbf{D}\| = \sqrt{\mathbf{D} : \mathbf{D}}$. The matrices L_{ij} and N_i become tensors of the fourth and the second order, respectively, and depend on the effective stress tensor (and its invariants) and on the void ratio [32].

3.4 Model calibration

Quality of experiments

Calibration of the model parameters is mostly based on laboratory experiments. Nowadays, the quality of test results is considered to be high (although the precision of measurements is limited) and consequently many researchers try to check constitutive equations by reproducing the experimental curves numerically as closely as possible. However, looking more in detail at basic element tests performed in different apparatuses, substantial scatter of experimental results can be observed [2, 18]. This is not a point against experiments (!), it merely suggests that the development and the validation of constitutive models should be based mainly on the qualitative aspects of the soil behaviour. The realistic solution of boundary value problems does not depend only on the reproduction of calibration curves.

Oversimplified models

Material parameters are always linked to a particular constitutive model. In other words, if the constitutive model does not describe particular aspects of the material behaviour properly, one cannot determine corresponding model parameters.

A classical example of this problem is the calibration of the Mohr-Coulomb model. Non-linear stress-strain curves are approximated by a straight line underpredicting thus the initial stiffness and overpredicting the tangential stiffness at higher strains. Moreover, one often does not inspect the volumetric behaviour which is uniquely linked to the stress-strain curve. Such an inspection would show that the initial (elastic) volumetric compression is grossly overpredicted (Figure 2).

Recognizing the above mentioned problem, some people try to fit the parameters of a constitutive model directly to in situ measurements, referring to the observational method [20]. Beside the fact that this approach disables a class A prediction, it usually fails to reproduce the realistic pattern of deformation. Regarding e.g. a tunnel excavation, one manages to fit the surface

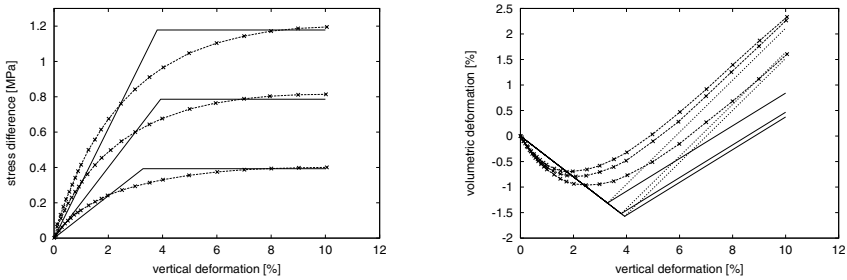


Figure 2. Questionable calibration of the Mohr-Coulomb model from the outputs of triaxial tests (Young modulus adjusted to the stress level)

settlement but departs from the measured tunnel deformation and vice versa [4].

The situation is even worse with respect to the subgrade reaction model. In many countries this model is standard for modelling the soil-structure interaction (design of excavation or tunnel support) in spite of its doubtful parameter, modulus of subgrade reaction. This parameter cannot be a soil constant since it is system dependent and thus it cannot be measured (except in a real 1:1 prototype in the same soil conditions).

Complex models

The calibration of complex models is a hard task since they usually have many material parameters. Nevertheless, the number of parameters is of a minor importance if a reliable calibration procedure is at the disposal. Unfortunately, this is rarely true.

Most model developers claim that one should be able to determine the model parameters from standard types of laboratory tests and they recommend usually triaxial and oedometer tests. Still, the majority of parameters are interrelated because they represent coefficients in highly non-linear functions which cannot be uniquely determined from experimental output. It is difficult to define a measure of an optimum approximation of the experimental data with such models which are also sensitive to the initial conditions in calibration tests.

A robust calibration should be based preferably on asymptotic states. Asymptotic states are e.g. critical states, proportional compression with constant ratio of the components of the stretching tensor or states with pressure-dependent minimum density due to cyclic shearing with a small amplitude. Such states are insensitive to initial conditions, eliminating thus the influence of sample preparation. They can be usually described by relatively simple equations with a low number of parameters. These parameters can be considered as true material constants because they do not depend on the actual state of soil.

Novel ways of model calibration seem to be optimization procedures including artificial neural networks. However, they are still under development and are thus not suitable for routine applications.

3.5 Initial state

It is astonishing how often an appropriate determination of the initial state is neglected in numerical simulations of BVP. It is standard to start calculations with a K_0 stress state, taking $K_0 = 1 - \sin\varphi$, although the geological history of the site can be reproduced at least qualitatively [17]. The initial state can be influenced even more dramatically by construction processes, e.g. old excavations, compaction or wall installation [33, 19].

The situation becomes worse with unsuitable constitutive models which usually do not distinguish between soil constants (i.e. material parameters) and state variables. An approximation of non-linear stress-strain curves with straight lines inevitably yields material "constants" which depend on the current state of soil and thus contradict the notion of a material parameter.

4 Mathematical and numerical aspects

Mathematical and numerical aspects in simulations of BVP are certainly linked to constitutive models. The general rule — the more complicated model, the more serious numerical problems — is valid.

Mathematicians have known for a long time that satisfactory results can be expected only in the case of a well-posed BVP, for which existence, uniqueness and stability of solution is guaranteed. It is not possible to perform such proofs in practical cases but one should be aware that there is sophisticated mathematics behind the contour plots and that to obtain a realistic solution is by no means obvious. According to the classification by Belytschko [1], most of the geotechnical calculations would still fall into the category *computer games*!

On the other hand, one cannot consider numerical difficulties as mere obstacles in applications. Reality is neither unique nor fully deterministic [21]. Bifurcation points in our equations often coincide with instabilities in the material response, which are far away from the limit states of stress, and a slight change in the initial conditions may result in a different material response. An instructive demonstration of this fact is obtained by analyzing the controllability of laboratory tests [11]. Some stress paths, which become out of control in laboratory apparatuses (e.g. behaviour after the peak of the stress-strain curve from undrained triaxial compression of water-saturated sand), cannot be controlled in numerical simulations as well.

5 Concluding remarks

I gave some hints on several pitfalls of numerical methods in calculation of deformations. In particular, we should be aware of

1. correct calculation of strains from displacements,
2. simultaneous loading and unloading in most cases,
3. absence of linear elasticity in the soil behaviour,
4. scatter in experimental results,
5. unsatisfactory deformation patterns calculated with simple models,
6. ambiguous calibration procedures,
7. considerable role of initial state,
8. difference between material parameters (constants) and state variables,
9. indeterministic components in the soil behaviour resulting in ill-posed BVP.

There are other important factors which have not been discussed here but which can significantly affect results of numerical calculations, among them

- scatter of soil conditions in situ,
- averaging procedures in multiphase continua,
- inertial and damping effects,
- drawbacks of iterative solution of the set of equations and of time integration.

The inability to get reliable numerical predictions could seem almost hopeless. So why do we sometimes succeed in obtaining realistic results? Is it just because we look only at particular curves we want to see? In the most cases we do not perform true predictions. We rather apply the observational method in the sense of the adaptation of numerical results to measured data. If we do not have data, we try at least to adapt the numerical results from analogical cases from our experience or intuition. The prediction competitions teach us that in unknown (new) situations the reliability of our predictions is poor. Consequently, questions may arise: Should we stop with the development of constitutive models and numerical tools since they do not bring us desired results? Should we rely only on experiments (model tests, centrifuge, prototypes) and observations?

I think that the opposite is true. We should abandon the tradition that a soil is characterized by E, φ and c, and intensify further research. We cannot proceed only based on physical modelling since most geotechnical constructions can be considered as prototypes and model tests with granular materials bring even more questions (e.g. how to scale the sand grains, how to eliminate systematic and random errors or how to measure stresses at small scale).

References

[1] T. Belytschko. On difficulty levels in non linear finite element analysis of solids. *IACM Expressions*, (2):6–8, 1996.

[2] G. Bianchini, P. Puccini, and A. Saada. Test results. In Saada and Bianchini, editors, *Constitutive Equations for Granular Non-Cohesive Soils*, pages 89–97. Balkema, 1988.

[3] J. Burland. "Small is beautiful" — the stiffness of soils at small strains. *Canadian Geotechnical Journal*, 26:499–516, 1989.

[4] M. Doležalová, V. Zemanová, and J. Danko. An approach for selecting rock mass constitutive model for surface settlement prediction. In *Int. Conf. on Soil-Structure Interaction in Urban Civil Engineering*, volume 1, pages 49–65, Darmstadt, 1998.

[5] G. Gudehus. A comprehensive constitutive equation for granular materials. *Soils and Foundations*, 36(1):1–12, 1996.

[6] G. Gudehus, F. Darve, and I. Vardoulakis, editors. *Results of the International Workshop on Constitutive Relations for Soils*, Grenoble, 1984. Balkema.

[7] I. Herle, T. Doanh, and W. Wu. Comparison of hypoplastic and elastoplastic modelling of undrained triaxial tests on loose sand. In D. Kolymbas, editor, *Constitutive Modelling of Granular Materials*, pages 333–351, Horton, 2000. Springer.

[8] I. Herle and P.-M. Mayer. Verformungsberechnung einer Unterwasserbetonbaugrube auf der Grundlage hypoplastisch ermittelter Parameter des Berliner Sandes. *Bautechnik*, 76(1):34–48, 1999.

[9] I. Herle and K. Nübel. Hypoplastic description of the interface behaviour. In G. Pande, S. Pietruszczak, and H. Schweiger, editors, *Int. Symp. on Numerical Models in Geomechanics, NUMOG VII*, pages 53–58, Graz, 1999. A.A.Balkema.

[10] I. Herle and J. Tejchman. Effect of grain size and pressure level on bearing capacity of footings on sand. In A. Asaoka, T. Adachi, and F. Oka, editors, *IS-Nagoya'97: Deformation and Progressive Failure in Geomechanics*, pages 781–786. Pergamon, 1997.

[11] S. Imposimato and R. Nova. An investigation on the uniqueness of the incremental response of elastoplastic models for virgin sand. *Mechanics of Cohesive-Frictional Materials*, 3:65–87, 1998.

[12] D. Kolymbas. An outline of hypoplasticity. *Archive of Applied Mechanics*, 61:143–151, 1991.

[13] D. Kolymbas. The misery of constitutive modelling. In D. Kolymbas, editor, *Constitutive Modelling of Granular Materials*, pages 11–24, Horton, 2000. Springer.

[14] T. Lambe. Stress path method. *Journal of the Soil Mechanics and Foundations Division ASCE*, 93(SM6):309–331, 1967.

[15] T. Lambe. Predictions in soil eingineering. *Géotechnique*, 23(2):149–202, 1973.

[16] I. Lydon. Behaviour problems. *Ground Engineering*, 33(11):31, 2000.
[17] M. Mähr. Berechnung primärer Spannungsfelder nach der Methode der Finiten Elemente und mit dem hypoplastischen Stoffgesetz. IGT, University of Innsbruck, 2000. Diploma Thesis.
[18] D. Muir Wood. The role of models in civil engineering. In D. Kolymbas, editor, *Constitutive Modelling of Granular Materials*, pages 37–55, Horton, 2000. Springer.
[19] K. Nübel, P.-M. Mayer, and R. Cudmani. Einfluß der Ausgangsspannungen im Boden auf die Berechnung von Wandverschiebungen tiefer Baugruben in Berlin. In *Ohde-Kolloquium*, pages 183–207. TU Dresden, 1997.
[20] R. Peck. Advantages and limitations of the observational method in applied soil mechanics. *Géotechnique*, 19(2):171–187, 1969.
[21] I. Prigogine and I. Stengers. *Order Out of Chaos.* Bantam Books, New York, 1983.
[22] K. Roscoe. The influence of strains in soil mechanics. *Géotechnique*, 20(2):129–170, 1970.
[23] A. Saada and G. Bianchini, editors. *Constitutive Equations for Granular Non-Cohesive Soils*, Cleveland, 1988. Balkema.
[24] H. F. Schweiger. Results from two geotechnical benchmark problems. In A. Cividini, editor, *Proc. 4th European Conf. Numerical Methods in Geotechnical Engineering*, pages 645–654. Springer, 1998.
[25] H. F. Schweiger. Comparison of finite element results obtained for a geotechnical benchmark problem. In C. Desai, T. Kundu, S. Harpalani, D. Contractor, and J. Kemeny, editors, *Computer methods and advances in geomechanics — Proc. of the 10th International Conference*, Tucson, 2001. A.A.Balkema.
[26] B. Simpson, N. O'Riordan, and D. Croft. A computer model for the analysis of ground movements in London Clay. *Géotechnique*, 29(2):149–175, 1979.
[27] F. Tatsuoka, S. Goto, T. Tanaka, K. Tani, and Y. Kimura. Particle size effects on bearing capacity of footings on granular material. In A. Asaoka, T. Adachi, and F. Oka, editors, *IS-Nagoya'97: Deformation and Progressive Failure in Geomechanics*, pages 133–138. Pergamon, 1997.
[28] F. Tatsuoka, R. Jardine, D. Lo Presti, H. Di Benedetto, and T. Kodaka. Theme lecture: Characterising the pre-failure deformation properties of geomaterials. In *Proc. XIV ICSMFE, Hamburg*, volume 4, pages 2129–2164. A.A.Balkema, 1999.
[29] F. Tatsuoka, M. Siddiquee, T. Yoshida, C. Park, Y. Kamegai, S. Goto, and Y. Kohata. Testing methods and results of element tests and testing conditions of plane strain model bearing capacity tests using air-dried dense Silver Leighton Buzzard sand. Report prepared for class-a prediction of the bearing capacity performance of model surface footing

on sand under plane strain conditions, Institute of Industrial Science, University of Tokyo, 1994.

[30] J. Tejchman and I. Herle. A "class A" prediction of the bearing capacity of plane strain footings on sand. *Soils and Foundations*, 39(5):47–60, 1999.

[31] G. Viggiani and C. Tamagnini. Ground movements around excavations in granular soils: a few remarks on the influence of the constitutive assumptions on FE predictions. *Mechanics of Cohesive-Frictional Materials*, 5:399–423, 2000.

[32] P.-A. von Wolffersdorff. A hypoplastic relation for granular materials with a predefined limit state surface. *Mechanics of Cohesive-Frictional Materials*, 1:251–271, 1996.

[33] P.-A. von Wolffersdorff. Verformungsprognosen für Stützkonstruktionen. Veröffentlichungen des Institutes für Bodenmechanik und Felsmechanik der Universität Fridericiana in Karlsruhe, 1997. Heft 141.

[34] C. P. Wroth. The predicted performance of soft clay under a trial embankment loading based on the Cam-clay model. In G. Gudehus, editor, *Finite Elements in Geomechanics*, pages 191–208. Wiley, 1977.

FE ultimate load analyses of pile-supported pipelines - tackling uncertainty in a real design problem

Hermann Lehar, Gert Niederwanger, Günter Hofstetter

Institut für Baustatik, Festigkeitslehre und Tragwerkslehre, Universität Innsbruck

Summary

This study addresses the prediction of the behavior of pile-supported buried pipelines when they are considered as shell structures. Results, obtained on the basis of the numerical model are compared with experimental results, conducted both in the laboratory and on-site. Special attention is paid to the scatter of geometrical, material and loading parameters. The experimental and numerical investigations served as the basis for the development of a design table for pile-supported buried pipelines.

1 Introduction

Pile-supported buried pipelines, made of ductile cast iron, represent a technically reliable and economic solution for sewage disposal in regions with poor soil conditions. Such a pipeline consists of individual pipes put together in sleeves and supported at discrete points, i.e. at the piles. The span between two adjacent piles is equal to the length of the individual pipes of 5 or 6 m, respectively. Typical pipe diameters range between 200 and 600 mm with wall thicknesses from 5 to 11 mm. A pile-supported pipeline can be viewed as a statically determinate beam structure with internal hinges (Fig. 1), consisting of several pipes, each of which is supported by a pile, except the first and the last one. The first pipe has two supports, one at the sewer shaft and a second on a pile. The last pipe is only supported by the sewer shaft at the other endpoint of the pipeline. The sleeve connections of adjacent pipes act as hinges. A pipe is supported on a pile by means of a saddle made of ductile cast icon (Fig. 2). The loads on the pipes mainly consist of earth pressure, resulting from dead load of the soil, surface loads and traffic loading.

From the reliability point of view, pipelines are series systems which are characterized by a poor redundancy, especially in the case of statically determinate structures. E.g., in a pipeline of 1 km length, consisting of 200 pipes,

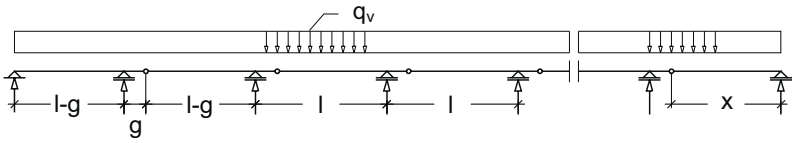


Fig. 1. Structural system and loading of a pile-supported pipeline.

just one single defective pipe may cause progressive failure of the system. Due to the severe possible consequences in terms of both property and environmental damage and resulting repair costs, pipeline designers are placing more emphasis on providing adequate resistance against failure. Nonlinear material behavior due to stress concentrations in the vicinity of the supports and ovalization of the pipes, known as *Brazier's effect* [3], result in strongly non-linear structural behavior. These effects require the use of a non-linear shell model, whereas the application of beam theory would lead to a severe overestimation of the load carrying capacity. Hence, in order to simplify the design of such pile-supported buried pipelines, a design table [13] was developed. It relies on non-linear finite element analyses, conducted for many available diameters and wall thicknesses of such pipes. The consideration of both geometrical and material non-linearity allows to simulate accurately the structural behavior up to failure by means of the numerical model.

The accuracy of the employed model is verified by laboratory tests and by in-situ measurements. The former consist of experiments, simulating the transmission of the pile force to the pipe. Strain gages, applied at several locations of the pipe, allow to measure the strains with increasing pile force and thus provide a thorough check of the numerical model for this most important part of the pipeline. The measurements on site focus on the determination of the pile forces and of the strains in the pipeline in the vicinity of the supports

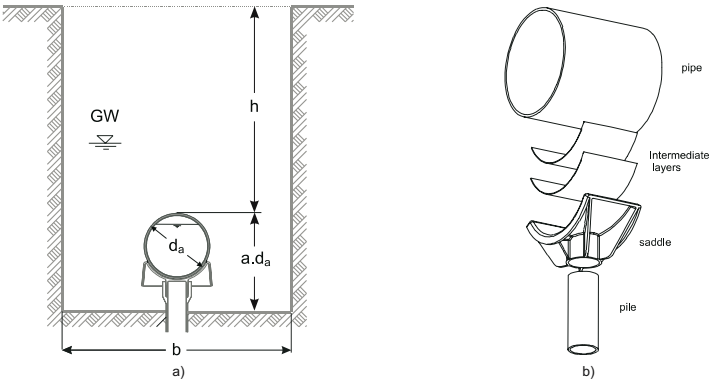


Fig. 2. Schematic diagram of a) the pipe trench and the pile-supported pipeline and b) the support of a pipe on a pile.

on the piles as well as midway between both ends of a pipe. These measurements serve as a means to check the assumptions for the design loads of buried pipelines, contained in the codes.

2 Pilot study

First of all, in order to determine the most critical regions of pile-supported pipelines and to gain better insight into the collapse mechanism of such structures, a three-dimensional FE-analysis was carried out using the general purpose program *ABAQUS* [11]. Taking advantage of symmetry, two adjacent pipe segments are modelled with solid and beam elements, respectively. To model the coupling, a beam-to-solid submodelling technique was used. The loading on the pipes was assumed to be equally distributed in axial direction. Fig. 3 shows the assumed distribution of the loading in the circumferential direction. For a given load q_v, contained for instance in codes [20], acting on

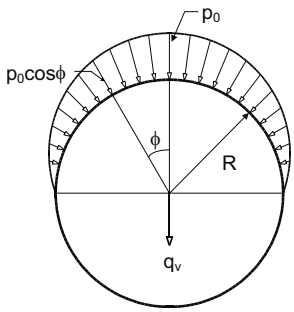


Fig. 3. Assumed circumferential distribution of the loading.

the pipes per unit length in axial direction, one obtains the pressure p_0, acting at the crown of the pipes, as follows:

$$p_0 = \frac{q_v}{R \int_{-\pi/2}^{\pi/2} \cos^2 \phi \, d\phi} = \frac{2\, q_v}{R\, \pi} \ . \tag{1}$$

The non-linear material behavior was approximated by a trilinear stress-strain relation. The contact between the saddle and the pipe and between the sleeve and the rubber sealing and the rubber sealing and the pipe, respectively, was accounted for in the numerical model. Three distinct locations were considered for a potential collapse:

- the pipe in the vicinity of the support,
- the sleeve connection,

- the section of the pipe with the largest bending moment.

Summarizing briefly the key results of the pilot study, it can be concluded that the pipe collapses due to ovalization and wrinkling in the vicinity of support. Fig. 4 shows that the maximum equivalent stress according to the v. Mises yield criterion occurs in the horizontal midplane section above the saddle and in the vicinity of the edges of the saddle. The equivalent stresses at the sleeve connection (Fig. 5) and at the cross-section of the largest bending moment are smaller and do not attain critical values.

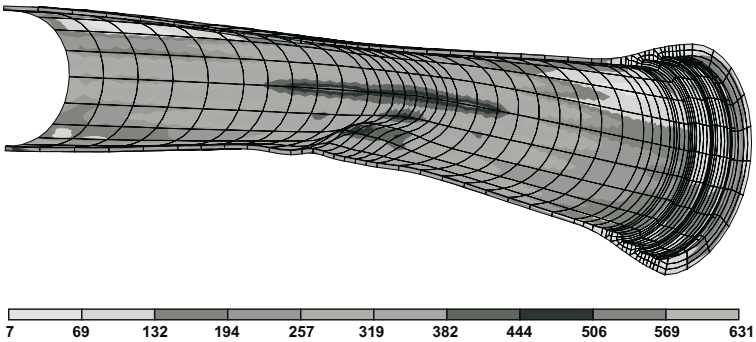


Fig. 4. Contour plot of the v. Mises equivalent stress [N/mm^2] of the pipe in the vicinity of the support on the pile.

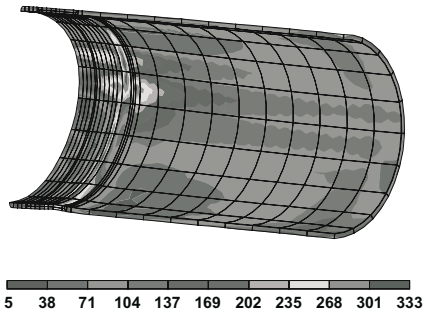


Fig. 5. Contour plot of the v. Mises equivalent stress [N/mm^2] of the pipe in the vicinity of the sleeve

3 Laboratory tests

3.1 Test set-up

For the validation of the numerical model two series of tests were carried out in a load frame investigating the behavior of a pipe segment in the region close to the support on the pile (Fig. 6). Regarding the assessment of geometric imper-

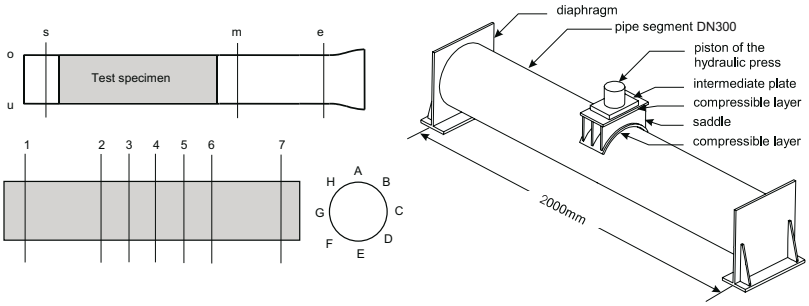


Fig. 6. Pipe segments and test set-up.

fections, the wall thickness and the thickness of the casting skin, respectively, were measured at cross-sections *1-7* (Fig. 6) prior to testing. Taking cutouts from cross-sections *s, m* and *e* at locations *o* and *u*, six specimens per pipe with standard geometries according to DIN standards were prepared to determine stress-strain curves for tension and compression loading. Strain gages were installed at the outer surface of the pipe segment to measure longitudinal as well as circumferential strains. Inductive pickups were used to measure the change of diameter and the deflection of the pipe during the experiment. The force applied by the piston of the hydraulic press, simulating the pile force transmitted from the pile to the pipe, was increased in a displacement controlled manner. Note that in the lab tests the pile is simulated by the piston of the load frame. Hence, the test set-up is rotated by 180° with respect to a pile-supported pipeline. The data acquisition system was a low speed scanner with a rate of 1 sample per second. The total number of channels was 20, each channel corresponding to a displacement transducer and a strain gage, respectively.

Tests were conducted for two different pipe classes specified as DN300/K9 and DN500/K9. The first one had a nominal outer diameter of 326 mm and a nominal wall thickness of 7.2 mm, the respective values for the second were 532 mm and 9 mm. Each of the test series involved the testing of three pipe segments. The following results refer to the pipe class DN300/K9.

3.2 Material tests

Grey cast iron is a material with a microstructure consisting of distributed graphite flakes embedded in a steel matrix (Fig. 7a). For tensile loading, the

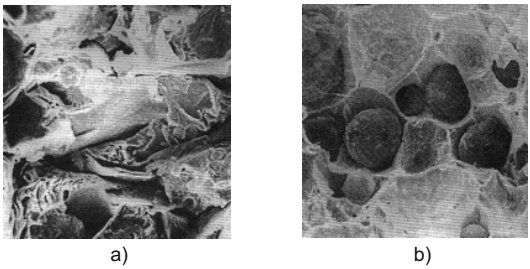


Fig. 7. Microstructure of a) grey cast iron and b) spheroidal graphite cast iron; (REMx700).

flakes are the source of stress concentrations, resulting in yielding as a function of the maximum principal stress, followed by brittle behavior. In the case of compressive loading, the graphite flakes do not have an appreciable effect on the macroscopic response, which is characterized by ductile behavior similar to that of steel [9, 10]. In contrast to grey cast iron, the spheroidal graphite cast iron with ball-shaped graphite particles, which is employed for the pipes, shows a different material behavior (Fig. 7b). In order to determine

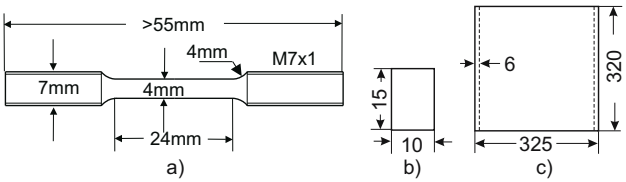


Fig. 8. Standard specimens for: a) tensile tests; b) compression tests; c) stub-column tests.

the required material parameters for the numerical analyses, tension and compression tests according to the standards DIN EN 10002-1 [21] and DIN 50106 [22] were performed. Fig. 8 shows the geometric properties of the specimens.

Comparing the experimental stress-strain curves for uniaxial tension and compression loading (Fig. 9 and Tab. 1), the following properties are observed:

- large scatter in the elastic region of the stress-strain curves for compression loading,
- relatively large deviations from the mean tensile response for some tensile tests,

- Young's modulus in tension, E_t, almost twice as large as Young's modulus in compression, E_c,
- yield stress in compression, σ_{cy}, exceeds the yield stress in tension, σ_{ty}, by 17 %,
- different hardening behavior in tension and compression,
- considerable differences between the ultimate stresses in compression and tension, σ_{cu} and σ_{tu}, respectively, and
- ductile behavior, characterized by a relatively large value of the ultimate strain ε_f.

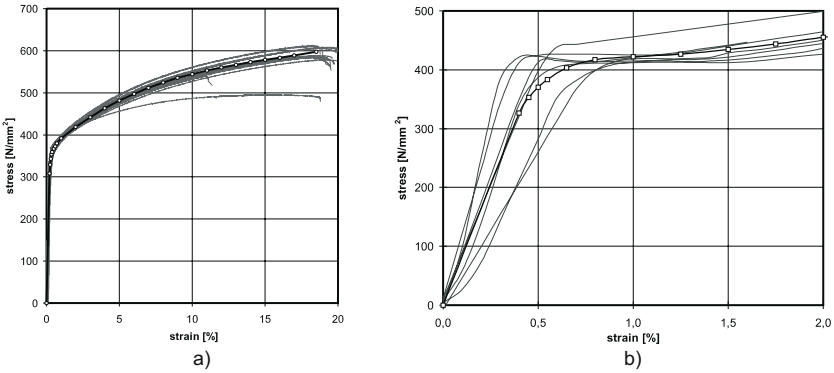


Fig. 9. True stress-strain curves from a) tension and b) compression tests.

Table 1. Material parameters (nominal values)

Parameter	unit	μ	δ	max	min
E_t	MPa	160 491	11.5 %	202 280	134 100
σ_{ty}	MPa	361	3.09 %	383	348
σ_{tu}	MPa	498	1.97 %	515	483
ε_f	%	21.1	13.1 %	23.61	12.05
E_c	MPa	86 757	33.6 %	136 540	53 546
σ_{cy}	MPa	424	13.0 %	451	413
σ_{cu}	MPa	903	5.00 %	942	853

Most of these phenomena can be explained by the microstructure of the spheroidal graphite cast iron (Fig. 7) and are in close agreement with results published in the literature [9, 10] on grey cast iron. However, the large differences of Young's modulus in tension and compression are not plausible (moreover, grey cast iron behaves quite conversely).

Concluding that test conditions and test specimens according to DIN 50106 are inadequate for compression tests in the present case, compression tests on two stub-column specimens (Fig. 10), which were cut out between sections m and e (Fig. 6) of the pipe segments, were performed to check the elastic modulus in compression. In order to measure circumferential and longitudinal strains, four equidistant strain gages were installed at the outer surface as well as at the inner surface of the pipe specimen. The measured elastic stress-strain curves in compression are shown in Fig. 11a. Determination of the slopes of the regression lines in the straight

Fig. 10. Axially loaded pipe with L/D=1 after testing.

parts of the loading and unloading branches of the stress-strain curves (Fig. 11b) yields the elastic modulus in compression. The statistical evaluation of the elastic moduli for compressive loading and unloading, E_{cl} and E_{cu}, respectively, and Poisson's ratio ν, is summarized in Tab. 2. Fig. 12a shows the stress-strain curve obtained from an individual tension test. The evaluation of the elastic modulus by means of the above described procedure, yields the values of 161 695 MPa, 140 875 MPa and 150 679 MPa

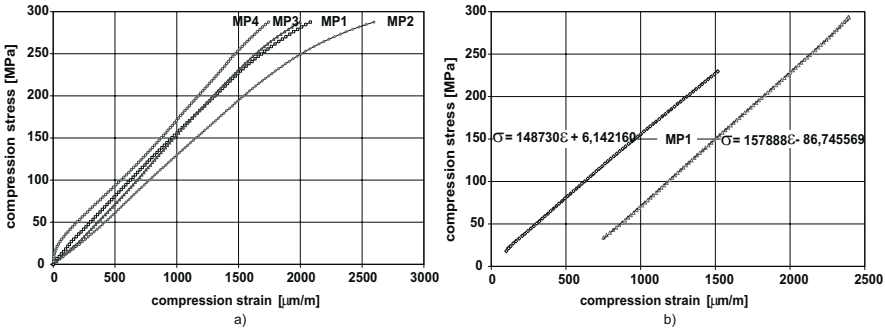


Fig. 11. Compressive stress-strain curves from stub-column tests; a) loading branches; b) individual loading and unloading branch.

Table 2. Material parameters from stub-column tests

Parameter	unit	μ	δ	max	min
E_{cl}	MPa	162 356	7.68 %	159 351	133 790
E_{cu}	MPa	149 637	7.40 %	172 277	147 377
ν		0.3918	6.51 %	0.4280	0.3720

for the loading, unloading and reloading branches, respectively. These values correlate very well with the corresponding values from stub-column tests according to Tab. 2.

In Fig. 9a the averaged tensile stress-strain curve (bold line with circular tickmarks) is shown, computed from the individual stress-strain curves by a procedure outlined in the following. The most convenient curve fit for a

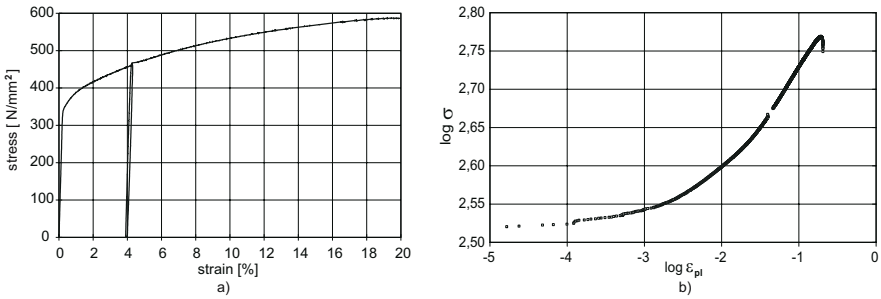


Fig. 12. Material behavior in tension: a) true stress-strain curve for loading, unloading and reloading; b) logarithmic diagram of plastic strains.

non-linear stress-strain curve is a power function. The resulting stress-strain equation is known as the *Ramberg-Osgood* equation:

$$\varepsilon = \varepsilon_{el} + \varepsilon_{pl} = \frac{\sigma}{E} + \frac{\sigma^n}{F}. \tag{2}$$

The first term of Eq. 2 represents the linear part of the stress-strain curve. As already outlined, the modulus E is determined as the slope of the linear part of the stress-strain curve. In the plastic domain, the plastic strain is represented in Eq. 2 by

$$\varepsilon_{pl} = \frac{\sigma^n}{F}. \tag{3}$$

For a given stress the plastic strain is obtained from the measured total strain as $\varepsilon_{pl} = \varepsilon - {}^{\sigma}/_{E}$. Taking the logarithm in Eq. 3 yields:

$$\log \varepsilon_{pl} = n \log \sigma - \log F. \tag{4}$$

Hence, in the logarithmic scale, experimental data should be represented by a straight line. However, as can be seen in (Fig. 12), this is not the case for the present material. Consequently, the constitutive relation according to *Ramberg-Osgood* cannot be applied.

Thus, the plastic region of the experimental stress-strain diagrams was subdivided into two parts, fitting a cubic polynomial for each part. Averaging the coefficients of the individual polynomials provided the averaged function for the plastic strain in terms of the stress.

It remains to determine the material behavior in the plastic compression range. At a first glance, it seams reasonable to use averaged stress-strain curves from axisymmetric stub-column tests. However, as depicted in Fig. 10, these tests show a characteristic behavior, also documented in [15, 16], characterized by the

- development of a single wrinkle over a small part of the specimen length,
- softening of the load-displacement relationship due to wrinkling,
- decrease of the compressive strain in regions remote from the wrinkle,
- increase of the compressive strain in the vicinity of the wrinkle.

Stresses calculated by the overall equilibrium condition from $\sigma = {}^{F}/_{A}$, F being the force and A being the sectional area, and strains measured by strain gages vary over the specimen length and hence do not represent material properties. Making use of the latter would result in a misinterpretation of experimental data. Hence, further experiments would have been necessary to

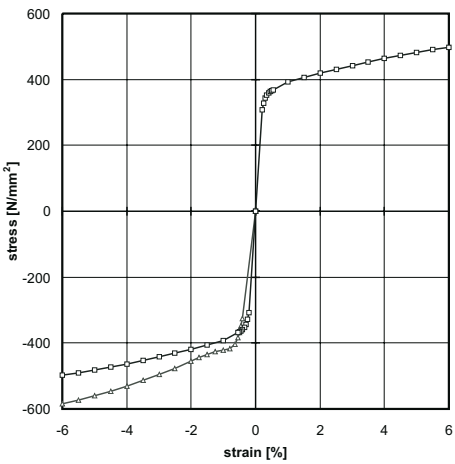


Fig. 13. Adopted true stress-strain curve of ductile cast iron.

eliminate this uncertainty. Nevertheless, because of economic reasons, it was decided to model the ductile cast iron as an elastic-plastic material assuming isotropic strain-hardening, i.e., identical behavior in tension and compression.

This simplification can be justified by the restriction to monotonic loading. It follows from Fig. 13, showing the averaged true stress-strain curve obtained from tensile tests (rectangular tickmarks) as well as the averaged stress-stain curve obtained from compression tests according to DIN 50106 [22] (triangular tickmarks), that this assumption seems to be conservative, at least for the plastic compression range. Furthermore, no experiments were performed to investigate the material behavior under cyclic tension-compression loading leading to kinematic hardening due to the *Bauschinger* effect.

3.3 Wall thickness measurements

Ductile cast iron pipes are manufactured in an inclined rotating chill-mold. As a result of this specific manufacturing process, the wall thickness is varying both in circumferential and longitudinal direction of the pipe. Thus, prior to testing, the wall thickness and the thickness of the casting skin of the pipes were measured for the cross-sections *1-7* at the locations *A-E* (Fig. 6) by an ultrasonic device. The measured thicknesses are summarized in Tab. 3 and

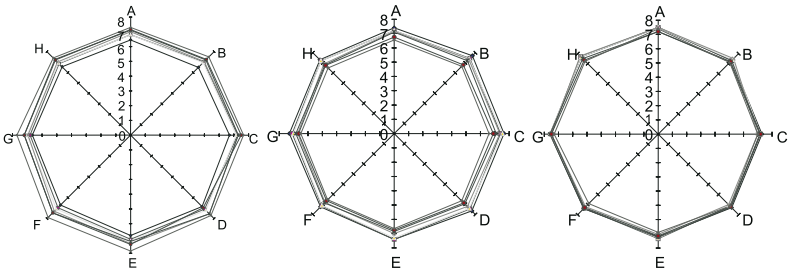


Fig. 14. Measured wall thicknesses at different cross sections of three pipes.

are illustrated in Fig. 14. Each octagon represents the measured wall thicknesses at one of the cross-sections *1-7* (Fig. 6). Hence, octagons lying closely together represent a small scatter of the wall thickness in axial direction and regular octagons indicate a small scatter of the wall thickness in circumferential direction.

Analyzing the data in Tab. 3, it is recognized that the wall thickness correlates with the cylindrical coordinates of the measurement point x and ϕ, x being the distance measured from some reference section to the measuring point and ϕ being the angle between the normal to the outer surface and a reference normal at a chosen origin. The distribution of the wall thickness s can be approximated by

$$s(x, \phi) = a_0 + a_1 x + a_2 \cos(\phi - \phi_0) + a_3 x \cos(\phi - \phi_0) \ , \tag{5}$$

Table 3. Measured wall thicknesses

	location	sect. 1	sect. 2	sect. 3	sect. 4	sect. 5	sect. 6	sect. 7	mean
	A	7.30	7.15	7.10	7.00	6.80	6.80	6.45	6.94
p	B	7.50	7.25	7.20	7.05	6.95	7.00	6.50	7.06
i	C	7.65	7.50	7.30	7.30	7.15	7.25	6.70	7.26
p	D	7.70	6.95	7.55	7.55	7.20	7.10	6.85	7.27
e	E	7.85	7.40	7.20	7.35	7.15	7.05	6.80	7.26
	F	7.90	7.45	7.45	7.30	7.10	6.95	6.70	7.26
I	G	7.70	7.20	7.20	6.95	7.00	6.80	6.65	7.07
	H	7.40	7.20	7.25	7.00	6.90	6.85	6.55	7.02
	mean	7.63	7.26	7.28	7.19	7.03	6.98	6.65	7.14
	A	7.40	7.20	7.20	7.15	7.10	6.75	6.55	7.05
p	B	7.70	7.45	7.40	7.35	7.20	6.80	6.65	7.22
i	C	7.65	7.40	7.55	7.35	7.10	6.90	6.60	7.22
p	D	7.65	7.45	7.40	7.20	6.95	6.85	6.65	7.16
e	E	7.40	7.40	7.30	7.00	6.85	6.75	6.65	7.05
	F	7.25	7.15	7.10	6.80	6.85	6.65	6.50	6.87
II	G	7.30	7.10	7.00	6.70	6.85	6.65	6.50	6.87
	H	7.30	7.20	7.20	6.90	6.80	6.75	6.50	6.95
	mean	7.46	7.29	7.27	7.06	6.96	6.76	6.58	7.05
	A	7.20	7.35	7.30	7.05	7.25	7.15	6.90	7.17
p	B	7.40	7.35	7.25	7.20	7.40	7.30	6.90	7.26
i	C	7.60	7.50	7.60	7.50	7.60	7.50	7.30	7.51
p	D	7.65	7.70	7.65	7.75	7.50	7.35	7.25	7.55
e	E	7.50	7.50	7.50	7.50	7.20	7.20	7.40	7.40
	F	7.30	7.30	7.40	7.50	7.15	7.20	7.10	7.28
III	G	7.30	7.20	7.20	7.00	7.15	7.20	7.00	7.15
	H	7.30	7.15	7.05	7.00	7.25	7.20	7.15	7.16
	mean	7.41	7.38	7.37	7.31	7.31	7.26	7.13	7.31

where ϕ_0 denotes the angle between the reference normal and the normal given by the diametrically opposed points of extreme thickness and the coefficients a_i and the angle ϕ_0 are random variables.

The variation of the wall thickness results from the inclined and, due to imperfections, slightly curved chill-mold in connection with both the gravitational and centrifugal forces. Because of the varying thickness, strictly speaking, the pipe loses its axisymmetric geometry. This can be seen in Fig. 15, where the measured thickness, averaged in longitudinal direction, is plotted for the three pipes of Table 3. The variation of the wall thickness in both longitudinal and circumferential direction of a pipe is checked by the quality management of the manufacturer. Pipes with wall thickness smaller than a specified minimum value are rejected.

In a final step, the effective thickness was determined. The outer surface of the pipe is wavy and the inner surface exhibits a casting skin. Extensive measurements with an ultrasonic device, verified by measurements with a micrometer screw, showed that the total thickness has to be reduced by 0.38 mm

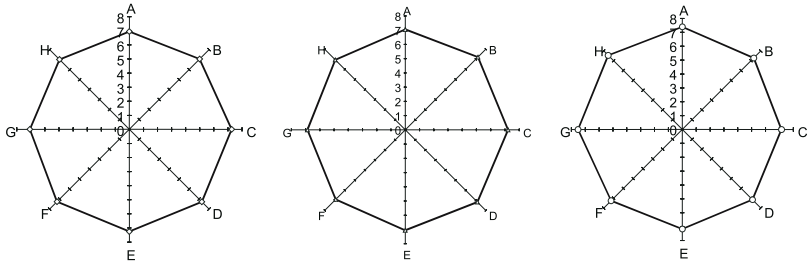


Fig. 15. Wall thickness in circumferential direction for three pipe segments, obtained from averaging in longitudinal direction.

to obtain the effective wall thickness of the pipe segments for the structural analysis.

After these preliminary studies, the bulk of statistical information for the random variables such as elastic modulus, Poisson's ratio, yield stress and uniaxial ultimate tensile stress and the corresponding strain was available.

However, there are further open questions. They refer for instance to the problem of localized imperfections, whose statistical distribution remains unknown, e.g. sinkholes which have been detected in some tests. They may trigger localized rupture or wrinkling (Fig. 16).

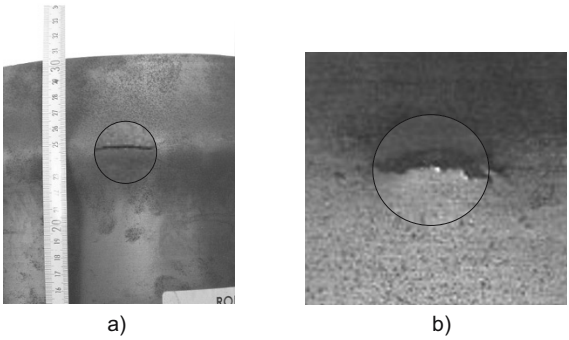


Fig. 16. Fracture at the a) outer surface of a stub-column test specimen due to b) a sinkhole at the inner surface of the specimen.

3.4 Pipe tests

For the validation of the numerical model three tests on pipe segments were conducted. They were loaded in the middle section by the piston of the hydraulic press (Fig. 6). The test results served for a comparison with the numerical predictions and for the confirmation of the employed boundary conditions. Strain gages were applied at 18 locations on the outer surface of the

pipe, two inductive pickups were used to measure both horizontal and vertical deflections of the pipe in the symmetry plane normal to the pipe axis. The measurement points are shown in Fig. 21.

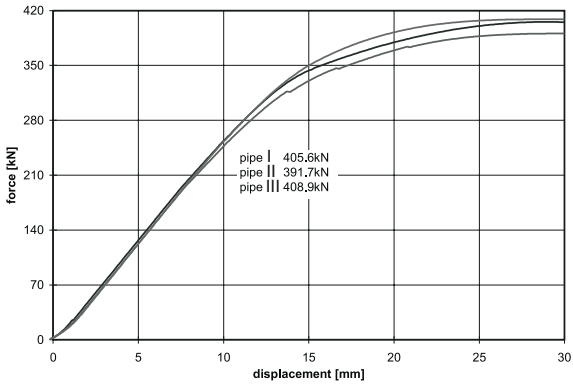


Fig. 17. Load-displacement curves.

Fig. 17 shows the measured load-displacement diagrams of the pipe segments. The load and the displacement refer to the force and to the vertical displacement of the piston. Excellent correspondence of the curves can be noticed. The ultimate loads of 405.6 kN for pipe I, 391.7 kN for pipe II and 408.9 kN for pipe III, correlate with the mean values of the wall thicknesses of the respective pipes of 7.14 mm, 7.05 mm and 7.31 mm, respectively (Tab. 3). The longitudinal strains, measured at points MP1 to MP4 at the pipe crown, as well as the circumferential strains, measured at points MP5 to MP9 in a horizontal plane containing the pipe axis, are depicted in Fig. 18.

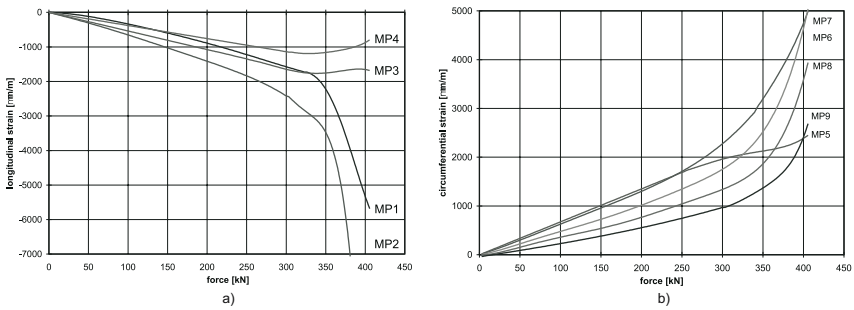


Fig. 18. Measured a) longitudinal and b) circumferential strains.

Fig. 19 shows the longitudinal and circumferential strains at points which are symmetric with respect to the symmetry planes normal and parallel to

the pipe axis. Up to a load level of 300 kN, the curves lie closely together. The subsequent deviations, especially for the circumferential strains, may be attributed to the observed scatter of the wall thickness.

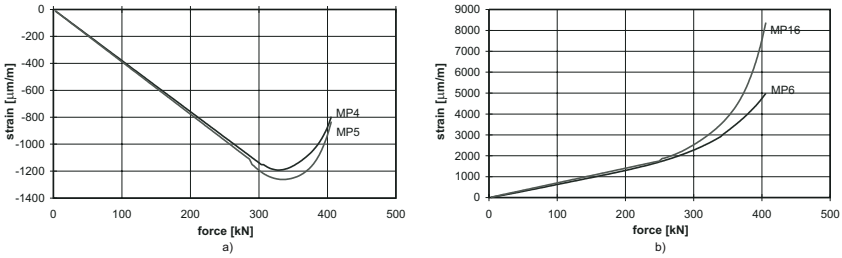


Fig. 19. a) Longitudinal strains and b) circumferential strains at symmetric measurement points.

The structural behavior can be interpreted from the shape of the strain curves shown in Fig. 18 and Fig. 19. Up to a load of about 250 kN, the structure shows linear elastic behavior. Upon further loading, ovalization increases considerably, leading to pronounced contact stress concentrations nearby the edges of the saddle extending parallel to the pipe axis. This in turn results in a decrease of the circumferential strains in the saddle area (Fig. 19a) at the expense of increasing circumferential strains above (Fig. 19b). Due to the reduction of the lever arm of the internal forces, caused by ovalization of the pipe, the longitudinal stresses increase. At a load level of 330 kN, wrinkling starts (see the longitudinal strains at MP3 and MP4 in Fig. 18a). Finally, the pipe fails (Fig. 20) due to the combined action of longitudinal compressive stresses and circumferential tensile stresses nearby the saddle. This is exactly the failure mechanism predicted in the pilot study by the distribution of the v. Mises equivalent stress (Fig. 4). Note, that in the laboratory tests, the pile is replaced by the piston of the load frame. Thus, the test set-up is rotated by 180° with respect to a pile-supported pipeline.

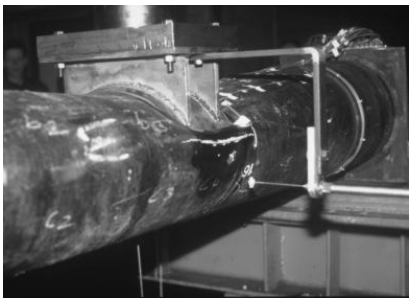


Fig. 20. Pipe segment after failure.

3.5 Validation of the numerical model by means of the pipe tests

The FE-mesh used for the numerical analysis of the pipe tests in order to validate the numerical model is shown in Fig. 21. Due to twofold symmetry, only one quarter of the test set-up, consisting of the pipe segment, the end-diaphragms, the saddle and the intermediate compressible layer, was discretized with shell elements. The saddle was modelled as rigid cylindrical surface with a modified pressure-contact relation in order to account for the soft contact between the saddle and the pipe due the intermediate compressible layers. The mean values of Young's modulus, the tensile strength and the ultimate tensile strain of the ductile cast iron according to Tab. 1, were used for the FE-analyses. The averaged stress-strain relation, consisting of the linear-elastic region and the elastic-plastic region, approximated by two cubic polynomials, was provided in tabular form. The observed scatter of the

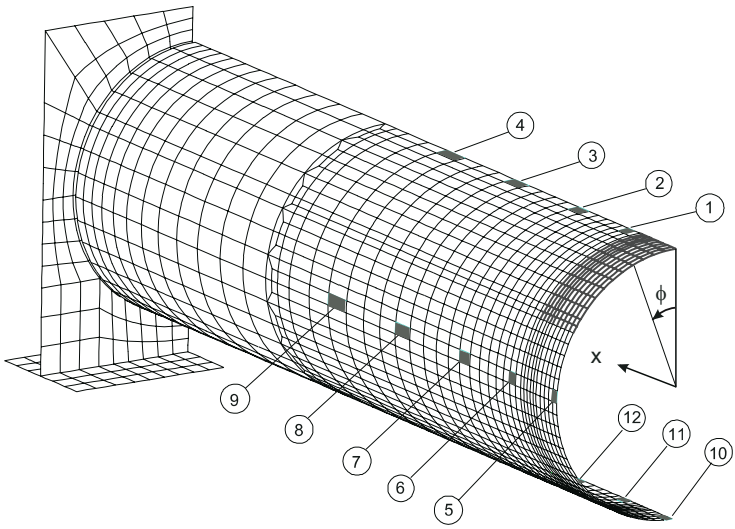


Fig. 21. FE-mesh for the pipe segment.

wall thickness was not taken into account in the computational model. Rather the mean value of 7.19 mm obtained from the mid-section of the three pipe segments (Tab. 3) was used in the first analysis. In a second run, the averaged thickness was reduced by 0.38 mm to account for the casting skin and the wavy surface.

In Tab. 4, the ultimate loads, determined from both experiments and FE-analyses, are listed. Excellent correspondence between the mean value of the measured ultimate loads and the computed ultimate load can be observed on the condition that the effective wall thickness is employed in the analysis. In Fig. 22, a contour plot of the equivalent stresses is depicted for the deformed

Table 4. Measured (pipe I, II and III) and computed (analysis I and II) ultimate loads

	wall thickness [mm]	ultimate load [kN]
pipe I	7.19	405.6
pipe II	7.06	391.7
pipe III	7.31	408.9
mean	7.19	402.1
analysis I	7.19	440.0
analysis II	6.81	409.1

mesh. The maximum values occur near the saddle edges. The computed deformation pattern is close to the experimental one, e.g. ovalization and wrinkling observed in the experiments are also predicted by the computational model.

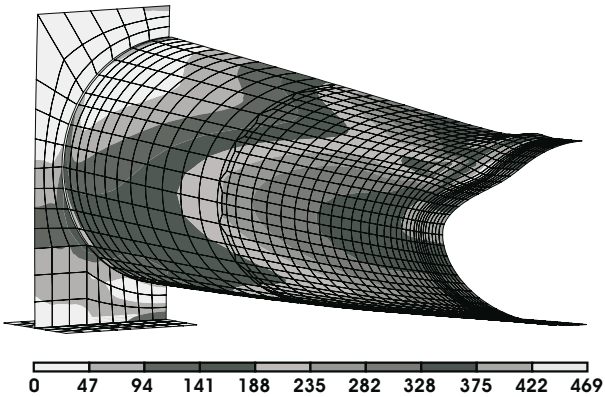


Fig. 22. Contour plot of the v. Mises equivalent stress for the pipe segment.

The accuracy of the numerical model is further confirmed by comparing the strains, measured at selected points, with the respective computed values. For the measurement points MP2 and MP6, shown in Fig. 21, both the measured and the computed longitudinal and circumferential strains are plotted in Fig. 23.

By contrast to Fig. 23, the correspondence between the computed and the measured displacements, shown in Fig. 24, is not as good. At a first glance, this seems to be a contradiction with the observed correspondence of the strains which are computed by differentiating the displacement field. However, the difference can be attributed to the fact that the measured displacement refers to the displacement of the piston whereas the computed displacement refers to the crown of the pipe beneath the piston. Because of the compressible

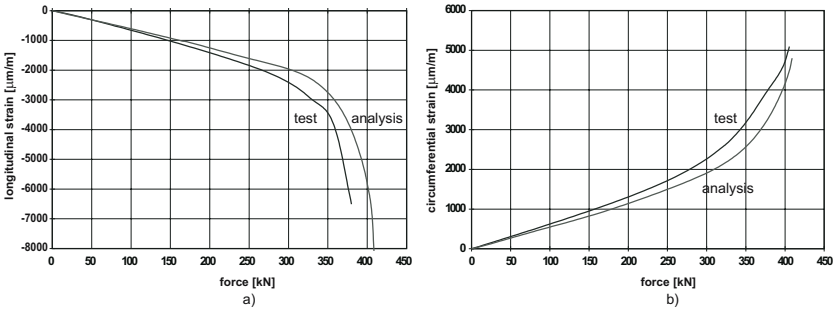


Fig. 23. Comparison of measured and computed a) longitudinal and b) circumferential strains.

layers between the saddle and the pipe as well as between the piston and the saddle, the vertical displacement of the pipe is smaller than the one of the piston. Concluding this section, a close correspondence of the results of the laboratory experiment and the FE-analysis can be stated. The shell model is perfectly adequate to predict the behavior of the pipe in the most critical region in the vicinity of the saddle provided that all non-linear effects are taken into account. Neglecting the geometrical non-linearity for instance would lead to an overestimation of the ultimate load by 30%.

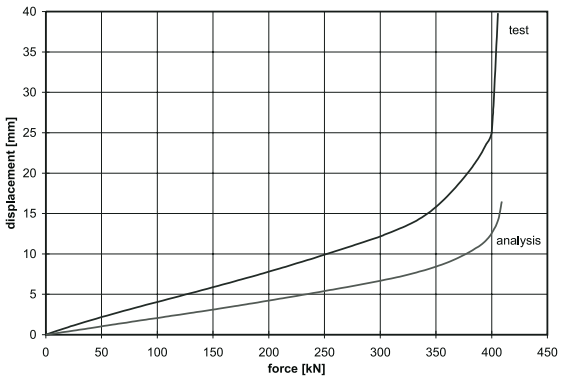


Fig. 24. Measured and computed load-displacement curve for pipe I.

Whereas for laboratory experiments the loads and boundary conditions are well-defined, additional uncertainties associated with the latter are introduced in engineering reality. The step from the validation of the FE-model by means of laboratory tests to the application of the numerical model in engineering practice will be performed subsequently.

4 Numerical model of the pile-supported pipeline

4.1 FE-discretization

In order to accurately reflect the structural behavior of a pile-supported pipeline in a numerical model, two adjacent pipes are discretized by second-order shear flexible shell elements, whereas the remaining pipes are discretized by second-order shear flexible beam elements (Fig. 25). The shell elements are quadrilateral elements with five degrees of freedom per node. Use of nine-noded shell elements was restricted to the contact zone of the pipe in the support region. Both shell elements and beam elements are formulated for large rotations but small strains. The shell elements are coupled with the adjacent beam elements by formulating suitable kinematic constraints. The features of the numerical model (Fig. 26) can be summarized as follows:

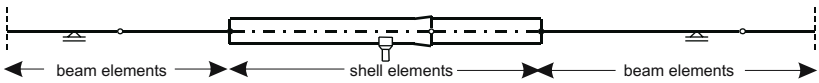


Fig. 25. Numerical model of a pile-supported pipeline.

- the solid elements, used in the pilot study, are replaced by shell elements,
- the sleeve is modelled in an approximate manner by shell elements with variable thickness combined with eccentric beam elements,
- the rubber sealing ring is discretized by spring elements, whereas the interaction between the sleeve and the sealing ring is modelled by simple gap elements,
- the support on the saddle is modelled as rigid cylindrical surface,
- the non-linear behavior of those pipes, which are represented by beam elements, is considered by a non-linear moment-curvature relation.

The latter is determined by means of a non-linear shell model of a pipe segment subjected to pure bending (Fig. 27). Taking advantage of symmetry, this model consists of six eight-noded shell elements in circumferential direction. It remains to prescribe appropriate boundary conditions.

Referring to Fig. 27, for the nodes in the plane of symmetry $z = 0$, the boundary conditions

$$u_z = \varphi_x = \varphi_y = 0 \tag{6}$$

hold. In pure bending the deformation is symmetric with respect to any rotated sectional plane. To prevent rigid body rotations about the z-axis, the boundary conditions

$$u_x = \varphi_y = \varphi_z = 0 \tag{7}$$

are prescribed for the plane $x = 0$. Hence, the latter is considered as a plane of symmetry. Identical boundary conditions have to be prescribed for the

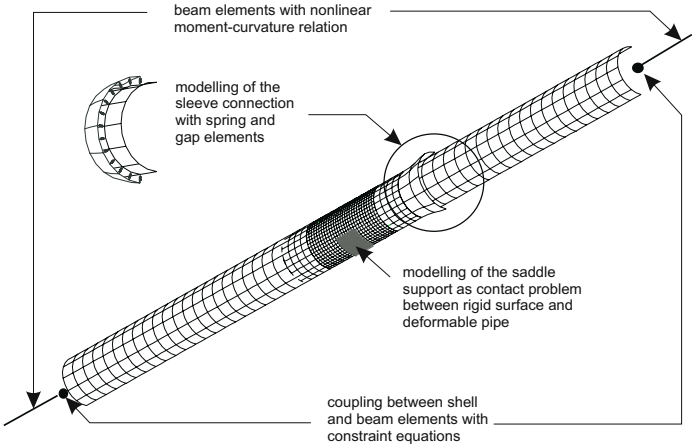


Fig. 26. FE-mesh for two adjacent pipes.

plane $x = l_e$, however, they have to be formulated with respect to the rotated coordinate system. To this end, a beam node, representing the motion of the plane $x = l_e$, is introduced with the degrees of freedom consisting of the displacement components u_x^b, u_y^b and the rotation φ_z^b. Pure bending of the shell model is simulated in a deformation controlled manner by prescribing the rotation φ_z^b of the beam node. Referring to [12], the following constraints

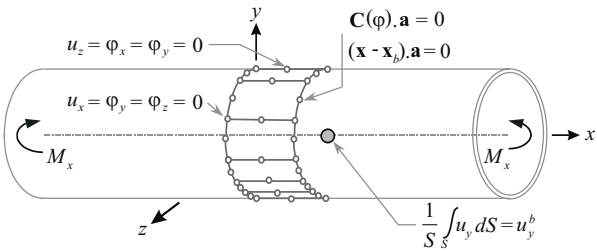


Fig. 27. Kinematic boundary conditions for a pipe segment in pure bending.

for the shell nodes in the rotated section plane $x = l_e$ can be written in terms of the tangent vectors of the undeformed pipe axis and the rotated pipe axis at $x = l_e$, denoted as $\mathbf{A}$ and $\mathbf{a}$, respectively:

- The shell nodes in the rotated plane at $x = l_e$ must remain in a plane normal to the pipe axis. This constraint is expressed by

$$(\mathbf{x} - \mathbf{x}^b) \cdot \mathbf{a} = 0 \tag{8}$$

  with $\mathbf{x}$ denoting the position vector of an arbitrary point of the rotated section plane and $\mathbf{x}^b$ representing the position vector of the mentioned

beam node. The rotation φ_z^b of the beam node must be equal to the rotation of the sectional plane $x = l_e$. Hence

$$\mathbf{C}(\varphi_z^b) \cdot \mathbf{A} = \mathbf{a} \tag{9}$$

holds, where $\mathbf{C}$ is the rotation matrix associated with the rotation φ_z^b. Since these constraints involving finite rotations are non-linear they have to be linearized in the context of the employed incremental-iterative FE solution strategy.

- In order to remove the rigid body mode in the direction of the y-axis a constraint is formulated, requiring that the averaged y-displacement of the shell nodes in the rotating cross-section at $x = l_e$ is equal to the y-displacement of the beam node, u_y^b. Thus, it follows that

$$\frac{1}{S} \int_S u_y \, dS = u_y^b \tag{10}$$

with S being the centerline of the cross section. u_y^b is then set to zero.

These non-linear multi-point constraint equations must be programmed in a FORTRAN user subroutine. From Eq. 8 it is evident, that the best position for the coupling node in the computational model of the pile-supported pipeline is the point of the largest bending moment midway between the piles.

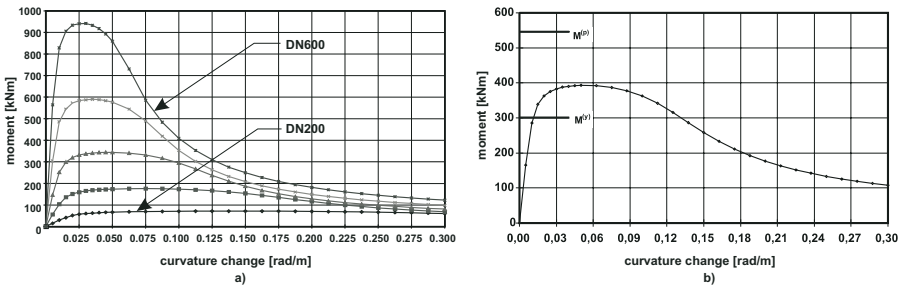


Fig. 28. Moment-curvature relations for a) pipe classes DN200 - DN500 and b) for the pipe class DN400/K9.

Fig. 28 contains the computed moment-curvature relations for different pipe diameters. Since the analysis was conducted in a displacement controlled manner, also the descending part of the moment curvature relation beyond the ultimate bending moment is obtained. For comparison, the elastic bending moment at first yield, $M^{(y)}$, and the fully plastic bending moment, $M^{(p)}$, both determined by a geometric linear beam theory, are shown in Fig. 28b for a particular pipe. It is emphasized that the ultimate bending moment, obtained on the basis of the shell model, is considerably smaller than the

ultimate bending moment, predicted by beam theory. The overestimation of the ultimate bending moment by beam theory can be attributed to the neglect of the deformation of the pipe section.

4.2 Loading

Pile-supported pipelines consist of three different structural members, namely the piles, the pipes and the subsoil. The soil pressure distribution acting on the pipe reflects the complex interaction between the soil and the pipe. It depends on the soil parameters, the ratio of the stiffness of the surrounding soil over the stiffness of the pipe and the dimensions and geometry of the pipe trench. Additionally, the soil pressure on the pipe may change with time due to environmental influences such as an unsteady groundwater table or freezing and thawing of the soil. According to [14, 17] and referring to Fig. 29, the pressure distribution in some pipe cross-section can be subdivided into three parts, the pressure p_1 acting on the upper half of the pipe, the pressure p_2 acting on the lower half of the pipe and the lateral pressure p_3. Each of

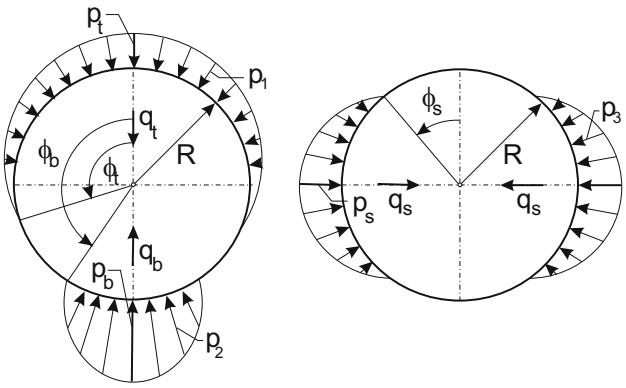


Fig. 29. Circumferential soil pressure distribution.

them could be approximated by:

$$p_i(\phi) = p_j \cos \frac{\pi(\pi - \phi)}{2(\pi - \phi_j)} \ , \quad i = 1, 2, 3, \quad j = t, b, s. \tag{11}$$

From equilibrium, the following relation between the pressure amplitudes p_t, p_b and p_s and the resulting line loads q_t, q_b and q_s, holds:

$$p_j = q_j \frac{\pi^2 - 4(\pi - \phi_j)^2}{4R\pi(\pi - \phi_j)\cos\phi_j} \ . \tag{12}$$

In a continuously supported pipeline, assuming homogeneous soil conditions, the amplitudes p_t, p_b and p_s and the angles ϕ_t, ϕ_b and ϕ_s are constant along

the pipe length. By contrast, in a pile-supported pipeline, these parameters vary periodically with the axial coordinate x (Fig. 30). Hence, they may be represented by a Fourier series as

$$p_j(x) = p_{j0} + \sum_{1}^{\infty} p_{jn} \cos \alpha_n x \ , \qquad \alpha_n = \frac{2n\pi}{l}. \tag{13}$$

It should be noticed that Eq. 13 is only valid for normal pressure components p_r with a symmetric distribution both in circumferential and longitudinal direction neglecting friction between the pipe and the surrounding soil. In the most general case, accounting also for frictional effects, the pressures acting on the outer surface of the pipe must be represented by a double Fourier series [5]:

$$\begin{aligned} p_x &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p_{xmn} \cos m\phi \, \cos \tfrac{n\pi x}{l} \ , \\ p_\phi &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} p_{\phi mn} \sin m\phi \, \sin \tfrac{n\pi x}{l} \ , \\ p_r &= \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} p_{rmn} \cos m\phi \, \sin \tfrac{n\pi x}{l} \ . \end{aligned} \tag{14}$$

The coefficients of the Fourier series could be determined, at least hypothetically, either by a numerical analysis, modelling the soil-pipe interaction as contact problem or by experiments. However, the result of such an analysis or experiment reflects the pressure distribution only for a particular case in a deterministic manner. This is in contradiction to the aims of a design table allowing the determination of the load-carrying capacity of pile-supported pipelines in different soils and at different depths. In order to assess the distribution of the line load q_b, acting along the bottom of the pipes, a pile-supported pipeline is modelled as an elastic beam on an elastic foundation of the Winkler type, fixed at discrete locations at a distance of l, representing the span between two adjacent piles. Due to repeated symmetry and neglecting for simplicity the hinge representing the sleeve connection, it is sufficient to consider a single span between two piles (Fig. 31). Let EI denote the flexural stiffness and b the effective width of the pipe, c the bearing coefficient of the

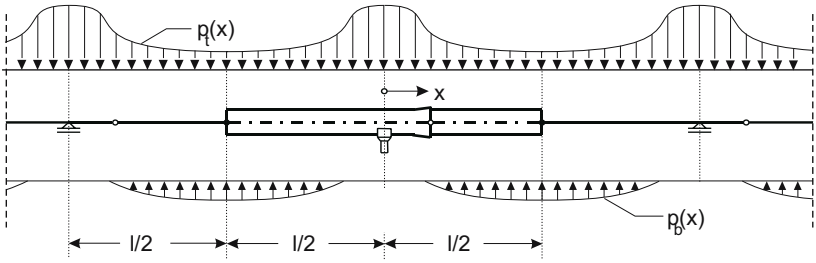


Fig. 30. Soil pressure distribution in axial direction.

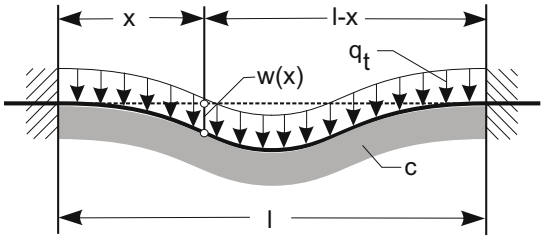


Fig. 31. Beam on elastic foundation.

subsoil and q_t the loading exerted by the soil above the pipe due to refilling the pipe trench. The solution of the differential equation for the deflection w

$$EI\frac{d^4w}{dx^4} + bc\,w = q_t \tag{15}$$

yields for a beam with fixed ends, subjected a uniformly distributed load q_t [8]:

$$\begin{aligned} &w(x) = \frac{q_t}{bc}\left[1 - \frac{1}{\sinh\lambda l + \sin\lambda l}\,(f_1 + f_2 + f_3 + f_4)\right] , \quad \lambda = \sqrt[4]{\frac{bc}{4EI}} , \\ &f_1(x) = \sinh\lambda x\cos\lambda(l-x) , \quad f_2(x) = \sin\lambda x\cosh\lambda(l-x) , \\ &f_3(x) = \sinh\lambda(l-x)\cos\lambda x , \quad f_4(x) = \sin\lambda(l-x)\cosh\lambda x . \end{aligned} \tag{16}$$

Making use of the relation for the shearing force

$$Q(x) = -EI\frac{d^3w}{dx^3}, \tag{17}$$

the pile force P is computed from $P = Q(0) + Q(l)$ as

$$P = 2\frac{q_t}{\lambda}\frac{\cosh\lambda l - \cos\lambda l}{\sinh\lambda l + \sin\lambda l}. \tag{18}$$

Here, the reaction force of the subsoil per unit length $q_b = bc\,w$ and the pile force P are compared for two pipelines with different pipe diameters: the first one is denoted as DN200/K9 with an outer diameter of $d = 222$ mm, a bending stiffness of EI $= 3\,933$ kNm^2 and a span of $l = 5$ m, the second one is denoted as DN400/K9 with $d = 429$ mm, EI $= 37\,140$ kNm^2 and $l = 6$ m. The outer diameter d is taken as the effective width b. For each pipeline two different soil types, namely a freshly poured sand with a bearing capacity of $c_1 = 3\,000$ kN/m^3 and a densely packed sand with a bearing capacity of $c_2 = 60\,000$ kN/m^3, are taken into account. The distributions of q_b, resulting from Eq. 16, are presented in Fig. 32. The diagram shows that the maximum value of q_b for the pipe DN400/K9 and poor soil conditions, only amounts to about 11% of the line load q_t. In contrast, the respective value of q_b for a pipe DN200/K9 with a significant lower bending stiffness and good soil conditions reaches about 100% of q_t.

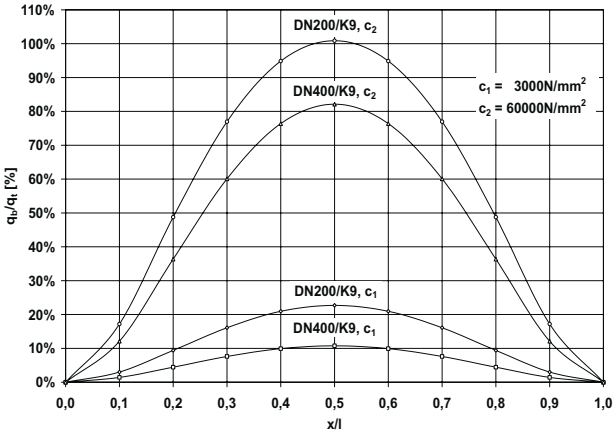


Fig. 32. Normalized reaction pressure distribution

Computing the pile forces P according Eq. 18 and relating them to the respective pile forces P_0 of a pile-supported pipeline without elastic foundation (Tab. 5), allows to conclude that neglect of the soil bedding is fully justified, since pile-supported pipelines are solely used in poor soil.

Table 5. Ratio of the pile forces P and P_0 with and without elastic foundation of the pipes.

pipe class	span [m]	c [kN/m^3]	P/P$_0$ [%]
DN200/K9	5	3 000	87.85
DN200/K9	5	60 000	42.34
DN400/K9	6	3 000	94.25
DN400/K9	6	60 000	54.78

Without doubt, some lateral pressure p_3 (Fig. 29) will be present. However, its magnitude depending on the soil parameters and on the compaction of the backfilled soil, remains uncertain. Keeping in mind that the lateral pressure has a rather beneficial effect on the pipes, diminishing the ovalization effect, neglect of the lateral pressure is a conservative assumption.

Because of the above mentioned reasons, for the computation of the design table only the pressure distribution p_1, acting on the upper half of the pipe from $\phi = -\pi/2$ to $\phi = \pi/2$ with the amplitude p_t at $\phi = 0$ is taken into account.

4.3 Benchmarks

As benchmark test for the performance of the computational model and, in particular, for the beam-to-shell submodelling technique, two pipelines, denoted as DN200/K9 and DN600/K8, with diameters of 222 and 635 mm, wall thicknesses of 4.8 and 6.9 mm and pipe lengths of 5 and 6 m, respectively, were analyzed. The vertical line loads according to Figs. 1 and 3 of $q_v = 30.8$ kN/m and $q_v = 83.8$ kN/m, respectively, were chosen close to the ultimate loads. As expected, the structural behavior of pipes with small diameters strongly

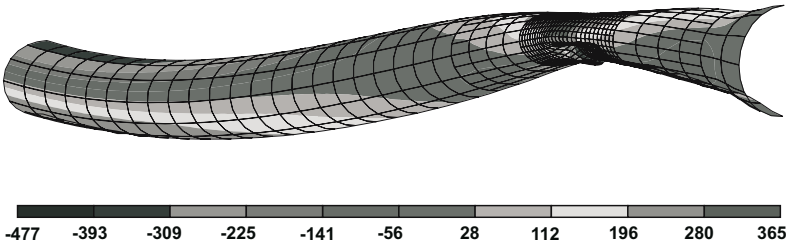


Fig. 33. Axial stress [N/mm^2] for a pile-supported pipe DN200/K9.

differs from that of pipes with large diameters. Fig. 33 and Fig. 34 show the deformed pipes with contour plots of the axial stress. Only the part of the pipe between the sleeve and the coupling node, containing the support on the pile, is shown. The shape of deflections, obtained for the pipe with a diameter of 222 mm (Fig. 33) is similar to predictions made on the basis of beam theory. A similar argument holds for the distribution of the axial stress, except for the regions in the vicinity of the support on the pile. This region is characterized by severe stress concentrations, resulting in a significant decrease of the load carrying capacity. Fig. 34 shows a very different structural behavior for a pipe with a diameter of 635 mm. Because of the considerably increased ratio of the diameter over the length of the pipe, the latter does not behave like a

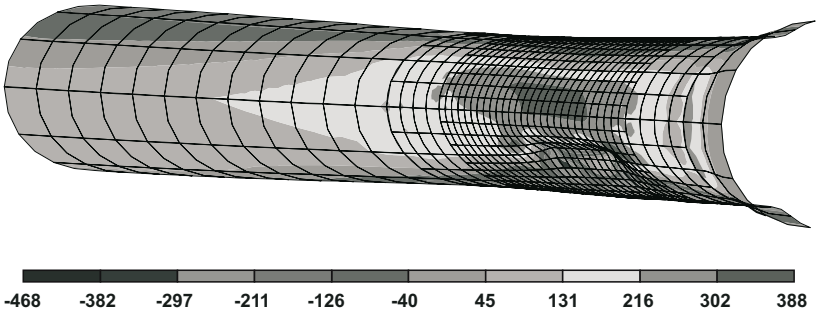


Fig. 34. Axial stress [N/mm^2] for a pile-supported pipe DN600/K8.

beam but like a shell. In this case the ovalization of the pipe is by far more important than the vertical deflection of the pipe axis.

Referring to Figs. 33 and 34 it is apparent, that the application of a nonlinear shell model is mandatory in order to correctly predict the complex stress and deformation state in the vicinity of the support. These special structural features are essential for the design, even for pipes with small ratios of the diameter over the span.

5 On-site measurements

The laboratory tests, described in subsection 3.4, are complemented by measurements on a construction site of a pile-supported sewer line in an Alpine valley with very poor soil conditions (Fig. 35). The specific weight of the soil can be assumed as 20 kN/m^3. The pipes are characterized by a length of 6 m, a diameter of 429 mm and a wall thickness varying between 7 and 10 mm. Supported on piles by means of a saddle, which is also made of ductile cast iron, they are placed 4 m below the ground surface (Fig. 2). The groundwater table is located close to the ground surface. During construction the side walls of the pipe trench were secured by sheet piles. The on-site measure-

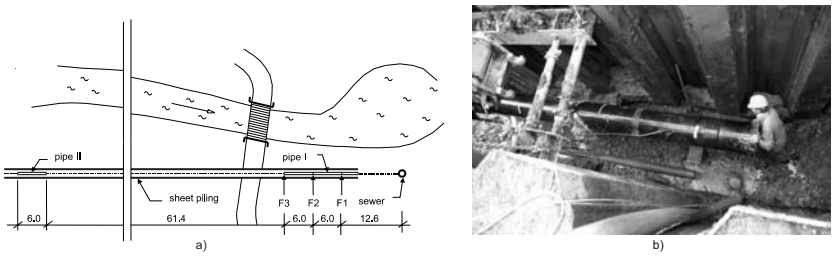


Fig. 35. On-site measurements: a) site plan; b) installation of a pipe.

ments allow to check the assumptions for the loads exerted from the ground on the pile-supported pipeline. Such loading assumptions can be found in recommendations [20] and codes for continuously supported buried pipelines. In addition, the influences of different construction stages were studied in the experimental program. In particular, the changes of the pile forces, caused by removing the sheet piles after refilling the pipe trench and by heavy traffic loading, were measured.

Strain gages were applied for the cross-section of a pipe, directly above the axis of the pile, as well as for the cross-section located midway between both ends of the pipe. The strain gages and the wiring were protected carefully from humidity and from mechanical damage during refilling the pipe trench. The pile forces of three adjacent piles, denoted as F1, F2 and F3 in Fig. 35a,

and the strains in two piles, denoted as pipe I and pipe II in Fig. 35a, were measured.

Fig. 36 shows the forces, measured for the three adjacent piles from the beginning of refilling the pipe trench until completion of the construction work by restoring the original shape of the ground surface. The differences of the

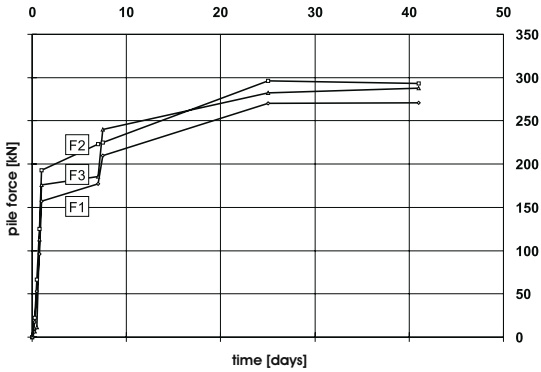


Fig. 36. Forces measured in three adjacent pipes.

measured pile forces are relatively small. Immediately after refilling the pipe trench the average pile force is 175 kN. The difference between the smallest and the largest pile force is 20%. Assuming an equally distributed loading in the longitudinal direction of the pipes, a line load $q_v = 29.2$ kN/m yields the average pile force of 175 kN. Removal of the sheet piles, 8 days after refilling the pipe trench, results in an increase of the average pile force to 225 kN. The reason for the further increase of the sheet piles to the final average value of 283 kN, measured 18 days after removal of the sheet piles, cannot be identified completely. It is attributed to the restoration of the original shape of the ground surface and to possible changes of the groundwater level. At this point of time the difference between the smallest and the largest pile force is reduced to 14%. The average pile force of 283 kN yields an equally distributed line load $q_v = 47.2$ kN/m. Between the last two measurements, no further increase of the pile force was observed. It is noted that the linear increase of the pile forces between two measurements, as shown in Fig. 36, is only obtained from connecting measured values at consecutive points of time. Hence, it does not necessarily reflect the actual time-dependent evolution of the pile forces. The removal of sheet piles causes a redistribution of the soil stresses, which may result in a considerable increase of pile forces.

Fig. 37a shows the short-term increase of a pile force of about 35 kN, measured during removal of a sheet pile close to the respective pile. After completion of removing the sheet piles, a permanent increase of the pile force of about 23 kN remains.

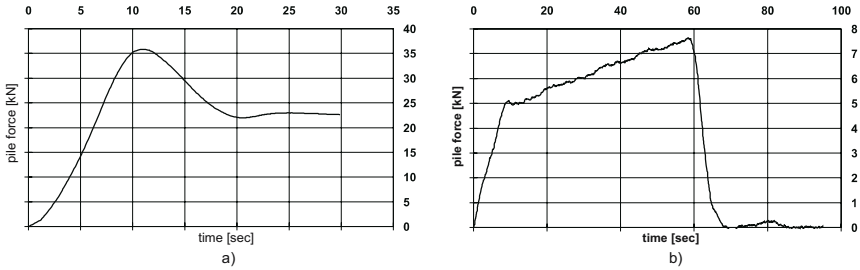


Fig. 37. Changes of the pile force due to a) removal of a sheet pile b) surface loads heavy truck.

Since the pipeline is located 4 m below the ground surface, the effect of a slowly moving heavy truck of 30 tons on the pile forces is relatively small. As can be seen in Fig. 37b, the maximum increase of the pile force does not exceed 8 kN.

Tab. 6 contains a comparison of the strains in a pipe, measured after completion of the construction work, with the respective computed values. The latter have been obtained on the basis of an uniformly distributed line load $q_v = 47.2$ kN/m, which corresponds to the measured average pile force of 283 kN.

Table 6. Comparison of measured and computed strains (* $\sim$ failure of strain gage).

measurement point	strain component	measured [μm/m]	computed [μm/m]
mid cross-section			
crown	longitudinal	-612	-646
crown	circumferential	-117	-199
bottom	longitudinal	+531	+667
bottom	circumferential		-445
support cross-section			
crown	longitudinal		+185
crown	circumferential		-718

According to Tab. 6, the largest deviations between measured and computed strains are noticed for the circumferential strains. This shortcoming of the computational model can be attributed to the neglect of the lateral pressure p_3 as pointed out in subsection 4.2.

Taking into account the difficult conditions for measurements on a construction site and the uncertainties of the assumptions with respect to the loads, acting on the pipes, the correspondence of the computed strains and the respective strains, measured on site, can be viewed as satisfactory.

The validation of the computational model by in-situ experiments closes the last open gap before the preparation of the design-table, which will be outlined in the last section.

6 Design

6.1 Probabilistic design approach

Probabilistic methods for the analysis and design of structures taking into account uncertainties of the geometry, the material behavior, the loading and, consequently, of the structural response, disclose possibilities for a theoretical foundation of existing standard design methods and the development of new, more economic design methods [1, 2, 6, 18]. Commonly, the design of a structure requires the consideration of the following limit states:

- *ultimate limit states* corresponding to collapse or other types of structural failure and
- *serviceability limit states* corresponding to specified service requirements for a structure (e.g. deformations affecting the use of a structure).

In the context of this study, the *ultimate limit state* is identified as failure of the pile-supported pipeline, which is induced by failure of the pipe in the vicinity of the support on the pile. The *serviceability limit state* is defined by an allowable limit of the ovalization of the pipe.

The design problem is usually formulated in the form of a *critical inequality*:

$$S \leq R \ , \tag{19}$$

where S and R are two random variables which represent a generalized *action* and a generalized *resistance*. In other words, the *probability of failure* P_{fail} is given by the probability of violating the critical inequality

$$P_{fail} = \text{Prob}[S > R] \ . \tag{20}$$

The basic design condition is to verify

$$P_{fail} \leq P_f \ , \tag{21}$$

P_f being a sufficiently small number which is established in the codes (for sewer systems usually $10^{-5} \leq P_f \leq 10^{-3}$ holds).

In geotechnical engineering it is often difficult to clearly distinguish between *action* and *resistance*. E.g., for pile-supported pipelines the pressure exerted by the soil on the outer surface of the pipe, i.e. the loading, is not known a priori but depends on a number of parameters, such as the material parameters of the surrounding soil, the stiffness of the pipe and the installation procedure. In addition, as can be seen in Fig. 36 and Fig. 37, the loading

is time-dependent. Hence, the *action* is a random process. The same applies to the *resistance*. For instance, due to corrosion effects during the *design life* of the pipeline, a decrease of the wall thickness may have to be taken into account. Thus, the exact determination of the failure probability is not feasible.

Hence, having in mind Bolotin's [2] position: *'Two dangers exist here: the danger of overestimating the role of statistical methods in structural mechanics as a whole or in individual applications, and the danger of underestimating them'*, which appears like a decision between *Scylla* and *Charybdis*, the design table for pile-supported pipelines was based on codes, following the semi-probabilistic approach.

6.2 Reliability in a code framework (Level 1 methods)

Level 1 reliability methods require to subdivide the relevant variables into *action* quantities W_i and *resistance* quantities χ_j. The *performance requirement* is expressed by

$$\mathcal{G}_1[\gamma_F^i, W_{ki}] \leq \mathcal{G}_2[\gamma_M^j, \chi_{kj}] \ , \tag{22}$$

where $\mathcal{G}_1$ and $\mathcal{G}_2$ are appropriate functions and γ_F^i and γ_M^j are *partial safety factors* which are specified by the codes. W_{ki} and χ_{kj} are the *characteristic values* of the corresponding *action* and *resistance* parameters, respectively. The *characteristic values* are defined as appropriate *fractiles* of the relevant random variables by the relations

$$\begin{aligned} &Prob\ [W_i < W_{ki}] = 1 - p \ , \\ &Prob\ [\chi_j < \chi_{kj}] = p \ , \end{aligned} \tag{23}$$

where the value of p is usually 0.05. Hence, W_{ki} is the 95% fractile of W_i and χ_{kj} the 5% fractile of χ_j. The structural reliability is controlled by the values of the *partial safety factors*, which are available in Level 1 codes. They were determined by taking into account probabilistic information.

6.3 Design according to DIN 18800

Specific codes of practice for the design of pile-supported pipelines, made of ductile cast iron, do not exist. For this reason the design was based on the German code DIN 18800 [23] for the design of structural steelwork.

The verification of the ultimate limit state according to DIN 18800 basically follows a Level 1 method. To this end, in Eq.(22) the characteristic values of the actions and of the resistance are factorized by partial safety factors, each of them reflecting the degree of confidence in the particular contributing effect. Depending of the load type, the partial safety factors, referring to the actions are, in general,

- for dead loads $\gamma_F^G = 1.35$ and
- for variable loads $\gamma_F^Q = 1.50$.

The respective value of the partial safety factor, referring to the resistance, i.e. to strength and stiffness, amounts to $\gamma_M = 1.10$. According to DIN 18800, it is permitted to account for γ_M in the context to the actions, resulting in

$$\gamma_F^G = 1.1 \times 1.35 \approx 1.5 \ , \quad \gamma_M = 1.1 \times 1.50 \approx 1.7 \ . \tag{24}$$

The ultimate state can be verified in terms of the line load, which is exerted by the soil on the pipes, by requiring

$$\frac{q_v^u}{q_v} \geq 1.7, \tag{25}$$

with q_v^u as the ultimate vertical line load, at which the bearing capacity of the pipeline is attained, and q_v denoting the respective line load under normal operating conditions. The latter is contained in recommendations and codes for continuously supported pipelines [20, 26]. The ultimate line load q_v^u was computed by means of the FE-model presented in section 4 for 25 different pipe classes [13].

Table 7. Material parameters for ductile cast iron from EN 598

parameter		unit	value
yield point	σ_y	[MPa]	300
ultimate tensile strength	σ_{tu}	[MPa]	420
ultimate strain	ε_f	[%]	10
elastic modulus	E_{cl}	[MPa]	170 000

Thereby, the characteristic values of strength and stiffness were replaced by the respective guaranteed minimal values, which are contained in the code EN 598 [24] and are checked by the quality management of the manufacturer of the pipes. They are summarized in Tab.7. Deviating from the code specifications, an elastic modulus of 156 000 MPa was employed in the numerical analysis. The latter represents the mean value obtained from the lab tests. The averaged true stress-strain curve determined from tests was scaled according to the code values. The latter, representing nominal values, were then converted to true (*Cauchy*) stresses and natural (*logarithmic*) strains (Fig. 38).

The characterstic value of the wall thickness was replaced by the guaranteed minimal wall thickness, given as

$$e_{\min} = K(0.5 + 0.001\,DN) - (1.3 + 0.001\,DN) \tag{26}$$

according to EN 545 [25], with K denoting the pipe class and DN the nominal outer diameter.

The design table, which was drawn up on the basis of the FE ultimate load analyses for 25 different types of pipes, contains the vertical line load q_v^o,

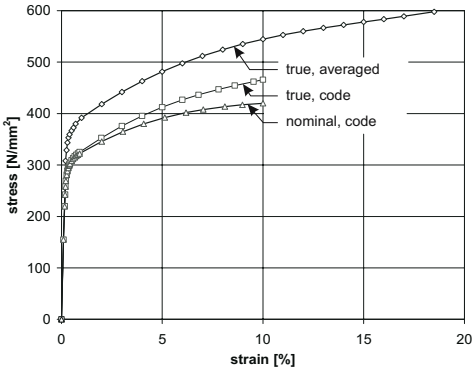


Fig. 38. Stress-strain diagrams.

at which the allowable limit of ovalizazion is attained as well as the vertical line load q_v^u, at which the bearing capacity of the pipeline is attained. Hence, the design table allows to verify both the serviceability limit state and the ultimate limit state of pile-supported buried pipelines, made of ductile cast iron, in an efficient manner.

7 Conclusions

An approach for the development of a design table for pile-supported buried pipelines was presented, which is characterized by evaluating uncertainties of the geometric properties, the material behavior and the loading of such structures by means of a synthesis of numerical modelling and experimental investigations. The material and geometric non-linear finite element model for ultimate load analyses of pile-supported buried pipelines was validated by laboratory experiments for a few selected types of pipes and by measurements on-site. Excellent correspondence of the experimental results, obtained in the laboratory and the respective computed values was achieved. Taking into account the difficult conditions for measurements on a construction site, even the correspondence of the computed strains in the pipes and the measured strains on site is rather good.

The close correspondence of the experimental and numerical results was achieved by the approach of conducting measurements and tests accompanying the development of the numerical model. The relatively close correspondence between the strains in the pipe measured on-site and the respective computed strains can at least partly be attributed to the fact that the line load q_v employed in the FE-analysis was determined from the pile forces measured on-site. The latter reflect in a broad sense the averaged *actual* soil-structure interaction, i.e. the loading. Such a favorable situation, characterized by the elimination of some of the uncertainties, associated with the loading,

is only encountered in geotechnical engineering, if on-site measurements are conducted.

Following [19], describing the design as a *decision process* in which 'various uncertainties have to be accounted for in assessing the design variables in order to obtain an acceptable failure probability', the presented synthesis of numerical modelling and experimental investigations turned out as a valuable design approach. It represents an alternative to the traditional way of gaining experience, which was expressed by Mark Twain by the well-known statement *'Good judgement comes from experience, experience comes from bad judgement'* [Mark Twain].

Acknowledgements

Financial support of this research project, granted by the Austrian company *Tiroler Röhren- und Metallwerke AG, A-6060 Hall, Tyrol, Austria*, a leading producer of pile-supported pipelines, made of ductile cast iron, and the Austrian *Forschungsförderungsfonds für die gewerbliche Wirtschaft* is gratefully acknowledged.

References

[1] Augusti G., Baratta A., Casciati F. (1984), Probabilistic Methods in Structural Engineering, Chapman and Hall Ltd. , London, New York

[2] Bolotin V.V. (1969), Statistical Methods in Structural Mechanics, Holden-Day, Inc., San Francisco, Cambridge, London, Amsterdam

[3] Brazier L.G. (1927), On the Flexure of Thin Cylindrical Shells and Other 'Thin' Sections, Proceedings of the Royal Society, London, Series A, vol. 116: 104–114

[4] Chen Y., Fuchs W., Netzer W. (1994), The Reliability of Buried Piping (Calculation of Safety Coefficients for Standard Önorm B5012 (in German), Österreichische Wasser- und Abfallwirtschaft, Jahrgang 46, Heft 11/12: 280–300

[5] Flügge W. (1967), Stresses in Shells, Fourth Printing, Springer-Verlag, Berlin, Heidelberg New York

[6] Freudenthal A.M. (1947), Safety of Structures, Trans. ASCE 112: 125–180

[7] Henzinger J. (1993), Pfahlbelastung bei der Gründung von Gussrohrleitungen auf Pfählen, geotechnik 16/4: 199–201

[8] Hetényi M.I (1946), Beams on elastic foundation; Theory with Applications in the Fields of Civil and Mechanical Engineering, The University of Michigan Press, Ann Arbor, Ninth printing, 1971

[9] Hjelm H.E. (1993), Elastoplasticity of Grey Cast Iron, FE-Algorithms and Biaxial Experiments. Ph.D. Thesis, Chalmers University of Technology, Division of Solid Mechanics, Sweden

[10] Hjelm H.E. (1994), Yield Surface for Grey Cast Iron under Biaxial Stress, Journal of Engineering Materials and Technology 116:148–154
[11] Hibbitt, Karlsson & Sorensen (1995), ABAQUS User's Manual, Version 5.5, HKS, Inc., Pawtucket, RI., USA
[12] Hibbitt, Karlsson & Sorensen (1995), ABAQUS Example Problems Manual, Version 5.5, HKS, Inc., Pawtucket, RI., USA
[13] Lehar H., Hofstetter G. (1997), Design table for pile-supported pipelines made by TRM AG. Research Report, Institute for Structural Analysis and Strength of Materials, University of Innsbruck, Innsbruck (in German)
[14] Mang F. (1966), Berechnung und Konstruktion ringversteifter Druckrohrleitungen, Springer, Berlin, Heidelberg New York
[15] Mohareb M., Alexander S.D.B., Kulak G.L., Murray D.W. (1993), Laboratory testing of line pipe to determine deformational behavior, Proc. 12th Int. Conf. on OMAE, Vol. V - Pipeline Technology, ASME: 109–114
[16] Murray David W. (1997), Local buckling, strain localization, wrinkling and postbuckling response of line pipe, Engineering Structures, Vol. 19, No. 5: 360–371
[17] Reusch D. (1961), Beitrag zur Berechnung eingeerdeter Rohre, Bauingenieur, 36
[18] Schuëller G.I. (1981), Einführung in die Sicherheit und Zuverlässigkeit von Tragwerken, W. Ernst & Sohn , Berlin
[19] Joint Committee on Structural Safety (1976), Common unified rules for different types of construction and material, CEB Bulletin d'Information, 116E
[20] Arbeitsblatt A-127 (1988), Richtlinie für die statische Berechnung von Entwässerungskanälen und -leitungen, Regelwerk der Abwassertechnischen Vereinigung (ATV), Gesellschaft zur Förderung der Abwassertechnik (GFA), St. Augustin, 2. Auflage
[21] DIN EN 10002-1 (1990), Metallic materials; Tensile Testing; Part 1: Method of testing (at ambient temperature), CEN (European committee for standardization), Brussels
[22] DIN 50106 (1978), Testing of materials; Compression test, DIN (German institute for standardization), Berlin
[23] DIN 18800 (1990), Teil 1: Stahlbauten-Bemessung und Konstruktion; DIN (Deutsches Institut für Normung), Berlin
[24] EN 598 (1994), Ductile iron pipes, fittings, accessories and their joints for sewerage application - requirements and test methods, Österreichisches Normungsinstitut, Wien
[25] EN 545 (2003), Ductile iron pipes, fittings, accessories and their joints for water conduits - requirements and test methods, Österreichisches Normungsinstitut, Wien
[26] Önorm B 5012 (1990): Statische Berechnung erdverlegter Rohrleitungen im Siedlungs- und Industriewasserbau, Österreichisches Normungsinstitut, Wien

The Implications of Geotechnical Model Uncertainty for Construction Management

Heimo Lessmann and Robert Vieider

Institut für Baubetrieb, Bauwirtschaft und Baumanagement, Universität Innsbruck

1 The "last"

The "last" turns off the light.

In the long chain of responsibilities for a building project, the actual execution represents the last link. She can turn off the light, but only if everything has run smoothly.

However, if failures occur, like for example the damage in Munich's Trudering on 20th, September 1994, the "last" can easily find himself under the harsh spotlight of the law, see Fig. 1. [9].

Fig. 1. Munich Trudering collapse, Subway line 2 - East Section 2 [9]

It's therefore understandable that the executors want to know what can come to light under scrutiny.

A responsible site engineering will always (not just in the event of failure) want to know how the building system will behave in the various phases of construction and what has been assumed as being the cause of a particular failure pattern.

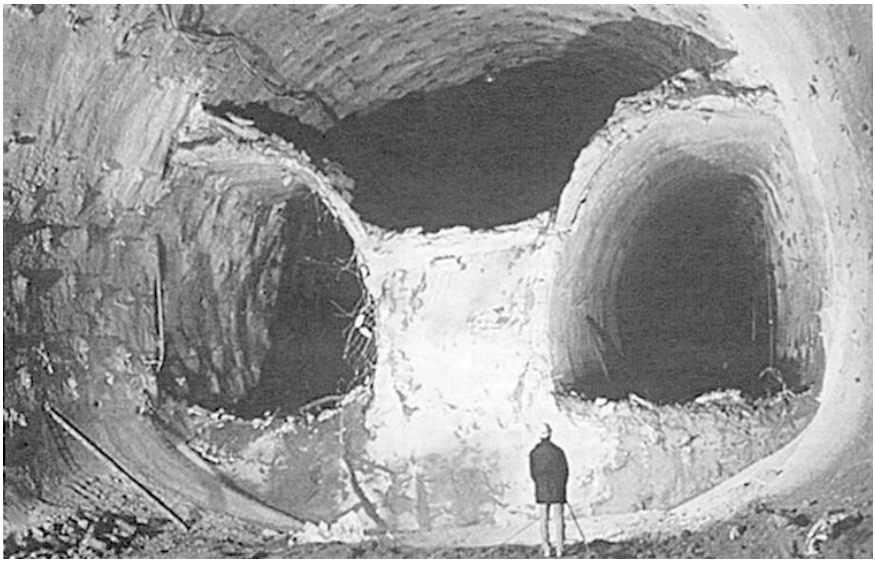


Fig. 2. The site engineer in front of the face [5]

Fig. 2. symbolises the site engineer's situation in front of his construction work and at the same time the responsibility that he bears [5].

The following worrying questions always remain [5]

- How will the soil respond to technical intrusion?
- How can the prediction – made as early as when the design decisions are taken – correctly cover the interaction between the soil and building?
- What has been overlooked?
- Will the planned work pass the test of reality?

2 The soil / building interaction

Critical patterns of behaviour can arise in these areas in particular.

Special attention should be paid to such situations in the case of shallow tunnels in soil, both in terms of construction as well as of course carrying out the work.

The approved plan with all the necessary documentation is the basis for carrying out every construction project.

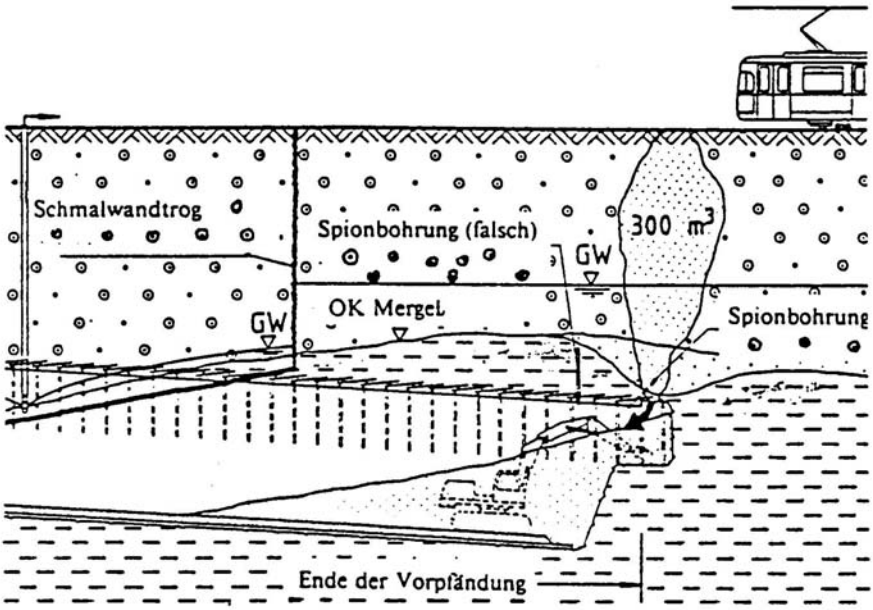


Fig. 3. Critical area in front of the face caused by weakening of the marl roof [3]

In critical areas, for example as shown graphically in Fig. 3. and Fig. 4., we can find that the actual behaviour of the system may differ significantly from the prediction of the planning and computation [3, 12].

It seems almost obvious, to look at first for the cause of such a deviation in the execution that does not go according to plan, because it is simple to believe that the computational model corresponds with the real behaviour and that the cause of the deviation can therefore only lie in divergent execution.

It is simply assumed that the computation (the mathematical model) corresponds to reality (the physical truth).

However, H. Duddeck pointed out two results for the examples listed in Fig. 5. of possible collapses in the case of large tunnel cross sections that were built using shotcrete construction method and driven in partial excavations:

1 Almost all collapses occur due to soil influences and happen while the tunnel is being driven.
2 Many of the hazardous situations are very difficult to record with the usual methods of statics and geotechnical engineering or are beyond even computational analysis [5].

The references below are intended to show just how questionable the application of the model assumptions to real problems can be.

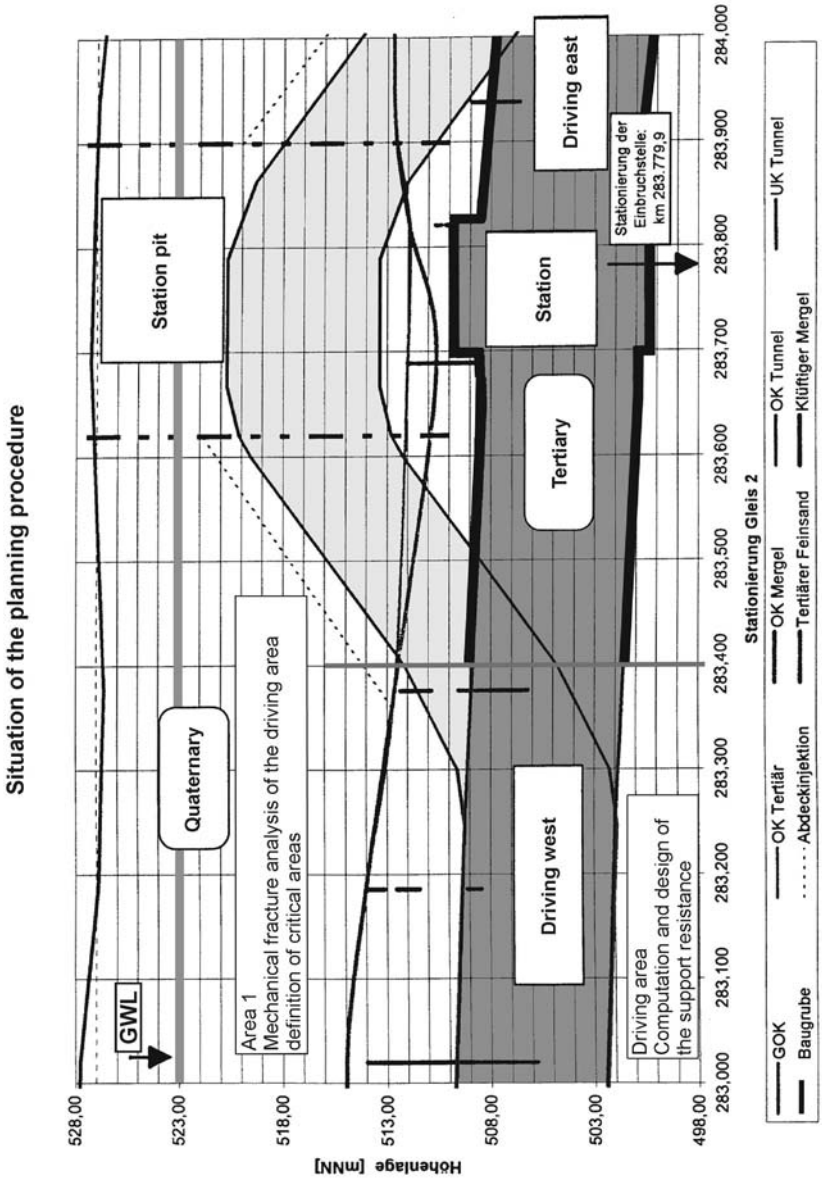


Fig. 4. Statutory Planning, Munich Trudering [12]

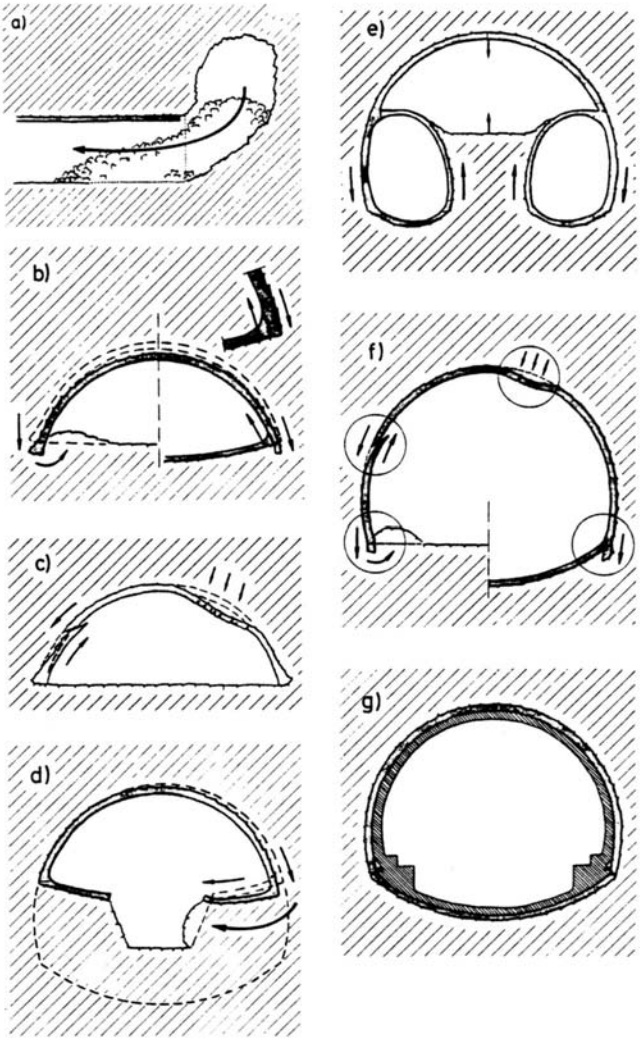


Fig. 5. Possible cases of collapse in tunnels built with the shotcrete construction method [5]

According to M. Wehr, in the field of pure mathematics the mechanisms of theorising are comparatively easy to recognise, since the complicated problem of applying the formulae to the measurement results is not a factor. "Despite this, it has recently become apparent that even that which is held to be mathematically true includes a subjective component, which means that everything is anything but evident".

"Even in modern mathematics, a theory is only that for what the scientists think of it. If we really ask in detail why certain attributes can be assigned to this while others cannot be assigned to it, it becomes apparent that here too we very quickly enter into a sphere of personal belief. Of course, this applies to the same extent to theories in which, like in chaos theory, mathematics is applied within the framework of a modelling process" [20].

Although the deviations between the measured data in the execution and the prediction of the computation can be significant, we derive the necessary conclusions from this only rarely. In his article "Difficulties related to geotechnical calculations" I. Herle clearly points out the possible deviations and in so doing arrives at the following result: "An overview of several prediction competitions shows an unsatisfactory state of the art" [11].

K. Poppers's demand that the theoretical model be called into question in the event of a lack of correspondence between assumption (prediction) and the experiment is, if necessary, replaced by adjusting the model assumption that does not apply. To draw an analogy here, the design computation represents the assumption (prediction) and the construction site represents reality and therefore the experiment.

Recognising fallibility and refuting the knowledge derived from assumptions are exactly the things that open up the possibility of making progress in our hypotheses.

This enables us to learn from our mistakes and come closer to the truth [10].

This enables us to draw up better theories and compare which theories contain more truth than others [10].

However, K. Poppers's opinion implies that scientific theories (unless they are falsified) will forever remain hypotheses and assumptions. No theory can claim to be the final truth, no matter how strong an explanation it provides.

We are looking for the truth but we do not possess it.

Science is the search for the truth: not the possession of truth, but the search for truth. Every theory is only a hypothesis [10].

Finally, we sometimes forget that some critical situations – for example in the area of the driven face – cannot be sufficiently comprehended, neither via computation based on the current state of the art nor by the current standards.

According to H. Duddeck, the main challenges faced by engineers when building tunnels lie in correctly understanding the soil. There is also the problem that the hazards arising from the soil properties for the construction work are so multi-faceted and complex that EURO Code safety rules and Finite Element Methods are barely able to cover them all [5].

G. Spaun used several examples of tunnel collapses to illustrate just how important or how dangerous various geological details can be and what consequences can arise from not recognising or not paying attention to different geological factors.

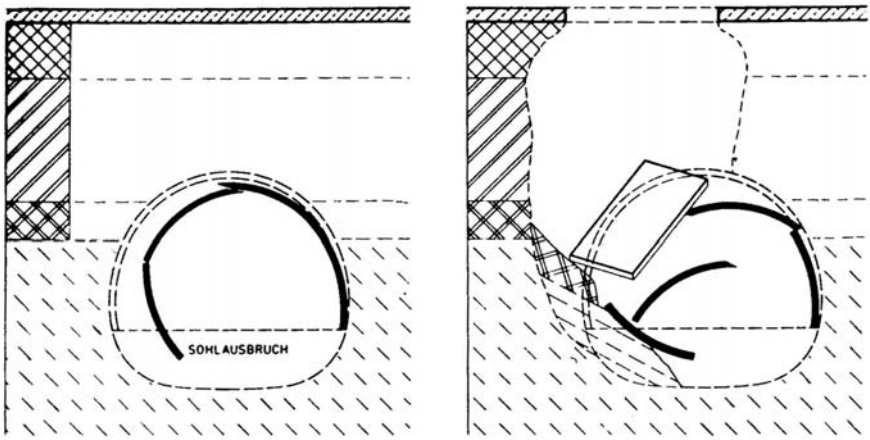


Fig. 6. Municipal railway tunnel, Bochum. Diagram of the tunnel collapse (after L. Müller 1978) [16]

Fig. 6. shows the pattern of tunnel collapse from excavation work for the arch invert during construction of the contract sections A2 of the Bochum municipal railway [16].

"The entire discussion in the regulations about safety coefficients and probabilities ignores the much more important question: how do we even find the potential risks and how can we estimate them correctly? In underground construction in particular, it becomes apparent that the computed safety is not the most important thing, but that which comes before, namely the hazardous situations to which attention must be paid" [5].

Although the efficiency of structure computations has increased greatly in recent years along with the further development of numerical methods[5], the continuum models (for example the Finite Element - and Boundary Element Methods) are hardly capable of determining states of collapse or complicated failure mechanisms, because this requires simulating the changeover to discrete discontinua which must first of all be found via the computation [5].

Using the example of the Krieberg tunnel collapse shown in Fig. 7., H. Duddeck has illustrated the main aspects of the crucial importance of the soil. The very specific soil conditions involved in the interaction between clay, sand, water and unannounced failure mechanisms were not incorporated into

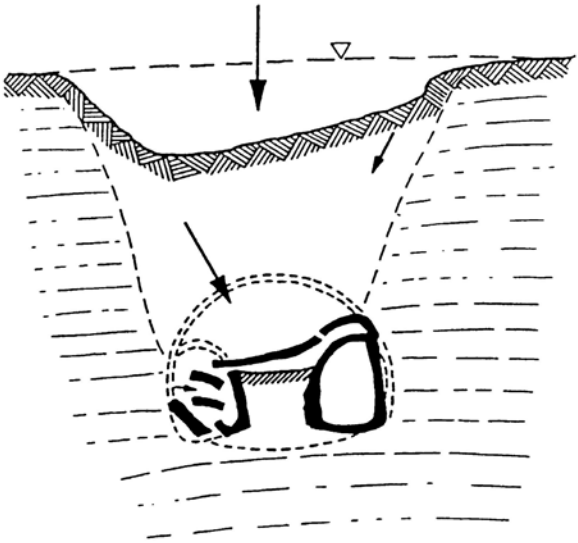


Fig. 7. Krieberg tunnel collapse [5]

the static computations. The convergence measurements carried out during the execution also did not illustrate the acute danger.

It was not the concrete (for which there are standards, a Euro Code and safety factors) that finally failed, but the soil with its unannounced sudden failure mechanisms [5].

3 The computational model

Many aspects of the present model assumptions can be called into question.

The diaphragm used as a computational model for a tunnel, as shown in Fig. 8. (whether two-dimensional or extended to a three-dimensional representation) is only an assumption, in the case of which it is not possible to conclusively examine whether this would enable the real behaviour to be described correctly [15].

Why are current model assumptions questionable?

- The soil is not a continuum [13].
- The soil parameters provide no complete information about the exposed soil. It is hardly possible to accurately determine the parameters, because the model reflects the soil behaviour only approximately [13].When examined in the laboratory, the soil samples provide a wide scatter of results and the choice of characteristic soil parameters has a crucial influence on

the result of the design. Due to the precise computation rules in the standards, the scatter of the soil parameters and their effects on the design tend to be forgotten [7].

- The various constitutive models are only approximations, something that seems to be undisputed [13]. The model / reality distance of approximation is largely unknown here.

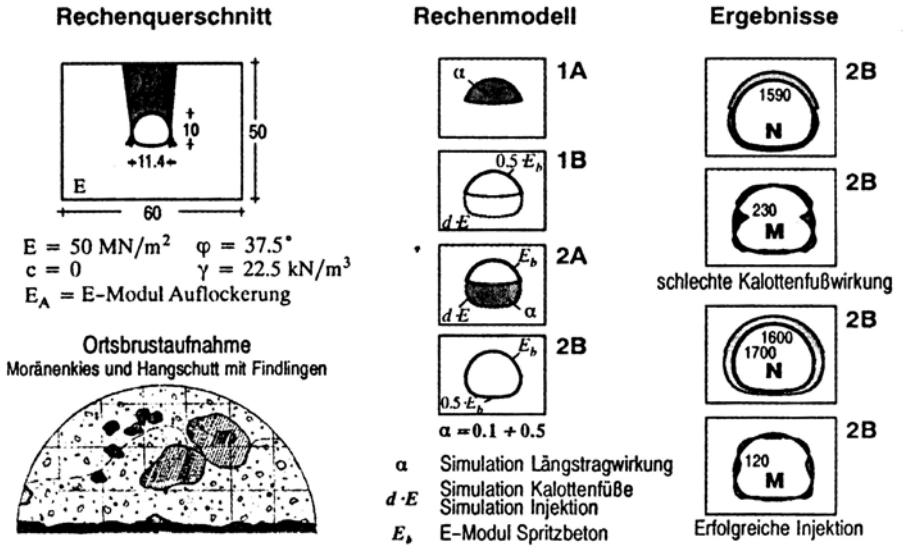


Fig. 8. Computational model for tunnel driving in soil, Farchant tunnel [15]

The fact that critical situations – for example in the area of the driven face – can be sufficiently detected neither by a computation based on the current state of the art nor by the applicable standards, cannot be ignored.

R. P. Feynman described this situation of approximating knowledge as follows: "The touchstone of all knowledge is the experiment. Experimentation is the sole measure of scientific "truth" [8].

Fig. 9. shows the experiment conducted in Munich in 1979 with compressed-air for the shotcrete construction method. As required by K. Popper and R. P. Feynman, the correspondence between assumption (prediction) and experiment was tested using a measurement program [1].

As Fig. 10. and Fig. 11. below show, a useable excavation system was only developed from the structural model after the experiment - and that is still the case today [6].

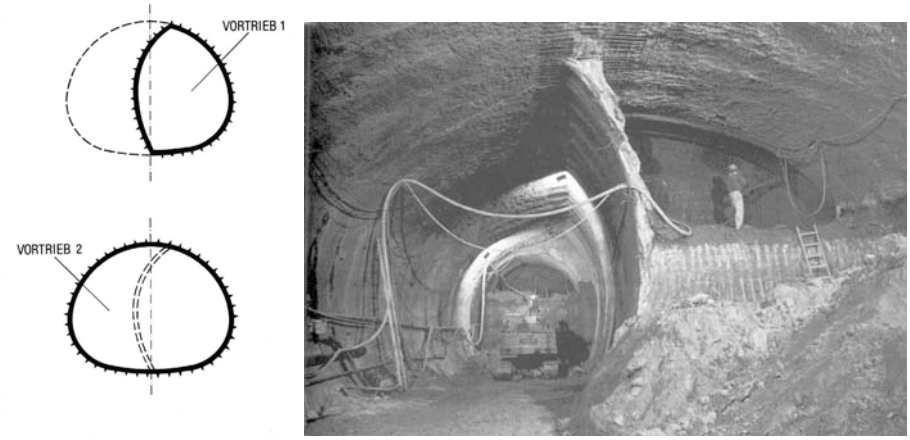


Fig. 9. Experiment – Compressed-air tunnel driving in the shotcrete construction method, 1979 Line5/9 Munich Odeonsplatz [1]

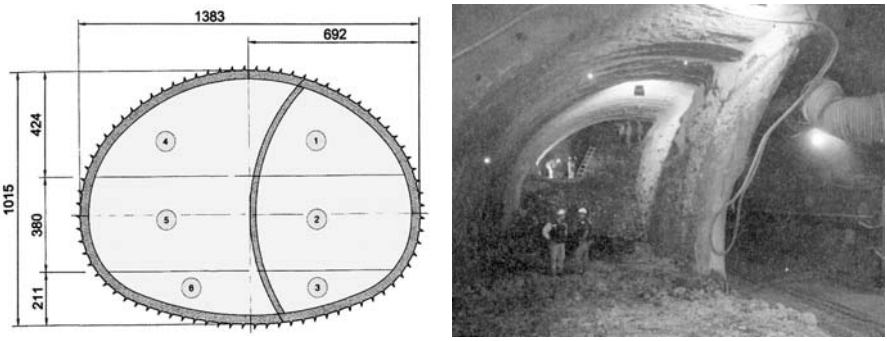


Fig. 10. Single sidewall drift station excavation, Metro de Santiago, Chile 2003 [6]

4 Safety

Under the conditions outlined above, even the most careful discrete computation with Finite Elements cannot provide a deterministic prediction that is certain to apply.

This computation model provides no indication as to how close a situation is to a state of collapse, and we also obtain no indication as to what conclusions the executor is supposed to derive from the deviations of the measurement results and the prediction.

However, this would be a necessary indication for carrying out an execution carefully.

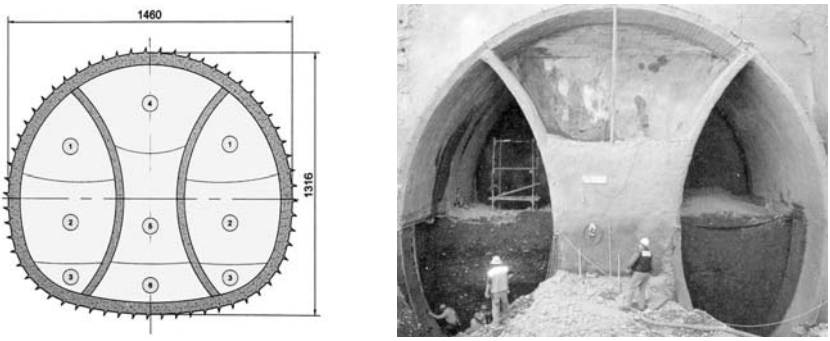


Fig. 11. Double sidewall drift station excavation, Metro de Santiago, Chile 2003 [6]

The damage that occurred on 20th/ 21st October, 1994 at London Heathrow Airport is shown in Figure 12 as an example of the consequences that can arise from a lack of reaction indication to the deviations between the measurement results and the prediction [18].

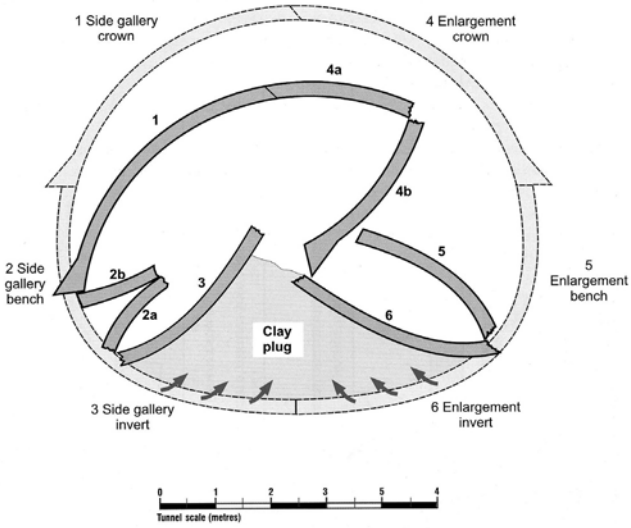


Fig. 12. Pattern of tunnel collapse, London Heathrow Airport; Damage on 20th / 21st October, 1994 [18]

Because of the responsibility for the safety of people at the construction site, the responsibility for the quality of the construction work and also because of the consequences of a criminal prosecution, a "critical" attitude to-

wards deterministic predictions derived from the computational models is appropriate.

In the final analysis, the deviations that occur can also be an indication of an incorrect computational model theory, because K. Popper's demands concerning the need for an experimental examination are in most cases not met.

Under these conditions, the meaning of some safety concepts is – if the veracity of the model is not examined – very questionable.

The question arises as to what importance can be attached to a failure probability statement if it is not possible to derive any clear distribution function from the soil parameters.

5 The probabilistic approach

Newer computation models that adopt a probabilistic approach attempt to give the impression that this mathematical method would be able to remedy the uncertainty in the description of the soil as a building material. It must not be forgotten here that probability theory by itself provides no information about the numerical values that can be assigned to concrete events [2].

What use (regarding his building) to the person responsible for a particular construction work is a statement that only applies if the law of large numbers is valid?

To this end, Fig. 13. shows as an example of a material collapse from an unknown gravel channel, which despite careful and tight exploratory drilling cannot always be detected. Due to the marl covering of less than 1.50 m, the hazard assessment stipulated securing with tube halls. A collapse occurred on the steep slope of the gravel channel when the halls were set [19].

In the event of the collapse of the gravel / water mixture, the hole in the excavation funnel was sealed with additional material such as timbering and wood wool. This enabled a collapse on the surface to be prevented [19].

Even the standards with the semi-probabilistic approach – a concept that does not exist in probability theory – cannot be helpful.

A statement according to which a geological formation can be "comprehensively covered" by a probabilistic model and tunnel driving in a particular "geological formation can be understood completely, comprehensibly and objectively" using the computational approach by a legal point of view outlined therein should be thought over after all [17].

In the probabilistic field in particular, the application of mathematics to physical phenomena is due to certain considerations of usefulness, which in turn are orientated towards special targets. This tool character which is subordinate to mathematics and which is well supported historically, is supposed to keep us from subordinating "nature" in a fatal Pythagorean Reverse ("the world of the number") to properties that are merely an expression of the systems of symbols used to describe them (Gericke 1993) [20].

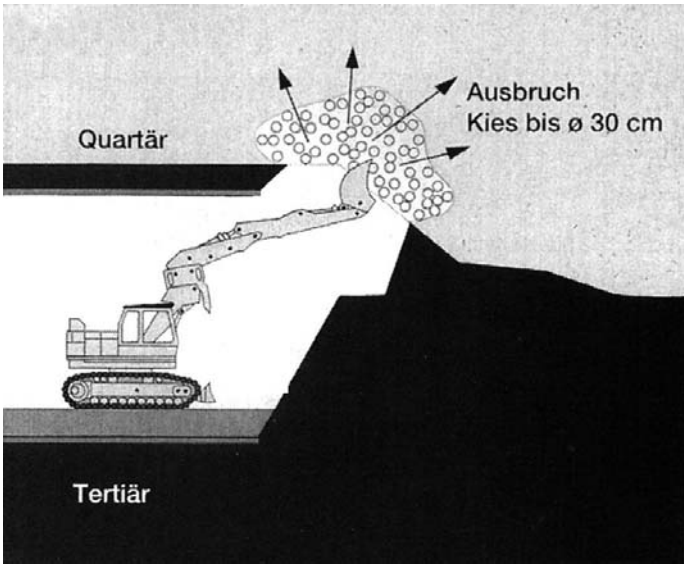


Fig. 13. Munich 1989 – 1994 subway line Candidplatz section, Compressed-air tunnel driving [19]

Probability theory is a mathematical branch that is just as precise as geometry, algebra or analysis, and it must not be mixed up with the conclusions derived from the application of the probabilistic model to the world in which we live [2].

It is the apparent weakness, but in reality the actual strength of the Kolmogoroff probability concept that it provides no information about the numerical values that can be assigned to concrete events. This is not a question of probability definition, but one of probability interpretation for the actual area of application. It is in fact advisable to differentiate between probability terms, primarily between the mathematical one and the one for describing the world. Countless discussions have been held about whether probabilities exist in nature or not [2].

It is entirely possible to observe the validity of the model for a large number of phenomena and to thereby obtain for it empirical confirmation [2]. However, if we consider that every construction work represents a unique item, the question arises as to how far probability theory is then valid at all.

The facts of the case, known as the *law of large numbers* after J. Bernoulli, allow still unknown probabilities to be inferred from relative frequencies. "The law of large numbers is therefore a *limit value statement* and calls for particular caution to be exercised when drawing conclusions" [2]. It must be noted here that even the precondition of the law of large numbers cannot be met, since it is known that every construction work is a unique thing.

"We are concerned with understanding and not just the knowledge or a description or a prediction. Yet theories that enable us to understand can display a high degree of general validity" [4].

"Theories are explanations, not just prediction" [4].

6 Information for the site engineer

In the case of some extremely complex theories relating to computational models, one can get the impression that they are an end in themselves and regard the main task – namely providing the necessary information for the executors – as negligible.

Executors require less deterministic predictions, and instead more understandable interval statements about possible measurable deformations, more statements about understandable behaviour patterns and, in particular, statements on possible collapse criteria and situations of collapse.

The intention should be to outline these statements in such a way that they can also be acted on later at the construction site.

Site engineer information as shown in Fig. 15., including the necessary explanations, is urgently required for the situation in the cross section shown in Fig. 14. [12].

7 Conclusion

The primary concern is the legal and contractual responsibility borne in the first instance by the site engineer. It appears too simple to assume that the computation (the mathematical model) corresponds with reality (the physical truth) and therefore to look first for the cause of such deviations possibly lying in the execution not having been carried out according to plans. Rather, it is the correspondence between the computation model and real behaviour that should be examined.

For the purpose of criminal law and economic assessment of failures, it will no longer be possible to assume that the computation model theory represents the sole criterion for assessment.

Particularly not if the evidence – based on K. Popper – of its usability cannot be furnished to a sufficient degree.

Too much is known and has also been written about the limits of many computation models.

However, the responsibility for failure is only part of the problem.

The much more important task is to consider how progress can be made with the uncertainties in the mechanical behaviour.

The potential possibility of failures should be further restricted. Undisputed here is the need to describe the existing uncertainties with mathematical model examinations and to show the effects in the model. However, it would

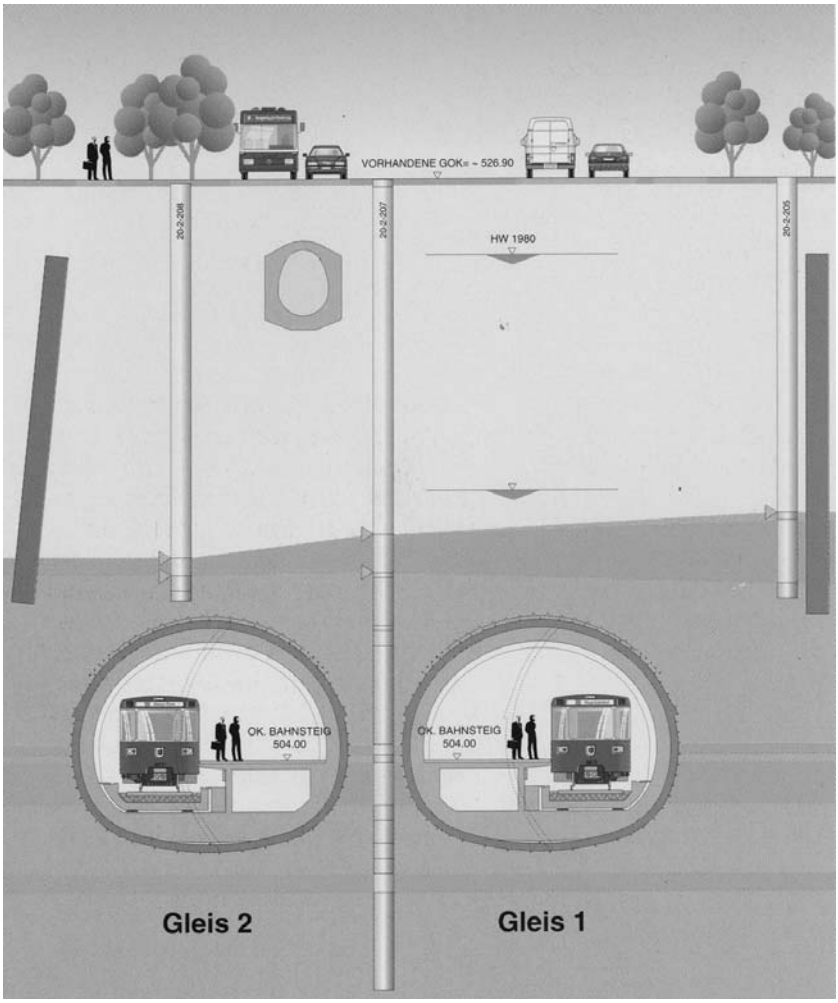


Fig. 14. Munich Trudering station – cross section

not be constructive to mask existing uncertainties. The strategy institute of the Boston Consulting Group has issued information about how one should handle uncertainties. They refer to historical knowledge like C. v. Clausewitz has it acquired in his famous book "About the War".

C. v. Clausewitz explains basic ideas on how to strategically deal with existing uncertainties: "Decisions that are made with the awareness of doubt are in most cases better than decisions in the case of which the doubt is pushed into the background. Pushing doubt into the background will backfire; correctly fostered doubt on the other hand can be used to support well thought-through actions" [14].

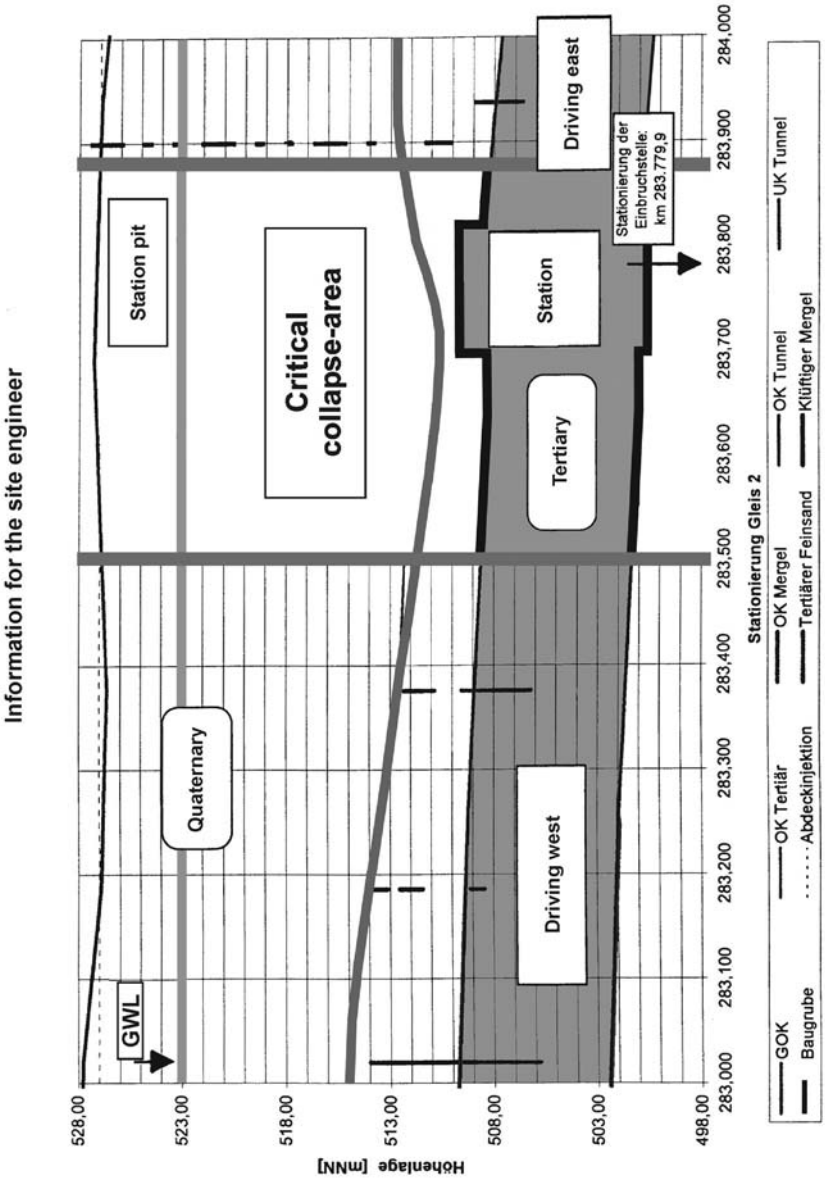


Fig. 15. Site engineer information that SHOULD be provided about the critical area, Munich Trudering [12]

The idea of particularly emphasising doubt and uncertainties could also be helpful to achieve comprehensible and useable solutions in the case of complex tasks, for example in geotechnics.

In the event of uncertainties, one should refrain from claiming mastery.

References

[1] Aicher M Baresel C Bilfinger+Berger et al, 25 Jahre U-Bahn-Bau in München, Schottenheim & Giess Offsetdruck KG, München
[2] Basieux P (1999) Abenteuer Mathematik, Brücken zwischen Wirklichkeit und Fiktion, Rowohlt Taschenbuch Verlag, Hamburg
[3] Brem G (1989) Analyse des Tagbruchs Witzelshöhetunnel der Neubaustrecke Hannover-Würzburg, Dissertation, Universität für Bodenkultur Wien, Institut für Bodenforschung und Baugeologie, Wien
[4] Deutsch D (2000) Die Physik der Welterkenntnis, Auf dem Weg zum universellen Verstehen, Deutscher Taschenbuchverlag, München
[5] Duddeck H (1994) Die wesentliche Herausforderung des Tunnelbauingenieurs, der Baugrund, Geotechnik, Verlag Glückauf, Essen
[6] Egger K Mercado C Chamorro G (2003) Santiagos U-Bahn expandiert, Felsbau, Nr 5, Verlag Glückauf, Essen
[7] Fellin W (2001) Zur Auswirkung der Festlegung charakteristischer Bodenkennwerte am Beispiel einer Grundbruchberechnung, Geotechnik 24, Nr 3, Verlag Glückauf, Essen
[8] Feynman R P (2003) Sechs physikalische Fingerübungen, Piper Verlag, München Zürich
[9] Friedrichsen G (1998) Wos für ein Einbruch, Der Spiegel Nr 19, Spiegel-Verlag, Hamburg
[10] Geier M (1994) Karl Popper – dargestellt von Manfred Geier, Rowohlt Taschenbuch Verlag, Reinbek bei Hamburg
[11] Herle I Difficulties related to geotechnical calculations, Added in this publication
[12] Jessberger H L (1997) Gutachten zum Schadensfall U-Bahn Linie 2-Ost Los 2 München Trudering Teil 1, Bochum
[13] Kolymbas D (2003) Theoretische Bodenmechanik, Vorlesung 23.01.2003, Universität Innsbruck
[14] Oetinger B Ghyczy T Bassford C (2003) Clausewitz Strategie denken, Deutscher Taschenbuch Verlag, München
[15] Schikora K Eierle B (2000) Tunnel Farchant - Geologie Vortrieb Berechnungen und Messungen, Der Bauingenieur Nr 75, Springer-Verlag
[16] Spaun G Beispiele von Verbrüchen im Tunnelbau (Teil 1), Vorträge zu der Festtagung der Tiefbau-Berufsgenossenschaft, Tiefbauberufsgenossenschaft
[17] Schweiger H F Pöttler R Thurner R (2002) Verknüpfung von Probabilistik und Numerik im Tunnelbau, Felsbau, Nr 20, Verlag Glückauf, Essen

[18] The collapse of NATM tunnels at Heathrow Airport, A report on the investigation by the Health and Safety Executive into the collapse of New Austrian Tunnelling Method (NATM) tunnels at the Central Area of Heathrow Airport on 20/21 October 1994, HSE Books, Sudbury 2000
[19] Tunnelbau-Symposium (1995) Erfahrungsaustausch zu aktuellen Fragen im Tunnelbau, TU München
[20] Wehr M (2002) Der Schmetterlingsdefekt, Turbulenzen in der Chaostheorie, Klett-Cotta Verlag, Stuttgart

Fuzzy, probabilistic and stochastic modelling of an elastically bedded beam

Michael Oberguggenberger[1] and Francesco Russo[2]

[1] Institut für Technische Mathematik, Geometrie und Bauinformatik, Universität Innsbruck
[2] Institut Galilée, Université Paris-Nord, Villetaneuse

Summary. This article sets out to compare the effects of modelling uncertainty using fuzzy sets, random variables and stochastic analysis. With the aid of an example from civil engineering - the bending equation for an elastically bedded beam - we discuss what each model is or is not capable of capturing. All models may be adequately used for variability studies, but may fail to detect the effect of localized parameter fluctuations on the response of the system. In the stochastics setting, we show an instance of the linearization effect of large noise which says that under large stochastic excitations, the contributions of nonlinear terms may be annihilated.

1 Introduction

This article addresses the question of how to model parameter uncertainties in engineering problems. Among the many possibilities we wish to compare three approaches: fuzzy, probabilistic and stochastic. The question which approach is the most adequate can only be decided by looking at the consequences of using one or the other model - in particular, whether the desired or exptected effects are adequately captured. This we intend to demonstrate in a typical application in civil engineering - the computation of the displacement of an elastically bedded beam.

Suppose the response of an engineering structure is described e. g. by a differential equation. Apart from the design dimensions, certain parameters describing exterior and material properties enter as input data. These parameters exhibit variations and are only imprecisely known. The question is how the variability of the data affects the response and consequently the safety of the structure. To be more specific, assume that a certain soil parameter c has been measured or estimated from a geological assessment at various locations on a construction site, so that a list of values $c_1, \ldots, c_n$ is available. This could be interpreted as n realizations of a constant parameter c, or as realizations of a spatially variable field $c(x)$, where x denotes a spatial coordinate. Commonly, in civil engineering, the size of the sample is

Published originally in: G. de Cooman, T. Fine, T. Seidenfeld (eds.), *ISIPTA'01, Proceedings of the Second Symposium on Imprecise Probabilities and Their Applications.* Shaker Publ BV, Maastricht, 2001, pp. 293–300.

small and the data are augmented by prior expert knowledge. In the first, spatially constant case, the parameter c might thus be modelled either as a fuzzy number or as a random variable in the sense of subjective probability (a frequentist interpretation appears inappropriate with data as described above). In the second case, $c(x)$ might be modelled as a stationary random field, where the choice of correlation function depends on prior knowledge with possible adaptation using the data. We will address the following issues:

- What is the solution concept appropriate to the fuzzy, probabilistic or stochastic approach?
- Is the variability and the sensitivity of the response of the system adequately predicted?
- Can the model describe localization effects, that is, the possibly deviating response due to a localized, rapid change in the parameter c, even though within the boundaries given by its fuzzy or random range?
- How are nonlinear terms in the differential equation affected by a stochastic input?

Our principal findings which we wish to report here are: the fuzzy, probabilistic and stochastic approach all can - in their respective terms - describe the variability of the outcome and admit a study of its sensitivity. However, the response due to localization may considerably exceed the boundaries predicted by the fuzzy or probabilistic model. What concerns the stochastic model, we wish to point out the nonlinear trivialization effect of large noise: if the stochastic excitation is derived from a multiple of white noise which becomes larger and larger, the response of the nonlinear system may tend to the response of the linear part of the system in the mean.

The plan of the paper is as follows: In Section 2 we present the engineering example underlying our investigation and collect its properties when the parameters are deterministic. Section 3 is devoted to the fuzzy and probabilistic approach. Section 4 contains a typical stochastic model. We prove existence and uniqueness of the solution and show that the mentioned nonlinear triviality effect occurs in the example. In the last section we summarize the conclusions for engineering design suggested by our observations.

In the current engineering literature, especially in geotechnics, the problems addressed here are of high concern. The mechanisms of failure have been investigated, for example, in [3, 9, 22] using a fuzzy, probabilistic, and stochastic approach, whereas questions of modelling uncertainty in engineering problems are addressed e. g. in [8, 11, 12, 15, 21]; for the localization phenomenon we refer to [6, 7, 13]. The nonlinear trivialization effect was discoverd by [1, 2] and has been studied in a number of papers by the authors [17, 18, 19, 23]. For an exposition of various models of uncertainty in connection with fuzzy sets we refer to [14].

2 The elastically bedded beam

The displacement $w(x), x \in \mathbb{R}$ of an infinite beam on a linear elastic bedding is described by the bending equation

$$EI\, w^{IV}(x) + bc\, w(x) = q(x),\ -\infty < x < \infty\,,$$

see e. g. [5, Section 61]. Here EI is the flexural rigidity of the beam, b its effective width, c the bearing coefficient of the foundation and $q(x)$ the loading. One may imagine that the beam describes a buried pipeline, the loading $q(x)$ resulting from the covering soil. The parameters EI and b of the beam may be considered as precisely known, whereas the soil properties c and q vary in an imprecisely known fashion. A more refined model would allow for a nonlinear elasto-plastic response of the founding soil, say

$$EI\, w^{IV}(x) + bc\, w(x) + F(w(x)) = q(x)\,.$$

In this situation, a typical nonlinear function would be antisymmetric and monotonically increasing with limits $\pm L$ at $\pm\infty$. The terms $bc\, w(x) + F(w(x))$ together would signify a hardening property of the foundation. We will study the singular boundary value problem for the standardized equation

$$u^{IV}(x) + 4k^4 u(x) = p(x),\ -\infty < x < \infty \tag{1}$$

requiring that the solution should remain bounded at $\pm\infty$. In case k is a constant and $p(x)$ is an integrable function, both deterministic, its unique classical solution is given by

$$u(x) = \int_{-\infty}^{\infty} G(x,y)\, p(y)\, dy$$

in terms of its Green function

$$G(x,y) = \frac{1}{8k^3} e^{-k|x-y|} \big(\sin k|x-y| + \cos k|x-y| \big)\,.$$

In case the standardized load $p(x) \equiv p$ is constant, the displacement is constant as well and simply given by

$$u(x) \equiv \frac{p}{4k^4}\,.$$

For the computational examples to follow we let the parameters vary around central moduli of $k = 10^{-2},\ p = 10^{-8}$. Approximately, this corresponds to the case of a buried steel pipeline with an effective diameter of 6 [cm], covered by about 100 [cm] of top soil and bedded in loosely packed sand. The resulting overall displacement would amount to $u(x) \equiv 0.25$ [cm] in the deterministic case.

3 Fuzzy and probabilistic modelling

The collection of fuzzy subsets of a set X will be denoted by $\mathbb{F}(X)$. A fuzzy set $A \in \mathbb{F}(X)$ is determined by its membership function $m_A(x), 0 \le m_A(x) \le 1$, describing the degree of membership of the value x in A, respectively the degree of possibility that the object described by A assumes the value x. The sets $[A]^\alpha = \{x : m_A(x) \ge \alpha\}$ are the α-level sets of A. Given a function $f : X \to Y$, the Zadeh extension principles allows to extend it to a function $f : \mathbb{F}(X) \to \mathbb{F}(Y)$ by

$$m_{f(A)}(y) = \sup \big\{ m_A(x), x \in f^{-1}(y) \big\}\,.$$

An element $A \in \mathbb{F}(\mathbb{R}^d)$ is called a fuzzy vector, if each of its α-level sets is convex and compact $(0 < \alpha < 1)$, and $[A]^1$ contains exactly one point. In the case of

$d = 1$ we refer to A as a fuzzy number. If $f : \mathbb{R}^d \to \mathbb{R}$ is continuous and A a fuzzy vector, the function value $f(A)$ is a fuzzy number. In addition, its level sets are obtained as the images of the level sets of the data A under the function f, that is, $[f(A)]^\alpha = f([A]^\alpha)$. For more details on this and on fuzzy sets in general we refer to [4, 10].

We return to the model (1) of the elastically bedded beam, assuming that the parameter k is described by a fuzzy number K, whereas the load $p(x) = \sum_{i=1}^n a_i p_i(x)$ is given by a combination of fixed shape functions $p_i(x)$ with parameters a_i, which in turn will be interpreted as fuzzy numbers $A_i, i = 1, \ldots, n$. We model the data as a non-interactive fuzzy vector in $\mathbb{R}^{(1+n)}$. This notion corresponds to independence in probability theory and signifies that the joint membership function is given by the formula $m_{(K,A_1,\ldots,A_n)}(k, a_1, \ldots, a_n) = \min\{m_K(k), m_{A_1}(a_1), \ldots, m_{A_n}(a_n)\}$. We shall compute the fuzzy point values of the fuzzy response $u(x)$ by applying the Zadeh extension principle to the solution operator

$$(k, a_1, \ldots, a_n) \longrightarrow u(x) = L_x(k, a_1, \ldots, a_n)\,. \tag{2}$$

By the discussion above, an α-level sets of $L_x(K, A_1, \ldots, A_n)$ is computed as the collection of the values of the solutions attained when the parameters vary in the respective level sets $[K]^\alpha, [A_1]^\alpha, \ldots, [A_n]^\alpha$.

In fact, the Zadeh extension principle allows to derive a fuzzy solution concept for which this is the unique fuzzy solution. To achieve this, equation (1) is rewritten as a prolongued system for the variables $(u, h, b_1, \ldots, b_n)$,

$$u^{IV}(x) + 4h^4 u(x) - \sum_{i=1}^n b_i p_i(x) = 0\,, \tag{3}$$

$$h' = 0, b_1' = 0, \ldots, b_n' = 0\,, \tag{4}$$

$$h(0) = k, b_1(0) = a_1, \ldots, b_n(0) = a_n\,. \tag{5}$$

The parameters $h = k, b_i = a_i, i = 1, \ldots, n$ are actually constants. The device incorporated in (4), (5) allows to interpret them as variables, so formally equation (3) contains no "parameters". This is important, because substitution of a function with fuzzy parameters in a term depending on the same parameters need not preserve equalities (see e. g. the discussions in [10, 16]). The equation operator $E : (\mathcal{C}_b^4(\mathbb{R}))^{2+n} \to (\mathcal{C}_b^0(\mathbb{R}))^{2+n}$ is defined by the left hand sides of (3), (4) on the space of bounded and four times differentiable functions, the restriction operator $R : (\mathcal{C}_b^4(\mathbb{R}))^{2+n} \to \mathbb{R}^{1+n}$ by the left hand sides of (5). The solution operator is

$$L : \mathbb{R}^{1+n} \to (\mathcal{C}_b^4(\mathbb{R}))^{2+n}\,,$$
$$(k, a_1, \ldots, a_n) \to (u, h, b_1, \ldots, b_n),$$

where, by construction, $h = k$ and $b_i = a_i$. An element $X = (U, H, B_1, \ldots, B_n) \in \mathbb{F}\big((\mathcal{C}_b^4(\mathbb{R}))^{2+n}\big)$ is called a fuzzy solution of system (3), (4), (5), if

$$E(X) = 0 \text{ in } \mathbb{F}\big((\mathcal{C}_b^0(\mathbb{R}))^{2+n}\big)\,,$$
$$R(X) = (K, A_1, \ldots, A_n) \text{ in } \mathbb{F}(\mathbb{R}^{1+n})\,.$$

Here 0 denotes the crisp zero function; note that the derivatives of the fuzzy function X occurring in $E(X)$ are defined by means of the extension principle, as are the other operations performed on it.

Proposition 1 *Given* $(K, A_1, \ldots, A_n) \in \mathbb{F}(\mathbb{R}^{1+n})$*, problem (3), (4), (5) has a unique solution* $X = (U, H, B_1, \ldots, B_n) \in \mathbb{F}\big((\mathcal{C}_b^4(\mathbb{R}))^{2+n}\big)$*. It is given by* $U = L(K, A_1, \ldots, A_n)$*.*

The proof follows by adapting the arguments given in [16] for the pure initial value problem. In addition, one can show that the restriction of the fuzzy solution X to a spatial point x is simply given by fuzzifying (2), thus by the fuzzy number $L_x(K, A_1, \ldots, A_n)$.

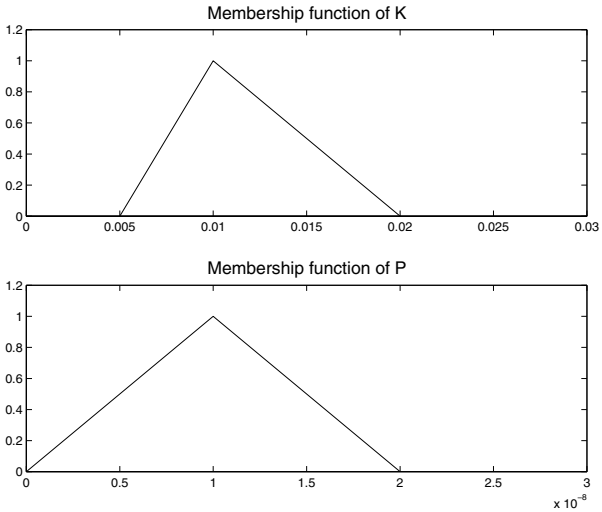


Fig. 1. Fuzzy parameters K and P.

In the example to follow, we assume that p is constant. For the sake of exposition, we take both K and P as triangular fuzzy numbers, centered around the values indicated in Section 1, namely

$$K = \langle \tfrac{1}{2}, 1, 2 \rangle \cdot 10^{-2}, \tag{6}$$

$$P = \langle 0, 1, 2 \rangle \cdot 10^{-8}, \tag{7}$$

see Figure 1. Then the (constant) fuzzy solution is simply given by the fuzzy number $L_x(K, P) = P/4K^4$. It is depicted in Figure 2.

We observe that the fuzzy model correctly describes the fluctuations of the response in dependence of the data variability. In addition, the α-level structure provides a good picture of the sensitivity of the result.

On the other hand, assume now that the parameter p varies in the range given by its fuzzy description, in the example its support $[0, 2 \cdot 10^{-8}]$, but its realization is not actually constant. For example, it might jump from 0 to $2 \cdot 10^{-8}$ at the point $x = 0$. This means that we have to solve equation (1) with a load

$$p(x) = \begin{cases} 0, & x < 0, \\ p = 2 \cdot 10^{-8}, & x > 0. \end{cases}$$

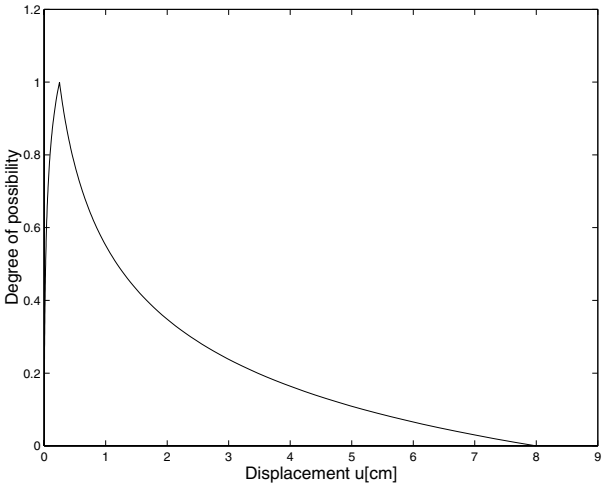


Fig. 2. Fuzzy displacement u.

The corresponding displacement is

$$w(x) = \begin{cases} \frac{p}{8k^4}\, e^{kx} \cos kx, & x < 0, \\ \frac{p}{8k^4} \left(2 - e^{-kx} \cos kx\right), & x > 0. \end{cases}$$

Taking the admissible value $k = \frac{1}{2} \cdot 10^{-2}$, it is seen that the the graph of w exceeds the band described by the fuzzy displacement with constant parameters. This is an example of the effect, observed in other circumstances as well, that a localized parameter fluctuation may produce a response not predicted by the fuzzy model. In Figure 3, the 0-level set of the fuzzy displacement is indicated by the horizontals $u = 0$ and $u = 8$, the curve is the graph of w, and the other horizontals indicate level sets of the fuzzy displacement u for $\alpha = 0.2, 0.4, 0.6, 0.8$. Degree of possibility equal to one occurs at $u = 0.25$.

We now turn to a probabilistic modelling of system (1). We assume again that the load is of the form $p(x) = \sum_{i=1}^{n} a_i p_i(x)$ as before. In the probabilistic approach, we describe the parameters $k, a_1, \ldots, a_n$ as random variables on a probability space (Ω, Σ, μ). In this case we obtain the solution as a random process

$$(x, \omega) \in \mathbb{R} \times \Omega \longrightarrow u(x, \omega)\,.$$

It is well known and easy to prove that, for almost all $\omega \in \Omega$, a pathwise solution exists. These pathwise solutions are almost surely unique and differentiable. The point values $u(x)$ are random variables. In continuation of the computational example above, we take p constant with a triangular distribution with defining boundaries as in (7), and similarly for the parameter k with boundaries as in (6). We approximate the distribution of the solution $u(x) \equiv p/4k^4$ numerically by simulation (using 10^5 realizations), the result being depicted in Figure 4.

For reasons of comparison, we have chosen the same shape for the probability distribution of k and p as for the possibility distribution in the fuzzy case. As expected, the probabilistic result centers around the probable outcome, whereas

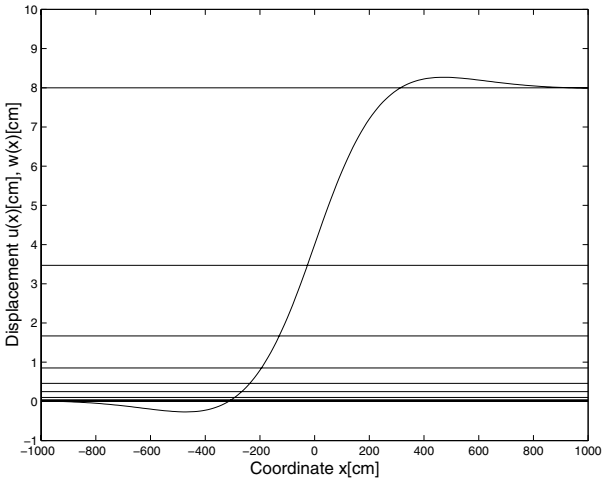


Fig. 3. Fuzzy solution versus effect of localization.

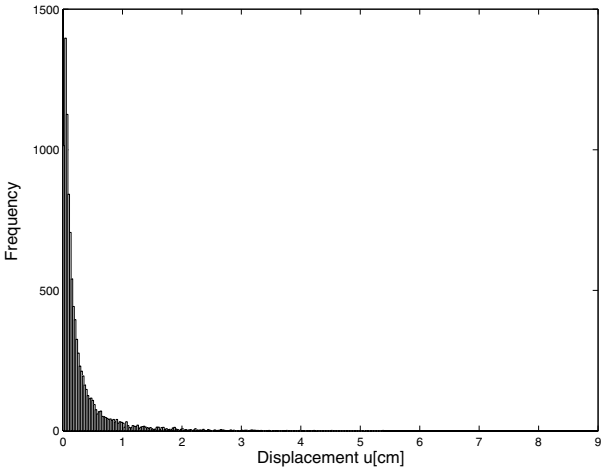


Fig. 4. Frequency distribution of random displacement.

the fuzzy result exhibits the possible variation of the outcome. The picture needs no further explication; it is clear that the probabilistic result fails to cover the response due to a localized parameter fluctuation all the more.

4 Stochastic modelling

By a stochastic model in the context of the elastically bedded beam we mean that the parameters are described by spatial random fields or, in one dimension, random

processes. We begin with the linear case. To fix ideas, we assume that the parameter k in (1) is constant, and the load is modelled by a multiple of Gaussian white noise, $p(x) = a\dot{W}(x)$. To adapt the model to real data, one might actually take a more general diffusion process and add a deterministic mean value. However, this does not change the effects we want to describe and so we shall use this simpler model, namely,

$$v^{IV}(x) + 4k^4 v(x) = a\dot{W}(x)\,. \tag{8}$$

As noted, white noise $\dot{W}$ is a stationary Gaussian process. Its mean and covariance are given by

$$\mathrm{E}\big(\dot{W}(x)\big) = 0,$$
$$\mathrm{COV}\big(\dot{W}(x), \dot{W}(y)\big) = \delta(x-y)\,,$$

where $\delta(.)$ denotes the Dirac measure. Thus $\dot{W}$ has to be interpreted as a generalized stochastic process, in fact, it is the generalized derivative of Wiener process W. In turn, the stochastic differential equation is interpreted as an Itô-differential equation (see e. g. [20] as a reference). Its (almost surely unique) solution is given by the Itô-integral

$$v(x) = a\int_{-\infty}^{\infty} G(x,y)\,dW(y)\,,$$

where $G(x,y)$ is the Green function. The Itô-integral can be defined as the mean square limit

$$\mathrm{L}^2 - \lim_{\Delta y \to 0} a \sum_{j=-\infty}^{\infty} G(x, j\Delta y)\,\Delta W_j\,,$$

where $\Delta W_j = W((j+1)\Delta y) - W(j\Delta y)$ denote the increments of Wiener Process. These are known to be independent and normally distributed according to $\mathcal{N}(0, \Delta y)$. This in turn indicates how the solution can be approximated by simulation. An approximate realization of white noise is shown in Figure 5, a number of sample paths of the Itô-solution are depicted in Figure 6. The constant a has been fixed in this computational example so that the standard deviation of the solution $v(x)$ coincides with the value 0.25 [cm] used for the deterministic displacement in Section 2. For this as well as later use we record the following result; the star denotes convolution.

Lemma 1. *The covariance of the Itô-solution v to problem (8) is*

$$\mathsf{COV}\big(v(x), v(z)\big) = a^2 S * S(|x-z|)$$

where $S(y) = G(0,y)$ denotes the fundamental solution of equation (1).

Proof: By the isometry property of Itô-integrals and using that $G(x,y) = S(x-y) = S(|x-y|)$, we have that

$$\mathsf{E}\big(v(x)v(z)\big) = a^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(x-y)S(z-y)\,dy$$
$$= a^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(x-z+y)S(y)\,dy = a^2 S * S(|x-z|)\,,$$

which proves the result, since $\mathsf{E}(v(x)) \equiv 0$.

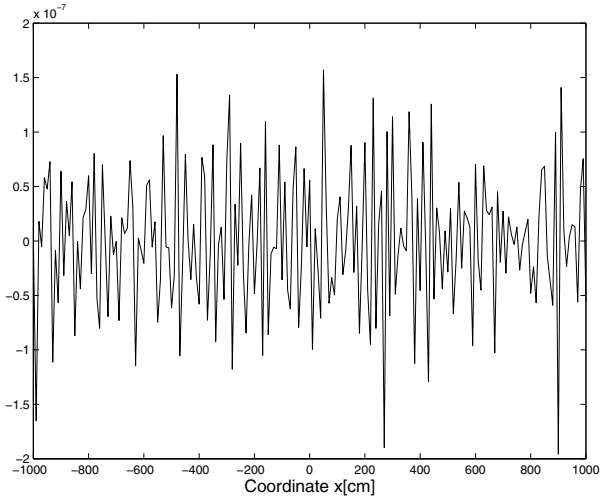


Fig. 5. An approximate sample path of white noise.

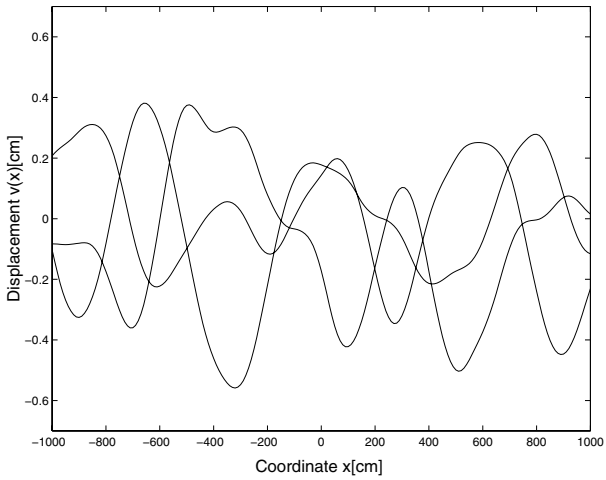


Fig. 6. Three sample paths of the solution v.

In particular,

$$\mathsf{V}\big(v(x)\big) = a^2 S * S(0) > \mathsf{COV}\big(v(x), v(z)\big) \tag{9}$$

whenever $x \neq z$. In fact, a little Fourier analysis shows that

$$S * S(r) = \frac{1}{\pi} \int_0^\infty \cos r\xi \, \frac{d\xi}{(\xi^4 + 4k^4)^2} \, ,$$

whence

$$S * S(0) - S * S(r) = \frac{1}{\pi} \int_0^\infty \left(1 - \cos r\xi\right) \frac{d\xi}{(\xi^4 + 4k^4)^2}$$

is strictly greater than 0 whenever $r \neq 0$.

Our next task is the investigation of the nonlinear, stochastic bending equation

$$u^{IV}(x) + 4k^4 u(x) + F(u(x)) = a\dot{W}(x)\,, \tag{10}$$

where we assume that F is Lipschitz continuous with a global Lipschitz constant $\mathrm{Lip}(F)$. We interpret (10) as an integral equation, namely

$$\begin{aligned} u(x) = &- \int_{-\infty}^{\infty} G(x,y)\, F(u(y))\, dy \\ &+ a \int_{-\infty}^{\infty} G(x,y)\, dW(y)\,. \end{aligned} \tag{11}$$

We shall construct a solution in the space $\mathcal{C}_b = \mathcal{C}_b^0(\mathbb{R} : \mathrm{L}^2(\Omega))$ of stochastic processes with bounded second moments. This is a Banach space, equipped with the norm

$$||u|| = \sup_{x\in\mathbb{R}} \left(\mathrm{E}|u(x)|^2\right)^{1/2}.$$

Proposition 2 *Under the condition Lip$(F) < 4k^4$, the integral equation (11) has a unique solution u in the space $\mathcal{C}_b$.*

Proof: First, the operator $\mathcal{L}$ defined by the right hand side of (11) maps the space $\mathcal{C}_b$ into itself. To see this, one has to employ the Lipschitz property of F, the mentioned isometry property of Itô-integrals, and the fact that the Green function $y \to G(x,y)$ is locally uniformly square integrable due to the inequality $e^{-|x-y|} \leq e^{|x|-|y|}$. The arguments are similar to the ones proving the contraction property which we now present in more detail:

$$\begin{aligned} &\sup_{x\in\mathbb{R}} \left(\mathsf{E}\left|\mathcal{L}u(x) - \mathcal{L}\tilde{u}(x)\right|^2\right)^{1/2} \\ &= \sup_{x\in\mathbb{R}} \left(\mathsf{E}\Big(\int_{-\infty}^{\infty} G(x,y)(F(u(y)) - F(\tilde{u}(y)))\, dy\Big)^2\right)^{1/2} \\ &\leq \sup_{x\in\mathbb{R}} \int_{-\infty}^{\infty} |G(x,y)|\, \mathrm{Lip}(F) \left(\mathsf{E}|u(y)) - \tilde{u}(y)|^2\right)^{1/2} dy \\ &= \frac{1}{4k^4} \mathrm{Lip}(F) \sup_{y\in\mathbb{R}} \left(\mathsf{E}|u(y)) - \tilde{u}(y)|^2\right)^{1/2} \end{aligned}$$

since the integral of the Green function at fixed x equals $1/4k^4$. Thus $\mathcal{L}$ is a contraction on $\mathcal{C}_b$ if the Lipschitz constant is less than $4k^4$; the integral equation (11) has a unique fixed point $u \in \mathcal{C}_b$, as required.

We remark that this fixed point $u(x)$ provides a pathwise solution of the differential equation (10) in the sense of generalized functions (Schwartz distributions). For the notions from distribution theory which we need here and below we refer to [24].

We now turn to the nonlinear trivialization effect of large noise. This refers to a comparison of the solutions of the nonlinear equation

$$u^{IV}(x) + 4k^4 u(x) + F(u(x)) = a\dot{W}(x)\,,$$

and the linear equation

$$v^{IV}(x) + 4k^4 v(x) = a\dot{W}(x)\,.$$

as a tends to infinity. We shall prove that - for a large class of nonlinearities F - the difference $u - v$ tends to zero in the mean. This class of nonlinearities has been introduced in [23]. For its definition we need the Fourier transform of bounded functions, hence the Schwartz theory of distributions is again required. Recall that $\mathcal{S}'(\mathbb{R})$ denotes the space of tempered distributions, the collection of continuous functionals on the space $\mathcal{S}(\mathbb{R})$ of rapidly decreasing functions. The action of a distribution T on a function ψ is denoted by $\langle T, \psi\rangle$.

Definition 1. *A distribution $T \in \mathcal{S}'(\mathbb{R})$ is said to be massless at zero, if $\langle T, \psi(\cdot/\varepsilon)\rangle$ tends to zero as $\varepsilon \to 0$ for the function $\psi(y) = \exp(-y^2/2)$.*

We shall be concerned with bounded functions F whose Fourier transform $\mathcal{F}F$ is massless at zero. This is the case if and only if

$$\lim_{\varepsilon\to 0} \varepsilon \int_{-\infty}^{\infty} F(y)\, e^{-\varepsilon^2 y^2/2}\, dy = 0\,.$$

Here are some examples of functions F whose Fourier transform $\mathcal{F}F$ is massless at zero (see [23] for details and further examples):

(a) Every bounded, continuous function F such that $\lim_{y\to-\infty} F(y) = -L$ and $\lim_{y\to\infty} F(y) = +L$;

(b) every continuous, periodic function F whose first derivative is locally square integrable;

(c) every element $F \in \mathrm{L}^p(\mathbb{R})$ when $1 \le p \le 2$.

The significance of this notion lies in the following result, the proof of which can be found in [23] as well.

Proposition 3 *Let F be a bounded function whose Fourier transform $\mathcal{F}F$ is massless at zero. Let $(v_a, w_a), a > 0$, be non-degenerate Gaussian vectors and assume that the variance $V(v_a)$ tends to infinity as $a \to \infty$. Then*

$$\mathsf{E}\big(F(v_a)F(w_a)\big) \to 0 \text{ as a} \to \infty\,.$$

The nonlinear trivialization effect of large noise is expressed in the following result.

Theorem 1 *Assume that the function F is bounded with Fourier transform $\mathcal{F}F$ massless at zero. Further, assume that Lip(F) $< 2k^4$. Let u be the solution constructed above to the nonlinear stochastic differential equation (10) and v the solution to the linear equation (8). Then*

$$\mathsf{E}\Big(\int_K |u(x) - v(x)|\, dx\Big) \to 0 \text{ as } a \to \infty$$

for every compact interval $K \subset \mathbb{R}$.

Proof: The difference of u and v satisfies the equation

$$(u-v)^{IV} + 4k^4(u-v) + F(u) - F(v) = -F(v)\,.$$

Applying the Green function and taking the expectation we get

$$\begin{aligned}\mathsf{E}\Big(u(x)-v(x) + \int_{-\infty}^{\infty} G(x,y)\big(F(u(y))-F(v(y))\big)dy\Big)^2 \\ = \mathsf{E}\Big(\int_{-\infty}^{\infty} G(x,y)F(v(y))\,dy\Big)^2 \\ = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x,y)G(x,z)\,\mathsf{E}\Big(F(v(y))F(v(z))\Big)dydz\,.\end{aligned}$$

Denote the last term by $h^2(x)$. Due to Proposition 3, the boundedness of F and the integrability of the Green function, $h^2(x)$ tends to zero as $a \to \infty$ and remains bounded. By the triangle and Hölder's inequality it follows that

$$\begin{aligned}\mathsf{E}\,|u(x)-v(x)| \\ \le \mathsf{E}\,\Big|\int_{-\infty}^{\infty} G(x,y)\big(F(u(y))-F(v(y))\big)dy\,\Big| + h(x)\,.\end{aligned}$$

Using the fact that $G(x,y) \le e^{-|kx-ky|}/4k^3$ and the elementary inequality $-\frac{|kx|}{1+b} - |kx-ky| \le -\frac{|ky|}{1+b} - \frac{b}{1+b}|kx-ky|$ for $b>0$, we obtain

$$\begin{aligned}\mathsf{E}\int_{-\infty}^{\infty} e^{-|kx|/(1+b)}|u(x)-v(x)|\,dx \\ \le \mathrm{Lip}(F)\,\mathsf{E}\,\Big(\frac{1}{4k^3}\int_{-\infty}^{\infty} e^{-b|kx-ky|/(1+b)}\,dx \\ \cdot \int_{-\infty}^{\infty} e^{-|ky|/(1+b)}|u(y)-v(y)|\,dydx\Big) \\ + \int_{-\infty}^{\infty} e^{-|kx|/(1+b)}h(x)\,dx \\ \le \frac{1+b}{2bk^4}\,\mathrm{Lip}(F)\,\mathsf{E}\int_{-\infty}^{\infty} e^{-|ky|/(1+b)}|u(y)-v(y)|\,dy \\ + \int_{-\infty}^{\infty} e^{-|kx|/(1+b)}h(x)\,dx\,.\end{aligned}$$

Thus if the coefficient in front of the first integral is less than one, which is the case when $\mathrm{Lip}(F) < 2k^4$ for b chosen large enough, then

$$\mathsf{E}\int_{-\infty}^{\infty} e^{-|kx|/(1+b)}|u(x)-v(x)|\,dx \to 0 \text{ as } a \to \infty$$

due to the properties of $h(x)$ derived above. This, in particular, implies the assertion of the theorem.

The theorem asserts that under quite general conditions on the nonlinear term, the stochastic solution to the nonlinear equation behaves like the stochastic solution to the corresponding linear equation in the mean. Thus large noise may annihilate nonlinearities in the model. This is the trivialization effect. We note that the condition on the Lipschitz constant of F can be improved by considering other norms. In addition, if the noise is concentrated on a compact interval, the trivialization effect occurs for all F as in the theorem with $\mathrm{Lip}(F) < 4k^4$.

5 Summary and Conclusions

We have investigated the assertions of a fuzzy, probabilistic and a stochastic model in a typical example from civil engineering. The approaches formulate uncertainty in different ways; the modelling assumptions differ, as do the interpretations of the results. One may state that all models capture and process the uncertainty and provide insight into the sensitivity of the system. They highlight different aspects and neglect others. For example, the fuzzy and the probabilistic approach evaluate the width of fluctuations quite differently; the stochastic approach may suppress the effect of nonlinearities (this, by the way, is an example of a certain smoothing observed otherwise as well). But most importantly, neither of the models - if set up in the way as done here - fully detects the effects of a localized parameter fluctuation on the response of the system.

As a consequence for engineering practice and design, we conclude that a safety analysis must include both a study of the variability, as can be achieved by the models under discussion, and an investigation into specific failure mechanisms that may arise from local disturbances. We focused our discussion on the fuzzy, probabilistic and stochastic approach for reasons of exposition. A similar analysis can be carried through in the framework of other theories of imprecise probabilities as well.

Acknowledgement. We thank H. Lehar and R. Stark for valuable advice on engineering questions.

References

[1] S. Albeverio, Z. Haba and F. Russo. Trivial solutions for a non-linear two-space dimensional wave equation perturbed by space-time white noise. *Stochastics and Stochastics Reports,* 56:127-160, 1996.

[2] S. Albeverio, Z. Haba, F. Russo. A two-space dimensional semilinear heat equation perturbed by (Gaussian) white noise. *Probability Theory and Related Fields,* 121:319-366, 2001.

[3] G. Auvinet and J.L. González. Three-dimensional reliability analysis of earth slopes. *Computers and Geotechnics,* 26:247-261, 2000.

[4] H. Bandemer, W. Näther. Fuzzy Data Analysis. *Kluwer Acad. Publ.,* 1992.

[5] V. V. Bolotin. Statistical Methods in Structural Mechanics. *Holden-Day, Inc.,* San Francisco, 1969.

[6] R. I. Borja, R. A. Regueiro and T. Y. Lai. FE modeling of strain localization in soft rock. *Journal of Geotechnical and Geoenvironmental Engineering,* 126:335-343, 2000.

[7] R. Chambon, D. Caillerie and N. El Hassan. One-dimensional localisation studied with a second grade model. *European Journal of Mechanics A/Solids,* 17:637-656, 1998.

[8] O. Ditlevsen, N.J. Tarp-Johansen and H. Denver. Bayesian soil assessments combining prior with posterior censored samples. *Computers and Geotechnics,* 26:187-198, 2000.

[9] G.R. Dodagoudar and G. Venkatachalam. Reliability analysis of slopes using fuzzy sets theory. *Computers and Geotechnics,* 27:101-115, 2000.

[10] D. Dubois, H. Prade. Possibility Theory. *Plenum Press,* New York, 1988.
[11] Th. Fetz, J. Jäger, D. Köll, G. Krenn, H. Lessmann, M. Oberguggenberger and R. Stark. Fuzzy models in geotechnical engineering and construction management. *Computer-Aided Civil and Infrastructure Engineering,* 14:93-106, 1999.
[12] Th. Fetz, M. Oberguggenberger and S. Pittschmann. Applications of possibility and evidence theory in civil engineering. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems,* 8:295-309, 2000.
[13] M. A. Guitiérrez, R. De Borst. Numerical analyis of localization using a viscoplastic regularization: influence of stochastic material defects. *International Journal for Numerical Methods in Engineering,* 44:1823-1841, 1999.
[14] G. J. Klir and M. J. Wiermann. *Uncertainty-Based Information. Elements of Generalized Information Theory.* Physica-Verlag, Heidelberg, 1998.
[15] B. S. L. P. de Lima and N. F. F. Ebecken. A comparison of models for uncertainty analysis by the finite element method. *Finite Elements in Analysis and Design,* 34:211-232, 2000.
[16] M. Oberguggenberger, S. Pittschmann. Differential equations with fuzzy parameters. *Mathematical and Computer Modelling of Dynamical Systems,* 5:181-202, 1999.
[17] M. Oberguggenberger and F. Russo. Nonlinear SPDEs: Colombeau solutions and pathwise limits. In: L. Decreusefond, J. Gjerde, B. Øksendal and A.S. Üstünel (Eds.). *Stochastic Analysis and Related Topics VI.* Birkhäuser, Boston, 1998, 319-332.
[18] M. Oberguggenberger and F. Russo. Nonlinear stochastic wave equations. *Integral Transforms and Special Functions,* 6:71-83, 1998.
[19] M. Oberguggenberger and F. Russo. Singular limiting behavior in nonlinear stochastic wave equations. In: A. B. Cruzeiro, J.-C. Zambrini (Eds.), *Stochastic Analysis and Mathematical Physics.* Birkhäuser-Verlag, Basel, 2001, 87-99.
[20] B. Øksendal. *Stochastic Differential Equations. An Introduction with Applications.* Fifth Edition, Springer, Berlin, 1998.
[21] T. L. L. Orr. Selection of characteristic values and partial factors in geotechnical designs to Eurocode 7. *Computers and Geotechnics,* 26:263-279, 2000.
[22] R. Rackwitz. Reviewing probabilistic soils modelling. *Computers and Geotechnics,* 26:199-223, 2000.
[23] F. Russo and M. Oberguggenberger. White noise driven stochastic partial differential equations: triviality and non-triviality. In: M. Grosser, G. Hörmann, M. Kunzinger and M. Oberguggenberger (Eds.), *Nonlinear Theory of Generalized Functions.* Chapman & Hall/CRC, Boca Raton, 1999, 315-333.
[24] L. Schwartz. *Théorie des Distributions.* Nouvelle Édition, Hermann, Paris, 1966.

Queueing models with fuzzy data in construction management

Michael Oberguggenberger

Institut für Technische Mathematik, Geometrie und Bauinformatik, Universität Innsbruck

Summary. Queueing problems arise in civil engineering, e. g., in earth work at large construction sites when an excavator serves a number of transport vehicles. Due to a large number of fuzzy side conditions, it is not plausible that a precise estimate for the input parameters can be given, as required in standard probabilistic queueing models. In this article, two alternatives are described that allow to incorporate data uncertainty: a probabilistic queueing model with fuzzy input and fuzzy probabilities, and a purely fuzzy queueing model formulated in terms of network theory.

1 Introduction

There is increasing awareness in the engineering community [1, 20, 24, 28] that probability theory alone does not suffice for modelling the uncertainties arising in engineering problems. Indeed, the data commonly available, say in soil mechanics or construction management, are often scarce, vague, ambiguous or in any case in need of interpretation. This necessitates the development of more flexible tools for assessing and processing subjective knowledge and expert estimates.

Using risk analysis, it is usually possible for the planning engineer to provide ranges for the fluctuations of the parameters involved at various risk levels. This opens the door for employing fuzzy sets, possibility theory or random set theory. When these types of methods are employed for describing the input data, it is essential that arithmetical processing is possible in the engineering models and results in output data of the same type. In probability theory, functions of random variables are again random variables. In fuzzy set theory, the extension principle (reducing the computations to evaluating the solution operators on the level sets) guarantees computability. In random set theory, the evaluation of arithmetic operations can be performed directly with the focal sets. Thus all these approaches are applicable in computational models.

The purpose of this contribution is to demonstrate two applications of these concepts in civil engineering: Both deal with a queueing problem as typically arising in earth work at larger construction sites, see e. g. [17], demonstrated in Fig. 1. To set the stage, we briefly describe this situation formally. The problem is a closed loop transport system, consisting of a single excavator (server) and N transport

Fig. 1. A waiting queue in earth work.

vehicles (customers). After loading by the server, the vehicles transport and unload the material and return to the excavator. Due to variations in the service and return times, a waiting queue will build up in front of the server, see Fig. 2. The engineering problem is to design the system in a cost efficient way. Having chosen a certain excavator, this will chiefly be decided by the number of transport vehicles employed: Too few vehicles will incur costs due to idle time of the server, while too many will incur costs due to waiting time spent in the queue. The transportation and excavation costs per unit time given, the essential performance index is the average time needed by each transport vehicle to complete a full run, or the average number of completed runs in a given period.

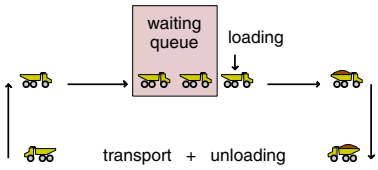


Fig. 2. Scheme of the queueing system.

The input parameters in the system are the service time (loading of a single transport vehicle by the excavator) and the travel or return time of the transport vehicles. In the project planning phase, the designing engineer has to determine these parameters in order to calculate the required capacity of the equipment. The loading time of the excavator depends on a large number of uncontrollable conditions: soil parameters, like grain structure, angle of internal friction, loosening; accessibility of construction site; effective slewing angle of the excavator; meteorological conditions, and so on. Similarly, the return time of the transport vehicles depends

on road and traffic conditions, conditions at the unloading site, and so on. In the planning phase, these data are only vaguely known. Statistical data - if collected from previous projects at all - are not assertive under the always different and only partially predictable circumstances arising in new building projects.

On the other hand, the standard probabilistic queueing models, like the Markov (M/M/1):N - model [16], assume certain probability distributions on the service and return times and require precise estimates for the average service rate and average return rate as input. Such a model is incapable of dealing with the data uncertainty as described above, unless augmented by a second level model for the data uncertainty. It is the aim of the first part of this paper to use fuzzy set theory to describe the parameter uncertainty in a standard probabilistic queueing model. This will result in fuzzy state probabilities from which the required performance measures can be estimated. As the method of computation, we use fuzzy differential equations here.

In the second part of the paper, we present a queueing model exclusively formulated in terms of fuzzy set theory. Here the input data consist of the service and return times proper, described by fuzzy numbers. The performance parameters are computed as fuzzy numbers as well, using a network model for the closed loop loading and transportation process. For example, the number of completed runs in a given period will be a discrete fuzzy number.

Two points should be emphasized: First, the phenomenon of queueing is due to uncontrollable fluctuations in the service and arrival times and thus cannot be captured in a deterministic model. Therefore, it is mandatory to employ methods that allow to incorporate the uncertainties. Second, neither a probabilistic, a fuzzy nor the imprecise probability model we propose here can serve as a quantitative predictive model in the construction management problem addressed here. The uncertainties about the side conditions as well as the mechanism producing the queues are just too big (to be sure, queues do arise in practice!) and it is not plausible that a stationary state is ever reached. In the situation described here, there is no algorithm on which exact specification of the capacity of the equipment could be based; all the more so, a search for a *cost optimal* layout makes no sense. What is gained, however, is an *understanding* of how queues may arise, and what system parameters influence them in which way. These investigations are in alignment with the spirit of *Operations Research*, understood as the science that develops and provides *aids in decision making* - decision making based on understanding rather than black boxes, to be sure.

The models we present in this article should be seen in this light: they provide a framework which allows the designing engineer to describe and to study the relevant input uncertainties and their effects in a qualitative way (this includes qualitative sensitivity analysis). They should not be misinterpreted as making conclusive quantitative assertions about the system behavior (which is simply an impossible task).

2 The fuzzy parameter probabilistic queueing model

2.1 Setting up the probabilistic model

As outlined in the introduction, we consider a closed loop queueing system, consisting of a single server (excavator) and N customers (transport vehicles). We adopt

a first-come-first-served queueing discipline. The required input parameters are the average service time $1/\mu$ (service rate μ) and the average return time $1/\nu$ (driving rate ν) of each vehicle. We note that the arrival rate λ is not an input variable in a closed system; it rather establishes itself during the process.

What concerns the theoretical analysis of the behavior of queues, we adopt the standard Markov (M/M/1):N - queueing model (see e. g. [16]). The model assumes that the service time and the return time are exponentially distributed with expectation values $1/\mu, 1/\nu$. The decisive state variables are the probabilities $p_k(t), k = 0, \ldots, N$ that k customers are present in the queueing system at time t. These probabilities give preliminary information on the initial behavior of the system, but will chiefly be used to compute the stationary state with limiting probabilities $q_k = \lim_{t\to\infty} p_k(t)$. The stationary state will serve as a good approximation to the behavior of the system (simulations with actual data from construction management indicate that it is usually reached within one to two hours).

Following standard arguments (e.g. [16]) one can deduce the system

$$\begin{aligned} p_0'(t) &= -N\nu p_0(t) + \mu p_1(t)\,, \\ p_k'(t) &= (N-k+1)\nu p_{k-1}(t) - (\mu + (N-k)\nu)p_k(t) + \mu p_{k+1}(t) \\ &\qquad \text{for} \quad k = 1, \cdots, N-1\,, \\ p_N'(t) &= \nu p_{N-1}(t) - \mu p_N(t) \end{aligned} \tag{1}$$

with constraint

$$\sum_{k=0}^{N} p_k(t) = 1\,. \tag{2}$$

We normally will assume "deterministic" initial data $p_j(0) = 1$ for some j, $p_i(0) = 0$ for $i \neq j$. We note that $(\sum p_k(t))' = 0$ so that the constraint (2) is automatically satisfied for all times iff satisfied initially. There is a unique equilibrium state given by

$$q_0 = \left(\sum_{n=0}^{N} \frac{N!}{(N-n)!} \left(\frac{\nu}{\mu}\right)^n \right)^{-1} \tag{3}$$

$$q_k = q_0 \frac{N!}{(N-n)!} \left(\frac{\nu}{\mu}\right)^n \quad \text{for} \quad k = 1, \ldots, N\,. \tag{4}$$

The convergence of the probabilities $p_k(t)$ to the equilibrium probabilities q_k is a well known result in the theory of Markov processes. What we need later will be the stronger result of uniform convergence with respect to the parameters μ and ν (when they vary in bounded intervals). This stronger result is proved e. g. in [11].

Once the equilibrium probabilities are found, other performance indicators can be computed by well known arguments from queueing theory. For instance, the average number of vehicles in the queue (in the equilibrium state) is $L = N - (\mu/\nu)(1 - q_0)$. By equating the arrival with the departure rate one gets for the arrival rate $\lambda = \nu(N - L) = \mu(1 - q_0)$. Finally, the average total run time (travel plus waiting time in the queue plus service time) of each vehicle equals $T = N/\lambda$.

2.2 Tools from fuzzy set theory

We briefly recall some notions from fuzzy set theory which we need. More details can be found e. g. in [22]. Let X be a set (of numbers, vectors, matrices, functions

or the like). A *fuzzy set* A in X is defined by means of a function $\pi_A : X \to [0,1]$ which assigns to each a a value $\pi_A(a)$ between 0 and 1. The primary interpretation is that $\pi_A(a)$ is the membership degree to which a belongs to the fuzzy set A. In classical set theory, the membership degree is either 0 or 1; fuzzy set theory allows gradual membership as well.

The level sets of a fuzzy set A are the (classical) subsets of X given by

$$A^\alpha = \{a \in X : \pi_A(a) \geq \alpha\},$$

$0 \leq \alpha \leq 1$. The level sets are nested, that is, $A^\beta \subset A^\alpha$ if $\alpha \leq \beta$. Knowledge of the level sets allows to reconstruct the membership function by means of the rule $\pi_A(a) = \sup\{\alpha : a \in A^\alpha\}$. This leads to the second interpretation of a fuzzy set as a family of parametrized level sets $\{A_\alpha : \alpha \in [0,1]\}$. The notion that reconciles these two interpretation is the concept of *possibility*: if we view A as representing the outcome of a certain variable, then $\pi_A(a)$ is interpreted as the degree of possibility that the variable takes the value a. Equivalently, the level set A^α is the range of fluctuation of the variable corresponding to degree of possibility at least α.

The membership function of a fuzzy set defines a *possibility measure* on X, which assigns a degree of possibility not only to the individual elements of X, but also to all subsets S of X by means of the prescription

$$\pi_A(S) = \sup\{\pi_A(a) : a \in S\}. \tag{5}$$

Occasionally, it will be convenient to use the language of *events*, especially when A represents the uncertain outcome of a variable. We then identify a subset $S \subset X$ with the event that *the outcome A lies in S* and speak of the degree of possibility of this event, denoted by

$$\pi(A \in S) = \pi_A(S).$$

By a *fuzzy number* R we mean a fuzzy subset of the real numbers $\mathbb{R}$ such that its core R^1 consists of a single point and all its level sets R^α as well as its support are compact intervals (the *core* of R is the subset $R^1 = \{r \in \mathbb{R} : \pi_R(r) = 1\}$; the *support* of R is the closure of the subset $\{r \in \mathbb{R} : \pi_R(r) > 0\}$). We will use the distance

$$d(Q,R) = \sup_{\alpha \in (0,1]} d_H(Q^\alpha, R^\alpha) \tag{6}$$

between fuzzy numbers Q, R where d_H denotes the Hausdorff distance (see e. g. [7] for details).

Finally, we need to explain how the uncertainty of a fuzzy input A is propagated through a function F. This will be done with the aid of the *Zadeh extension principle* [30]. If the input is a fuzzy set with membership function $\pi_A(a)$, say, the output will also be a fuzzy set. Its membership function $\pi_{F(A)}(b)$ is given by

$$\pi_{F(A)}(b) = \sup\left\{\pi_A(a) : F(a) = b\right\}. \tag{7}$$

In case the input consists of a vector of parameters $A = (A_1, \ldots, A_m)$, the extension principle takes the form

$$\pi_{F(A)}(b) = \sup\left\{\min\left(\pi_{A_1}(a_1), \ldots, \pi_{A_m}(a_m)\right) : F(a_1, \ldots, a_m) = b\right\}.$$

In case F is continuous and the level sets of $A_1, \ldots, A_m$ are compact $(0 < \alpha \leq 1)$, this corresponds exactly to computing the range of the function F on each level set,

$$F(A)^\alpha = F(A^\alpha)\,, \quad \text{respectively,} \quad F(A_1,\ldots,A_m)^\alpha = F(A_1^\alpha,\ldots,A_m^\alpha)\,.$$

When the A_j^α are intervals, in addition, the set $F(A_1,\ldots,A_m)^\alpha$ is an interval as well and the computation of its boundaries is a task of global optimization: finding the minimum and maximum value of F on the set A^α.

When computing the time dependent state probabilities of the queueing system with fuzzy input parameters, we will have to solve the system (1) of differential equations with fuzzy parameters. Thus we briefly recall the basics from the theory of fuzzy differential equations. For our purpose, it suffices to consider a linear system of the form

$$x'(t) = M(\gamma)x(t) \tag{8}$$

where $M(\gamma)$ is an $(n \times n)$-matrix depending smoothly on a parameter $\gamma \in \mathbb{R}^m$ and $x(t) \in \mathbb{R}^n$. We fix (crisp) initial data and denote the value of the solution at time t by $x(t) = S_t(\gamma)$. When some of the components of the parameter γ are fuzzy, we can view it as a fuzzy subset Γ of $\mathbb{R}^m$ (temporarily denoting the fuzzy parameters by upper case letters and their realizations by lower case letters). Our approach will be to apply the Zadeh extension principle (7) to the (continuous) map $\gamma \to S_t(\gamma) : \mathbb{R}^m \to \mathbb{R}^n$, and we will consider $S_t(\Gamma)$ as the *fuzzy solution* at time t. In this way, the α-level sets of $S_t(\Gamma)$ are precisely the images of the α-level sets of Γ under the solution operator S_t. This approach to fuzzy differential equations is along the lines of [3] and [25]. It was developed in detail in [23], where it was shown that this solution is unique (as a solution to an equivalent prolonged system). We also employ the numerical methods that were developed in [23].

We remark that other - nonequivalent - approaches to fuzzy differential equations have been undertaken: imbedding fuzzy sets into metric spaces [7, 12, 13, 27], differentiation of bounding curves of α-level sets [12, 14, 29], parametrized fuzzy numbers [26]; see also [5, 8, 12] for a study of the interrelations.

2.3 The probabilistic model with fuzzy input and fuzzy output

In this section, we adhere to the probabilistic (M/M/1):N - queueing model constructed in Subsection 2.1, but assume that the distribution parameters μ and ν are fuzzy. In the terminology and spirit of [19], this is an (FM/FM/1):N - queueing system. In contrast to the Markov chain methods there, our approach is based on the differential equations for the probabilities $p_k(t)$ that k customers are present in the system at time t. We apply the extension principle to the solution operator $(\mu,\nu) \to S_t(\mu,\nu)$ of system (1) with constraint (2). When μ and ν are given by fuzzy numbers, the extension principle yields the fuzzy solution

$$p(t) = S_t(\mu,\nu), \tag{9}$$

consisting of the fuzzy vector of time-dependent state probabilities. The k-th component of the solution operator produces the fuzzy probabilities that k vehicles are in the system at time t:

$$p_k(t) = S_t^k(\mu,\nu). \tag{10}$$

We note that the fuzzy equilibrium probabilities q_k could be computed from formula (3) and (4) with the aid of the extension principle. However, having our machinery at hand, it is simpler to pass to the limit in the expressions (10) for the time-dependent

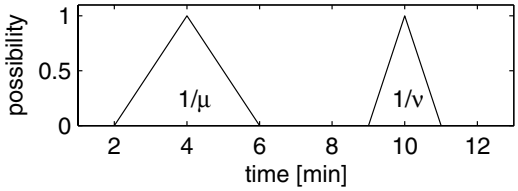


Fig. 3. Fuzzy input parameters.

fuzzy marginals as $t \to \infty$. The convergence holds in the sense of the distance (6), as has been proven in [11].

The performance indicators mentioned at the end of Subsection 2.1 can now easily be computed, again by the extension principle, from the fuzzy output probability q_0 and the fuzzy service rate. The total time needed by each vehicle on average to complete a full run is given as the fuzzy number

$$T = N/\mu(1 - q_0); \tag{11}$$

due to the fact that μ and q_0 are interactive, this is actually an upper estimate. Various further information can be extracted from there. First, the total cost per transported unit mass is proportional to T and $C_V + C_S/N$, where C_V and C_S denote the cost per unit time of a single vehicle and the server, respectively. Thus estimates (with their degrees of possibility) of how the total cost emanates depending on the number N of vehicles can be obtained. The average number of completed runs by all vehicles together in a given period T_0 is $T_0 N/T = T_0\mu(1 - q_0)$. Thus the degree of possibility of achieving a certain performance in a given period of time can be computed as well, giving the basis for assessing the risk of not achieving a required threshold.

2.4 Example

To illustrate the analysis above we now present a computational example with three transport vehicles ($N = 3$). The fuzzy mean serving time is assumed to be given by a triangular fuzzy number with supporting interval $[2, 6]$ and center at $z = 4$, while the average return time is modelled as a triangular fuzzy number with support $[9, 11]$ and center at $z = 10$ (see Fig. 3). The initial state is taken deterministic as described above with $p_3(0) = 1$ (all three transport vehicles present at start). Fig. 4 shows the α-level sets of the fuzzy time-dependent probability $p_0(t)$ for $0 \le t \le 30$; Fig. 5 displays the contour of its membership function as a three-dimensional graph with the degree of possibility as third axis. One can read off that at time $t = 30$ the equilibrium state is almost reached; the probability q_0 can be approximated by a triangular fuzzy number with supporting interval $[0.13, 0.56]$ and center at 0.28. The average total time to complete a full run is given by the fuzzy number depicted in Fig. 6; the approximate bounds for its support are $[6.9, 40.9]$, while its central value lies approximately at 16.7. Similarly, the average number of completed runs in unit time, $\mu(1 - q_0)$, is roughly computed as the triangular fuzzy number with supporting interval $[0.88, 5.22]$ and central value at 2.88. Thus the vagueness of the input data leads to possibly large fluctuations in the performance of the system, in

the range of ±70% of the central value. Such a range of variation is quite realistic, as is seen e. g. from the collected project data in [17].

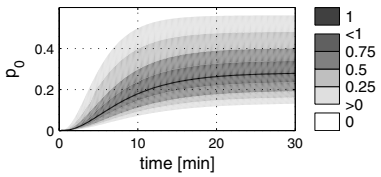


Fig. 4. α-level sets of $p_0(t)$.

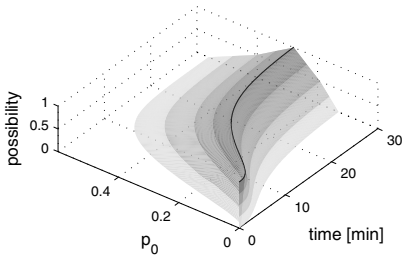


Fig. 5. Contour of fuzzy set $p_0(t)$.

3 Fuzzy service and return times

3.1 Setting up the model

In this section we discuss a *direct* approach to fuzzy data queueing systems. We do not presume any probabilistic queueing model, but study a closed loop system with fuzzy service times T_s and fuzzy return times T_r and N customers. In the language of [4, 15, 19, 21], this is an (F/F/1):N - queueing system. We are again interested in the number of completed runs at a given time T_0 as the primary performance measure.

Our basic tool will be the observation that the behavior of a closed loop queueing system can be represented by a network. The nodes of the network correspond to the activities

- k-th service,
- i-th travel of vehicle number j.

Here we make the additional simplifying assumption that the vehicles return and are served in the same order through the whole observation period. This has the pleasant consequence that standard network techniques, like the *critical path method*, suffice to compute the earliest starting and finishing times of the k-th service and l-th travel. In general, one would have to use *Petri-nets* instead [2, 6].

It is not difficult to derive a formula for the earliest starting times. We adhere to the following notation:

- $S(k)$ = earliest starting time of k-th service,
- $D[j](i)$ earliest starting time of i-th travel of vehicle number j,
- $R(l)$ = total duration of l travels.

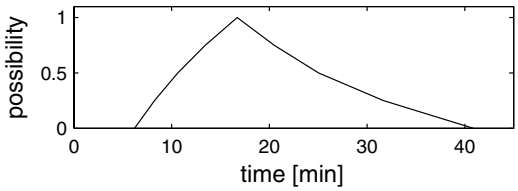


Fig. 6. Fuzzy average total run time T.

Assuming that service and travel times are the same for each vehicle and putting $S(1) = 0$ we get:

$$\begin{aligned} D[1](1) = T_s, D[2](1) = 2T_s, \ldots, D[N](1) = NT_s, \\ S(N+1) = \sup\{S(N) + T_s, D[1](1) + T_r\}, \\ D[1](2) = S(N+1) + T_s, \\ S(N+2) = \sup\{S(N+1) + T_s, D[2](1) + T_r\} \end{aligned}$$

and so on. The general formula is

$$S(Ni + j) = \sup\{S(Ni + j - 1) + T_s, D[j](i) + T_r\} \tag{12}$$

for $1 \leq j \leq N$ and $i = 1, 2, 3, \ldots$ The total duration of l complete travels is

$$R(l) = D[j](i + 1) + T_r$$

when $l = Ni + j$. The formulas can be easily verified by checking the structure of the example given in Fig. 7 where $N = 3$ and $l = 12$.

We now turn to fuzzy input data. For the sake of presentation, we shall work with triangular fuzzy numbers. These are defined by a supporting interval $[a, b]$ and a central value $z, a \leq z \leq b$; we shall use the notation $\langle a, z, b\rangle$. We put a partial order on the set of triangular fuzzy numbers by defining

$$\langle a_1, z_1, b_1\rangle \leq \langle a_2, z_2, b_2\rangle$$

iff

$$a_1 \leq a_2,\ z_1 \leq z_2,\ b_1 \leq b_2.$$

The supremum in formula (12) is understood in the sense of this partial order, that is, it is given as the smallest triangular number larger than the two given triangular numbers. The addition of two triangular numbers follows Zadeh's extension principle and is just done by summing the respective values a, z, b, see e.g. [9]. With these operations as defined above, the earliest starting and finishing times of the activities in a network can be obtained by the standard *forward computation* from network theory. For more details on these methods and applications of fuzzy network techniques in civil engineering we refer to [10, 18].

The same calculations could be applied with polygonal fuzzy numbers. For general fuzzy numbers, the notion of supremum that follows from Zadeh's extension principle would have to be used. This introduces some computational complications, because the comparison must be done on each α-level set, and not only on the ones defined by the nodes of a polygonal number. From a computational point

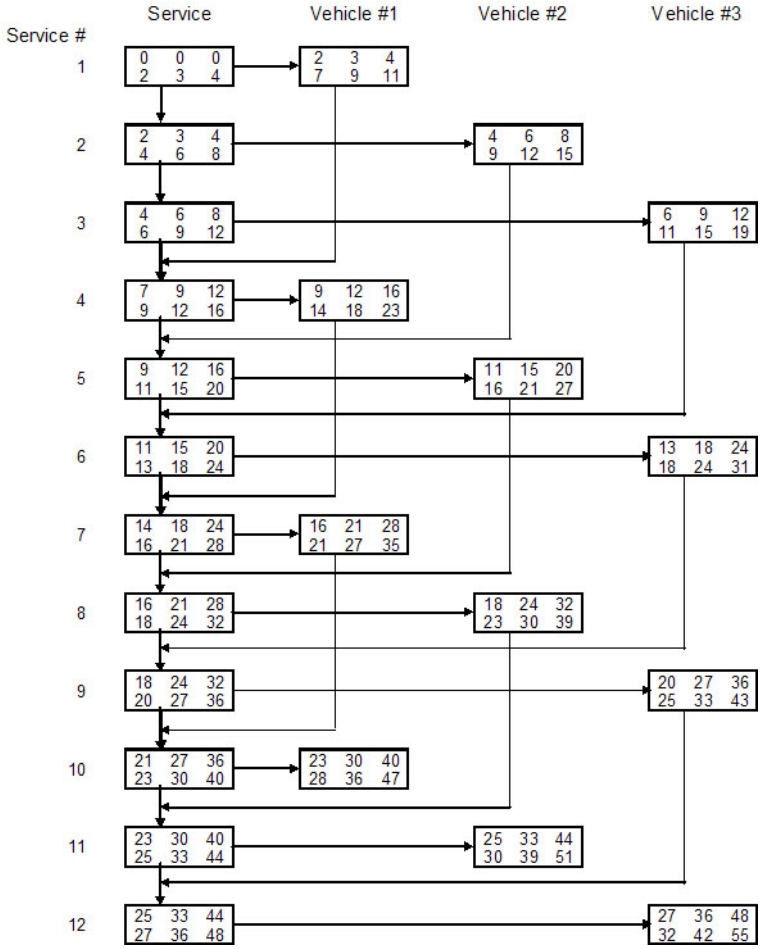


Fig. 7. Network computation of queueing system.

of view, it is therefore recommended to use polygonal approximations which can be as close as desired to the original fuzzy numbers.

The fuzzy network computation results in the fuzzy total durations $R(l)$ of l completed runs. The engineer who wants to analyze the capacity of the system is interested in the number C of completed runs (of all vehicles together) in a given period T_0. We have that

$$C = l \text{ if and only if } R(l) \leq T_0 \text{ and } R(l+1) > T_0.$$

If $\pi_{R(l)}(r)$ denotes the membership function of the fuzzy set $R(l)$, the degree of possibility of the event $R(l) \leq T_0$ is

$$\pi\big(R(l) \leq T_0\big) = \pi_{R(l)}\big((-\infty, T_0]\big) = \sup\big\{\pi_{R(l)}(r) : r \in (-\infty, T_0]\big\},$$

see (5), and similarly for the event $R(l+1) > T_0$. Defining the degree of possibility of the intersection of two fuzzy sets by the minimum of the two degrees of possibility, we arrive at

$$\pi\big(C = l\big) = \pi_C\big(\{l\}\big) = \min\big\{\pi\big(R(l) \le T_0\big), \pi\big(R(l+1) > T_0\big)\big\}. \tag{13}$$

This way we obtain C as the discrete fuzzy number of completed runs up to time T_0. This information can then be used by the designing engineer to decide whether the capacity of the system is sufficient.

3.2 Example

To illustrate the approach, we take an example with three vehicles (N=3) and service and travel times given by the triangular fuzzy numbers

$$T_s = \langle 2, 3, 4\rangle,\ T_r = \langle 5, 6, 7\rangle.$$

The computation is done in Fig. 7 for $l = 1, \ldots, 12$ runs. The nodes contain the triangular numbers for the earliest starting and finishing times. The left column refers to the services, the right three columns to the return travels of vehicles 1,2 and 3. The total durations $R(l)$ can be read off from Fig. 7 for $l = 1, \ldots, 12$ and are given by the following triangular numbers:

$$\begin{aligned}
R(1) &= \langle\ 7,\ \ 9, 11\rangle, & R(2) &= \langle\ 9, 12, 15\rangle, & R(3) &= \langle 11, 15, 19\rangle,\\
R(4) &= \langle 14, 18, 23\rangle, & R(5) &= \langle 16, 21, 27\rangle, & R(6) &= \langle 18, 24, 31\rangle,\\
R(7) &= \langle 21, 27, 35\rangle, & R(8) &= \langle 23, 30, 39\rangle, & R(9) &= \langle 25, 33, 43\rangle,\\
R(10) &= \langle 28, 36, 47\rangle, & R(11) &= \langle 30, 39, 51\rangle, & R(12) &= \langle 32, 42, 55\rangle.
\end{aligned}$$

We choose $T_0 = 30$ and ask for the degree of possibility that the number of completed runs up to time 30 is l. From formula (13) and the computed data above one reads off (see also Fig. 8) that

$$\begin{aligned}
\pi\big(C = \ \ 4\big) &= \min\{1, 0\} = 0, & \pi\big(C = \ \ 5\big) &= \min\{1, 1/7\} = 1/7,\\
\pi\big(C = \ \ 6\big) &= \min\{1, 5/8\} = 5/8, & \pi\big(C = \ \ 7\big) &= \min\{1, 1\} = 1,\\
\pi\big(C = \ \ 8\big) &= \min\{1, 1\} = 1, & \pi\big(C = \ \ 9\big) &= \min\{5/8, 1\} = 5/8,\\
\pi\big(C = 10\big) &= \min\{2/8, 0\} = 1/4, & \pi\big(C = 11\big) &= \min\{0, 1\} = 0.
\end{aligned}$$

The resulting discrete fuzzy number is depicted in Fig. 9 (left picture).

It is interesting to compare the outcome of this calculation with the one that the model from Section 2 would have produced. In this model, the average number of completed runs in a given period T_0 is $T_0\mu(1 - q_0)$, as stated at the end of Subsection 2.3, with q_0 given by Eq. (3). Taking T_s, T_r as in Example 3.2, we can compute $T_0\mu(1-q_0)$ by means of the extension principle, observing the relation $\mu = 1/T_s, \nu = 1/T_r$. The resulting approximation to the fuzzy number C of completed runs in the period $T_0 = 30$ is depicted in Fig. 9 (right picture). Due to the nature of the model of Section 2, it is a real fuzzy number rather than a fuzzy integer and thus has to be interpreted as a continuous approximation to a discrete fuzzy number.

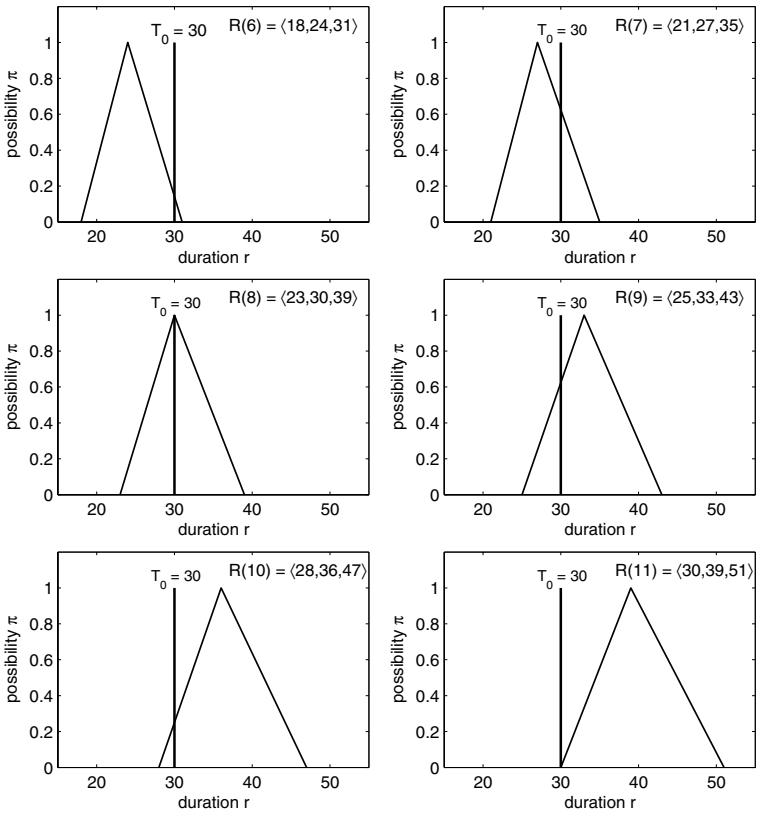


Fig. 8. Duration $R(l)$ of l runs.

References

[1] H. Adeli and B. Möller (Eds.). *Special Issue: Fuzzy Modeling.* Vol. **14**(2) of *Computer Aided Civil and Infrastructure Engineering*, 1999.

[2] B. Baumgarten. Petri-Netze. Grundlagen und Anwendungen. Spektrum Akad. Verlag, Heidelberg 1996.

[3] A. Bonarini and G. Bontempi. A qualitative simulation approach for fuzzy dynamical models. *ACM Trans. Modeling Comp. Sim.*, **4** (1994), 285–313.

[4] J. J. Buckley. Elementary queueing theory based on possibility theory. *Fuzzy Sets and Systems*, **37** (1990), 43–52.

[5] J. J. Buckley. Solving fuzzy equations. *Fuzzy Sets and Systems*, **50** (1992), 1–14.

[6] J. Cardoso and H. Camargo (Eds.). Fuzziness in Petri Nets. Physica-Verlag, Heidelberg 1999.

[7] P. Diamond and P. Kloeden. *Metric Spaces of Fuzzy Sets.* World Scientific, Singapore, 1994.

[8] D. Dubois and H. Prade. Towards fuzzy differential calculus, part 3: differentiation. *Fuzzy Sets and Systems*, **8** (1982), 223–233.

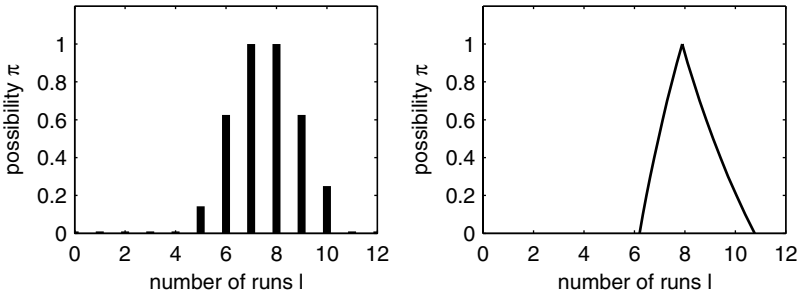


Fig. 9. Fuzzy number C of runs up to time $T_0 = 30$ and continuous approximation.

[9] D. Dubois and H. Prade. *Possibility Theory.* Plenum Press, New York, 1988.

[10] Th. Fetz, J. Jäger, D. Köll, G. Krenn, H. Lessmann, M. Oberguggenberger, and R. Stark. Fuzzy models in geotechnical engineering and construction management. *In this volume.*

[11] Th. Fetz, M. Oberguggenberger, S. Pittschmann. Applications of possibility and evidence theory in civil engineereing. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, **8** (2000), 295–309.

[12] O. Kaleva. Fuzzy differential equations. *Fuzzy Sets and Systems*, **24** (1987), 301–307.

[13] O. Kaleva. The Cauchy problem for fuzzy differential equations. *Fuzzy Sets and Systems*, **35** (1990), 389–396.

[14] A. Kandel. Fuzzy Mathematical Techniques with Applications. Addison-Wesley, Reading, 1986.

[15] C. Kao, C.-C. Li, S.-P. Chen. Parametric programming to the analysis of fuzzy queues. *Fuzzy Sets and Systems*, **107** (1999), 93–100.

[16] L. Kleinrock. *Queueing Systems, vol. I: Theory.* John Wiley, New York, 1975.

[17] G. Kühn. *Der maschinelle Erdbau.* Teubner-Verlag, Stuttgart, 1984.

[18] H. Lessmann, J. Mühlögger, and M. Oberguggenberger. Netzplantechnik mit unscharfen Methoden. *Bauingenieur* **69** (1994), 469–478.

[19] R.-J. Li and E. S. Lee. Analysis of fuzzy queues. *Computers Math. Appl.*, **17** (1989), 1143–1147.

[20] B. Möller. Fuzzy-Modellierung in der Baustatik. *Bauingenieur*, **72** (1997), 75–84.

[21] D. S. Negi, E. S. Lee. Analysis and simulation of fuzzy queues. *Fuzzy Sets and Systems*, **46** (1992), 321–330.

[22] M. Oberguggenberger. The mathematics of uncertainty: models, methods and interpretations. *In this volume.*

[23] M. Oberguggenberger and S. Pittschmann. Differential equations with fuzzy parameters. *Math. Mod. Systems*, **5** (1999), 181–202.

[24] M. Oberguggenberger, G. I. Schuëller, K. Marti (Eds.). *Special Issue: Fuzzy Sets, Imprecise Probability and Stochastics in Engineering. Zeitschrift f. Angewandte Mathematik u. Mechanik*, to appear 2004.

[25] K. N. Otto, A. D. Lewis, and E. K. Antonsson. Membership induced on manifolds by vector fields and flows. *Fuzzy Sets and Systems*, **70** (1995), 15–29.

[26] R. Palm and D. Driankov. Fuzzy inputs. *Fuzzy Sets and Systems*, **70** (1995), 315–335.

[27] M. L. Puri and D. Ralescu. Differentials of fuzzy functions. *J. Math. Anal. Appl.*, **91** (1983), 552–558.

[28] T. J. Ross, J. M. Booker, W. J. Parkinson (Eds.). *Fuzzy Logic and Probability Applications - Bridging the Gap*. SIAM, Philadelphia PA, ASA, Alexandria VA, 2002.

[29] S. Seikkala. On the fuzzy initial value problem. *Fuzzy Sets and Systems*, **24** (1987), 319–330.

[30] L. A. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning, Part 1 - 3. Information Sci **8** (1975), 199–250; 301–357; **9** (1975), 43–80.

Fuzzy models in geotechnical engineering and construction management

Thomas Fetz[1], Johannes Jäger[2], David Köll[3], Günther Krenn[3], Heimo Lessmann[3], Michael Oberguggenberger[1], and Rudolf F. Stark[4]

[1] Institut für Mathematik und Geometrie, Universität Innsbruck
[2] Beton- und Monierbau Innsbruck GesmbH
[3] Institut für Baubetrieb und Bauwirtschaft, Universität Innsbruck
[4] Institut für Festigkeitslehre, Universität Innsbruck

Summary. This article is devoted to a variety of applications of *fuzzy models* in civil engineering, presenting current work of a group of researchers at the University of Innsbruck. With *fuzzy methods* and *possibility theory* as an encompassing framework, the following areas are addressed: uncertainties in geotechnical engineering, fuzzy finite element computation of a foundation raft, fuzzy dynamical systems, processing uncertainty in project scheduling and cost planning.

1 Introduction

Traditional engineering models are deterministic: sharp inputs are processed in a structurally determined system to produce a sharp output, which is taken as an approximate prediction of reality. To account for fluctuations, probability theory has been introduced. This shifts the emphasis from sharp input data to sharp probabilities of sharply defined events. Among engineers, there is increasing discomfort with the observed fact that predictions obtained in this way may deviate to such an extent from reality as to render them useless.

For example, a case study [17] concerning the design of a sheetpile wall, in which a number of reputable European engineering companies had agreed to participate, resulted in a corresponding number of seemingly precise predictions. These, however, differed drastically from each other and from the observed behavior of the completed structure.

Conspicuous differences between prediction and reality may arise in many areas in the course of the realization of an engineering structure: from site investigations and analysis to scheduling and cost planning. We should be clear about the facts that

Published originally in: *Computer-Aided Civil and Infrastructure Engineering* **14**(2) (1999), 93–106 (H. Adeli, B. Möller (eds.), Special Issue: Fuzzy Modeling).

- models are approximations to and conjectures about reality;
- the input parameters are known only imprecisely;
- probabilistic methods may fail to capture the information available about the deviations.

The inherent vagueness of modelling procedures can be traced to various reasons, for example: lack of knowledge of boundary conditions; simplification in complex circumstances forcing a single parameter to cover a wider range of situations; lack of a precisely quantifiable definition of some verbally defined variable; uncertainty about future dispositions. However, there is clearly no alternative to employing rational models in the three central activities of engineering: design, construction, and control. Rather, the engineer should face the limitations of the modelling process, put the range of imprecision into the open and make it accessible to responsible assessment by all participants in the construction process. This will involve processing not only data but also the available objective and subjective information on their uncertainty.

We believe that fuzzy set theory provides a framework for accomplishing this task. The power of fuzzy set theory is that it allows a formalization of vague data, a representation of their fuzziness which can be entered into computations and a possibility theoretic interpretation. Assigning degrees of possibility as a replacement of probability appears to be more adapted to formalizing expert knowledge, due to the relaxation of axioms. Working with fuzzy methods forces and allows the engineer to address the uncertainties, see and judge the possible range of outputs the fuzzy model predicts, and gain understanding of the possible behavior of the system, given the imperfect description formulated initially.

The purpose of this article is to demonstrate that fuzzy formulation and computation is possible in a number of engineering tasks ranging from geotechniques to dynamics to project planning. It presents the ongoing work of a group of researchers in construction management, strength of materials, mathematics and numerics at the University of Innsbruck; we refer to the papers [4, 5, 9, 11].

The plan of exposition is as follows. After a short account of the basic notions of fuzzy set theory, we address questions of modelling and uncertainties in geomechanics. This is followed by an investigation of the effect of fuzzy parameters in a raft foundation, the corresponding fuzzy finite element computation, and a discussion of the interpretation. Next, we exhibit some fuzzy ideas on dynamical systems. Finally, we turn to project planning: scheduling and cost estimation. We show that suitable methods of presentation allow a clear exhibition of fuzziness in a network structure, providing means of control under risk. The fuzzy approach gives a lucid picture of the complexities involved when duration dependent costs enter into network analysis. It explains why cost optimization is impossible, but must be replaced by the search for a satisficing solution.

A methodological remark seems in order. We use fuzzy set theory to characterize vague data by intervals of variation, supplied with a valuation. This valuation provides the additional degrees of freedom needed to model uncertainty, to describe and interpret its effects. Numerically, the results are computed by repeated application of deterministic algorithms to all relevant data combinations. The increased computational effort can currently be handled in problems of moderate size.

2 Fuzzy sets

This section serves to briefly collect what we need from fuzzy set theory: definition of and computation with fuzzy quantities.

Given a basic set X of discourse, a fuzzy subset A is characterized by (and can be identified with) its membership function $m_A(x)$, $0 \leq m_A(x) \leq 1$, defined for each element x of X. The value $m_A(x)$ can be interpreted as

- the membership degree of the element x belonging to A;
- the degree of possibility that the variable A takes the value x.

Introducing the α-level sets ($0 \leq \alpha \leq 1$)

$$[A]_\alpha = \{x \in X : m_A(x) \geq \alpha\} \tag{1}$$

we arrive at the interpretation mentioned in the Introduction:

The variable A fluctuates in the range $[A]_\alpha$ with possibility degree α.

A fuzzy real number a is defined as a fuzzy subset of the basic set X of real numbers with the property that each level set $[a]_\alpha$ is a compact interval, $0 < \alpha \leq 1$. See Fig. 1 for an example of a triangular fuzzy real number a; depicted is the level set $[a]_{1/2}$.

In general, the graph of the membership function will be curved. However, as an approximation it often suffices to work with polygonal fuzzy numbers.

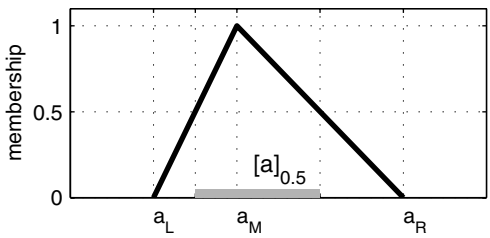


Fig. 1. Triangular fuzzy number.

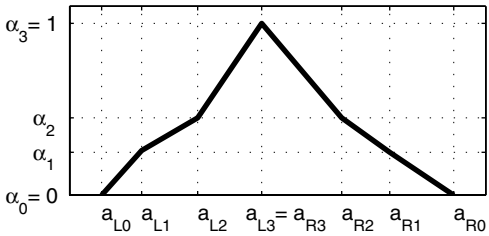


Fig. 2. Polygonal fuzzy number.

Their graphs are piecewise linear with corner points at $a_{L0} \leq a_{L1} \leq \ldots \leq a_{Ln} \leq a_{Rn} \leq \ldots a_{R1} \leq a_{R0}$ and corresponding levels $0 = \alpha_0 < \alpha_1 < \ldots < \alpha_n = 1$, see Fig. 2.

Polygonal fuzzy numbers are denoted by

$$a = \langle a_{L0}, a_{L1}, \ldots, a_{Ln}, a_{Rn}, \ldots, a_{R0} \rangle.$$

Frequently used special cases are triangular fuzzy numbers $\langle a_L, a_M, a_R \rangle$ as well as trapezoidal shapes. Compact intervals $[a_L, a_R]$ can be viewed as fuzzy numbers with rectangular membership function. Among the more general fuzzy quantitites are fuzzy vectors, in case the basic set X is the n-dimensional Euclidean space, or even fuzzy functions, in case X is a space of functions.

In scientific computations, parameters have to be inserted in functions. Thus the necessity arises of evaluating functions on fuzzy numbers (or vectors). This is achieved by means of the Zadeh extension principle [18], exhibited here for the case of a function $z = f(x, y)$ of two variables. If a, b are fuzzy numbers, so will be the result $f(a, b)$. Its fuzzy value is determined by prescribing its membership function:

$$m_{f(a,b)}(z) = \sup_{z=f(x,y)} \min \{m_a(x), m_b(y)\}. \tag{2}$$

When using the possibilistic interpretation, the Zadeh extension principle is especially intuitive: In order to determine the membership degree of the dependent variable z, one considers all possible combinations (x, y) leading to $z = f(x, y)$. Each single combination arises with degree of possibility $\min\{m_a(x), m_b(y)\}$. For the final result, the maximal degree of possibility (the supremum) is decisive.

An essential computational tool is interval analysis on each α-level set: Indeed, in case f is a continuous function, the α-level set of $f(a, b)$ is obtained by evaluating the image of the α-level sets of a, b under the function f:

$$[f(a, b)]_\alpha = f([a]_\alpha, [b]_\alpha). \tag{3}$$

In case f is an arithmetic operation, this leads to simple formulae. For example, addition $f(a,b) = a + b$ of triangular or polygonal fuzzy numbers is achieved by simply adding the abscissae of their corner points:

$$\langle a_L, a_M, a_R \rangle + \langle b_L, b_M, b_R \rangle = \langle a_L + b_L, a_M + b_M, a_R + b_R \rangle,$$

the result being again a triangular or polygonal fuzzy number. In the case of subtraction $f(a,b) = a - b$, a crosswise interchange is required:

$$\langle a_L, a_M, a_R \rangle - \langle b_L, b_M, b_R \rangle = \langle a_L - b_R, a_M - b_M, a_R - b_L \rangle.$$

For general operations f, the resulting membership function can be curved; the bounds of the α-level sets are obtained by solving an optimization problem.

In network planning, the need arises to compare two or more fuzzy numbers. Here we face the difficulty that there is no total order on fuzzy real numbers, that is, it cannot be decided in general whether $a \leq b$ or $b \leq a$ (in contrast to the case of usual real numbers). A closer look at network planning shows that we actually need to compute $\max(a,b)$ at nodes where two paths meet. This can be done by applying the Zadeh extension principle to the function $f(a,b) = \max(a,b)$. In the case of triangular or polygonal fuzzy numbers we approximate the result by taking the maximal value at each corner point of the membership function (in order to avoid the introduction of additional levels):

$$\max(a,b) = \langle \max(a_L, b_L), \max(a_M, b_M), \max(a_R, b_R) \rangle$$

and correspondingly for the minimum needed in the backward computation.

3 An application of fuzzy set theory in geotechnical engineering

3.1 Preliminary remarks

In foundation engineering the design of mat foundations is just an ordinary task and it is certainly far from being classed as a special geotechnical job. At first glance one might be tempted to assume that the result of a raft design with well defined structural input data will fall into a narrow band of solutions. However, whereas the loading, the material parameter of the foundation structure and the requirements the foundation should meet, like allowable differential settlement, are well defined, there is generally a substantial lack of information concerning the soil. In some cases this deficiency of information serves as an excuse for simplified modelling and it is often argued that it is not worth using a possibly more appropriate and expensive model for the analysis of the problem at hand. On the other hand, even the most sophisticated model will not guarantee that the solution based on it will correctly predict all desired aspects of the foundation response. Hence, the decision

about which model will be appropriate with respect to the importance of the project and the available soil data will very much depend on the expertise of the geotechnical engineer. Irrespective of both the vagueness of input data and the uncertainties with respect to the model, in today's practice the engineer will come up with crisp results. Here, fuzzy set theory opens new opportunities to reflect the lack of information and uncertainties in the results of engineering computations in a rational manner. In this section we will demonstrate the application of fuzzy set theory to the analysis of a foundation raft.

3.2 Material modelling

Constitutive equations or inequalities describe the response of materials. To specify a certain class of material with desired accuracy a sufficient number of such relations is required. The purpose of engineering constitutive models is not to give a mirror image of realistic material behavior but rather to describe the main mechanical properties, which are important for the design of constructions.

Constitutive relations consist of equations and inequalities which contain the basic principles of continuum mechanics (conservation of mass, momentum and energy) and of material theory (principle of frame invariance etc.). They must be satisfied exactly by the constitutive law. As these principles are not sufficient for the determination of all the variables involved in the material model, additional assumptions are required. These assumptions are based on the interpretation of material test data and thus they introduce an amount of uncertainty into the material model. For example elasticity, plasticity and damage characterize properties of idealized material behavior. In reality the existence of an elastic domain, the transition from elasticity to plasticity or the initiation of damage cannot be defined exactly. The theory of fuzzy sets offers a framework to model these various uncertainities in a consistent manner [6].

General constitutive relations are tensor valued functions of stress, strain, strain-rate and additional variables, which describe the mechanical state of the material. In order to simplify the constitutive law, material parameters are introduced replacing complicated functions of the variables by one single constant. As a consequence the material properties described by the function which is replaced by the constant will be captured by the numerical value of one single material parameter. Thus material parameters of simple models have to cover a larger bandwidth of values than material parameters of more refined models.

An additional source of uncertainity arises when the numerical values for material parameters are determined by carrying out material tests. Usually the constants introduced into the model are identified by minimizing a measure of error between theoretical (calculated) and experimental (measured) variables. Due to the mathematical structure of most nonlinear constitutive equations a global minimum of error is rarely found. In reality several equivalent solutions exist. Further, considering that experimental data are burdened by scatter

and errors the results gained by the calibrated model contain some additional uncertainty.

3.3 Numerical model

It should be emphasized very clearly that using fuzzy set theory will in no way release the engineer from his job of establishing a correct mechanical model and using adequate numerical techniques to tackle the problem.

It is common sense that soil behavior is far from being adequately described as linear elastic, however, for the analysis of raft foundations it is usually justified to assume linear elastic soil response within the range of working loads. Although yielding eventually might occur in some confined regions of the soil they usually remain small compared with the overall area of the soil-structure interface. In current engineering practice the Winkler model is probably most widely used to analyze rafts on elastic foundations. This is based much more on the fact that this model is very easy to handle and inexpensive both in terms of discretization and computation cost, rather than due to its mechanical merits. More reliable results are obtained by using a continuum model. Some of these models may even account for the increasing stiffness of the soil with depth.

In this example the soil is assumed to be adequately modelled as an isotropic non-homogeneous elastic medium. We therefore will adopt a continuum model with a power variation of Young's modulus E with depth like the one proposed by Booker et al [2], i.e., the elastic constants of the soil medium are given by

$$\begin{aligned} E(z) &= E_1 z^{\rho}, \qquad 0 \leq \rho \leq 1 \\ \nu &= \text{const.} \end{aligned} \tag{4}$$

where E_1 is a constant which determines Young's modulus at the depth $z = 1$, ρ is referred to as the non-homogeneity parameter and ν denotes Poisson's ratio. Eqn. (4) could be used for sand deposits where E is likely to vary nonlinearly with the overburden pressure. The raft, which also is assumed to behave elastically, is modeled by finite elements.

Numerically we are combining the finite element procedure with a boundary element solution. One of the main tasks is to determine the displacements of the non-homogeneous elastic half-space caused by a constant surface traction acting on an arbitrarily shaped area. In the context of the applied algorithm the shape of the loaded area depends on the shape of the finite element. Assuming that the boundary of the loaded domain Ω can be approximated by a polygon consisting of k segments, as shown in Fig. 3, the vertical displacement u_z of a point $(x, y, 0)$ on the soil surface due to a vertical traction p_z is given by

$$u_z(x, y) = \frac{Bp_z}{E_1} \sum_{i=1}^{k} \int_{L_i} \frac{un_\xi + vn_\eta}{(\rho - 1)R^{\rho+1}} ds. \tag{5}$$

In the above equation B is a function of ρ and ν; u, v and R measure the distance between source point and field point and n_ξ and n_η are the direction cosines of the actual segment. The integral in (5) can be evaluated numerically in a standard manner by applying Gauss-Legendre quadrature. Solutions for other displacement components due to loading in any direction are given by [15].

For rectangular shaped elements the displacement u_z can be found without performing any numerical integration at all making the procedure quite efficient. In the case of vertical loading within the domain Ω, u_z is given by

$$u_z\,(x,y) = \frac{Bp_z}{E_1}\lambda \tag{6}$$

where λ is a function of (x,y) and is determined by evaluating κ for the boundaries of the loaded rectangle (see Fig. 4), i.e.

$$\begin{aligned}\lambda = \kappa(x-x_1, y-y_1) - \kappa(x-x_1, y-y_2) + \\ +\kappa(x-x_2, y-y_2) - \kappa(x-x_2, y-y_1).\end{aligned} \tag{7}$$

κ is given by

$$\kappa = \frac{v\,\mathrm{sgn}(u)\,|v|^{-\rho}\,\mathrm{B}_\zeta(a,b) + u\,\mathrm{sgn}(v)\,|u|^{-\rho}\,\mathrm{B}_{(1-\zeta)}(a,b)}{2(1-\rho)} \tag{8}$$

where $\mathrm{B}_\zeta(a,b)$ is the incomplete Beta function, its parameters given by

$$\zeta = \frac{u^2}{u^2+v^2}\,, \quad a = \frac{1}{2}\,, \quad b = \frac{\rho}{2}. \tag{9}$$

For a more detailed derivation of all displacement components due to vertical and horizontal loading the reader is referred to [16].

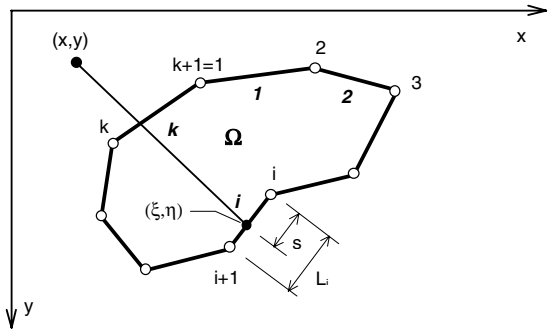


Fig. 3. Polygonal approximation of the loaded area.

3.4 Numerical example

A rectangular raft with a flexural rigidity of 20460 kNm and an aspect ratio of 8 to 4 m on granular soil is considered. The elastic constants E_1 and ν and the exponent ρ determining the degree of non-homogeneity of the soil are assumed to be fuzzy numbers $\tilde{E}_1$, $\tilde{\nu}$ and $\tilde{\rho}$ shown in Fig. 3.4. The raft is subjected to a uniform loading of 100 kPa. The interface is assumed to be smooth, i.e. no shear stresses are transmitted between raft and soil.

3.5 Calculation of the fuzzy solution

The numerical algorithm for crisp data has been presented in Section 3.3. Our task is to perform it with fuzzy data, producing a fuzzy output. Thus we shall have to evaluate functions $f_{(x,y)}(\tilde{E}_1, \tilde{\nu}, \tilde{\rho})$ where

- f represents an output quantity such as displacement, stress or bending moment,
- (x, y) is a point of the raft under consideration,
- $\tilde{E}_1$, $\tilde{\nu}$ and $\tilde{\rho}$ are the fuzzy parameters.

As shown in Section 2 we can reduce this to a calculation of the image

$$[f_{(x,y)}(\tilde{E}_1, \tilde{\nu}, \tilde{\rho})]_\alpha = f_{(x,y)}([\tilde{E}_1]_\alpha, [\tilde{\nu}]_\alpha, [\tilde{\rho}]_\alpha) \tag{10}$$

using the α-cuts. In general the function $f_{(x,y)}$ does not depend monotonically on the parameters E_1, ν and ρ, so it is not sufficient to use only the bounds of the α-cuts for computing the bounds of $[f_{(x,y)}(\tilde{E}_1, \tilde{\nu}, \tilde{\rho})]_\alpha$. Instead of this we have to solve a global optimization problem. This has to be done for each desired point (x, y) – and also for each desired output quantity f.

To reduce the computional effort we approximate the function $f_{(x,y)}$ on a 3-dimensional grid. The gridpoints are given by the Cartesian product $S_{E_1} \times$

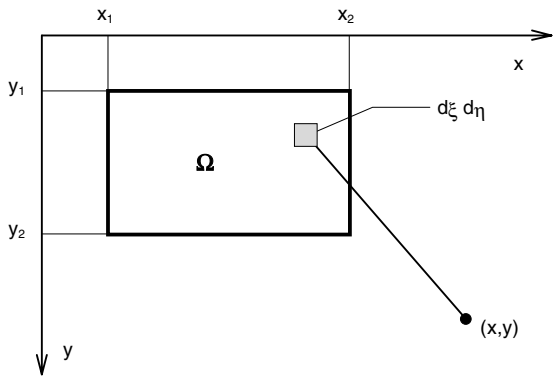


Fig. 4. Uniformly distributed load over a rectangular area.

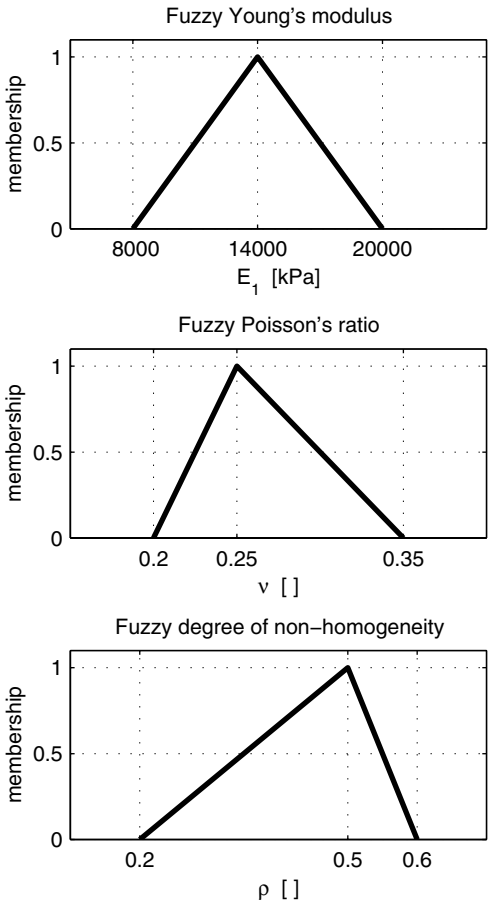


Fig. 5. Membership functions of $\tilde{E}_1$, $\tilde{\nu}$ and $\tilde{\rho}$.

$S_\nu \times S_\rho$ where S_{E_1}, S_ν and S_ρ are discretizations of intervals which include the supports of the fuzzy parameters. For the above example the following sets are taken:

$$S_{E_1} = \{8000, 10000, 12000, \ldots, 20000\} \tag{11}$$

$$S_\nu = \{0.2, 0.25, 0.3, 0.35, 0.4\} \tag{12}$$

$$S_\rho = \{0.0, 0.1, 0.2, \ldots, 0.9\} \tag{13}$$

The FE-computation for a crisp triple $(E_1, \nu, \rho) \in S_{E_1} \times S_\nu \times S_\rho$ using the above method results in $f_{(x,y)}(E_1, \nu, \rho)$ for all desired points (x, y) and for all output quantities f. Doing this for all $(E_1, \nu, \rho) \in S_{E_1} \times S_\nu \times S_\rho$ we get an approximation $\hat{f}_{(x,y)}$ of $f_{(x,y)}$ using linear interpolation. Using this approxi-

mation it is easy to compute approximative bounds of $[f_{(x,y)}(\tilde{E}_1, \tilde{\nu}, \tilde{\rho})]_\alpha$ for all points (x, y).

After performing this for all α-cuts, e.g. $\alpha \in \{0, 0.2, 0.4, \ldots, 1\}$, we get a fuzzy number $f_{(x,y)}(\tilde{E}_1, \tilde{\nu}, \tilde{\rho})$ for each point (x, y), e.g. for each point on a grid on the raft.

Caution: We have to treat each fuzzy number $f_{(x,y)}(\tilde{E}_1, \tilde{\nu}, \tilde{\rho})$ separately from fuzzy numbers at other points, because we have neglected the interactions. So in general fuzzy numbers $f_{(x',y')}(\tilde{E}_1, \tilde{\nu}, \tilde{\rho})$ at additional points (x', y') may not be calculated by interpolation, but interpolation can be used at the stage of the crisp FE-computation.

3.6 Visualizing the fuzzy solution

At single points (x, y)

To represent the value of a single fuzzy quantity at a single point, it suffices to plot the membership function. See Fig. 6 where the fuzzy bending moment $\tilde{M}_x$ at the center of the raft is shown.

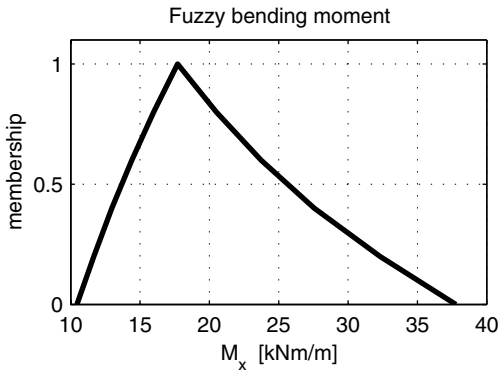


Fig. 6. Membership function of $\tilde{M}_x$.

Along a section

Along the section we plot at each point, for which we have calculated the fuzzy output, the function values indicating the degree of membership by a gray-scale. White represents 0, light gray lower and dark gray higher degree of membership and finally black represents 1. See Fig. 7 where the fuzzy bending moment $\tilde{M}_x$ on a section in x-direction through the middle of the raft is plotted. Taking the membership values of the fuzzy bending moment in Fig. 7 at the location $x = 4$ leads to Fig. 6.

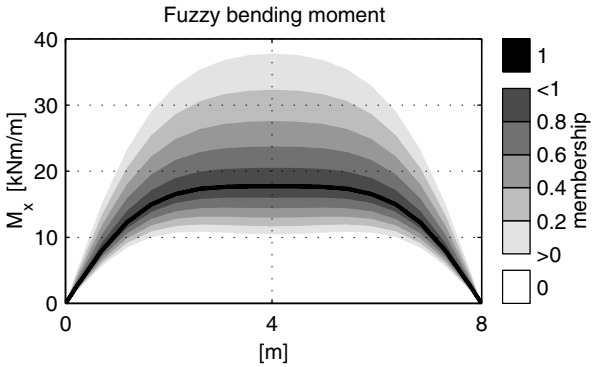


Fig. 7. Fuzzy bending moment $\tilde{M}_x$ along a section.

On the whole raft

To get an overall picture of the stresses or bending moments these quantities f are visualized on the whole raft using contour plots. Usually areas such as $A = \{(x, y) : c_1 \leq f_{(x,y)}(E_1, \nu, \rho) \leq c_2\}$ for fixed E_1, ν and ρ are colored to indicate lower or higher values of f. Here we use a fuzzy extension of this visualizing concept. Let $\tilde{f}(x, y) := f_{(x,y)}(\tilde{E}_1, \tilde{\nu}, \tilde{\rho})$ be the fuzzy value at an arbitrary point (x, y) on the raft. We define the degree of membership of the fuzzy value $\tilde{f}(x, y)$ to the crisp interval $C = [c_1, c_2]$ by

$$m_C(\tilde{f}(x, y)) = \sup_{a \in C} m_{\tilde{f}(x,y)}(a)\,, \tag{14}$$

cf. Fig. 8 showing an example with $m_C(\tilde{f}(x, y)) = 0.5$.

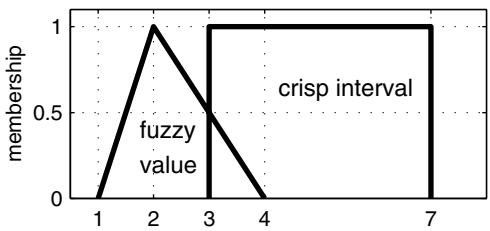


Fig. 8. Fuzzy element of a crisp interval.

Performing this for all (x, y) on a grid on the raft we get a fuzzy set or area which itself is visualized by a contour plot. Fig. 9 and Fig. 10 show the

degree of membership of the fuzzy bending moment $\tilde{M}_x$ to the interval $[7.5, 10]$ kNm/m and $[10, 12.5]$ kNm/m, respectively.

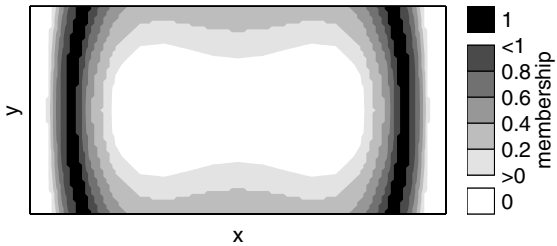


Fig. 9. $\tilde{M}_x \in [7.5, 10]$ kNm/m.

Remark: The areas plotted black in Fig. 7, 9 and 10 represent the output with membership degree 1. This corresponds to the crisp output M_x for $E_1 =$ 14000 kPa, $\nu = 0.25$ and $\rho = 0.5$, cf. Fig. 3.4. In Fig. 6 the value $M_x = 17.71$ kNm/m with membership degree 1 also represents the crisp solution using the above parameters. This shows the great loss of information if uncertainties are neglected and the analysis is done just using crisp soil parameters.

3.7 Determining membership functions

In the numerical example of Section 3.4, membership functions of the fuzzy parameters $\tilde{E}_1$, $\tilde{\nu}$, and $\tilde{\rho}$ have been taken as triangular fuzzy numbers, for the sake of exposition. In practical applications, the membership functions can be constructed from sample data obtained from site investigation and laboratory testing as well as geotechnical judgement provided by the engineer. Various methods for fuzzy data aggregation have been suggested in the literature (see e.g. [1, 3]). We are going to elaborate on one of the approaches, say, for the case of Young's modulus E_1. Assume that in a cross section parallel to the x-axis of the part of soil in question of length L three measurements have been taken at points x_1, x_2, x_3, yielding three intervals I_1, I_2, I_3 for the parameter E_1. If no additional information on the distribution of E_1 over the length L is given, one might define the membership function $m_{\tilde{E}_1}(E_1')$ of the fuzzy parameter $\tilde{E}_1$, at a given value E_1', as $\frac{1}{3}$ the number of intervals I_j in which E_1' is contained. Thus, $m_{\tilde{E}_1}(E_1') = 1$ if E_1' appears as a value in all measurement intervals I_1, I_2, I_3, $m_{\tilde{E}_1}(E_1') = \frac{2}{3}$ if E_1' appears in two of the intervals, and so on, see Fig. 11.

On the other hand, in case more geotechnical information on the approximate distribution of E_1 is available or can be estimated from further indications, one might describe the conjectured distribution over the length L by a

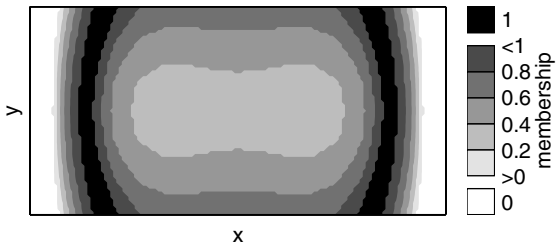


Fig. 10. $\tilde{M}_x \in [10, 12.5]$ kNm/m.

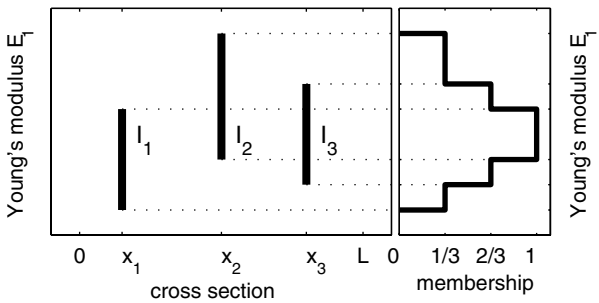


Fig. 11. Determination of fuzzy number $\tilde{E}_1$.

band in the (x, E_1)-plane. A piecewise linear extrapolation is sketched in Fig. 12.

To determine the membership degree $m_{\tilde{E}_1}(E_1')$ of a given value E_1', the length of the broken horizontal at height E_1' cut out by the boundaries of the shaded band is measured and divided by L. For simplification, the result can be approximated by a triangular or trapezoidal fuzzy number. For computational purposes, one could place a rectangular subdivision on the (x, E_1)-plane and just count the number of rectangles met by the shaded region at each height. Here we have presented the procedure in a one-dimensional case; it extends in an obvious fashion to the 2- and 3-dimensional situation.

At this point, an important modelling question arises. Namely, in the fuzzy FE-model, we work with a constant, albeit fuzzy parameter $\tilde{E}_1$. On the other hand, measurements of E_1 have exhibited fluctuations in certain intervals I_j. The question is the meaning of the measurements in the context of the model. There are two possibilities:

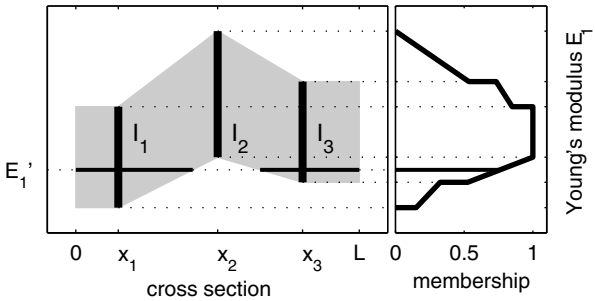


Fig. 12. Determination of fuzzy number $\tilde{E}_1$.

1. The parameter E_1 is approximately constant in reality, and the variations I_j arise from measurement uncertainties. In this case, the use of a constant, fuzzy $\tilde{E}_1$ is fully justified. The membership function simply reflects our information on the measurement uncertainties.
2. The parameter E_1 is not constant, and the fluctuations in our measurements indicate physical variations over the length L. In this case, the actual displacements and stresses might not be covered by a computation using a fuzzy constant $\tilde{E}_1$, even on α-level zero. This problem has to be faced by any modelling procedure, fuzzy or not. It concerns the uncertainty or lack of knowledge about the actual functional dependence of the parameter E_1 on the location along length L, which, in our model, is absorbed in the fuzziness of the assumed constant $\tilde{E}_1$. Preliminary results of a comparative case study under way show that linear variations of E_1 are covered by fuzzy constants, but more drastic fluctuations might lead to bending moments numerically outside the fuzzy range, yet still qualitatively predicted.

4 Fuzzy differential equations

In this section we report on work in progress on dynamical problems. We are concerned with models described by systems of differential equations

$$\begin{aligned} \frac{dx(t)}{dt} &= F(t, x(t), b) \\ x(0) &= a. \end{aligned} \tag{15}$$

Here t denotes time, and $x = (x_1, \ldots, x_n)$ is the state vector of the system. The function F is assumed to be known formally, but may contain fuzzy parameters $b = (b_1, \ldots, b_m)$; the initial state $x(0) = a$ could be a fuzzy vector

as well. Typically, these systems arise in structural mechanics, the vector $x(t)$ comprising displacements and velocities of the nodes of the structure. Further engineering applications include dynamical problems in continuum mechanics (subjected to a spatial discretization), problems of heat conduction described by Newton's law or the time evolution of Markovian probabilities in queueing, maintenance and reliability models.

The parameters and data in problem (15) being given by fuzzy vectors, the solution $x(t)$ at any point of time t will be a fuzzy vector as well and the trajectory of the system $t \to x(t)$ in state space will be a fuzzy function. Our method of computation is by means of the Zadeh extension principle. Assuming that the initial value problem (15) can be solved uniquely when the data and parameters (a, b) are usual numbers, we can assert that the solution is given as a continuous function of the data and parameters

$$x(t) = L_t(a, b). \tag{16}$$

It is this function L_t to which we apply the extension principle. Thus the α-level set of the fuzzy solution $x(t)$ at time t is given by

$$[x(t)]_\alpha = L_t([a]_\alpha, [b]_\alpha). \tag{17}$$

This shows that the solution concept above produces information of engineering interest: if the data and parameters fluctuate in the sets $[a]_\alpha, [b]_\alpha$ with degree of possibility α, then the state of the system at time t is confined to the level set $[x(t)]_\alpha$. This way the assessment of the variations of the input parameters is faithfully processed and reflected in the fuzzy output. In addition, $x(t) = L_t(a, b)$ can be seen as the unique fuzzy solution corresponding to a fuzzy solution concept for system (15), obtaining further mathematical justification this way [11].

As a simple illustrative example, consider the displacement $x(t)$ of a mass m under the influence of a linear spring with stiffness k. It is described by the second order differential equation

$$m\frac{d^2x(t)}{dt^2} + kx(t) = 0. \tag{18}$$

Assuming that parameters m and k are fuzzy while the initial data $x(0) = 0$, $dx(0)/dt = 1$ are known precisely, the fuzzy solution is given by

$$x(t) = L_t(m, k) = \frac{k}{m}\sin(\frac{m}{k}t). \tag{19}$$

Fig. 13 depicts a certain fuzzy parameter m/k, Fig. 14 and 15 the corresponding fuzzy displacement $x(8)$ at time $t = 8$ (polygonal approximation), respectively α-level sets for part of the fuzzy trajectory.

The computation proceeds as follows: First, trajectories $t \to L_t(m.k)$ are computed corresponding to an array of values m/k. Next, the bounding curves

of the level trajectories $[t \to x(t)]_\alpha$ are constructed as envelopes of the trajectories arising from parameter values in level set $[m/k]_\alpha$. Finally, the membership function of $x(t)$ at fixed points of time t is obtained by projection. More efficient numerical algorithms involving interpolation, evolution properties and piecewise monotonicity of the membership functions are currently being developed [12].

As a final observation, we note that from the fuzzy trajectory as a primary object, various secondary quantities of interest can be computed, for example: maximal displacement at level α, resonance frequency, spectral coefficients and so on.

5 Fuzzy data analysis in project planning and construction management

Fuzzy methods can be a valuable help in network planning. They enable the engineer to incorporate his information on the uncertainties of the available data and his assessment of future conditions. They provide a tool for moni-

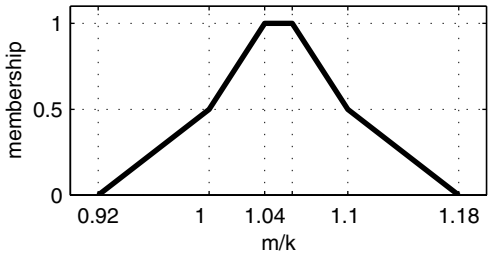


Fig. 13. Fuzzy parameter m/k.

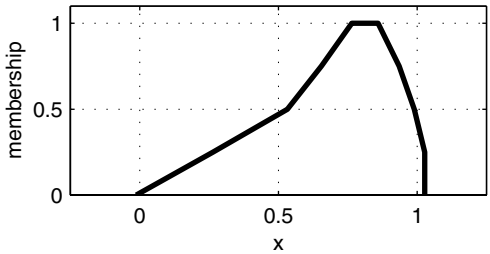


Fig. 14. Fuzzy displacement $x(8)$.

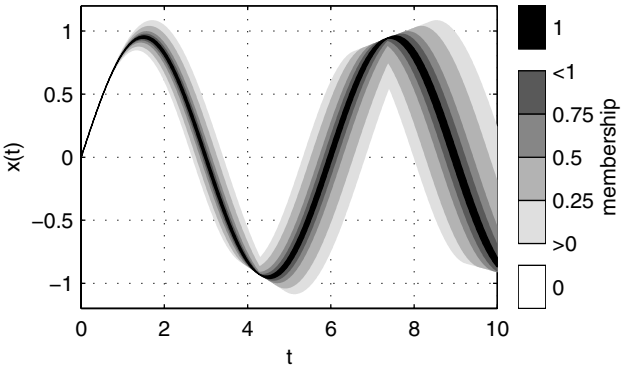


Fig. 15. Fuzzy trajectory.

toring and control, and do not, after all, give the false impression of precision in the project schedule that can never be kept up in reality.

Basic tools in fuzzy network planning and its application to engineering projects have been developed in [9]. In this section we shall elaborate on three additional topics of practical interest: possibilistic modelling of geological data in a tunnelling project; aids for project monitoring; questions of time/cost optimization.

5.1 Time estimate for a tunnelling project: an example

Data for the subsequent example come from a preliminary investigation at the site of a projected road tunnel at the German/Austrian border in geologically challenging terrain. We contrast our approach [7] with a previous study [13], in which the probabilistic PERT-technique had been employed. The total extension of the tunnel of approximately 1250 m was divided into 16 sections, according to geological criteria. For each section, a geologist had provided a verbal description plus estimated percentages of the rock classes to be expected. For example, section 2 of 340 m length was classified as 80% slightly fractured rock and 20% fractured rock. A deterministic engineering estimate yielded driving times for each section and rock class. For example, completion of section 2 was estimated at 60 days, provided only slightly fractured rock was encountered, and at 81 days under conditions of fractured rock. In the PERT-analysis, the duration of each section was interpreted as a discrete random variable with elementary probabilities defined by the given percentages. For example, this way the expected duration for section 2 was computed to $0.8 \cdot 60 + 0.2 \cdot 81 \approx 64$ with standard deviation ≈ 8.4. In section 8, risk of tunnel failure was presumed and in sections 7 and 9, risk of water

inrush. In the PERT-analysis, the corresponding delays were modelled in a similar probabilistic way.

Tunnelling durations as triangular fuzzy numbers: We argue that possibility theory provides a viable alternative. The key to our approach is the interpretation of the percentages as fuzzy ratios. In our treatment of the tunnelling project, we assigned membership degree 1 to the ratios proposed by the geologist, that is 80 : 20 in section 2, for example. In the subsequent analysis, we estimated the ratios defining the bounds for the domain of membership degree zero; in section 2, 90 : 10 and 50 : 50. This resulted in the triangular fuzzy number $\langle 62, 64, 70 \rangle$ for the tunnelling duration in section 2, and similarly for the other 15 sections. Risk of tunnel failure in section 8 was analyzed separately. In particular, the possible occurence of one major tunnel failure and up to two minor ones was taken into account, with corresponding delays modelled by triangular fuzzy numbers as shown in Fig. 16.

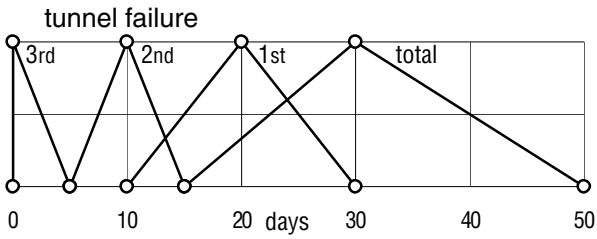


Fig. 16. Delay caused by tunnel failure.

The sum of these prognoses for possible delays was added to the driving duration of section 2. Further delays due to water inrush in sections 7 and 9 were estimated in a similar way and the respective durations were modified accordingly.

We would like to point out that the fuzzy approach allows the incorporation of information going beyond probability distributions. For example, at each transition of the rock classes, individual delays are due to change of the cross-sectional area of the tunnel, change in equipment, safety regulations to be observed and so on. The planning engineer can assess these individual circumstances from previous experience, from specific enterprise data, from discussions with experts involved in the construction process. The information gathered in this way can be subsumed under a formulation by means of fuzzy numbers.

Total duration: The linear structure of the tunneling process is described by a serial network consisting of one path only and 16 nodes for the sections. The total project duration is obtained by simply adding the individual durations, resulting here in a triangular fuzzy number. In contrast to this, in the probabilistic approach, expectation values and variances are added, assuming

stochastic independence of the individual activities. As is customary in the PERT-technique, the total duration is assumed to be normally distributed. The two results are contrasted in Fig. 17.

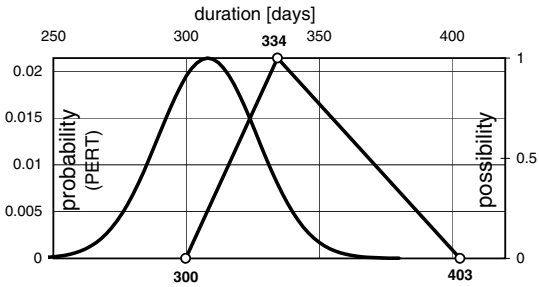


Fig. 17. Probability/possibility distribution of total duration.

The deviation of the central value (membership degree 1) in the fuzzy approach from the expectation value in the probabilistic approach is mainly a consequence of different handling of the exceptional risks in sections 7, 8 and 9. We emphasize that the meaning of the two graphs is totally different; the determining nodes of the triangular fuzzy number directly reflect the risk assessment performed in the analysis, the height of the curve representing the estimated degree of possibility associated with each duration. On the other hand, the area under the probability distribution curve defines the probability that the total duration appears in a certain interval. In view of the various artificial stochastic hypotheses entering in the PERT-algorithm, it is questionable whether it allows a direct interpretation relevant for managerial decisions.

As with most construction projects, the underlying uncertainties are not of a statistical nature. The percentages provided by the geologist are neither samples from a large number of completed tunnels, nor are they statistical averages from a large number of exploratory borings along the prospective tunnel route. They are nothing but subjective estimates based on expertise. Thus it appears more appropriate to translate them into a possibilistic rather than a probabilistic formulation.

5.2 Aids for monitoring

In serial networks such as, for example, arise in tunnelling projects, a fuzzy time/velocity diagram may be taken as a monitoring device for the unfolding of the construction. It simply describes the fuzzy point of time when a certain position along the tunnel route should be reached. The diagram in Fig. 18 reflects all uncertainties taken into account when the project starts.

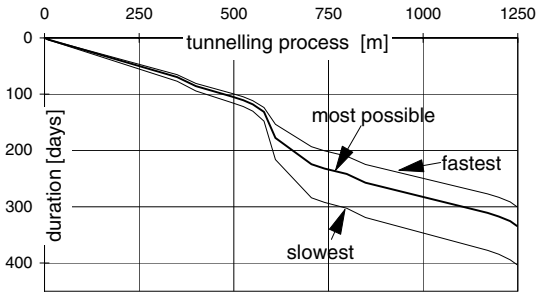


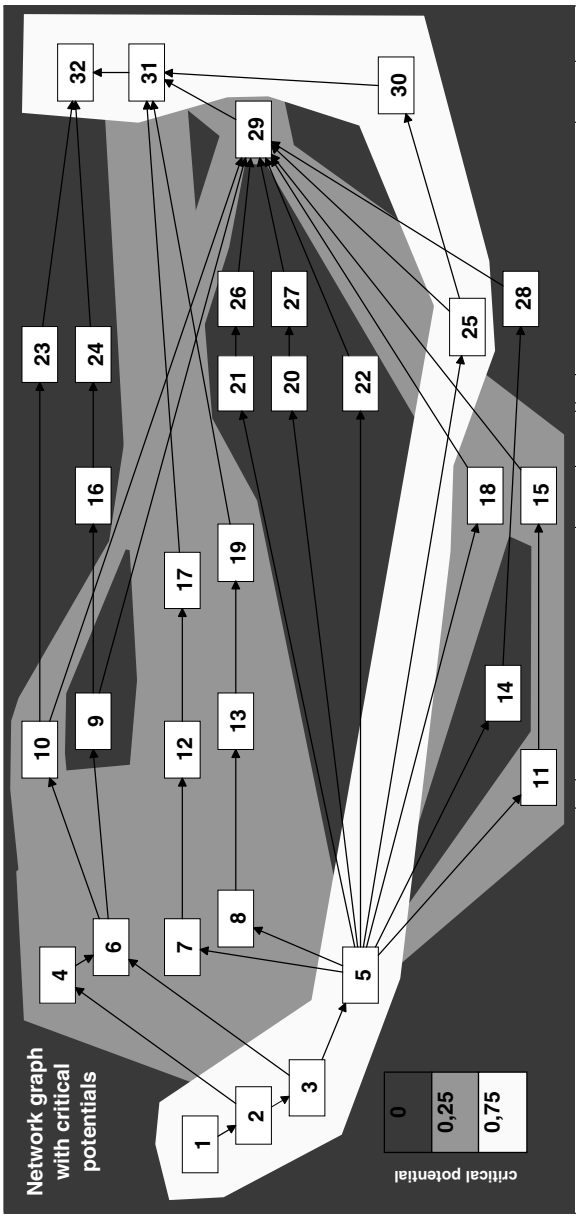
Fig. 18. Time/velocity diagram.

As construction progresses, the uncertainties are narrowed down step by step. The diagram can be actualized by a simple cancellation of the design uncertainties, once a definite state has been reached.

In a branching network, a time/velocity diagram might not contain enough information. Here it is essential to assess the criticality of activities or branches. In deterministic network planning, the slack time for each activity is computed by a backward pass through the graph from the desired completion date, and criticality means slack zero. As explained in detail in [9], the backward pass with fuzzy durations no longer yields the slack of each activity, but rather its critical potential. We prefer to call the fuzzily computed slack *range of uncertainty*. Negative range of uncertainty implies a certain possibility that delay of the completion date may be caused at the respective activity. Numerically, the *critical potential* is defined as the possibility degree of zero range of uncertainty.

This opens the way for enhancing the network presentation by shading (or coloring) areas of different criticality in the network (see Fig. 19, where an example of a project plan for a sewage plant is presented).

Such a presentation uncovers and emphasizes the uncertainties and can help the construction manager to assess rapidly which activities may become critical. In the course of the realization of the project, the diagram can be updated continually, thereby recording shifts in criticality and making it possible to recognize trends early. This is the central objective of monitoring and control, and it is the basis for taking adequate measures in order to avert developments endangering the timely completion of the project or its economic success. As opposed to deterministic planning methods, the project uncertainties do not disappear in the "black box" of an algorithm. The fuzzy network representation may aid all persons concerned with the project in strategic considerations.



Sewage plant (duration in months)

job	description	duration	crit.pot.	job	description	duration	crit.pot.	job	description	duration	crit.pot.
1	start	0	0.83	11	screening plant, grit chamber, part 1	4	0.83	22	Archimedian screw pump station	5	0
2	site installations, river divertion	1,1,2	0.83	12	primary settling tank, part 2	5	0.26	23	digestion towers finishes	4	0
3	bridge, access	2	0.83	13	activated sludge tank, part 2	5	0.83	24	management building finishes, part 2	5	0
4	deep foundation digestion towers, part 1	1.5,2,3	0.26	14	gas plant, part 1	3	0.28	25	small structures, canalization, part 1	8,9,11	0.83
5	filtering well, dewatering	2,2,4	0.83	15	screening plant, grit chamber, part 2	4	0.28	26	secondary settling tank 2	3	0
6	deep foundation digestion towers, part 2	1.5,2,3	0.28	16	management building finishes, part 1	5	0.28	27	secondary sludge thickener	2	0
7	primary settling tank, part 1	1,1,1.5	0.28	17	primary settling tank, part 3	7.5,8,9	0.28	28	gas plant, part 2	4	0.06
8	activated sludge tank, part 1	1,1,1.5	0.28	18	engine house	7.5,8,9.5	0.34	29	asphalt, recultivation, fence, part 1	5.5,6,7	0.56
9	management building carcass	6.5,7,8	0.83	19	activated sludge tank, part 3	7.5,8,9	0.28	30	small structures, canalization, part 2	7	0.83
10	digestion towers carcass	7.5,8,9	0.83	20	primary sludge thickener	3	0	31	asphalt, recultivation, fence, part 2	5	0.83
				21	secondary settling tank 1	3	0	32	end	0	0.83

Fig. 19. Fuzzy network.

5.3 Construction time and cost

In this subsection we address the following: Is it possible to design and implement a cost optimal project plan? There are many parameters to be varied:

1. we can change the overall method of construction;
2. given a construction method, we can change its internal structure, the temporal and causal interdependencies of the individual activities;
3. given a fixed project structure, we can vary the costs of individual activities by acceleration, deceleration and resource modification.

It is clear right away that possibilities (i) and (ii) allow an infinity of variations, depending on the inventiveness of the designing engineer. No single solution can be guaranteed to be optimal; an absolutely optimal construction method or project structure simply does not exist. For the discussion to follow, we therefore concentrate on point (iii) which already features all difficulties (see [8]).

Thus we assume a fixed network structure chosen for the project plan. We want to optimize costs by changing the duration of the activities. As a precondition, we must know the effects of resource modification on the time/cost relation of the activities. This is a major source of uncertainty and will be discussed below. We first have a look at the standard deterministic approach to this optimization problem. The basic assumption is that for each of the activities $A_i, i = 1, \ldots, n$, the time/cost relation is known. We denote by $C_i(D_i)$ the cost of activity A_i when performed at duration D_i. Further, the duration D_i of activity A_i ranges between certain upper and lower bounds:

$$D_{iL} \leq D_i \leq D_{iR}. \tag{20}$$

Each choice of duration $D_1, \ldots D_n$ for the activities results in a total duration T of the project, determined by the activities on the corresponding critical paths. The smallest and largest possible project durations $T_{\min}$, $T_{\max}$ are obtained by performing all activities in minimal time D_{iL}, or maximal time D_{iR}. There is an external time limit on the total duration T, given by the required deadline T_e and the commencement date T_b: T must be smaller than the difference $T_e - T_b$. This results in the constraint on the total duration:

$$T_L \leq T \leq T_R \tag{21}$$

where $T_L = T_{\min}$ and T_R is the smaller of the two values $T_{\max}$ and $T_e - T_b$. The standard optimization proceeds in two steps.

Step 1: cost optimization at fixed project duration T. The duration T is attained by many different combinations of individual durations $D_1, \ldots, D_n$. For each combination, we get a corresponding total cost

$$C_T(D_1, \ldots, D_n) = \sum C_i(D_i). \tag{22}$$

The objective is to minimize $C_T(D_1, \ldots, D_n)$ subject to constraint (20): This results in the minimal cost $C(T)$ at fixed project duration T.

Step 2: optimizing the total project duration. In this step, we simply choose the project duration T^*, subject to constraint (21), such that the corresponding total cost $C(T^*)$ is the least among all minimized costs $C(T)$.

Under various assumptions on the time/cost relationships $C_i(D_i)$, this standard optimization procedure has been extensively dealt with in the literature (see e.g. [10]). It seems we have solved the problem. However, we shall see in the simple example below that the combinatorial structure of the possibilities leading to optimal duration T^* becomes tremendously complex with the increasing size of the project. The construction manager would be required to control the construction process in such a way that each activity runs precisely the duration ultimately producing minimal costs $C(T^*)$. This is impossible. Therefore, the information obtained by the standard optimization procedure is useless.

Time/cost analysis: We turn to fuzzy modelling of the duration dependent costs of a single activity A_i. The costs $C_i(D_i)$ required to complete this activity in a certain time D_i are described by a triangular fuzzy number, see Fig. 20.

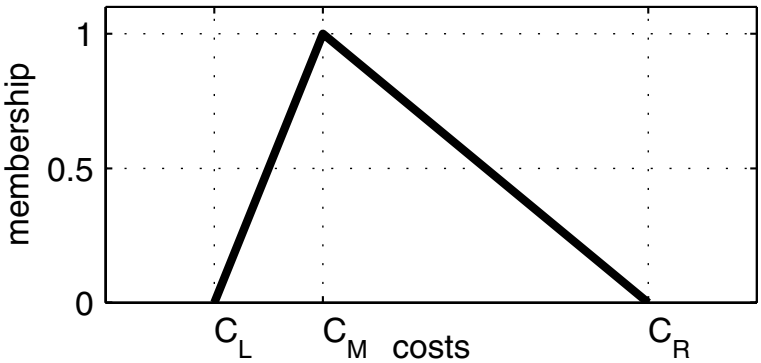


Fig. 20. Membership function for cost of activity at fixed duration.

The planning engineer might first arrive at the central value C_M with membership degree 1 by employing the standard deterministic computations from construction management data. Then a risk analysis might provide the lowest possible costs C_L and a largest bound C_R for the estimated costs. Of course, further subdivision of risk levels can provide a refined analysis (this was carried through e.g. in [9]), but as a first approximation, a triangular fuzzy number may satisfactorily reflect the cost fluctuations under risk.

Next, we discuss the cost distribution as the duration D_i of activity A_i varies in its bounds D_{iL}, D_{iR}. An example of such a diagram is shown in Fig. 21, exhibiting 0-level and 1-level curves of the fuzzy cost $C_i(D_i)$.

We can distinguish three regimes of the time/cost dependence:

1. *Normal area*: This is the normal range for completing the activity. Costs for equipment, labor and material and costs for site overhead are balanced.

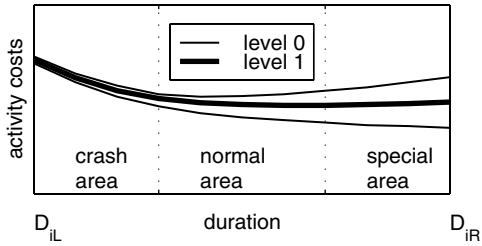


Fig. 21. Fuzzy time/cost dependence.

2. *Crash area*: The acceleration of the activity causes the costs for equipment, labor and material to predominate. The increase of capacities may cause interference and unintended obstructions, thereby further raising costs uneconomically. In any case, a progressive cost development is to be expected when the activity duration is pushed to its lower limit.
3. *Special area*: The largest uncertainties arise in this range. Costs for maintaining the construction site at minimum capacity may be high or low, depending very much on the specific project and circumstances.

We point out that the diagram in Fig. 21 depicts but one of the many time/cost relationships that can arise. They typically are nonlinear, but may even have discontinuities. For example, to achieve a certain acceleration, the construction process may have to be changed, additional machinery may be required, or shift-work may have to be introduced.

Example: As an illustrative example, we consider a simple serial network consisting of the three activities: excavation (A_1), foundation (A_2) and mat construction (A_3), see Fig. 22.

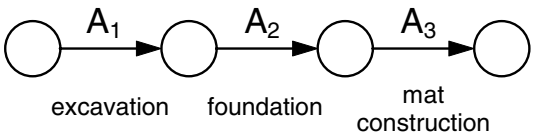


Fig. 22. Serial network.

The durations D_i of each activity vary in an interval $[D_{iL}, D_{iR}]$ and thus can be described by a rectangular fuzzy number. We choose $[D_{1L}, D_{1R}] = [3, 8], [D_{2L}, D_{2R}] = [D_{3L}, D_{3R}] = [7, 12]$. For each activity, a time/cost relationship as in Fig. 21 is assumed. The resulting total duration is the sum of the three duration intervals and thus may vary in the interval $[17, 32]$. For

computational simplicity, we allow only integer values for each duration. Following the pattern of the deterministic optimization algorithm, we first choose a fixed duration T in the interval $[17, 32]$ and determine the combinations of individual durations D_1, D_2, D_3 summing up to T. Each combination requires a cost of $C_T(D_1, D_2, D_3) = C_1(D_1) + C_2(D_2) + C_3(D_3)$, represented by a triangular number. Superposition of these triangular numbers shows the cost variations that can arise if the total project duration is T ($T = 29$ in Fig. 23).

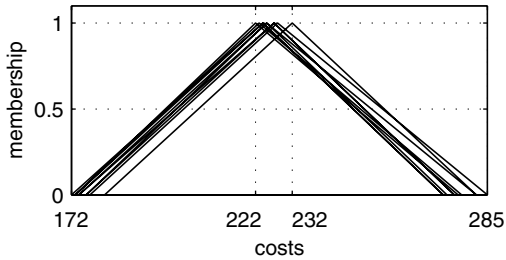


Fig. 23. Fuzzy cost variations at duration $T = 29$.

The envelope of these triangular numbers can be approximated by a trapezoidal number, see Fig. 24.

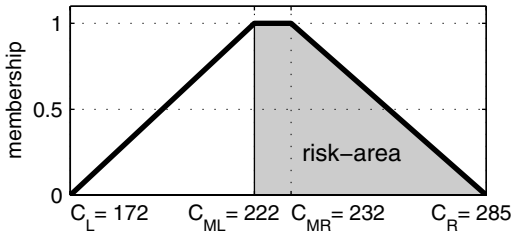


Fig. 24. Risk assessment of range of costs.

We note that the boundaries C_{ML}, C_{MR} of the central plateau arise already from the combinatorial possibilities when all activities run at deterministic costs C_{iM} (membership degree 1). Thus the shaded region can be considered as an indicator of the economic risk the designing engineer has to face. To assess the risk of economic failure, the engineer should examine the combinations of activity durations leading up to the characterizing values C_{ML}, C_{MR}, C_R, respectively. The number of these combinations grows fast

with the project size. Thus in practice this is an impossible task. One must be content with the indications extractable from the diagram and estimates from a detailed study of a few extremal cases.

Three-dimensional presentation: To each attainable total duration T in the interval $[T_L, T_R]$ there are corresponding costs described by a trapezoidal fuzzy number as in Fig. 24. We can collect these in a 3-dimensional diagram, see Fig. 25.

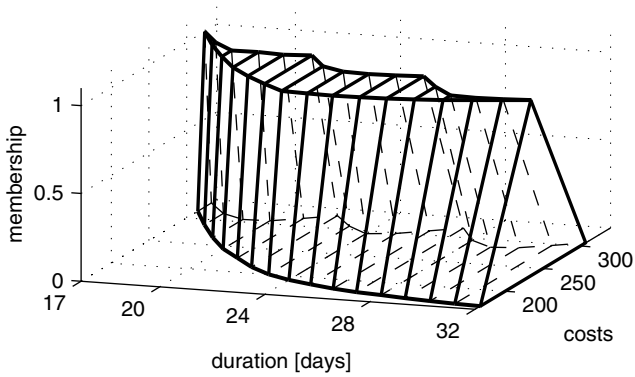


Fig. 25. Time/cost/possibility diagram.

The height of the resulting surface over a point (T, C) shows the degree of possibility that the project duration is T with project costs C. The trapezoidal number in Fig. 24 arises as a cross-section of the surface at fixed duration T. The plateau area of possibility degree 1 embraces the time/cost combinations attained when all activities run at deterministic cost C_{iM}. This 3-dimensional graph puts in evidence the domain in which cost and duration of the project may vary, when the duration of each single activity has been modeled by a rectangular fuzzy number, single costs by a triangular fuzzy number and the time/cost relation by a diagram as in Fig. 21. Accelerating or decelerating single activities will move the result within the boundaries of this domain.

Final Remark: We realize from these considerations that in a time/cost analysis with fuzzy data a large variety of possible results with different membership degrees arise, already in a simple example involving three activities only. In view of this observation it is clear that the goal of classical optimization – the search for and the implementation of a cost optimal project plan – cannot be achieved. Not only is it legitimate from a modelling perspective to assume that cost and duration data are fuzzy, but in real life construction projects the only type of data available are fuzzy data. We may conclude that with respect to cost and duration of construction projects, one cannot strive

for optimization, but rather should attempt to achieve a reasonable solution within the limitations of the given risks and uncertainties.

These considerations also show that text book strategies to accomplish a certain result do not exist, but every measure and its effects have to be evaluated in each specific situation. Frequently heard statements such as "Reduction of the construction period will reduce costs" have no validity, with the possible exception of specific projects where they may have resulted from a thorough investigation of the determining factors. It is essential that the designing engineer consider the data and project uncertainties from the earliest planning phase onwards, so as to have a firm basis for the assessment of the risk of economic failure.

As the economist H. A. Simon puts it [14],

> "...exact solutions to the larger optimization problems of the real world are simply not within reach or sight. In the face of this complexity the real-world business firm turns to procedures that find good enough answers to questions whose best answers are unknowable. Thus normative microeconomics, by showing real-world optimization to be impossible, demonstrates that economic man is in fact a satisficer, a person who accepts 'good enough' alternatives, not because he prefers less to more but because he has no choice."

References

[1] H. Bandemer and W. Näther. *Fuzzy Data Analysis.* Kluwer, Dordrecht, 1992.

[2] J.R. Booker, N.P. Balaam, and E.H. Davis. The behaviour of an elastic, non-homogeneous half-space. Part I – Line load and point loads. *Int. J. num. analytical Meth. Geomechanics*, 9:353–367, 1985.

[3] D. Dubois and H. Prade. *Possibility Theory.* Plenum Press, New York, 1988.

[4] Th. Fetz. Finite element method with fuzzy parameters. In I. Troch and F. Breitenecker, editors, *Proceedings IMACS Symposium on Mathematical Modelling, Vienna 1997*, volume 11 of *ARGESIM Report*, pages 81–86, 1997.

[5] Th. Fetz, M. Hofmeister, G. Hunger, J. Jäger, H. Lessmann, M. Oberguggenberger, A. Rieser, and R. Stark. Tunnelberechnung – Fuzzy? *Bauingenieur*, 72:33–40, 1997.

[6] M. Klisinski. Plasticity theory based on fuzzy sets. *Journal of Engineering Mechanics*, 114(4):563–583, 1988.

[7] D. Köll. Netzplanberechnung mit unscharfen Zahlen. Diplomarbeit, Universität Innsbruck, 1997.

[8] G. Krenn. Kosten und Bauzeit. Eine Untersuchung über Zusammenhänge mit Fuzzy-Methoden. Diplomarbeit, Universität Innsbruck, 1996.

[9] H. Lessmann, J. Mühlögger, and M. Oberguggenberger. Netzplantechnik mit unscharfen Methoden. *Bauingenieur*, 69:469–478, 1994.

[10] K. Neumann. *Operations Research Verfahren*, volume 3. Carl Hanser Verlag, München, 1975.

[11] M. Oberguggenberger. Fuzzy differential equations. In I. Troch and F. Breitenecker, editors, *Proceedings IMACS Symposium on Mathematical Modelling, Vienna 1997*, volume 11 of *ARGESIM Report*, pages 75–80, 1997.

[12] S. Pittschmann. Lösungsmethoden für Funktionen und gewöhnliche Differentialgleichungen mit unscharfen Parametern. Diplomarbeit, Universität Innsbruck, 1996.

[13] S. Plankensteiner. Unsicherheiten im Projektablauf, Fallbeispiel: Grenztunnel Füssen-Vils. Diplomarbeit, Universität Innsbruck, 1991.

[14] H.A. Simon. *The Science of the Artificial.* MIT Press, Cambridge, 1981.

[15] R.F. Stark and J.R. Booker. Surface displacements of a non-homogeneous elastic half-space subjected to uniform surface tractions. Part I – Loading on arbitrarily shaped areas. *Int. J. num. analytical Meth. Geomechanics*, 21(6):361–378, 1997.

[16] R.F. Stark and J.R. Booker. Surface displacements of a non-homogeneous elastic half-space subjected to uniform surface tractions. Part II – Loading on rectangular shaped areas. *Int. J. num. analytical Meth. Geomechanics*, 21(6):379–395, 1997.

[17] P.-A. Von Wolffersdorf. Feldversuch an einer Spundwand im Sandboden: Versuchsergebnisse und Prognosen. *Geotechnik*, 17:73–83, 1994.

[18] L. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

Authors (alphabetical listing)

Wolfgang Fellin
Institut für Geotechnik und Tunnelbau
Universität Innsbruck, Technikerstraße 13
A-6020 Innsbruck, Austria
e-mail: Wolfgang.Fellin@uibk.ac.at

Thomas Fetz
Institut für Technische Mathematik, Geometrie und Bauinformatik
Universität Innsbruck, Technikerstraße 13
A-6020 Innsbruck, Austria
e-mail: Thomas.Fetz@uibk.ac.at

Ivo Herle
Institut für Geotechnik
Technische Universität Dresden
D-01062 Dresden, Germany
e-mail: Ivo.Herle@mailbox.tu-dresden.de

Günter Hofstetter
Institut für Baustatik, Festigkeitslehre und Tragwerkslehre
Universität Innsbruck, Technikerstraße 13
A-6020 Innsbruck, Austria
e-mail: Guenter.Hofstetter@uibk.ac.at

Johannes Jäger
Edith-Stein-Weg 2
A-6020 Innsbruck, Austria
e-mail: Johannes.Jaeger@bemo.co.at

David Köll
Jakob-Wibmer-Straße 2
A-9971 Matrei in Osttirol, Austria
e-mail: D.Koell@tirol.com

Günther Krenn
Bayrisch-Platzl-Straße 15/74
A-5020 Salzburg, Austria
e-mail: Guenther.Krenn@porr.at

Hermann Lehar
Institut für Baustatik, Festigkeitslehre und Tragwerkslehre
Universität Innsbruck, Technikerstraße 13
A-6020 Innsbruck, Austria
e-mail: Hermann.Lehar@uibk.ac.at

Heimo Lessmann
Starkenbühel 304
A-6073 Sistrans, Austria

Gert Niederwanger
Institut für Baustatik, Festigkeitslehre und Tragwerkslehre
Universität Innsbruck, Technikerstraße 13
A-6020 Innsbruck, Austria
e-mail: Gerhard.Niederwanger@uibk.ac.at

Michael Oberguggenberger
Institut für Technische Mathematik, Geometrie und Bauinformatik
Universität Innsbruck, Technikerstraße 13
A-6020 Innsbruck, Austria
e-mail: Michael.Oberguggenberger@uibk.ac.at

Alexander Ostermann
Institut für Technische Mathematik, Geometrie und Bauinformatik
Universität Innsbruck Technikerstraße 13
A-6020 Innsbruck, Austria
e-mail: Alexander.Ostermann@uibk.ac.at

Francesco Russo
Département de Mathématiques, Institut Galilée
Université Paris-Nord, 99, Avenue Jean-Baptiste Clément
F - 93430 Villetaneuse, France
e-mail: russo@math.univ-paris13.fr

Rudolf Stark
Institut für Baustatik, Festigkeitslehre und Tragwerkslehre
Universität Innsbruck, Technikerstraße 13
A-6020 Innsbruck, Austria
e-mail: Rudolf.Stark@uibk.ac.at

Robert Vieider
Vieider Ingenieur GmbH
Rebschulweg 1/E
I-39052 Kaltern an der Weinstraße, Italy
e-mail: vieider.ingenieur@rolmail.net