



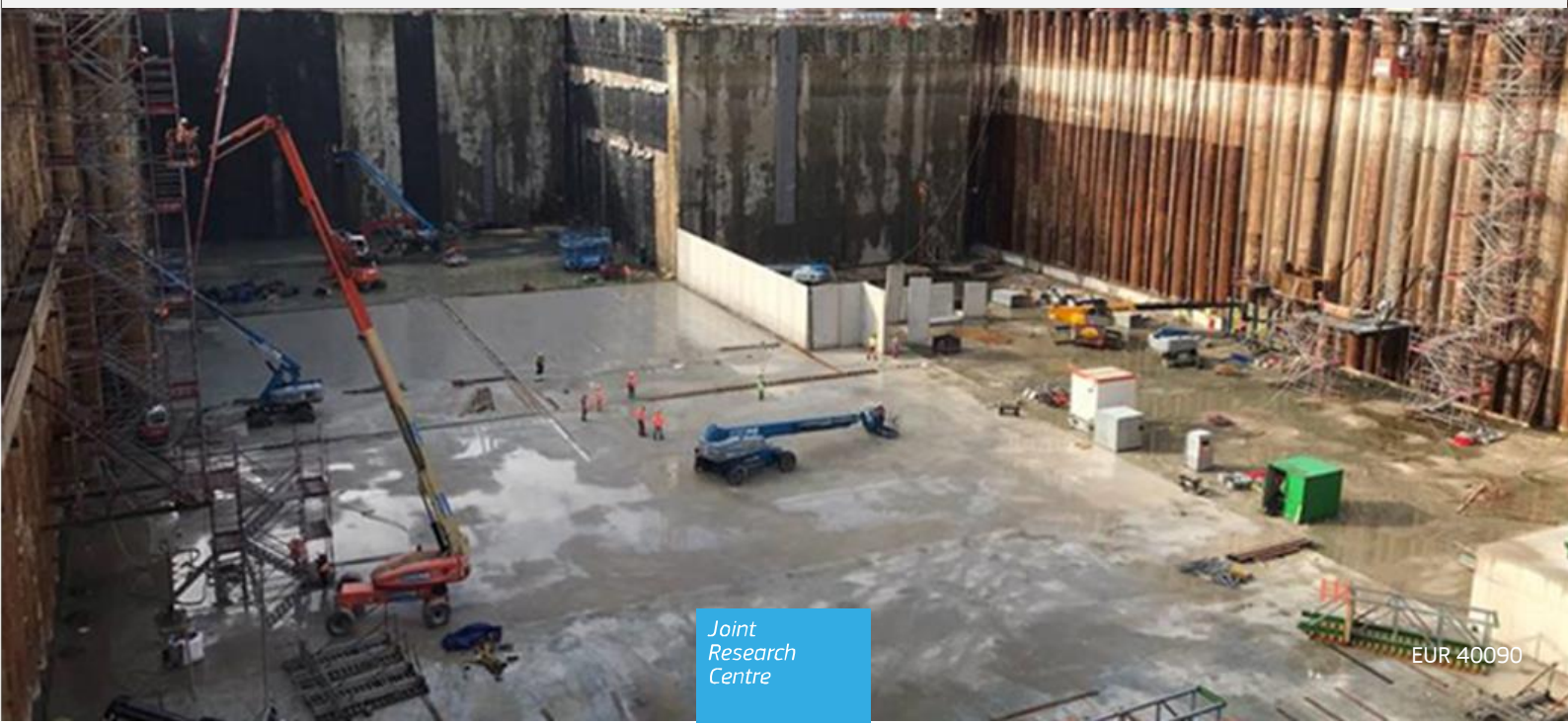
Reliability-based verification of limit states for geotechnical structures

Guidelines for the application of the 2nd generation of Eurocode 7: Geotechnical design

van Den Eijnden, B., Knuuti, M., Lesny, K., Löfman, M., Mavritsakis, A., Roubos, A., Schweckendiek, T., Sciarretta, F., Ebener, A., Escher, K., Spross, J., Commend, S., Hehenkamp, M., Arnold, P., Wilhelm, S., Ene, A., Rimoldi, P., Pereira, R.

Edited by Schweckendiek, T.

2024



This document is a publication by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The contents of this publication do not necessarily reflect the position or opinion of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. For information on the methodology and quality underlying the data used in this publication for which the source is neither Eurostat nor other Commission services, users should contact the referenced source. The designations employed and the presentation of material on the maps do not imply the expression of any opinion whatsoever on the part of the European Union concerning the legal status of any country, territory, city or area or of its authorities, or

EU Science Hub

JRC139408

EUR 40090

Print ISBN 978-92-68-21385-8 ISSN 1018-5593 doi:10.2760/0991189 KJ-01-24-103-EN-C
PDF ISBN 978-92-68-21384-1 ISSN 1831-9424 doi:10.2760/1342542 KJ-01-24-103-EN-N

Luxembourg: Publications Office of the European Union, 2024

© European Union / van Den Eijnden, B. / Knuuti, M. / Lesny, K. / Löfman, M. / Mavritsakis, A. / Roubos, A. / Schweckendiek, T. / Ebener, A. / Escher, K. / Spross, J. / Commend, S. / Hehenkamp, M. / Arnold, P. / Wilhelm, S. / Ene, A. / Rimoldi, P. / Pereira, R., 2024



The reuse policy of the European Commission documents is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Unless otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

For any use or reproduction of photos or other material that is not owned by the European Union permission must be sought directly from the copyright holders.

- Cover page illustration, © Adriaan Van Seters, 2000

How to cite this report: European Commission, Joint Research Centre, van Den Eijnden, B., Knuuti, M., Lesny, K., Löfman, M., Mavritsakis, A., Roubos, A., Schweckendiek, T., Sciarretta, F., Ebener, A., Escher, K., Spross, J., Commend, S., Hehenkamp, M., Arnold, P., Wilhelm, S., Ene, A., Rimoldi, P. and Pereira, R., *Reliability-based verification of limit states for geotechnical structures*, Schweckendiek, T. editor, Publications Office of the European Union, Luxembourg, 2024, <https://data.europa.eu/doi/10.2760/1342542>, JRC139408.

Contents

Abstract	8
Forewords.....	9
Acknowledgements.....	13
1 Introduction	14
1.1 Context.....	14
1.2 Objectives.....	14
1.3 Target audience.....	14
1.4 Scope	15
1.5 Application of this guideline	15
1.6 Outline.....	16
2 Reliability background of the Eurocodes	17
2.1 Reliability aspects in the Eurocodes.....	17
2.2 Uncertainties and probability.....	18
2.2.1 Sources and types of uncertainties.....	18
2.2.2 Probability interpretation	18
2.3 Principles of structural reliability.....	18
2.3.1 Risk-based approach.....	19
2.3.2 Reliability-based approach.....	19
2.3.3 Semi-probabilistic approach.....	19
2.3.4 Observed failure frequency, notion of hidden safety	20
2.3.5 System reliability	21
2.3.6 Time effects.....	21
2.4 Target reliability	22
2.4.1 Safety and economy.....	22
2.4.2 Reliability differentiation.....	23
2.4.3 Annual versus lifetime targets.....	23
2.4.4 Minimum versus average reliability	23
2.4.5 Member versus system	23
2.4.6 Brittle versus ductile failure	24
2.5 Assessment of existing structures.....	24

3	Reliability verification procedure.....	25
3.1	General.....	25
3.2	Problem definition.....	26
3.3	Uncertainty characterization.....	26
3.4	Reliability analysis.....	27
4	Reliability requirements.....	28
4.1	Target values for reliability index (ULS).....	28
4.1.1	Target values in EN 1990-1.....	28
4.1.2	Consequence classes.....	29
4.1.3	Risk-informed approach.....	29
4.1.4	Other sources.....	30
4.2	Member versus system reliability.....	31
4.3	Reliability targets for the resistance or load only.....	31
4.4	Annual versus lifetime reference period.....	33
4.5	Serviceability limit states (SLS).....	34
4.6	Assessment of existing structures.....	35
4.7	Temporary structures.....	35
5	Uncertainties in geotechnical design and assessment.....	37
5.1	Principles and definitions.....	37
5.1.1	Sources of uncertainty.....	37
5.1.2	Uncertainties in geotechnical models.....	37
5.1.3	Probability interpretation.....	40
5.2	Geotechnical units in Ground Model.....	42
5.2.1	Stratification (ground layers).....	42
5.2.2	Geological and man-made anomalies.....	45
5.2.3	Groundwater.....	45
5.3	Ground properties.....	46
5.3.1	Introduction to treating uncertainty in ground properties.....	46
5.3.2	Uncertainty components.....	47
5.3.3	Inherent variability.....	48
5.3.3.1	Assessment of inherent variability from data.....	50

5.3.3.2	Estimation of scale of fluctuation.....	53
5.3.4	Spatial averaging.....	55
5.3.4.1	Weak zone seeking.....	58
5.3.4.2	Non-averaging properties.....	59
5.3.5	Measurement error.....	60
5.3.5.1	Measurement error in spatially averaged ground properties.....	61
5.3.6	Statistical uncertainty.....	61
5.3.7	Transformation uncertainty.....	63
5.3.8	Total uncertainty.....	65
5.3.8.1	Recommended indicative values of total uncertainty.....	65
5.3.8.2	Total uncertainty – general equation.....	66
5.3.8.3	Directly measured ground properties (site-specific dataset).....	67
5.3.8.4	Indirectly measured ground properties (using transformation models).....	67
5.4	Defining probability distributions.....	68
5.4.1	Useful probability distributions.....	68
5.4.2	Practical approaches to defining probability distributions.....	70
5.4.2.1	Approach A: Fitting to site-specific data.....	70
5.4.2.2	Approach B: Combining site-specific data and literature guidelines.....	70
5.4.2.3	Approach C: Physical constraints and engineering judgement.....	72
5.4.2.4	Approach D: Bayesian methods.....	73
5.4.2.5	Approach E: Prescribed distribution types.....	73
5.4.3	Cross-correlation between ground properties.....	74
5.4.4	Depth-dependency.....	76
5.5	Model uncertainty.....	79
5.5.1	Model factor (definition).....	79
5.5.2	Estimation.....	79
5.5.3	Modelling.....	79
5.5.4	Recommended indicative values.....	80
5.6	Actions.....	80
6	Reliability analysis for geotechnical applications.....	83
6.1	Principles and definitions.....	83

6.1.1	Limit state and limit state function	83
6.1.2	Random variables and joint probability distribution.....	84
6.1.2.1	Continuous variables.....	84
6.1.2.2	Discrete variables.....	84
6.1.2.3	Correlations	85
6.1.3	Probability of failure and reliability index	87
6.1.4	Time-dependent variables.....	88
6.1.5	System reliability aspects.....	88
6.2	Reliability analysis methods.....	89
6.2.1	Basic notions	89
6.2.2	First-order second moment method (FOSM)	91
6.2.3	Point estimate method (PEM).....	92
6.2.4	First order reliability method (FORM).....	92
6.2.5	Monte Carlo simulation (MCS).....	94
6.2.6	Other sampling methods.....	95
6.2.6.1	Importance sampling (IS).....	95
6.2.6.2	Subset simulation (SubSim).....	96
6.2.6.3	Directional sampling (DS).....	97
6.2.6.4	Surrogate models.....	98
6.3	Practical recommendations	98
6.3.1	Limit state functions.....	98
6.3.1.1	Explicit (analytical) versus implicit (numerical) limit states.....	99
6.3.1.2	System and component limit states.....	99
6.3.2	Selection of random variables	101
6.3.3	Selection of reliability method.....	101
6.3.3.1	Computation time of the model	103
6.3.3.2	Number of random variables (dimensionality)	104
6.3.3.3	(Non-)Linearity.....	104
6.3.3.4	(Numerical) model stability and existence of solution	104
6.3.4	Verification of results (as part of design check).....	104
6.3.4.1	Design point (FORM).....	105

6.3.4.2	Uncertainty bounds for sampling-based reliability methods.....	105
6.3.5	Point estimates.....	107
6.3.6	Reliability-based design values.....	107
6.3.7	Conditional reliability analysis / fragility curve approach.....	109
6.4	Benchmark example.....	111
6.4.1	Limit state function.....	112
6.4.2	Reliability methods.....	113
6.4.3	FORM design point and influence factor.....	115
6.5	Random finite element method (RFEM).....	115
6.5.1	Software.....	116
7	Bayesian analysis of geotechnical data.....	118
7.1	Introduction.....	118
7.2	Principles and definitions of Bayes' theorem.....	118
7.3	Steps for performing Bayesian inference.....	120
7.3.1	Establishing prior distributions.....	120
7.3.2	Formulating the likelihood function.....	122
7.3.3	Bayesian inference (updating).....	123
7.3.3.1	Numerical integration.....	123
7.3.3.2	Conjugate priors (analytical solution).....	124
7.3.3.3	Sampling algorithms.....	124
7.3.3.4	Variational Inference.....	124
7.3.3.5	Selecting a method.....	125
7.3.4	Interpreting the posterior distribution.....	125
7.3.4.1	Posterior and posterior predictive distributions.....	126
7.4	Bayesian inference in geotechnical applications.....	127
7.5	Bayesian parameter estimation.....	128
7.6	Reliability updating.....	132
7.7	Observational method.....	134
8	Reliability-based partial factors.....	136
8.1	Semi-probabilistic approach.....	136
8.1.1	Verification formats.....	136

8.1.2	Representative and characteristic values	138
8.2	Calibration procedures.....	138
8.2.1	Minimizing the scatter in reliability achieved.....	139
8.2.2	Design value method.....	141
8.2.3	Quantile-based method.....	143
8.3	Project-specific partial factors.....	143
8.4	Calibration examples.....	144
9	Closing remarks.....	145
	References	146
	List of abbreviations and definitions.....	154
	List of boxes.....	156
	List of figures.....	158
	List of tables.....	164
	Annexes	167
	Annex A. Literature summary of inherent variability and uncertainties.....	167
	A.1. Inherent variability.....	167
	A.2. Measurement error.....	168
	A.3. Transformation uncertainty.....	168
	References	169
	Table summaries for transformation uncertainty of clay, sand, and rock (intact and rock mass) extracted from ISSMGE-TC304 report.....	170
	A.4 Model uncertainty.....	171
	Table summary for of model factor statistics for various geotechnical structures.....	172
	A.5. Groundwater level (background).....	173
	A.6. Statistical uncertainty.....	174
	Annex B. Examples.....	176
	B.1. Introduction.....	176
	B.2. Slopes, cuttings, and embankments.....	176
	B.2.1 Settlement analysis for embankments using preloading without surcharge	177
	B.3. Spread foundations.....	185
	B.3.1 Vertically loaded spread foundation.....	185
	B.4. Piled foundations.....	190

B.4.1 Comparison of reliability levels achieved by the pile design method of EN 1997-3 with the reliability targets of EN 1990.....	190
B.4.2 Reliability updating of driven pile using pile tests and Bayesian inference.....	196
B.5. Retaining structure.....	207
B.5.1 Probabilistic FEM analysis of a retaining wall.....	207
B.5.2 Soldier pile wall.....	220
B.5.3 Propped embedded retaining wall.....	227
B.6. Reinforced fill structures.....	233
B.6.1 Geogrid reinforced soil wall.....	233
B.7. Rock Engineering.....	246
B.7.1 Rock slope stability.....	246
B.8. Further readings.....	255

Abstract

Safety in geotechnical design has evolved from experience and judgment using overall factor of safety approaches to design methods in modern codes of practice which apply partial factors to loads and to strength parameters or resistances. Explicit treatment of uncertainties using probabilistic methods has been receiving more attention recently, both in the justification of partial factors in codes of practice as well as in design and assessment. The Eurocodes are fundamentally reliability-based, even though not all safety-relevant aspects are explicitly tied to the underlying reliability concepts as explained in the JRC Technical Report 'Reliability background of the Eurocodes' (European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al., 2024).

EN 1997-1 explicitly states reliability-based methods as one of the options to verify limit states of geotechnical structures alongside the partial factor method, prescriptive rules, testing, and the observational method. The present document serves as a guideline for reliability-based verification of limit states in design and assessment of geotechnical structures within the safety and reliability concepts of EN 1990-1 and EN 1997-1.

Forewords

JRC Foreword

The construction ecosystem is of strategic importance to the European Union (EU), as it delivers the buildings and infrastructures needed by the rest of the economy and society, having a direct impact on the safety of persons and the quality of citizens' life. The construction ecosystem includes activities carried out during the whole lifecycle of buildings and infrastructures, namely design, construction, maintenance, refurbishment and demolition. The industrial construction ecosystem employs around 25 million people in the EU and provides an added value of EUR 1 158 billion (9.6% of the EU total)^{1,2,3}.

The construction ecosystem is a key element for the implementation of the European Single Market and many other important EU strategies and initiatives. The European Green Deal ([COM\(2019\) 640 final](#)) aims to achieve climate neutrality for Europe by 2050, and relies on numerous initiatives, noteworthy:

- the New Circular Economy Action Plan ([COM\(2020\) 98 final](#)) and the New Industrial Strategy for Europe ([COM\(2020\) 102 final](#)) intending to accelerate the transition of the EU industry to a sustainable model based on the principles of circular economy;
- the revision ([COM\(2022\) 144 final](#)) of the Construction Products Regulation ([Regulation \(EU\) No 305/2011](#)) aiming to enable the construction ecosystem's contribution to meeting climate and sustainability goals and embrace the digital transformation of the built environment;
- the New EU Strategy on Adaptation to Climate Change ([COM \(2021\) 82 final](#)) supported by the recent Commission Communication on managing climate risks ([COM\(2024\) 91 final](#)) that reinforces the need to address climate change concerns to guarantee resilience and sustainability of built structures and infrastructures and to ensure regular science-based risk assessments;
- the first European Climate Risk Assessment ([EUCRA](#)) report which highlights the importance of EU policies for the built environment, including updating construction standards and related European datasets.

Furthermore and recognizing that the EU's ambitions towards a climate neutral, resilient and circular economy cannot be delivered without leveraging the European standardization system, the European Commission presented a new Standardization Strategy ([COM\(2022\) 31 final](#)). The strategy spots standards as “*the silent foundation of the EU Single Market and global competitiveness*”.

The EU has put in place a comprehensive legislative and regulatory framework for the construction sector, including European standards (EN). Within this framework, the [Eurocodes](#) are a series of 10 European standards, EN 1990 to EN 1999, providing common technical rules for the design of buildings and other civil engineering works. In fact, the Commission Communication on managing

¹ Commission staff working document: Scenarios for a transition pathway for a resilient, greener and more digital construction ecosystem (<https://ec.europa.eu/docsroom/documents/47996>)

² Council of the EU, Press release 30 June 2023, <https://www.consilium.europa.eu/en/press/press-releases/2023/06/30/council-adopts-position-on-the-construction-products-regulation/>

³ Transition Pathway For Construction, European Commission, DG GROW, <https://ec.europa.eu/docsroom/documents/53854>

climate risks directly mentions the Eurocodes, highlighting the role of building and infrastructure standards in integrating climate adaptation and resilience.

The [Commission Recommendation 2003/887/EC](#) on the implementation and use of the Eurocodes for construction works and structural construction products recommends undertaking research to facilitate the integration into the Eurocodes of the latest developments in scientific and technological knowledge. In this context, the so-called second generation of the Eurocodes is under development under Mandate M/515 and expected to be available by 2026. The second generation Eurocodes incorporates improvements to the existing standards and extends their scope by embracing new methods, new materials, and new regulatory and market requirements, including considerations for climate change impact on structural design.

In order to support the implementation of the second generation EN 1997 “Geotechnical Design”, CEN Technical Committee 250/Sub-Committee 7 (TC 250/SC 7) produced a series of guidelines addressing the most important new aspects in the standard. The series of guidelines contains the following documents:

- Guideline C1 – “Determination of representative values from derived values for verification of limit states with EN 1997”
- Guideline C2 – “Assembling the Ground model and the derived values”
- Guideline C3 – “Reliability-based verification of limit states for geotechnical structures”
- Guideline C4 – “Implementation of design in execution and service life”

Within the framework of Administrative Arrangements between the European Commission’s Joint Research Centre (JRC) and DG GROW on support to policies and standards for the construction ecosystem, JRC is engaged in activities facilitating the implementation and practical use of the second generation Eurocodes. In this context, the guidelines by TC 250/SC 7 are published as JRC technical reports, part of the series “*Support to the implementation, harmonization and further development of the Eurocodes*”.

We hope that this report will provide a sound and helpful basis for the implementation and use of the second generation EN 1997 “Geotechnical Design” and contribute to training and education of the professionals engaged in geotechnical design, supporting further skills development for individuals' careers and also EU's competitiveness. The report is available to download from the “Eurocodes: Building the future” website (<http://eurocodes.jrc.ec.europa.eu>).

The authors have sought to present useful and consistent information in this report. However, users of the information contained in this report must assess if such information is suitable for their purposes.

Ispra, September 2024

François Augendre, Head of Unit
Georgios Tsionis, Deputy Head of Unit
E. 3 Built Environment Unit
Directorate E – Societal Resilience and Security
Joint Research Centre (JRC)
European Commission

CEN/TC 250/SC 7 Foreword

With the adoption of the 2nd Generation of Eurocode 7 – Geotechnical design - Member States will need to implement new procedures in many different topics, not considered in the first generation of the Code such as the assessment of Representative values, the Ground Model, the use of Reliability methods and the implementation of design in the execution phase. To facilitate the implementation of Eurocode 7, CEN Technical Committee 250/Sub-Committee 7 (TC 250/SC 7) therefore decided to produce a suite of Guidelines, one for each of the most relevant new aspects, to help ensure that the objectives of the code writers are reached in practice.

According to Clause 4 of Eurocode 7, Basis of design, the fundamental step to ensure that the prescribed reliability in geotechnical design is reached, is the development of a representative Geotechnical Design Model being the combination of the Ground Model and the set of design values of relevant geotechnical properties needed for verifications.

The first Guideline of the suite addresses the fundamental process of determining design values of geotechnical properties from derived values, obtained from a variety of activities of the ground investigation. Once the representative values of properties, either characteristic (through statistical evaluation) or nominal (cautious estimate), are determined, design values are obtained by applying the partial factors for a design situation.

The process of assembling the Ground Model is addressed in the second Guideline, where the importance of the progressive upgrading of the Model with an increase of knowledge of the ground within the Zone of Influence of the specific structure is highlighted. Note that the concept of the Zone of Influence in the second generation of Eurocode 7 is sensibly widened after environmental and seismic aspects have become central in planning the Ground Investigation and processing the results.

A significant novel aspect is the use of Eurocode 7 in combination with reliability-based methods. This is likely to lead to a very important evolution of safety assessments in geotechnical design in the coming years. Eurocode 7 like other Eurocodes is fundamentally reliability-based although safety verifications are tied to the application of partial factors. However, in the second generation of Eurocode 7, Clause 4 explicitly states that reliability-based design is only one of the options to verify limit states in geotechnical design. As these methods are not usually addressed in most of the teaching programs, it has been decided to dedicate a specific guideline to reliability-based verification of limit states coherently with the safety concepts of the Eurocodes. Moreover, the objective of this third Guideline is to provide information for code developers to perform reliability calibration of partial factors given the preparation of National Annexes.

It is well known by geotechnical designers that a great contribution to the reliability of a geotechnical construction relies upon its execution and its real performance. This is why Eurocode 7 now specifically dedicates a full clause in EN 1997-1 to the implementation of design during execution and service life. The fourth Guideline presents measures to ensure that the design is correctly implemented in the different construction phases and how to document the activities carried out to this scope on the construction site. After a general description of the suggested rules and methods, the guideline describes good practice for establishing a Supervision Plan, Inspection Plan, Monitoring Plan and Maintenance Plan, how to establish acceptance criteria and limit values and gives contingency measures that might be utilized when an acceptance criterion/limit value is reached. This is implemented for typical geotechnical constructions such as embankments, bored piles, rigid inclusions, and groundwater control.

The upgrading of Eurocode 7 to the methods of modern Geotechnics was a very difficult task that has been achieved thanks to the strong involvement of many dedicated people from all European countries. However, this process cannot be considered concluded with the publication of the new Eurocode documents alone and there will be a long route to implementing Eurocode 7 into the engineering practices of the many countries involved.

The involvement of the many members of the Task Group C1 to C4 of SC7, who prepared these Guidelines is very gratefully acknowledged. These Task Groups have performed tremendous work over almost 4 years to compile knowledge and experience in European geotechnical engineering to draft these Guidelines.

It is therefore strongly believed that for the transition from the 1st to the 2nd Generation of Eurocode 7, these Guidelines will be very helpful in clarifying the new concepts and methods. The Guidelines will also provide didactic background material that could not be presented in the Code.

Giuseppe Scarpelli, Coordinator Task Group C1 to C4

Adriaan van Seters, Chairman

CEN/TC 250/SC 7 “Geotechnical Design”

Acknowledgements

This guideline was produced by a dedicated task group of the Eurocode 7 development subcommittee (CEN/TC 250/SC 7) with the following members:

Patrick Arnold, Marcos Arroyo, Jens Bergan-Haavik, Witold Bogusz, Conrad Boley, Nezam Bozorgzadeh, Nicola Brusa, Ulrich Burbaum, Stefano Collico, Stephane Commend, Nicolai Droniuc, Andra Ebener, Gudmund Reidar Eiksund, Alexandra Ene, Karl Escher, Jose Estaire, Hakan Garin, David Hard, John Harrison, Max Hehenkamp, Robert Heinz, Casimir Katz, Mika Knuuti, Geert Krajima, Kerstin Lesny, Monica Löfman, Antonis Mavritsakis, Rita Migliazza, Lars Olsson, Trevor Orr, Renato Pereira, Philippe Reiffsteck, Pietro Rimoldi, Alfred Roubos, Giuseppe Scarpelli, Timo Schweckendiek, Joao Serra, Heena Sheth, Julia Sorgatz, Johan Spross, Geir Svano, Bram van den Eijnden, Joel Vimpeli, Herbert Walter, Richard Yuan, Moulay Zerhouni, Alexandra Ålenius.

The detailed technical review by Gunilla Franzén, Vice-Chair of CEN/TC 250/SC 7, Geoverkstan, Kungsbacka (SE), and Andrew Bond, Geocentrix, Banstead (UK), is gratefully acknowledged.

Authors (guideline)

Bram van den Eijnden, Delft University of Technology, the Netherlands

Mika Knuuti, University of Tampere, Finland

Kerstin Lesny, University of Siegen, Germany

Monica Löfman, Aalto University, Finland

Antonis Mavritsakis, Deltares, the Netherlands

Alfred Roubos, Port of Rotterdam, the Netherlands

Timo Schweckendiek, Deltares, the Netherlands

Francesca Sciarretta, Joint Research Centre of the European Commission, Ispra, Italy

Authors (examples in Annex B)

Andra Ebener, University of Siegen, Germany

Karl Escher, Royal Institute of Technology, Sweden

Johan Spross, Royal Institute of Technology, Sweden

Stephane Commend, GeoMod SA, Lausanne, Switzerland

Maximilian Hehenkamp, Deltares, the Netherlands

Patrick Arnold, GuD Geotechnik und Dynamik Consult GmbH, Berlin, Germany

Sigrid Wilhelm, GuD Geotechnik und Dynamik Consult GmbH, Berlin, Germany

Alexandra Ene, University of Bucharest, Romania

Pietro Rimoldi, Pietro Rimoldi Consultant, Milano, Italy

Renato Pereira, National Laboratory of Civil Engineering, Lisbon, Portugal

Editor

Timo Schweckendiek

1 Introduction

1.1 Context

Safety in geotechnical design has evolved from experience and judgment using overall factor of safety approaches to design methods in modern codes of practice which apply partial factors to loads and to strength parameters or resistances. Explicit treatment of uncertainties using probabilistic methods has been receiving more attention recently, both in the justification of partial factors in codes of practice as well as in design and assessment. The Eurocodes are fundamentally reliability-based, even though not all safety-relevant aspects are explicitly tied to the underlying reliability concepts as explained in the JRC Technical Report ‘Reliability Background of the Eurocodes’ (European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al., 2024).

EN 1997-1 explicitly states reliability-based methods as one of the options to verify limit states of geotechnical structures alongside the partial factor method, prescriptive rules, testing, and the observational method. The present document serves as a guideline for reliability-based verification of limit states in design and assessment of geotechnical structures within the safety and reliability concepts of EN 1990-1 and EN 1997-1.

1.2 Objectives

The overarching objective of this guideline is:

To provide guidance for (full probabilistic) reliability-based verification of limit states for geotechnical structures within the safety concepts of the Eurocodes

To this end, the following sub-objectives will be addressed:

1. To explain the reliability-based safety concept of the Eurocodes in relation to geotechnical design and assessment according to EN 1990 (all parts) and EN 1997-1.
2. To provide recommendations for target reliability values.
3. To provide guidance for modelling the uncertainties involved with geotechnical analysis, including typical statistics for ground properties and model uncertainties.
4. To describe methods and tools for geotechnical reliability analysis.
5. To describe how to derive partial factors using reliability theory.
6. To formulate best-practices and provide examples.

1.3 Target audience

The target audience of this guideline is two-fold:

1. Practitioners applying reliability-based methods in geotechnical design and assessment, in particular related to the Eurocodes.
2. Code developers using reliability analysis to substantiate nationally determined parameters (NDP) in the implementation of Eurocode 7 in member states.

The implication of addressing this target audience is that the guideline focuses on practical guidance rather than extensive theoretical backgrounds. Users of the guideline are assumed to have basic knowledge of probability and reliability concepts; though some background is provided, since the topic will be rather new to many practitioners.

1.4 Scope

The present guideline document addresses all aspects involved with enabling practitioners to perform reliability-based (full probabilistic) verification of limit states for the design of new geotechnical structures according to EN 1990 and EN 1997-1, as well as the assessment of existing ones. Furthermore, the contents align with the reliability backgrounds of the Eurocodes as stated in European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al. (2024).

When referring to reliability, it is important to realize that the reliability management in EN 1997 is a broader concept than the reliability (verification) elaborated on in this guideline. Reliability management also includes quality control, minimum required site investigation and requirements to personnel involved with the design, among others. In this guideline, the verification is restricted to assessing the probability of exceeding limit states based on the uncertainties in the ground model, material properties, actions and calculation models

In addition to guidance for applying (full-probabilistic) reliability-based methods, this document contains information relevant for code developers to perform reliability calibration of partial factors, primarily targeting the national annexes to EN 1990-1 and EN 1997-1. The presented approach to deriving partial factors may also be relevant in situations when project-specific partial factors need to be derived.

Based on the objectives and target audience as stated above, the contents of this guideline aim at practical application. Generally speaking, textbook material has been avoided, except when deemed necessary for coherence, readability and providing essential background knowledge. The approaches in this document have a proven track record in practical application as much as possible. Ongoing developments in academia to improve the state-of-the-art of geotechnical reliability analysis may be mentioned in references but are not extensively discussed. They may be included in future updates of this guideline.

1.5 Application of this guideline

The application of reliability-based methods in EN 1997-1 to verify the safety with respect to the exceedance of limit states is not likely to replace the conventional methods, nor does this guideline document intend to do so. Using reliability-based methods is usually more labor-intensive than conventional methods (e.g. the partial factor method) and requires additional experience with probabilistic concepts, on top of the geotechnical expertise required to perform any geotechnical design or assessment. There are, however, situations in which reliability-based methods can be preferable to conventional methods. The list below contains several examples:

- design of special structures for which the standard partial factors may not be appropriate;
- design or assessment of structures with explicit reliability requirements (e.g. acceptable probability of failure) such as storage of hazardous substances;
- assessment of existing structures for which the standard partial factors are not intended;

- assessment of existing structures using performance information.

There are also limitations to approaching designs and assessments with reliability-based methods. For example, in situations in which:

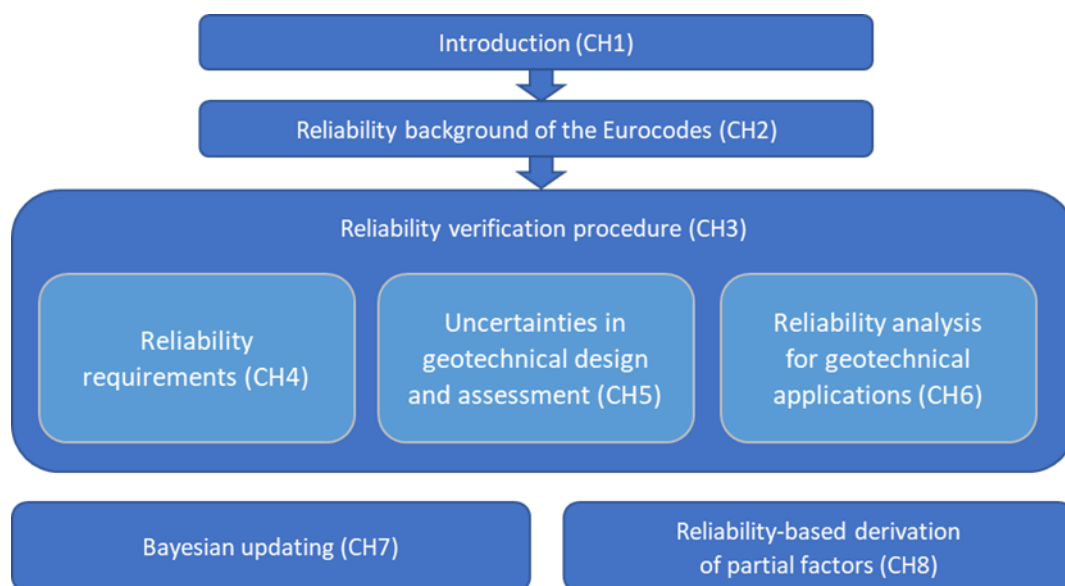
- the geotechnical design models are very complex and computationally expensive;
- problems are so simple that it is not worth the effort; etc.

The contents of this guideline are recommendations, not requirements. Given the relative novelty of the reliability-based approach, it is paramount that the application is performed by and/or reviewed by professionals with appropriate expertise.

1.6 Outline

This guideline (**Figure 1**) starts by summarizing relevant elements of the reliability background of the Eurocodes in chapter 2, followed by outlining the reliability verification procedure in chapter 3. Chapter 4 gives considerations for choosing reliability targets and their application. Subsequently, chapter 5 describes the characterization and quantification of uncertainties, which serve as input for the reliability analysis as addressed in chapter 6. Chapters 7 and 8 are about the special topics of Bayesian updating and reliability-based derivation of partial factors, respectively.

Figure 1. Visual outline of the guideline



Source: Authors' own work

2 Reliability background of the Eurocodes

This chapter provides the reliability backgrounds of the Eurocodes, focusing on elements relevant for design and assessment of geotechnical structures with EN 1990 and EN 1997-1, drawing largely from European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al. (2024).

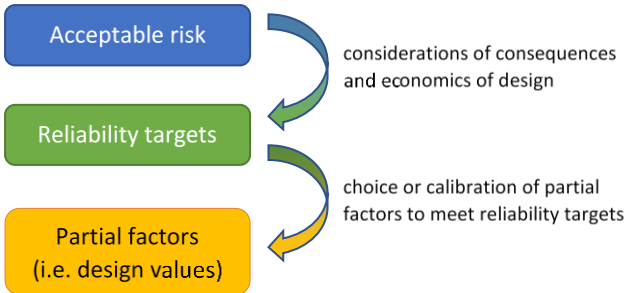
2.1 Reliability aspects in the Eurocodes

EN 1990 defines structural reliability as the ability of a structure to fulfil the specified requirements during the service life for which it has been designed; the notion covers safety, serviceability and durability. According to this definition reliability is the concern of all parts of the whole suite of Eurocodes. In a more narrow sense of the word, reliability refers to the way uncertainties are dealt with in design and assessment by the use of partial factors or more advanced probabilistic methods. This document deals primarily with this more narrow definition.

Reliability analysis is part of the science and practice of engineering today, not only with respect to the safety of structures, but also for questions of serviceability and other requirements of technical systems that might be impacted by uncertainties. Treatment of uncertainties may take place on several levels, which are often referred to as (1) the risk-based approach, (2) the reliability-based approach and (3) the semi-probabilistic or partial factor-based approach (**Figure 2**, see ISO 2394). Appropriate calibration procedures should guarantee that the various levels give consistent results. In addition to this relative calibration of the various design approaches there is also a need to compare the results to design and assessment methods used successfully in the past.

Geotechnical design with the Eurocodes employs the limit state design approach. The design (or assessment) verifications address the exceedance of (failure) limit states, while the design values are chosen such that an appropriate level of reliability (i.e. probability of not exceeding the limit state) is achieved. The present document focuses on full-probabilistic reliability-based verification of limit states, and briefly touches upon how semi-probabilistic and full-probabilistic verifications are related.

Figure 2. Relationship between acceptable risk, reliability targets and partial factors (i.e. the semi-probabilistic approach)



Source: Authors' own work

2.2 Uncertainties and probability

Reliability concepts are founded in probability theory. To properly apply the approaches and methods described in this document, it is therefore essential how we define and interpret uncertainties and probability, along with the notions of variability and frequency.

2.2.1 Sources and types of uncertainties

Design and assessment of structures involves uncertainties of different kinds. The most obvious uncertainties are related to the natural variability or randomness of actions, material properties and geometric data. On top of that, in geotechnical engineering we have to deal with the uncertainty in ground properties resulting from spatial variability and limited site investigation coverage. When establishing the statistical properties of relevant variables, we often deal with unclear sets of subpopulations and substantial shortage of data. This leads to the important class of statistical uncertainties.

In addition, our theories and models about environmental and structural behavior are approximations of reality and make predictions with some higher or lower degree of accuracy. The related uncertainties are referred to as model uncertainties. Similarly, we have to deal with uncertainties involved in measurements and observations. Finally, in practice we also have to deal with uncertainties resulting from human errors of several kind as well as unforeseen (or even unforeseeable) events.

2.2.2 Probability interpretation

In reliability literature uncertainties are divided into aleatory uncertainties resulting from (natural) variability and epistemic uncertainties resulting from incomplete knowledge. The essential difference is that epistemic uncertainties can be reduced by acquiring additional data while aleatory cannot.

In the *classical (frequentist) school* of statisticians a clear distinction in the treatment is made. The basic philosophy is that the quantification of epistemic uncertainties in principle requires subjective estimates and should be kept carefully separated from the objective part.

In the alternative *Bayesian school*, the notion of probability is understood as a “degree of belief” which enables equal treatment of objective and subjective uncertainties. In applications in structural engineering the Bayesian approach has been adopted as the most appropriate one from the beginning; in geotechnical engineering a frequentist interpretation of probabilities is problematic, and a subjective interpretation makes much more sense (Baecher & Christian, 2003).

A pronounced difference in application is visible in the statistical evaluation procedures for design by testing. For example, in the frequentist formulas for the 5% characteristic values we first find the best estimate for the 5% fractile (natural uncertainty) and an additional 90% confidence interval (statistical uncertainty). In the Bayesian approach (see EN 1990 Annex D) the 5% fractile is found including the statistical uncertainty and no confidence interval is given.

2.3 Principles of structural reliability

As already mentioned and illustrated in **Figure 2**, risk, reliability and partial factors are related. The following paragraphs allude to the very principles underlying each concept, and how they relate to each other.

2.3.1 Risk-based approach

The full risk-based approach entails an economic optimization within safety constraints (e.g. human safety). In its most simplified form, the optimization is formulated as the minimization of the sum of the direct costs of the design decision and the risk due to failure. That is, leaving out discounting and other complicating factors like maintenance, we want to minimize the total cost expectation C_{tot} :

$$\min C_{tot} = C_{con} + P_f C_f$$

Equation 1.

with the constraint

$$P_f < P_{f,t}$$

Equation 2.

where C_{tot} are the expected total costs, C_{con} is the cost of the (design) decision, C_f are the costs of the consequences of unsatisfactory performance (failure) and P_f is the probability of unsatisfactory behaviour during the design service life (or assessment period). The constraint ($P_f < P_{f,t}$) is based on ethical and societal aspects of human safety where relevant. The risk-based approach is generally used for individual structures, not for a portfolio of structures.

The approach implicitly assumes that the structural behaviour can be subdivided into a satisfactory performance and an unsatisfactory performance or failure. Note that the term “failure” in reliability theory is not restricted to collapse but may refer to any form of unintended behaviour (e.g. excessive deformations).

2.3.2 Reliability-based approach

In the reliability-based approach the optimization reduces to minimizing the construction cost:

$$\min C_{tot} = C_{con}$$

Equation 3.

with the constraint

$$P_f < P_{f,t}$$

Equation 4.

The target reliability $P_{f,t}$ now also includes (next to immaterial constraints like human safety) also the economic considerations that were left to an explicit optimization procedure in the risk-based approach.

2.3.3 Semi-probabilistic approach

For the semi-probabilistic approach, used in daily practice, the requirements in **Equation 3** and **Equation 4** are transferred to:

$$\min C_{tot} = C_{con}$$

Equation 5.

with the constraint

$$g(\mathbf{X}_d) > 0$$

Equation 6.

where $g()$ is the limit state function. \mathbf{X}_d represents the so-called design values of the random variables. In terms of reliability theory, the design values represent the values of \mathbf{X} with the highest probability density of failure (after a transformation to standard normal space). If X has a normal distribution, the value \mathbf{X}_d can be found from:

$$X_d = \mu_x + \alpha_x \beta_t \sigma_x$$

Equation 7.

where μ_x and σ_x are the mean and standard deviation of X respectively, α_x is the (FORM) importance factor of the variable X , β_t is the target reliability index corresponding to the target failure probability $P_{f,t}$. For other than normal distributions the value of X_d may be found from $F(X_d) = \Phi(\alpha_x \beta_t)$ with $F()$ being the cumulative distribution function of the variable X . As the result must be conservative for a large group of structures, the selected α_x -values will usually be conservative (uneconomic) in individual cases.

In the Eurocodes, the design values are derived by multiplying (or dividing) the characteristic (or in the 2nd Generation Eurocodes 'representative') values \mathbf{X}_k with the respective partial and combination factors.

$$X_d = X_k / \gamma \text{ for resistance}$$

Equation 8.

$$X_d = \gamma X_k \text{ or } \gamma \Psi X_k \text{ for actions}$$

Equation 9.

These partial factors are ideally calibrated in such a way that the design or assessment decision for a large class of structures fulfils the reliability target.

2.3.4 Observed failure frequency, notion of hidden safety

The question is often raised whether or not the theoretical probability of failure corresponds or should correspond to the frequency of structural failures in practice. In this discussion the notion of hidden safety plays an important role, among other factors such as gross/human errors or negligence during design, execution or operation. Gross/human errors are supposed to be covered by quality control and operational requirements (e.g. execution standards); they are not to be covered by engineering models and physical and model uncertainties related to uncertain parameters.

Generally speaking, assumptions and choices in design and the related engineering modelling tend to be conservative, or safe-sided. The list below contains several examples of hidden safety in construction design:

- Resistance modelling: We may opt for elastic analysis instead of plastic analysis. Model choices are often based on the wish to make simplifications (ease of use). The simplified models are generally safe-sided and even robust.
- Load modelling: The load patterns assumed for modelling the spatial distribution of imposed and climatic loads are often chosen with very safe assumptions, or the seismic action is described through a uniform hazard spectrum, rather than a less conservative conditional spectrum.

- Design type or structural system: In many design situations, a structure is designed according to its critical section or member, while the rest is oversized. Also, quite often not the ULS is governing the structural dimensions, but the SLS. In these cases, the safety margin can be considerable. For example, anchors of retaining walls are designed such that neighbouring anchors could cope with the failure of an in-between anchor.
- Structural behaviour modelling: The structural behaviour is simplified using conservative assumptions. For example, a hinge is assumed where in fact an encasement at least partially exists. Such system redundancies are often not explicitly exploited.
- Material strength modelling: For example, the usual steel design is for yielding, while strain hardening and the margin to the ultimate strength typically neglected.
- Practical issues rounding off structural dimensions to the safe side and producers aiming at better quality then required to avoid a negative acceptance test.

The list above is not exhaustive but meant to provide examples of types of hidden safety. It should be clear that the estimation of actual, observable failure rates in reality using engineering models is difficult, given the unavoidable bias introduced by hidden safety elements, even though some types can be covered by taking model bias explicitly into account.

As consequence of hidden safety elements in design and assessment, the observed failure rates (e.g. ULS exceedance) appear in general to be small compared to the estimations with engineering models (excluding gross/human errors).

2.3.5 System reliability

The analysis of reliability and risk should in principle be applied to the total structural system, including its interactions with the environment. In practice however the verification of the reliability is often performed on the level of single members (beams, columns, connections), single modes (shear, bending) or even on the level of a single cross section or individual point. The verification rules in the Eurocodes are in most cases formulated at the member level and also the target reliability is used for the individual failure mode at member level (see detailed discussion in European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al., 2024).

2.3.6 Time effects

Probabilities of failure are always defined for a reference period (e.g. annual or 50 years). The joint influence of (largely) time-varying loads and (largely) time-invariant resistance properties leads to time-dependency of the failure probability in the period of time under consideration. In addition, a structure may lose resistance due to degradation processes like creep or corrosion.

If we include time, the random variables as well as the limit state function itself may become time-dependent. Subdividing the total period t_L in subintervals (for instance periods of one year) we may write:

$$P_f(t_L) = P_1 + P_2 + P_3 + \dots + P_n$$

Equation 10.

where P_i is the probability of failure exactly in year i , which may be written in terms of limit state functions as the so-called unconditional annual failure probabilities:

$$P_1 = P(g_1 < 0)$$

$$P_2 = P(g_2 < 0 \text{ and } g_1 > 0)$$

$$P_3 = P(g_3 < 0 \text{ and } g_2 > 0 \text{ and } g_1 > 0)$$

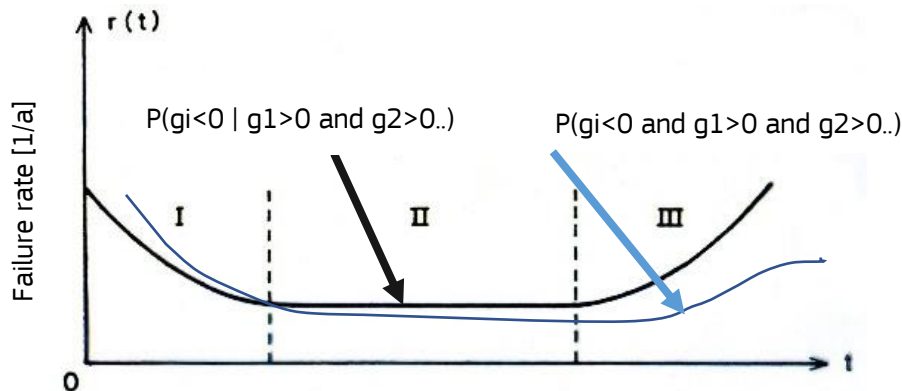
etc.

Equation 11.

where g_i is the (minimum) value of the limit state function in year i . The logic is that failure in the second year is applicable only if the structure has survived first year.

The time effects for typical structures are illustrated by the so-called 'bath tub curve' as depicted in **Figure 3**.

Figure 3. Bathtub curve (black line) showing the conditional failure probability as a function of time, with zones of (I) increasing reliability (dominant time independent uncertainties like resistance and self-weight), (II) constant reliability (dominating time dependent uncertainties, usually loads) and (III) decreasing reliability (dominating deterioration). The blue line shows the unconditional failure probabilities.



Source: European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al. (2024)

2.4 Target reliability

2.4.1 Safety and economy

For the reliability-based approach and the partial factor approach, the target reliability is a prescribed requirement. Usually, the target reliability level is proposed by the engineering community and endorsed by the National Authorities, who finally are responsible. The common starting point is that a structure should be safe as well as economic during the design or remaining service time. In general, the accepted reliability level with respect to ULS-failure depends on:

1. the consequences of failure (which may depend on the failure mode);
3. the costs related to additional reliability;
4. the anticipated service life;
5. the variability and uncertainty in resistance and load, including the likelihood of mistakes (as far as not covered by quality control).

In most codes, only the first factor is used for explicit reliability differentiation in terms of different levels of target reliability levels. The consequences of failure, however, are not elaborated in detail,

but only a rough classification of structural types (dwellings, high-rise buildings, grand stands, etc) is given.

2.4.2 Reliability differentiation

Reliability differentiation depends on the consequences of failure (related to economy losses, environmental effects and losses of human lives). For various reasons it does not seem appropriate to distinguish in reliability requirements between various materials. All structures or structural elements within a certain class of use and failure consequences should have the same reliability.

A special application of the reliability differentiation concerns assessment of existing structures. Strengthening an existing structure often requires more effort (i.e. costs) than building new structures. Also, the (remaining) design service life might be (much) shorter than 50 years. Hence, the reliability targets for existing structures are reasonably lower than for new ones.

2.4.3 Annual versus lifetime targets

Reliability targets always have to be specified for a reference period. Annual values are commonly used when dealing with human safety. The lifetime option is convenient if only economic optimization is considered in the derivation. In the Eurocodes, the lifetime option is connected directly to a design service life of 50 years. Rules for structures for a different design service life (e.g. temporary structures, large civil engineering works) are not specified.

When annual reliability targets are formulated, the maximum annual value during the design or remaining service life is decisive.

If failure events in various years are independent and have small, similar probabilities, the relation between an annual (P_1) and a 50 year-target is approximately:

$$P_f(50) \approx 50 \cdot P_1$$

Equation 12.

The formula still holds for moderate correlation, but if there is a large degree of correlation between years (e.g. due to dominating permanent loads or other time-invariant variables) more sophisticated analysis is required to determine the time-development of reliability. Notice that for the other extreme of full correlation in time (i.e. all variables are time-invariant), $P_f(50) = P_1$.

2.4.4 Minimum versus average reliability

In principle, the target reliability may be formulated for a portfolio of structures (say all bridges). In that case it may be questioned whether the target refers to the average reliability of the portfolio or to the minimum reliability (average or maximum failure probability respectively). If aiming at an economic optimum the average values make sense; in order to fulfil requirements related to human safety the minimum reliability level seems to be the most appropriate.

2.4.5 Member versus system

In most practical applications target reliabilities are used for designing individual members (including connections) for separate failure modes. However, consequences are usually related to system failures and the system failure probability may be the result of contribution of a set of

system failure modes. For a part this has to be considered as the state of the art. However, in some cases we encounter exceptions are possible as in the design of pile groups.

2.4.6 Brittle versus ductile failure

Brittle failures on the member level usually have more serious consequences for the structural system integrity than ductile failure. This might be a reason to use a higher consequence class and higher reliability targets for brittle failure modes. An additional aspect is the degree of warning: ductile failure modes may give some kind of pre-warning and thus time to evacuate a site or for emergency interventions.

2.5 Assessment of existing structures

There is an increasing demand for assessing the reliability of existing structures when these face new demands or extension of the remaining service life needs to be substantiated. In chapter 4 of European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al. (2024), some guidance for the assessment of existing structures is provided, addressing target reliabilities, updating information (e.g. performance observations) and verification methods. Reliability methods are mentioned as the most accurate way of assessment, along with risk-based decision procedures.

3 Reliability verification procedure

3.1 General

Reliability verification in general is based on the comparison of the calculated reliability of the system with a target value. The reliability is expressed either by the reliability index β or, equivalently, by the probability of failure. Hence, the basic requirement for reliability verification is to fulfil the following inequality:

$$\beta \geq \beta_T$$

Equation 13.

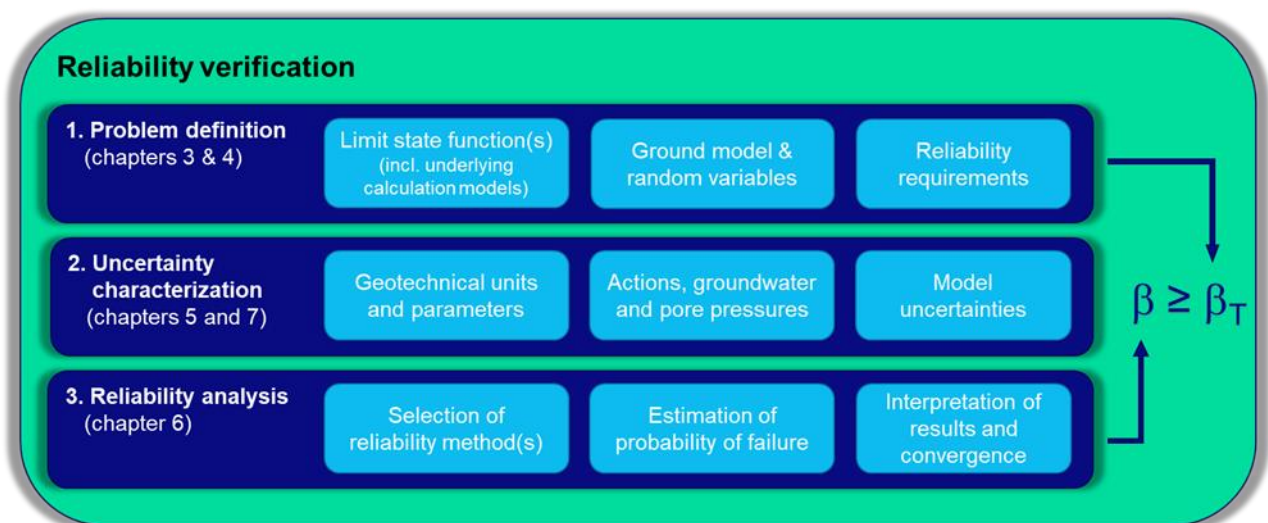
or equivalently, $P_f \leq P_{f,T}$. Notice that the probability of failure and the related reliability index will always pertain to a specific reference period and failure definition (e.g. component versus (sub)system failure).

This chapter outlines the general reliability verification procedure that can be broadly subdivided into the following steps:

- Problem definition: limit state function, ground model, target reliability (β_T or P_T).
- Uncertainty characterization: geotechnical units and parameters; loads, groundwater and pore pressures; model uncertainties.
- Reliability analysis: selection of appropriate reliability method, calculation of system reliability, interpretation of results.

Figure 4 illustrates the sequence of these steps. The remainder of this chapter provides a brief description of these main steps; detailed treatment of the components of the reliability assessment follows in the remaining chapters as indicated.

Figure 4. Overview of the workflow for the reliability verification procedure (with indications of the chapter containing in-depth treatment)



Source: Authors' own work

3.2 Problem definition

The problem definition step entails:

— **Limit state functions and calculation model:**

Definition of the limit state function is ideally as much as possible in accordance with the formulation of the verification with partial factors in EN 1997-1 and -3.

Evaluating the limit state function requires calculation models to evaluate the resistance or load effects. The complexity of the calculation models to evaluate the resistance or load effects affects the performance of the reliability analysis.

— **Ground model and random variables:**

The ground model needs to be established, including the related uncertainties, just like with the partial factor method.

The limit state function and the underlying calculation model contain uncertain variables. Ideally all uncertain variables are modeled probabilistically. However, this is practically often not feasible nor necessary. The choice to model variables as random or deterministic can be based on engineering judgement or sensitivity analysis.

— **Reliability requirements:**

The reliability requirements entail the target value of the reliability index, with a given reference period and pertaining typically to a consequence class (see Chapter 4).

3.3 Uncertainty characterization

The uncertainty characterization entails expressing all relevant and significant uncertainties in probabilistic terms, e.g. by probability distributions, scenarios or point estimates, covering the following categories:

— **Geotechnical units and ground properties:**

The uncertainties in the presence and geometry of the geotechnical units in the ground model need to be quantified or estimated, accounting for the various sources of uncertainty (i.e. spatial variability, measurement and transformation error, statistical uncertainty).

— **Actions, groundwater and pore pressures:**

The uncertainty and variability in actions, groundwater and pore pressures needs to be quantified if relevant and significant for the reliability problem.

— **Model uncertainties:**

Calculation models involved with the evaluation of the limit state on either the resistance or the load effect side involve model errors which need to be accounted for.

Chapter 5 is entirely devoted to uncertainty characterization.

3.4 Reliability analysis

Having completed the steps in the problem definition and the uncertainty characterization, all inputs for the reliability analysis are ready. Chapter 6 provides detailed recommendations for selecting a reliability method, formulating the reliability problem properly and interpreting the results of the reliability analysis, including illustrative examples.

4 Reliability requirements

Reliability verification of geotechnical structures compares the assessed reliability with the reliability target, i.e. $\beta \geq \beta_T$ should hold (see 3.1). This chapter provides guidance on the selection and the use of reliability targets.

Notice that EN 1997-1 contains several elements of reliability management, with reliability interpreted in the wider sense of safety. The geotechnical category (GC) is determined through a combination of the geotechnical complexity class (GCC) and the consequence class (CC). The reliability targets in this chapter refer to the narrower definition of reliability (i.e. probability of failure), and therefore have a direct relation with the consequence class. The GCC and GC serve other purposes such as defining the required amount of ground investigation or quality management in execution.

Box 1. Background: Target failure probability according to ISO 2394:2015

According to ISO 2394:2015 the target probabilities of failure should be selected taking into account the consequence and the nature of failure, the economic losses, the societal inconvenience, effects to the environment sustainable use of natural resources, and the amount of expense and effort required to reduce the probability of failure. If there is no risk of loss of human lives associated with structural failures, the target failure probabilities can be selected solely on the basis of an economic optimization. If structural failures are associated with risk of loss of human lives, the marginal lifesaving costs principle applies and is recommended. In all cases, the acceptable failure probabilities should be calibrated against well-established cases that are known from past experience to have adequate reliability.

4.1 Target values for reliability index (ULS)

4.1.1 Target values in EN 1990-1

EN 1990-1 Annex C provides target values for the reliability index for a reference period of 50 years (**Table 1**). The corresponding note indicates that different target reliability values for different reference periods can be set by the National Annexes (i.e. the table is NDP).

Table 1. Target reliability values for reliability index β (ultimate limit state) for different consequence classes according to EN 1990-1 Table C.3.2 (NDP). Corresponding note: Table C.3.2 gives target values β for the 50 years reference period assumed within this standard. Different target reliability values β for different reference periods can be set by the National Annex for use in a country.

Consequence class	1-year reference period β	50-year reference period	
		β	$P_{f,50}$
CC3	5.2	4.3	$\sim 10^{-5}$
CC2	4.7	3.8	$\sim 10^{-4}$
CC1	4.2	3.3	$\sim 10^{-3}$

Source: EN 1990-1

European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al. (2024) clarifies that these target values were originally derived for buildings and bridges, while the reliability calibration study in the same report suggests that these targets may also be

applicable to geotechnical structures (as shown for spread foundations and for piled foundations). For the specification in National Annexes and related calibration studies it may be sensible to differentiate between types of (geotechnical) structures and/or different reference periods. A reference period of 1 year seems more appropriate for the assessment of existing structures, or for the design of temporary structures.

Deltares (2024) contains additional material supporting that the central target value of $\beta_T = 3.8$ in **Table 1** for a 50-year reference period is also applicable to geotechnical structures, and provides information for deriving corresponding annual target values.

4.1.2 Consequence classes

The consequences of failure of a structure or a structural member are classified into one of the five consequence classes specified in EN 1990 Table 4.1 (NDP), based on a qualification of the consequences of failure (see **Table 2**). National Annexes can give different qualifications for use in a country.

Table 2. Qualification of consequence classes according to EN 1990-1 Table 4.1 (NDP)

Consequence class	Indicative qualification of consequences	
	Loss of human life or personal injury ¹	Economic, social or environmental consequences ¹
CC4 – Highest	Extreme	Huge
CC3 – Higher	High	Very great
CC2 – Normal	Medium	Considerable
CC1 – Lower	Low	Small
CC0 - Lowest	Very Low	Insignificant

(¹) The consequence class is chosen based on the more severe of these two columns

Source: EN 1990-1

The classification into a consequence class may also be influenced by the type of failure mode, e.g. ductile versus brittle.

Notice that consequence classes CC4 and CC0 are not covered by the target reliability values in EN 1990-1. CC0 would likely not be amenable for reliability-based methods; CC4 is rather uncommon and would require a risk-informed approach.

4.1.3 Risk-informed approach

When appropriate reliability targets are not readily available, the risk-informed approach may be applied. Examples would be structures with very high consequences of failure (i.e. CC4) or structures which cannot be assigned to the consequence classes for other reasons, for example when working outside of the Eurocode framework.

According to EN 1990-1, reliability requirements can be formulated in terms of minimum reliability requirements and/or service life cost optimal target reliability requirements:

- the minimum reliability requirements depend on the societal capacity and preferences to invest into life safety (minimum reliability requirements are normally compulsory);
- target reliability requirements depend on the expected failure consequences and on all costs associated with the design, construction costs, operation, inspection, maintenance and renewal of structures over the time period for which they are needed.

The most common risk acceptance criteria for civil structures in order to address these requirements are:

- individual risk (to life);
- economically optimal risk (i.e. cost-benefit considerations);
- group risk (considering incidents with larger numbers of fatalities);
- Life Quality Index (LQI).

More information on establishing reliability targets using risk acceptance criteria can be found in ISO 2394:2015, in which Annex G also specifically elaborates on the LQI and the underlying marginal lifesaving cost principle (MLSC).

4.1.4 Other sources

When working outside the Eurocode framework, other sources may be applicable or useful to define or choose the reliability target for the design or assessment of (geotechnical) structures, such as other codes of practice or literature recommendations. For some (geotechnical) structures, specific reliability targets have been formulated in the pertinent codes of practice, see **Table 3** for examples.

Table 3. Examples of codes and standards specifying or recommending reliability targets

Code or standard	Reference	Remarks
ENW, flood defences, Netherlands	ENW (2017)	Annual target failure probabilities of flood defense segments (systems) ranging from 10^{-2} to 10^{-6} ($\beta = 2.3$ to 4.8)
USBR, dams, USA	FERC (2015)	Annual targets based on FN-Curves including individual risk and group risks, with target failure probabilities ranging from 10^{-4} to 10^{-6} ($\beta = 3.7$ to 4.8).
ISO 2394, General	ISO (2015)	Annual target reliabilities for structures of 4.2 to 4.7 depending on consequence class.
Probabilistic Model Code, General	JCSS (2001)	Annual target reliabilities for structures of 4.2 to 4.7 depending on consequence class (for small relative cost of safety measures).
USACE, Geotechnical, USA	USACE (1999)	Annual geotechnical target reliabilities.
ASCE, structures, USA	ASCE (2010)	Lifetime target reliabilities for structures of 2.5 to 4.5 depending on consequence class.
OCDI, maritime, Japan	OCDI (2009)	Lifetime target reliabilities for marine structures of 2.19 to 3.65 depending on consequence class.

Code or standard	Reference	Remarks
CUR 166, sheet piles, NL	CUR 166 (2012)	Lifetime target reliabilities sheet piles of 2.5 – 4.2.
CUR 211, quay walls, NL	CUR 211 (2013)	Lifetime target reliabilities quay walls of 3.3 -4.3.

Source: Authors' own work

4.2 Member versus system reliability

The verification rules in the Eurocodes mostly formulated at the level of individual failure modes of structural members (or components). Likewise, the target reliability is typically used for the individual failure mode at member level (European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al., 2024). The relation between the structural member or local failure and failure of the total structural system is typically not considered.

Following the same reasoning, also the reliability of geotechnical structures can be verified for individual limit states (or failure modes) at member level, for example:

- exceedance of bearing capacity of a spread foundation;
- embankment slope instability;
- loss of equilibrium of a retaining structure;
- exceedance of the yield stress in a sheet pile wall; etc.

In situations with significant interaction between structural components (and the ground), it can make more sense to consider failure modes of structural (sub-)systems or a combination of failure modes for the same structural member in the reliability verification. Examples are:

- exceedance of bearing capacity of a pile group (load re-distribution);
- failure of a sheet pile wall by buckling or yielding (series system);
- deep excavations analysed using numerical methods (several failure modes are modelled and checked simultaneously); etc.

Considering system effects generally comes with increased complexity of the reliability analysis; at the same time, the reliability assessment becomes more accurate by accounting for the favourable and/or unfavourable effects of systems behaviour. Considerations of structural (sub-)systems are also mentioned in ISO 2394:2015.

4.3 Reliability targets for the resistance or load only

In most geotechnical design problems, the uncertainty in the resistance dominates the reliability estimates, while the influence of (variable) loads is relatively low (e.g. foundation pile verification). It can then be practical and sensible to consider only the resistance side of the problem probabilistically, while working with design values for the loads.

Similarly, there may be situations in which the uncertainty in the (effects of) actions dominates the reliability estimates and where we would consider only the load (effect) side of the problem

probabilistically for practical reasons, while working with design values for the resistance (e.g. retaining structures).

The approach is based on the premise that (FORM) design point values (see 6.2.4) are the optimal design values for semi-probabilistic design. The acceptable probability of (not) exceeding the design resistance R_d or design value of load S_d is given by:

$$P(R \leq R_d) = \Phi(-\alpha_R \cdot \beta_T)$$

Equation 14.

$$P(S \geq S_d) = \Phi(-\alpha_S \cdot \beta_T)$$

Equation 15.

In which α_R is the influence coefficient (FORM) of the resistance (i.e. the effect of all components of uncertainty on the resistance side), α_S is the influence coefficient (FORM) of the load (i.e. the effect of all components of uncertainty on the load side); and $\Phi(\cdot)$ is the standard normal cumulative density function (CDF).

The values of α_R and α_S are ideally based on experience from reliability assessments (representative FORM calculations of comparable structures) or alternatively on literature recommendations. ISO 2394:2015 recommends the values in **Table 4**.

Table 4. Recommended influence coefficients (α) from ISO 2394:2015 (Table E.3)¹

X_i	α_i
Dominating resistance parameter	0.8
Other resistance parameters	$0.4 \times 0.8 = 0.32$
Dominating load parameter	-0.7
Other load parameters	$-0.4 \times 0.7 = 0.28$
USBR, dams, USA	FERC (2015)

(¹) The principle of standardised α -values was already present in ISO 2394:1998, where the same α -values as in Table E.3 were proposed.

Source: ISO 2394:2015

Box 2. Example 4.1: Pile resistance

Suppose we are interested in assessing the reliability of a pile foundation, and we are able to characterize the uncertainties in the pile resistance (i.e. related to the soil conditions and the resistance model), while we do not have proper information to formulate a probabilistic load model based on the information available from the structural engineers. In this case, we can still assess the probability of the actual resistance R being lower than the design value R_d (or S_d , if the unity check is less than 1) by considering **Equation 14**.

For a CC2 structure ($\beta_T=3.8$), considering the pile resistance to be the dominant resistance parameter ($\alpha_R=0.8$), we obtain an acceptable probability of not exceeding the design value of $P(R \leq R_d) = \Phi(-\alpha_R \cdot \beta_T) = \Phi(-0.8 \cdot 3.8) = \Phi(-3.04) = 1.2 \cdot 10^{-3}$.

The main advantage of this approach is that it does not require probabilistic load modelling, for which data are often harder to attain for the geotechnical engineer. The disadvantage is that the

simplification comes at the cost of accuracy, which mostly depends on the suitability of the chosen influence coefficient α_R for the design or assessment problem at hand. The recommended values in ISO 2394:2015 are considered conservative for most cases, but in some geotechnical problems higher values may need to be applied. In case of doubt, $\alpha_R = 1$ or $\alpha_S = 1$, respectively, can be applied as conservative upper bound.

Likewise, the procedure can be applied by only considering the uncertainties in the (effects of) actions in comparison to the material resistance. An example would be the stresses in steel sheet piles that are mostly influenced by the uncertainties in the soil-structure interaction, in which case the probability of exceedance of the (design) yield stress of the steel can be assessed.

4.4 Annual versus lifetime reference period

The target values for the reliability index β in EN 1990-1 are based on a reference period of 50 years (see 4.1.1), whereas the annual values are derived (assuming independence between years). It can be advantageous or necessary to work with different reference periods, for example 1 year. This is typically the case with the assessment of existing structures.

Box 3. Reference period

The reference period is the period of time that is used as a basis for statistically assessing degradation/deterioration of materials and extreme realizations of variable actions and possibly for accidental actions.

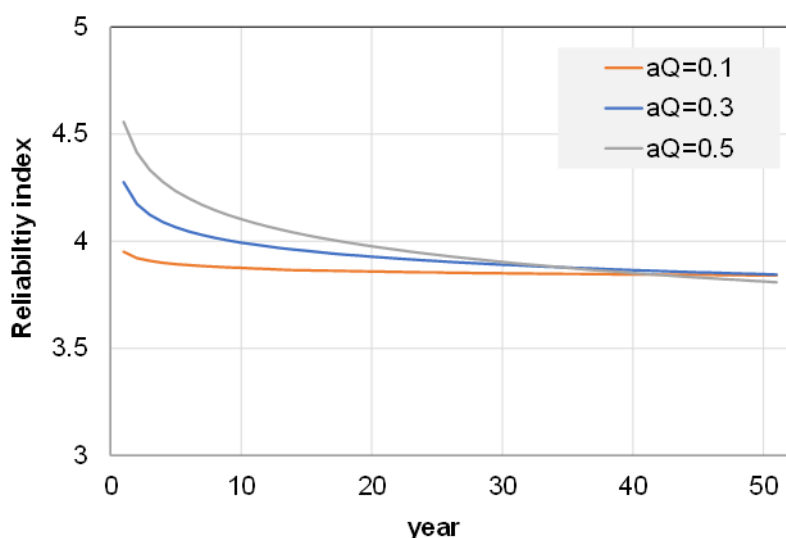
For a failure probability, a reference period of 50 years means that the probability refers to at least one failure of the structure (and the member and limit state under consideration) in 50 years. An annual reference period means that the probability refers to at least one failure in 1 year.

The reference period must not be confused with the design service life of a structure, which is often also 50 years. A structure with a design service life of 50 years can be assessed using probabilities with an annual reference period. In the latter case each year within the considered design lifetime needs to comply with the annual reliability targets (see 2.3.6).

The correspondence of annual reliability targets with the 50-year targets in EN 1990-1 depends on the correlation of failure events between years, which in turn depends on the influence of time-invariant versus time-variable random variables on the design or assessment problem. Most ground-related uncertainty, permanent loads and model uncertainty can be considered (pre-dominantly) time-invariant, while variable loads are naturally time-variant.

The analyses in Deltares (2024) with a generalized limit state function suggest that the annual reliability indices corresponding to a 50-year target of $\beta_T = 3.8$ for CC2 range between 3.9 and 4.5, depending on the contribution of the variable load (a_D); see **Figure 5**.

Figure 5. Reliability index as a function of the considered reference period



Source: Deltares (2024)

For practical use, it seems sensible to keep the reliability differentiation of plus/minus 0.5 β for CC1 and CC3, overall resulting in the following recommended values for a reference period of 1 year in **Table 5**.

Table 5. Recommended reliability target values for geotechnical structures, differentiating between situation with low, moderate or high influence of the load (i.e. time-dependent variables)

Consequence class	β (annual)			β (50 years)
	load influence			
	low	moderate	high	
CC3	4.4	4.7	5.0	4.3
CC2	3.9	4.2	4.5	3.8
CC1	3.4	3.7	4.0	3.3

Source: Authors' own work

4.5 Serviceability limit states (SLS)

For serviceability limit states (SLS) there are no formal minimum reliability requirements, since risk to life is not applicable for SLS by definition. Reliability targets for SLS verifications should be chosen by the client, designer and/or contractor (depending who is responsible for SLS exceedance) according to the risk profile of exceeding the serviceability limit state, because the consequence of SLS exceedance is typically the cost of repair, indirect costs due to loss of functionality or a contractual fine (see Example 4.2).

SLS limit values are mostly based on experience or provided (contractually or based on recommendations). An example would be allowable displacements of a retaining structure of a deep

excavation adjacent to existing buildings. The SLS limit values will then be established as a function of the damage-proneness of the adjacent buildings to displacements.

Recommended reliability target values in the literature are typically in the range of 1 to 5 % probability of failure (i.e. exceedance of SLS), regardless of the reference period, which roughly corresponds to reliability index values in the range of 1.5 to 2.3 (as also mentioned in the JCSS Probabilistic Model Code). The previous version of EN 1990 (EN 1990 1st Generation Eurocodes) recommended $\beta_T = 1.5$ for a reference period of 50 years ⁽⁴⁾; ISO 13822:2010 also recommends $\beta_T = 1.5$ for the assessment of existing structures with a reference period of the remaining service life.

Box 4. Example 4.2: Risk-based reliability target for road settlement (SLS)

Suppose a contractor makes an SLS design for a road embankment with an expected project benefit of 300,000 Euros. The residual settlement may not exceed 10 cm, subject to a contractual fine in case of non-compliance of 3 million Euros. A sensible SLS reliability would then be an acceptable probability of exceeding the residual settlement limit of 10 %, for which the expected loss equals the expected benefit.

Of course, this example is highly simplified, since the expected benefit may be influenced by the design measures taken in order to comply with the SLS target. A cost-benefit analysis of construction costs versus risk cost could accommodate such more detailed information.

4.6 Assessment of existing structures

According to EN 1990-2, the target reliability for an existing structure should take account of the essentially the same relevant factors as for new structures. The target reliability for existing structures can be lower than that for new structures as the relative costs of safety measures to increase the reliability of an existing structure is greater than of a new structure.

General recommendations for reliability targets for existing geotechnical structures are beyond the scope of this guideline ⁽⁵⁾. However, in the absence of specific reliability targets, the target values for new structures as specified in EN 1990-1 (see **Table 1**) can be assumed to represent a safe upper bound.

4.7 Temporary structures

Temporary structures may have different risk profiles compared to permanent structures, and hence may require specific reliability targets. For example, the design lifetime of building pits is several months or years. Furthermore, the use of temporary structures is different to permanent ones, implying potentially different consequences in case of failure. For example, at a construction site many people can be present simultaneously during a short period of time, potentially requiring a higher consequence class (CC).

⁴ EN 1990:2022 Table C2 also states an annual value of $\beta_T = 2.9$, which is questionable because it relates to the 50 year values of $\beta_T = 1.5$ assuming independence of failure events between years, which is hardly justifiable for geotechnical structures generally speaking

⁵ Discussions on how to derive reliability targets for existing structures can be found in the literature, for example in Steenbergen et al. (2018) or Roubos et al. (2019).

When using generally recommended target values for the reliability index, it is recommended to use values for a reference period of 1 year (see 4.4), to be close to the intended design lifetime.

This topic has not been examined yet in the literature, and tailored recommendations for reliability targets are beyond the scope of this guideline ⁽⁶⁾.

⁶ The assessment of temporary structures in the related field of scaffolding is discussed in Vereecken et al. (2020).

5 Uncertainties in geotechnical design and assessment

This chapter elaborates on the characterization of uncertainties involved with the verification of limit states in geotechnical design and assessment. Some contents are closely related to the guidelines on the ground model (European Commission: Joint Research Centre, Garin, H., Baldwin, M., Reiffsteck, P. van der Made K-J. et al., to be published) and representative values (European Commission: Joint Research Centre, Orr, T., Sorgatz, J., Estaire, J., Prästings, A. et al., to be published)

5.1 Principles and definitions

5.1.1 Sources of uncertainty

Uncertainties in geotechnical or civil engineering problems may be categorized according to the source of uncertainty as follows:

- **Inherent variability** (natural or intrinsic variability) is the natural randomness of a quantity, such as the natural variability of the soil strength within a certain soil unit or the time-variability of groundwater levels.
- **Measurement error** is uncertainty caused by imperfect measurement tools and/or sample disturbance effects.
- **Model uncertainty** is uncertainty due to imperfection and simplifications inherent to model formulations. In geotechnical problems we distinguish two types:
 - uncertainty in **calculation models** of the physical behavior of (geotechnical) structures reflecting the error in model predictions with respect to the real behaviour;
 - uncertainty in transformation models (**transformation uncertainty**) used to obtain a certain geotechnical parameter from measurements (e.g. undrained shear strength from CPT through an N_{kt} factor), also called ‘correlations’ (not to be confused with the statistical correlation).
- **Statistical uncertainty** is uncertainty due to limited information, such as a limited number of observations in a site investigation. Furthermore, there is statistical model uncertainty related to the selection of probability distributions.

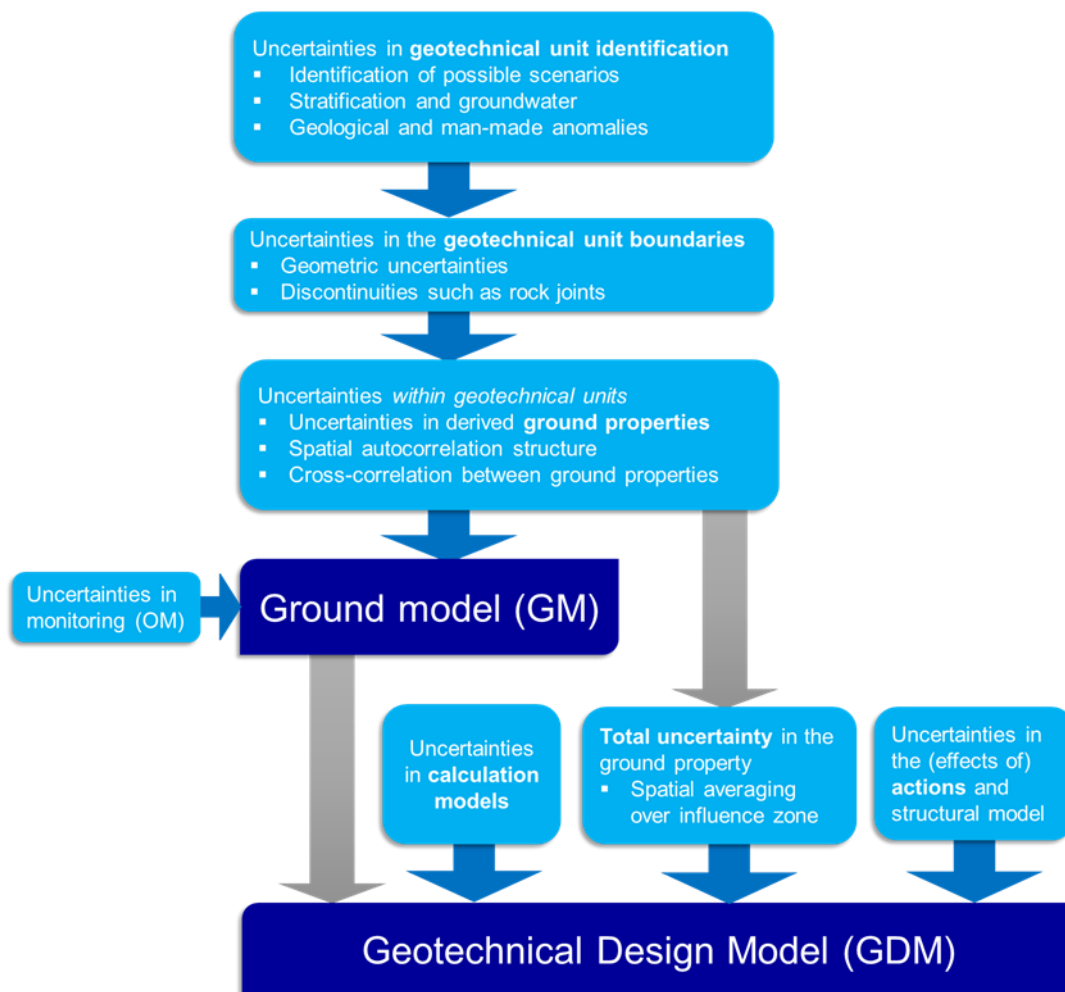
All the above uncertainties are to be accounted for on both the load side (i.e. actions and effects of actions) and the resistance side (e.g. ground properties).

Uncertainties are represented in reliability analysis by modelling the relevant variables as random variables. Any probability distributions used in a reliability analysis should, as far as possible, be documented based on statistical analysis of available background data, or be motivated on other grounds such as literature references or expert judgment.

5.1.2 Uncertainties in geotechnical models

For site characterization, EN 1997-1 distinguishes the Ground Model (GM) and the Geotechnical Design Model (GDM) used for limit state verification. The relation of various uncertainties involved and discussed in this document are illustrated in **Figure 6**, indicating their relation to the GM and GDM.

Figure 6. How the uncertainties in each chapter relate to GM and GDM.



Source: Authors' own work

Uncertainties in the Ground model involve three hierarchies of uncertainties: (1) Uncertainties in the identification of geotechnical units and possible stratification scenarios (e.g. presence or absence of soil layers or anomalies), (2) uncertainties in the boundaries of ground layers (i.e. geometric uncertainties), and (3) Uncertainties in the ground properties within these geotechnical units. The need to address the uncertainties in geotechnical unit identification varies depending on the geotechnical structure, limit state, and local geology.

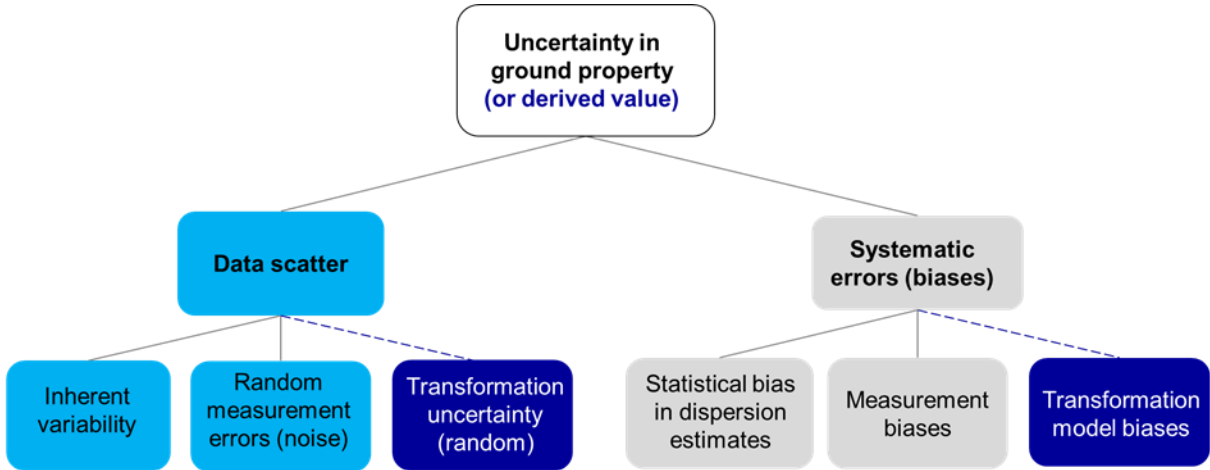
Geotechnical units, such as ground layers (i.e. stratification), are identified by means of ground investigation. Within each geotechnical unit, the ground should be sufficiently homogeneous to be modelled as (statistically) homogeneous (Baecher 2021). That is, the geotechnical unit should be homogeneous enough so that the ground property within that unit can be represented by a single random variable or a stationary random field.

If discontinuities within the geotechnical unit needs to be modelled explicitly, such as for stability problems in rock engineering, geometric uncertainties of the discontinuities can also be modelled as random variables. The division into geotechnical units affects the assessment of uncertainties in the

estimates ground properties. For example, inherent variability within a geotechnical unit may become smaller if the subsoil is divided into multiple homogeneous ground layers.

The estimation of ground properties that represent a geotechnical unit is affected by various sources of uncertainty, some of which are due to systematic errors (see e.g. Baecher 2019), as shown in **Figure 7**. The extent to which the various uncertainties need to be assessed depends on (1) the relevant geotechnical parameters and zone of influence for the considered limit state, (2) the sensitivity of the system reliability to the uncertainty in question, and (3) the target reliability.

Figure 7. Categories of uncertainty in estimated ground property.



Source: Authors' own work

The inherent (natural) variability in ground properties is caused by geological and/or historical processes. Inherent variability in ground properties is usually visible in ground investigation measurements in a form of data scatter even within a seemingly homogeneous ground layer (i.e. geotechnical unit). The inherent variability can be divided into two categories:

- spatial variability (variability with respect to location), and
- temporal variability (variability with respect to time).

Temporal variability in ground properties is related to changes caused by ongoing geological processes or seasonal effects. Compared to temporal variability, spatial variability is usually dominant for ground properties and, hence, in this guideline “inherent variability” refers to spatial variability unless stated otherwise. On the other hand, groundwater level and actions are often marked by temporal variability.

Inherent variability in ground properties is usually characterized by **spatial autocorrelation**, meaning that the same ground property values, measured at close proximity, show significant correlation.

Statistical uncertainty is introduced when limited data are used to estimate properties of probability distributions (e.g. mean, standard deviation or coefficient of variation). Statistical uncertainty can be reduced by collecting more data.

Measurement error arises from imperfect equipment (e.g., lack of calibration), procedural-operator effects (human error) and random measurement noise. The measurement errors may cause both systematic biases and random errors. “Observed variability” (data scatter) in in-situ or laboratory measurements is a result of two sources of uncertainty: inherent variability and measurement error.

Transformation uncertainty is related to situations where in-situ measurements or laboratory results are transformed into *derived values of a ground property* by means of transformation models such as empirical correlations. These transformation models may be biased when applied to specific sites (resulting in systematic errors). Since empirical correlations are always marked by some degree of data scatter in the relationship between measurements and derived values, some transformation uncertainty always prevails even after careful calibration with relevant data. On the other hand, if the ground property is determined via direct measurement (e.g. undrained shear strength determined from direct simple shear test instead of from a CPTu), there is no transformation uncertainty involved.

The **total uncertainty in a ground property** (within one geotechnical unit) is a combination of inherent variability, measurement error, statistical uncertainty and transformation uncertainty (if applicable). Moreover, when ground properties used in geotechnical design are averaged over some dimension (such as sliding surface), variance reduction due to spatial averaging needs to be accounted for.

Correlation may be present between variables involved with the geotechnical site characterization. Correlation (or dependence) refers to any kind of statistical dependence observed between two variables. This also occurs with geotechnical data, and as a result, the random variables in the geotechnical calculation model may be cross-correlated.

5.1.3 Probability interpretation

Most uncertainties in geotechnical design and assessment are related to lack of knowledge rather than inherent variability. Hence, they represent “epistemic uncertainty” that can be reduced by collecting more data. For example, the epistemic uncertainty related to the mean value of soil strength in different geotechnical units can be reduced by performing more ground investigations. On the contrary, aleatory uncertainties are related to inherent variability that cannot be eliminated; for example, the variations in actions and groundwater pressures during the lifetime of a structure represent such aleatory uncertainties.

Geotechnical assessment usually deals with estimating the probability of a single event (e.g. failure of an embankment), and hence adopting a **(Bayesian) subjective interpretation of probability** is sensible in geotechnical engineering (as has been adopted for the reliability concepts of the Eurocodes, see 2.2.1 and European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al., 2024). Frequentist and Bayesian interpretations of probability are compared in **Table 6**.

Table 6. Comparison between Frequentist and Bayesian interpretations of probability

Approach	How probability is defined	Example – Heads or tails?	Objective of statistical inference	Approach to statistical uncertainty
Frequentist	Relative frequency of an outcome in a long run of identical "trials".	Frequentist cannot estimate the probability of a single event (e.g. whether the next toss of a coin will be heads or tails).	Direct probability: Given the state of nature (the hypothesis), what data would be expected?	Separate analysis with <i>confidence intervals</i> .
Bayesian	Degree of belief. Subjective probability (about the state of nature).	Based on prior knowledge, Bayesian can estimate that the probability of heads is $P = 0.50$.	Inverse probability: Given the observed data, how probable are the states of nature (e.g. values of shear strength)?	Statistical uncertainty included in the parameters and description with <i>credible intervals</i> .

Source: Authors' own work

Judgement-based assessment of probabilities can be expressed using verbal classification schemes, see example in **Table 7**. Use of subjective probabilities in geotechnical engineering is further discussed in e.g. Ang and Tang (2006), Baecher (2019), and Vick (2002).

Table 7. Probability estimates based on verbal classification (based on Table 6-7 in Vick 2002).

Probability	Verbal description
0.001	<i>Virtually impossible</i> , due to known physical condition or process that can be described and specified with almost complete confidence.
0.01	<i>Very unlikely</i> , although the possibility cannot be ruled out on the basis of physical or other reasons.
0.10	<i>Unlikely</i> , but it could happen.
0.50	<i>As likely as not</i> , with no reason to believe that one possibility is more or less likely than the other.
0.99	<i>Very likely</i> , but not completely certain.
0.999	<i>Virtually certain</i> , due to known physical conditions or process that can be described and specified with almost complete confidence.

Source: Vick 2002

The determination of uncertainties in geotechnical design and assessment often relies on engineering judgement and utilization of prior knowledge (e.g. literature) and/or comparable experience. That is, the amount of geotechnical data usually poses limitations to the extent to which site-specific uncertainty quantification can be performed on a rigorous statistical basis. Bayesian probability theory enables us to combine data and knowledge from different sources through Bayesian updating, for example engineering judgement-based prior estimates and limited site-specific data, as also acknowledged in EN 1990 (Annex C.3).

5.2 Geotechnical units in Ground Model

Defining the Ground model is explained in European Commission: Joint Research Centre, Garin, H., Baldwin, M., Reiffsteck, P. van der Made K-J. et al. (to be published). This document presents some approaches and examples on how to quantify uncertainties related to the geotechnical units in the Ground model.

5.2.1 Stratification (ground layers)

The characterization of the ground model (i.e. identification of geotechnical units) is essentially a mapping problem involving the following assessments (Baker and Calle 2006):

1. the main pattern of (statistically) homogeneous ground layers,
6. locally present smaller ground units or other local phenomena such as discontinuities,
7. classification of each geotechnical unit.

If uncertainties in this assessment, such as the boundaries of ground layers or other discontinuities, have significant effect on the performance of geotechnical structure, these uncertainties should be considered explicitly in the reliability analysis. The accuracy by which a geotechnical unit can be defined, depends on the prior geological model, and the extent and quality of the ground investigation. Defining geotechnical usually involves considerable engineering judgement. For practical application, the uncertainty in stratification may be considered in three alternative ways:

(1) Explicit probabilistic modeling

There are different options for explicit probabilistic modelling of the uncertainty in ground layer geometry such as Kriging (Example 5.1), (conditional) random field simulation, or by treating certain geometrical properties as random variables (e.g. the vertical coordinates of layer boundaries). Often the geometric uncertainties are not that influential to the reliability and can be modeled as best estimates deterministically.

Box 5. Example 5.1: Kriging of layer boundaries

Kriging is a probabilistic interpolation method for predicting unknown values from observations at known locations, such as locations of soil layer boundaries between ground investigation profiles. In this example, we make a best linear unbiased estimation (BLUE) of bedrock elevation based on 8 boreholes as summarised in the following table, where z is the bedrock elevation.

borehole	x [m]	y [m]	z [m]	borehole	x [m]	y [m]	z [m]
B1	18.70	3.78	-4.31	B5	29.91	22.29	-5.24
B2	23.08	4.19	-4.29	B6	25.42	24.53	-5.34
B3	20.99	7.31	-4.65	B7	10.50	36.90	-3.11
B4	23.20	9.93	-4.93	B8	0.89	30.78	-3.83

We can estimate the semi-variogram representing the auto-correlation structure from the data of the 8 boreholes with to the following covariance model (with $\delta_{\tau=0}$ being the Dirac delta function for adding measurement uncertainty, **Figure 8**):

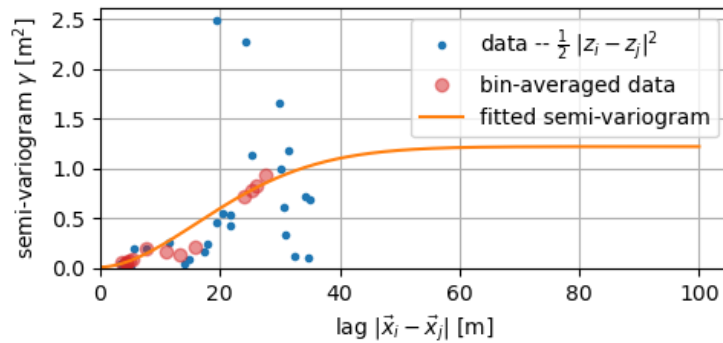
$$C(\tau) = \sigma_{inh}^2 \exp(-\tau^\nu) + \sigma_{meas}^2 \delta_{\tau=0}$$

$$\tau = \frac{1}{\theta} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

The fitted parameters of this covariance model are:

- $\sigma_{inh}^2 = 1.1^2$: (stationary) variance in the bedrock elevation
- $\sigma_{meas}^2 = 0.1^2$: variance of the measurement error
- $\nu=1.8$: shape parameter (controlling the smoothness)
- $\theta=15$ m : horizontal auto-correlation length (i.e. scale of fluctuation)

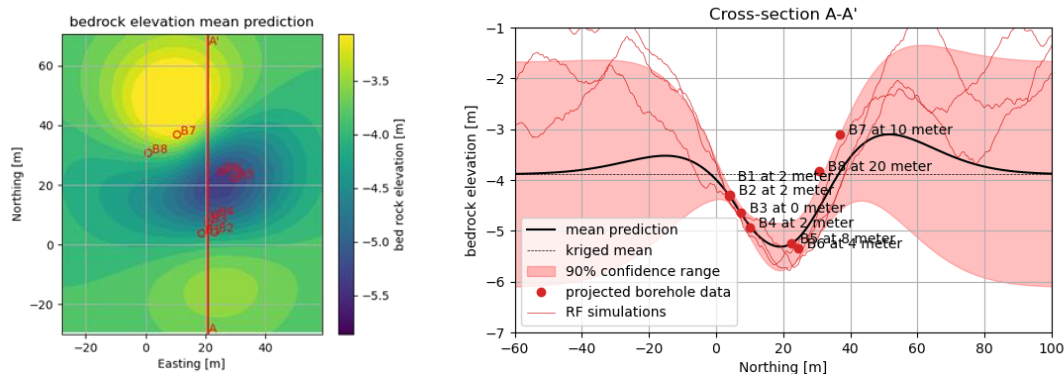
Figure 8. Calibrated semi-variogram $\gamma(\Delta x) = \sigma_{tot}^2 - C(\Delta x)$ against the data of table above.



Source: Authors' own work

Figure 9 presents the kriging predictions in plain view and in cross-section A-A'. The kriging prediction matches the conditioning data at the borehole locations (with a tolerance of σ_{meas}) and converges to the kriged mean $\mu_{OK} = -3.88$ m with increasing distance from the borehole locations. The cross-sectional prediction shows the 90% confidence range based on the ordinary kriging variance, indicating the depth range in which the bedrock can be expected. Three realisations of conditional random fields (RF simulations) are included as examples of sampling-based simulation of the bedrock elevation. Alternatively, the location-dependent combination of kriging mean and variance can be used for stochastic characterisation of the bedrock elevation.

Figure 9. Predicted bedrock elevation using ordinary kriging with the above-specified data and covariance model in plain view (left) and on cross-section A-A' (right).



Source: Authors' own work

The complete formulation of (ordinary) kriging can be found in the literature on geostatistics (e.g. Wackernagel (2003)).

Whenever the geotechnical unit must be considered as a discontinuous media, geometric properties, such as orientation and spacing of discontinuities in rock masses, should be considered explicitly as random variables with probabilistic properties derived from the statistical analysis of geometric data.

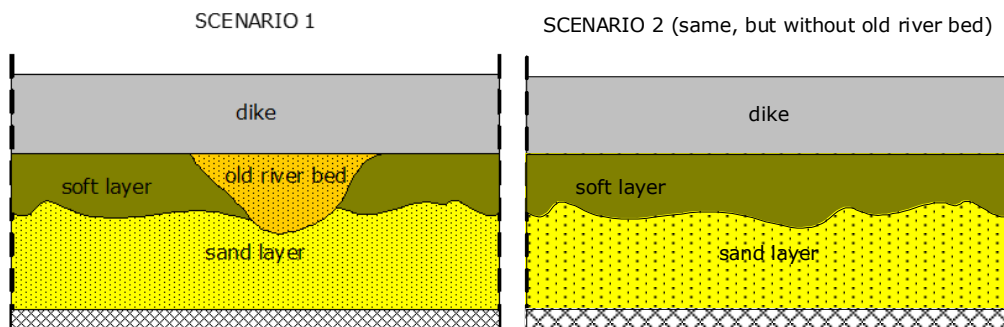
(2) Stratification scenarios

An alternative option for dealing with uncertainty in the ground model is to define discrete stratification scenarios. That is, to generate different versions of the ground model that are each plausible with the site investigation and prior knowledge of the site available. Subjective estimates of these probabilities can be based on verbal classification schemes (see **Table 7**). See Example 5.2 for an illustrative example of defining stratification scenarios.

Box 6. Example 5.2: Stratification scenarios

Suppose we are assessing the stability of a dike (flood embankment) for internal erosion piping. The dike (embankment) rests on a soft soil layer, and at several meters depth there is an aquifer (sand) (**Figure 10**).

Figure 10. Illustration of stratification scenarios for a dike on a soft soil deposit with potential sand lenses



Source: Authors' own work

In deltaic areas, old riverbeds of historically meandering rivers may cut through the blanket layer and form weak spots or dike sections in the considered reach. In this example engineering geologists estimate the probability of encountering at least one old riverbed in the considered reach to be 5% based on their geological knowledge of the area and the general spatial frequency of these features occurring in this geological environment (e.g. based on identification by hand-borings or high-resolution digital elevation models).

If no other scenarios besides the base scenario (without old riverbed) and the river bed scenario are considered, the base scenario has a probability of 95% of being true in the considered reach, since the scenarios need to be mutually exclusive (i.e. if one is true, the other is not) and exhaustive (the probabilities of all scenarios sum up to 1).

(3) Conservative characterization

Another alternative is to use a conservative characterization of the stratification (i.e. deterministic approach) based on engineering judgement. This approach is comparable to using conservative point estimates for variables (see 6.3.5). In order not to bias the reliability analysis, the degree of conservatism of the chosen ground model should be comparable to choosing a representative or

characteristic value for a ground property. Consequently, using the definitions from **Table 7**, the characterization should include features or values that are ‘unlikely, but could be true’.

5.2.2 Geological and man-made anomalies

Geotechnical units may contain anomalies such as faults, lenses or fills. Such geological anomalies, if not revealed by ground investigations, may have a significant effect on the geotechnical structures. Based on the knowledge of the regional geology and the structure to be designed, the designer evaluates whether the possibility of certain anomalies may be disregarded or need to be accounted for. The properties and the dimensions of the structure to be designed governs what kinds of geological and man-made anomalies are significant in terms of functional performance.

The modelling options in reliability analysis for such local phenomena are essentially the same as discussed for stratification in 5.2.1. The detection of this type of anomaly feature using a finite number of soil samples is addressed in the field of search theory (Baecher and Christian, 2003; Benkoski et al., 1991). Examples of approaches to boulder detection can be found from Ang and Tank (2006), Tang (1987), Tang and Halaim (1988), Tang and Quek (1986), and Tang and Saadeghvaziri (1983).

5.2.3 Groundwater

In EN 1990-1 (6.1.3.2) and 1997-1 (6.4 and **Figure 30**), the uncertainty in groundwater actions is treated in terms of probability of exceedance and the corresponding return period, as summarized in **Table 8**.

Table 8. Probability-based groundwater actions according to EN 1990

Value of variable water action	Symbol	Probability of exceedance	Return period (years)
Characteristic	Q_{wk}	2% per annum	50
Combination	$Q_{w,comb}$	10% per annum (2022 draft)	10
Frequent	$Q_{w,freq}$	Fraction of time exceeded = 1%	
Quasi-permanent	$Q_{w,qper}$	Fraction of time exceeded = 50%	
Accidental	$Q_{w,rep}$	0.1% per annum	100

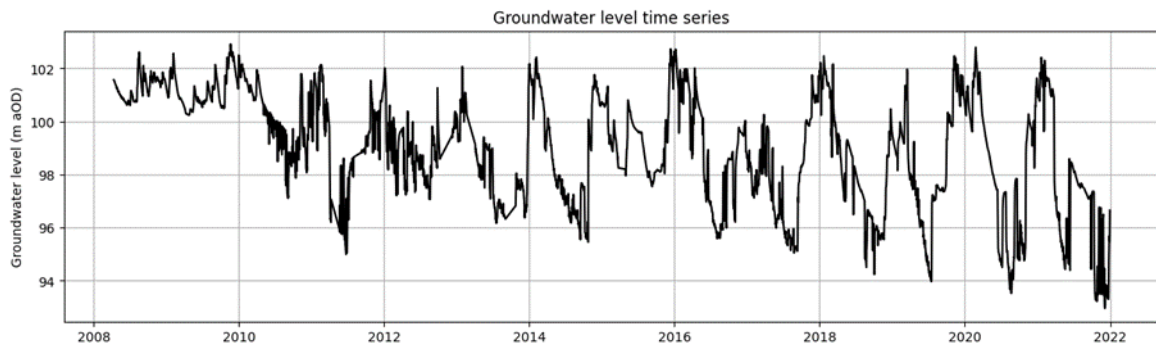
Source: Authors’ own work

The values with a prescribed probability of exceedance can be obtained by fitting extreme value distributions to time series of groundwater level measurements (see Example 5.3), if available. In the absence of data, judgement-or experience-based estimates will be necessary (similar to nominal value estimation of ground properties).

Box 7. Example 5.3: Fitting extreme value distribution to groundwater level timeseries

In this example, we determine the probability-based groundwater levels by fitting an extreme value distribution to the time series below (**Figure 11**), with a record of roughly 14 years.

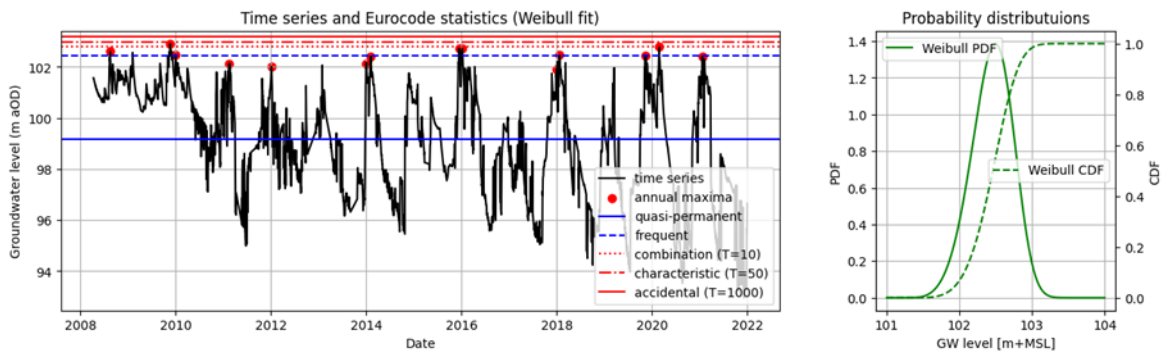
Figure 11. Time series of measured groundwater levels.



Source: Authors' own work

After extracting the annual maxima from the record, we fit extreme value distributions to these values. In this case, the Weibull distribution (for maxima) gave a better fit than the Generalized Extreme Value (GEV) or Gumbel distribution. The Weibull distribution and the related values are shown in **Figure 12**.

Figure 12. Time series and the fitted values (left) and the Weibull distribution (right)



Source: Authors' own work

Note that the quasi-permanent and frequent value (both blue) do not require the extreme value distribution, but are based on the 50% (median) and 99%-quantiles of the time series (if readings are evenly spaced). Furthermore, the analysis could be improved by defining the annual maxima not on calendar years but between summers for example, since some maxima belong to the same event. Comparing the extreme value distributions can be done by comparing with the empirical CDF visually, using goodness-of-fit tests or other judgement-based criteria. In general, using Gumbel or Weibull distributions will be sufficient, with Gumbel being the more conservative choice (fatter right tail).

5.3 Ground properties

5.3.1 Introduction to treating uncertainty in ground properties

The minimum requirements for defining a random variable are estimates of the mean (μ_x) and standard deviation (σ_x) or variance (σ_x^2) in the case of two-parameter distributions such as Gaussian (normal) distribution or lognormal distribution. The dispersion (standard deviation or variance) can also be defined by means of coefficient of variation (V_x), which is defined as the standard deviation divided by the mean ($V_x = \sigma_x / \mu_x$); as such it represents the relative dispersion or uncertainty.

Ground properties are inherently spatially variable. There are fundamentally two approaches to deal with spatial variability in uncertainty modelling: (a) explicit modelling of spatial variability using random fields and (b) implicit consideration of spatial variability by using single random variables per ground property and geotechnical unit, while accounting for the effects of spatial variability in the estimation of the probability distribution (parameters).

This guideline focuses on the latter, the single random variable approach, using so-called mobilized values (Phoon et al., 2024), such that the result of the analysis in terms of the response ideally resembles the outcome of an analysis with the Random Finite Element Method (RFEM). In producing the distributions of mobilized values, the main effect to account for is spatial averaging.

This choice was made because at the time of writing reliability analysis with standard geotechnical models is considered the most realistically accessible option for practitioners, as opposed to RFEM which is mostly restricted to academic studies and special applications due to its computationally expensive nature.

5.3.2 Uncertainty components

The uncertainty in a ground property typically originates from various sources. The total uncertainty $V_{X,tot}$ in a (spatial average) ground property may be presented as (Ching et al. 2020):

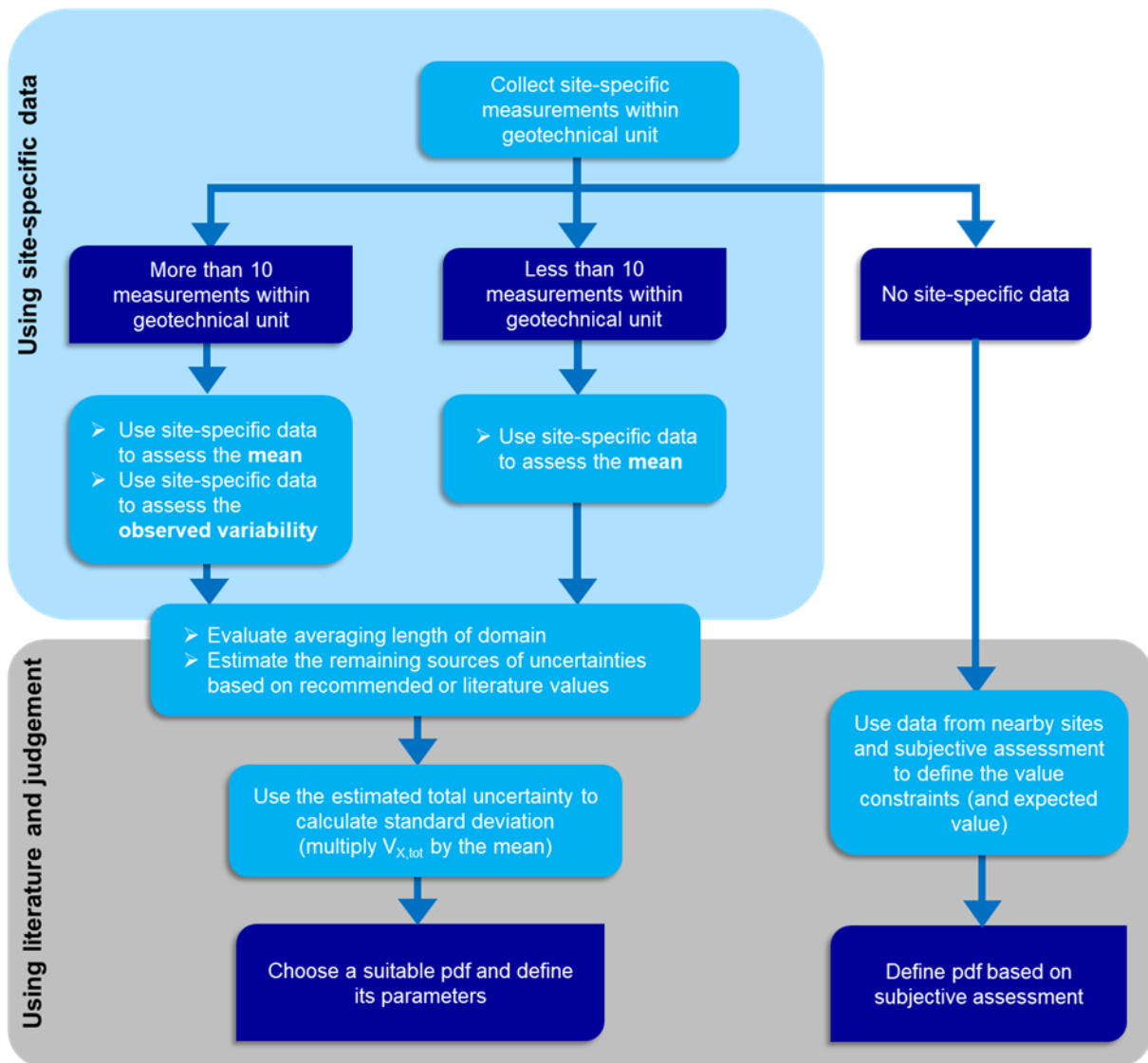
$$V_{X,tot} = \sqrt{V_{X,inh}^2 \Gamma^2 + V_{X,trans}^2 + V_{X,meas}^2 + V_{X,stat}^2}$$

Equation 16.

where $V_{X,inh}$, $V_{X,meas}$, $V_{X,trans}$ and $V_{X,stat}$ are coefficients of variation for inherent variability, measurement error, transformation uncertainty, and statistical uncertainty, respectively. Γ^2 is variance reduction factor that accounts for spatial averaging effects (see O). It is noteworthy that also Annex A (EN 1997-1) acknowledges these four sources of uncertainty (note that Annex A refers to measurement error by $V_{x,quality}$).

The available site-specific data affects the process of uncertainty characterization, as illustrated in **Figure 13**. Note that even when there is enough site-specific data ($n > 10...30$) to evaluate the coefficient of variation, the calculated value should also be compared with literature ranges and recommended indicative values.

Figure 13. Steps in uncertainty characterization depending on the number of site-specific measurements.



Source: Authors' own work

5.3.3 Inherent variability

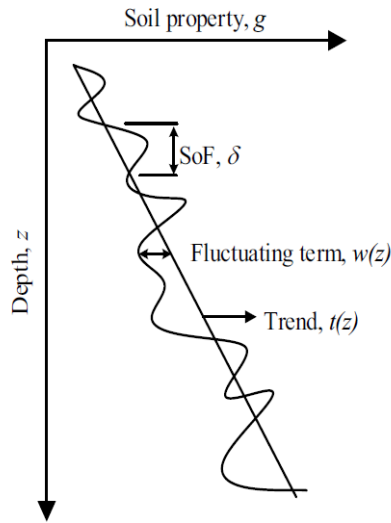
The inherent variability (i.e. spatial variability) of a ground property in a certain direction may be decomposed into a trend function (describing the mean) and a random variable that represents the fluctuation around the mean trend). For example, a soil property $X(z)$ which exhibits a trend with depth (z) , can be described by a trend function $t(z)$ and a random fluctuation $w(z)$:

$$X(z) = t(z) + w(z)$$

Equation 17.

Using this model, the coefficient of variation for inherent variability $V_{x,inh}$ may be defined using the standard deviation of the term $w(z)$. The (vertical) scale of fluctuation θ_v , describes the distance over which the residuals $w(z)$ are correlated (as illustrated in **Figure 14**).

Figure 14. Inherent variability of soil and trend with depth (Nadim 2015)



Source: Nadim 2015

The inherent variability in the horizontal direction may be modelled analogously, using horizontal coordinates instead of depth coordinates. Scale of fluctuation in the horizontal direction (θ_h) is usually one or two orders of magnitude greater than the vertical θ_v (Cami et al. 2020).

Accumulating literature on the magnitude of inherent variability has shown that certain physical soil properties tend to exhibit $V_{x,inh}$ values that vary within the same order of magnitude (Phoon and Kulhawy 1999; Lacasse et al. 2007; Uzielli et al. 2006). Inherent variability can be considered universal in a sense that factors like the geological age of the soil have little effect on $V_{x,inh}$ (Uzielli et al. 2006). Therefore, the literature ranges of $V_{x,inh}$ can be used at sites where the amount of data are not sufficient for statistical analysis of the inherent variability (Phoon and Kulhawy 1999). Indicative values of $V_{x,inh}$ are given in **Table 9**, based on literature values collected in Annex A to this report, which in turn is largely based on TC 304, 2021. The recommended indicative values can be useful in situations when the expected value of a ground property can be based on site-specific data (or experience), while there are insufficient data to estimate the variance (or coefficient of variation).

Table 9. Recommended indicative coefficients of variation for inherent variability ($V_{x,inh}$) for various ground properties (based on ranges reported in TC 304 (2021), see Annex A to this report).

Property	Soil	Recommended $V_{x,inh}$	Range of COV from EN 1997-1 ¹
γ (kN/m ³) (total)	Clay and sand	0.05	0.05-0.10
ϕ' (°)	clay	0.15	0.05-0.15
ϕ' (°)	sand	0.08	
c' (kPa)	clay	0.30	0.30-0.50
s_u (kPa)	clay	0.30	0.30-0.50
s_u/σ'_v	clay	0.20	-

Property	Soil	Recommended $V_{X,inh}$	Range of COV from EN 1997-1 ¹
OCR	clay	0.20	-
C_c	clay	0.35	-
C_{ur}	clay	0.40	-
K_0	clay	0.15	-
K_0	sand	0.30	-
SPT-N	clay	0.30	0.15-0.45
SPT-N	sand	0.35	
q_c (MPa)	clay and sand	0.10	0.05-0.15
E (MPa) (in-situ)	sand	0.35	0.20-0.70

(¹) The indicative ranges in EN 1997-1 Annex A do not specify whether they refer to inherent variability, or to observed or total variability. Hence, they can be considered as upper bound values for inherent variability.

Source: Authors' own work

5.3.3.1 Assessment of inherent variability from data

Inherent variability of ground properties is usually the primary cause for the data scatter observed within ground layers. However, scatter can also be caused by measurement errors. Hence, both inherent variability ($V_{X,inh}$) and measurement error ($V_{X,meas}$) contribute to the observed variability ($V_{X,obs}$) that can be assessed from the in-situ or laboratory measurements. Inherent variability $V_{X,inh}$ may be evaluated from $V_{X,obs}$ and $V_{X,meas}$ according to (Orchant, Kulhawy, and Trautmann 1988; Phoon and Kulhawy 1999; Müller, Larsson, and Spross 2014):

$$V_{X,inh}^2 = V_{X,obs}^2 - V_{X,meas}^2$$

Equation 18.

It should be noted that the above equation assumes independence (i.e. untransformed data). That is, if the data contains derived ground properties estimated using transformation models, the transformation uncertainty also contributes to the data scatter within the geotechnical unit (see 5.3.7). If $V_{X,meas}$ is significantly smaller than $V_{X,inh}$, the observed variability $V_{X,obs}$ may be taken as an approximation for the actual inherent variability $V_{X,inh}$.

The observed variability $V_{X,obs}$ may be larger than the actual inherent variability $V_{X,inh}$ (Phoon and Kulhawy 1999), if:

- data from different geological layers are mixed, or
- $V_{X,meas}$ is large due to poor quality equipment and insufficient procedural controls, or
- spatial trends with depth are not considered (data not de-trended), or
- data are collected over a long time period.

If the mean ground property can be assumed to be constant within the geotechnical unit (e.g. there is no trend with depth or in horizontal direction), $V_{x,obs}$ may be estimated by means of the sample standard deviation s_x :

$$s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - m_x)^2}$$

Equation 19.

where x_i is the measured ground property, n is the number of independent measurements x_i , and m_x is the sample mean (arithmetic mean) of all measurements x_i . As n increases, s_x provides more accurate estimate for the actual standard deviation of the whole population, σ_x . $V_{x,obs}$ can then be calculated from:

$$V_{X,obs} = \frac{s_x}{m_x}$$

Equation 20.

If there is a trend with depth, the mean trend of a ground property is first defined. Usually, linear regression can be used, resulting in the following linear trend function $t(z)$:

$$\hat{t} = t(z) = a_0 + a_1 z$$

Equation 21.

where a_0 is the intercept coefficient, a_1 is the slope (gradient) coefficient, and z is the depth (or elevation). Then, the de-trended standard deviation (assumed to be independent of depth) can be estimated from (Lacasse et al. 2007; DNV 2021):

$$s_{x,detrended} = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (x_i - \hat{t}_i)^2}$$

Equation 22.

where x_i is the measured ground property value at depth z_i , \hat{t}_i is the value predicted by the trend function (**Equation 21**) at that depth, n is the number of independent measurements x_i , and $(n-2)$ is the degrees of freedom. Least-squares regression analysis usually assumes that the residuals are normally distributed. The natural logarithm (ln) may be used to transform the trend value into a normally distributed value (see 5.4.4).

After defining $s_{x,detrended}$, $V_{x,obs}$ may be calculated using the ground property value predicted by the trend function, at the middle of the affected volume within the ground layer (Phoon and Kulhawy 1999). If the measured values are distributed evenly across the trend line, using the mean ground property m_x or the mean value of all trend values leads to an almost identical result.

If the number of measurements n within the geotechnical unit is very small (e.g. $n < 10...15$), the estimate s_x is marked by significant statistical uncertainty. For small samples, the observed value range ($x_{max}-x_{min}$) may be utilized to acquire an estimate for σ_x . For normally distributed data, an estimate for σ_x is given by (e.g., Baecher and Christian 2003):

$$\widehat{\sigma}_x = N_n(x_{max} - x_{min})$$

Equation 23.

where N_n is the correction factor for estimating the standard deviation that depends on the sample size n (see **Table 10**).

Table 10. Correction factor N_n for estimating the standard deviation from sample range for normally distributed variable as a function of the sample size n .

n	N_n	n	N_n	n	N_n
2	0.886	9	0.337	16	0.283
3	0.510	10	0.325	17	0.279
4	0.486	11	0.315	18	0.275
5	0.430	12	0.307	19	0.271
6	0.395	13	0.300	20	0.268
7	0.370	14	0.294	30	0.244
8	0.351	15	0.288		

Source: Authors' own work

Box 8. Example 5.4: Assessment of observed variability from field vane data

Consider a soft clay layer (depth $z = 0.78\text{--}4.32$ m) with $n = 22$ measurements combined from three field vane shear strength $s_{u,FV}$ profiles (data from Lehtonen et al. 2015). The soft clay layer was identified by means of both $s_{u,FV}$ and classification tests. The 22 $s_{u,FV}$ observations are assumed independent (uncorrelated). In this example, the field vane measurements were reduced with a plasticity-dependent correction factor of 0.94 to correct for strain-rate effects, leading to the derived s_u values presented in the table below.

z	0.78	1.19	1.28	1.32	1.69	1.78	1.82	2.19	2.28	2.32	2.69
$s_{u,FV}$	9.3	12.2	10.2	13.6	9.3	11.3	12.5	9.9	6.4	16	10.4
s_u	8.7	11.5	9.6	12.8	8.7	10.6	11.8	9.3	6.0	15.0	9.8
z	2.78	2.82	3.19	3.28	3.32	3.69	3.78	3.82	4.19	4.28	4.32
$s_{u,FV}$	11.3	12.8	10.7	11.6	14.8	12.8	9.9	15.4	11.3	11.6	16.8
s_u	10.6	12.0	10.1	10.9	13.9	12.0	9.3	14.5	10.6	10.9	15.8

Option (a) mean $s_{u,FV}$ constant with depth

The sample mean is $m_x = 11.1$ kPa. The sample standard deviation (**Equation 19**) and coefficient of variation (**Equation 20**) are:

$$s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - m_x)^2} = 2.30 \text{ kPa} \quad (\rightarrow V_{x,obs} = 0.207)$$

To compare, the range estimate for standard deviation (**Equation 23**) is given by:

$$\widehat{\sigma}_x = N_n(x_{max} - x_{min}) = 0.268 \times (15.8 - 6.0) \text{ kPa} = 2.62 \text{ kPa} \quad (\rightarrow V_{x,obs} = 0.236)$$

Option (b) mean $s_{u,FV}$ as linear trend with depth

The mean can be defined using linear regression (**Equation 23**) as:

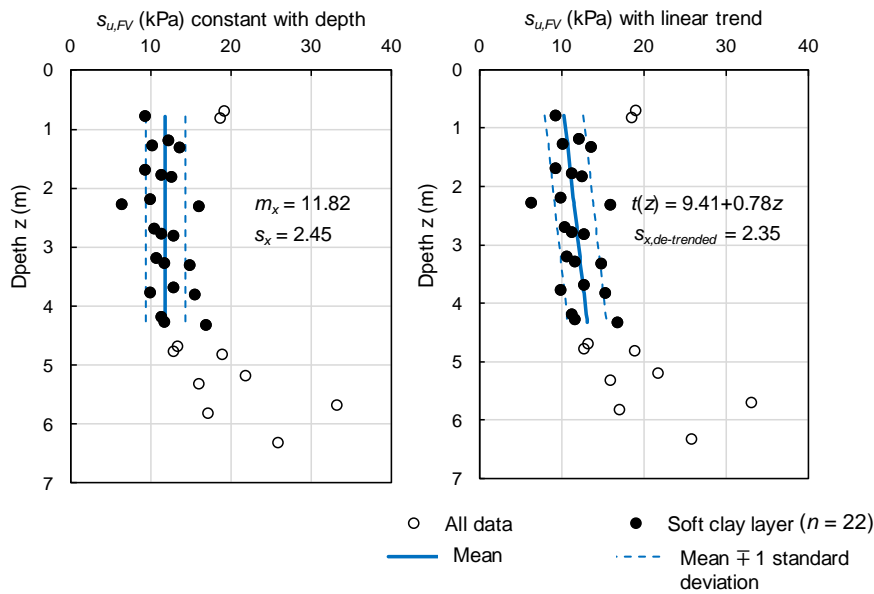
$$\hat{t} = t(z) = a_0 + a_1 z = 9.16 + 0.73z$$

The de-trended standard deviation (**Equation 22**) can be defined from:

$$s_{x,detrended} = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (x_i - \hat{t}_i)^2} = 2.21 \text{ kPa} \quad (\rightarrow V_{x,obs} = 0.201)$$

The corresponding $V_{x,obs}$ was calculated using the trend value \hat{t} at the middle of the soft clay layer (11.0 kPa). If the mean value of all \hat{t} is used, $V_{x,obs} = 0.20$ is obtained. In this example, the mean can be assumed constant with depth, as the difference in the results is very small (**Figure 15**).

Figure 15. Estimates of mean and standard deviation for field vane strength for soft clay.



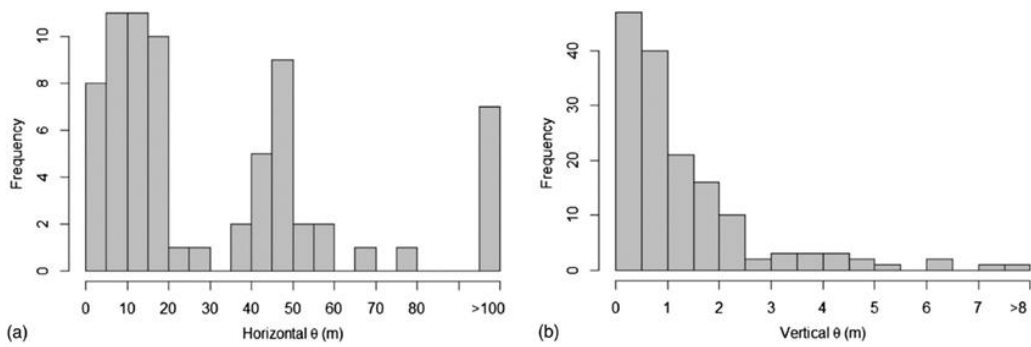
Source: Authors' own work

Notice that this example does not yet consider statistical or transformation uncertainty (see Example 5.5).

5.3.3.2 Estimation of scale of fluctuation

Typical value ranges for the vertical and horizontal scales of fluctuation θ_v and θ_h for soil properties have been recently summarized (Arnold, 2016; Cao et al., 2016; Cami et al., 2020 and TC304, 2021). The TC304 (2021) report points out that most studies utilized CPT profiles due to the amount and frequency of data obtained. As a result, the θ_v estimates (compared to the θ_h estimates) can be considered more reliable. **Figure 16** illustrates histograms of both and shows that mainly due to geological deposition θ_v tends to be much smaller compared to θ_h .

Figure 16. Horizontal and vertical scales of fluctuations for soil properties (Cami et al. 2020).

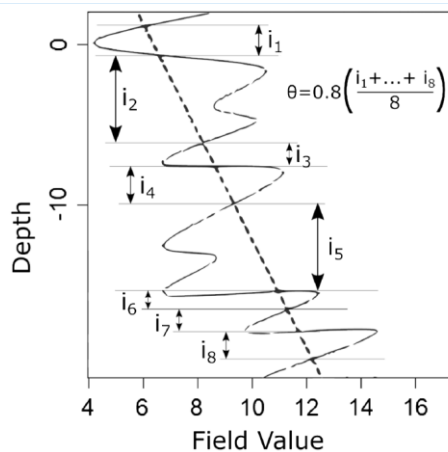


Source: Cami et al. 2020

The practical implication is that soil layers can be interpolated between ground investigation profiles in the horizontal direction with rather high reliability (assuming that no anomalies have been detected, see 5.2.2) for most soils, possibly with the exception of soft organic soils.

Methods to estimate the vertical scale of fluctuation from site-specific CPT data have been described and compared by Lloret Cabot et al. (2014) and Cami et al. (2020). The latter state-of-the-art report includes reviews on (1) typical value ranges for vertical and horizontal θ , (2) worst-case θ values for different geotechnical problems, (3) methods to estimate θ from CPT data and examples and (4) different autocorrelation models. The ‘rule of thumb’ method Cami et al. (2020) used to estimate θ_v is depicted in **Figure 17**. The linear dashed line represents the mean trend of the measured property. Values i are defined as the distances between intersections of the mean trend and the fluctuating property. An approximation for θ_v is then given by the average distance multiplied by factor 0.8 (Vanmarcke, 1977).

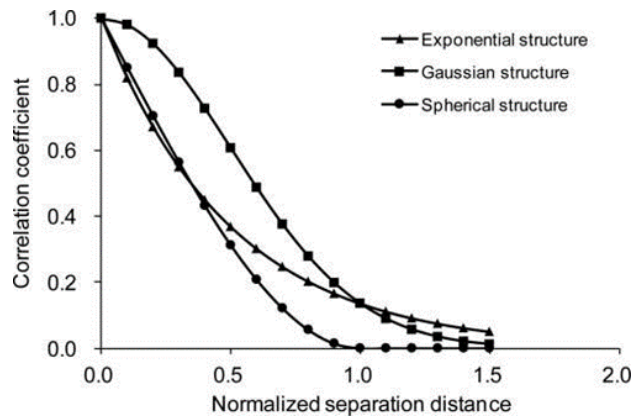
Figure 17. “Rule of thumb” method to estimate scale of fluctuation δ (Cami et al. 2020).



Source: Cami et al. 2020

The scale of fluctuation is a parameter that describes the autocorrelation of inherent variability and may be modelled via an autocorrelation function (or covariance function), see **Figure 18**. The autocorrelation function describes the coefficient of correlation of a soil property between two points as a function of their distance (DNV, 2021).

Figure 18. Illustration of auto-correlation functions



Source: Zhu and Zhang (2013)

Table 11 summarizes ranges for typical values of the scale of fluctuation. As outlined above, spatial correlation and as such the scales of fluctuation are not fixed properties but rather dependent on the observation scale. The values presented in **Table 11** represent the engineering scale for geotechnical structures.

If no or little information on the scales of fluctuation or degree of anisotropy is available, the geotechnical design should account for the “worst-case” scale of fluctuation which may be present at site. This can be determined within a sensitivity analysis and by engineering judgement with respect to the investigated limit states and associated failure volumes. Larger scales of fluctuations tend to result in more conservative estimates, as less local averaging takes place.

Table 11. Ranges for typical scale of fluctuation.

	Likely range	Potential range	Remarks
Vertical scale of fluctuation θ_v [m]	0.2 ... 1.0	0.1 ... 2.5	In exceptional cases in the order of $\theta_v = 5$ to 10 m have been found.
Horizontal scale of fluctuation θ_h [m]	5.0 ... 30.0	2.0 ... 100	Very low values are typically found for organic soft soils. High values are likely for marine deposits.
Degree of anisotropy $\zeta = \theta_h / \theta_v$	5.0 ... 30.0	2.0 ... 100	Potentially positive correlation with the vertical scale of fluctuation.

Source: Authors' own work

5.3.4 Spatial averaging

The geotechnical parameter of interest is the ground property value affecting the occurrence of the limit state. Therefore, averaging over an affected volume or a failure surface is typically involved, leading to variance reduction. The spatially averaged parameter should be the actual physical property that defines the ground resistance (e.g. $\tan \phi$ instead of the angle ϕ). The averaging effect can be determined in terms of the variance reduction factor Γ , which is governed by the relationship between the averaging dimensions (i.e. affected volume), and the spatial auto-correlation. For one-dimensional averaging over a length L , the variance of the averaged random field is given by (Nadim 2015):

$$\sigma_{X,av}^2 = \Gamma^2(L)\sigma_X^2$$

Equation 24.

where Γ^2 is the variance reduction factor ($0 \leq \Gamma^2 \leq 1$), is the variance of parameter X , and L is the length over which X is averaged. The resulting distribution with σ_X^2 is sometimes referred to as the distribution of the mobilized value (Phoon et al., 2024). For an in-depth treatment of spatial averaging over lengths, areas and volumes refer to Vanmarcke (2010).

For practical applications, the following approximation of the variance reduction factor Γ^2 in the vertical direction has been proposed by Vanmarcke (1983):

$$\Gamma^2 = \begin{cases} \Gamma^2(L) \approx 1 & \text{if } L \leq \delta_v \\ \Gamma^2(L) \approx \frac{\delta_v}{L} & \text{if } L > \delta_v \end{cases}$$

Equation 25.

where L is the averaging length and δ_v is the vertical scale of fluctuation (if averaging occurs in horizontal direction, δ_h should be used instead).

The extent of variance reduction depends on the averaging length L (or other domain) and the scale of fluctuation δ . Effectively, full variance reduction from spatial averaging effect (i.e. $\Gamma^2 \approx 0$) would occur if δ is a small fraction of averaging length L ($L/\delta \gg 1$). On the contrary, no significant variance reduction occurs (i.e. $\Gamma^2 \approx 1$) if δ is equal to or larger than L ($L/\delta \approx 1$). In the intermediate case, some variance reduction occurs but also the mean may be reduced by weak zone seeking. These three cases (Hicks et al. 2019) are compared in **Table 12** and also addressed in European Commission: Joint Research Centre, Orr, T., Sorgatz, J., Estaire, J., Prästings, A. et al. (to be published).

The averaging length L (or area or volume) depends on the geotechnical problem and the limit state considered (see **Figure 19**), and therefore the averaging effect should be addressed for each Geotechnical Design Model separately. For instance, for s_u , the averaging length may be defined from the mobilization zone (resulting in mobilized undrained shear strength, $s_{u,mob}$). If the calculation model is discretized into thinner calculation layers (e.g. to consider change in mean ground property or to consider different layer stiffnesses in a settlement problem), the averaging length is the thickness of the calculation layer (TC304, 2020).

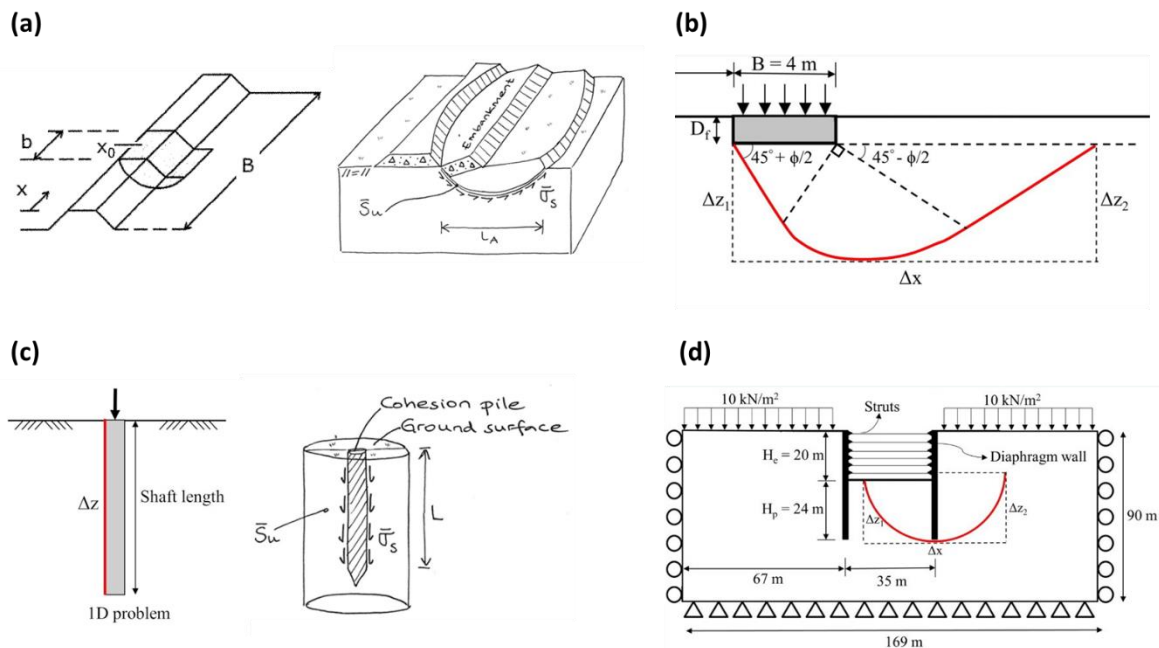
Table 12. Different cases of spatial averaging

Case	Description (applicable to strength variables)	Ratio L/δ	Variance reduction factor Γ^2	Case in Fig. 21	Corresponding case for characteristic value in EN 1997
A	Considerable averaging of soil property value which leads to narrower mobilized value distribution, centered about the mean of the underlying distribution	$L/\delta \gg 1$ (δ is only a small portion of averaging domain)	$\Gamma^2 \approx 0$ (full variance reduction)	Spatially averaged shear strength	Case A: Estimate of the mean value

Case	Description (applicable to strength variables)	Ratio L/δ	Variance reduction factor Γ^2	Case in Fig. 21	Corresponding case for characteristic value in EN 1997
B	Failure tends to be local and the mobilized value distribution resembles the underlying distribution (no averaging)	$L/\delta \leq 1$ (δ is equal to or larger than averaging domain)	$\Gamma^2 \approx 1$ (no variance reduction)	Point shear strength	Case B: Estimate of the inferior or superior value
C	Extent of spatial averaging is problem-dependent and there are two competing factors: (i) mobilized value distribution becomes narrower due to variance reduction, and (ii) mean is reduced compared to the underlying distribution due to weak zone seeking	intermediate values of L/δ	$0 < \Gamma^2 < 1$ (some variance reduction)	Mobilised shear strength	Case C: intermediate value

Source: Authors' own work

Figure 19. Examples of averaging lengths: (a) 3D slope failure, (b) Strip footing subjected to vertical loading, (c) friction pile under compression, and (d) basal heave for excavation in clay



Source: (a) Vanmarcke 2011; Müller 2013, (b) Tabarrok et al. 2022, (c) Tabarrok et al. 2022; Müller 2013, and (d) Tabarrok et al. 2021

Box 9. Background: Practical assumption of full spatial (depth-)averaging

A common practical assumption for spatial averaging is that

- a. full averaging over depth applies (since the vertical dimension of the affected volume typically is a multiple of the vertical scale of fluctuation);

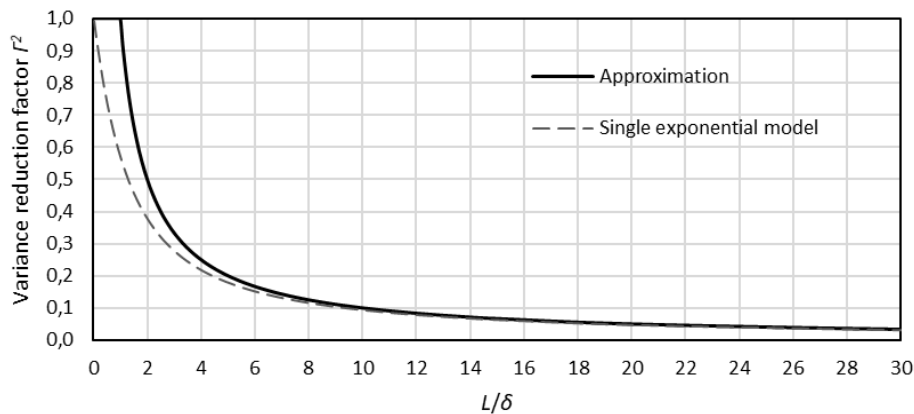
b. no averaging in the horizontal plane applies (since the horizontal dimensions are typically less than the horizontal scale of fluctuation).

When the above conditions apply (mainly a), we may apply full spatial (depth-)averaging and a variance reduction of $\Gamma^2 = 0$.

Notice that the same assumptions are also underlying equations for characteristic values, as explained in Deltares (2021).

Equation 25 implies that the variance reduction factor may be estimated using δ only, without the need to choose a specific auto-correlation function. Nonetheless, definitions of factor Γ^2 for specific autocorrelation models are also available (see e.g. Cami et al., 2021). **Figure 20** illustrates Γ^2 as a function of ratio L/δ for the approximation method and single exponential model, implying that for roughly $L/\delta \geq 10$ the practical assumption of full spatial averaging seems appropriate, since then other sources of uncertainty (e.g. statistical due to limited number of data) will usually dominate.

Figure 20. Relationship between variance reduction factor Γ^2 and ratio L/δ .

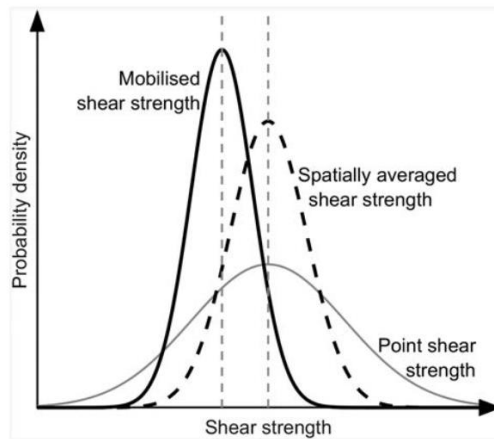


Source: Authors' own work.

5.3.4.1 Weak zone seeking

Spatial averaging as conceptualized here presumes a fixed or known failure surface (e.g. along the shaft of a foundation pile). When the geometry of the failure surface is not fixed, weak zones can influence the failure mechanism and lower the mean value of the mobilized strength value. **Figure 21** illustrates that while the variance of the ground property is reduced by the averaging affect, this weak zone seeking can cause the distribution to shift to the left (i.e. mean value has decreased) (Case 3 in **Table 12**). The Random Finite Element Method (RFEM) can be used in such conditions (e.g. Varkey, Hicks and van den Eijnden 2020; Tabarroki et al. 2021); see section 6.5.

Figure 21. Probability distributions of point and mobilized shear strength



Source: Tabarrokhi et al. 2021

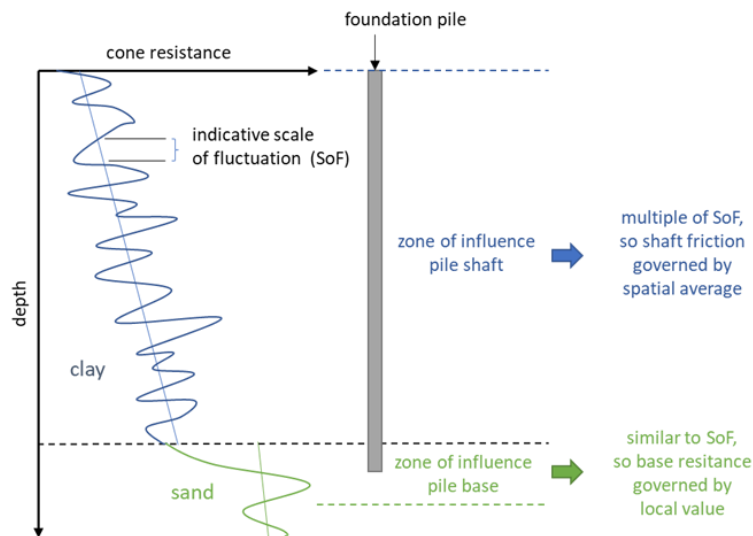
Weak zone seeking is considered to be marginal for most applications, even though for some applications it can be considerable and possibly be accounted for by reduction of the property mean (Tabarrokhi et al., 2021).

Notice that the above concerns ‘random’ weak zones in otherwise statistically homogeneous ground layers which fall within the normal spatial variability patterns, as opposed to (not identified) weak layers. The latter should be considered as separate geotechnical units.

5.3.4.2 Non-averaging properties

Some problems are not governed by average ground properties, but are instead affected by local weak zones, in which case variance reduction is not applicable. For example, seepage may occur through weak (high permeability) zones and hence averaging effects are less significant. As illustrated in **Figure 22**, the pile base resistance may be governed by local and hence inferior value, as opposed to the shaft friction, which is often governed by the average.

Figure 22. Spatial averaging for a foundation pile



Source: Authors' own work

5.3.5 Measurement error

Measurement error represents the deviation of the measured value from the actual value. The error stems from several sources such as imperfect equipment or (human) procedural–operator effects. Measurement error can include both systematic error (bias) and random measurement error (noise). Bias is the (systematic) difference between the actual mean and the mean of the measurements. Bias may be introduced by poorly calibrated equipment or systematic error in the procedure. In laboratory testing, sample disturbance often causes both systematic bias and random errors (Baecher and Christian 2003). Increasing the number of ground investigation data does not necessarily decrease bias, since the bias might remain the same throughout testing. Systematic bias is not reduced by averaging, while random errors are. If the magnitude of the bias is known, it should be accounted for by correcting the measurements. Uncertainties related to equipment and operator effects are often random. Such random measurement errors cause scatter around the mean.

Gross errors can be caused by faulty equipment or procedure, and they are supposed to be avoided by means of quality control and operational requirements (e.g. by following the recommendations in execution standards). Gross errors related to interpretation of ground investigations should be avoided by a thorough control scheme and an expert review (Baker and Calle 2006). Gross errors are not usually included stochastically in the reliability analysis.

Some examples of evaluated total measurement error, defined through coefficient of variation $V_{X,meas}$, are collected in **Table 13**. A more comprehensive literature collection can be found in Annex A and ISSMGE-TC304 (2021). When choosing estimates for $V_{X,meas}$ from literature, it is important to make sure that $V_{X,meas}$ is smaller than the observed variability $V_{X,obs}$ (**Equation 18**). For example, the unit weight for homogeneous clay may have $V_{X,obs}$ as low as 1-2 % (Löfman and Korkiala-Tanttu 2019, 2022), and in such cases $V_{X,meas} = 0\%$ may be assumed despite the literature range of $V_{X,meas} = 1-2\%$.

Table 13. Recommended indicative measurement error in laboratory and in-situ tests (largely based on Phoon and Kulhawy 1999).

Testing type	Test	Recommended indicative $V_{X,meas}$	Estimated range of total $V_{X,meas}$ (%)
Laboratory tests	s_u for clays and silts	0.10	5-38
	φ (°) for clays, silts and sands	0.05	3-56
	$\tan(\varphi)$ for clays, silts and sands	0.05	2-22
	Atterberg limits (%) for fine-grained soils	0.05	3-18
	Water content w (%) for fine-grained soils	0.05	6-12
	Unit weight γ (kN/m ³) for fine-grained soils	0.00	1-2
In-situ tests	Standard penetration test (SPT)	0.20	15-45
	Electric cone penetration test (CPT)	0.10	5-15
	Field vane shear test (VST)	0.15	10-20
	Dilatometer test (DMT)	0.10	5-15

Source: Authors' own work

Box 10. Measurement error in estimating total uncertainty

Measurement error should be applied explicitly in estimating the total uncertainty when inherent uncertainty is estimated from literature or recommended indicative values.

When inherent uncertainty is estimated from site investigation data, the measurement error is already included in the observed variability.

When indirect measurement (e.g. CPT-based correlations) are used, measurement error is already covered by the transformation uncertainty, and not considered separately.

5.3.5.1 Measurement error in spatially averaged ground properties

When the mean value of ground property is evaluated, the related measurement error decreases with increasing number of independent measurements n due to averaging (Spross and Larsson 2021; Müller, Larsson, and Spross 2014):

$$V_{X,avemeas}^2 = V_{X,meas}^2 \frac{1}{n}$$

Equation 26.

Notice that assuming high measurement error is not necessarily conservative in estimating the total uncertainty (see 5.3.8). Deltares (2021) recommends that measurement errors are not accounted for in the statistical estimation of characteristic values of soil strength properties, unless it is quite certain which part of the measurement error is random and which part is systematic.

5.3.6 Statistical uncertainty

Statistical uncertainty is introduced when estimating statistics of a population based on a sample with limited observations or data. When the total uncertainty in a spatially averaged ground property is assessed, the statistical uncertainty in the mean (i.e. spatial average) is taken into account, while the uncertainty in dispersion measures (e.g. standard deviation) is usually ignored.

Statistical uncertainty is treated differently manner in frequentist and Bayesian statistics. From a Bayesian perspective statistical uncertainty is included in the analysis by default, because the probabilistic model parameters (e.g. the distribution parameters) are modelled as random variables. In a frequentist approach, model parameters are fixed, and statistical uncertainty is addressed separately through confidence intervals.

This section is limited to classical frequentist approaches to statistical uncertainty. Regarding geotechnical data, we often have prior knowledge, and the sample sizes are typically small and therefore Bayesian statistics are generally advantageous as any available prior knowledge can be considered, as covered in chapter 6.

The statistical uncertainty ($V_{X,stat}$) related to assessing the (spatially averaged) mean ground property from limited data (n independent observations) can be estimated from (e.g., Müller at al., 2014):

$$V_{X,stat}^2 = V_{inh}^2 \psi$$

Equation 27.

where the statistical factor is $\psi = 1/n$ if there is no trend with depth (i.e. constant mean), $V_{x,inh}$ is uncertainty related to inherent variability. As the sample size n increases, $V_{x,stat}$ is reduced and can usually be ignored if $n \geq 20$.

The above equation is based on the standard deviation of sample mean, also called standard error of the mean, which is defined as the population standard deviation σ_x divided by \sqrt{n} . Hence, the standard deviation of the sample mean σ_{m_x} is given by:

$$\sigma_{m_x} = \frac{\sigma_x}{\sqrt{n}} \approx \frac{s_x}{\sqrt{n}}$$

Equation 28.

where s_x is the sample standard deviation of measurements x_i and n is the number of independent observations. If the 'true' standard deviation of the population (σ_x) is unknown, the Student t-distribution may be used when estimating the uncertainty in sample mean: if the underlying distribution is Gaussian (normal), but σ_x is unknown, the resulting estimated distribution follows the Student t-distribution which has 'fatter' tails compared to normal distribution. The advantage of the simple definition in **Equation 28** above is that it applies to any probability distribution (e.g., Baecher 2019).

In the case of the mean being a linear trend with depth, the statistical factor ψ is a function of n and also of depth z (for a normally distributed variable; e.g. Spross and Larsson, 2021):

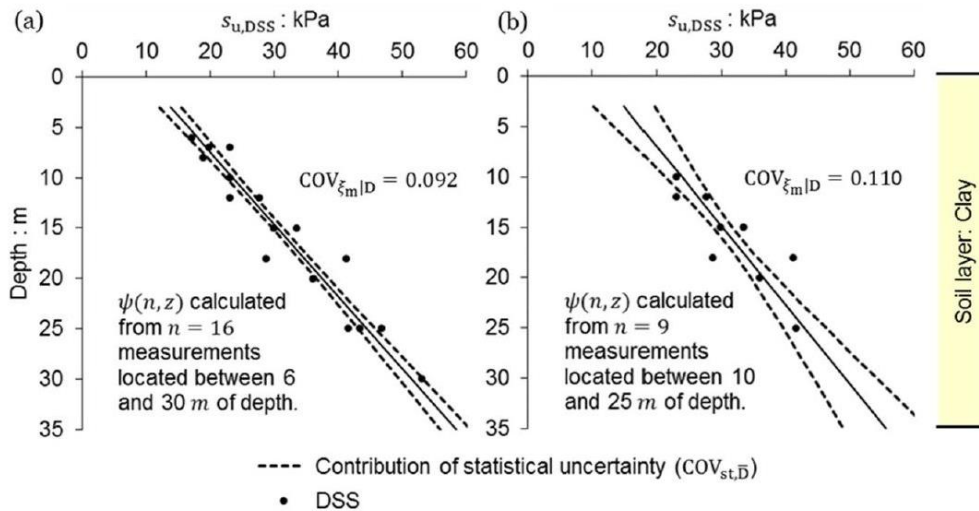
$$\psi = \psi(n, z) = \frac{n-1}{n-3} \left\{ \frac{1}{n} \left[1 + \frac{n}{n-1} \frac{(z - m_z)^2}{s_z^2} \right] \right\}$$

Equation 29.

where m_z is the sample mean of depths z_i where the measurements were taken and s_z^2 is the sample variance at the respective depths. This factor $\psi(n,z)$ represents the statistical uncertainty in the estimation of the regression coefficients (intercept a_0 and slope a_1) and the variance σ_x^2 (Prästings 2019). **Figure 23** compares this statistical uncertainty (the dashed line marks \pm one standard deviation) in two different measurements sets: the statistical uncertainty increases with smaller sample size and greater extrapolation of the trend line (black line). Furthermore, if the variance of the ground property X can be assumed to be known (e.g. based on prior knowledge), the term $(n-1)/(n-3)$ is omitted (Müller 2013).

Compared to estimating the mean, estimation of standard deviation σ_x (and variance) or the coefficient of variation V_x tend to require larger sample sizes n . In general, the variance underestimated if n is small, and in such case literature values should be utilized. According to TC304 (2021), there should be at least $n = 10$ observations if sample V_x is estimated from site-specific data, whereas statistics with $n = 30$ can be considered very reliable (TC304, 2021). For further information about the statistical uncertainty in dispersion estimates, see Annex A.

Figure 23. Examples of contribution from statistical uncertainty in the assessment of inherent variability of undrained shear strength



Source: Prästings 2019

5.3.7 Transformation uncertainty

Transformation uncertainty is involved with the estimation of geotechnical parameters (i.e. derived values) when applying transformation models to the results of in-situ or laboratory measurements. Many transformation models are based on empirical correlations like linear regression equations, but they also frequently incorporate some element of theory. The correlation is usually obtained with scatter, which is one component of the transformation uncertainty. Also, biases may be involved; that is, the transformation model may on average overestimate or underestimate a geotechnical parameter, sometimes bias may be on purpose (e.g. conservative estimation). Transformation models defined with “global data” (i.e. consisting of various sites and soil types) tend to be less biased globally, but also have more scatter. Nevertheless, for a specific site the bias in the global transformation models may be significant. Meanwhile, site-specific transformation models are typically unbiased for the investigated site, but may suffer from lack of data (i.e. statistical uncertainty), **Figure 24**.

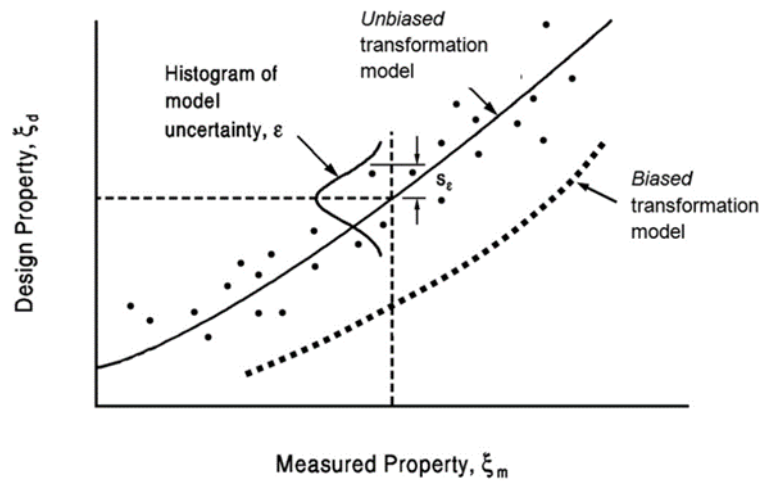
Based on the definition by Ching and Phoon (2014), the transformation uncertainty is usually described using the following ratio (TC304, 2021):

$$\varepsilon = \frac{\text{measured value}}{\text{predicted value}}$$

Equation 30.

The mean value of the ratios ε is the bias (b), and the coefficient of variation is denoted as V_ε .

Figure 24. Transformation uncertainty resulting from pairwise correlation between measured (target) value and the derived value of a ground property (“design property” in the figure; not to be confused with the design value)



Source: Ching et al. 2017

With known bias and uncertainty, the true value can be modeled as (Ching and Phoon 2014):

$$\text{true value} = \text{predicted value} \times b \times \varepsilon$$

Equation 31.

where the variability term ε for the transformation model may be assumed lognormal with mean 1 and coefficient of variation V_ε . This characterization of transformation uncertainty is referred to as the multiplicative form (Ching et al. 2017). An alternative characterization of transformation uncertainty would be an additive form with a normal distributed zero-mean error term (see e.g. Ching et al. (2017) and Phoon and Kulhawy 1999b).

If the data used to derive the transformation model is not available, approximate visual estimation may be used to evaluate the standard deviation of the transformation model (s_ε): roughly, two thirds of the data should fall within \mp one s_ε (Phoon and Kulhawy 1999b).

Indicative values of $V_{x,\text{trans}}$ (i.e. V_ε above) are given in **Table 14**, based on average literature values as collected in Annex A. If available, transformation uncertainty assessed using local (regional) data should be preferred over indicative values. Transformation uncertainties and biases may be based on global databases (see e.g. TC304, 2021) in the absence of other information. More specific transformation models carefully fitted using high-quality regional data have the smallest δ .

Empirical studies suggest that transformation uncertainty is not affected by spatial averaging (TC304 2021): Ching et al. (2016) studied the CPTU-su transformation model and concluded that transformation uncertainty does vary in space. However, the vertical scale of fluctuation of the transformation uncertainty was so large that it could be assumed that the transformation uncertainty is fully correlated in space in practical terms. Site-specific transformation models (Van der Krogt et al, 2018) can behave differently.

Table 14. Indicative values for coefficients of variation for transformation uncertainty for various ground property pairs (largely based on ranges reported in TC 304 (2021), see Annex A).

Transformation model	Soils	Global $V_{\text{trans}} (\delta)$	Regional $V_{\text{trans}} (\delta)$ [carefully fitted model]
$s_u - \sigma'_p$	clays	0.50	0.30
$\left(\frac{s_u}{\sigma'_v}\right) - OCR$	clays	0.50	0.30
$\left(\frac{s_u}{\sigma'_v}\right) - CPT$	clays	0.50	0.30 [0.20]
$OCR - CPT$	clays	0.40	0.20
$\sigma'_p - CPT$	clays	0.40	0.20 [0.15]
$C_c - LL$	clays	0.90	0.70
$C_c - e_0$	clays	0.70	0.50 [0.35]
$C_s - e_0$	clays	0.70	0.40
$\varphi' - SPT$	sands	0.10	0.10
$\varphi' - CPT$	sands	0.10	0.10
$D_r - SPT$	sands	0.20	0.20
$D_r - CPT$	sands	0.30	0.20

Source: Authors' own work

5.3.8 Total uncertainty

To begin with, the total uncertainty of a (spatial average) ground property can be based on recommended indicative values of the coefficient of variation based on the literature in the absence of sufficient, adequate data. The other option is to consider all individual components, if relevant, to compose the total uncertainty. Both options are described below.

5.3.8.1 Recommended indicative values of total uncertainty

Phoon and Ching (2015) present ranges of total coefficient of variation for reliability calibration purposes (see **Table 15** and Annex A).

Table 15. Ranges of total soil property variability (for reliability calibration) according to Phoon & Ching (2015) and Phoon & Kulhaw (2008)

Geotechnical parameter	Property variability	COV (%)
Undrained shear strength	Low ¹	10-30
	Medium ²	30-50
	High ³	50-70

Geotechnical parameter	Property variability	COV (%)
Effective stress friction angle	Low ¹	5-10
	Medium ²	10-15
	High ³	15-20
Horizontal stress coefficient	Low ¹	30-50
	Medium ²	50-70
	High ³	70-90

(¹) Typical of good-quality direct laboratory or field measurement.

(²) Typical of indirect correlations with good field data, except for the SPT.

(³) Typical of indirect correlations with SPT field data and with strictly empirical correlations.

Source: Phoon and Kulhawy 2008

5.3.8.2 Total uncertainty – general equation

Combining the elements from the previous sections, the total uncertainty in the spatial average (mean) ground property may be calculated from (Müller, Larsson, and Spross 2014):

$$V_{X,tot}^2 \approx V_{X,inh}^2 \Gamma^2 + \overbrace{V_{X,inh}^2 \psi}^{V_{X,stat}^2} + \overbrace{V_{X,meas}^2 \left(\frac{1}{n}\right)}^{V_{X,avemeas}^2} + V_{X,trans}^2$$

Equation 32.

The measurement error $V_{X,meas}$ is reduced to $V_{X,avemeas}$ only if the n measurements are independent from each other (e.g. no systematic measurement error). Hence, n should be taken as the number of laboratory tests or number of sounding profiles. If averaging of measurement error is not considered, the above equation provides the upper bound estimate for total uncertainty.

Box 11. Background: Total uncertainty in ground properties

Measurement error should be applied explicitly in estimating the total uncertainty when inherent. When the total uncertainty in the estimated (spatial average) ground property is assessed, the combination of inherent variability, measurement error, statistical uncertainty and transformation uncertainty (if applicable) should be considered. The total uncertainty in a spatially averaged ground property $\sigma_{X,tot}^2$ can be defined using the sum of variances while assuming independence between the sources of error (based on Ching et al. 2020):

$$\sigma_{X,tot}^2 = \sigma_{X,inh}^2 \Gamma^2 + \sigma_{X,stat}^2 + \sigma_{X,avemeas}^2 + \sigma_{X,trans}^2$$

where $\sigma_{X,inh}^2$, $\sigma_{X,stat}^2$, $\sigma_{X,meas}^2$, and $\sigma_{X,trans}^2$ are variances of inherent variability, statistical uncertainty, measurement error, and transformation uncertainty, respectively; and Γ^2 is the variance reduction factor that accounts for spatial averaging. The above sum of variances applies regardless of distribution type. If the components are independent, normally distributed random variables, the sum is also normally distributed. Likewise, if the components are independent, lognormally distributed random variables, their product is also lognormally distributed. Frequently, there is more information on the coefficients of variation than on the variances. Analogously, the total uncertainty $V_{X,tot}^2$ in a spatial average ground property may then be presented as (Ching et al. 2020):

$$V_{X,tot}^2 \approx V_{X,inh}^2 \Gamma^2 + V_{X,stat}^2 + V_{X,meas}^2 + V_{X,trans}^2$$

where $V_{X,inh}$, $V_{X,stat}$, $V_{X,meas}$, and $V_{X,trans}$ are the coefficients of variation. The above approximation applies if the total uncertainty is composed of independent random variables coefficients of variation of which are small (less than 0.30) (Melchers and Beck, 2018).

The various components in **Equation 19** may be evaluated according to the previous subsections. Below some simplifications of the equation for certain conditions or special cases are considered.

5.3.8.3 Directly measured ground properties (site-specific dataset)

In case of directly measured ground properties, $V_{X,trans} = 0$ and drops out. If there are enough site-specific measurements to estimate dispersion ($n > 10 \dots 30$), we may assume that the calculated observed variability $V_{X,obs}$ is the sum of inherent variability $V_{X,inh}$ and measurement error $V_{X,meas}$, thus resulting in (Müller, Larsson, and Spross 2014):

$$V_{X,tot}^2 \approx \left(\frac{V_{X,inh}^2}{V_{X,obs}^2 - V_{X,meas}^2} \right) \left(\Gamma^2 + \frac{1}{n} \right) + \frac{V_{X,avemeas}^2}{V_{X,meas}^2} \left(\frac{1}{n} \right)$$

Equation 33.

For $V_{X,meas}$, recommended literature values may be used in the absence of project-specific information, with a maximum of $V_{X,meas} \leq 0.3 V_{X,obs}$. It should be noted that in the above equation, larger measurement error can be 'beneficial' (i.e. unconservative) in terms of total uncertainty.

For small sample sizes ($n < 10$), using a Student-t distribution would be more appropriate to also account for the statistical uncertainty in the variance, and to be consistent with the recommendation for assessing characteristic values in European Commission: Joint Research Centre, Orr, T., Sorgatz, J., Estaire, J., Prästings, A. et al. (to be published).

5.3.8.4 Indirectly measured ground properties (using transformation models)

The total uncertainty in a spatially averaged value of a derived ground property using a transformation model (e.g. an empirical correlation) can be acquired from:

$$V_{X,tot}^2 \approx V_{X,inh}^2 \left(\Gamma^2 + \frac{1}{n} \right) + V_{X,trans}^2$$

Equation 34.

where $V_{X,inh}$ is simply taken as the sample coefficient of variation (i.e. equal to $V_{X,obs}$) calculated using the derived ground property values (i.e., the geotechnical parameters which have been estimated using certain transformation models) and $V_{X,trans}$ is the related transformation uncertainty.

If the transformation uncertainty is defined using the additive form (zero-mean normal variable), the equation for $V_{X,tot}$ depends on the form of the transformation model. Examples of $V_{X,tot}$ for different transformation models can be found in e.g. Phoon and Kulhawy (1999b).

Box 12. Example 5.5: Total uncertainty in undrained shear strength from field vane tests

Consider the soft clay layer (depth $z = 0.78\text{--}4.32$ m) with 22 measurements of field vane shear strength $s_{u,FV}$ presented in Example 5.4. The observed variability was $V_{X,obs} = 0.207$, which is taken to be equal to the inherent variability $V_{X,inh}$.

When s_u is estimated from $s_{u,FV}$ measurements, a correction factor is needed to consider strain-rate and anisotropy effects. The transformation uncertainty related to field vane testing of soft clays is estimated to have the range $V_{X,trans} = 0.075\text{--}0.15$ (Phoon and Kulhawy 1999b). A mid-range value is chosen, i.e. $V_{X,trans} = 0.11$.

The variance reduction factor Γ^2 is estimated while assuming that the averaging length is the thickness of the soft clay layer, $L = 3.54$ m. For the vertical scale of fluctuation, $\delta_v = 1$ m is chosen based on the upper bound of the likely value range of clays. Vanmarcke's approximation gives $\Gamma^2 \approx 0.28$.

The total uncertainty $V_{\bar{X},tot}$ related to spatially averaged s_u can then be evaluated from **Equation 34**

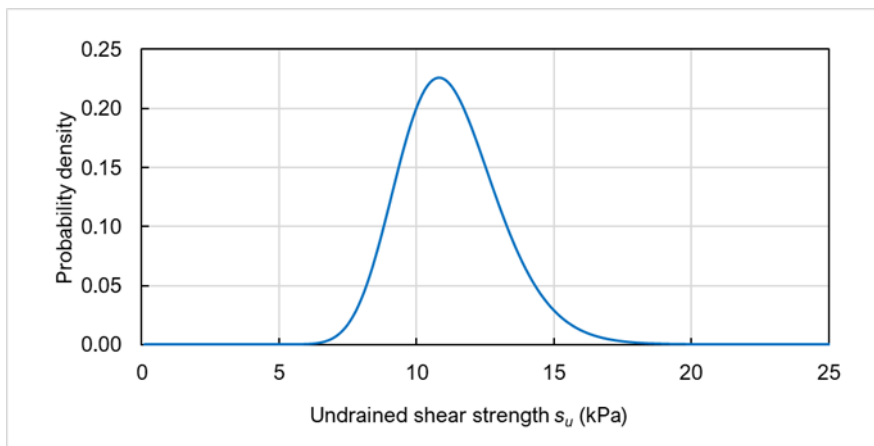
$$V_{\bar{X},tot}^2 = V_{X,inh}^2 \left(\Gamma^2 + \frac{1}{n} \right) + V_{X,trans}^2 = 0.207^2 \left(0.28 + \frac{1}{22} \right) + 0.11^2 = 0.026 \rightarrow V_{\bar{X},tot} = 0.162$$

Now that the total uncertainty has been assessed, the pdf can be constructed. A lognormal distribution is assumed (**Figure 25**).

The standard deviation corresponding to the total uncertainty (σ_{su}) is given by:

$$\sigma_{su} = \mu_{su} \times V_{\bar{X},tot} = 11.1 \text{ kPa} \times 0.162 = 1.80 \text{ kPa}$$

Figure 25. Total uncertainty in undrained shear strength represented by lognormal distribution.



Source: Authors' own work

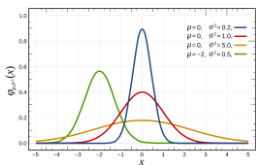
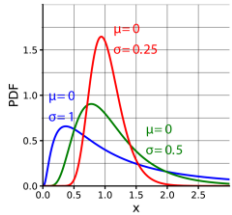
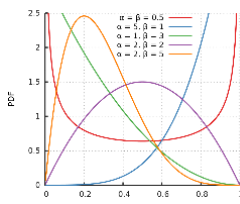
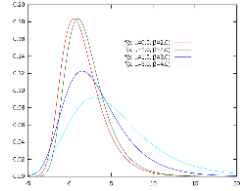
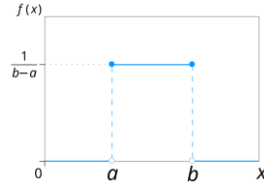
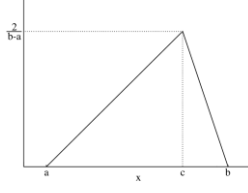
5.4 Defining probability distributions

Analysis of available data will not always give enough information, so the choice of probability distribution will often involve engineering judgement. This section provides some guidance on how to choose an appropriate one.

5.4.1 Useful probability distributions

Examples of continuous probability density distributions commonly used to represent the variability in ground properties are collected in **Table 16**.

Table 16. Commonly used continuous probability distributions

Probability distribution	Probability density function (pdf)	Examples (pdf)	Notes
Normal (Gaussian)	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)^2\right]$ $-\infty < x < \infty$	 <p>Source: Wikipedia – public domain</p>	mean = μ , standard deviation = σ
Lognormal	$f_X(x) = \frac{1}{x\sigma_{\ln x}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \mu_{\ln x}}{\sigma_{\ln x}}\right)^2\right]$ $0 \leq x < \infty$	 <p>Source: Wikipedia – public domain</p>	$\ln(x)$ is normal distributed
Beta	$f_X(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$ $x \in [0,1] \text{ or } x \in (0,1)$	 <p>Source: Wikipedia – public domain</p>	α and β are shape parameters ($\alpha, \beta > 0$)
Gumbel (extreme value distribution)	$f_X(x) = \frac{1}{\beta} e^{-(z+e^{-z})}$ <p>where $z = \frac{x-\mu}{\beta} \quad -\infty < x < \infty$</p>	 <p>Source: Wikipedia – public domain</p>	$Mode \approx \mu;$ $\beta = \frac{\sigma\sqrt{6}}{\pi};$ $Mean \approx \mu + 0.5772\beta$
Uniform	$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$	 <p>Source: Wikipedia – public domain</p>	All values are equally likely
Triangular	$f(x) = \begin{cases} 0 & x < a \\ \frac{2(x-a)}{(b-a)(c-a)} & a \leq x \leq c \\ \frac{2(b-x)}{(b-a)(b-c)} & c \leq x \leq b \\ 0 & b < x \end{cases}$	 <p>Source: Wikipedia - CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=182090</p>	mode = c , mean = $(a+b+c)/3$

Source: Authors' own work

Some uncertainties in geotechnical design are better represented by discrete distributions such as binomial and Poisson (**Table 17**). For example, the Poisson distribution can be used to model the number of fracture zones in a given tunnel length (Stille et al. 2003).

Table 17. Commonly used discrete probability distributions.

Probability distribution	Probability mass function (pmf)	Examples (pdf)	Notes
Binomial	$f_X(k, n, p) = \binom{n}{k} p^k (1-p)^{n-k}$ where $p \in [0,1]$ is the probability of getting “success” in each Bernoulli trial	pmf gives the probability of getting exactly k successes in n independent Bernoulli trials	In Bernoulli trials, there can be either the occurrence (“success”) or the non-occurrence of an event, e.g. the exceeding of an allowable vibration limit when blasting (Stille et al. 2003).
Poisson	$f_X(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$ where k is the number of occurrences ($k = 0, 1, 2 \dots$), e is Euler’s number, and $!$ is the factorial function. Parameter λ is the expected rate of occurrences	pmf gives the probability of k events occurring in a fixed interval of time or space if these events occur with a known constant mean rate λ and independently of the time since the last event	Example: Calculating the probability of k overflow floods to occur in 100 years, while the average rate is one overflow flood per 100 years ($\lambda = 1$). Note that pmf can be adapted to have $\lambda = rt$ where r is the average rate at which events occur and t is the studied time interval.

Source: Authors’ own work

5.4.2 Practical approaches to defining probability distributions

A suitable probability density distribution (pdf) for a given ground property (or other random variable) can be chosen via different approaches:

5.4.2.1 Approach A: Fitting to site-specific data

If the amount of site-specific data is sufficient (e.g. $n \geq 15$) for distribution fitting, a method like maximum likelihood (MLE) or method of moments can be applied. After finding one or more candidate distributions, their suitability for the dataset can be evaluated by means of graphical assessment or goodness-of-fit tests (like Kolmogorov–Smirnov or Anderson–Darling). One common graphical tool is the quantile-quantile graph; another simple approach is to compare the relative frequency histogram of the data and the pdfs. Another recommended visual method is to compare the empirical cumulative distribution function (ecdf) constructed using the dataset (i.e. based on relative frequencies) with the cdf of the distribution candidate (Benjamin and Cornell 1970).

It should be noted that goodness-of-fit tests consider the data as a whole, and therefore the distribution with highest scores might not be the best fit in the tails; from the reliability point of view the fit to the tails is often more important (Baecher 2019).

5.4.2.2 Approach B: Combining site-specific data and literature guidelines

In this simplified approach, recommended indicative coefficients of variation are utilized to define a normal or lognormal distribution by method of moments. The process entails the following steps:

1. Calculate or estimate the site-specific mean ground property (within geotechnical unit)
2. Use indicative literature V_x ranges and engineering judgement to find an appropriate V_x for the

3. ground property (and compare with site-specific sample)
4. Select a normal or lognormal distribution (see e.g. **Table 18**)
5. Determine the parameters (mean and standard deviation) for the distribution.

Table 18. Suggested probability distributions for different soil and rock properties.

Ref. ¹	Ground property	Soil type or rock	Distribution (N = normal distribution; LN = lognormal distribution)
A	Cone resistance	Sand	LN
A		Clay	N/LN
A	Undrained shear strength s_u	Clay (triaxial tests)	LN
A, B		Clay (index tests)	LN
A		Clayey silt	N
A, B	Stress-normalized s_u	Clay	N/LN
A	Friction angle	Sand	N
A	Plastic limit	Clay	N
A, B	Sub-merged γ	All soils	N
B	Natural water content w_n	Clay	N
A, B	Void ratio, porosity	All soils	N
A, B	OCR	Clay	N/LN
B, B	Compressibility	Clay	N/LN
C	Hydraulic conductivity	Rock mass	LN
C	Rock strength	Rock mass	LN

(¹) References: Ref A = Uzielli et al. (2006), Lacasse and Nadim (1996); Ref B: Lofman (2022); Ref C = Stille et al. (2003).

Source: Authors' own work

If the lognormal distribution is chosen, the parameters can be calculated from the site-specific mean and estimated V_x or standard deviation using the following transformations. If the mean μ_x and standard deviation σ_x are known, the corresponding lognormal parameters (i.e. of the lognormal-transformed values), $\mu_{\ln x}$ and $\sigma_{\ln x}$, can be determined from:

$$\sigma_{\ln x}^2 = \ln \left(1 + \frac{\sigma_x^2}{\mu_x^2} \right) = \ln(1 + V_x^2)$$

Equation 35.

$$\mu_{\ln x} = \ln(\mu_x) - \frac{1}{2} \sigma_{\ln x}^2$$

Equation 36.

Reversely, the coefficient of variation can be obtained from the standard deviation of the log-transformed values (independent of the mean):

$$V_X = \sqrt{\exp(\sigma_{\ln X}^2) - 1}$$

Equation 37.

5.4.2.3 Approach C: Physical constraints and engineering judgement

Engineering judgement and prior knowledge on the limits of the ground property values may be used in choosing and defining probability distributions. For example, to exclude negative values, lognormal or truncated normal distributions may be chosen. Since normal distribution can produce negative values, using lognormal is recommendable if the $V_{x,tot}$ exceeds 0.30 in case the value is known to be non-negative. The amount of information that is available for subjective assessment of a probability distribution should be considered (see **Table 19**).

Table 19. Needed information and constraints to assign certain pdfs.

Constraint	Assigned pdf
Upper bound, lower bound	Uniform
Minimum, maximum, mode	Triangular
Mean, standard deviation	Normal
Range, mean, standard deviation	Beta
Mean occurrence rate	Poisson

Source: Mishra 2002

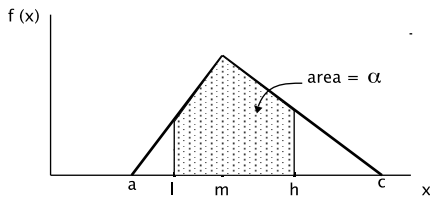
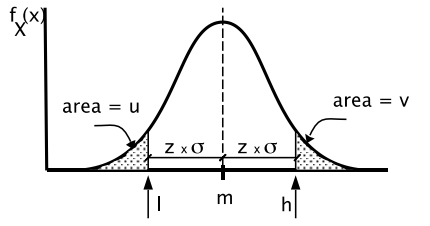
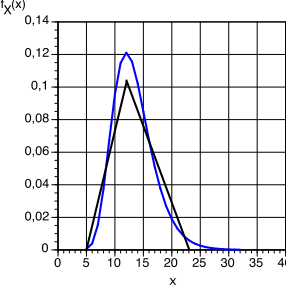
Most sophisticated expert opinion procedures are time-consuming, whereas the three-point estimate -method can be used by a small group or an individual engineer while still minimizing biases. In this method, the assessor gives estimates for low, high and most probable value, and the following principles should be followed in order to minimize the possibility of biases (SGF 2022):

1. The most probable value should be assessed last in order to avoid anchoring bias;
2. The assessor should not try to evaluate the absolute smallest and largest values, but instead give low and a high value that with a small probability might be exceeded, for example the 5% and 95% fractiles. The absolute values are calculated based on the fractiles (and the assessed most probable value).

Some three-point estimate methods are listed in **Table 20**.

Table 20. Three-point estimate methods based on engineering judgement.

pdf	Description	Illustrative figure
Triangular – Approach 1	The assessor is willing to state the probabilities of exceedance, i.e. define which probabilities are used. This case is described by Hudak (1994) (note that there is an error in the paper) and by Kotz & van Dorp (2004).	<p>Source: SGF (2022)</p>

pdf	Description	Illustrative figure
Triangular – Approach 2	The assessor assesses the probability that the true value falls between the low and the high value, see Garvey et al. (2016).	 <p>Source: SGF (2022)</p>
Normal distribution	A normal distribution can be assessed from one percentile and the mode. The mean is equal to the mode (normal distribution is symmetric). The standard deviation can be calculated using a table of z-values.	 <p>Source: SGF (2022)</p>
Lognormal distribution	<p>A lognormal distribution can be assessed by first assessing a triangular distribution, see Nederlof (2022) and Chang & Ko (2017). The mean μ_{lnX} and variance σ_{lnX}^2 of the lognormal distribution can be calculated from the mean μ and the mode m of the triangular distribution:</p> $\mu_{lnX} = \frac{\ln(m) + 2 \ln(\mu)}{3}$ $\sigma_{lnX}^2 = \frac{2(\ln(\mu) - \ln(m))}{3}$	 <p>Source: SGF (2022)</p>

Source: Authors' own work

Alternatively, any sort of pdf can be defined by means of the cumulative distribution method (e.g., Baecher 2019) may be used to plot a distribution based on engineering judgement. In the 'fixed value approach', the expert is given a ground property value and asked what is the probability that the true value is less than that value. These value-probability points are connected by a curve that is the approximation of the cumulative distribution function (CDF).

On the other hand, if there is no knowledge on which values are more frequent than others, the uniform distribution may be chosen which is defined by the minimum and maximum value only, while all the value between are equally likely to be true.

5.4.2.4 Approach D: Bayesian methods

The use of Bayesian methods is discussed in Chapter 7. The selection of probability distribution can favor conjugate distributions if Bayesian methods are to be used.

5.4.2.5 Approach E: Prescribed distribution types

The type of distribution can be prescribed in codes and regulations.

Box 13. Example 5.6: Triangular and lognormal distributions based on estimated constraints

A settlement calculation for embankment on clay requires pdf for coefficient of consolidation c_v (m^2/a). Let's assume that one oedometer test is available from a nearby site, but no site-specific data is available. The nearby site's data and subjective assessment are used to estimate the constraints to define pdf (triangular and lognormal).

Step 1: Triangular pdf

Triangular pdf (parameters minimum a , mode m , and maximum c) is defined by estimating the mode and the estimated low and high values which define the range where the true value is with 95 % probability ($\alpha=0.95$).

$$m = 1 \text{ m}^2/a \text{ (estimate of mode, based on nearby site's data and subjective assessment)}$$

$$l = 0.25 \text{ m}^2/a \text{ (estimate of the low value, based on subjective assessment)}$$

$$h = 4 \text{ m}^2/a \text{ (estimate of the high value, based on subjective assessment)}$$

$$a = m - \frac{m-l}{1-\sqrt{1-\alpha}} = 0.03 \text{ m}^2/a \qquad c = m + \frac{h-m}{1-\sqrt{1-\alpha}} = 4,9 \text{ m}^2/a$$

Step 2: Lognormal pdf from triangular pdf

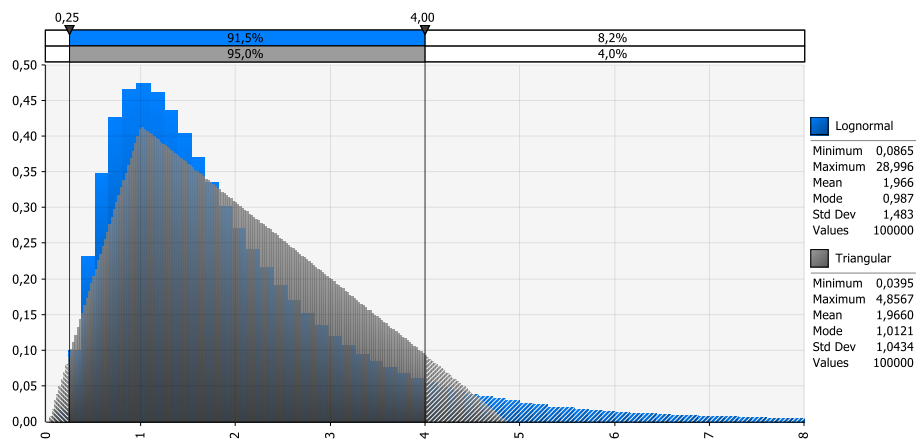
The estimated triangular pdf is then transformed into lognormal pdf. (The mean of the triangular pdf is $\mu = (a+m+c)/3$). The mean and variance of the lognormal pdf are given by:

$$\mu_{\ln X} = \frac{\ln(m)+2 \ln(\mu)}{3} = 0.45$$

$$\sigma_{\ln X}^2 = \frac{2(\ln(\mu) - \ln(m))}{3} = 0.45$$

Figure 26 below shows the histograms simulated using with these triangular and lognormal pdfs.

Figure 26. Histograms simulated using triangular and lognormal pdfs.



Source: Authors' own work

5.4.3 Cross-correlation between ground properties

Cross-correlation refers to the probabilistic dependency between different ground properties. If cross-correlation is ignored, the reliability index may be over- or underestimated, depending on the limit state considered. For the same ground property pair, the magnitude of the cross-correlation might vary depending on the dataset being used (e.g. site-specific or global data).

The magnitude of cross-correlation between two random variables can be described with the correlation coefficient ρ using the categories suggested by Evans (1996) (ISSMGE-TC304, 2021):

- Very strong when $|\rho| \geq 0.8$
- Strong when $0.6 \leq |\rho| < 0.8$
- Moderate when $0.4 \leq |\rho| < 0.6$
- Weak when $0.2 \leq |\rho| < 0.4$
- Very weak when $|\rho| < 0.2$

The level of correlation that becomes practically relevant depends on how sensitive the system response is to the dependence between ground properties. In most cases, it is safe to assume that weak to very weak cross-correlations ($|\rho| < 0.4$) may be ignored for the sake of simplicity. The relevance of moderate to very strong cross-correlations can be investigated via sensitivity studies.

In general, the cross-correlation between ground properties is estimated by taking samples that can be assumed to be from the same population and statistically comparing them by pairs (Fenton and Griffiths, 2008). The correlation coefficient ρ can then be estimated from the observations (data pairs). The number of data pairs should be at least ten (ISSMGE-TC304, 2021). The most common formulation for sample ρ is the Pearson product-moment correlation coefficient (often denoted by r), which measures linear correlation. In the case of non-linear correlation, rank correlation coefficients (like Spearman or Kendall's tau correlation coefficients) are more appropriate.

One of the most well-known cross-correlations is the negative correlation between cohesion c' and friction angle ϕ' . Values of ρ ranging from -0.24 to -0.70 have been reported by various researchers (based on review by Forrest and Orr, 2010).

A literature summary of site-specific correlation coefficients is presented in the TC304 (2021) report. For non-site-specific cross-correlation coefficients, a wide range of ρ may imply large variability between sites, and more ground investigation efforts could then be directed to estimate the site-specific ρ , if deemed essential.

Table 21 provides cross-correlation coefficients of some common soil property values. The coefficients have been mostly extracted from the ISSMGE-TC304 report (2021), where more correlations can be found. The ranges are representative of the averaged 25%- and 75% percentile of the observed correlation coefficients (Pearson product-moment correlation coefficient and the Spearman rank correlation).

Table 21. Recommended indicative cross-correlation coefficients. Parameters: c' = effective cohesion; C_c = compression index; C_s = swelling index; D_r = relative density; K_{DMT} = dilatometer horizontal stress index; K_0 = earth pressure at rest; K_s = saturated hydraulic conductivity; L_L = liquid limit; M_n = normalized effective constrained modulus determined by oedometer; P_i = plasticity index; w_n = gravimetric water content; φ = porosity; ϕ' = effective friction angle

Parameters		Recommended value (-)	Range ¹		Soil	Remarks
w_n	C_c	0,7	0,3	0,9	cohesive	From TC304 (2021)
LL		0,7	0,4	0,9	cohesive	From TC304 (2021)
PI		0,55	0,3	0,75	cohesive	From TC304 (2021)
w_n	C_s	0,35	0,1	0,6	cohesive	From TC304 (2021)

Parameters		Recommended value (-)	Range ¹		Soil	Remarks
LL		0,4	0	0,6	cohesive	From TC304 (2021)
PI		0,4	0	0,9	cohesive	From TC304 (2021)
c'	ϕ' $\tan(\phi')$	-0,3	-0,6	0,0	cohesive / non-cohesive	From Arnold (2016)
ϕ	K_s	0,5	0,0	0,7	cohesive / non-cohesive	From Arnold (2016)
D_r	M_n	0,3	0,1	0,5	non-cohesive	From TC304 (2021)
K_0		0,75	0,6	0,9	non-cohesive	From TC304 (2021)
K_{DMT}		0,7	0,6	0,8	non-cohesive	From TC304 (2021)

(¹) range is representative of the averaged 25%- and 75% percentile of the observed correlation coefficients

Source: Authors' own work

5.4.4 Depth-dependency

When a ground property exhibits significant depth-dependency, regression analysis can be applied to explicitly model the trend of the mean ground property as a function of depth. In that case, the standard deviation of the ground property values around the trend (i.e. error residuals) may be assumed to be either constant (standard deviation constant with depth) or variable (coefficient of variation constant with depth).

Also uncertainty in the mean stemming from statistical uncertainty in the linear regression trend can be modelled as a function of depth, see **Equation 29** in 5.3.6.

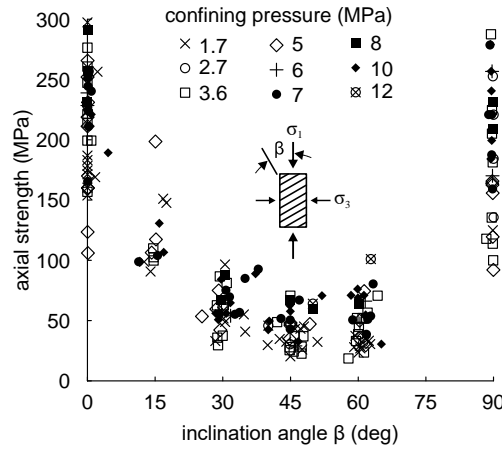
The principles described for depth-dependent ground properties can be applied to other variables too, such as stress. Example 5.7 illustrates how linear regression is used to characterize the uncertainty in rock strength as a function of confining pressure.

Box 14. Example 5.7: Defining distribution parameters for linear regression

Background

This example is adapted from Bozorgzadeh and Harrison (2014), who analyzed a data set of triaxial strength of Delabole slate (Brown et al. 1977), a transversely anisotropic rock. The strength of such rocks with planes of anisotropy (e.g. foliation, cleavage, schistosity or sedimentary structures) is well-understood to vary with the orientation of the applied major principal stress relative to the anisotropic structure as shown in Fig. 1. At $\beta = 0^\circ$, the failure mechanism is dominated by fracturing across the cleavage and is due to strain localization in the rock matrix. At cleavage inclinations between 45° and 60° , shearing along a single or small number of anisotropy planes dominates. At inclinations close to 90° , splitting of one or more cleavage planes dominates. A hybrid of the cross- and along-cleavage mechanisms occurs at all other orientations. Here, a subset of data shown in **Figure 27** with nominal $\beta = 90^\circ$ are analyzed in isolation using simple linear regression, focusing on quantifying and communicating uncertainties (statistical uncertainty in particular) and variabilities.

Figure 27. Delabole slate strength data (from Bozorgzadeh and Harrison, 2014).



Source: Bozorgzadeh and Harrison, 2014

Simple linear regression

Figure 28 shows the axial strength data vs. confining pressure for Delabole slate data with nominal $\beta = 90^\circ$ along with the fitted simple linear regression model which is discussed below.

For $i = 1, 2, \dots, n$ pairs of observations $(\sigma_{3[i]}, \sigma_{1[i]})$, a simple linear regression model is formulated as:

$$\sigma_{1[i]} = \alpha_0 + \alpha_1 \times \sigma_{3[i]} + \varepsilon_{[i]}$$

where σ_1 is the axial strength which is the response variable in the regression model, σ_3 is the confining pressure which is the predictor (also known as covariate or explanatory variable) in the regression model, and α_0 and α_1 are the intercept and slope, which are the parameters of the regression model to be estimated from the data; $\mu_{\sigma_1} = \alpha_0 + \alpha_1 \times \sigma_{3[i]}$ is the *mean* σ_1 when $\sigma_3 = \sigma_{3[i]}$. The error residuals $\varepsilon_{[i]}$ are assumed to be independent and normally distributed with zero mean and unknown variance σ^2 . The least-squares estimates of the slope and intercept are:

$$a_1 = \frac{\sum_{i=1}^n (\sigma_{3[i]} - \bar{\sigma}_3)(\sigma_{1[i]} - \bar{\sigma}_1)}{\sum_{i=1}^n (\sigma_{3[i]} - \bar{\sigma}_3)^2} \quad \text{and} \quad a_0 = \bar{\sigma}_1 - a_1 \bar{\sigma}_3$$

Note the change in notation, from " α " (unknown population parameter) to " a " (least-squares estimate, which is subject to statistical uncertainty). The estimate of σ^2 is the average squared residual:

$$s^2 = \frac{\sum_{i=1}^n (\sigma_{1[i]} - \hat{\sigma}_{1[i]})^2}{n-2}$$

where $\hat{\sigma}_{1[i]}$ is the estimated axial strength at $\sigma_{3[i]}$. The below table summarizes the least-squares parameter estimates and the estimated standard errors (SEs) of the estimated parameters. The SE is the standard deviation of the sampling distributions of the parameter and can informally be thought of as a measure of statistical uncertainty in the estimates. Estimates of the SEs of the slope and intercept are obtained as:

$$SE_{a_1} = \frac{s}{\sqrt{\sum_{i=1}^n (\sigma_{3[i]} - \bar{\sigma}_3)^2}}$$

$$SE_{a_0} = s \sqrt{\frac{1}{n} + \frac{\bar{\sigma}_3^2}{\sum_{i=1}^n (\sigma_{3[i]} - \bar{\sigma}_3)^2}}$$

The SEs can be used to construct confidence intervals (CIs) for the slope and intercept:

$$a_1 \pm t \cdot SE_{a_1} \quad , \quad a_0 \pm t \cdot SE_{a_1}$$

Here, t is the critical value for the t -distribution with $(n - 2)$ degrees of freedom and a specified confidence level. For a 95% confidence interval and 34 observations, $t = -2.034$; the resulting 95% confidence intervals are reported in the table. These intervals can be thought of as margins of error for the estimated parameters. It should be noted that there are many computer packages included in freely available programming languages such as R, Python and Octave, as well as the LINEST() function in Excel that return least-squares parameter estimates, SEs and confidence intervals.

Parameter	Point estimate _(SE)	95% confidence interval
$\hat{\alpha}_0$	133.61 _(21.39)	(90.04, 177.19)
$\hat{\alpha}_1$	10.10 _(3.40)	(3.17, 17.04)
\hat{s}	50.35	-

Figure 28 depicts the fitted model. At any specific value of σ_3^* , the estimated *mean* axial strength is $\hat{\mu}_{\sigma_1} = a_0 + a_1 \times \sigma_3^*$. This is shown as the solid line. The two dashed lines mark the 95% CI for μ_{σ_1} and add a margin of error to the estimate. A confidence interval for mean axial strength (i.e. mean response variable) at a given value of σ_3^* (i.e. the predictor) is constructed as:

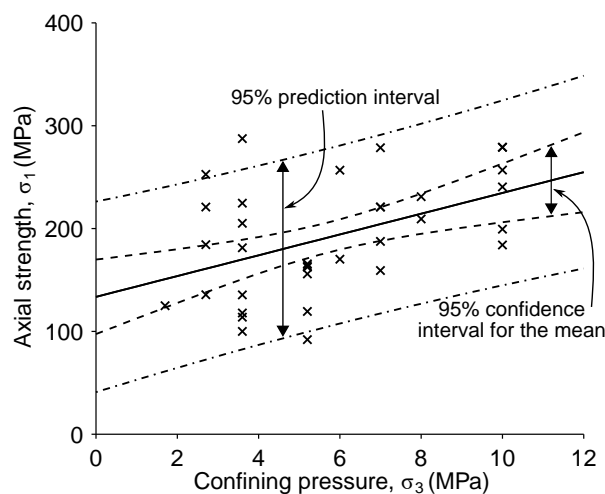
$$\hat{\mu}_{\sigma_1} \pm t \cdot SE_{\hat{\mu}_{\sigma_1}} \quad \text{where} \quad SE_{\hat{\mu}_{\sigma_1}} = s \sqrt{\frac{1}{n} + \frac{(\sigma_3^* - \bar{\sigma}_3)^2}{\sum_{i=1}^n (\sigma_{3[i]} - \bar{\sigma}_3)^2}}$$

The critical t value is the same as before. It is emphasised again that the above confidence interval is a margin of error for the estimated mean response and effectively reflects uncertainty in estimated slope and intercept. The regression model could also be used to predict a future value of axial strength $\hat{\sigma}_1$ at a given confining pressure σ_3^* . The margin of error for this prediction is the prediction interval (PI) and calculated as:

$$\hat{\sigma}_1 \pm t \cdot SE_{\hat{\sigma}_1} \quad \text{where} \quad SE_{\hat{\sigma}_1} = s \sqrt{1 + \frac{1}{n} + \frac{(\sigma_3^* - \bar{\sigma}_3)^2}{\sum_{i=1}^n (\sigma_{3[i]} - \bar{\sigma}_3)^2}}$$

The PI accounts for the additional variation of the individual observations about the mean; this is shown as the dashed-dotted lines in **Figure 28**.

Figure 28. Simple linear regression model fitted to the subset of data with $\beta = 90^\circ$



Source: Bozorgzadeh and Harrison, 2014

5.5 Model uncertainty

Calculation model uncertainty is related to imperfections and idealizations made in applied engineering models for representation and prediction of quantities such as pile resistance and foundation capacity (DNV 2021). Calculation models may be biased (e.g. conservative on average) and marked with scatter in the relationship between predicted response (e.g. calculated resistance) and the actual response (e.g. full-scale load test). In geotechnical engineering, calculation model uncertainties can be significant or even dominant in a reliability assessment. Therefore, it should always be included in the analysis.

5.5.1 Model factor (definition)

The model uncertainty can be quantified via random variable referred to as model factor M , which is the ratio of measured ('actual' or 'true') experimental value r_e and the theoretical value (i.e. quantity predicted by the model) r_t :

$$M = \frac{r_e}{r_t}$$

Equation 38.

Value r can be a quantity like load, resistance, or displacement. The mean of value $M_i (= r_{e,i} / r_{t,i})$ is the bias. A mean value equal to 1 means that the model is unbiased. The standard deviation or the coefficient of variation of M describes the variability of the model predictions. This statistical treatment of model uncertainty is analogous to the approach described in Annex D (Design assisted by testing) of EN 1990.

5.5.2 Estimation

The statistics for the model factor (i.e. model uncertainty) may be assessed on the basis of (Baker and Calle 2006):

- empirical or experimental data (such as the compilations by ISSMGE-TC304 (2021) and Tang and Phoon (2021)),
- comparison with more advanced computation models,
- in the case of lack of specific or explicit data, based on engineering judgement (expert opinion) (DNV 1997).

To decrease the uncertainty, full-scale testing at the site or the Observational Method (see section 8.3) may be considered.

5.5.3 Modelling

A simplified approach to considering the uncertainty in calculation model is to multiply the calculated response r_t with a random variable M (Baker and Calle 2006, ISSMGE-TC304 2020):

$$r = r_t \times M$$

Equation 39.

where M may be defined as a lognormal random variable with mean and CoV equal to the corresponding statistics (sample mean, i.e. the bias and sample CoV) calculated for the model

factors M_i (ISSMGE-TC304, 2020). The additive form of calculation model uncertainty may be used alternatively (similarly as for transformation uncertainty).

5.5.4 Recommended indicative values

Model factor statistics are scarce and expensive to obtain. Annex A provides an overview of available statistics, largely based on the state-of-the-art report by ISSMGE-TC304 (2020). **Table 22** provides indicative model factors for practical use in reliability assessments for the most important geotechnical problems and calculation models.

Table 22. Recommended indicative model factors for geotechnical computation models*

Type of problem	Type of calculation model	Expected value μ_M	Coefficient of variation V_M
Slope stability	LEM or 2D-FEM	1.0	0.10
Retaining walls (effects of actions)	Elastic/plastic spring supported beam (bending moments and shear forces)	1.0	0.10
Spread foundations (resistance)	Stability, analytical (e.g. Brinch Hansen)	1.0	0.15
	Settlements (SLS)	1.0	0.20
Piled foundations (resistance)	Model pile method (CPT-based)	1.0	0.15
	Other methods	1.0	0.15
Embankment settlement	Analytical consolidation and creep models	1.0	0.20

(*) The values in this table are judgement-based and largely taken over from the JCSS Probabilistic Model Code (Baker and Calle, 2006). Notice that the higher coefficients of variation for many geotechnical models as reported in Annex A are mostly the result not being able to separate pure model uncertainty from other variability or uncertainty in the measured and predicted values. In other words, the reported statistics seemingly over-estimate model uncertainty because other components of variability and uncertainty are still included.

Source: Authors' own work

These indicative values can be used in the absence of (better) model uncertainty data or estimates.

When using model with large model uncertainty ($CoV > 30\%$), it can make sense to work with model uncertainty only, and to ignore parametric uncertainty, because the contribution of the latter will be insignificant compared to the model uncertainty.

5.6 Actions

The reliability of geotechnical structures is also controlled by uncertainties of (external) actions (e.g., wind load or surcharge load) and the self-weight of construction materials such as concrete or steel.

Reliability analyses in the Eurocodes context are usually formulated as time-invariant reliability problems, as opposed to time-dependent ones. In time-dependent problems, variability in time, such as the variability in actions acting on the structure, is modeled explicitly (e.g., through time series).

In the time-invariant approach, variability in time is captured implicitly through probability distributions, such as extreme value distributions (see example on groundwater levels in 5.2.3). The probability distribution is then, for example, used to model the uncertainty in the maximum load in the reference period considered.

Typically, geotechnical reliability problems are dominated by the uncertainty in the resistance (e.g., ground properties), but the uncertainty in actions can be important. We distinguish three situations:

3. Uncertainty in actions insignificant: If the uncertainty in actions has no significant effect on the reliability estimate (and no probabilistic model is available), the design values can be used as (deterministic) point estimates (see 6.3.5).
4. Uncertainty in actions dominant: If the uncertainty in actions dominates the reliability estimates, the actions should be modeled probabilistically based on statistical analysis of data (including potential correlations).
5. Uncertainty in actions significant but not dominant: If the uncertainty in actions has a significant effect on the reliability estimates yet not a dominant one (i.e. some influence), the action models should be modeled probabilistically, ideally based on data. In the absence of suitable data, the probabilistic action models can be based on approximate probability distributions fit to the known design values (see example below).

For the probabilistic modelling of actions, both permanent and variable, guidance can be found from other publications, such as the JCSS Probabilistic Model Code, Part 2: Load Models (Joint Committee on Structural Safety).

Box 15. Example 5.8: Approximate probability distributions for actions based on design values

In this example we derive approximate probability distributions for the permanent actions (G) and the variable actions (Q) for a situation in which the design values are known but no (underlying) data are available. The design values are $G_d = 100$ kN and $Q_d = 50$ kN, respectively, both for consequence class CC2.

Since the relevant partial factors for CC2 are $\gamma_G = 1.00$ and $\gamma_Q = 1.30$ (for VC3), we can derive the characteristic values as:

$$G_k = \frac{G_d}{\gamma_G} = \frac{100}{1.00} = 100 \text{ kN} \quad Q_k = \frac{Q_d}{\gamma_Q} = \frac{50}{1.30} = 38.5 \text{ kN}$$

According to EN 1990 Annex C, typical quantiles (p) of the underlying probability distributions of the actions are:

- permanent actions: $p = 0.5$;
- time-variable actions: $p = 0.98$ (referring to the distribution of the yearly extreme values!).

Furthermore, we assume typical coefficients of variation of $V_G = 0.05$ and $V_Q = 0.15$, as for example discussed in European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al. (2024). Now, for the permanent action, we assume a normal distribution (uncertainty with small V), and with the data above we can determine the mean value and standard deviation as:

$$\mu_G = F_G^{-1}(0.5) = 100 \text{ kN} \quad \sigma_G = \mu_G \cdot V_G = 100 * 0.05 = 5 \text{ kN}$$

For the variable action, we assume a Gumbel distribution, since the variable action considers maximum values occurring in the specified reference period. For obtaining the parameters of the Gumbel distribution, we have to combine two steps, since the design value is given for a reference period of 50 years, whereas the quantile p is specified for yearly extreme values. Firstly, the exceedance probability (CDF) of Q_k for a Gumbel distribution (for annual maxima) is defined as:

$$p = P(Q < Q_k) = \exp\left(-\exp\left(-\frac{Q_k - a_1}{b}\right)\right)$$

with location parameter a_1 (subscript 1 for one year reference period) and scale parameter b . Furthermore, the relations between the Gumbel parameters and the distributions moments are:

$$\text{Mean and location parameter: } \mu_{Q,1} = a_1 + 0.577 \cdot b \leftrightarrow a_1 = \mu_{Q,1} - 0.577 \cdot b$$

$$\text{Standard deviation and scale parameter: } \sigma_Q = \frac{\pi b}{\sqrt{6}} \leftrightarrow b = \frac{\sqrt{6} \mu_{Q,1} V_Q}{\pi}$$

knowing that the coefficient of variation is $V_Q = \frac{\sigma_Q}{\mu_{Q,1}} = 0.15$.

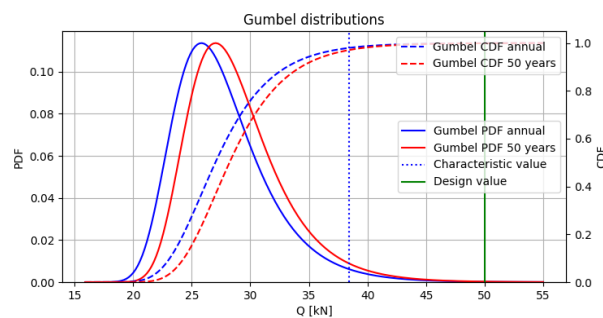
We can now find $\mu_{Q,1}$ (or a_1) iteratively or by optimization, and obtain for this case $\mu_{Q,1} = 27.7$ kN, which is the expected annual maximum value of Q .

The corresponding location parameter is $a_1 = 25.82$, which can be scaled to a reference period of n years by:

$$a_n = a_1 + \frac{\ln n}{b}$$

which results in $a_{50} = 27.03$ for 50 years. Notice that this scaling of reference periods is a special property of the Gumbel distribution, and results in a horizontal shift of the PDF and CDF, as illustrated below (**Figure 29**). As discussed, this approximate approach may be apt, when the actions are not dominant in the reliability problem.

Figure 29. Gumbel distributions with different reference periods and the design value.



Source: Authors' own work

6 Reliability analysis for geotechnical applications

6.1 Principles and definitions

6.1.1 Limit state and limit state function

In order to assess structural performance, structural responses are divided into desirable and undesirable states. The structural response is represented by a limit state function $g(x(t))$, which attains a positive value for desirable states and a negative value for undesirable states (EN 1990 Annex C.3). The boundary between these states is called the limit state and is defined as:

$$g(x(t)) = 0$$

Equation 40. EN 1990 Formula (C.1)

The domain of undesirable states or adverse (failure) events $\Omega(X(t))$, also called the failure domain, is given by

$$\Omega(X(t)) = \{g(X(t)) < 0\}$$

Equation 41. EN 1990 Formula (C.3)

where $x(t) \in X(t)$ is the vector of the random variables governing the exceedance of the limit state. These variables characterise the uncertainty as discussed in Chapter 5. As stated in section 5.6, most structural and geotechnical engineering problems are formulated as time-invariant (see 6.1.4). Therefore, in the remainder of the text we will mostly use x instead of $x(t)$, as the time-dependence is modelled implicitly in the variables.

Using these definitions, 'failure' entails any undesirable state, which can be ultimate limit state (ULS) or serviceability limit state (SLS), depending on the definition limit state function.

In the case where resistances R and effects of actions E are explicitly defined in the computational model, the limit state function can be written in terms of a safety margin $g(\mathbf{X}) = R - E$ or a factor of safety⁷ $g(\mathbf{X}) = R / E - 1$. Many of the geotechnical computational models do not explicitly define resistance and load terms, in which case the limit state function can be formulated by means of indirect variables. For example, when the strength reduction method is applied in a finite element analysis to determine a factor of safety, the resulting factor of safety FS can be used to formulate the limit state function as $g(\mathbf{X}) = FS(\mathbf{X}) - 1$ (i.e. factors of safety lower than one imply failure).

For serviceability limit states (SLS) the serviceability criterion can often be used to formulate the limit state function. For example, the limit state function for an acceptable displacement u_{acc} would be intuitively:

$$g(\mathbf{X}) = u_{acc} - u(\mathbf{X})$$

Equation 42.

⁷ Notice that in structural engineering the unity check $UC = E / R$ is commonly used, which is the inverse of the factor of safety commonly used in geotechnical engineering.

A limit state indicator $I_g(\mathbf{X})$ can be used to indicate whether a computational model results in a desirable or undesirable state:

$$I_g(\mathbf{X}) = \begin{cases} 1 & \text{if } g(\mathbf{X}) < 0, \quad \text{undesirable state (failure)} \\ 0 & \text{if } g(\mathbf{X}) \geq 0, \quad \text{desirable state (non – failure)} \end{cases}$$

Equation 43.

This formulation is convenient when the evaluation of the limit state results only (or easier) in failure versus non-failure rather than an exact value. For example, when the ULS is evaluated using a finite element model, loss of equilibrium implies failure and is sufficient to evaluate the limit state indicator $I_g(\mathbf{X})$, while the evaluation of the limit state function $g(x(t))$ may require the iterative shear strength reduction, which is computationally more demanding. Most of the sampling-based reliability methods (e.g. Monte Carlo simulation) only require failure versus non-failure information.

6.1.2 Random variables and joint probability distribution

In order to account for the uncertainties involved in the reliability analysis, they need to be captured in a joint probability density function $f_{\mathbf{X}}(x)$. This density function describes the probability density of the vector of the basic random variables $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$ using marginal distributions $f_{X_i}(x)$ for each individual component X_i and the correlation structure between the different components. The individual components X_i represent the individual random variables of the geotechnical structure such as material properties; the correlation describes any (partial) dependence between the individual components.

6.1.2.1 Continuous variables

For continuous marginal distributions of random variables, which are used to represent most uncertainties in geotechnical reliability problems, the uncertainty is modelled with probability density function (PDF) $f_{X_i}(x)$ and the cumulative distribution function (CDF) $F_{X_i}(x)$.

6.1.2.2 Discrete variables

In some cases, discrete distributions arise in addition to the continuous distributions. Relevant discrete distributions are those describing probabilities for subsurface scenarios (e.g. the presence of a weak or conductive thin layer). The reliability methods described below are mostly capable of dealing with continuous distributions. Therefore, discrete variables are often treated as scenarios. The total probability is then computed based on the conditional probabilities per scenario S_i , using the law of total probability:

$$P_f = \sum_{i=1}^N P_f(\mathbf{X}|S_i)P(S_i) = P(g(\mathbf{X}) < 0|S_i) \cdot P(S_i)$$

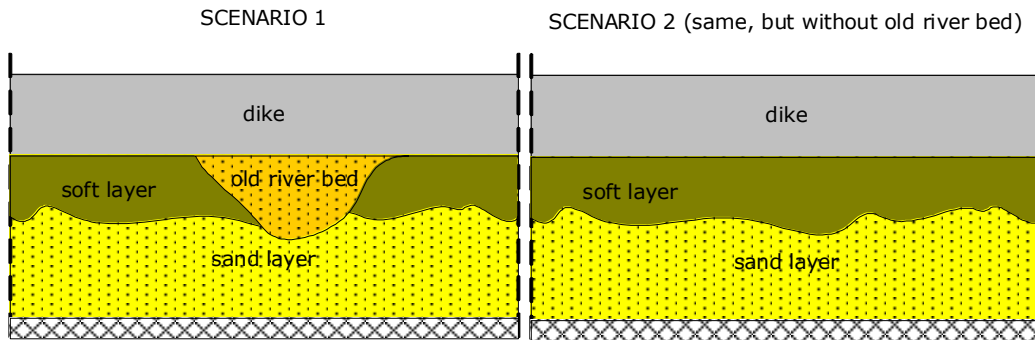
Equation 44.

Example 6.1 illustrates this treatment of discrete scenarios using the subsurface scenarios described in Example 5.2.

Box 16. Example 6.1: Total probability in stratification scenarios

Here we reconsider Example 5.2 with two possible stratification scenarios in the ground model (**Figure 30**):

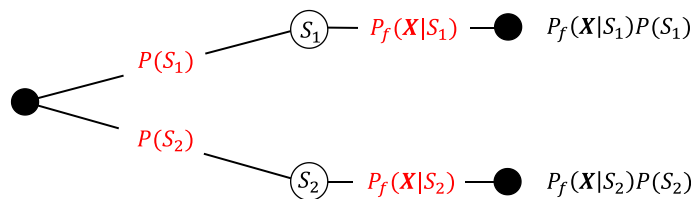
Figure 30. Example of two stratification scenarios.



Source: Authors' own work

Probabilities for scenarios 1 and 2 are denoted as $P(S_1)$ and $P(S_2)$; they are mutually exclusive and supposed to be exhaustive. The scenarios can be presented in a probability tree (i.e. event tree, **Figure 31**), defining a probability of failure for each branch:

Figure 31. Probability event tree to account for scenarios.



Source: Authors' own work

The total probability can then be calculated by summing the probabilities of each branch: $P_f(\mathbf{X}) = P(S_1) * P_f(\mathbf{X}|S_1) + P(S_2) * P_f(\mathbf{X}|S_2)$. The approach can be extended to include multiple discrete variables by sub-branching scenarios.

If there is correlation between the discrete and continuous variables, the continuous variables can be expressed in terms of conditional stochastic variables $\mathbf{X}_1 \sim f_{\mathbf{X}}(\mathbf{x}|S_1)$ and $\mathbf{X}_2 \sim f_{\mathbf{X}}(\mathbf{x}|S_2)$. However, for subsurface scenarios that is rarely the case.

6.1.2.3 Correlations

Ground properties or other (geotechnical) parameters may be (cross-)correlated (see 5.4.3). The most common method for modelling correlation in a structural reliability context is through an iso-probabilistic transformation (for details refer to the literature on Rosenblatt and Nataf transformations, e.g. Liu & Der Kiureghian (1986)). The Nataf transformation is illustrated in Example 6.2. The advantage of this transformation is that besides the marginal probability distributions only (Pearson) correlation coefficients are required as input for the cross-correlations. Most relevant software packages use this option. More sophisticated correlation models such as

copulas are uncommon and hardly justifiable given the sparse data on correlation in geotechnical applications.

Box 17. Example 6.2: Standard Normal (Nataf) transformation

This type of transformations is implemented and used in most of the relevant software packages for reliability analysis. As a user, you can consider this as background information, which is not necessary for practical application.

Consider a joint probability with marginals $X_1 \sim U(0,1)$ and $X_2 \sim LN(1,0.2)$ and a Pearson correlation coefficient $\rho = 0.6$. Samples in parameter-space can be generated from standard-normal random samples in two steps. First by correlating the standard normal random variables \mathbf{U} to create correlated standard normal variable \mathbf{V} , and secondly by applying the iso-probabilistic transformation of \mathbf{V} to introduce the marginal distributions for each variable:

$$\mathbf{V} = \mathbf{A}\mathbf{U}$$

$$\mathbf{X} = F_X^{-1}(\Phi(\mathbf{V}))$$

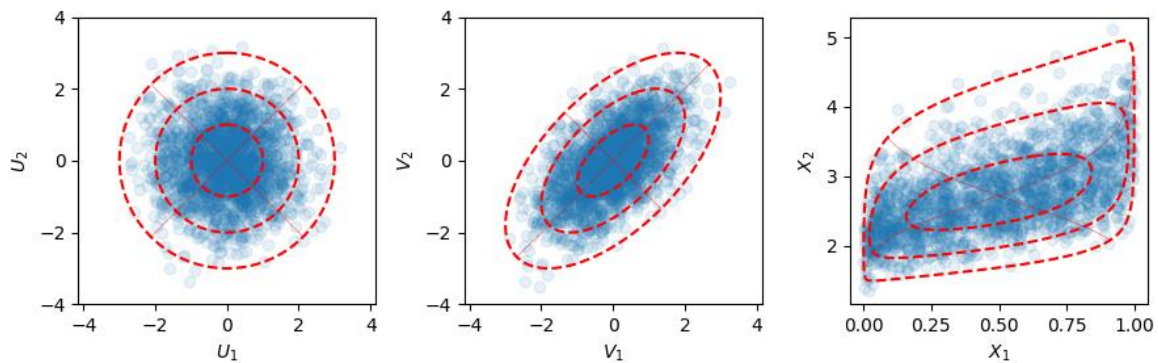
\mathbf{A} is some factorisation of the correlation matrix \mathbf{R} , such that $\mathbf{R} = \mathbf{A}\mathbf{A}^T$, such as the Cholesky decomposition. The components of the correlation matrix R_{ij} are the Pearson correlation coefficients ρ_{ij} between normalised stochastic variables V_i and V_j . Other decompositions (e.g. Eigen decomposition) can also be used resulting in an equivalent transformation, but with a different rotation in the transformation between U-space and V-space. The iso-probabilistic transformation is applied variable-by-variable and reads $X_i = F_{X_i}^{-1}(\Phi(V_i))$. Applied to the two marginal distributions, this results in:

$$X_1 = \Phi(V_1)$$

$$X_2 = \exp(1 + 0.2 * V_2)$$

The combination of the two steps can be referred to as transformation T , such that $\mathbf{X} = T(\mathbf{U})$. The transformation from parameter space (\mathbf{X}) to standard-normal space (\mathbf{U}) can be performed by the inverse of the two operations $\mathbf{U} = T^{-1}(\mathbf{X})$. In this way, the limit state function can be formulated in standard-normal space to facilitate probabilistic interpretation. **Figure 32** gives a graphical representation of the two transformation steps.

Figure 32. Step-wise transformation from U-space to parameter space by correlation (U to V) and iso-probabilistic transformation (V to X). Red dashed lines indicate the (mapping of) the contour lines of equal distance to the mean in standard normal space (i.e. β_{HL} -contours).



Source: Authors' own work

6.1.3 Probability of failure and reliability index

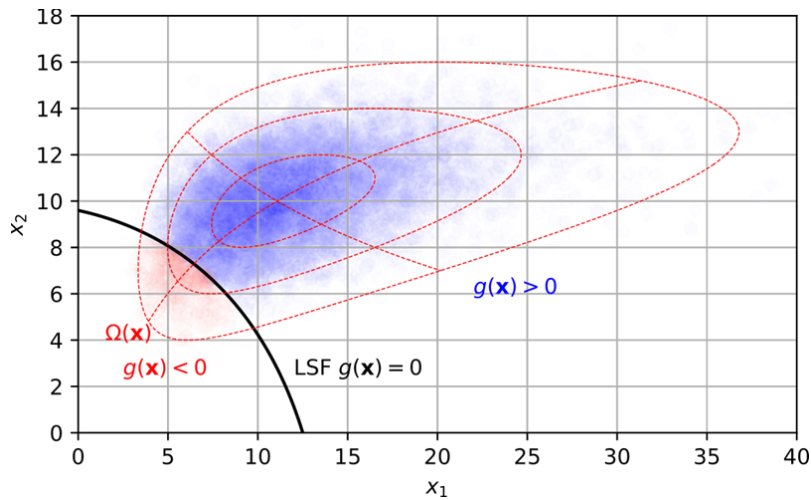
The probability of failure is defined as the probability of undesirable performance $P(g(\mathbf{X}) < 0)$, formulated consistently with 6.1.2 as:

$$P_f = \int_{\Omega(x)} f_X(x) dx$$

Equation 45. EN 1990 Formula (C.4)

where $\Omega(x)$ is the domain of undesirable events where $g(x) < 0$ and $f_X(x)$ is the joint probability density function for vector of random variables \mathbf{X} (see **Figure 33**).

Figure 33. Joint distribution $f_X(x)$, limit state surface $g(x) = 0$ and failure domain $\Omega(x)$.



Source: Authors' own work

The reliability index β and the probability of failure P_f are used as standard metrics to express structural reliability. The functional relationship between the failure probability and the reliability index is given as:

$$P_f = \Phi(-\beta)$$

Equation 46. EN 1990 Formula (C.5)

with $\Phi(\cdot)$ being the standard normal cumulative distribution function. **Table 23** shows some example combinations of β and P_f .

Table 23. Relationship of reliability index β and failure probability P_f illustrated with numerical examples

Reliability index β	Probability of failure P_f
1.0	1.5×10^{-1}
2.0	2.3×10^{-2}
3.0	1.4×10^{-3}
4.0	3.2×10^{-5}

Source: Authors' own work

6.1.4 Time-dependent variables

Most of the ground model and geotechnical parameters are time-invariant, meaning that the probability distributions model the lack of knowledge about a true value (i.e. epistemic uncertainty) that does not change in time. Therefore, geotechnical reliability analyses are generally performed as time-invariant. In a time-invariant reliability analysis, time-dependent variables (e.g. for external actions or ground water levels) can be included consistently within the reliability-based approach by use of extreme value distributions in the following cases:

- only a single time-dependent variable exists;
- multiple time-dependent variables exist with full correlation between them, such that the extreme events can be expected to occur simultaneously, or can be linked to a single time-dependent latent variable.

Accounting for multiple time-dependent variables requires simplifying assumptions or more advanced approaches accounting for the dependence between the variables. See JCSS (2001) for more details and practical approaches.

When extreme value distributions are used, the corresponding random variables often represent the (annual) extreme values, in which case the calculated probability of failure has a reference period of one year (denoted as β_a and $P_{f,a}$). If the reliability requirements are formulated with a different reference period, for example 50 years to represent the lifetime of the structure (see 2.4.3), the extreme value distributions will need to refer to the 50 year maxima or minima. The corresponding probabilities of failure or reliability indices may be referred to as β_{50} and the annual probability of failure $P_{f,50}$ respectively. An alternative approach to evaluate the reliability for longer periods is to work out and sum conditional probabilities over a series of shorter time intervals. This allows for explicit differentiation in the variable distributions over time and can give a more complete insight into the probability of failure over time (see 2.3.6).

6.1.5 System reliability aspects

In the Eurocodes context, reliability analyses will mostly be carried out for individual failure modes (of individual structural members, if applicable), because reliability targets are formulated (or interpreted) at that level (see 0). Hence, system reliability analysis is rarely required.

When system reliability does need to be evaluated and the system performance is not already formulated in a single limit state function, the components of the system can be identified as individual members (e.g. multiple piles as part of a single foundation) or failure modes (e.g. sliding or overturning of a gravity retaining wall).

When evaluating the reliability of a system with multiple components, an important aspect to be accounted for is the re-distribution of loads over the different members (redundancy). This redistribution can take place before as well as after the failure of a single member and is strongly linked to the type of failure (yielding, brittle failure) of the individual members.

When evaluating the reliability of a system with multiple failure modes (for which individual limit states can be defined), the correlation between the failure modes is an important aspect.

Some approaches to combine the individual reliability of different components are:

- If practically feasible, the structural performance as a whole, involving all (interacting) members and failure modes, is modelled in a single model. In this way, the exceedance of any of the involved limit states directly provides the system reliability.
- As a conservative estimate of the system reliability, failure probabilities of the individual failure modes can be summed (for a series system). This forms a straightforward and relatively accurate approach for systems with only few, distinct and independent (low-correlated) failure modes.
- An alternative approach in case of correlation between failure modes is to account for this correlation (i.e. the overlap of the failure domains of the individual failure modes) in the summation of the component failure probabilities. A possible approach based on FORM design points (see) is given by Hohenbichler and Rackwitz (1982).

A more extensive discussion on system reliability can be found in Der Kiureghian (2022). Many software packages already include capabilities to combine limit states in a system reliability analysis.

6.2 Reliability analysis methods

This section gives an overview and basic descriptions of the reliability methods most commonly used in practice, to give a basic understanding to practitioners. The methods are implemented in most relevant software packages and explained in more detail in their documentation. Implementations, documentation and examples of open-source and/or free software are available here⁸:

- COSSAN (<https://cossan.co.uk/>)
- OpenTURNS (<https://openturns.github.io/www/>)
- Probabilistic Toolkit (<https://www.deltares.nl/en/software/probabilistic-toolkit-ptk/>)
- UQLab (<https://www.uqlab.com/>)
- UQPy (<https://uqpyproject.readthedocs.io>).

More in-depth Algorithmic details, examples and evaluation of reliability methods can be found in the literature, such as Der Kiureghian (2022), Ditlevsen & Madsen (1996) or Phoon & Ching (2017).

6.2.1 Basic notions

Starting from the structural performance described by a limit state function $g(x)$, defined such that $g(x) = 0$ represents the limit state, the probability of failure P_f , or the reliability index β , is to be evaluated. Example 6.3 deals with this exercise for a simple two-variable problem.

⁸ The list of software contains packages known to be useful to the authors at the time of writing which are either free or open-source. Commercial packages are also available but not listed here to avoid generating competitive (dis)advantages. The list is in alphabetical order, not in any order of preference.

In practice, limit state functions are often non-linear and random variables may be non-normal or cross-correlated. The two most widely-used methods for reliability analysis are briefly described in the sections below, namely:

- First-order second moment method (6.2.2)
- Point estimate methods (6.2.3)
- First-order reliability method (6.2.4)
- Monte Carlo simulation (6.2.5)
- Other sampling techniques (6.2.6).

Response surface methods and metamodels or surrogate models will be briefly discussed in Section 6.2.7, since they have shown to potentially overcome accuracy, robustness and efficiency problems of the conventional methods.

Box 18. Example 6.3: Reliability index for a simple limit state (analytical solution)

Suppose that a safety margin, given by the difference between resistance and load effects ($g = M = R - S$), describes the structural performance of a retaining wall, a pile, or other geotechnical structure. In that case, the mean value μ_M and variance σ_M^2 are, respectively given by

$$\mu_M = \mu_R - \mu_S$$

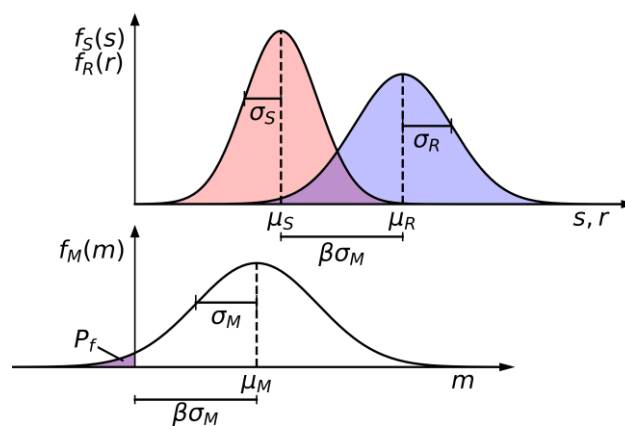
$$\sigma_M^2 = \sigma_R^2 + \sigma_S^2 - 2 \cdot \rho_{RS} \cdot \sigma_R \cdot \sigma_S$$

where ρ_{RS} is the coefficient of correlation between R and S . Assuming normal distributed load and resistance, the reliability index β is estimated by (Cornell, 1969)

$$\beta = \frac{\mu_M}{\sigma_M}$$

Figure 34 gives a graphical interpretation of the reliability index when evaluated in this way in the case where loads S and resistances R are uncorrelated. In this representation, the reliability index represents the number of standard deviations that the mean safety margin is removed from the limit state (i.e. failure).

Figure 34. Graphical representation of the reliability index for independent ($\rho_{RS} = 0$) and normal distributed load S and resistance R .



Source: Authors' own work

6.2.2 First-order second moment method (FOSM)

A simple option to approximate of the probability of failure is using the first-order second moment (FOSM) method. This method only requires information related to the second moment statistics (mean value and covariance) of the random variables, while approximating the limit state function by the first order term of the corresponding Taylor series expansion around its mean, i.e.,

$$g(x_1, \dots, x_n) \approx g(\mu_{x_1}, \dots, \mu_{x_n}) + \sum_{i=1}^n (x_i - \mu_{x_i}) \cdot \frac{\partial g}{\partial x_i}$$

Equation 47.

The evaluation of the partial derivatives of the limit state function can be done analytically (for example in the case of analytical limit state functions) or by a finite difference approximation. Depending on the type of finite difference approximation used, the computational model is evaluated $n+1$ times (in case of forward difference or backward difference) or $2n+1$ (in the case of central difference). The applied perturbation Δx_i in the finite difference approximation needs to be large enough to avoid numerical instabilities in case of numerical modelling result. For FOSM, $\Delta x_i = 0.1\sigma_{X_i}$ is a reasonable starting point. Larger perturbations can lead to better approximations, under the condition that the direction of perturbation (forward-difference or backward-difference) is chosen towards the limit state. In this way, part of the potential non-linearities in the limit state functions are accounted for in the differentiation. Mean value μ_g and variance of the limit state function can be directly approximated by

$$\mu_g \approx g(\mu_{x_1}, \dots, \mu_{x_n})$$

Equation 48.

$$\sigma_g^2 \approx \sum_{i=1}^n \sum_{j=1}^n \rho_{X_i X_j} \cdot \sigma_{X_i} \cdot \sigma_{X_j} \cdot \frac{\partial g}{\partial X_i} \cdot \frac{\partial g}{\partial X_j}$$

Equation 49.

The reliability index and the probability of failure can then be estimated using $\beta = \mu_g / \sigma_g$. Due to its simplicity, the method can be easily implemented in spreadsheets or scripts. Note that **Equation 49** simplifies to the following form for $\Delta x_i = \sigma_{X_i}$ (and assuming independence between the random variables):

$$\sigma_g^2 \approx \sum_{i=1}^n (g(\mu_{x_1}, \dots, \mu_{x_i} - \sigma_{X_i}, \dots, \mu_{x_n}) - \mu_g)^2$$

Equation 50.

FOSM leads to exact results for linear limit state functions with normal distributed, independent random variables. In general, however, it only provides a rough approximation of the probability of failure, because this method ignores important information on the random variables, such as their distribution, and introduces significant errors in case of non-linear limit state functions.

Improvements to FOSM proposed in the literature like the second order second moment method (Baecher and Christian, 2003) and the quantile-based first order second moment method (Yang &

Ching 2019) have advantages under certain conditions, but have not found wide use in geotechnical applications.

6.2.3 Point estimate method (PEM)

The point estimate method (PEM) is based on the estimation of the central moments (mean and standard deviation) of the limit state function $g(\mathbf{X})$, i.e. the safety margin, to estimate the probability of failure:

$$\mu_g \approx \sum_{i=1..N} w_i \cdot g(\mathbf{x}_i)$$

Equation 51.

$$\sigma_g^2 \approx \sum_{i=1..N} w_i \cdot (g(\mathbf{x}_i) - \mu_g)^2$$

Equation 52.

with \mathbf{x}_i being the N points at which the limit state function is evaluated and w_i being the weights by which $g(\mathbf{x}_i)$ contributes. The optimal choice for the combination of points \mathbf{x}_i and weights w_i follows the principles of Gaussian quadrature, and depends on the distribution of \mathbf{X} and the desired accuracy in the estimation of the moments. Recommendations can be found in Baecher & Christian (2003). Having obtained the mean μ_g and variance σ_g^2 of the LSF-distribution, we can estimate the reliability index by $\beta = \mu_g / \sigma_g$.

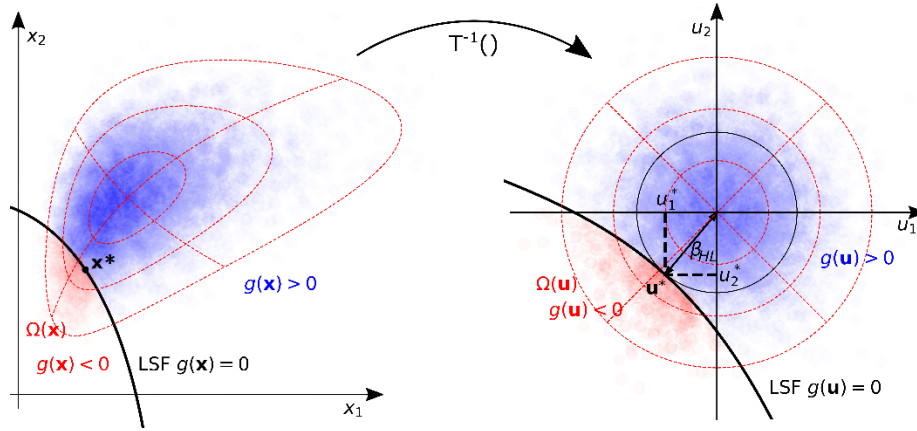
When using only the first two moments (mean and variance), two points \mathbf{x}_i are needed for one stochastic variable. For larger numbers of variables, the number of points for which the limit state function needs to be evaluated increases rapidly. For example, when $n = 10$, a total of $2^{10} = 1024$ evaluations are needed. Another limitation of the PEM based on two central moments, is that it only yields exact results for linear limit states. This leads to significant errors in the calculated reliability in the case of (strong) non-linearities of the limit state function (see Christian & Baecher, 1999).

Due to the limitations, the PEM is not frequently chosen for application in geotechnical problems. Its accuracy is similar to FOSM, while FOSM is easier to use for first approximations.

6.2.4 First order reliability method (FORM)

In the First-order reliability method (FORM) method, the reliability problem is formulated in the standard normal variable space, and linearized in the so-called design point instead of in the mean values. The design point (\mathbf{x}^*), which is the central concept in FORM, loosely speaking is the parameter combination for which the limit state is most likely to be exceeded. Or more rigorously, it is the point in parameter space on the limit state with the highest probability density (after transformation to standard normal space, see **Figure 35** and background in 6.1.2).

Figure 35. Transformation from physical space (left) to standard normal space (right) for a joint probability distribution with correlated parameters and non-linear limit state function. The definitions of design point $\mathbf{x}^* = T(\mathbf{u}^*)$, Hasofer–Lind reliability index β_{HL} , and ingredients for the importance factors u_i^* are given in standard normal space. The design point can be represented in parameter space as point \mathbf{x}^* .



Source: Authors' own work

In standard normal space, the Hasofer-Lind reliability index β_{HL} is defined as the minimum distance from the origin to the limit state (Hasofer and Lind, 1974), i.e.,

$$\beta_{HL} = \min_{g(\mathbf{u})=0} \sqrt{\mathbf{u}^T \cdot \mathbf{u}} = \|\mathbf{u}^*\|$$

Equation 53.

which can be geometrically interpreted as the number of standard deviations of the safety margin between expected performance and failure, in terms of the limit state function.

Another important element in the first order reliability method are the influence factors α , which are defined as:

$$\mathbf{u}^* = -\boldsymbol{\alpha} \cdot \beta_{HL}$$

Equation 54.

where $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_n\}$ is the unit length direction cosine vector, expressing the sensitivity of the standardized limit state to changes on each random variable. The influence factors provide valuable information on the FORM results due to their following properties:

- Influence factors lie in the range $-1 \leq \alpha_i \leq 1$. Negative values indicate an unfavorable influence of an increasing value of the variable in question (i.e. load); positive values indicate a favourable influence (i.e. resistance).
- The absolute value of an influence factor lies between zero and one, zero indicating no influence on the reliability and one indicating dominant influence.
- The sum of the squared influence factors equals one (i.e. $\sum \alpha_i^2 = 1$), implying that α_i^2 indicates the contribution of the uncertainty of a random variable to the probability of failure. For example, a variable with $\alpha_1 = 0.8$ has a contribution of 64%.

Due to the transformation from standard normal to physical space, the alpha factors for correlated parameters need to be treated with caution, especially in the presence of strong correlations.

FORM is available in many commercial and non-commercial software packages, most often in combination with detailed algorithmic documentation. An example of a FORM analysis, including discussion of influence factors is given in Section 6.4.3.

6.2.5 Monte Carlo simulation (MCS)

Monte Carlo simulation (MCS) is a sampling based method, which in its simplest form (Crude Monte Carlo) implies drawing random samples from the joint probability distribution $f_X(x)$ of the random variables involved (see 6.1.2), and evaluating the limit state function for those samples. The probability of failure is then estimated as the ratio of the number of realizations resulting in failure (N_f) and the total number of realizations (N):

$$P_f = \frac{N_f}{N}$$

Equation 55.

Noteworthy, design point estimates (as obtained in with FORM) can also be produced for Crude Monte Carlo simulation (and most other sampling techniques), aiding the interpretation of the results. Most software packages dedicated to (structural) reliability analysis comprehend this feature.

The accuracy of the estimate depends on the number of (failed) realizations, which can be expressed in terms of the coefficient of variation of the estimated probability of failure:

$$V_{P_f} = \sqrt{\frac{1 - P_f}{N \cdot P_f}}$$

Equation 56.

A coefficient of variation V_{P_f} of 5 to 10% is considered sufficiently accurate for estimating structural reliability. Hence, the V_{P_f} can be used as a convergence or stopping criterion for Crude Monte Carlo simulation and in addition forms the basis for the derivation of uncertainty bounds of the estimated values (see Section 6.3.4.2. **Equation 56** implies that the number of required samples can be estimated using an a-priori estimate of the probability of failure by

$$N_{req} \approx \frac{1}{V_{P_f}^2 \cdot P_f}$$

Equation 57.

For example, if the probability of failure is 10^{-4} (which corresponds to $\beta = 3.7$), the required number of samples for a sufficiently reliable answer would be one million (10^6) for a coefficient of variation of $V_{P_f} = 10\%$.

The implication is that Crude Monte Carlo simulation, while very attractive due to its accuracy, simplicity and robustness, is only tractable for limit state functions which can be evaluated with very low calculation times (e.g. closed-form, analytical formulae). The method is largely prohibitive for computationally expensive limit state functions such as problems involving finite element analysis. This efficiency problem is addressed by other sampling methods and surrogate models as discussed in the next sections.

6.2.6 Other sampling methods

Several variants of Monte Carlo simulation offer better efficiency than Crude Monte Carlo (i.e. fewer required evaluations of the limit state function), in some cases at the cost of accuracy or robustness. The list below contains widely used sampling methods:

- Latin hypercube sampling (LHS)
- Importance sampling (IS)
- Subset simulation (SS)
- Directional sampling (DS)
- Line sampling (LS).

The practical significance of these sampling methods is that they are more accurate and less prone to convergence problems compared to FORM, while being more efficient than Crude Monte Carlo. The main characteristics of the more widely used variance reducing sampling methods are given below.

6.2.6.1 Importance sampling (IS)

Importance sampling (IS) uses prior knowledge of the failure region (i.e. the parameter combinations leading to failure) to define an alternative joint PDF $f_{\mathbf{X}}^{IS}(x)$, the importance sampling distribution, which will be more efficient in generating samples around the limit state compared to the original joint PDF.

The probability of failure estimate is obtained analogously to Crude Monte Carlo by counting the failure samples, only for IS these are weighted to correct for being generated from the IS probability distribution instead of the original one:

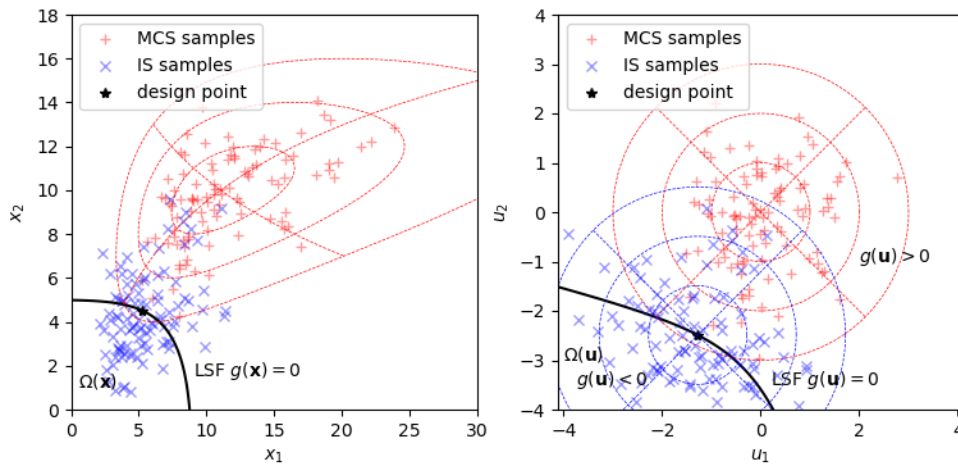
$$\hat{P}_f \approx \frac{1}{N} \sum_{i=1}^N w_i I_f[\mathbf{x}_i], \quad w_i = \frac{f_{\mathbf{X}}(\mathbf{x}_i)}{f_{\mathbf{X}}^{IS}(\mathbf{x}_i)}$$

Equation 58.

with $I_f[\mathbf{x}_i]$ being the indicator function returning the value 1 for failure with sample \mathbf{x}_i and zero for non-failure. Notice that for $f_{\mathbf{X}}^{IS}(\mathbf{x}_i) = f_{\mathbf{X}}(\mathbf{x}_i)$ this boils down to Crude Monte Carlo.

The importance sampling distribution can be chosen completely arbitrarily, but will be more effective if centred around the design point (see **Figure 36**), producing realizations practically equally on both sides of the limit state. Hence, the combination of first running the first order reliability method (FORM) to obtain an estimate of the design point, and subsequent importance sampling is an attractive option (e.g. to deal with non-linearity). Alternatively, an importance distribution can be defined using insight into the limit state function and the safety margin. The most important parameters are often easily identified based on experience, after which the importance sampling distribution can be shifted in the general direction of the failure domain. This can be done for a selection of the most important parameters. However, an incorrect design point as the centre of the importance distribution can lead to overestimation of the reliability or reduced efficiency (see Example 6.6). Updating the position of the importance distribution during the simulation can be an effective strategy to limit the impact of this problem. This method of improving the sampling strategy during simulation is known as adaptive importance sampling (AIS).

Figure 36. Illustration of importance sampling (IS) compared to Monte Carlo simulation (MCS) with the parameter space of a multivariate distribution (left) and the corresponding standard normal space (right). Compared to Monte Carlo simulation, importance sampling is very effective in sampling from the failure domain with the same total number of samples.

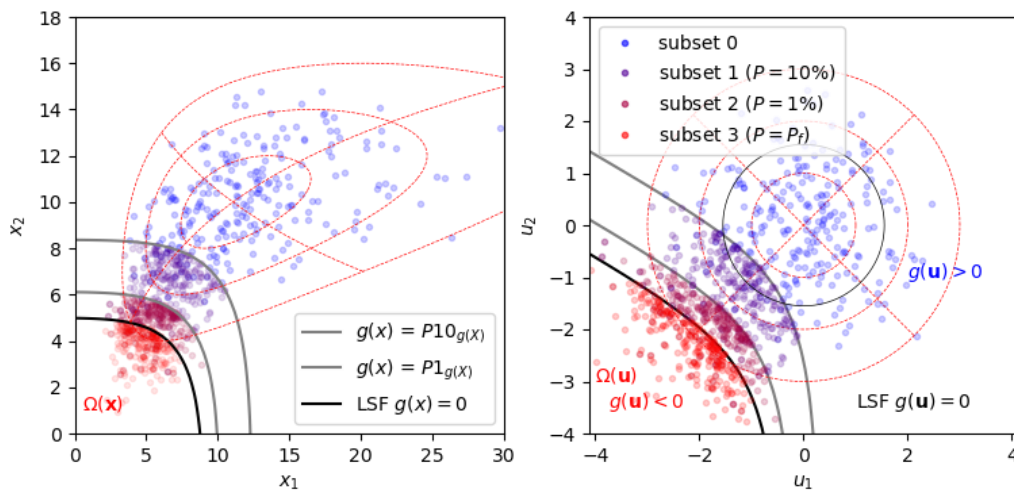


Source: Authors' own work

6.2.6.2 Subset simulation (SubSim)

Subset simulation is a sampling-based reliability method in which a series of conditional Monte Carlo samples is generated. Samples are conditioned to fall in the lower p_0 -percentile of the previous subset, typically the 10-percentile. In this way, the samples are narrowing in on the limit state in subsequently generated subsets, after which the final probability of exceeding the limit state is estimated as the product of the conditional probabilities of the subsets. **Figure 37** shows an example of a subset simulation.

Figure 37. Schematic example of a subset simulation with the parameter space of a multivariate distribution (left) and the corresponding standard normal space (right). Subsets are generated with subsequent conditional probabilities of 10% (i.e. 10-percentile, 1-percentile, ...) until the limit state condition $g(x) = 0$ is reached. The probability of failure is the production of the subsequent conditional probabilities between subsets. Note that Subset 0 is a standard Monte Carlo sample set and that Subset 3 itself does not need to be sampled for the evaluation of P_f .



Source: Authors' own work

Subset simulation is mostly insensitive to non-linearities, dimensionality and smoothness of the limit state function. The efficiency in terms of number of model evaluations scales approximately linear with $\log(P_f)$, and the method typically requires several thousands of realisations. The robustness is influenced by the performance of the Markov chain Monte Carlo sampler that is being employed for the conditional samples of the subsets. Further details and evaluation of subset simulation in the context of geotechnical engineering can be found in Au & Wang (2014).

6.2.6.3 Directional sampling (DS)

In directional sampling (DS, **Figure 38**), the limit state is iteratively evaluated in randomly sampled directions in standard normal space by a line search algorithm. Along each of the randomly sampled directions θ_i (i.e. along the search lines starting at the origin), a conditional reliability index $\beta_{|\theta_i}$ is established as the distance in standard normal space between the origin and the limit state. The probability of failure is then evaluated as

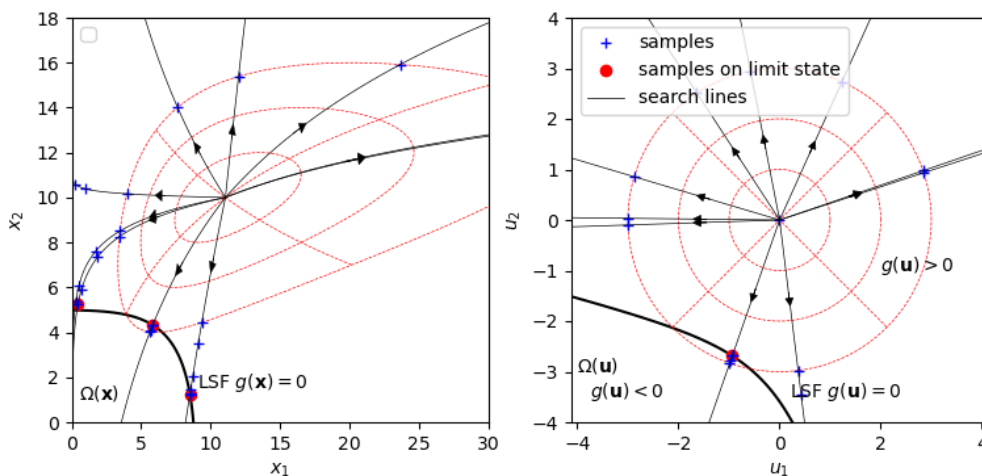
$$\hat{P}_f = 1 - \frac{1}{N} \sum_{i=1..N} \chi^2(\beta_{|\theta_i}^2 | D)$$

Equation 59.

where $\chi^2(\cdot | D)$ is the chi-square distribution, N the number of directions in which the line-search is performed and D the number of random variables (i.e. dimensionality). Like all sampling-based reliability methods, directional sampling converges to the exact probability of failure with increasing N .

The computational cost of directional sampling scales approximately linear with both dimensionality and $\log(P_f)$. Depending on the employed line-search algorithm, directional sampling may be sensitive to strong non-linearity and noise in the limit state function. Many (adaptive) improvements are possible to enhance the efficiency of directional sampling, such as moving the origin of the directional line search in the direction of the (apparent) design point, or applying importance sampling on θ_i with an increased sampling density for the orientations in which $\beta_{|\theta_i}$ is expected to be small.

Figure 38. Schematic example of directional sampling with radial sampling from the origin in standard-normal space.



Source: Authors' own work

6.2.6.4 Surrogate models

The option of using surrogate models for reliability analysis is explicitly mentioned in EN 1990. A surrogate model (also referred to as response surface or metamodel) is an approximation of the original computational model fitted to a limited but carefully selected set of evaluations of the limit state function called the experimental design. The surrogate model itself is ideally a computationally inexpensive model which approximates the computational model (at least) in relevant parts of the parameter space (i.e. around the limit state). Typical options for surrogate models are polynomial expressions or Kriging models. In adaptive versions, the surrogate model is updated as more evaluations of the limit state function are performed and an informed decision can be made in the selection of parameter combinations to be evaluated. Once a surrogate model approximates the computational model sufficiently well, the reliability can be estimated by applying traditional reliability methods on the faster surrogate model instead of the original computational model.

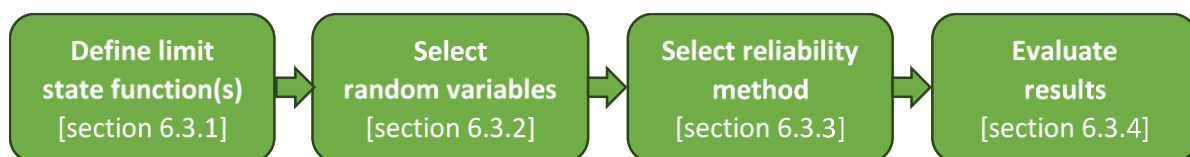
Since the research in adaptive surrogate modelling has increased in recent years, there are many methods available combining different surrogate models, learning functions, algorithms for reliability estimation and the stopping criteria. See Moustapha et al. (2022), and Teixeira et al. (2021) for reviews of such approaches.

Broadly speaking, simpler metamodels (e.g. polynomial) tend to perform adequately for low complexity limit-state functions, while more complex metamodels (e.g. Kriging) perform better for complex functions (e.g. highly non-linear, noisy and/or incomplete) as discussed in Van den Eijnden et al. (2021).

6.3 Practical recommendations

This section contains practical recommendations for reliability verification of geotechnical limit states, relating to the steps generally taken for the analysis as outlined in **Figure 39**.

Figure 39. Steps in reliability analysis (section outline)



Source: Authors' own work

The other practical options described subsequently are the use of point estimates (6.3.5), reliability-based design values (0) and conditional reliability analysis using fragility curves (6.3.7).

6.3.1 Limit state functions

In addition to EN 1990, EN 1997-1 establishes principles and requirements for the safety, serviceability, robustness, and durability of geotechnical structures. EN 1997-3 specifically provides principles and requirements for the design and verification of certain types of geotechnical structures, supporting elements and groundwater control. The formulation of limit state functions for reliability-based analysis is in principle identical to the formulation for partial factor-based analysis.

6.3.1.1 Explicit (analytical) versus implicit (numerical) limit states

EN 1997-3 mostly presents explicit limit state functions based on design equations. Due to their computational simplicity, these closed-form equations are very easily evaluated in a fully probabilistic way, even by the computationally most demanding sampling methods (i.e. Crude Monte Carlo). Many limit states in engineering practice can be formulated using these design equations, even for more complex designs, but others need numerical models, which often introduce a significant computational load and result in an implicit limit state function.

Once the numerical model response is reformulated in the form of a limit state function, implicit limit state functions are, in principle, treated the same way as explicit limit state functions when applied in the reliability method. However, implicit limit state functions can be (very) time consuming and may exhibit (numerical) 'noise'. Improving the numerical accuracy of the computational model (to improve differentiability/smoothness of the implicit limit state function) as much as possible is advisable.

It is essential to prepare a robust calculation model for the probabilistic analysis, accounting for possible limitations of the software used or the formulations of the constitutive laws. For example, extreme values of the random variables or combinations of these might be incompatible for the software or lead to numerical issues. For this reason, prior tests of the model by evaluating the model for possible extreme parameter combinations are advisable. This should also reveal if other calculation adjustments are necessary (e.g. adjusting calculation precision or steps, adding small cohesion). Moreover, for computationally expensive models (e.g. FEM), the model should be as simple as possible without compromising the accuracy of the results. For example, some aspects to be analyzed and accounted might be: dimensions of the model, simplification of the geometry, mesh optimization, number of calculation stages/ phases. This is always good practice but even more important in probabilistic analysis.

6.3.1.2 System and component limit states

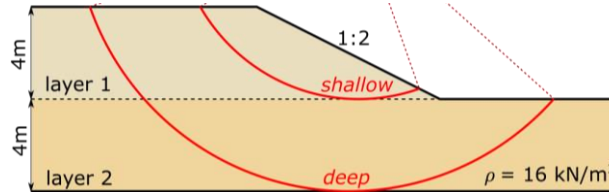
For system reliability problems, it can be beneficial to define the limit state for the overall structural performance by combining all components into a single limit state function. In this way, each model evaluation directly gives the performance of the system and the reliability method evaluates the system reliability rather than the reliability of single members or failure modes. This approach works with reliability methods that are insensitive to the (non-)linearity of the limit state function (e.g. Monte Carlo simulation).

With reliability methods that are sensitive to the linearity of the limit state function (e.g. FOSM or FORM), it is better to define and evaluate separate limit state functions for each individual component or failure mode. Each component can then be evaluated separately, and the combined reliability can be determined as explained in section 6.1.5, as illustrated in Example 6.4 below.

Box 19. Example 6.4: Limit states for multiple failure modes

The reliability problem of slope stability in undrained conditions contains two layers with independent stochastic soil strength $s_{u,1}$ and $s_{u,2}$ (all other variables are deterministic), **Figure 40**.

Figure 40. Geometry of slope in undrained condition for the definition of the ultimate limit state for slope stability $g(\mathbf{X}) = FS - 1$. The two failure mechanisms are the deep and shallow sliding circles.



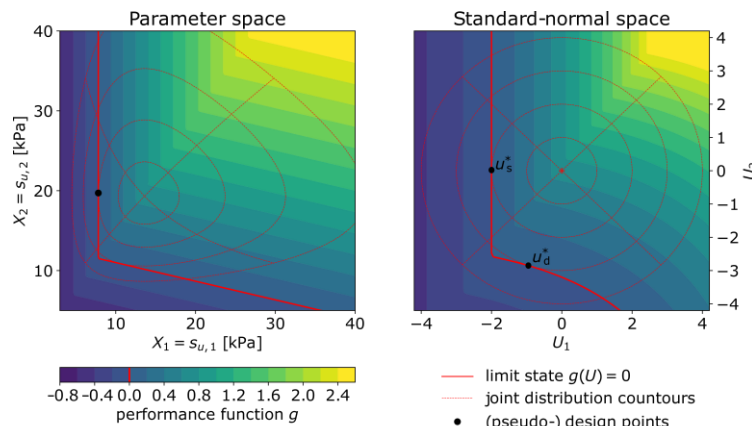
Source: Authors' own work

Using Bishop's limit equilibrium method, either a deep or a shallow slip circle can be found critical, depending on the undrained shear strength in each layer. The system limit state function $g_{sys}(\mathbf{X})$ can be formulated as the combination of the component limit state functions $g_d(\mathbf{X})$ and $g_s(\mathbf{X})$ for deep and shallow failure modes respectively:

$$g_{sys}(\mathbf{X}) = \min(g_s, g_d) \quad \text{with} \quad \begin{cases} g_s = FS_s - 1 \\ g_d = FS_d - 1 \end{cases}$$

The undrained shear strength (in kPa) in each layer is lognormal distributed with $s_{u,1} = X_1 \sim LN(\mu_X = 14.14; \sigma_X = 4.0)$ and $s_{u,2} = X_2 \sim LN(\mu_X = 19.71; \sigma_X = 4.0)$; they are assumed independent. The performance function can be plotted in parameter space and u-space (**Figure 41**):

Figure 41. System limit state function for undrained slope stability involving two layers with independent shear strength.



Source: Authors' own work

When using a standard FORM implementation, the design point corresponding to a deep sliding surface ($\beta_{HL,d} = 3.00$) is found. When the reliability is evaluated component-wise, both design points ($\beta_{HL,d} = 3.0$ and $\beta_{HL,s} = 2.0$) are found, indicating that the system FORM analysis over-estimated the reliability. The system reliability can be estimated by combining the component reliabilities, e.g. $\beta_{sys} \approx -\Phi^{-1}[\Phi(-\beta_{HL,d}) + \Phi(-\beta_{HL,s})] = 1.98$.

6.3.2 Selection of random variables

Based on the characterisation of uncertainty of all relevant parameters and their cross-correlations involved in the deterministic model (geotechnical parameters, geometry, loads, model parameters), a joint distribution is to be formulated as input for the reliability analysis. Both for the performance and the interpretation of the analysis, it is beneficial to only model variables with significant uncertainty as random variables. The selection of random variables can be based on:

— Engineering judgement

(For example, the uncertainty in elastic properties may be large, but the relative contribution to the ultimate limit state may well be negligible compared to the contribution of the strength parameters.)

— Sensitivity analysis

— Preliminary reliability analysis using a less refined model.

(A less refined and computational model can be used first with a more efficient reliability method. The relative influence of the different random variables, e.g. based on FORM influence coefficients, can aid the selection.)

A simple check to verify if treating variables as deterministic and not random is to perform the reliability analysis twice: once using design values, and once using the mean values. The difference in the reliability index should be small (e.g. $\Delta\beta < 0.1$).

6.3.3 Selection of reliability method

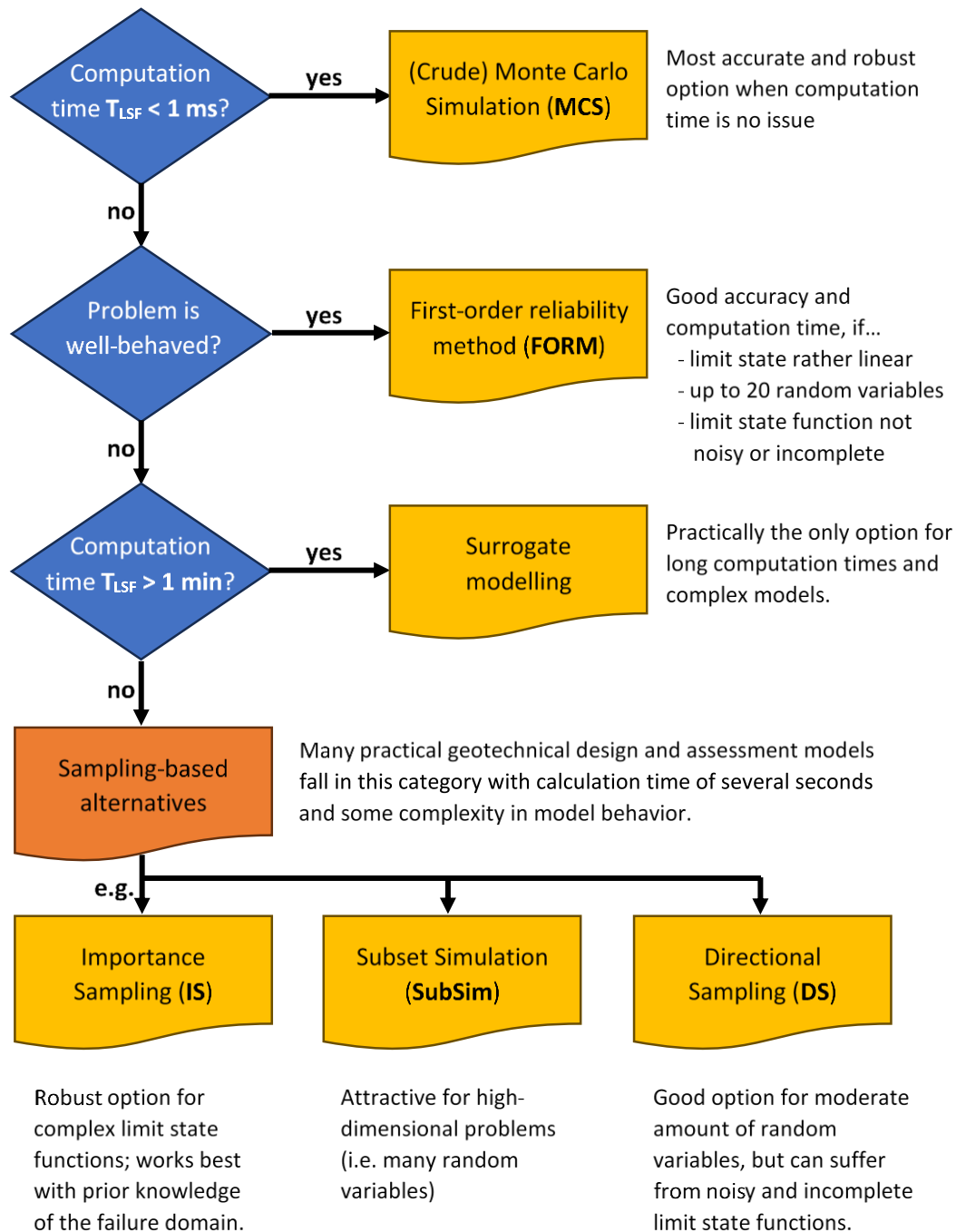
Several problem-specific aspects need to be considered to select an appropriate reliability method that can provide appropriate estimates of the reliability in a reasonable amount of time. Appropriate estimates in a practical context are those with an absolute estimation error in β below 0.1 (or a CoV in the estimated Pf below 0.1); a reasonable amount of computation time is considered to be several hours.

Figure 42 provides a flow-chart based on the following characteristics of the problem and the associated limit state function:

- the time T_{LSF} required to evaluate the limit state function once;
- the behaviour of the limit state model (i.e. well-behaved problems have limit state functions which are not strongly non-linear, noisy or incomplete);
- and to a lesser degree:
 - the number of random variables (dimensions) N_{dim} ;
 - insight into the failure domain (e.g. influence factors α and design point x^*).

All aspects affecting the selection of the reliability method are discussed more in detail below.

Figure 42. Flow chart for selecting a reliability method based on characteristics of the limit state function (LSF). The dotted arrows indicate optional verification or improvement steps, or an alternative approach as back-up.



Source: Authors' own work

The decision criteria should be considered as indicative, as strict criteria are impossible to formulate for general limit state functions. Specifically the selection of an appropriate sampling-based method may depend on specific characteristics of the limit state function, and (small) algorithmic enhancements of the reliability methods can have a significant impact on the efficiency and accuracy of the method.

6.3.3.1 Computation time of the model

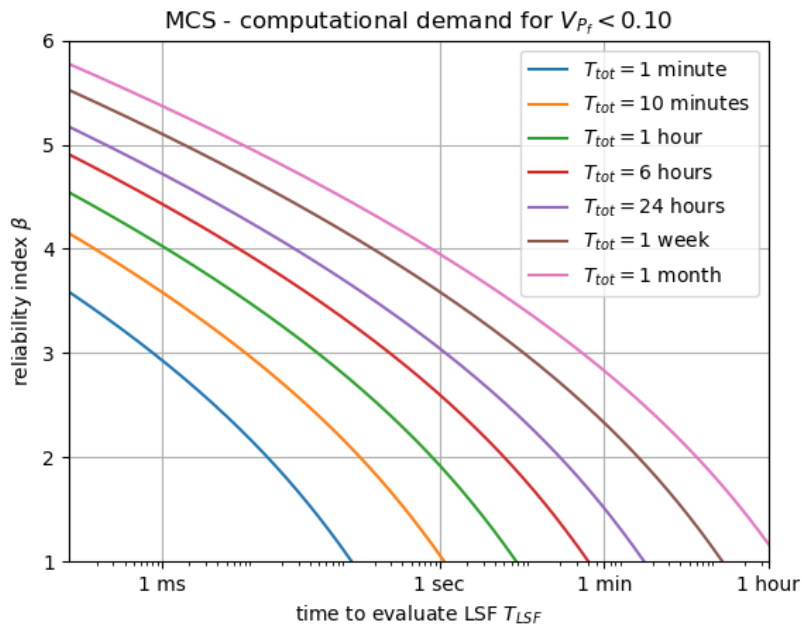
The time it takes to evaluate the limit state function has a major impact on the applicability of reliability methods. For example, a high computation time excludes Monte Carlo simulation as a suitable reliability method due to the high number of required calculations. Based on the computation time T for a single evaluation of the limit state (i.e. the computational model(s) used to evaluate the limit state function $g(x)$), the following guidelines can be followed:

- $T < 1$ ms All reliability methods can be used, even for high reliability (i.e. low probability of failure), with Crude Monte Carlo being the most robust method.
- $T \sim 1$ s Only sampling-based methods with variance reduction for efficiency (importance sampling, directional sampling), surrogate model-based and Level-II methods can be used.
- $T > 1$ min Only level-II methods and surrogate model-based methods can be used. In exceptional cases, sampling methods may be an option.

Box 20. Background: Required accuracy and computation times with Crude Monte Carlo

Monte Carlo simulation is the only reliability method for which the required number of limit state function evaluations only depends on the reliability level and its required estimation accuracy. Consequently, the total computation time for the estimation of the reliability with a required accuracy (e.g. $V_{\hat{p}_f} < 0.10$) can be evaluated analytically. **Figure 43** gives the required total computation time as a function of reliability index β and time T_{LSF} . For example, for limit state model with computation time $T_{LSF} = 1$ ms and a reliability index of $\beta = 4$ the total computation time is roughly $T_{LSF} = 1$ h.

Figure 43. Total computation time required for the estimation of the reliability index by MSC with estimation uncertainty $V_{\hat{p}_f} < 0.10$ as a function of the reliability index β and the time to evaluate the limit state function



Source: Authors' own work

6.3.3.2 Number of random variables (dimensionality)

High-dimensional reliability problems (i.e. many random variables) are problematic for gradient-based methods. For example, FORM requires more model evaluations for the derivation of the gradient of the limit state function for higher dimensionality, and the assumption of a linear limit state in the Hasofer-Lind reliability index loses accuracy in higher dimensions. Some sampling-based reliability methods are completely insensitive to the dimensionality of the problem (e.g. Monte Carlo simulation or subset simulation), and sampling-based methods are typically more effective in high-dimensional problems with strong non-linearities.

6.3.3.3 (Non-)Linearity

Reliability methods which are based on assumptions regarding the shape of the limit state in either parameter space or U-space can perform poorly when applied on strongly non-linear problems. In particular, the first order second moment (FOSM) and point estimate method (PEM) are very sensitive to non-linearity of the limit state function, and should only be applied when the limit state function is approximately linear. The first order reliability method (FORM) can suffer from non-linearities through the convergence at an unimportant design point as well as an incorrect approximation of the limit state by the assumption of linearity. In (geotechnical) practice, the latter is seldomly a problem and the approximation by FORM tends to be sufficiently accurate. Importance sampling (IS) can lead to incorrect results when non-linearities of the limit state form distinct failure domains, of which one or more are under-sampled. Only few reliability methods are completely insensitive to the shape of the limit state, of which Monte Carlo simulation is the most robust method.

6.3.3.4 (Numerical) model stability and existence of solution

Numerical models can suffer from instabilities in their formulation, often as a result of finite precision in their numerical convergence schemes. This can lead to stability issues resulting in erratic behaviour of the limit state function in the form of numerical noise or discontinuities. Noise and discontinuities can be problematic for the application of gradient-based reliability methods like FORM or DS.

Another form of instability in the limit state function may come from the validity of the computational model for certain (extreme) combinations of input parameters. Some combinations of parameters, although plausible according to the joint probability density function, may lead to physically impossible scenarios in the (numerical) model, such that a solution for the limit state function does not exist. The non-existence of solutions can cause difficulties in the evaluation of the reliability if the algorithm of the selected reliability method is not designed to deal with these incomplete model responses. Incomplete model responses are common in the more complex geotechnical analysis, for example when failure occurs during the initial stages of the analysis (e.g. the modelling of a construction phase before the evaluation of the actual limit state).

6.3.4 Verification of results (as part of design check)

The results from reliability methods are estimations of the reliability and need verification for accuracy. Inaccuracies may be due to the unsuccessful application of a method, such as failure to converge or the convergence on an unimportant design point. Verification of the results can address questions such as:

- Is the resulting reliability index reasonable?

- Is the failure mechanism (in the design point) plausible?
- Have all failure mechanisms been identified?
- Are the confidence bounds ($V_{\hat{p}_f}$) on the calculated reliability index acceptable?
- How sensitive is the result to the input parameters?

Below some practical recommendations are given.

6.3.4.1 Design point (FORM)

Practical checks that confirm the plausibility of first order reliability method (FORM) analysis outcomes are:

- The development of the reliability index (β) and the influence coefficients (α_i) over the iterations can indicate whether the analysis has converged satisfactorily.
- The influence coefficients (α_i) should be plausible for the problem at hand, i.e. the variables with the largest contributions should make sense. And the sign (+/-) should correctly indicate whether a variable has a favourable or unfavourable effect.
- The limit state function can be (re)evaluated in the design point (x^*) in order to:
 - check whether the design point is indeed close to the limit state;
 - examine the failure mode in the deterministic analysis (e.g. failure surface, stresses, deformations).

Whereas design point estimation is inherent to FORM, sampling-based methods can also provide design point estimates (for example the sample in the failure domain with the highest probability density in standard normal space). Most software packages for reliability analysis already include this option.

6.3.4.2 Uncertainty bounds for sampling-based reliability methods

Uncertainty bounds for sampling-based reliability methods can be used to assess the uncertainty in the estimation of the probability of failure. These bounds can be formulated based on the estimation variance, which can be evaluated without additional simulations for most of the sampling-based methods (see e.g. Crude Monte Carlo in 6.2.5).

Note that verifying the convergence of the reliability estimation through V_{P_f} is only meaningful if the method is applied successfully and may otherwise provide a false sense of accuracy. This is particularly relevant when sampling-based methods fail to cover relevant parts of the failure domain, and hence underestimate the probability of failure and at the same time give a low estimation variance (see Example 6.5).

Box 21. Example 6.5: Convergence of sampling-based methods

Consider the evaluation of the reliability of the slope stability problem involving 2 layers defined in Example 6.4, for which FORM converges to the incorrect design point with $\beta_{HL} \approx 3.0$, while the true reliability index $\beta = 1.98$. Here, the performance of sampling-based reliability methods Monte Carlo simulation (MCS) and importance sampling (IS) is assessed. A convergence criterion based on the

coefficient of variation in the predicted probability of failure is defined as $V_{P_f} < 0.10$. Equation 5.13 is used to evaluate V_{P_f} for Monte Carlo simulation, for importance sampling the following relation is used:

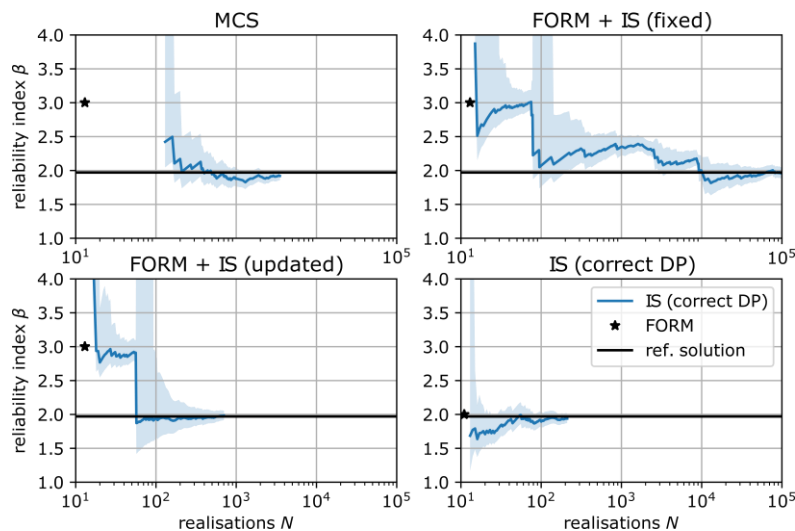
$$V_{P_f}^{IS} = \frac{\sigma_{P_f}}{\bar{P}_f}, \quad \sigma_{P_f}^2 = \frac{1}{N-1} \left(\frac{1}{N} \sum_i^N w_i^2 I_i - \bar{P}_f^2 \right)$$

Figure 44 below shows the convergence of the reliability estimation towards the correct solution, together with the 90% confidence interval:

- Monte Carlo simulation (top-left) converges in approximately $N = 4000$ realisations, with the reference solution falling within the confidence interval.
- Importance sampling with an importance distribution centered at the secondary design point found with FORM (top-right) shows very poor convergence, and an initial false tendency to converge towards $\beta = 3.00$.
- Importance sampling based on the same FORM design point, but with an update of the centre of the importance distribution whenever a failing sample closer to the origin becomes available (bottom-left) initially shows the same tendency of converging to the incorrect solution, while correct and efficient convergence is obtained after the importance distribution is updated, here around $N = 60$.

Importance sampling with the importance distribution centred at the true design point shows correct and very efficient ($N \approx 250$) convergence, with the correct solution within the confidence bounds.

Figure 44. Convergence rate for Monte Carlo and Importance-based sampling algorithms.



Source: Authors' own work

This example shows that although sampling-based methods converge to the correct solution with increasing number of realisations, care must be taken when interpreting the confidence in the results, especially when relatively few realisations are used. Nevertheless, variance reduction techniques like importance sampling can, when applied successfully, lead to significant improvements in efficiency (orders of magnitude) compared to Crude Monte Carlo simulation.

Cross-verification with multiple reliability methods can be applied to verify the consistency in the outcomes of reliability analyses. As recommended by EN 1990 (Annex B.5(4)), "design checking should cover, [...] where appropriate, separate calculations as alternatives to reviewing the design

calculations". When an initial reliability method is used that is not robust, the outcome can be verified by a second reliability method.

6.3.5 Point estimates

When probability distributions are hard to attain for variables which are supposed to have some (but not dominant) influence on the probability of exceeding the limit state, point estimates may be used instead. A point estimate is a deterministic value of a random variable, which ideally assumes the design point value (in the context of FORM) in case the variable had been modeled as random, in order not to affect the reliability estimate significantly.

The above implies:

- a. for variables with hardly any effect on the estimated probability of exceeding the limit state (due to little uncertainty, or little effect on the limit state function), the best estimate (or expected value) can be used;
- b. for variables with moderate effect on the estimated probability of exceeding the limit state, a cautious estimate can be used.

Case a above is actually equivalent to the selection of random variables described in 6.3.2; i.e. only variables with a significant effect on the reliability estimate need to be modeled explicitly as random variables. Furthermore, important random variables with a significant or even dominant effect on the reliability estimate, should not be reduced to point estimates, but always be modelled probabilistically.

A typical situation in which a point estimate is used is when experts may agree on a conservative or cautious estimate, but data are lacking, and no consensus is achieved on the probability distribution (see Example 6.6).

Box 22. Example 6.6: Point estimate of train load for railway embankment slope reliability

An example is the train load for the slope stability analysis of a railway embankment. The train load is often of moderate influence in the stability problem, which is typically dominated by the uncertainty in soil strength (and the groundwater conditions). The geotechnical engineer may have information on the design value of the train loads, but not the underlying statistics. A reliability analysis using the design value of the train loads should then give a reasonable or slightly conservative estimate of the reliability index.

6.3.6 Reliability-based design values

Occasionally, in reliability verifications only part of the problem may be analyzed probabilistically to produce reliability-based design values, which can then be used in an otherwise semi-probabilistic verification. Remember that the basic semi-probabilistic verification format of the Eurocodes reads:

$$E_d \leq R_d$$

Equation 60.

implying that the design values of the effects of actions E_d shall be less than or equal to the design value of the corresponding resistance.

We may now treat either E or R in a full probabilistic fashion, while assessing the design value of the other quantity conventionally. When assessed fully probabilistically, the design values are

assessed at a specific probability of exceedance according to the so-called design value method (which is otherwise used for calibration of partial factors, see EN 1990 Annex C):

$$E_d = F_E^{-1}(\Phi(-\alpha_E \beta_T))$$

Equation 61.

$$R_d = F_R^{-1}(\Phi(-\alpha_R \beta_T))$$

Equation 62.

where β_T is the target reliability index, and α_E and α_R are the FORM influence coefficients of the normal-distributed E and R , respectively. The probabilistic analysis of E or R would deliver the corresponding probability distributions F_E or F_R , from which the design value is obtained at the specified quantile or probability of exceedance (with Φ being the standard normal CDF). EN 1990 Annex C also presents the general case for any variable, providing specific solutions for commonly used probability distributions (Normal, Lognormal and Gumbel).

Since the actual FORM influence coefficients are unknown in the absence of a complete reliability analysis, EN 1990 Annex C provides approximate values which can be used, based on ISO 2394:2015 (see **Table 24**).

Table 24. Recommended influence coefficients (α)

X_i	α_i
Dominating resistance parameter	0.8
Other resistance parameter	$0.4 \times 0.8 = 0.32$
Dominating load parameter	-0.7
Other load parameters	$-0.4 \times 0.7 = -0.28$

Source: ISO 2394:2015 (Table E.3)

EN 1990 Annex C adds that influence factors of $\alpha_R = 1$ or $\alpha_E = -1$ should be used when the variable of interest dominates the reliability problem (conservative approach). Notice that this approach is equivalent to the reliability targets for resistance or actions as described in section 4.3, only embedded in a different format or presentation.

Box 23. Example 6.7: Design value of effects of actions for sheet pile wall

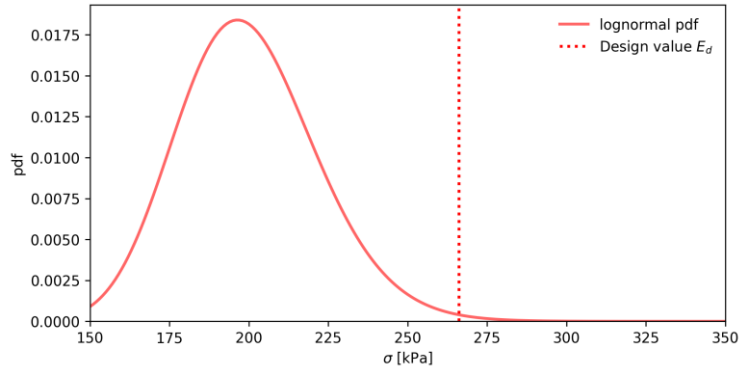
In this example we suppose we have done a probabilistic analysis of the stresses in a steel sheet pile wall, using Monte Carlo simulation. The resulting probability distribution of the stresses (maximum over depth) is characterized by a lognormal distribution with mean 200 kPa and standard deviation 22 kPa (i.e. coefficient of variation of 11 %; see **Figure 45** below).

We now want to compare the design value of the stresses (effects of actions) with the yield stress of the steel (corresponding strength), i.e. $E_d \leq R_d$. To obtain the design value E_d using **Equation 49**, we assume a target reliability index of $\beta_T = 3.8$ for CC2, and an influence coefficient of $\alpha_E = -0.7$. The resulting design value is determined as:

$$E_d = F_E^{-1}(\Phi(-\alpha_E \beta_T)) = F_E^{-1}(\Phi(0.7 \cdot 3.8)) = F_E^{-1}(0.9961) = 266.1 \text{ kPa}$$

In other words, the design value of the stresses (effects of actions) is determined at a probability of exceedance of $P(E > E_d) = 0.4\%$ in this case.

Figure 45. Comparison between the probability distribution of variable E and associated design value E_d , when accounting for the influence coefficient α_E .



Source: Authors' own work

6.3.7 Conditional reliability analysis / fragility curve approach

When evaluating small probabilities of failure for a multivariate problem in which a single independent (load) variable governs the exceedance of the limit state, the so-called fragility curve approach can be used. When an independent and significant variable X_i is identified, conditional probabilities $P_f(\mathbf{X} | X_i = x_i)$ can be evaluated for a series of values x_i of X_i . The resulting conditional probabilities can be presented in a fragility curve ($X_i - \beta$ plot). The total probability can then be evaluated by integrating out the random variable X_i :

$$P_f(\mathbf{X}) = \int P_f(\mathbf{X} | X_i = x) f_{X_i}(x) dx$$

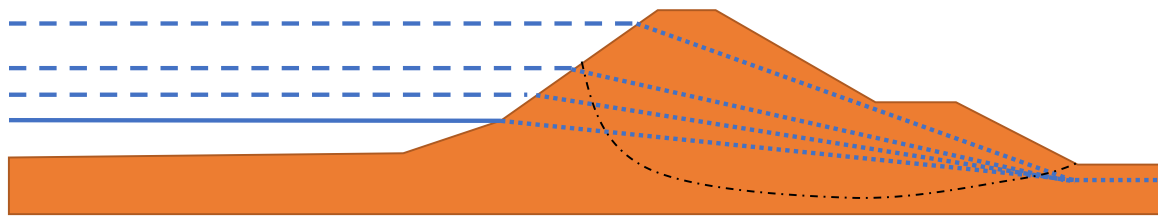
Equation 63.

This 1-dimensional integral can be evaluated by numerical integration. Besides a potential gain in the efficiency of the reliability method, the reliability curve gives insight into the sensitivity of the problem to variations in the single independent (load) variable. This approach is usually applied to load-related variables (water levels, foundation loads etc.), as illustrated in Example 6.8.

Box 24. Example 6.8: Fragility curves

Consider the following example of a flood embankment (river dike) as an illustration of working with fragility curves (**Figure 46**). When assessing slope stability of the embankment, we may find very different failure behaviour for different water levels (i.e. the main loads), for example deep versus shallow failure surfaces.

Figure 46. Schematisation of a flood embankment for the analysis of reliability conditional to different water tables.



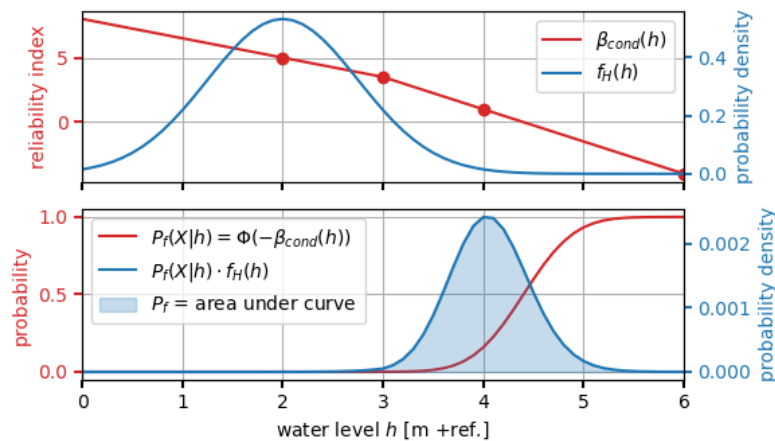
Source: Authors' own work

In such a situation, it can be beneficial to do conditional reliability analyses for discrete water levels to have better insight into the reliability and dominating uncertainties at each level of the load. Furthermore, reliability methods like FORM may suffer from convergence issues when the water level is considered as a continuous random variable, due to the changing failure modes for varying water levels. We assume that the reliability index (β_{cond}) belonging to the conditional probability of failure for given water levels ($H = h$) has been determined as shown in the table below:

water level h [m+REF]	2.0	3.0	4.0	6.0
conditional reliability index β_{cond}	5.0	3.5	1.0	-4.0

Furthermore, we assume that the probability distribution of the annual maximum water level is captured by a Normal distribution with mean $\mu_H = 2.0$ m and standard deviation $\sigma_H = 0.75$ m. Both, the conditional reliability index β_{cond} and the water level distribution $f_H(h)$ are depicted in **Figure 47**.

Figure 47. Fragility curve conditional to the water level h with water level distribution $f_H(h)$. The area under the weighted curve (shaded in blue) represents the total probability, showing that water levels around $h = 4$ m +ref. contribute most to the failure probability.



Source: Authors' own work

6.4 Benchmark example

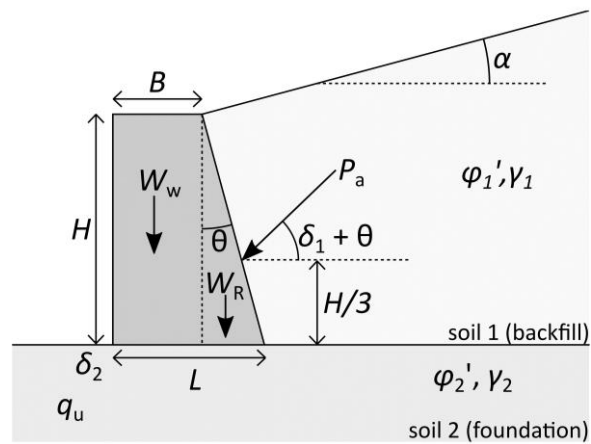
This example (**Figure 48**) is based on geotechnical Benchmark problem 3 of the GEOSNet⁹ and concerns a retaining wall problem with 4 independent random variables. Three limit state functions (g_1 , g_2 and g_3) are considered including sliding, bearing capacity and overturning. The purpose of this example is to examine the robustness of each reliability method for problems with multiple failure modes, when they are combined in a single limit state function $g_{sys}(\mathbf{X})$:

$$g_{sys}(\mathbf{X}) = \min(g_1(\mathbf{X}), g_2(\mathbf{X}), g_3(\mathbf{X}))$$

Equation 64.

Four deterministic dimensions are defined, together with four independent random variables for the soil properties $x = \{\gamma_1, \gamma_2, \phi_1, \phi_2\}$, **Table 25**

Figure 48. Problem sketch for gravity wall example



Source: Ching & Hsieh (2011)

Table 25. Deterministic and stochastic variables in the gravity wall example problem.

Variable	Description	Distribution	Statistics
B	Top width of the wall	deterministic	2.5 m
H	Height of the wall	deterministic	4.0 m
L	Bottom of the wall	deterministic	3.5 m
α	Slope of the backfill soil	deterministic	5.0°
γ_1	Unit weight of the backfill soil	Normal	$\mu = 19.0 \text{ kN/m}^3$ $V = 0.10$
γ_2	Unit weight of the foundation soil	Normal	$\mu = 17.0 \text{ kN/m}^3$

⁹ http://140.112.12.21/issmge/reliability_benchmark/GEOSNet_TG3_Example_3.doc

Variable	Description	Distribution	Statistics
			$V = 0.10$
φ_1	Friction angle of the backfill soil	Normal (truncated at α , i. e. PDF = 0 when $\varphi_1 < \alpha$)	$\mu = 35^\circ$ $V = 0.10$
φ_2	Friction angle of the foundation soil	Normal	$\mu = 35^\circ$ $V = 0.10$

Source: Authors' own work

6.4.1 Limit state function

The system limit state function is formulated as the minimum of three component limit states, expressing the three associated failure mechanisms as follows:

$$g_1(x) = [P_a \sin(\delta_1 + \theta) + W_w + W_R] \tan \delta_2 - P_a \cos(\delta_1 + \theta) \quad (\text{sliding})$$

$$g_2(x) = q_u \bar{L} - (W_w + W_R + P_a \sin(\delta_1 + \theta)) \quad (\text{bearing capacity})$$

$$g_3(x) = M_R - M_O \quad (\text{overturning})$$

with

P_a = active force

δ_1 = friction angle between the backfill and the back of the wall

δ_2 = friction angle between the foundation soil and base of the wall

φ_1 = friction angle of the backfill soil

φ_2 = friction angle of the foundation soil

$W_w + W_R$ = the weight of the retaining wall

B = top width of the retaining wall

L = bottom width of the retaining wall

α = slope of the backfill soil

q = bearing capacity of foundation soil

H = height of the retaining wall

θ = back angle of the retaining wall

M_R = resisting moment

M_O = overturning moment

where

$$W_w = 23.58 \cdot B \cdot H \quad (\text{unit weight of the retaining wall: } 23.58 \text{ kN/m}^3)$$

$$W_R = \frac{23.58(L-B)H}{2}$$

$$\theta = \tan^{-1} \left[\frac{L-B}{H} \right]$$

$$\delta_1 = \frac{2}{3}\phi_1, \quad \delta_2 = \frac{2}{3}\phi_2, \quad P_a = \frac{1}{2}K_a\gamma_1H^2$$

$$K_a = \frac{\cos^2(\phi_1 - \theta)}{\cos^2\theta \cos(\delta_1 + \theta) \left[1 + \sqrt{\frac{\sin(\delta_1 + \phi_1) \sin(\phi_1 - \alpha)}{\cos(\delta_1 + \theta) \cos(\theta - \alpha)}} \right]^2}$$

$$q_u = \frac{1}{2}\gamma_2\bar{L}N_\gamma F_{\gamma i}, \quad N_\gamma = 2(N_q + 1) \tan\phi_2, \quad N_q = e^{\pi \tan\phi_2} \tan^2\left(45 + \frac{\phi_2}{2}\right)$$

$F_{\gamma i} = \left(1 - \frac{\beta}{\phi_2}\right)^2$ ($F_{\gamma i}$ from Hanna and Meyerhof (1981)), where β is the inclination angle of the total foundation loading)

$$\beta = \tan^{-1} \left(\frac{P_a \cos(\delta_1 + \theta)}{(P_a \sin(\delta_1 + \theta) + W_W + W_R)} \right)$$

$$M_R = W_W \frac{B}{2} + W_R \left(B + \frac{1}{3}(L - B) \right) + P_a \sin(\delta_1 + \theta) \left(B + \frac{2}{3}(L - B) \right)$$

$$M_O = P_a \cos(\delta_1 + \theta) \frac{H}{3}$$

$$\bar{L} = L - 2e_L \quad e_L = \left| \bar{x} - \frac{L}{2} \right|, \quad \bar{x} = \frac{M_R - M_O}{W_W + W_R + P_a \sin(\delta_1 + \theta)}$$

The above equations make up the geotechnical model for the three limit states and are implemented in a single system limit state function $g_{sys}(\mathbf{x}) = \min(g_1(\mathbf{x}), g_2(\mathbf{x}), g_3(\mathbf{x}))$, assuming a series system.

6.4.2 Reliability methods

A series of reliability methods is applied to evaluate the (system) reliability of the retaining wall. For the sampling-base reliability methods, a convergence criterion based on the coefficient of variation in the estimated probability of failure $V_{P_f} < 0.10$ is used. This allows a relative comparison in terms of accuracy and efficiency (i.e. how many computations are needed). A total of 12 reliability methods are used, most of which are discussed in Section 6.2. In addition, the second order reliability method (SORM, Breitung (1984)¹⁰) is used, which is a second-order extension of FORM, accounting for the curvature of the limit state at the design point. A surrogate-model approach is used as well by applying two versions of AK-MCS, which is an active learning reliability method combining Kriging and Monte Carlo simulation (Echard et al., 2011). Here, a kriging-based surrogate model is constructed for efficiently performing a Monte Carlo analysis.

Table 26 contains the calculated reliability index and the required number of limit state function evaluations.

¹⁰ Breitung, K., 1984, "Asymptotic Approximations for Multinormal Integrals," J. Eng. Mech., 110(3), pp. 357–366.

Table 26. Results of the reliability analysis of GEOSNet benchmark problem 3 using different reliability methods.

Method	β_{sys} (relative error)	Number of LSF evaluations
Monte Carlo simulation	2.884 (-1.1%)	51,000
Importance sampling (around FORM design point)	2.889 (-0.9%)	358
Adaptive importance sampling	2.957 (+1.4%)	519
Subset simulation	2.977 (+2.1%)	7,228
Directional sampling	2.889 (-0.9%)	6,487
FOSM (forward difference)	5.481	5
FOSM (central difference)	5.460	9
Point estimate method	5.446	16
FORM	2.922 (+0.2%)	36
SORM	3.097 (+6.2%)	52
AK-MCS (zero-mean)	2.931 (+0.5%)	90
AK-MCS (linear trend)	2.901 (-0.5%)	82
Monte Carlo simulation (reference)	2.917 (-)	100,000,000

Source: Authors' own work

Monte Carlo simulation using $k = 10^8$ samples is used as a reference solution. The estimates of the reliability index indicate that, with the exception of the first-order second moment (FOSM) method and point estimate method (PEM), all methods give rather accurate results, with relative errors for sampling-based methods in line with the convergence criterion of $V_{P_f} < 0.10$.

FOSM and PEM seem inadequate for this type of problems: despite all variables following a normal distribution, the non-linearities in the limit state function $g_{sys}(x)$, probably arising from the system reliability aspect, result in unacceptable errors.

FORM gives very accurate results, suggesting that the limit state function is rather linear around the design point, and that the problem is governed by a single failure mechanism. In fact, for the given distributions, the limit state for bearing capacity $g_2(x)$ governs the system reliability with a reliability index $\beta_{g_2} = 2.917$, while the other reliability indices are much higher ($\beta_{g_1}, \beta_{g_3} > 5.0$).

In terms of efficiency, the results show clearly that (adaptive) importance sampling is the most efficient sampling-based method for this problem, which is partly because the design point can accurately be determined using FORM in order to define a suitable importance sampling distribution.

The evaluation of reliability methods on more geotechnical benchmark problems can be found on the website of TG304¹¹ and in Muhammad Rayyan (2021).

6.4.3 FORM design point and influence factor

The results of the first order reliability method analysis are used to showcase the design point as the most-likely combination of parameters leading to failure (see O). The design point values for the stochastic parameters x_i^* , the design point values in standard-normal space u_i^* , and the corresponding alpha influence factors α_i^* are given in **Table 27**.

Table 27. Design point values and influence factors for GEOSNet benchmark problem 3 based on the FORM analysis.

Random variable	Design point value x_i^*	Design point value u_i^*	Influence factor α_i^*
$\gamma_1 \sim N(19.0, 1.9)$	19.90 kN/m ³	0.471	-0.161
$\gamma_2 \sim N(17.0, 1.7)$	16.26 kN/m ³	-0.434	0.148
$\varphi_1 \sim N(35.0, 3.5)$	32.39 degree	-0.747	0.255
$\varphi_2 \sim N(35.0, 3.5)$	25.35 degree	-2.756	0.942

Source: Authors' own work

The uncertainty in the friction angle of the foundation soil φ_2 is dominating the limit state of the retaining wall, while the other stochastic parameters have a modest influence. The negative influence factor $\alpha_1 = -0.161$ for the unit weight of the backfill material γ_1 indicates that it acts as a load variable. In contrast, the unit weight of the foundation soil γ_2 has a positive alpha factor, and an increase in γ_2 meaning that the foundation soil body contributes to the resistance against failure.

6.5 Random finite element method (RFEM)

Random fields can be used to explicitly model the spatial variability of material properties and the effect on the limit state function. Using finite element models enhanced with random fields for the spatial variation of material properties is known as the random finite element method (RFEM). Random fields representing the spatial variability of soil properties are mapped on the integration points of the elements. In contrast to the single random variable (or effective distribution) approach, the random field approach results in a random variable for each integration point in the finite element model. This means that the dimensionality of the reliability problem increases to thousands of variables, and Monte Carlo analysis or subset simulation are the most suitable reliability methods for evaluating probabilities of failure.

The use of the explicit modelling of the spatial variability and the resulting structural response in RFEM, provides predictions of structural performance and the resulting reliability estimates with the least strong assumptions on the effects of spatial variability. As such, RFEM results can serve as reference solutions for approximating approaches with stronger assumptions, such as simplification

¹¹ <http://140.112.12.21/issmge/tc304.htm>

to a single random variable. Ideally, a single random variable approach is based on, or calibrated against, RFEM analyses.

Example 6.9 shows the results of an RFEM analysis in a reference situation in which the entire soil domain is considered as a single material in which the strength parameters (undrained shear strength in this case) are spatially variable. The resulting distributions of the factor of safety (one considering the point statistics combined for a homogeneous soil and one based on RFEM analysis) shows that the results of RFEM have a lower mean in combination with a lower variance. The lower mean is due to the weakest path seeking of the failure mechanism, the lower variance due to the spatial averaging over that failure mechanism.

6.5.1 Software

The random finite element method is an advanced modelling framework, requiring dedicated finite element software. The crucial part is the possibility of mapping random fields onto the finite element mesh, ideally to individual integration points. This can be implemented as an integral part of the software, or come as an interface option of the finite element software to external toolboxes or scripts. Open-source implementations of academic RFEM codes are available as software accompanying Fenton and Griffiths (2008). Many commercial (geotechnical) finite element codes have some form of random field analysis implemented (e.g. Optum+, DIANA), or allow work-arounds to introduce spatial variability in state variables (e.g. through the user-defined soil models in PLAXIS (Pană, 2022) or through the UMAT-files for user-defined constitutive relations in ABAQUS).

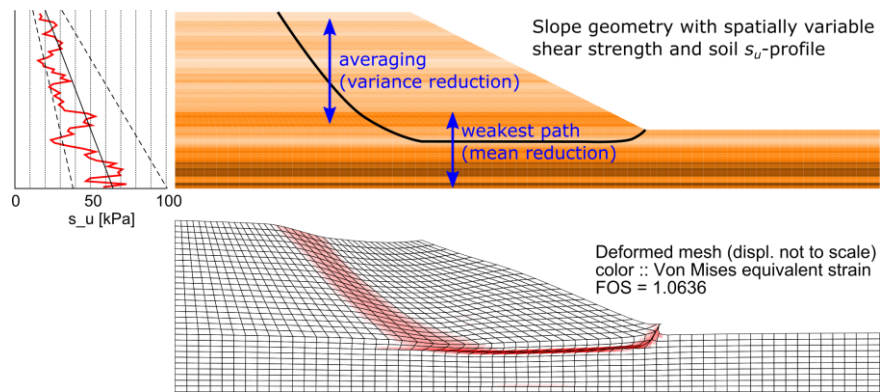
Box 25. Example 6.9: Spatial averaging and attraction to weak zones for slope stability

A 10 meter high 1:2 slope in a spatially variable cohesive soil is analysed under undrained conditions. The undrained shear strength is considered to have a mean trend equal to $s_{u,mean}(z) = 20 + 3 \cdot z$ kPa, with z the depth below the top of the slope. Around this mean trend, the undrained shear strength has a lognormal distribution with coefficient of variation $V_{s_u} = 0.3$ relative to the mean trend (i.e. the standard deviation is depth-dependent). Spatial variability of $\ln(s_u)$ is modelled by means of a Gaussian random field with a Markov correlation function $\rho(\Delta z) = \exp(-2 \cdot \Delta z / \delta_v)$, parameterized by the vertical scale of fluctuation $\delta_v = 2.0$ m. Variability of the soil in the horizontal direction is not considered. The random finite element method is used to evaluate the distribution of the factor of safety against slope failure due to the uncertainty in the undrained shear strength, explicitly accounting for the spatial variability of the soil. The results are compared with those based on homogeneous layers, in which the effects of heterogeneity is not accounted for (i.e. with a linear trend, $V_{s_u} = 0.3$ and $\Gamma^2 = 1$).

Figure 49 shows one realization evaluated within the RFEM framework, with the vertical profile of undrained shear strength and the failure mechanism in the deformed finite element mesh. In this realization, the failure surface cuts through the horizontally layered soil in the upper part of the domain, which implies averaging over the spatially variable soil strength profile. The lower part of the failure surface follows the weakest layer in the lower part of the domain. As a result of this attraction of the failure surface to the weakest path, the mobilised strength along the failure surface tends to be lower than the mean strength. The combined effect of this averaging and weakest path seeking results in a distribution of the structural performance (here in terms of FS) with a lower mean and a lower variance.

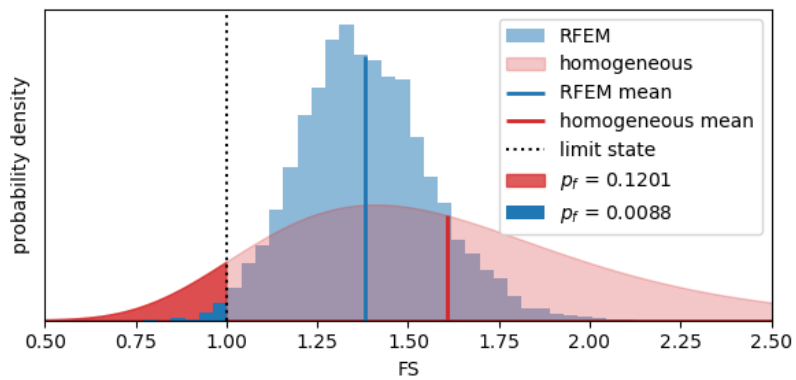
This can be seen in **Figure 50**, in which the distribution of FS is shown with and without accounting for spatial variability. The RFEM results shows a narrower distribution with a relative shift to the left. As a result, FS is on average lower than calculated based on homogeneous layers because of the mean reduction, but the calculated probability of failure is still lower due to the variance reduction.

Figure 49. One realization of a slope in RFEM with spatially variable undrained shear strength.



Source: Authors' own work

Figure 50. Distribution of FS with (RFEM) and without (homogeneous) accounting for spatial variability



Source: Authors' own work

7 Bayesian analysis of geotechnical data

7.1 Introduction

Geotechnical data or measurements of the structural behavior become available before design, during construction and during operation of a geotechnical structure. Such data represent evidence and reduce uncertainties. For example, site investigation data collected at the project site can provide a more refined description of the geotechnical parameter uncertainty, which corrects or updates the prognosis of a desk study. As well, performance observations like displacement measurements during construction or operation of a retaining wall are informative for the reliability of the structure during the (remaining) service life. Bayesian statistics is the paradigm that allows for integrating prior knowledge with available evidence (i.e. new data) to the end of updating the probabilistic model of the structural reliability.

This chapter starts with an introduction to Bayesian statistics (7.2). Then, it describes the steps for performing Bayesian inference 7.3. Consequently, it details typical applications of Bayesian statistics in geotechnical engineering and provides relevant worked examples (7.4-7.6). Lastly, the chapter describes how Bayesian statistics can be applied in the context of the Observational Method towards an integrated approach of feedback-oriented design and construction.

7.2 Principles and definitions of Bayes' theorem

In probability theory, Bayesian statistics is used to describe the probability of an event based on the available evidence and the prior probability of the event. The latter is also the main point of difference between Bayesian and Frequentist statistics, i.e. the consideration of prior knowledge. Bayesian inference is the statistical inference method adopting Bayes' theorem (**Equation 65**).

$$P(\mathbf{E}|\varepsilon) = \frac{P(\varepsilon|\mathbf{E})P(\mathbf{E})}{P(\varepsilon)}$$

Equation 65.

Bayes' theorem is composed of the following components:

- The examined event is denoted by \mathbf{E} . Such an event can be the failure of a slope, the exceedance of a displacement threshold for a retaining wall, or even a geotechnical parameter taking a specific value.
- The evidence, denoted by ε , can take the form of data such as the occurrence of failure, measurements of structural response or site investigation.
- The posterior probability $P(\mathbf{E}|\varepsilon)$ describes the updated knowledge about event occurrence, conditioned on the evidence.
- The prior probability $P(\mathbf{E})$ describes the initial knowledge about event occurrence, before any evidence has been considered.
- The likelihood function $P(\varepsilon|\mathbf{E})$ describes the probability that the evidence is true in case the event is true.
- The evidence term $P(\mathbf{E})$ states the probability of observing the evidence. It is treated as a normalizing constant, which formally is defined as $P(\mathbf{E}) = \int P(\varepsilon|\mathbf{E}) P(\mathbf{E})d\mathbf{E}$.

The presented description of Bayes' theorem updates the probability of event occurrence given some evidence and so is expressed in terms closer to probability theory. From the statistical standpoint, Bayes' theorem can be rephrased according to **Equation 66**. In this way, Bayes' theorem can be used in updating knowledge on the parameter set θ of a statistical model. The parameter set is conditioned to data y (which assumes the role of evidence in updating) that can be described by said statistical model through the likelihood function $L(y|\theta)$, which expresses the likelihood of obtaining the data given the parameter set θ is true. The prior and posterior distributions ($f'(\theta)$ and $f''(\theta|y)$) respectively) now define the knowledge on parameter set θ before and after updating.

$$f''(\theta|y) = \frac{L(y|\theta)f'(\theta)}{\int L(y|\theta)f'(\theta)d\theta}$$

Equation 66.

The process of developing knowledge from the prior distribution to the posterior distribution through Bayesian inference is called Bayesian updating. The emphasis of Bayesian updating lies in combining the prior probability and the likelihood. Ultimately, Bayesian inference provides a posterior distribution that combines the prior knowledge and the information contained in the data.

Bayesian updating reduces epistemic (reducible) uncertainty, while aleatory (inherent, irreducible) uncertainty remains by definition. It is therefore essential to take care of the distinction in the formulation or the probabilistic or statistical model.

Box 26. Example 7.1

An introduction to Bayesian analysis is offered through a simple example.

Suppose that a medical test has accuracy 99% for a disease. This means that the test will show a positive result 99% of the times the patient has the disease and a negative result 99% of the times the patient does not have the disease. Also, assume that the disease inflicts 0.01% of the general population. In case a test is positive, what is the probability that the patient actually suffers from the disease?

For this problem, the examined parameter is whether the patient suffers from the disease (D) and can be either *True* or *False*. The evidence is that the test is positive:

$$\varepsilon = P$$

The prior distribution of D comes from the percentage that the disease is met in the general population

$$P(D = T) = 0.01\% \text{ and } P(D = F) = 0.99\%$$

With Bayesian updating, the probability that $P(D = True)$, given the test is $\varepsilon = Positive$. Applying Bayes' theorem:

$$P(D = T, \varepsilon = P) = \frac{P(\varepsilon = P|D = T)*P(D=T)}{P(\varepsilon=P)} \rightarrow$$

$$P(D = T, \varepsilon = P) = \frac{P(\varepsilon = P|D = T)*P(D=T)}{P(\varepsilon = P|D = T)*P(D=T)+P(\varepsilon = P|D = F)*P(D=F)}$$

$$P(D = T, \varepsilon = P) = \frac{99\%*0.01\%}{99\%*0.01\%+1\%*0.99\%} = 0.98\%$$

Bayes' theorem describes that the probability of having the disease with a positive test is the ratio of two values. The first one is actually suffering from the disease and receiving the positive test. The second is the probability of receiving a positive test, which can happen if the patient suffers from the disease and the test is accurate and the patient does not suffer but the test is inaccurate. After updating, it is revealed that

the patient has a probability of ~1% of suffering from the disease, even though the result is positive. The reason of such a low probability lies in the prior. Essentially, the disease is rare in the general population.

7.3 Steps for performing Bayesian inference

This section provides suggestions for setting up Bayesian models relevant to geotechnical engineering problems, performing inference and interpreting the resulting posterior.

7.3.1 Establishing prior distributions

Prior distributions are used to incorporate pre-existing knowledge into the updating procedure. A prior distribution can be described with regards to its power in expressing prior knowledge: it can be informative, expressing definite information about the variables, weakly informative, expressing partial information about the variables, or uninformative, expressing general to no information about the variables. Usually, informative priors are adopted in geotechnical engineering.

Informative prior distributions can be used to affect the updating outcome in two ways. Firstly, informative priors allow for the incorporation of prior knowledge in the updating procedure, which is the principal contribution of the prior in Bayesian updating. For example, such knowledge can originate from pre-existing databases, experience in similar cases, regional datasets, or expert opinion. Secondly, informative priors can be used to impose physical limitations on the updating. As a result, the posterior distribution assigns zero probability to variable values that are physically impossible to occur, e. g. negative friction angle values. Likewise, although weakly informative prior distributions only provide loose information on the variable, they are largely used to the same end, bounding the range of possible variable values. Bounding priors can be achieved by selecting specific distribution types (e.g. the lognormal distribution assigns zero probability to negative values), or by truncation of the distribution.

Table 18 suggests appropriate prior distribution types for geotechnical parameters, according to general knowledge. Since most geotechnical parameters cannot take negative values, the informative prior distributions have been adjusted accordingly. Parameters that typically have a low coefficient of variation, which implies that the probability of negative values is trivial, are assigned a normal prior distribution. In case negative values become an issue, the truncated normal distribution can be adopted, but cautiously with respect to its impact on Bayesian updating techniques. On the other hand, in case the coefficient of variation is significantly high, the lognormal distribution is adopted, which allows only positive parameter values by default.

Box 27. Example 7.2: Comparison of posteriors retrieved by updating with weakly informative and informative priors

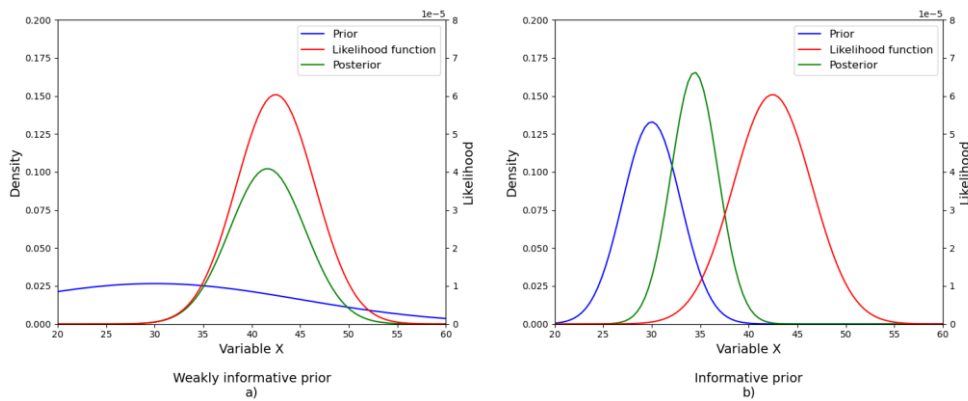
This example showcases the effect of prior distribution strength on the posterior using a simple Bayesian updating case. **Figure 51** illustrates the example, with plot a) showing the results when updating with a weakly informative prior and b) with an informative prior. Bayesian analysis focuses on updating variable X . The available data leads to the likelihood functions shown in **Figure 51**, which are the same in both plots. Both priors have the same mean, but the one of a) is considered as weakly informative for the problem, because its variance largely exceeds the area of interest for X . On the other hand, the prior in plot b) is stronger in conveying information for the updating, since most of its probability mass is located in a neighborhood relevant to the updating of X .

The impact of prior strength is evident in the plots. When a weakly informative prior is used, the likelihood function dominates updating. This means that data hold the greatest influence on the posterior, since the

prior information does not contribute decisive information for updating. As a result, the likelihood function constrains the posterior distribution in its neighborhood.

On the other hand, an informative prior is able to mitigate the influence of the likelihood function by introducing firm prior knowledge into the updating. As seen in plot b), the informative prior is able to draw the posterior distribution from the neighborhood of the likelihood function. This means that prior knowledge is actually contributive to updating, which is not solely determined by the data. Ultimately, updating is influenced by the relative strength of the prior knowledge and the information carried by the data, expressed through the likelihood function.

Figure 51. Comparison of posteriors after updating with a) weakly informative and b) informative priors.



Source: Authors' own work

Box 28. Example 7.3: Comparison of posteriors retrieved by unbounded and bounded priors.

This example showcases the effect of prior distribution on the posterior when the former is bounded at values that hold no geotechnical meaning.

Bayesian analysis is performed for the dry unit weight (γ_{dry}) of soft soil. Two different approaches are examined; in the first, a normal prior distribution is adopted, while the second case uses a normal distribution with truncation at $\gamma_{dry} = 0$ as a prior.

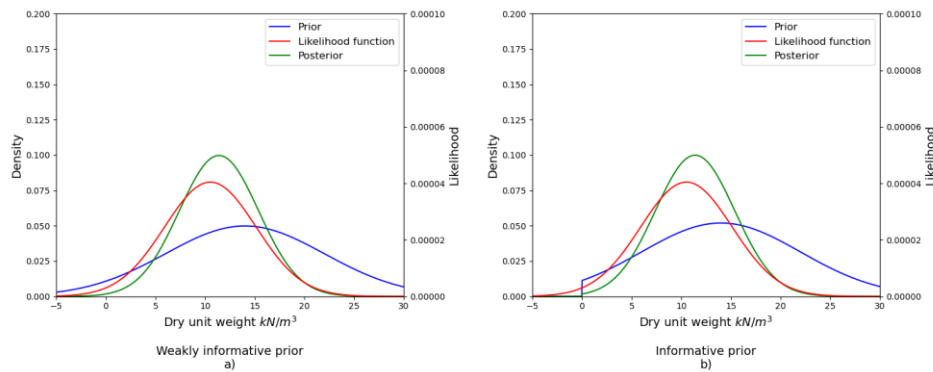
The results of updating are shown in

This example employed a normal truncated distribution for the sake of presenting the effect of using a bounded prior distribution by comparing two priors which apart from the truncation behave similarly. In practical settings, bounded prior distribution are achieved by utilizing other distribution types, such as the lognormal distribution.

Figure 52. In plot a) the likelihood function provides a non-zero likelihood score at γ_{dry} values that are negative. The normal prior distribution enables updating to consider this region of the γ_{dry} domain. Eventually, the posterior accredits a probability to negative γ_{dry} values. On the other hand, a normal distribution truncated at $\gamma_{dry} = 0$ prohibits updating from assigning probabilities to that region. As shown is plot b), even though the likelihood function enables negative γ_{dry} values, the prior does not. Thus, the posterior assigns a probability density of zero at the $\gamma_{dry} < 0$ region.

This example employed a normal truncated distribution for the sake of presenting the effect of using a bounded prior distribution by comparing two priors which apart from the truncation behave similarly. In practical settings, bounded prior distribution are achieved by utilizing other distribution types, such as the lognormal distribution.

Figure 52. Comparison of posteriors after updating with a) untruncated and b) truncated normal prior distributions.



Source: Authors' own work

7.3.2 Formulating the likelihood function

The likelihood function, typically formulated as a probability mass or density function (PDF/PMF), describes the information contained in the data. Essentially, the likelihood function assesses how likely the parameters are, given a set of observations, by evaluating the probability density of the observations for all possible distribution parameter sets. In other words, the likelihood describes the (relative) probability that the parameters are true. Additionally, the likelihood function can incorporate information on the error of the observed value.

The distribution parameter set is defined by the variables of the Bayesian statistical model. In many cases, the variables are not in the same terms as the observations, which means that a transformation is imminent. For example, in case the variable of the model is the compressibility of the soil and the observation is the settlement at the end of the soil improvement phase, the settlement model must be evaluated to transform a compressibility value to a settlement value. Thus, the settlement model is part of the likelihood function evaluation and introduces engineering performance into statistical inference.

Box 29. Example 7.4: Implementation of a likelihood function

This example elaborates how the likelihood function is formed in a case of Bayesian updating for the volumetric weight (γ) of a soft soil.

The example assumes that 5 measurements of γ are available at a site, which takes the values: $y = [11.48, 22.66, 20.55, 16.26, 10.68]$. The goal of updating is to derive the distribution of the mean unit weight in the site ($\bar{\gamma}$). The measurements are assumed to be independent, which means that the total likelihood is the product of the individual likelihoods per measurement. Thus, the likelihood function can be defined according to

Equation 67, where σ_y is taken as constant and equal to the standard deviation of the measurements ($\sigma_y = 4.70$).

$$L_y(\bar{\gamma}) = \prod_{i=1}^5 N(y_i | \mu_{\bar{\gamma}}, \sigma_y)$$

Equation 67.

Additionally, measurements may not be independent. This scenario is highly relevant when spatial variability is considered. In this case, a joint distribution should be used for the likelihood function. **Equation**

68 displays a case where measurements follow a multivariate normal distribution, described by $\mu_{\bar{y}}$ and a covariance matrix C , which assumed to be known.

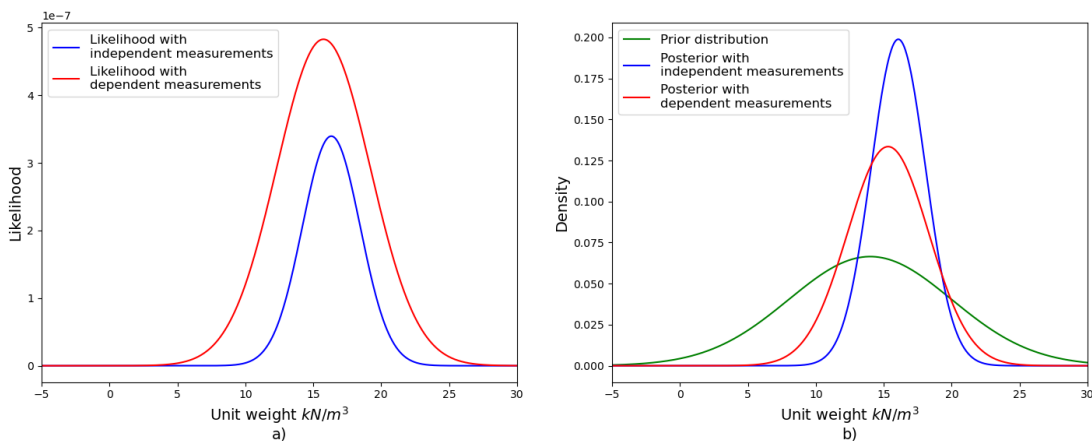
$$L_y(\bar{y}) = MVN(y|\mu_{\bar{y}}, C)$$

Equation 68.

$$C = \begin{bmatrix} 1 & 0.18 & 0.31 & 0.17 & 0.87 \\ 0.18 & 1 & 0.59 & 0.79 & 0.16 \\ 0.31 & 0.59 & 1 & 0.54 & 0.28 \\ 0.17 & 0.79 & 0.54 & 1 & 0.15 \\ 0.87 & 0.16 & 0.18 & 0.15 & 1 \end{bmatrix} * \sigma_y^2$$

Figure 53 illustrates the likelihood function per case. A common prior is assumed for both cases, which is normally distributed with mean $\mu_{\bar{y}} = 16 \text{ kN/m}^3$ and standard deviation $\sigma_{\bar{y}}' = 6 \text{ kN/m}^3$. Plot a) exhibits that the likelihood function of the independent measurement case is narrower and thus inflicts stronger restrictions to updating. Since the correlation matrix is known, the using dependent measurements accredits relatively high likelihood scores to a broader range of $\mu_{\bar{y}}$. Thus, as shown in plot b), the case with dependent measurements leads to a wider posterior distribution than the one achieved in the independent measurement case. Moreover, adopting dependent measurements translates the center of the likelihood function and posterior distribution. This effect originates from the covariance matrix.

Figure 53. a) Likelihood functions and b) posterior distributions for independent and dependent measurements.



Source: Authors' own work

7.3.3 Bayesian inference (updating)

Several methods are available for performing Bayesian inference in differing effectiveness, applicability and complexity. Selecting an updating method is specific to the problem examined and the available computational resources. The most common methods are listed below.

7.3.3.1 Numerical integration

Numerical integration is the numerical approximation of Bayes' theorem. A mesh is created over the variable domain and the likelihood function and the prior distribution are calculated at every node. The posterior density of at each node of the mesh is the product of the likelihood and the prior values, normalized by the integral of the product over the mesh. Numerical integration can be

appropriate for low numbers of variables, but becomes computationally less attractive for say more than 2 variables and/or very dense grids.

7.3.3.2 Conjugate priors (analytical solution)

In some cases, an analytical formulation exists for the posterior distribution through the use of conjugate priors (Ang & Tang, 2015; Howard Raiffa & Robert Schlaifer, 1961), when the combination of the prior distribution and the likelihood function leads to a posterior distribution of known type with analytical formulation. Valid combinations are available in literature.

A simple case of Bayesian updating using conjugate priors, and apt for application to soil properties, is inferring the posterior mean (μ'') using n normally distributed observations with a sample mean of $\bar{\varepsilon}$. Firstly, a normal prior distribution parametrized by μ' and σ'^2 is assumed, representing prior knowledge of the mean (e.g. the layer average). Secondly, a normal likelihood function is adopted with a known variance of σ^2 (representing the site variability). The posterior distribution of μ' is then normal distributed and described by the mean (μ'') and standard deviation (σ'') given by **Equation 69** and **Equation 70**.

$$\mu'' = \frac{\frac{\mu'}{\sigma'^2} + \frac{n\bar{\varepsilon}}{\sigma^2}}{\frac{1}{\sigma'^2} + \frac{n}{\sigma^2}}$$

Equation 69.

$$\sigma'' = \left[\frac{1}{\sigma'^2} + \frac{n}{\sigma^2} \right]^{-1/2}$$

Equation 70.

7.3.3.3 Sampling algorithms

Sampling algorithms perform Bayesian inference by directly sampling from the posterior distribution. After collecting the sample, the posterior distribution can be approximated in parametric form. The most common algorithm meeting is Markov Chain Monte Carlo (MCMC), especially the Metropolis-Hastings and Gibbs samplers. Hamiltonian Monte Carlo is another efficient sampling algorithm. While more efficient than numerical integration, sampling algorithms may still be prone to long calculation times. Also, for application in a reliability analysis context, it can be challenging to obtain good posterior distributions in the relevant distribution tails. Some open-source packages that provide sampling algorithms are: Stan (Stan Development Team, 2020), PyMC (Salvatier et al., 2016) and Tensorflow Probability (Google research, 2015).

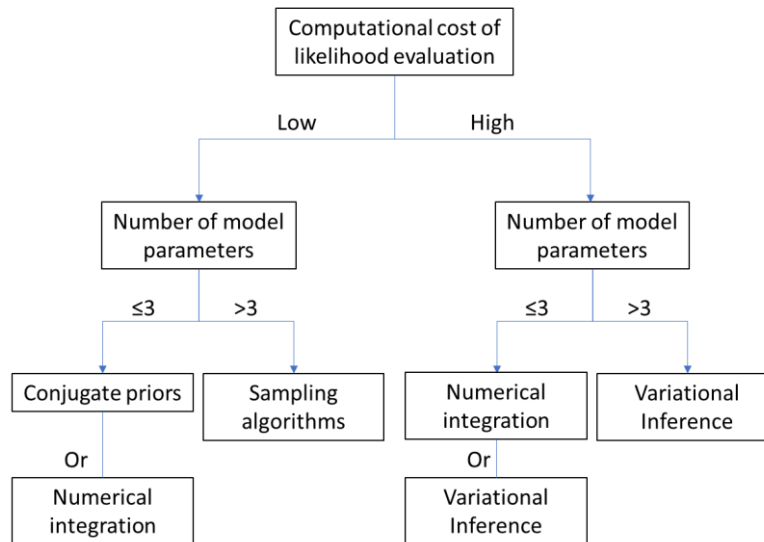
7.3.3.4 Variational Inference

Variational inference is an approximation technique that replaces the unknown posterior distribution with a new distribution of known formulation. It then reframes Bayesian inference as an optimization problem, aiming for the best approximate solution. The objective of optimization is the maximization of information gain from the evidence.

7.3.3.5 Selecting a method

Selecting a method depends on the number of parameters in the statistical model, its complexity, the assumptions that can be adopted and the practical requirements of the performance of the method, such as computational power and time restrictions. Method suggestions with regards to characteristics of the examined updating problem are given by the flowchart of **Figure 54**.

Figure 54. Flowchart for Bayesian inference method selection.



Source: Authors' own work

In inference, typically the likelihood function is evaluated repeatedly and requires the calculation of a geotechnical model. Hence, the computational cost of evaluating the likelihood function can be considerable depending on the model used. Surrogate modelling can then be an efficient solution. The additional uncertainty introduced can be reflected as observation uncertainty (similar to measurement error) in the likelihood function.

7.3.4 Interpreting the posterior distribution

The posterior distribution is the outcome of the inference process and describes the updated knowledge on the parameters, after incorporating information from the data, the likelihood function and the prior distribution.

Multiple Bayesian statistical models should be examined, since several models can achieve adequate description of the data. These models need to be compared, to the end of selecting the most appropriate one. A technique for comparing the statistical models involves withholding part of the dataset from the inference step and preserving it for model checking (test data). Ultimately, the model that provides the best description of the test data should be selected, or in other words, the model that assigns the greatest posterior predictive values to the test data (posterior predictive check) (Gelman et al., 2013). Another metric for model comparison is the Bayesian Information Criterion (Claeskens & Hjort, 2001), which penalizes the fit of the data by the complexity of the statistical model. Essentially, if compared models achieve a similar description of the data, simpler models are favoured.

7.3.4.1 Posterior and posterior predictive distributions

The posterior and the posterior predictive distributions are distinguished. The posterior distribution provides the probability of the updated parameter conditioned on the observed evidence. On the other hand, the posterior predictive distribution describes the probability of possible unobserved or future data (y'), based on the posterior distribution and so, the observed evidence. The posterior predictive $f'(y'|y)$ is given by **Equation 71**, where $f(\theta|y)$ is the posterior distribution of θ for evidence y and $f'(y|\theta)$ is the predictive distribution of new data conditional to θ . The distinction between the posterior and posterior predictive distributions is important to geotechnical engineering. This is explained in example 7.5.

$$f'(y'|y) = \int_{\Omega_{\theta}} f(y'|\theta) * f''(\theta|y) d\theta$$

Equation 71.

Box 30. Example 7.5: The importance of posterior and posterior predictive in geotechnical engineering

This example aims to clarify the distinction between the posterior and posterior predictive distributions and highlight their importance through application in a geotechnical setting.

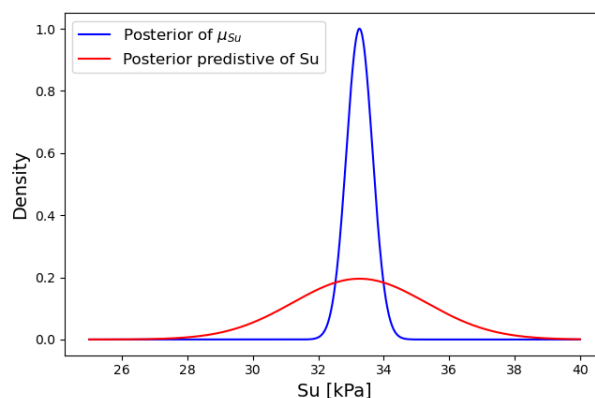
The example assumes a case of updating the mean of the undrained shear strength (S_u) of a site (μ_{S_u}). The data (y_{S_u}) is assumed to follow a normal distribution with a fixed standard deviation (σ_{S_u}), so the likelihood function for the model is given by:

$$L(\mu_{S_u}) = N(y_{S_u} | \mu_{S_u}, \sigma_{S_u})$$

The likelihood function quantifies how well a value of μ_{S_u} can describe the data. The distribution used in the likelihood function is the predictive distribution and connects the data to the statistical model parameters. Essentially, this distribution can be used to generate new data or predict the future data. In order to derive the posterior predictive distribution, the posterior distribution needs to be integrated with the predictive distribution to account for uncertainty in μ_{S_u} .

After performing Bayesian updating, the posterior distribution of μ_{S_u} is retrieved (**Figure 55**). Using **Equation 74**, the predictive distribution is formed. Its illustration is also provided in **Figure 55**. But why is the distinction between the two important to geotechnical engineering?

Figure 55. Posterior distribution of μ_{S_u} and posterior predictive distribution of S_u .



Source: Authors' own work

In this example, the posterior distribution expresses the belief regarding the μ_{S_u} of the site, while the posterior predictive distribution indicates the probability that S_u values appear upon further sampling from the same site. In geotechnical terms, the posterior distribution reflects the mean statistics for S_u , while the posterior predictive reflects the point statistics.

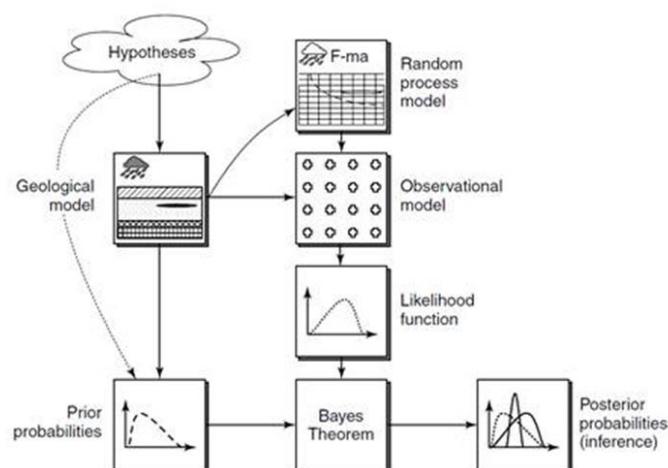
In case a suitable characteristic value has to be selected, the effect of averaging determines the relevant distribution. The two distributions represent the effective S_u distribution for the two extreme cases of averaging. The mean statistics are effective when the soil is heterogenous and averaging is strong, while the point statistics are effective for homogenous soil with no averaging. Then, the characteristic value can be selected as an appropriate estimate of the relevant distribution. In non-extreme cases, determining the effective distribution is not as straightforward and further evaluation is required.

7.4 Bayesian inference in geotechnical applications

Since ground is a highly variable material, considerable levels of epistemic uncertainty remain with the typical amounts of site investigation available. Bayesian inference can reduce that uncertainty. Moreover, since geotechnical structures are typically one-of-a-kind and engineering judgment plays an important role, the frequentist approach is of limited use (see e.g Vrouwenvelder (2002), Baecher (2021), Baecher & Christian (2003)) and Bayesian techniques have clear advantages.

Site investigation is an obvious application for Bayesian inference, as pointed out by Baecher and Christian (2003) and illustrated in **Figure 56**. Existing knowledge of or experience with ground parameters can be updated by performing Bayesian updating using site investigation data, thus merging the soft expert knowledge with the hard data from the site. Bayesian applications on site investigation data can be divided in two groups: (i) Bayesian parameter estimation, which aims to update ground parameter distributions; and (ii) data-driven site characterization (DDSC), which focuses on establishing relationships between ground parameters and map parameter distributions over the subsurface of the site.

Figure 56. Schematic logic of probabilistic site characterization (Source: Baecher & Christian 2003)



Source: Baecher & Christian 2003

Data-driven Site Characterization encompasses techniques that use data of geotechnical site investigation towards deriving relevant geotechnical parameters, as well as producing a parameter map of the subsurface. Unfortunately, restrictions such as limited site investigation data availability

and sparse measurements render most traditional methods inadequate. The Bayesian framework can deal with small site investigation datasets and robustly quantifies the uncertainty in inferred parameters, as well as the prediction. This feature is essential for subsequent uncertainty quantification in the response of the structure and the assessment of its reliability. DDSC techniques have been focusing on inferring the parameters of the statistical model and eventually predicting geotechnical parameters over a subsoil domain. To that end, several DDSC methods focus on the recognition of spatial variability patterns, which enable such predictions. The reader is pointed to the work of (Ching & Phoon, 2019; K. K. Phoon & Ching, 2021; K.-K. Phoon et al., 2022) for further information on DDSC methods.

Bayesian analysis is the possibility to incorporate performance observations in a quantitative fashion. Performance observations can be observations, measurements or inspection data of the structural response to (loading) conditions. A classical example is the observed survival of a (significant) load, leading to increased reliability estimates through reduced uncertainties. Also application of the observational method can be implemented in a Bayesian framework.

Some typical applications of Bayesian updating in geotechnical design with potential benefits are presented in more detail in the following sections.

7.5 Bayesian parameter estimation

This section is based on examples 7.6–7.8 and presents an example of Bayesian parameter estimation in a practical setting. Specifically, the example demonstrates how several updating methods (conjugate priors, numerical integration, sampling) can be used to update the mean friction angle (φ) distribution of a soil layer when lab test data become available.

Box 31. Example 7.6: Bayesian updating of the friction angle with site investigation data using the conjugate priors method

This example combines prior knowledge with site investigation data (y) of the friction angle φ through Bayesian updating using the conjugate prior modelling. The quantity of interest is the mean friction angle of the site.

The mean friction angle φ of a non-cohesive soil layer at a geotechnical site is modelled with a normal prior distribution with a mean of $\mu_{\varphi}' = 27.0^\circ$ and standard deviation $\sigma_{\varphi}' = 3.0^\circ$, according to regional data (measurements at nearby sites in the same geological deposit). Soil samples are collected from the specific site and 6 direct shear tests are performed with the following results: $y = [30.7^\circ, 29.6^\circ, 27.7^\circ, 28.3^\circ, 31.8^\circ, 29.7^\circ]$. The dataset is assumed to be normally distributed and the sample standard deviation is $\sigma_y = 2.22^\circ$.

As a reference, the characteristic value as derived from the prior is equal to 22.07° , being the 5%-quantile of the prior distribution.

Before performing Bayesian parameter estimation, the knowledge on the friction angle is updated using the Maximum Likelihood approach, which derives the mean friction angle value that is most likely to be true according solely to the data. In this simple case, the Maximum Likelihood implementation is equivalent to deriving the moments of the lab test sample. The mean of sample is $\bar{y} = 30.53^\circ$ and the standard deviation of the mean is: $\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{n}} = \frac{2.22^\circ}{\sqrt{6}} = 0.91^\circ$, where n is the number of tests in the sample. Eventually, the characteristic value is calculated as: $\bar{y}_{char} = \bar{y} - 1.64 \sigma_{\bar{y}} = 29.04^\circ$

The resulting posterior distribution of the mean friction angle is again a normal distribution with posterior mean:

$$\mu_{\varphi}'' = \frac{\frac{\mu_{\varphi}' + n\bar{y}}{\sigma_{\varphi}'^2 + \sigma_y^2}}{\frac{1}{\sigma_{\varphi}'^2} + \frac{n}{\sigma_y^2}} = \frac{\frac{27.00 + 6 \cdot 30.53}{3.00^2 + 2.22^2}}{\frac{1}{3.00^2} + \frac{6}{2.22^2}} = 29.55^{\circ}$$

And the standard deviation of the posterior mean is:

$$\sigma_{\varphi}'' = \left[\frac{1}{\sigma_{\varphi}'^2} + \frac{n}{\sigma_y^2} \right]^{-1/2} = \left[\frac{1}{3.00^2} + \frac{6}{2.22^2} \right]^{-1/2} = 0.55^{\circ}$$

Finally, the characteristic value of the posterior mean friction angle amounts to $\varphi_k = 28.64^{\circ} \gg 29^{\circ}$.

Box 32. Example 7.7: Bayesian updating of the friction angle with site investigation data using the numerical integration method

This example derives the posterior probability of the mean friction angle by applying Bayes theorem through numerical integration. The lab test data is the same as in the previous example. The variable of interest is the mean of the friction angle (μ_{φ}).

Following the assumptions made in the previous example, the prior distribution is normal, centered around $\mu_{\varphi}' = 27.0^{\circ}$ and has a standard deviation of $\sigma_{\varphi}' = 3.0^{\circ}$. The likelihood function uses a standard deviation equal to the one of the sample ($\sigma_y = 2.22^{\circ}$). A grid 1,000 points is defined for μ_{φ} in the range of $[15.0^{\circ}, 45.0^{\circ}]$. Determining the mesh for the parameter considers aspects such as: areas where the variable is expected to give meaningful probability and likelihood scores, sufficient grid discretization, computational effort. If computation time is not an issue, a regular (evenly spaced) grid can be used.

Following, the prior distribution and likelihood function are calculated per mesh point. Since observations are assumed to be i.i.d. (i.e. no correlation between lab tests), the multivariate normal likelihood function can be evaluated as the product of univariate normal likelihood functions, with each one defined per observation.

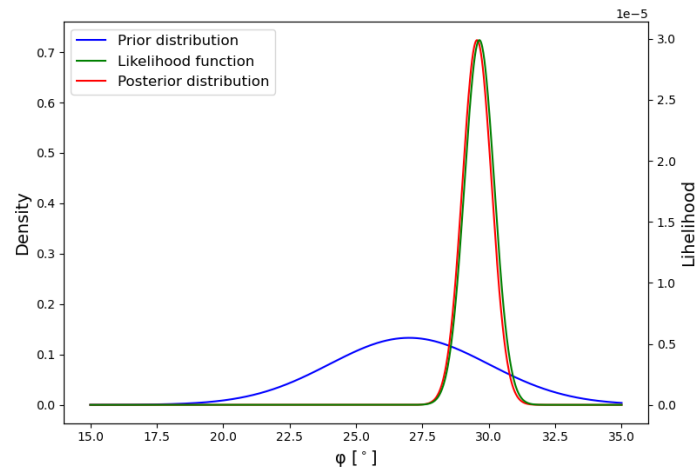
$$\text{Prior: } P(\mu_{\varphi}) = N(\mu_{\varphi} | \mu_{\varphi}', \sigma_{\varphi}')$$

$$\text{Likelihood function: } L_y(\mu_{\varphi}) = \prod_{i=1}^n N(y_i | \mu_{\varphi}, \sigma_y)$$

The calculation results are presented in **he impact** of the lab test data.

Figure 57. The posterior distribution moves away from the prior distribution and closer to the neighborhood of the mesh where the greatest likelihood scores are met. This means that the likelihood function is significantly stronger than the prior distribution; essentially, the prior knowledge on φ is not influential enough to μ_{φ}'' when compared to the impact of the lab test data.

Figure 57. Prior and posterior distribution and likelihood function over the calculation mesh.



Source: Authors' own work

Again, the characteristic value of the mean friction angle is estimated as the 5th percentile of the posterior distribution of μ_{φ} . This leads to: $\varphi_k = 28.63^\circ \gg 29^\circ$.

Box 33. Example 7.8: Bayesian updating of the friction angle with site investigation data using a sampling method

This example performs Bayesian parameter estimation by drawing samples directly from the posterior distribution using the Markov Chain Monte Carlo (MCMC) sampling algorithm. The lab test data is the same as in the previous examples.

The formulation of the model is same as in the previous example. **Equation 72** sets the prior distribution for μ_{φ} , while **Equation 73** shows the implementation of the likelihood function.

$$\mu_{\varphi}' \sim N(\mu_{\varphi}', \sigma_{\varphi}')$$

Equation 72

$$L_y(\mu_{\varphi}) = \prod_{i=1}^n N(y_i | \mu_{\varphi}, \sigma_y)$$

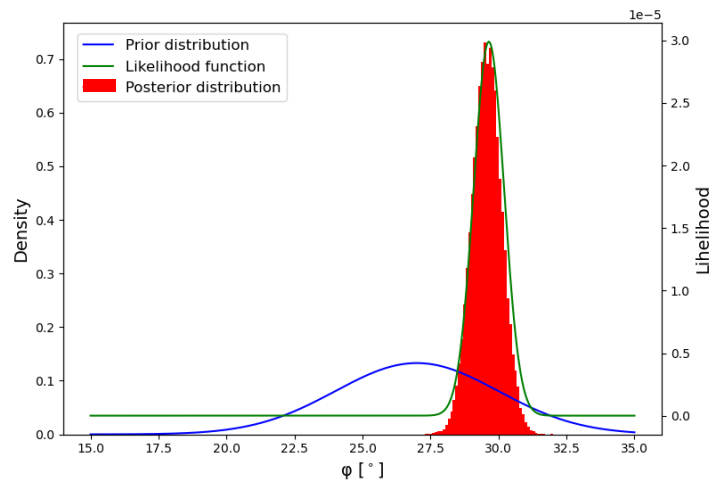
Equation 73

The resulting histogram of the posterior is shown in

Figure 58 and approximates the solution given in the previous example. Essentially, the MCMC sampler is able to draw samples directly from the prior, without the need to calculate the prior and likelihood and integrate over a mesh. For high-dimensional problems, the MCMC can lead to considerably less computational effort.

The resulting characteristic value of the friction angle is $\varphi_k = 28.67^\circ \gg 29^\circ$.

Figure 58. Histogram of posterior sample



Source: Authors' own work

The characteristic value results of the examples are collected in **Table 28**. All inference methods offer the same characteristic value, which is greater than the one suggested by the prior, because the data lies at greater φ values. Maximum Likelihood Estimation (MLE) (Wasserman, 2004) leads to the same characteristic value as Bayesian updating methods, even though it does not incorporate the influence of the prior distribution. This happens because the likelihood appears to convey strong information about φ and eventually overwhelms the effect of prior knowledge.

Table 28. Characteristic values according to the different Bayesian updating methods of the examples, assuming normally distributed data.

Characteristic value calculation method	Characteristic value (°)
Prior (expert judgment only)	22
MLE(data only)	29
Conjugate Priors	29
Numerical Integration	29
Sampling algorithm	29

Source: Authors' own work

As a result, Bayesian inference does not incorporate significant influence from the prior and is mostly based on the likelihood. The influence of the prior can only be observed as the difference between MLE and Bayesian methods in the second decimal of the characteristic value. It should be noted that theoretically the MLE leads to the exact same results as Bayesian methods if the latter were performed using uninformative priors. All Bayesian methods have led to the same characteristic values, with some approximation error being noticeable only in the second decimal. This example showcases how Bayesian analysis can integrate prior beliefs and available data for the site in a robust fashion and how the dynamics between the prior distribution and likelihood resolve.

7.6 Reliability updating

When (performance) information (ε) is obtained, the failure probability can be updated by **Equation 74**, which is a re-formulation of **Equation 45** by using the posterior distribution $f_{\theta|\varepsilon}$, as estimated with the evidence ε . This two-step procedure of first updating the probability distribution and subsequently re-evaluating the reliability is often called the ‘indirect method’.

$$P_{F|\varepsilon} = \int_{g(\theta) < 0} f_{\theta|\varepsilon}(\theta) d\theta$$

Equation 74.

The updated probability of failure $P_{F|\varepsilon}$ then can also be obtained using the definition of conditional probability (**Equation 75**), where $h(\cdot)$ is the observation function, which describes the relationship between the evidence and the engineering model. Two types of information are distinguished. In equality information the evidence implies that the observation is equal to a function of the variable state, such that $\varepsilon \equiv \{h(\mathbf{X}) = 0\}$. Examples are measured displacements or pore water pressures. On the other hand, inequality information the evidence implies that the observation is greater or lower than a function of the variable state, such that $\varepsilon \equiv \{h(\mathbf{X}) < 0\}$. Examples of inequality information are survival of a loading condition, or the (non-)exceedance of threshold values.

$$P_{F|\varepsilon} = \frac{P(F \cap \varepsilon)}{P(\varepsilon)} = \frac{P(g(\theta) < 0 \cap h(\theta) < 0)}{P(h(\theta) < 0)}$$

Equation 75.

This approach is called the ‘direct method’ because it uses a one-step procedure, without updating the joint probability distribution first. While the direct and indirect updating are mathematically equivalent, the direct method is relatively easier to implement, especially with simulation type of reliability analysis methods such as Monte Carlo simulation or Bayesian Updating with Structural reliability methods (BUS) (Straub & Papaioannou, 2015). The latter is able to focus and sample from areas of the posterior distribution that bear significant information for reliability estimation.

The following example showcases the application of direct updating in a pile survival testing case. A more elaborate example of pile survival testing is given in Annex B to this report.

Box 34. Example 7.9: Direct updating of pile reliability in axial loading

The following example demonstrates how the direct updating method can be used to update the reliability of a pile in axial loading. The prior capacity of the pile R follows a normal distribution with a mean of $150kN$ and a standard deviation of $30kN$. The load F on the pile follows a normal distribution with a mean of $80kN$ and a standard deviation of $20kN$. A proof load F_p is applied on the pile and the pile survives.

In this example, updating is possible with analytical solutions. However, a Monte Carlo simulation is chosen instead for showcasing the ease of application in problems with more complex definitions of failure. Pile capacity and load are sampled according to their respective distributions. The prior probability of failure is estimated as given by Crude Monte Carlo simulation; the resulting probability of failure is $Pf = 2.6\%$, which corresponds to a reliability index of $\beta = 1.94$.

1st case: Deterministic proof load (no measurement error)

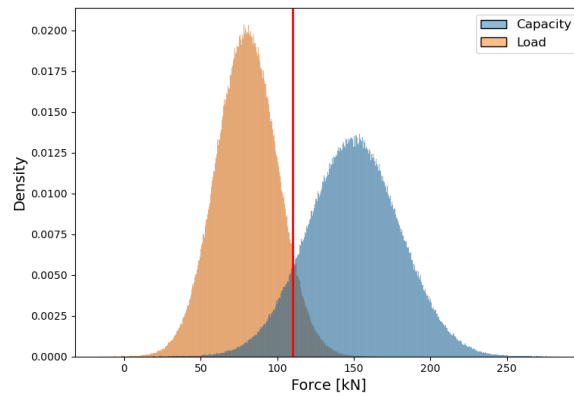
In this case, the proof load is measured to be $110kN$. Since the pile has survived this value, it means that its capacity is at least equal to the proof load and there is no probability of failure for loads lower than the proof load.

The sampling of the Monte Carlo simulation can be the same as for the estimation of the prior probability of failure. Modifying the direct updating formula to operate with Monte Carlo samples instead of analytical probability density functions yields the expression below, where S is the proof load measurement.

$$P_{F|\varepsilon} = P(F|\varepsilon) = \frac{P(R < F) \cap P(R > F_p)}{P(R > F_p)} = \frac{P(R - N < 0) \cap P(R - F_p > 0)}{P(R - F_p > 0)}$$

The updated probability of failure is: $P'_f = 0.48\%$, which corresponds to a reliability index of $\beta = 2.58$. The histograms of the analysis are visualized in **Figure 59**.

Figure 59. Histogram of pile load and capacity of Monte Carlo samples for case 1 (red line is for the proof load value).



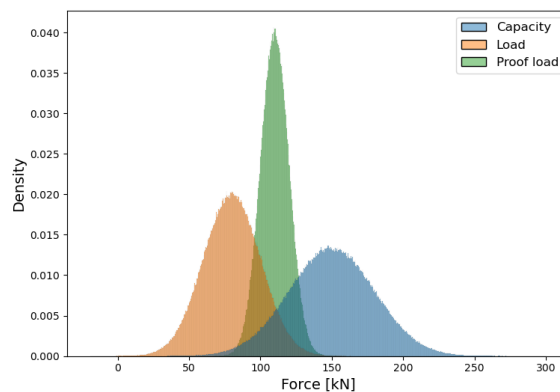
Source: Authors' own work

Box 35. Example 7.10: Direct updating of pile reliability in axial loading (continued)

2nd case: Stochastic proof load (with measurement error)

The second case solution follows the implementation described in the first case. However, this time the proof load measurement is also a variable, that follows a normal distribution with a mean of 110 kN and a standard deviation of 10 kN , and the Monte Carlo simulation should draw samples for it. Following, the same formula is applied to estimate the updated probability of failure as $P_{F|\varepsilon} = 0.58\%$, which corresponds to a reliability index of $\beta = 2.52$. The histograms of the analysis are visualized in **Figure 60**.

Figure 60. Histogram of pile load, capacity and test load of Monte Carlo samples for case 2.



Source: Authors' own work

7.7 Observational method

The Observational Method is a design verification method, according to which monitoring data collected during the construction phase of a project can be used to implement design changes during construction. According to the Observational Method (EN 1997-1, clause 4.7), a range of possible design variants shall be established, covering all foreseeable relevant ground responses and ground-structure interaction. Comparing monitoring results to established thresholds during construction, the appropriate design variant is implemented. A key component is the preplanning of contingency measures that ensure safe transition between design variants, should a threshold be violated. Typically, the initial design setting includes a considerable level of epistemic uncertainties, which are the primary reason to the initial design variant not meeting the requirements of a conventional design verification method. During construction, the ground behavior becomes apparent through monitoring data. Hence, uncertainty is reduced, and more accurate predictions of structural behavior can be made. Thus, the main premise of the Observational Method is highlighted: flexible design and feedback loops using monitoring data can lead to design optimization, as opposed to adopting a conservative design a-priori.

Bayesian updating can enrich the Observational Method with probabilistic information on the evolution of estimates and uncertainties. Monitoring data can be used to update the reliability of the structure. In this way, actual ground behavior is utilized to reduce epistemic uncertainty and raise prediction accuracy. Iterative application of this method means that the posterior estimated at one step of the project is used at the prior of the following step. Developing further, reliability can be connected to the cost of design measures, to define risk and enable decision making, using Bayesian decision theory (Spross & Johanson, 2017; Löfman & Korkiala-Tanttu, 2022).

Moreover, **Equation 76** enables the establishment of reliability-based thresholds, to be used when verifying structural safety with a reliability-based observational method, as detailed in (Spross & Johansson, 2017) and (Spross & Gasch, 2019). The threshold is obtained by equating $P_{f|\epsilon}$ with the target failure probability, P_{f_T} , and solving the equation for the threshold x_{alarm} , this being the only unknown.

$$P(g(X) \leq 0 | X \leq x_{\text{alarm}}) = P_{f_T}$$

Equation 76.

Conceptually, having a reliability-based threshold implies that the target failure probability (relevant for the construction phase) is violated only when the threshold is. Monitoring during construction can thereby be used as a means to verify that the structure achieves sufficient reliability, allowing for a design less conservative than the ones attained by conventional approaches.

Box 36. Example 7.11: Application of the Observational Method for the vertical deformation of a rock pillar

A very long rib pillar is to be excavated in rock (**Figure 61 a**). The limit state concerns its vertical strain ϵ_1 , which must not exceed a maximum strain $\epsilon_{1_{max}}$, i.e. $G(X) = \epsilon_{1_{max}} - \epsilon_1$, where X collects a number of underlying rock mass properties as random variables (the reader is referred to *Spross and Johansson (2017)* for the complete geotechnical model). In this case P_{f_T} is set equal to 0.001. Performing a reliability-based analysis (e.g. using Monte Carlo simulation) it is found that for an unsupported pillar, the probability of limit state violation at the end of the excavation process is $P_f = 0.0046$, which exceeds the target failure probability. However, if four horizontal rock anchors are added to the model to provide confinement, P_f becomes negligible. The decision maker now stands before two alternatives:

- Adopt the more expensive design using the rock anchors (i.e. a conventional design),

- Apply the Observational Method to reduce the uncertainty about the rock mass properties by monitoring the deformation of the unreinforced pillar during the excavation process.

The excavation is assumed to be carried out stepwise, making the pillar thinner and thinner with each step. Measurements are made after each excavation step, providing a number of measurements from which a trendline is derived to predict the final deformation, including the scatter.

If the actual ground conditions turn out to be favorable, rock anchors will not be needed after all. The alternatives are illustrated as a decision tree in **Figure 61 b)**, together with their corresponding costs. The cost components are given by (MU = monetary unit):

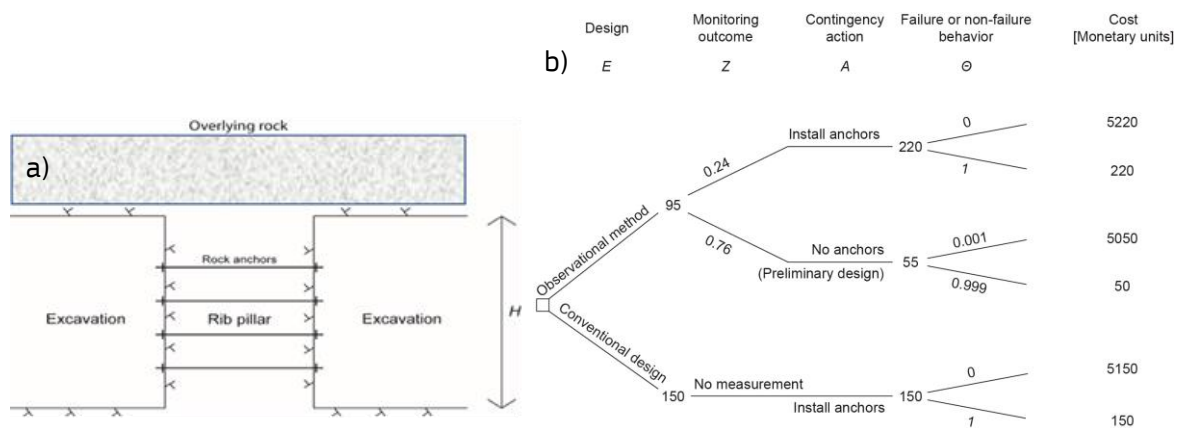
- anchor system: 150 MU is applied directly or 170 MU if applied later,
- advanced monitoring system: 20 MU
- cost increase due to more careful excavation using the observational method: 30 MU
- failure: 5000 MU

The threshold for the vertical strain without anchors is found from **Equation 76**, leading to $\varepsilon_{1alarm} = 0.33 \cdot 10^{-3}$ (corresponding to 2.7mm deformation). This allows evaluation of the probability of the unsupported design variant being successful: $P(\varepsilon_1 \leq \varepsilon_{1alarm}) = 76\%$. Hence, threshold violation that triggers the rock anchor installation occurs with 24% probability.

A so-called pre-posterior analysis is visualized by a decision tree in **Figure 61 b)** and shows that the expected cost of applying the observational method is 95 MU, while the expected cost of the conventional design is 150 MU. The analysis is performed by weighing the costs with respect to their probability of occurrence in the decision tree. Based on the result, a risk-neutral decision maker should choose the observational method. Note however that if the observations trigger the contingency measure to install rock anchors, the final cost becomes larger (220 MU) than that of the conventional design alternative. A risk-avoiding decision maker may therefore consider the conventional design better, as it comes with a known cost.

Note that in this case the cheapest solution is “to do nothing” and accept the high probability of failure leading to a cost expectation of $0.0046 \cdot 5000MU = 23MU$. Whether this is legally allowed may depend on the type of limit state (SLS or ULS) and the National Annex, as $P_f = 0.0046$ exceeds P_{fT} .

Figure 61. a) Section of the analysed rib pillar. b) For the analysed rib pillar, the observational method is found more favorable, because the expected cost of that design alternative (95 Monetary Units) is less than that of the conventional design (150 MU)



Source: Spross and Johansson (2017), CC-BY, <http://creativecommons.org/licenses/by/4.0/>

8 Reliability-based partial factors

This chapter describes the derivation or calibration of reliability-based partial factors for two purposes:

1. Code calibration: National standard bodies (NSB) or other competent authorities may support the choice of partial factors using reliability-based calibration. The aim is to find generally applicable partial factors which can be used for certain set of structure types and limit states. For example, partial factors for retaining wall design should work generally well on average in different design situations (wall material, soil conditions etc.).
2. Project-specific (or site-specific) partial factors: When the general-purpose partial factors in the Eurocodes (and national annexes) do not seem appropriate for a design or assessment, for example because the structure or uncertainties in design are unusual, partial factors can be derived for a specific project (or site) such that the underlying reliability requirements are met. Of course, project-specific calibration needs to be accepted by the relevant parties (e.g. the client, or competent authorities).

8.1 Semi-probabilistic approach

Verification of limit states using representative (or characteristic) values and partial factors is also referred to as the semi-probabilistic approach (see section 2.3.3). As laid out in section 2.1 and illustrated in **Figure 2**, the partial factor approach can be related to the reliability-based approach through the reliability targets, and to the risk-based approach through the potential consequences of failure.

8.1.1 Verification formats

The general reliability verification method in Eurocode (prEN 1990) is the semi-probabilistic approach via partial factor design format. The partial factor design must satisfy the following inequalities (**Equation 77** and **Equation 78**) for each specific design situation and corresponding limit states (ULS), load combination and load arrangement:

$$E_d \leq R_d$$

Equation 77. Ultimate limit states ULS

$$E_d \leq C_{d,ULS}$$

Equation 78. ULS caused by excessive deformation

where E_d is the design value of an action effect (ULS) or the relevant displacement or deformation (ULS), R_d is the design value of the corresponding resistance and $C_{d,ULS}$ is a corresponding limit value.

The design value for resistance R_d can be calculated according to EN 1990 as:

$$R_d = \frac{R \left\{ \frac{hX_k}{\gamma_m}; \alpha_d; \sum F_{Ed} \right\}}{\gamma_{Rd}}$$

Equation 79.

where:

- γ_{Rd} is partial factor accounting for uncertainty in the resistance model;
- η is a conversion factor accounting for scale effects, effects of moisture and temperature, effects of ageing of materials, and any other relevant parameters (not commonly used for geotechnical structures; i.e. $\eta = 1$);
- X_k represents the characteristic values of material of product properties (see prEN 1990 6.2(5));
- γ_m is the (material) partial factor ;
- $R\{\dots\}$ denotes the output of the resistance model;
- a_d denotes the design values of geometrical parameters;
- F_{Ed} denotes design values of actions used in assessment of E_d (F_{Ed} appears in the function as in some cases the resistance is dependent on actions).

For simplicity, the separate partial factors for resistance model uncertainty (γ_{Rd}) and material or product properties (γ_m) are usually combined into single partial material factor ($\gamma_M = \gamma_m \times \gamma_{Rd}$) or into a single partial resistance factor ($\gamma_R = \gamma_m \times \gamma_{Rd}$).

The design value of the effect of actions E_d for certain combination of actions can be calculated according to prEn 1990:

$$E_d = \gamma_{Sd} E \left\{ \sum (\gamma_f \gamma F_k); a_d; X_{Rd} \right\}$$

Equation 80.

where:

- γ_{Sd} is a partial factor that takes account of uncertainties in modelling the effects of actions;
- ψ is a combination factor either equal to 1.0 for permanent actions or as defined in prEn 1990 6.1.2.3 for variable actions;
- F_k is the characteristic value of an action;
- γ_f is a partial factor that takes account of unfavorable deviation of an action from its characteristic value;
- $E\{\dots\}$ denotes the combined effect of the enclosed variables;
- $\Sigma(\dots)$ denotes the combination of actions;
- a_d denotes the design values of geometrical parameters, defined in prEN 1990 8.3.7;
- X_{Rd} denotes the values of material properties used in the assessment of R_d (see prEn 1990 8.3.6). The term is in the equation because in some situations effects of actions depend on material properties e.g. earth pressures.

For simplicity, the separate partial factors for uncertainty in modelling action effects (γ_{Sd}) and unfavorable deviation in individual loads (γ_f) are usually combined into single partial factor on actions ($\gamma_F = \gamma_f \times \gamma_{Sd}$) or into a single partial factor on effects of actions ($\gamma_E = \gamma_{Sd} \times \gamma_f$).

8.1.2 Representative and characteristic values

In geotechnical engineering the representative value of ground property is determined either by (prEN 1990:2021, European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al., 2024):

1. using a nominal value based on engineering judgement as a cautious estimate of the value affecting the occurrence of the limit state;
2. determining a characteristic value using statistical methods such that the calculated probability of a worse value is not greater than 5%.

In both cases, the degree of confidence should be consistent in a reliability sense.

In the semi-probabilistic approach, uncertainties related to variables (soil properties, loads) are covered by the specification of representative values (characteristic values) of these variables and partial safety factors that are applied to those variables. In other words, the design value of soil property is obtained by dividing the representative value by the corresponding partial factor. Thus, the representative value and partial factor both are elements in achieving the target reliability.

Guidance for choosing appropriate project-specific representative values of ground properties is presented in European Commission: Joint Research Centre, Orr, T., Sorgatz, J., Estaire, J., Prästings, A. et al. (to be published). However, for code calibration purposes the representative value is often the characteristic value of a typical or sensible probability distribution for the calibration case. The characteristic value is chosen as a certain quantile value from its statistical distribution, for example:

1. 5 % -fractile for resistance variables,
2. 50 %- fractile (mean) for permanent actions, or
3. 98 %-fractile for time-variable actions (distribution for 1-year extreme value).

8.2 Calibration procedures

There are two common calibration procedures to derive partial factors for semi-probabilistic design:

1. the reliability optimization method (e.g., ISO 2394, Sorensen 2010), which essentially minimizes the scatter in reliability achieved, see 8.2.1; or
2. the design value method (e.g., EN 1990 Annex C; European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al., 2024), see 8.2.2.

A third possible code optimization method is the quantile-value method, introduced by Phoon and Ching (2011, 2013, 2014). This method is discussed briefly in 8.2.3.

All methods use full-probabilistic reliability analyses in the procedure in some form, for which the guidance in this report can be used, particularly chapter 5 on uncertainty modeling and chapter 0 on reliability analysis.

The recommended partial factors presented in current Eurocodes typically include model uncertainties “behind” the variables (see 8.1.2). Thus, also in the partial factor calibration model uncertainties should be included, ideally by defining dedicated partial factors.

8.2.1 Minimizing the scatter in reliability achieved

This method is an optimization method (e.g., ISO 2394, Ditlevsen and Madsen 1996) where the aim is minimizing the difference between the target reliability and the calculated reliability. The calculated reliability is obtained by analyzing designs made with a certain sets of partial factors, and then comparing the performance of the tested sets of partial factors. The most common optimization function is:

$$\min \sum w_i (\beta_i(\gamma) - \beta_T)^2$$

Equation 81.

The summation is over the selected portfolio of structures and design scenarios. Parameter w_i is a weight factor, which, for example, can take economic considerations or type of design scenarios into account (or the relative frequency of the scenarios encountered in practice).

The calibration procedure consists of following steps (see ISO 2394):

1. Select a set of comparable reference structures, design scenarios, materials and relevant failure modes. For example, a bearing resistance of spread foundations on sand with varying loading conditions.
2. Select the objective of the calibration. The code objective may be to obtain uniform target probability of failure or reliability index (see chapter 4).
3. Select and specify a set of reliability elements (e.g. partial factors) to calibrate. For geotechnical code calibration within the Eurocode context, usually only partial factors for material or resistance are calibrated whereas load factors and combination factors are unaltered.
4. Design the chosen structures according to the selected set of reliability elements. Designs should meet the semi-probabilistic assessment requirements (i.e. $E_d = R_d$).
5. Calculate the reliability indices β_i for the designed structures.
6. Calculate the difference between target reliability index and calculated reliability indices.
7. Repeat steps 3-5 to find the minimum for **Equation 81**.

The procedure is illustrated as a flowchart in **Figure 62**.

Notice that the designs in step 4 need to be based on the same (real or fictitious representative) data as the reliability analysis in step 5, in order to make the comparison meaningful. That implies that the characteristic values used in the semi-probabilistic design need to be the relevant quantiles of the probability distributions in the full probabilistic reliability analysis.

The optimization problem usually has constraints, such as:

1. Minimum beta value, β_{\min} :
The calculated reliability indices for different design scenarios in calibration domain should be above certain threshold value.
2. Maximum deviation COV_{β} :

Narrower range in calculated reliability indices indicates also better accuracy. The accuracy may be improved by partition of structure types or uncertainty ranges (i.e. differentiation of sets of partial factors).

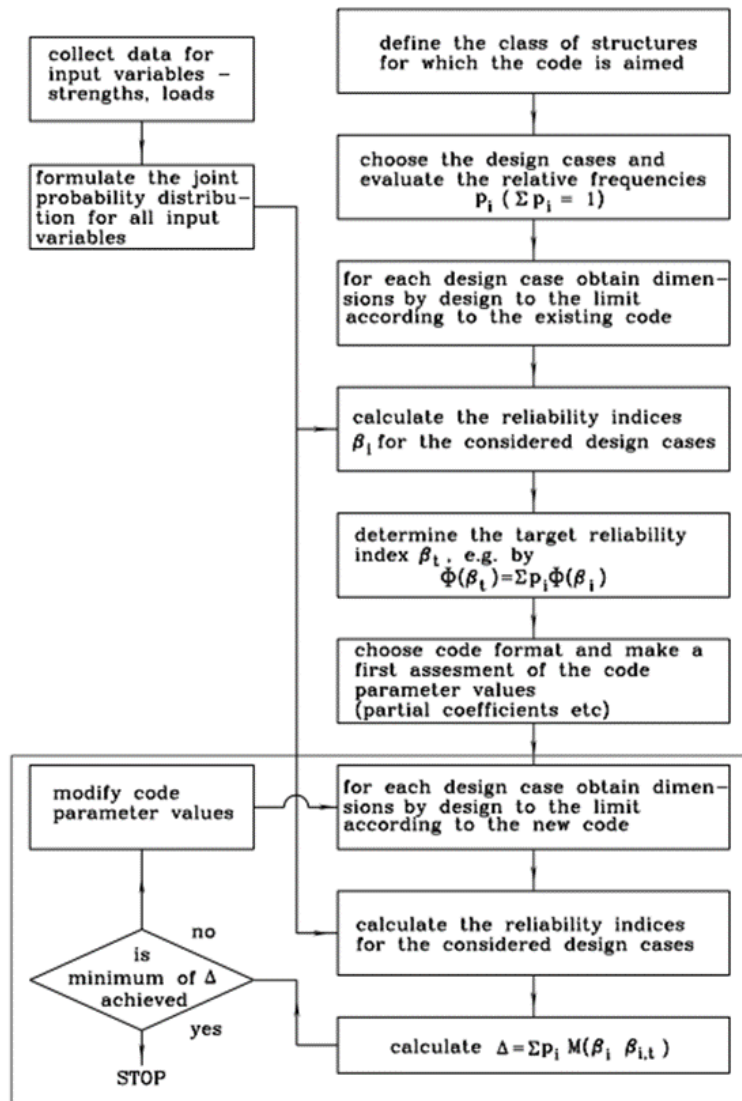
3. Minimum or maximum partial factors for some parameters:

For example, partial factors for resistance / strength parameters are set to be at least equal to 1.

4. Material independent load factors:

The load factors are not usually altered in geotechnical code calibration (in the context of the Eurocodes). Only material or resistance factors are calibrated, conditional on the load (effect) factors which hold for all materials (Eurocodes).

Figure 62. Workflow of code optimization by using “Minimizing the scatter in reliability achieved” method



Source: Ditlevsen and Madsen (1996)

8.2.2 Design value method

The design value method (as also discussed in EN 1990 Annex C) is based on first-order approximation methods like FORM (see O), which produce a design point and influence coefficients besides the reliability index. The fundamental idea is that the values of the random variables in the design point are the most suitable to be used as design (or assessment) values in a semi-probabilistic verification. The design value for one independent, normal distributed parameter in relation to its mean value is given by:

$$x_d = \mu(1 \pm \alpha\beta_T V)$$

Equation 82.

Hence, the design value x_i^* of variable X_i depends on:

- the type and parameters of the probability distribution of X_i ;
- the target reliability index β_T for the limit state and design situation considered;
- the influence coefficient α_i describing the sensitivity to variations in X_i with regard to attaining the limit state considering the uncertainty in X_i .

Partial factors for semi-probabilistic assessment can be calculated as the ratio between mean and design value of the parameter. If a characteristic (or representative) value is used, the partial factor can be determined as the ratio between characteristic and design value.

For example, the partial resistance factor for a normal distributed resistance variable with 5%-quantile characteristic values is given by:

$$\gamma_R = \frac{R_k}{R_d} = \frac{\mu_R - 1.645 \cdot \sigma_R}{\mu_R - \alpha_R \beta_T \sigma_R} = \frac{1 - 1.645 \cdot V_R}{1 - \alpha_R \beta_T V_R}$$

Equation 83.

Recommended values for α_i in the absence of suitable analysis results are presented in EN 1990 and were discussed in section 4.3 (specifically **Table 4**), although estimating the influence coefficients from reliability analyses of the structures and limit states under consideration is clearly preferable. **Table 29** below illustrates how the partial factor varies as a function of the influencing variables. Notice that partial factors below 1.0 would probably be rounded to 1.0 in practice.

Table 29. Partial resistance factor γ_R for a normal distributed resistance variable as a function of the uncertainty (V_R) and the relative influence (α_R)

Target reliability $\beta_T = 3.8$	Influence coefficient α_R					
Coefficient of variation V_R	0.4	0.5	0.6	0.7	0.8	0.9
0.05	0.99	1.01	1.04	1.06	1.08	1.11
0.10	0.99	1.03	1.08	1.14	1.20	1.27
0.15	0.98	1.05	1.14	1.25	1.38	1.55
0.20	0.96	1.08	1.23	1.43	1.71	2.12
0.25	0.95	1.12	1.37	1.76	2.45	4.06

Source: Authors' own work

More generally, the design value is related to the FORM influence coefficient α and the target reliability index β_T by:

$$X_d = F_X^{-1}(\Phi(-\alpha_X\beta))$$

Equation 84.

Likewise, the general expression for a characteristic value with p -quantile is:

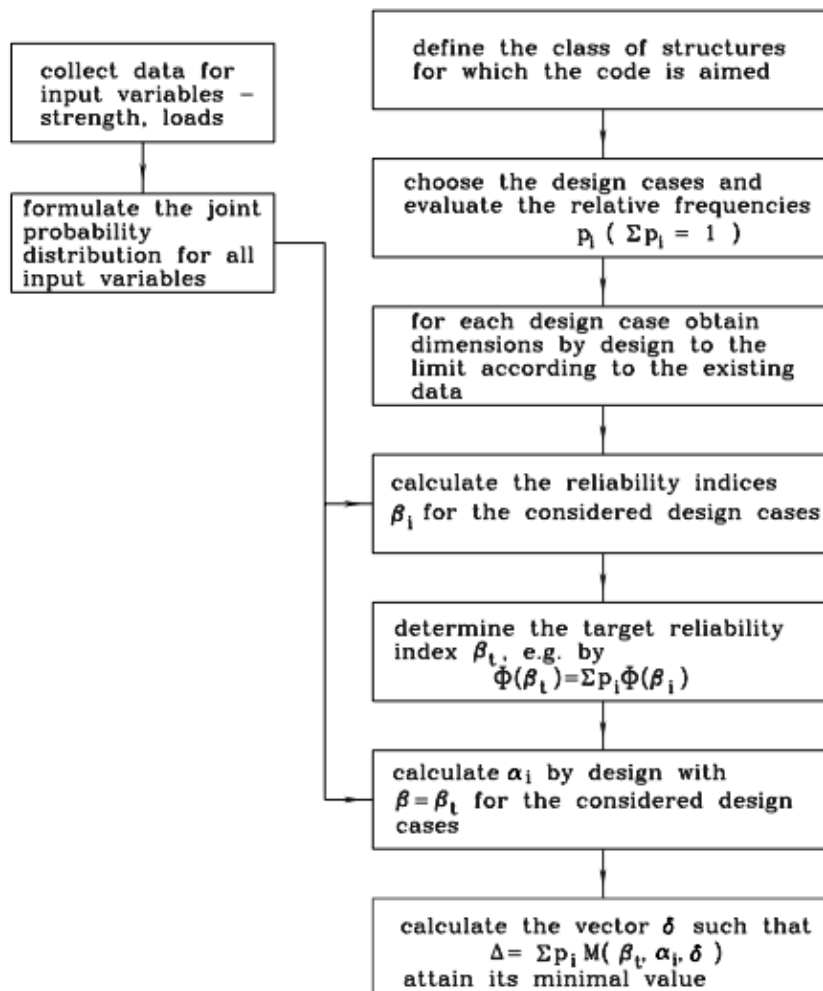
$$X_k = F_X^{-1}(p)$$

Equation 85.

Both expressions can be used to derived partial resistance factors as $\gamma_R = R_k / R_d$ and partial load factors as $\gamma_F = F_d / F_k$

A workflow of design value method is shown in **Figure 63**.

Figure 63. Workflow of code optimization by using “Design-value method”



Source: Ditlevsen and Madsen (1996)

8.2.3 Quantile-based method

The quantile-based method was introduced by Ching and Phoon (2011). The idea of the quantile-based method is to reduce the resistance to its η quantile (small value) and increase the load to its $1 - \eta$ quantile. The parameter η is called probability threshold and the same value of η is used for both random variables.

The objective is to obtain a quantile value η that fulfills the reliability requirement in different design scenarios, or at least minimizes the scatter in reliability indices within calibration domain. Example calculations and comparison to other code calibration methods are provided in Ching and Phoon (2011, 2013 and 2014). The obtained quantile value can also be used to calculate partial factors.

8.3 Project-specific partial factors

It may be beneficial to calibrate partial factors for certain project (or site) specifically. For example, the uncertainty of material resistance can be so high that fixed partial factors do not cover it appropriately. The designer can then derive project-specific partial factors based on data collected from site and project-specific considerations.

For project-specific partial factors, most commonly the design-value method (8.2.2) is used, since alternative methods typically require more effort (e.g. a larger set of reference structures to be analyzed). The design value method does not necessarily require (full) reliability analysis, but the required sensitivity factors and target reliability indices can be taken from Eurocode recommended values, even though dedicated reliability analyses will improve the quality of the design value estimates (i.e. the influence coefficients α).

$$Y_d = F_y^{-1} \left(\Phi(-\alpha_y \beta) \right)$$

Equation 86.

As indicated in EN 1990 Annex C the following values can be used as an approximation (see also **Table 4**):

- $\alpha = 0.8$ for a dominant resistance variable
- $\alpha = -0.7$ for a dominant load variable
- $\alpha = 0.32$ for non-dominant strength variables (e.g., if the resistance model consists of several resistance variables)
- $\alpha = -0.28$ for accompanying loads (non-dominant; e.g., if the load model consists of several variables)
- $\alpha = 1.0$ the variable is dominating the whole reliability problem ($\alpha = 1.0$ for resistance and $\alpha = -1.0$ for loads)

The above values should be used only if the range of the ratio of load over resistance uncertainty is within $0.16 \leq \sigma_E / \sigma_R \leq 7.6$. Notice that resistance here does not have any other meaning than that an increasing value has a positive effect on the reliability and for the load a negative effect.

8.4 Calibration examples

While providing in-depth examples of partial factor calibration is outside the scope of this document, the list below contains references which should be useful for familiarizing with the topic:

- The ‘Reliability Backgrounds of the Eurocodes’ (European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al., 2024) contains:
 - more background information,
 - a calibration study of reliability indices achieved for various materials and limit states, and
 - a simple (didactive) code calibration example in the annex.
- Faber, M.H. & Sørensen, J.D. “Reliability Based Code Calibration – The JCSS Approach”. Proceedings of the 9th International Conference on Applications of Statistics and Probability. San Francisco. 2003.
- Michael R. Lodahl, Kristian T. Brodbeck, Carsten S. Sorensen (2014). Calibrating partial factors for Danish railway embankments using probabilistic analyses. *Journal of Rock Mechanics and Geotechnical Engineering*, Volume 6 (2), pages 150-155. DOI: 10.1016/j.jrmge.2014.01.008.
- Rijkswaterstaat (2017). WBI code calibration - Reliability-based code calibration and semi-probabilistic assessment rules for the WBI 2017. For the assessment of dikes in the Netherlands. Dutch Ministry of Infrastructure and Environment, August, 2017.
- Hehenkamp, M. (2022). Reliability Analysis of Foundation Pile Designs from Eurocode 7. MSc thesis, Delft University of Technology, June 2022.

Furthermore, querying the literature with the following keywords should point to more sources: code calibration, reliability calibration, partial factor derivation, semi-probabilistic verification.

9 Closing remarks

The present document aims to provide guidance to practitioners and code-writers for performing reliability-based assessment of limit states for geotechnical structures, including reliability-based derivation of partial factors. Reliability analysis for geotechnical or other civil engineering structures is not new; developments have been ongoing for several decades in the academic and code-writing communities. Practical application, however, has been limited so far to some specific applications and geographical regions. The reasons for reliability-based design and assessment not being applied more widely until now are numerous, but certainly one of them has been the lack of guidance and authoritative references. Producing this document should contribute to closing this gap.

As discussed in the introduction, reliability assessment has advantages over the partial factor method in a variety of applications. Prospective users of this guideline are encouraged to explore and exploit these benefits. At the same time, the new 'freedom' comes with the responsibility to apply reliability concepts in a sound and thorough manner, implying that users and/or reviewers should be well-trained and experienced.

The method and techniques of reliability assessment are under constant development, which will make updates of this guideline highly desirable. Noteworthy potential improvements are:

- more insight into model uncertainties;
- development of homogenization approaches (i.e. equivalent homogeneous properties based on spatially variable fields); or
- better accessibility of the random finite element method (RFEM).

There are also topics that have not been included in this first version of this guideline simply due to time restrictions and priorities, which however are very worthwhile addressing in future versions, such as:

- design (assisted) by testing;
- geophysical explorations (i.e. how to combine geophysics with other site investigation data); or
- data-driven site characterization.

References

- Ayyub, B. M. (2001a), *Elicitation of Expert Opinions for Uncertainty and Risks*, Taylor & Francis, Boca Raton
- Ayyub, B. M. (2001), *A Practical Guide on Conducting Expert-Opinion Elicitation of Probabilities and Consequences for Corps Facilities*. U.S. Army Corps of Engineers Institute for Water Resources. IWR Report 01-R-01.
- Ahmadi, S. H. and Sedghamiz, A. (2007), 'Geostatistical Analysis of Spatial and Temporal Variations of Groundwater Level', *Environmental Monitoring and Assessment*, No 129, pp. 277-294.
- Ang, A. H.-S., and Tang, W. H. (2006), *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*, John Wiley & Sons, New York.
- Arnold, P. (2016), *Probabilistic modelling of unsaturated slope stability accounting for heterogeneity*, Doctoral thesis, University of Manchester.
- ASCE (2010), *Minimum Design loads for buildings and other structures*. Reston, USA: American Society of Civil Engineers. ISBN 978-0-7844-1085-1.
- Au, S.-K., Wang, Y. (2014), *Engineering risk assessment with subset simulation*, John Wiley & Sons, New York.
- Baecher, G. B. (1972), *Site exploration: A probabilistic approach*, Doctoral thesis, Massachusetts Institute of Technology.
- Baecher, G. B. (2019), 'Putting Numbers on Geotechnical Judgment', Companion whitepaper to the 27th Buchanan Lecture, presented at Texas A&M University, College Station, October 18, 2019.
- Baecher, G. B. (2021), *59th Terzaghi Lecture. Geotechnical systems, uncertainty and risk*
- Baecher, G. B. (1983), 'Simplified Geotechnical Data Analysis', in: *Reliability Theory and Its Application in Structural and Soil Mechanics*, edited by P Thoft-Christensen, Springer, Dordrecht
- Baecher, G. B. and Christian J. T. (2003), *Reliability and Statistics in Geotechnical Engineering*, John Wiley & Sons, New York.
- Baker, J. and Calle, E. (2006), *JCSS probabilistic model code, Section 3.7: Soil properties*, Joint Committee on Structural Safety, Zürich
- Benjamin, J. R. and Cornell, A. (1970), *Probability, Statistics, and Decision for Civil Engineers*, McGraw-Hill, Inc., New York
- Benkoski, S. J., Monticino, M. G., and Weisinger, J. R. (1991), 'A survey of the search theory literature', *Naval Research Logistics*, No 38, 469-494.
- Bloomfield, J. P. and Marchant, B. P. (2013), 'Analysis of groundwater drought building on the standardized precipitation index approach', *Hydrology and Earth System Sciences*, No 17, pp. 4769-4787
- Bojke, L., Soares, M., Claxton, K., Colson, A., Fox, A., Jackson, C., ... & Taylor, A. (2021), 'Developing a reference protocol for structured expert elicitation in health-care decision-making: a mixed-methods study', *Health Technology Assessment*, Vol. 25, No 37, pp. 1-124

- Cami, B., Javankhoshdel, S., Phoon, K. and Ching, J. (2020), 'Scale of Fluctuation for Spatially Varying Soils: Estimation Methods and Values', *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, Vol. 6, No 4
- Cao, Z., Wang, Y., and Li, D. (2016), 'Quantification of Prior Knowledge in Geotechnical Site Characterization', *Engineering Geology*, Vol. 203, pp. 107-116
- Chang, C. and Ko, J. W. (2017), 'New Approach to Estimating the Standard Deviations of Lognormal Cost Variables in the Monte Carlo Analysis of Construction Risks', *Journal of Construction Engineering and Management*, Vol. 143, No 1, Article 06016006,
- Cherubini, C., (1997), 'Data and considerations on the variability of geotechnical properties of soils', *Advances in Safety and reliability*, in: *Proceedings of ESREL 97*, edited by C. Guedes Soares, Vol. 2, pp. 1583-1591
- Christian, J. T., Ladd, C. C., and Baecher, G. B. (1994), 'Reliability Applied to Slope Stability Analysis', *Journal of Geotechnical Engineering*, Vol. 120, No 12, pp. 2180-2207.
- Ching, J. and Phoon, K.K. (2014), 'Transformations and correlations among some parameters of clays the global database', *Canadian Geotechnical Journal*, Vol. 51, No 6, pp. 663-685.
- Ching, J., Wu, T.J., and Phoon, K.K. (2016), 'Spatial correlation for transformation uncertainty and its applications', *Georisk*, Vol. 10, No 4, pp. 294-311.
- Ching, J., Lin, G. H., Chen, J. R., & Phoon, K. K. (2017), 'Transformation models for effective friction angle and relative density calibrated based on generic database of coarse-grained soils', *Canadian Geotechnical Journal*, Vol. 54, No 4, pp. 481-501.
- Ching, J., Phoon, K.-K., Chen, K. F., Orr, T. L. L. and Schneider H. R. (2020), 'Statistical Determination of Multivariate Characteristic Values for Eurocode 7', *Structural Safety*, Vol. 82, No 1, 101893.
- Cheng, Y., and He, D. (2020), 'Slope reliability analysis considering variability of shear strength parameters', *Geotechnical and Geological Engineering*, No. 38, pp. 4361-4368.
- Chilès, J.-P. and Delfiner, P. (2012), *Geostatistics: Modeling Spatial Uncertainty. 2nd edition*, John Wiley & Sons, New York.
- Cooke, R. M. (1991), *Experts in Uncertainty: Opinion and Subjective Probability in Science*, Oxford University Press, New York.
- CUR 166 (2012). Damwandconstructies (deel 1 en 2).
- CUR 211 (2015). Quay Walls. Download link: <https://www.crow.nl/publicaties/quay-walls/>
- Deltares (2021), *Characteristic values of soil properties in Dutch codes of practice - Theoretical backgrounds and assumptions*. Authors: E.O.F. Calle, W. Kanning & T. Schweckendiek. Deltares report no. 11206883-014-GEO-0001.
- Deltares (2024), *Reliability targets for geotechnical structures in the Eurocode framework*. Deltares report no. 11210291-008-GEO-0002.
- Der Kiureghian, A. (2022) *Structural and system reliability*. Cambridge University Press, New York
- Ditlevsen, O. and Madsen, H.O. (1996) *Structural reliability methods*, John Wiley & Sons, New York.
- DNV (1992), *Structural reliability analysis of marine structures*, Classification notes No. 30.6, Det Norske Veritas (DNV).

- DNV (2021), *Statistical representation of soil data. September 2021. Recommended practice DNV-RP-C207. Edition September 2019. Amended September 2021*, Det Norske Veritas (DNV).
- van den Eijnden, A. P., Schweckendiek, T., & Hicks, M. A. (2022), 'Metamodelling for geotechnical reliability analysis with noisy and incomplete models', *Georisk*, Vol. 16, No 3, pp. 518-535.
- ENW – Expertise Netwerk Waterveiligheid (2017). *Fundamentals of Flood Protection*. Download link: <https://www.enwinfo.nl/>
- Evans, J. D. (1996), *Straightforward Statistics for the Behavioral Sciences*, Brooks/Cole Publishing Company, Pacific Grove.
- Fenton, G. A, and Griffiths, D. V. (2008), *Risk Assessment in Geotechnical Engineering*, John Wiley & Sons, New York.
- FERC (2015) *Federal Guidelines for Dam Safety Risk Management*. Download link: <https://www.fema.gov/>.
- Forrest, W. S. and Orr, T. L. L. (2010), 'Reliability of shallow foundations designed to Eurocode 7', *Georisk*, Vol. 4, no 4, pp. 186-207
- Garvey, P. R., Book, S. A., & Covert, R. P. (2016), *Probability methods for cost uncertainty analysis: A systems engineering perspective*. CRC Press, Rotterdam
- Hanna, A.M. and Meyerhof, G.G. (1981), 'Experimental evaluation of bearing capacity of footings subjected to inclined loads', *Canadian Geotechnical Journal*, Vol. 18, No 4, pp. 599-603.
- Hofmann, Proske and Zeck (2021). *Vergleich der Einsturzhäufigkeit und Versagenswahrscheinlichkeit von Stützbauwerken*. DOI: 10.1002/bate.202000084.
- Hohenbichler, M., & Rackwitz, R. (1982), 'First-order concepts in system reliability', *Structural safety*, Vol. 1, No 3, pp. 177-188.
- Hudak, D. G. (1994), 'Adjusting triangular distributions for judgmental bias', *Risk Analysis*, Vol. 14, No 6, pp. 1025-1031.
- ISSMGE-TC205/ISSMGE-TC304 (2017), Joint TC205/ISSMGE-TC304 Working Group on 'Discussion of statistical/reliability methods for Eurocodes' – *Final Report* (September 2017)
- ISSMGE-TC304 (2020), *Probabilistic solutions for survey questions in 'Are we overdesigning? – A survey of international practice'*. ISSMGE Technical Committee ISSMGE-TC304: Risk management in engineering practice.
- ISSMGE-TC304 (2021), *State-of-the-Art Review of Inherent Variability and Uncertainty in Geotechnical Properties and Models*. ISSMGE Technical Committee ISSMGE-TC304: Risk management in engineering practice.
- ISO 13822:2010. *Bases for design of structures — Assessment of existing structures*.
- ISO 2394 (2015), *General Principles on Reliability for Structures*.
- JCSS, Joint Committee on Structural Safety (2001) *JCSS Probabilistic model code*.
- Johnson, D. (2002), 'Triangular approximations for continuous random variables in risk analysis', *Journal of the Operational Research Society*, Vol. 53, No 4, pp. 457-467.

European Commission: Joint Research Centre, Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J. et al. (2024), Reliability background of the Eurocodes – Support to the implementation, harmonization and further development of the Eurocodes, Publications Office of the European Union, <https://data.europa.eu/doi/10.2760/9482837>

European Commission: Joint Research Centre, Garin, H., Baldwin, M., Reiffsteck, P van der Made K-J. et al. (to be published), Assembling the Ground Model and the derived values. Guidelines for the application of the 2nd Generation of Eurocode 7 – Geotechnical design.

European Commission: Joint Research Centre, Orr, T., Sorgatz, J., Estaire, J., Prästings, A. et al. (to be published), Determination of representative values from derived values for verification of limit states with EN 1997. Guidelines for the application of the 2nd Generation of Eurocode 7 – Geotechnical design.

Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (Eds.), (2001), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, New York

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021), *Noise: A flaw in human judgment*, Little, Brown Spark.

Kelley, Ken. (2007), 'Sample Size Planning for the Coefficient of Variation from the Accuracy in Parameter Estimation Approach', *Behavior Research Methods*, Vol. 39, No 4, pp. 755–66.

Kotz, S., & Van Dorp, J. R. (2004), *Beyond beta: other continuous families of distributions with bounded support and applications*, World Scientific Printers (S) Pte Ltd, Singapore.

Lacasse, S., Nadim, F., Rahim, A., and Guttormsen, T. R. (2007), *Statistical Description of Characteristic Soil Properties*, Offshore Technology Conference.

Lehtonen, V. J., Meehan, C. L., Länsivaara, T. T. and Mansikkamäki, J. N. (2015), 'Full-scale embankment failure test under simulated train loading', *Géotechnique*, Vol. 65, No 12, pp. 961-974.

Lloret Cabot, M., Fenton, G. A. and Hicks, M. A. (2014), 'On the estimation of scale of fluctuation in geostatistics', *Georisk*, Vol. 8, No 2, pp. 129–140

Löfman, M. S. and Korkiala-Tanttu, L. (2019), 'Variability and typical value distributions of compressibility properties of fine-grained sediments in Finland', In: Ching, J., Li, D. Q. & Zhang, J. (Eds.), *Proceedings of the 7th International Symposium on Geotechnical Safety and Risk (ISGSR 2019), 11–13 December 2019, Taipei, Taiwan*, Research Publishing (S) Pte. Ltd

Löfman, M. S., & Korkiala-Tanttu, L. K. (2022), Transformation models for the compressibility properties of Finnish clays using a multivariate database. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, vol 16 no 2, 330-346.

Mackay, J. D., Jackson, C. R., & Wang, L. (2014), 'A lumped conceptual model to simulate groundwater level time-series', *Environmental modelling & software*, No 61, pp. 229-245.

Machiwal, D., Mishra, A., Jha, M. K., Sharma, A., and Sisodia, S.S. (2012), 'Modeling Short-Term Spatial and Temporal Variability of Groundwater Level Using Geostatistics and GIS', *Natural Resources Research*, Vol. 21, No 1, pp. 117-136

Marsden, E. (2022), 'Uncertainty in risk engineering: concepts' in: Risk Engineering. E-learning website, <https://risk-engineering.org>

- Melchers, R. E. and Beck, A. T. (2018), *Structural Reliability Analysis and Prediction*, John Wiley & Sons, New York
- Mikkola, P., Martin, O. A., Chandramouli, S., Hartmann, M., Pla, O. A., Thomas, O., ... & Klami, A. (2021), *Prior knowledge elicitation: The past, present, and future*. arXiv preprint arXiv:2112.01380.
- Mishra, S. (2002), *Assigning probability distributions to input parameters of performance assessment models*, Swedish Nuclear Fuel and Waste Management Co. Technical Report TR-02-11
- Moustapha, M., Marelli, S. & Sudret, B. (2022), 'Active learning for structural reliability: Survey, general framework and benchmark', *Structural Safety*, No 96, 102174
- Muhammad Rayyan, R. (2021), *Reliability Benchmarking of Eurocode 7 Design Examples: CIE5050-09*, Additional Thesis Project 2021, Delft University of Technology, <http://resolver.tudelft.nl/uuid:0d6af648-9338-41ba-8202-fc8f95a07864>
- Müller, R. (2013), *Probabilistic stability analysis of embankments founded on clay*, Doctoral thesis, KTH Royal Institute of Technology, Stockholm.
- Müller, R., Larsson, S., and Spross, J. (2014), 'Extended multivariate approach for uncertainty reduction in the assessment of undrained shear strength in clays', *Canadian Geotechnical Journal*, Vol. 51, No 3, pp. 231–245.
- Nadim, F. (2015), 'Accounting for Uncertainty and Variability in Geotechnical Characterization of Offshore Sites', *Geotechnical Safety and Risk V*, T. Schweckendiek, A. F. van Tol, D. Pereboom, M. T. van Staveren, and P. M. C. B. M. Cools, eds., IOS Press BV, Amsterdam, 23–34.
- Nederlof, M.H. (2022), Lognormal Distribution. URL: www.mhnederlof.nl/lognormal.html.
- OCDI (2009), *Technical Standards and Commentaries for Port and Harbour Facilities in Japan*. Tokyo, Japan: The Overseas Coastal Area Development Institute.
- O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., ... and Rakow, T. (2006), *Uncertain judgements: eliciting experts' probabilities*, John Wiley & Sons, New York
- Orchant, C. J., Kulhawy F. H., and Trautmann, C. H. (1988), *Critical Evaluation of In-Situ Test Methods and Their Variability*, Report EL-5507(2)
- Pană, P.T. (2022), *Implementation of spatial variability in PLAXIS*, MSc thesis, Delft University of Technology, <http://resolver.tudelft.nl/uuid:81b2faa6-7234-4c81-92a6-aaea09dd6b07>
- Peterson, T. J., & Western, A. W. (2018), 'Statistical interpolation of groundwater hydrographs', *Water Resources Research*, No 54, pp. 4663– 680.
- Phoon, K-K., Shuku, T., Ching, J. (2024), *Uncertainty, Modeling, and Decision Making in Geotechnics*, Routledge, Taylor & Francis
- Phoon, K.-K. (Ed.), (2008), *Reliability-Based Design in Geotechnical Engineering. Computations and Applications*, CRC Press, Rotterdam
- Phoon, K.-K., and Ching, J. (Eds.), (2015), *Risk and reliability in geotechnical engineering*, Taylor & Francis.
- Phoon, K.-K., and Kulhawy, F. H. (1999a), 'Characterization of Geotechnical Variability', *Canadian Geotechnical Journal*, Vol 36, No 4, pp. 612–24.

- Phoon, K.-K., and Kulhawy, F. H. (1999b), 'Evaluation of geotechnical property variability', *Canadian Geotechnical Journal*, Vol 36, No 4, pp. 625-639.
- Phoon, K.-K., and Kulhawy, F. H. (2005), 'Characterisation of model uncertainties for laterally loaded rigid drilled shafts', *Géotechnique*, Vol. 55, No 1, pp. 45-54.
- Phoon, K.-K., and Tang, C. (2019), 'Characterisation of geotechnical model uncertainty', *Georisk*, Vol 13, No 2, pp. 101-130.
- Prästings, A. (2019), *Managing uncertainties in geotechnical parameters: From the perspective of Eurocode 7*. Doctoral thesis, KTH Royal Institute of Technology, Stockholm.
- Prästings, A., Spross, J., Müller, R., Larsson, S., Bjureland, W., and Johansson, F. (2017), 'Implementing the extended multivariate approach in design with partial factors for a retaining wall in clay', *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, Vol. 3, No 4, 04017015.
- Rippi, A. (2015). *Structural reliability analysis of a dike with a sheet pile wall - Coupling Reliability methods with Finite Elements*. MSc thesis TU Delft.
- Roberds, W. J. (1990), 'Methods for developing defensible subjective probability assessments', *Transportation Research Record*, No 1288, pp. 183-190
- Roubos (2019). *Enhancing reliability-based assessments of quay walls*. PhD dissertation, Delft University of Technology.
- Roubos A.A., Allaix D., Fischer K., Jonkman S.N. (2019), 'Target reliability indices for existing quay walls derived on the basis of economic optimization and human safety requirements', *Structure and Infrastructure Engineering*, Vol. 16, No 4, pp. 1-13.
- Rossi, N., Bacic, M., Kovacevic, M.S., Libric, L. (2021), 'Development of Fragility Curves for Piping and Slope Stability of River Levees', *Water*, Vol. 13, No 5, 738.
- Schneider, H. R, and Schneider M. A. (2013), 'Dealing with Uncertainties in EC7 with Emphasis on Determination of Characteristic Soil Properties', In: *Modern Geotechnical Design Codes of Practice*, edited by P. Arnold, T. Schweckendiek, G. A. Fenton, B. Simpson, and M. A. Hicks, IOS Press, pp. 87-101.
- SGI (2011), *Göta älvutredningen - delrapport 28* (in Swedish)
- SGF (2022), *Metodbeskrivning: Åsätta subjektiva sannolikheter i geotekniska projekt* (in Swedish) [*Methodology: Assigning subjective probabilities in geotechnical engineering projects*]. SGF Report 2:2022. Swedish Geotechnical Society, Linköping.
- Schweckendiek, T. and Calle E. O. F. (2010), 'A Factor of Safety for Geotechnical Characterisation', in: *Proceedings of the 17th Southeast Asian Geotechnical Conference* (Taipei, Taiwan, May 10-13 2010),
- Spross, Johan, and Stefan Larsson. 2021. 'Probabilistic Observational Method for Design of Surcharges on Vertical Drains', *Géotechnique*, Vol. 71, No 3, pp. 226-238.
- Steenbergen R.D.J.M., Rozsas A., Vouwenvelder A.C.W.M. (2018), 'Target reliability of new and existing structures - A general framework for code making', *Heron*, Vol. 63, No. 3.
- Stille, H., Andersson, J., Olsson, L. (2003), *Information Based Design in Rock Engineering*, Report 61, SveBeFo, Swedish Rock Engineering Research.

- Tabarroki, M., Ching, J., Phoon K.K., Chen, Y.Z. (2021), 'Mobilisation-based characteristic value of shear strength for ultimate limit states', *Georisk Assessment and Management of Risk for Engineered Systems and Geohazards*, Vol. 16, No 4, pp. 1-22
- Tabarroki, M., Ching, J., Phoon, K.-K., and Chen, Y. Z. (2022), 'Mobilisation-based characteristic value of shear strength for ultimate limit states', *Georisk*, Vol. 16, No 3, pp. 413-434
- Tang, W. H. (1987), 'Updating anomaly statistics--single anomaly case', *Structural Safety*, Vol. 4, pp. 151-163.
- Tang, W. H., and Halaim, I. (1988), 'Updating anomaly statistics--multiple anomaly pieces', *ASCE Journal of Engineering Mechanics*, Vol. 114, No 6
- Tang, W. H., and Quek, S. T. (1986), 'Statistical model of boulder size and fraction', *Journal of Geotechnical Engineering*, Vol. 112, No 1, pp. 79-90.
- Tang, C., and Phoon, K.-K. (2018), 'Statistics of model factors in reliability-based design of axially loaded driven piles in sand', *Canadian Geotechnical Journal*, Vol. 55, No 11, pp. 1592-1610.
- Tang, C., Phoon, K. K., Li, D. Q., & Akbas, S. O. (2020), 'Expanded Database Assessment of Design Methods for Spread Foundations under Axial Compression and Uplift Loading', *Journal of Geotechnical and Geoenvironmental Engineering*, Vol. 146, No 11, 04020119.
- Tang, C., and Phoon, K.-K. (2021), *Model Uncertainties in Foundation Design*, Taylor & Francis
- Tang, W. H., and Saadeghvaziri, M. A. (1983), 'Updating distribution of anomaly size and fraction', in: *Recent Advances in Engineering Mechanics and Their Impact on Civil Engineering Practice*, edited by W. F. Chen, ASCE
- Teixeira, R., Nogal, M., and O'Connor, A. (2021), 'Adaptive Approaches in Metamodel-based Reliability Analysis: A Review', *Structural Safety*, Vol. 89, 102019.
- TNO (2021). *Bewezen sterkte damwanden en kademuren - Stap 1*. Liesette la Gasse/Diego Allaix.
- USACE (1999), *Risk-Based Analysis in Geotechnical Engineering for Support of Planning Studies*. ETL1110-2-556. Washington DC, United States.
- Uzielli, M., Lacasse, S., Nadim, F. and Phoon K-K. (2006), 'Soil Variability Analysis for Geotechnical Practice', in: *Proceedings of the 2nd International Workshop on Characterisation and Engineering Properties of Natural Soils, Singapore*
- Van der Krogt, M. G., Schweckendiek, T., and Kok, M. (2019), 'Uncertainty in spatial average undrained shear strength with a site-specific transformation model', *Georisk*, Vol. 13, No 3, pp. 226-236.
- Vanmarcke, E. H. (1977), 'Probabilistic modeling of soil profiles', *ASCE Journal of the Geotechnical Engineering Division*, Vol. 103, GT11, pp. 1227-1246.
- Vanmarcke, E. H. (2010), *Random fields: analysis and synthesis*, World Scientific, Singapore.
- Vanmarcke, E. H. (2011), 'Risk of Limit-Equilibrium Failure of Long Earth Slopes: How it depends on Length', in: *Proceedings of GeoRisk 2011*, ASCE.
- Varkey, D., Hicks, M. A., van den Eijnden, A. P., & Vardon, P. J. (2020), 'On characteristic values for calculating factors of safety for dyke stability', *Géotechnique Letters*, Vol. 10, No 2, pp. 353-359.

Vereecken E., Botte W., Droogne D., Caspee R. (2020), 'Reliability-based calibration of partial factors for the design of temporary scaffold structures', *Structure and Infrastructure Engineering*, Vol. 16, No 4, pp. 642-658.

Vick, S. G. (2002), *Degrees of Belief. Subjective Probability and Engineering Judgement*, ASCE Press

Vrouwenvelder, A.C.W.M. (2002), 'Developments towards full probabilistic design codes', *Structural Safety*, Vol. 24, pp. 417-432.

Wackernagel, H. (2003), *Multivariate geostatistics: an introduction with applications*, Springer Science & Business Media.

Zellner, M., Abbas, A. E., Budescu, D. V., & Galstyan, A. (2021), 'A survey of human judgement and quantitative forecasting methods', *Royal Society Open Science*, Vol. 8, No 2, 201187.

List of abbreviations and definitions

The symbols used in this guideline correspond to those used in the Eurocodes EN 1990 and EN 1997. As EN 1990 describes reliability-based design in more detail than EN 1997, this guideline favors EN 1990 symbols. Parallel symbols from EN 1997 are indicated in parentheses.

Abbreviations	Definitions
LSF	Limit State Function
RBD	Reliability-Based Design
SLS	Serviceability Limit State
ULS	Ultimate Limit State
b	Bias, sample mean of transformation error
n	Number of observations, i.e. sample size
L	Averaging length
M	Model factor
P_f	Failure probability
s_x	Sample standard deviation of variable X
s_x^2	Sample variance of variable X
\hat{t}	Trend value predicted by a regression line
z	Depth from ground surface
m_x	Arithmetic sample mean of variable X
V	Coefficient of variation
V_x	Coefficient of variation of variable X (denoted Δ_x in EN 1997)
$V_{x,inh}$	Coefficient of variation for inherent variability of variable X
$V_{x,obs}$	Coefficient of variation for observed variability calculated from site data of variable X
$V_{x,meas}$	Coefficient of variation for measurement error of variable X

Abbreviations	Definitions
$V_{x,stat}$	Coefficient of variation for statistical uncertainty of variable X
$V_{x,tot}$	Coefficient of variation for total uncertainty of variable X
$V_{x,trans}$	Coefficient of variation for transformation uncertainty of variable X
X	Basic variable, for example a material property
Y	Basic or derived general variable
α_x, α_y	Sensitivity factor indicating the importance of X or Y in the reliability estimation
β	Reliability index
β_t	Target reliability index
β_n	Reliability index for a reference period of n years
γ	Partial factor
Γ^2	Variance reduction factor
θ_v	Scale of fluctuation in vertical direction
θ_h	Scale of fluctuation in horizontal direction
δ	Sample coefficient of variation of transformation error
μ_x	Mean of variable X
ρ	Correlation coefficient
σ_x	Standard deviation of variable X
σ_x^2	Variance of variable X
Φ	Cumulative distribution function of the standardised Normal distribution
ζ	Degree of anisotropy

List of boxes

Box 1. Background: Target failure probability according to ISO 2394:2015	28
Box 2. Example 4.1: Pile resistance.....	32
Box 3. Reference period.....	33
Box 4. Example 4.2: Risk-based reliability target for road settlement (SLS).....	35
Box 5. Example 5.1: Kriging of layer boundaries.....	42
Box 6. Example 5.2: Stratification scenarios	44
Box 7. Example 5.3: Fitting extreme value distribution to groundwater level timeseries	45
Box 8. Example 5.4: Assessment of observed variability from field vane data.....	52
Box 9. Background: Practical assumption of full spatial (depth-)averaging.....	57
Box 10. Measurement error in estimating total uncertainty	61
Box 11. Background: Total uncertainty in ground properties	66
Box 12. Example 5.5: Total uncertainty in undrained shear strength from field vane tests	68
Box 13. Example 5.6: Triangular and lognormal distributions based on estimated constraints.....	74
Box 14. Example 5.7: Defining distribution parameters for linear regression	76
Box 15. Example 5.8: Approximate probability distributions for actions based on design values	81
Box 16. Example 6.1: Total probability in stratification scenarios	85
Box 17. Example 6.2: Standard Normal (Nataf) transformation.....	86
Box 18. Example 6.3: Reliability index for a simple limit state (analytical solution).....	90
Box 19. Example 6.4: Limit states for multiple failure modes.....	100
Box 20. Background: Required accuracy and computation times with Crude Monte Carlo	103
Box 21. Example 6.5: Convergence of sampling-based methods	105
Box 22. Example 6.6: Point estimate of train load for railway embankment slope reliability.....	107
Box 23. Example 6.7: Design value of effects of actions for sheet pile wall.....	108
Box 24. Example 6.8: Fragility curves.....	109
Box 25. Example 6.9: Spatial averaging and attraction to weak zones for slope stability.....	116
Box 26. Example 7.1.....	119
Box 27. Example 7.2: Comparison of posteriors retrieved by updating with weakly informative and informative priors	120
Box 28. Example 7.3: Comparison of posteriors retrieved by unbounded and bounded priors.	121

Box 29. Example 7.4: Implementation of a likelihood function.....	122
Box 30. Example 7.5: The importance of posterior and posterior predictive in geotechnical engineering.....	126
Box 31. Example 7.6: Bayesian updating of the friction angle with site investigation data using the conjugate priors method.....	128
Box 32. Example 7.7: Bayesian updating of the friction angle with site investigation data using the numerical integration method.....	129
Box 33. Example 7.8: Bayesian updating of the friction angle with site investigation data using a sampling method.....	130
Box 34. Example 7.9: Direct updating of pile reliability in axial loading.....	132
Box 35. Example 7.10: Direct updating of pile reliability in axial loading (continued).....	133
Box 36. Example 7.11: Application of the Observational Method for the vertical deformation of a rock pillar.....	134
Box 37. Overestimation of model error.....	173
Box 38. Notes to the examples.....	176

List of figures

Figure 1. Visual outline of the guideline.....	16
Figure 2. Relationship between acceptable risk, reliability targets and partial factors (i.e. the semi-probabilistic approach).....	17
Figure 3. Bathtub curve (black line) showing the conditional failure probability as a function of time, with zones of (I) increasing reliability (dominant time independent uncertainties like resistance and self-weight), (II) constant reliability (dominating time dependent uncertainties, usually loads) and (III) decreasing reliability (dominating deterioration). The blue line shows the unconditional failure probabilities.....	22
Figure 4. Overview of the workflow for the reliability verification procedure (with indications of the chapter containing in-depth treatment).....	25
Figure 5. Reliability index as a function of the considered reference period.....	34
Figure 6. How the uncertainties in each chapter relate to GM and GDM.	38
Figure 7. Categories of uncertainty in estimated ground property.....	39
Figure 8. Calibrated semi-variogram $\gamma\Delta x = \sigma_{tot}^2 - C(\Delta x)$ against the data of table above.	43
Figure 9. Predicted bedrock elevation using ordinary kriging with the above-specified data and covariance model in plain view (left) and on cross-section A-A' (right).	43
Figure 10. Illustration of stratification scenarios for a dike on a soft soil deposit with potential sand lenses.....	44
Figure 11. Time series of measured groundwater levels.....	46
Figure 12. Time series and the fitted values (left) and the Weibull distribution (right).....	46
Figure 13. Steps in uncertainty characterization depending on the number of site-specific measurements.....	48
Figure 14. Inherent variability of soil and trend with depth (Nadim 2015).....	49
Figure 15. Estimates of mean and standard deviation for field vane strength for soft clay.....	53
Figure 16. Horizontal and vertical scales of fluctuations for soil properties (Cami et al. 2020).....	54
Figure 17. “Rule of thumb” method to estimate scale of fluctuation δ (Cami et al. 2020).	54
Figure 18. Illustration of auto-correlation functions.....	55
Figure 19. Examples of averaging lengths: (a) 3D slope failure, (b) Strip footing subjected to vertical loading, (c) friction pile under compression, and (d) basal heave for excavation in clay	57
Figure 20. Relationship between variance reduction factor Γ^2 and ratio L/δ	58
Figure 21. Probability distributions of point and mobilized shear strength.....	59
Figure 22. Spatial averaging for a foundation pile.....	59

Figure 23. Examples of contribution from statistical uncertainty in the assessment of inherent variability of undrained shear strength.....	63
Figure 24. Transformation uncertainty resulting from pairwise correlation between measured (target) value and the derived value of a ground property (“design property” in the figure; not to be confused with the design value)	64
Figure 25. Total uncertainty in undrained shear strength represented by lognormal distribution...	68
Figure 26. Histograms simulated using triangular and lognormal pdfs.....	74
Figure 27. Delabole slate strength data (from Bozorgzadeh and Harrison, 2014).....	77
Figure 28. Simple linear regression model fitted to the subset of data with $\beta = 90^\circ$	78
Figure 29. Gumbel distributions with different reference periods and the design value.....	82
Figure 30. Example of two stratification scenarios.....	85
Figure 31. Probability event tree to account for scenarios.....	85
Figure 32. Step-wise transformation from U-space to parameter space by correlation (U to V) and iso-probabilistic transformation (V to X). Red dashed lines indicate the (mapping of) the contour lines of equal distance to the mean in standard normal space (i.e. β_{HL} -contours).....	86
Figure 33. Joint distribution $f_{\mathbf{x}}(\mathbf{x})$, limit state surface $g(\mathbf{x}) = 0$ and failure domain $\Omega(\mathbf{x})$	87
Figure 34. Graphical representation of the reliability index for independent ($\rho_{RS} = 0$) and normal distributed load S and resistance R.....	90
Figure 35. Transformation from physical space (left) to standard normal space (right) for a joint probability distribution with correlated parameters and non-linear limit state function. The definitions of design point $\mathbf{x}^* = T(\mathbf{u}^*)$, Hasofer-Lind reliability index β_{HL} , and ingredients for the importance factors u_i^* are given in standard normal space. The design point can be represented in parameter space as point \mathbf{x}^*	93
Figure 36. Illustration of importance sampling (IS) compared to Monte Carlo simulation (MCS) with the parameter space of a multivariate distribution (left) and the corresponding standard normal space (right). Compared to Monte Carlo simulation, importance sampling is very effective in sampling from the failure domain with the same total number of samples.....	96
Figure 37. Schematic example of a subset simulation with the parameter space of a multivariate distribution (left) and the corresponding standard normal space (right). Subsets are generated with subsequent conditional probabilities of 10% (i.e. 10-percentile, 1-percentile, ...) until the limit state condition $g(\mathbf{x}) = 0$ is reached. The probability of failure is the production of the subsequent conditional probabilities between subsets. Note that Subset 0 is a standard Monte Carlo sample set and that Subset 3 itself does not need to be sampled for the evaluation of P_f	96
Figure 38. Schematic example of directional sampling with radial sampling from the origin in standard-normal space.....	97
Figure 39. Steps in reliability analysis (section outline).....	98

Figure 40. Geometry of slope in undrained condition for the definition of the ultimate limit state for slope stability $gX = FS - 1$. The two failure mechanisms are the deep and shallow sliding circles.....	100
Figure 41. System limit state function for undrained slope stability involving two layers with independent shear strength.....	100
Figure 42. Flow chart for selecting a reliability method based on characteristics of the limit state function (LSF). The dotted arrows indicate optional verification or improvement steps, or an alternative approach as back-up.....	102
Figure 43. Total computation time required for the estimation of the reliability index by MSC with estimation uncertainty $VPf < 0.10$ as a function of the reliability index β and the time to evaluate the limit state function.....	103
Figure 44. Convergence rate for Monte Carlo and Importance-based sampling algorithms.....	106
Figure 45. Comparison between the probability distribution of variable E and associated design value E_d , when accounting for the influence coefficient α_E	109
Figure 46. Schematisation of a flood embankment for the analysis of reliability conditional to different water tables.....	110
Figure 47. Fragility curve conditional to the water level h with water level distribution $f_H(h)$. The area under the weighted curve (shaded in blue) represents the total probability, showing that water levels around $h = 4 \text{ m} + \text{ref.}$ contribute most to the failure probability.....	110
Figure 48. Problem sketch for gravity wall example.....	111
Figure 49. One realization of a slope in RFEM with spatially variable undrained shear strength..	117
Figure 50. Distribution of FS with (RFEM) and without (homogeneous) accounting for spatial variability.....	117
Figure 51. Comparison of posteriors after updating with a) weakly informative and b) informative priors.....	121
Figure 52. Comparison of posteriors after updating with a) untruncated and b) truncated normal prior distributions.....	122
Figure 53. a) Likelihood functions and b) posterior distributions for independent and dependent measurements.....	123
Figure 54. Flowchart for Bayesian inference method selection.....	125
Figure 55. Posterior distribution of μSu and posterior predictive distribution of Su	126
Figure 56. Schematic logic of probabilistic site characterization (Source: Baecher & Christian 2003).....	127
Figure 57. Prior and posterior distribution and likelihood function over the calculation mesh.....	130
Figure 58. Histogram of posterior sample.....	131

Figure 59. Histogram of pile load and capacity of Monte Carlo samples for case 1 (red line is for the proof load value).	133
Figure 60. Histogram of pile load, capacity and test load of Monte Carlo samples for case 2.....	133
Figure 61. a) Section of the analysed rib pillar. b) For the analysed rib pillar, the observational method is found more favorable, because the expected cost of that design alternative (95 Monetary Units) is less than that of the conventional design (150 MU).....	135
Figure 62. Workflow of code optimization by using “Minimizing the scatter in reliability achieved” method.....	140
Figure 63. Workflow of code optimization by using “Design-value method”.....	142
Figure 64. Preloading an embankment without using a surcharge. The procedure aims at meeting the target height, h_{emb} , within the allotted preloading time, where the excess material at the beginning is sufficient to compensate for the occurred settlement, s_x	177
Figure 65. Ground model and its use in the probabilistic ground characterization and the analytical settlement model used to determine S_∞	179
Figure 66. Top row: data points of preconsolidation pressure, modulus number, and modulus M_L . Middle row: evaluation of linear trend of log-transformed data points (Equation 92). Bottom row: exponential trend lines and their uncertainty after transformation back to the physical space.	181
Figure 67. Top row: data points of undrained shear strength and unit weight of clay. Middle row: evaluation of linear trend of log-transformed data points (Equation 92). Bottom row: exponential trend lines and their uncertainty after transformation back to the physical space.	181
Figure 68. Generation of $\sigma'c$ and $\sigma'L$ satisfying the restriction $\sigma'c < \sigma'L$	182
Figure 69. Visualization of calculation procedure to find $t_{preload}$ for each generated settlement trajectory.....	183
Figure 70. Monte Carlo simulation of settlement trajectories, histogram of their S_∞ , and required preloading time T_G to satisfy the allowable residual settlement.	184
Figure 71. Probability of violating the serviceability limit state (Equation 90) for different chosen $t_{preload}$	184
Figure 72. Geometry and loading of strip footing.	186
Figure 73. Ground model - for friction angle (a) and shear modulus (b), taken from TC 304 (2020).	187
Figure 74. 2D mesh (with a vertical axis of symmetry on the left) and settlements for one of the ZSOIL realizations.	188
Figure 75. Prior and Posterior distributions of vertical settlement.	189
Figure 76. Plot of cone resistance per CPT over depth and soil stratigraphy.....	191
Figure 77. Estimated reliability index per probabilistic model for different sizes of the CPT sample and comparison to the EN 1990 requirement.	195

Figure 78. Example set-up: (a) soil profile, (b) CPT-tip resistance profile, (c) bearing-capacity design.	197
Figure 79. Quantile ranges of the density function for pile resistances from empirical values according to EA-Pfähle (2013): (a) based on the EA-Pfähle (2013), (b) representation of the variance of the pile skin friction $q_{s,rep}$ with q_c variance of the layers S1 and S2.	198
Figure 80. (a) A priori density and distribution functions. (b) Sensitivity of the power function G vs. the variation of skin friction q_s or tip pressure q_b	201
Figure 81. (a) Density functions of the variables, (b) density and distribution functions for performance function G	204
Figure 82. Excavation and retaining system layout (left) and characteristic section (right).	208
Figure 83. Geological section on the eastern side of Building 1 (taken from the Geotechnical Report).	211
Figure 84. Probability distributions of the main geotechnical parameters.	213
Figure 85. Probabilistic results for limit state function ZM ($MRd = 1.25 \cdot Mk$) using Erraga metamodel (screenshot Probabilistic Toolkit).	215
Figure 86. Probabilistic results for limit state function ZM ($MRd = 1.25 \cdot Mk$) using Importance Sampling around design point from Erraga metamodel (screenshot Probabilistic Toolkit).	216
Figure 87. Probabilistic results for limit state function ZS ($SRd = 1.25 \cdot Sk$) using Erraga metamodel (screenshot Probabilistic Toolkit).	217
Figure 88. Probabilistic results for limit state function ZS ($SRd = 1.25 \cdot Sk$) using Importance Sampling around design point from Erraga metamodel (screenshot Probabilistic Toolkit).	217
Figure 89. Reliability results for the structural ULS verification.	218
Figure 90. Principle of the coupling between PTK and PLAXIS (Deltares, 2019b).	219
Figure 91. Investigated system of soldier pile wall.	221
Figure 92. Ground model used for reliability analysis.	222
Figure 93. Vertical displacement field and displacement vectors corresponding to one of the 30 finite element computations with ZSWalls.	224
Figure 94. Sobol indices illustrating the degree of sensitivity of the random input variables on the bending moment of the wall.	225
Figure 95. Results of the reliability analysis of limit state equation $g_1(X)$ – pdf of the maximum bending moment (left); PCE realisations based on FE samples depending on combinations of h_{water} and ϕ'_{FG} (right).	226
Figure 96. Results of the reliability analysis of limit state equation $g_2(X)$ – pdf of settlement (left); PCE realisations based on FE samples depending on combinations of h_{water} and ϕ'_{FG} (right).	226
Figure 97. Sketch of the retaining wall.	228
Figure 98. CPT 1-5 combined results and selected representative values.	228

Figure 99. Investigated system of soldier pile wall.....	229
Figure 100. ZSWalls model.....	230
Figure 101. Failure mechanism associated with the global safety factor, for one of the ZSWalls realisations.....	230
Figure 102. Histogram of safety factor and associated P_f for $d = 2.5$ m and $COV(\varphi) = 20\%$	231
Figure 103. Probability of failure with respect to the embedment depth, for $COV(\varphi) = 20\%$	232
Figure 104. Example of the system geometry of a design problem.....	233
Figure 105. Schemes of the limit states for internal and external stability of GRSW: a. sliding; b. overturning; c. bearing resistance; d. pullout; e. tensile failure.....	239
Figure 106. Variation of $P(f)$ for sliding along the base with $CV(\varphi_s)$	243
Figure 107. Variation of $P(f)$ for sliding along the base with $CV(\varphi_f)$	243
Figure 108. Variation of the tensile strength with $CV(\varphi_R)$ while $CV(\varphi_s) = CV(\varphi_f) = 20\%$	244
Figure 109. Variation of the length L with $CV(\varphi_s) = CV(\varphi_f)$ when $CV(\varphi_R) = 5\%$	244
Figure 110. Lower hemisphere equal-angle projection of the joint sets identified on field.....	247
Figure 111. Realization of a discrete fracture network (DFN).....	248
Figure 112. Rock wedge highlighted in Figure 111 , intersected by 4 joints from joint set 3 forming 5 rock blocks.....	249
Figure 113. Outcome of a rock joint shear testing program.....	250
Figure 114. Representation of the water load effects on joints (equivalent two-dimensional problem).....	252
Figure 115. Wedge height vs number of joints from set F3 intersecting them, for the DFN in Figure 111	253
Figure 116. Output from reliability analysis.....	254

List of tables

Table 1. Target reliability values for reliability index β (ultimate limit state) for different consequence classes according to EN 1990-1 Table C.3.2 (NDP). Corresponding note: Table C.3.2 gives target values β for the 50 years reference period assumed within this standard. Different target reliability values β for different reference periods can be set by the National Annex for use in a country.....	28
Table 2. Qualification of consequence classes according to EN 1990-1 Table 4.1 (NDP).....	29
Table 3. Examples of codes and standards specifying or recommending reliability targets.....	30
Table 4. Recommended influence coefficients (α) from ISO 2394:2015 (Table E.3) ¹	32
Table 5. Recommended reliability target values for geotechnical structures, differentiating between situation with low, moderate or high influence of the load (i.e. time-dependent variables).....	34
Table 6. Comparison between Frequentist and Bayesian interpretations of probability.....	41
Table 7. Probability estimates based on verbal classification (based on Table 6-7 in Vick 2002)..	41
Table 8. Probability-based groundwater actions according to EN 1990.....	45
Table 9. Recommended indicative coefficients of variation for inherent variability ($V_{x,inh}$) for various ground properties (based on ranges reported in TC 304 (2021), see Annex A to this report).	49
Table 10. Correction factor N_n for estimating the standard deviation from sample range for normally distributed variable as a function of the sample size n	52
Table 11. Ranges for typical scale of fluctuation.....	55
Table 12. Different cases of spatial averaging.....	56
Table 13. Recommended indicative measurement error in laboratory and in-situ tests (largely based on Phoon and Kulhawy 1999).....	60
Table 14. Indicative values for coefficients of variation for transformation uncertainty for various ground property pairs (largely based on ranges reported in TC 304 (2021), see Annex A).....	65
Table 15. Ranges of total soil property variability (for reliability calibration) according to Phoon & Ching (2015) and Phoon & Kulhawy (2008).....	65
Table 16. Commonly used continuous probability distributions.....	69
Table 17. Commonly used discrete probability distributions.....	70
Table 18. Suggested probability distributions for different soil and rock properties.....	71
Table 19. Needed information and constraints to assign certain pdfs.....	72
Table 20. Three-point estimate methods based on engineering judgement.....	72
Table 21. Recommended indicative cross-correlation coefficients. Parameters: c' = effective cohesion; C_c = compression index; C_s = swelling index; D_r = relative density; K_{DMT} = dilatometer horizontal stress index; K_0 = earth pressure at rest; K_s = saturated hydraulic conductivity; L_L = liquid	

limit; M_n = normalized effective constrained modulus determined by oedometer; P_i = plasticity index; w_n = gravimetric water content; φ = porosity; ϕ' = effective friction angle.....	75
Table 22. Recommended indicative model factors for geotechnical computation models*.....	80
Table 23. Relationship of reliability index β and failure probability P_f illustrated with numerical examples.....	87
Table 24. Recommended influence coefficients (α).....	108
Table 25. Deterministic and stochastic variables in the gravity wall example problem.....	111
Table 26. Results of the reliability analysis of GEOSNet benchmark problem 3 using different reliability methods.....	114
Table 27. Design point values and influence factors for GEOSNet benchmark problem 3 based on the FORM analysis.....	115
Table 28. Characteristic values according to the different Bayesian updating methods of the examples, assuming normally distributed data.....	131
Table 29. Partial resistance factor γ_R for a normal distributed resistance variable as a function of the uncertainty (V_R) and the relative influence (α_R).....	141
Table 30. Recommended site-specific coefficient of variation (COV).....	167
Table 31. Total measurement error.....	168
Table 32. Measurement error of in-situ tests.....	168
Table 33. Indicative values for transformation uncertainty.....	169
Table 34. Deterministic and random variables in the analysis. All random variables are assumed lognormal, except TMO, which is assessed to be normally distributed.....	179
Table 35. Overview of deterministic and random variables within this example.....	187
Table 36. Input parameters.....	192
Table 37. Summary of input parameters for the example.....	198
Table 38. Summary of pile loading test.....	199
Table 39. Summary of failure probabilities for case A.....	204
Table 40. Summary of failure probabilities for case B.....	206
Table 41. Representative values for the geotechnical parameters used in the design.....	211
Table 42. Statistics of the stochastic variables used in the probabilistic analysis.....	212
Table 43. Statistics of the soil parameters considered as random variables for the “Bucharest Clay” layer.....	213
Table 44. Reliability results for the verification of structural Ultimate Limit State for the normal design situation (Design Case 1).....	215

Table 45. Probabilistic results for limit state function ZM ($MRd = 1.25 \cdot Mk$) using Erraga metamodel.	215
Table 46. Probabilistic results for limit state function ZM ($MRd = 1.25 \cdot Mk$) using importance sampling around design point from Erraga metamodel.	216
Table 48. Probabilistic results for limit state function ZS ($SRd = 1.25 \cdot Sk$) using Importance Sampling around design point from Erraga metamodel.	217
Table 49. Range of ground property values of the present soil layers.	222
Table 51. Probability density functions of random input parameters.	223
Table 52. Overview of deterministic and random variables within this example.	229
Table 54. Input data for one run of Monte Carlo simulation with example of selected values for a specific project.	234
Table 55. Coefficient of variation (COV) of geotechnical and man-made-materials.	240
Table 56. Example of Monte Carlo iterations.	242
Table 57. Example of trial-and-error calculation.	243
Table 58. Input parameters.	249
Table 59. Description of random variables related to geometric uncertainties.	250

Annexes

Annex A. Literature summary of inherent variability and uncertainties

This annex contains a collection of literature values on the various sources of uncertainty considered in the probabilistic modeling for reliability verification. The reported ranges underpin the recommendations on the main text and provide further context. Most of the material stems from the state-of-the-art report published by ISSMGE-TC304 (2021).

A.1. Inherent variability

The recommended values for COV in **Table 30** correspond to the mean COV, largely based on global summaries of site-specific statistics (TC304, 2021). Range of mean values indicates the applicability of the recommended COV.

In absence of sufficient local or site-specific data, these recommended values may be taken as *cautious estimates* for $V_{X,inh}$ because (1) site-specific variability can be greater than variability within one homogeneous geotechnical unit (i.e. may be only part of the site), and (2) most of the COV values in Table A.1 represent observed variability $V_{X,obs}$, which includes measurement error $V_{X,meas}$ in addition to inherent variability $V_{X,inh}$. Hence, if these values are used as $V_{X,inh}$, the chosen value for $V_{X,meas}$ should be chosen as a lower bound estimate (or even zero).

Table 30. Recommended site-specific coefficient of variation (COV)

Property	Soil	Recommended COV	Range of COV (%) in literature ^(a)	Range of mean values ^(a)	Remarks
γ (kN/m ³) (total)	clay and sand	0.10	3-20	13-24	From Cao et al. (2016). COV = 0.05 for homogeneous soils.
ϕ' (°)	clay	0.20	10-50	3-33.3	From TC304 (2021).
ϕ' (°)	sand	0.08	4.3-12.4	32.4-51.5	From TC304 (2021).
c' (kPa)	clay	0.20	0.02-0.7	–	From Arnold (2016)
s_u (kPa)	clay	0.30	9.9-53.5	7.2-558.4	From TC304 (2021).
s_u/σ'_v	clay	0.20	5.0-39.3	0.06-1.07	From TC304 (2021).
OCR	clay	0.20	1.5-38.8	0.90-3.11	From TC304 (2021).
C_c	clay	0.35	18.1-47.3	0.19-2.15	From TC304 (2021).
C_{ur}	clay	0.40	22.6-50.5	0.03-0.21	From TC304 (2021).
K_0	clay	0.15	2.4-22.0	0.48-2.88	From TC304 (2021).
K_0	sand	0.30	25.8-36.9	0.64-2.20	From TC304 (2021).
SPT-N	clay	0.30	15.9-57	1.75-75.3	From TC304 (2021).
SPT-N	sand	0.35	18.5-61.0	6.8-73.3	From TC304 (2021).
q_c (MPa)	clay	0.30	16-40	1.2-2.1	From TC304 (2021).
q_c (MPa)	sand	0.40	17.0-77.4	0.85-13.17	From TC304 (2021).
q_{t1} (MPa)	clay	0.20	5.8-39.7	2.04-13.13	From TC304 (2021).
E (MPa) (in-situ)	sand	0.35	8.7-73.0 ^b	5.24-62.0 ^b	From TC304 (2021). In-situ tests: DMT and PMT.

^(a) 95% confidence interval (TC304, 2021) or reported range (Arnold 2016; Cao et al. 2016)

^(b) Combined range based on the statistics for E_{DMT} and E_{PMT}

Source: Authors' own work

A.2. Measurement error

Table 31 and **Table 32** below provide summaries of measurement error from Phoon and Kulhawy (1999) for various laboratory and in-situ tests. We may assume that the measurement techniques have improved in the last two decades, and that the reported values do not refer to exclusively measurement error in all cases (since other error sources are hard to separate from measurement error in the assessment). Therefore, the reported values are considered rather high, and the recommended values in the main text are in the lower ranges.

Table 31. Total measurement error

Table 5. Summary of total measurement error of some laboratory tests (source: Phoon et al. 1995, p. 4-22).

Property ^a	Soil type	No. of data groups	No. of tests per group		Property value		Property COV (%)	
			Range	Mean	Range	Mean	Range	Mean
s_u (TC) (kN/m ²)	Clay, silt	11	—	13	7–407	125	8–38	19
s_u (DS) (kN/m ²)	Clay, silt	2	13–17	15	108–130	119	19–20	20
s_u (LV) (kN/m ²)	Clay	15	—	—	4–123	29	5–37	13
$\bar{\phi}$ (TC) (°)	Clay, silt	4	9–13	10	2–27	19.1	7–56	24
$\bar{\phi}$ (DS) (°)	Clay, silt	5	9–13	11	24–40	33.3	3–29	13
$\bar{\phi}$ (DS) (°)	Sand	2	26	26	30–35	32.7	13–14	14
$\tan \bar{\phi}$ (TC)	Sand, silt	6	—	—	—	—	2–22	8
$\tan \bar{\phi}$ (DS)	Clay	2	—	—	—	—	6–22	14
w_n (%)	Fine grained	3	82–88	85	16–21	18	6–12	8
w_L (%)	Fine grained	26	41–89	64	17–113	36	3–11	7
w_P (%)	Fine grained	26	41–89	62	12–35	21	7–18	10
PI (%)	Fine grained	10	41–89	61	4–44	23	5–51	24
γ (kN/m ³)	Fine grained	3	82–88	85	16–17	17.0	1–2	1

^aLV, laboratory vane shear test.

Source: Phoon and Kulhawy (1999)

Table 32. Measurement error of in-situ tests

Table 6. Summary of measurement error of common in situ tests (source: Orchant et al. 1988, p. 4-63; Kulhawy and Trautmann 1996, p. 283).

Test	Coefficient of variation, COV (%)				
	Equipment	Procedure	Random	Total ^a	Range ^b
Standard penetration test (SPT)	5–75 ^c	5–75 ^c	12–15	14–100 ^c	15–45
Mechanical cone penetration test (MCPT)	5	10–15 ^d	10–15 ^d	15–22 ^d	15–25
Electric cone penetration test (ECPT)	3	5	5–10 ^d	7–12 ^d	5–15
Vane shear test (VST)	5	8	10	14	10–20
Dilatometer test (DMT)	5	5	8	11	5–15
Pressuremeter test, prebored (PMT)	5	12	10	16	10–20 ^e
Self-boring pressuremeter test (SBPMT)	8	15	8	19	15–25 ^e

^aCOV(Total) = [COV(Equipment)² + COV(Procedure)² + COV(Random)²]^{0.5}.

^bBecause of limited data and judgment involved in estimating COVs, ranges represent probable magnitudes of field test measurement error.

^cBest to worst case scenarios, respectively, for SPT.

^dTip and side resistances, respectively, for CPT.

^eIt is likely that results may differ for p_o , p_t , and p_L , but the data are insufficient to clarify this issue.

Source: Phoon and Kulhawy (1999)

A.3. Transformation uncertainty

The values in **Table 33** for bias and coefficient variation (V_{trans}) are indicative values based on the average values reported in the literature. Transformation models calibrated with regional data (contains multiple sites) are typically marked with smaller transformation uncertainty and should therefore be preferred to statistics based on global data (if available). The smallest amount of transformation uncertainty is usually encountered in transformation models, which have been

carefully fitted to regional data (i.e., outliers removed and the most suitable predictor variable and function chosen).

Table 33. Indicative values for transformation uncertainty.

Transformation model	Soils	Global ^(a)		Regional data, examples ^(b)		
		Bias (<i>b</i>)	V_{trans} (δ)	Bias (<i>b</i>)	V_{trans} (δ)	Ref. ^(c)
$s_u - \sigma'_p$	clays	1.0	0.50	0.95	0.30	A
$\left(\frac{s_u}{\sigma'_v}\right) - OCR$	clays	1.1	0.50	1.0	0.30	A
$\left(\frac{s_u}{\sigma'_v}\right) - CPT$	clays	0.95	0.50	N/A	0.30 (0.20)	B
$OCR - CPT$	clays	1.0	0.40	0.7	0.20	C
$\sigma'_p - CPT$	clays	1.0	0.40	0.8 (1)	0.20 (0.15)	C (D)
$C_c - LL$	clays	1.2	0.90	1.5	0.70	E
$C_c - e_0$	clays	N/A	N/A	1.3 (1)	0.50 (0.35)	E (E)
$C_s - e_0$	clays	0.31	0.70	1.0	0.40	E
$\varphi' - SPT$	sands	1.1	0.10 ^(d)			
$\varphi' - CPT$	sands	0.95	0.10 ^(d)			
$D_r - SPT$	sands	1.0	0.20 ^(d)			
$D_r - CPT$	sands	0.8	0.30 ^(d)			

^(a) TC 304 (2021): Global calibration database for clays (CLAY/10/7490) and for sands (SAND/7/2794).

^(b) Values in parenthesis are transformation uncertainties for models carefully fitted with regional data (instead of general models calibrated with regional data).

^(c) A = D'Ignazio et al. (2016): databases of Finnish clays (F-CLAY/10/216) and Scandinavian clays (S-CLAY/10/168); B = Paniagua et al. (2019): high-quality database of Norwegian clays; C = D'Ignazio et al. (2019): global database with high-quality calibration data only (CLAY-9/249); D = Di Buò (2020): database of soft sensitive Finnish clays (homogeneous data); E = Löfman and Korkiala-Tanttu (2022): database of Finnish clays (FI-CLAY/14/822).

^(d) Global database SAND/7/2794 contains high-quality data and is comparable with regional transformation uncertainty.

Sources: see the above notes

References

D'Ignazio, M., Phoon, K. K., Tan, S. A., & Länsivaara, T. T. (2016). Correlations for undrained shear strength of Finnish soft clays. *Canadian Geotechnical Journal*, 53(10), 1628-1645.

D'Ignazio, M., Lunne, T., Andersen, K. H., Yang, S., Di Buò, B., & Länsivaara, T. (2019). Estimation of preconsolidation stress of clays from piezocone by means of high-quality calibration data. *AIMS Geosciences*, 5(2): 104-116.

Di Buò, B. (2020). "Evaluation of the preconsolidation stress and deformation characteristics of Finnish clays based on piezocone testing." Doctoral dissertation, Tampere University.

Löfman, M. S., & Korkiala-Tanttu, L. K. (2022). Transformation models for the compressibility properties of Finnish clays using a multivariate database. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 16(2), 330-346.

Paniagua Lopez, A. P., D'Ignazio, M., L Heureux, J. S., Lunne, T., & Karlsrud, K. (2019). CPTU correlations for Norwegian clays: an update. *AIMS Geosciences*, 5(2): 82-103.

Table summaries for transformation uncertainty of clay, sand, and rock (intact and rock mass) extracted from ISSMGE-TC304 report

Table 5.2. Transformation models for clay

Model	Literature	n	Transformation model	b	δ	Data restriction / Application range	Calibration database
$s_u^{re} - LI$	Locat and Demers (1988)	899	$s_u^{re}/P_a \approx 0.0144 \times LI^{-2.44}$	1.92	1.25		CLAY/10/7490
$S_t - LI$	Bjerrum (1954)	1279	$S_t \approx 10^{0.8LI}$	2.06	1.09		CLAY/10/7490
$\sigma'_p - LI - S_t$	Stas and Kulhawy (1984)	249	$\sigma'_p/P_a \approx 10^{1.11-1.62 \times LI}$	2.94	1.90	$S_t < 10$	CLAY/10/7490
	Ching and Phoon (2012b)	489	$\sigma'_p/P_a \approx 0.235 \times LI^{-1.319} \times S_t^{0.536}$	1.32	0.78		CLAY/10/7490
$s_u - \sigma'_p$	Mesri (1975, 1989)	1155	$s_u(mob)/\sigma'_p \approx 0.22$	1.04	0.55		CLAY/10/7490
$\left(\frac{S_u}{\sigma'_v}\right) - OCR$	Jamiolkowski et al. (1985)	1402	$s_u(mob)/\sigma'_v \approx 0.23 \times OCR^{0.8}$	1.11	0.53		CLAY/10/7490
$\left(\frac{S_u}{\sigma'_v}\right) - OCR - S_t$	Ching and Phoon (2012b)	395	$s_u(mob)/\sigma'_v \approx 0.223 \times OCR^{0.823} \times S_t^{0.121}$	0.84	0.34		CLAY/10/7490
$\left(\frac{S_u}{\sigma'_v}\right) - CPT$	Ching and Phoon (2012a)	423	$s_u(mob)/\sigma'_v \approx \frac{(q_t - \sigma_v)/\sigma'_v}{29.1 \times \exp(-0.513B_q)}$	0.95	0.49		CLAY/10/7490
$OCR - CPT$	Chen and Mayne (1996)	690	$OCR \approx 0.259 \times [(q_t - \sigma_v)/\sigma'_v]^{1.107}$	1.01	0.42		CLAY/10/7490
	Kulhawy and Mayne (1990)	690	$OCR \approx 0.32 \times (q_t - \sigma_v)/\sigma'_v$	1.00	0.39		CLAY/10/7490
$\sigma'_p - CPT$	Chen and Mayne (1996)	690	$\sigma'_p/P_a \approx 0.227 \times [(q_t - \sigma_v)/P_a]^{1.200}$	0.99	0.42		CLAY/10/7490
	Kulhawy and Mayne (1990)	690	$\sigma'_p \approx 0.33 \times (q_t - \sigma_v)$	0.97	0.39		CLAY/10/7490
$C_c - LL$	Skempton (1944)	3398	$C_c \approx 0.007 \times (LL - 10)$	1.59	0.90	Remolded clay	CLAY/8/12225
	Terzaghi and Peck (1967)	3398	$C_c \approx 0.009 \times (LL - 10)$	1.24	0.90		CLAY/8/12225
$C_c - PI$	Kulhawy and Mayne (1990)	2964	$C_c \approx PI/73$	1.31	0.91		CLAY/8/12225
$C_c - LL - G_s$	Nagaraj and Murty (1985)	1523	$C_c \approx 0.2343 \times (LL/100) \times G_s$	1.06	0.81		CLAY/8/12225
$C_s - e_0$	Peck and Reed (1954)	668	$C_s \approx 0.208 \times e_0 + 0.0083$	0.31	0.69		CLAY/8/12225
$C_s - \omega$	Azzouz et al. (1976)	771	$C_s \approx 0.003 \times (\omega + 7)$	0.47	0.67		CLAY/8/12225

$C_s - PI$	Kulhawy and Mayne (1990)	847	$C_s \approx PI/385$	0.96	0.63		CLAY/8/12225
$C_s - LL$	Isik (2009)	846	$C_s \approx 0.0007 \times LL + 0.0062$	1.55	0.64		CLAY/8/12225
$K_0 - OCR$	Mayne and Kulhawy (1990)	1009	$K_0 \approx [1 - \sin(\phi')] \times OCR^{\sin(\phi')}$	1.00	0.34		CLAY/12/3997
		124	$K_0 \approx 0.27 \times K_{DMT}$	0.92	0.45		CLAY/12/3997
$K_0 - K_{DMT}$	Powell and Uglow (1988)	124	$K_0 \approx 0.34 \times K_{DMT}^{0.55}$	1.22	0.30		CLAY/12/3997
	Kulhawy et al. (1989)	74	$K_0 \approx 0.073 \times N_{60}/(\sigma'_v/P_a)$	2.16	0.70		CLAY/12/3997
$K_0 - CPT$		74	$K_0 \approx 0.1 \times (q_t - \sigma_v)/\sigma'_v$	1.18	0.74		CLAY/12/3997
$E_{PMT} - N_{60}$	Kulhawy and Mayne (1990), Ohya et al. (1982)	812	$E_{PMT}/P_a \approx 19.3 \times N_{60}^{0.63}$	1.60	1.28		CLAY/12/3997
$M' - CPT$	Mayne (2007)	111	$M' \approx 5 \times (q_t - \sigma_v)$	1.57	0.75		CLAY/12/3997

Note: $s_u(mob)$ = undrained shear strength mobilized in embankment and slope failures (Mesri and Huvaj 2007) ; E_{PMT} = pressuremeter modulus; M' = effective constrained modulus.

Table 5.3. Transformation models for sand

Model	Literature	n	Transformation model	b	δ	Data restriction / Application range	Calibration database
$D_r - SPT$	Terzaghi and Peck (1967)	198	$D_r(\%) \approx 100 \times \sqrt{(N_1)_{60}/60}$	1.05	0.23	$(N_1)_{60} < 60$	SAND/7/2794
	Kulhawy and Mayne (1990)	199	$D_r(\%) \approx 100 \times \sqrt{\frac{(N_1)_{60}}{[60 + 25 \log_{10}(D_{50})] \times OCR^{0.18}}}$	1.01	0.21		SAND/7/2794
$D_r - CPT$	Jamiolkowski et al. (1985)	681	$D_r(\%) \approx 68 \times [\log_{10}(Q_{tn}) - 1]$	0.84	0.33	$Q_{tn} < 300$	SAND/7/2794
$\phi' - D_r$	Bolton (1986)	391	$\phi' \approx \phi'_{cv} + 3 \times (D_r [10 - \ln(p'_f)] - 1)$	1.03	0.052		SAND/7/2794
$\phi' - SPT$	Hatanaka et al. (1998)	58	$\phi' \approx \begin{cases} \sqrt{15.4 \times (N_1)_{60}} + 20 & (N_1)_{60} \leq 26 \\ 40 & (N_1)_{60} > 26 \end{cases}$	1.07	0.090	$(N_1)_{60} < 150$	SAND/7/2794

	Chen (2004)	59	$\phi' \approx 27.5 + 9.2 \times \log_{10}[(N_1)_{60}]$	1.00	0.095		SAND/7/2794
$\phi' - CPT$	Robertson and Campanella (1983)	99	$\phi' \approx \tan^{-1}[0.1 + 0.38 \times \log_{10}(q_t/\sigma'_v)]$	0.93	0.056		SAND/7/2794
	Kulhaway and Mayne (1990)	376	$\phi' \approx 17.6 + 11 \times \log_{10}(Q_{tn})$	0.97	0.081		SAND/7/2794
$E_{PMT} - N_{60}$	Kulhaway and Mayne (1990), Ohya et al. (1982)	1081	$E_{PMT}/P_a \approx 9.08 \times N_{60}^{0.66}$	2.24	1.05		SAND/10/4113
$E_{DMT} - N_{60}$	Mayne and Frost (1989)	591	$E_{PMT}/P_a \approx 22 \times N_{60}^{0.82}$	1.69	0.72		SAND/10/4113
$M' - CPT$	Mayne (2007)	113	$M' \approx 5 \times (q_t - \sigma'_v)$	1.57	0.76		SAND/10/4113
$K_0 - OCR$	Mayne and Kulhaway (1990)	1207	$K_0 \approx [1 - \sin(\phi')] \times OCR^{\sin(\phi')}$	1.05 ($\psi = 30^\circ$)	0.50 ($\psi = 30^\circ$)		SAND/10/4113

Note: E_{DMT} = pressurometer modulus; E_{PMT} = pressurometer modulus; M' = effective constrained modulus.

Table 5.4. Transformation models for intact rock

Model	Literature	n	Transformation model	b	δ	Data restriction / Application range	Calibration database
$\sigma_{ci} - n$	Kilic and Teymen (2008)	911	$\sigma_{ci} \approx 147.16 \times e^{-0.0835n}$	0.91	0.75	$0.16 < n < 37.81$	ROCK/9/4069
$\sigma_{ci} - R_L$	Karaman and Kesimal (2015)	664	$\sigma_{ci} \approx 0.1383 \times R_L^{1.743}$	0.78	0.53	$11.82 < R_L < 59.59$	ROCK/9/4069
$\sigma_{ci} - S_h$	Altindag and Guney (2010)	297	$\sigma_{ci} \approx 0.1821 \times S_h^{1.5833}$	1.15	0.65	$9 < S_h < 100$	ROCK/9/4069
$\sigma_{ci} - \sigma_{bt}$	Prakoso and Kulhaway (2011)	525	$\sigma_{ci} \approx 7.8 \times \sigma_{bt}$	1.31	0.50	$5 < \sigma_{bt} < 25.64$	ROCK/9/4069
$\sigma_{ci} - I_{s50}$	Mishra and Basu (2013)	1074	$\sigma_{ci} \approx 14.63 \times I_{s50}$	1.18	0.45	$1.15 < I_{s50} < 14.13$	ROCK/9/4069
$\sigma_{ci} - V_p$	Kahraman (2001)	1247	$\sigma_{ci} \approx 9.95 \times V_p^{1.21}$	1.26	0.63	$1.02 < V_p < 6.3$	ROCK/9/4069
$E_i - R_L$	Katz et al. (2000)	289	$E_i \approx 0.000113 \times R_L^{3.09074}$	1.47	0.99	$24.01 < R_L < 73.3$	ROCK/9/4069
$E_i - S_h$	Deere and Miller (1966)	197	$E_i \approx 0.739 \times S_h + 11.51$	0.61	0.71	$11 < S_h < 105$	ROCK/9/4069
$E_i - \sigma_c$	Deere and Miller (1966)	1152	$E_i \approx 0.303 \times \sigma_{ci} - 0.8745$	1.23	0.94	$21 < \sigma_{ci} < 1330$	ROCK/9/4069
$E_i - V_p$	Yasar and Erdogan (2004)	192	$E_i \approx 10.67 \times V_p - 18.71$	0.90	0.72	$3.11 < V_p < 5.6$	ROCK/9/4069

Note: σ_{ci} in MPa; E_i in GPa; I_{s50} in MPa; σ_{bt} in MPa; V_p in km/s; n in %.

Table 5.5. Transformation models for rock mass

Model	Literature	n	Transformation model	b	δ	Data restriction / Application range	Calibration database
$E_m - E_i - RQD$	Coon and Merritt (1970)	147	$E_m/E_i = 0.0231 \times RQD - 1.32$	1.26	1.09	$57 < RQD < 100$	ROCKMass/9/5876
	Zhang and Einstein (2004)	161	$E_m/E_i = 10^{0.0186 \times RQD - 1.91}$	1.54	0.89	$0 < RQD < 100$	ROCKMass/9/5876
$E_m - RMR$	Bieniawski (1978)	1091	$E_m = 2 \times RMR - 100$	0.57	1.48	$50 < RMR < 85$	ROCKMass/9/5876
	Gokceoglu et al. (2003)	1749	$E_m = 0.0736 \times e^{0.0755 \times RMR}$	1.53	1.21	$20 < RMR < 85$	ROCKMass/9/5876
	Serafim and Pereira (1983)	1749	$E_m = 10^{\frac{(RMR-10)}{40}}$	0.51	1.00	$20 < RMR < 85$	ROCKMass/9/5876
$E_m - Q$	Grimstad and Barton (1993)	288	$E_m = 25 \times \log_{10}(Q)$	0.58	0.88	$1.1 < Q < 1000$	ROCKMass/9/5876
$E_m - GSI$	Hoek and Diederichs (2006)	349	$E_m = 100 \times \frac{(1-D/2)}{1 + e^{\frac{75+25D-GSI}{11}}}$	0.83 ($D=0$)	1.00 ($D=0$)	$10 < GSI < 100$	ROCKMass/9/5876

Note: σ_{ci} in MPa; E_m in GPa; D is disturbance factor (Hoek and Diederichs 2006).

Source: ISSMGE-TC304 (2021)

A.4 Model uncertainty

Robust evaluation of statistics for the model factor M (see section 5.5) usually requires (Phoon and Kulhaway 2005):

- large-scale model tests (prototype tests)
- a database that is sufficiently large and representative
- high-quality testing with control for excessive uncertainties

The benefit in laboratory-scale load tests is that other sources of uncertainty (such as inherent variability) can be minimized. On contrary, field load tests provide a more diverse collection of geometries and soil parameters, but may also contain extraneous uncertainties. However, the statistics for model factors may be rather similar when compared, possibly due to normalization effect (see Phoon and Kulhaway 2005).

Adequate estimates for the statistics for the model factor may be available, based on sets of field measurements and predictions (DNV 1997). It should be noted that the defined bias is the average bias for the sets in the database: the calculation method can be unconservative for a specific case even though the mean bias would be greater than 1 (i.e. conservative on average). Once the

variance has been assessed, it is possible to estimate a probability of measured value being lower than the calculated value.

Calculation model uncertainties (bias and COV) have been recently summarized by ISSMGE-TC304 (2021) (table in Annex A): “The dataset used in calibration include laboratory (scaled model or prototype in a centrifuge facility) (representing controlled soil condition) or in situ (representing natural soil condition) load tests. The results cover various geotechnical structures (e.g., shallow foundations, offshore spudcans, pipes, anchors, drilled shafts, driven piles, rock sockets, helical piles, mechanically stabilized earth walls, soil nail walls, slopes and braced excavations) and a wide range of geomaterials from soft clay to soft rock. Two typical limit states (i.e., ULS and SLS) are calibrated. The mean and COV values and number of tests (N) averaged over n data groups that, belong to the same geotechnical structure, limit state, and geomaterial are presented in Figure 4.1.”

A comprehensive summary of uncertainties in calculation models is given in a book by Tang and Phoon (2021). A summary table of model factor statistics for various geotechnical structures (from Tang and Phoon 2021) can be found below.

Table summary for of model factor statistics for various geotechnical structures

Table 8.3 Model statistics for various geostructures (Source: data taken from Phoon and Tang 2019)

Geostructure	Limit state	Geomaterial	No. data groups	No. of tests per group		Design method	Mean		COV			
				Range	Mean		Range	Mean	Range	Mean		
Footings	Bearing Settlement	Sand	6	6–138	51	Vesic (1975)	0.99–1.67	1.33	0.23–0.47	0.35		
		Sand	4	90–131	106	D'Appolonia et al. (1970)	1.13–1.71	1.48	0.47–1	0.65		
	Punch-through	Sand-clay	1	—	49	Hough (1959)	—	0.66	—	0.45		
			2	27–95	61	Load spread (1:3)	1.49–1.81	1.65	0.27–0.31	0.29		
			2	27–95	61	Load spread (1:5)	2.36–2.37	2.37	0.33–0.38	0.36		
Rock sockets	Bearing Settlement	Rock	1	—	58	Punching shear	1.61–2.71	2.16	0.29–0.46	0.38		
			1	—	52	Okamura et al. (1998)	0.76–0.85	0.81	0.12–0.22	0.17		
			1	—	58	Hu (2015)	0.82–1	0.91	0.13–0.19	0.16		
			1	—	52	Goodman (1989)	—	1.23	—	0.54		
			1	—	61	Kulhaway (1978)	—	0.98	—	1.36		
Driven piles	Static	Rock	1	—	61	Goodman (1989)	—	1.52	—	0.54		
			1	—	37	Kulhaway (1978)	—	1.64	—	1.73		
			Compression	Clay	26	4–115	28	Total stress analysis	0.39–1.54	0.97	0.13–0.62	0.34
					24	5–71	29	Effective stress analysis	0.61–1.66	1.18	0.21–0.64	0.43
					17	13–80	40	Total/effective stress	0.48–1.81	0.91	0.31–0.59	0.45
Tension	Clay	8	4–69	28	Total stress analysis	0.74–1.43	1.01	0.13–0.39	0.27			
		8	5–51	20	Effective stress analysis	0.98–1.6	1.26	0.22–0.56	0.36			

Dynamic	Compression	Soil	1	—	125	CAPWAP (EOD)	—	1.63	—	0.49	
			1	—	162	CAPWAP (BOR)	—	1.16	—	0.34	
			5	34–175	90	WEAP (EOD)	1.27–1.94	1.67	0.52–0.77	0.69	
			4	34–175	87	WEAP (BOR)	0.9–1.12	1.03	0.36–0.55	0.47	
			3	90–135	102	FHWA Gates formula	0.77–1.07	0.89	0.29–0.53	0.42	
Drilled shafts	Compression	Clay	6	13–64	41	Brown et al. (2010)	0.84–1.15	0.99	0.25–0.5	0.44	
			Sand	11	9–46	30	Brown et al. (2010)	0.48–2.57	1.35	0.24–0.74	0.48
				9	10–90	28	Brown et al. (2010)	0.6–1.32	1.09	0.16–0.58	0.32
	Tension	Clay	2	13–32	22	Brown et al. (2010)	0.87–1	0.94	0.34–0.37	0.36	
			Sand	4	11–49	26	Brown et al. (2010)	0.83–1.25	1.06	0.32–0.54	0.45
				3	14–39	26	Brown et al. (2010)	1.07–1.25	1.16	0.29–0.48	0.4
Pile foundations	Lateral	Clay	1	—	72	Broms (1964a)	—	1.49	—	0.38	
			Sand	1	—	75	Broms (1964b)	—	1.22	—	0.4
				1	—	29	Poulos (1994)	—	1.11	—	0.65
	Settlement	Soil	2	29–31	30	Load transfer	0.94–1.18	1.06	0.5–0.78	0.64	
			Soil	4	22–62	44	t-z curve (“FZ” model)	1.23–1.41	1.29	0.44–0.66	0.58
				t-z curve (“AB1” model)	0.78–1.02	0.94	0.4–0.71	0.62			
t-z curve (“AB2” model)	0.66–0.89	0.81	0.67–1.09	0.88							
Plate anchors	Pullout	Sand	1	—	54	Limit equilibrium	—	1.16	—	0.23	
Pipes	Pullout	Sand	3	61–300	168	Limit equilibrium	0.81–1.41	1.06	0.23–0.39	0.3	
Slopes	Global stability	Soil	7	24–134	51	Bishop	0.89–1.27	1.01	0.15–0.28	0.22	
SNW	Nail tensile load	Weak rock	2	45–54	49	Lazarte et al. (2015)	0.66–0.95	0.8	0.38–0.52	0.45	
			2	30–74	52	Lazarte et al. (2015)	2.98–3.58	3.3	0.36–0.43	0.4	
	Facing tensile force	—	2	23–42	32	Lazarte et al. (2015)	0.77–0.85	0.81	0.43–0.67	0.55	
MSE	Maximum tensile load	Geosynthetic	6	41–143	83	Berg et al. (2009a, b)	0.16–0.66	0.4	0.73–1.46	1.01	
			4	29–104	70	Berg et al. (2009a, b)	0.85–1.36	1.14	0.39–0.58	0.47	
	Pullout	Geosynthetic	3	25–159	106	Berg et al. (2009a, b)	2.02–2.68	2.37	0.47–0.59	0.52	
9			17–129	63	Berg et al. (2009a, b)	1.12–2.73	1.9	0.33–0.69	0.48		
MAW	Maximum tensile load	c ≥ 0, φ > 0	1	—	36	PWRC (2002)	—	0.81	—	0.79	
			Pullout	c ≥ 0, φ > 0	1	—	28	PWRC (2002)	—	1.21	—

Source: Tang and Phoon 2021

Box 37. Overestimation of model error

Estimation of model uncertainty is hard, mainly due to various uncertainties playing a role in model predictions, and also due to the presence of measurement error in experiments or field observations. Ideally, we would single out the error exclusively due to model choices and simplifications in deriving model error statistics, because in the forward modeling as described in section 5.5 we include other sources of uncertainty separately.

Practically speaking, however, it is impossible to single out the actual model error only, which results in over-estimation of the model uncertainty (coefficient of variation) because other components of variability and uncertainty are still included.

For this reason, the indicative (i.e. recommended) values in Table 22 are relatively low compared to the ranges reported in the literature.

A.5. Groundwater level (background)

The estimated level of groundwater is affected by measurement error, time frame of the measurement series, and uncertainties related to the interpolation. In addition, the groundwater

level is typically marked with seasonal variations. Hence, the groundwater level is characterized by both spatial and temporal autocorrelation. It is noteworthy that the groundwater in rock (e.g., in heterogeneous crack systems) is often characterized by more complex spatial and temporal variability compared to soil aquifers. Moreover, the temporal variations of groundwater level may be affected by events such as groundwater recharge. Machiwal et al. (2012) observed that the monthly groundwater levels were normally distributed. This is not however always the case, especially for short periods (e.g., Bloomfield and Marchant 2013).

Geostatistical methods like kriging allow the groundwater level to be estimated in a time series (e.g., Ahmadi and Sedghamiz 2007, Machiwal et al. 2012). Other, often more useful methods include impulse-response functions and lumped parameter groundwater models (e.g., Peterson and Western 2018, Mackay et al. 2014). In practical applications (e.g. in the case of water pressure-related loads), it has been common practice to model the groundwater level deterministically as a conservative point estimate. Nonetheless, probabilistic modelling of groundwater level should be preferred since it allows to consider the various uncertainties related to the groundwater level measurements such as the measurement errors, frequency and observation period. For instance, it is evident that more frequent measurements (e.g., automatic daily monitoring) taken over a long time period lead are marked with smaller uncertainties compared to measurements taken once a month for a short time period.

The probability of exceeding certain groundwater level or pressure depends on the considered time period. The binomial distribution can be used to calculate the probability of exceedance p_e , i.e. the probability that a certain groundwater level or pressure is exceeded *at least once* during a time period of n years is given by:

$$p_e = 1 - \left[1 - \left(\frac{1}{T} \right) \right]^n$$

Equation 87.

where T is the return period (e.g. 50 years). The annual probability of exceedance is given by setting $n = 1$.

A.6. Statistical uncertainty

The equation for sample standard deviation s_x is based on the formula for *unbiased sample variance*. If the variable X follows normal distribution with parameters mean μ_x and variance σ_x^2 , the sample variance s_x^2 (i.e. standard deviation is $\sqrt{s_x^2}$) follows a scaled chi-squared distribution. The statistical uncertainty in estimating σ_x^2 of the population using a sample with n observations can be evaluated by means of variance of the estimate s_x^2 (e.g. NVD 2012):

$$\sigma_{s_x^2}^2 \approx \frac{2(s_x^2)^2}{n-1}$$

Equation 88.

Hence, a greater statistical uncertainty is introduced if variance is estimated for a ground property which is characterized by large inherent variability.

The sample coefficient of variation V_x tends to be too low when small sample is taken from a population. For normally distributed data, the *unbiased estimate* for V_x is given by $[1+1/(4n)] \cdot V_x$.

According to TC304 (2021), there should be at least $n = 10$ observations if sample V_x is estimated from site-specific data, whereas statistics with $n = 30$ can be considered very reliable (TC304, 2021). Approximate guidelines on appropriate sample sizes may also be drawn from confidence interval for V_x (Kelley, 2007): it can be shown that when considering a normally distributed random variable, its V_x follows a noncentral t-distribution. For example, if the population V_x is 0.20, the width of its 95 % confidence interval is $\omega = 0.10$ with 80 % assurance if the number of samples is $n = 45$. In other words, if $n = 45$ observations from population with $V_x = 0.20$ are used to calculate the sample V_x , we are 80 % certain that the 95 % confidence interval is from 0.15 to 0.25. Generally, a larger sample size is needed as (1) population V_x increases, (2) narrower confidence intervals are targeted (i.e., smaller ω), (3) greater assurance is targeted, or if (4) confidence interval with wider coverage is targeted (e.g. 95 % versus 99 % confidence interval).

Annex B. Examples

B.1. Introduction

This Annex provides practical examples on the application of reliability-based methods in the design of geotechnical structures. Annex B is divided into subsections, with each of the subsections B.2 to B.7 containing examples related to the geotechnical structures covered by Eurocode 7, part 3, as far as they were available. All examples have been processed following the procedure of reliability verification outlined in the guideline. The last section, B.8, contains a list of additional examples from the literature, sorted by topic.

Box 38. Notes to the examples

The examples presented in Annex B were not developed specifically for this guideline and were essentially derived from past practical projects based on the results of a questionnaire. Thus, there may be differences to the guideline recommendations, e.g., different definitions of target reliability or different, more pragmatic procedures to consider uncertainties. The reader therefore is advised to consider the relevant sections of the guideline for further information on the various methods and their application. Nevertheless, where specifically required detailed references will be provided. Further on, the software tools listed in the examples are purely informative and do not represent a recommendation for their use.

The following examples will be presented in detail:

Geotechnical Structure according to EC7-3		Example	
B.2	Slopes, cuttings, and embankments	B.2.1.	Settlement analysis for embankments using preloading without surcharge
B.3	Spread foundations	B.3.1.	Vertically loaded spread foundation
B.4	Piled foundations	B.4.1.	Comparison of reliability levels achieved by the pile design method of EN 1997 with the reliability targets of EN 1990
		B.4.2.	Reliability updating of driven pile using pile tests and Bayesian inference
B.5	Retaining structure	B.5.1.	Probabilistic FEM analysis of a retaining wall
		B.5.2.	Soldier pile wall
		B.5.3.	Propped embedded retaining wall
B.6	Reinforced fill structures	B.6.1	Geogrid reinforced soil wall
B.7	Rock Engineering	B.7.1	Rock slope stability

B.2. Slopes, cuttings, and embankments

Authors: Karl Escher, Johan Spross

Reference: Escher, K. 2022. Probabilistic settlement analysis for embankments using preloading without surcharge. MSc thesis. TRITA-ABE-MBT 22636. KTH Royal Institute of Technology, Stockholm. Link to publication: <http://urn.kb.se/resolve?urn=urn:nbn:se:kth.diva-317375>

B.2.1 Settlement analysis for embankments using preloading without surcharge

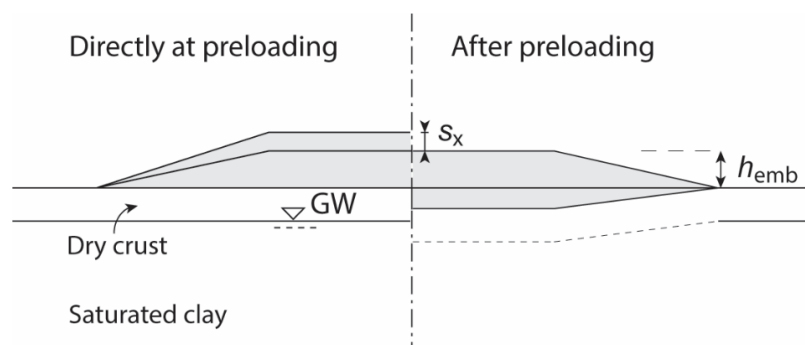
Problem definition

Description

In the design of an embankment on soft clay, residual settlement after completion of the embankment needs to be considered as a serviceability limit state. Various ground improvement methods are available, such as dry deep mixing columns or prefabricated vertical drains in combination with a surcharge. This example concerns another option: preloading without a surcharge; its principle is shown in **Figure 64**.

The embankment has a crest width of 18.5 m and a target height $h_{emb} = 1.8$ m. The slope inclination has a 1:2 ratio. The soil consists of 7.5 m clay, where the upper 1.5 m is dry crust. The clay overlies a firm till which is not depicted in **Figure 64**. Consolidation settlements are expected to be an issue in the saturated clay.

Figure 64. Preloading an embankment without using a surcharge. The procedure aims at meeting the target height, h_{emb} , within the allotted preloading time, where the excess material at the beginning is sufficient to compensate for the occurred settlement, s_x .



Source: Escher 2022

Available Data

A geotechnical site investigation provides data in terms of constant-rate-of-strain (CRS) oedometer tests from auger sampling taken from several different depths in the saturated clay. In addition, some parameters must be assessed from the literature, e.g. the unit weight of the embankment material, as it has not yet been constructed when the design is prepared.

Limit State function(s)

Defining residual settlement as $S_{res} = S_{\infty} - S(t_{preload})$, where S_{∞} is the long-term primary consolidation settlement and $S(t_{preload})$ is the occurred settlement at the end of the preloading time $t_{preload}$, a serviceability limit state can be formulated as:

$$G_s = s_{allow} - S_{res}$$

Equation 89.

where s_{allow} is the allowed residual primary consolidation settlement, which in this example is set to 5 cm. The settlement at t_{preload} is $S(t_{\text{preload}}) = U(t_{\text{preload}})S_{\infty}$, where $U(t_{\text{preload}})$ is the average degree of consolidation at time t_{preload} .

In this case, the designer is interested in finding the preloading time needed to satisfy the residual settlement criterion, i.e. determining how much preloading time is needed, so that the residual settlement after completion of the superstructure exceeds s_{allow} only with acceptable probability. Therefore, **Equation 89** is reformulated in terms of time:

$$G_t = t_{\text{preload}} - T_G$$

Equation 90.

where t_{preload} is the used preloading time, and T_G is the (inverse) function of S_{∞} and s_{allow} and is defined as the necessary preloading time to limit the residual settlement to $s_{\text{allow}} = 5$ cm. As the long-term S_{∞} is uncertain, T_G will also be uncertain. Setting a value to the target probability of limit state violation, $p_{f,T}$, the designer can however find the necessary t_{preload} by solving $P(G_t = t_{\text{preload}} - T_G < 0) = p_{f,T}$. The reliability-based analysis presented in the following aims at determining the distribution of T_G .

The effect of secondary consolidation settlement (creep) is not considered in this example.

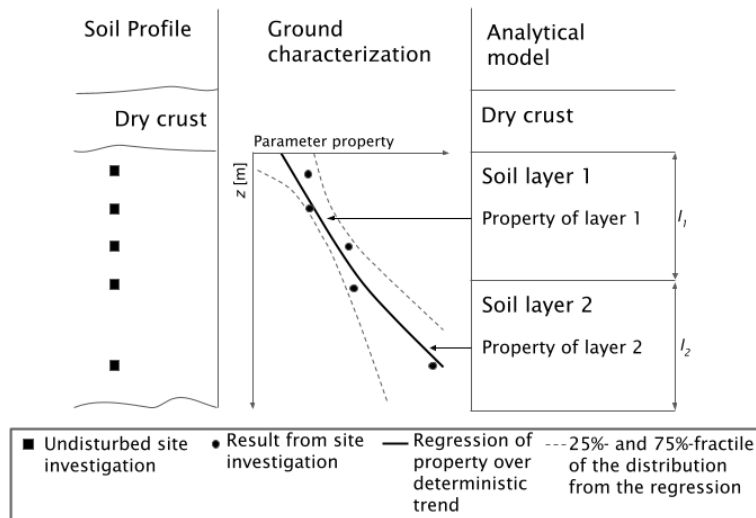
Target reliability

The target failure probability is set to $p_{f,T} = 5\%$ as this is a serviceability limit state, corresponding to a client's potential risk acceptance level for a violation of this limit state.

Ground Model

The ground is modelled as a 2D cross section of the soil under the embankment (**Figure 65**). The soil profile consists of very settlement susceptible soft clay, overlain by a stiffer dry crust. Under the clay, there is a stiff till. Based on 5 auger samples in the soft clay, the ground properties are characterized probabilistically (see section 2.1). The soft clay is divided into two separate 3 m thick layers in the settlement analysis, which for this is deemed sufficiently thick to assume that the settlement of each layer is a fully averaging process (i.e. the variance reduction factor is assumed $\Gamma^2 = 0$). The centre of each layer is located at depths $z = 3.1$ m and $z = 6.1$ m below the embankment foundation level. The deformation of the dry crust and till is assumed negligible.

Figure 65. Ground model and its use in the probabilistic ground characterization and the analytical settlement model used to determine S_{∞} .



Source: Escher 2022

Input Parameters (random and deterministic variables)

Table 34 shows the considered random variables which are characterized as described in section 2.1. The parameters of the limit states are functions of these random variables:

$$S_{\infty} = f(\gamma_{emb}, \gamma_{cl}, \sigma'_c, \sigma'_L, M_L, M', M_0) \text{ and } U = f(c_v)$$

Equation 91.

The geometries of the problem were treated as deterministic variables, as was the unit weight of the dry crust.

Table 34. Deterministic and random variables in the analysis. All random variables are assumed lognormal, except T_{M0} , which is assessed to be normally distributed.

Deterministic Variable	Unit	Description
b	18.5 m	Crest width of embankment
h_{emb}	1.8 m	Target embankment height
h_{clay}	6 m	Thickness of saturated soft clay
h_{crust}	1.5 m	Thickness of dry crust of clay
Random Variable	Unit	Description
γ_{cl}	kN/m ³	Unit weight of saturated clay
σ'_c	kPa	Preconsolidation pressure
σ'_L	kPa	Limit pressure
M_L	kPa	Modulus for $\sigma'_c < \sigma' \leq \sigma'_L$
M'	-	Modulus number $\sigma'_L < \sigma'$
τ_{fu}	kPa	Undrained shear strength
T_{M0}	-	Empirical transformation factor for M_0
γ_{emb}	kN/m ³	Unit weight of embankment
c_v	m ² /s	Coefficient of vertical consolidation
$M_0(*)$	kPa	Modulus for $\sigma' < \sigma'_c$

(*) M_0 is not a basic random variable but assessed from an empirical transformation (see section 2.3).

Uncertainty characterization

Geotechnical units and parameters

The uncertainty of the parameters $X = [\sigma'_c, M', M_L, \tau_{fu}, \gamma_{cl}]$ is assessed using the methodology developed by Müller et al. (2016), which is compatible with the procedure presented in this guideline. The methodology has previously been applied to a probabilistic embankment settlement analysis by Spross & Larsson (2021). An exponential trend line for the average value \bar{x} along depth z is determined for each parameter X_i (subscript i dropped for convenience):

$$\ln \bar{x} = \hat{a} + \hat{b}z$$

Equation 92.

where \hat{a} and \hat{b} are linear regression parameters evaluated from n data points which have been log-transformed in advance.

Assessing the total uncertainty, and assuming that, (i) $\Gamma = 0$ as the settlement process is assumed fully averaged within each separate clay layer and (ii) systematic measurement errors are small, the total uncertainty ($\sigma_{\ln X, tot}^2$) along the vertical trend (**Figure 65**) is reduced to the statistical uncertainty ($\sigma_{\ln X, st}^2$) only:

$$\sigma_{\ln X, tot}^2 \approx \sigma_{\ln X, st}^2 = \psi(n, z) \cdot \sigma_{\ln X, inh}^2$$

Equation 93.

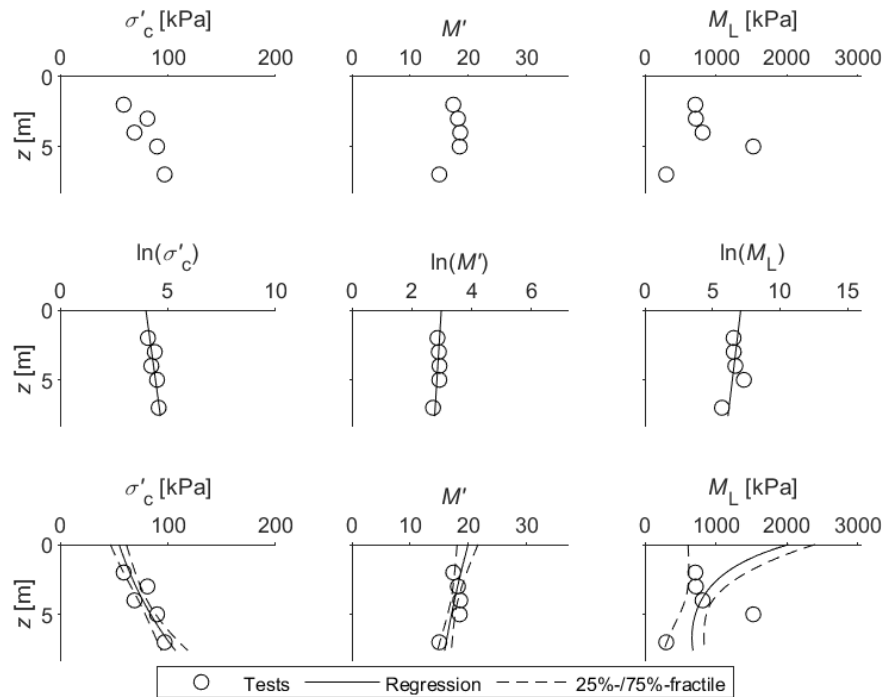
where $\sigma_{\ln X, inh}^2$ is the de-trended variance of the log-transformed data points, and $\psi(n, z)$ provides the statistical uncertainty along z :

$$\psi(n, z) = \frac{n-1}{n-3} \cdot \left[\frac{1}{n} \left(1 + \frac{n}{n-1} \cdot \frac{(z-\bar{z})^2}{s_z^2} \right) \right]$$

Equation 94.

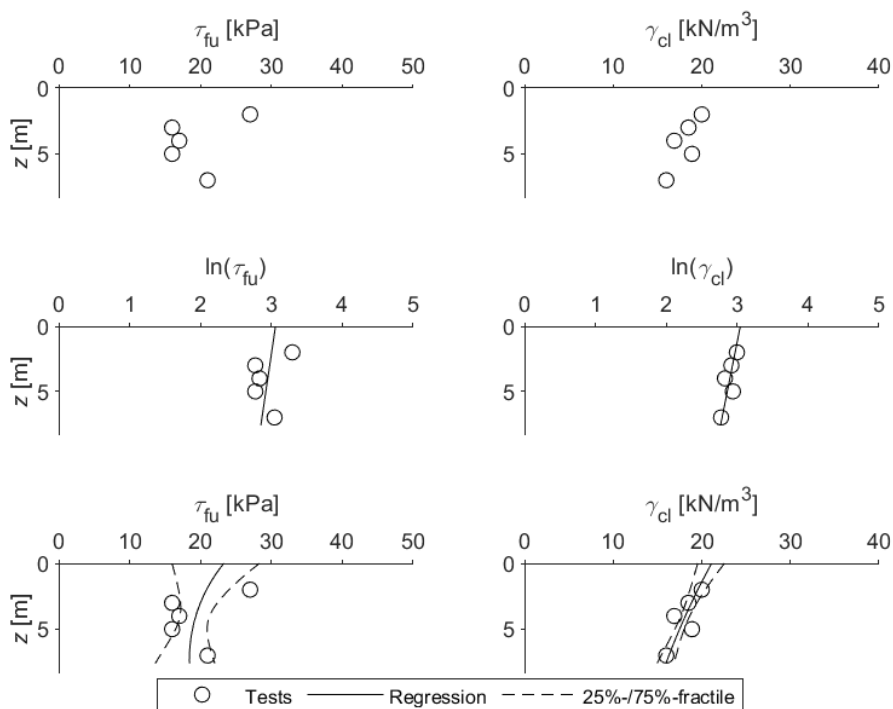
where \bar{z} is the sample mean of the depths of the measurements and s_z^2 is the sample variance of the depths. The $\ln \bar{x}$ and $\sigma_{\ln X, tot}^2$ are the parameters of the lognormal distribution, which straightforwardly can be transformed back to the physical space. The result of the uncertainty characterization along the depth is shown in **Figure 66** and **Figure 67**.

Figure 66. Top row: data points of preconsolidation pressure, modulus number, and modulus M_L . Middle row: evaluation of linear trend of log-transformed data points (**Equation 92**). Bottom row: exponential trend lines and their uncertainty after transformation back to the physical space.



Source: Escher 2022

Figure 67. Top row: data points of undrained shear strength and unit weight of clay. Middle row: evaluation of linear trend of log-transformed data points (**Equation 92**). Bottom row: exponential trend lines and their uncertainty after transformation back to the physical space.



Source: Escher 2022

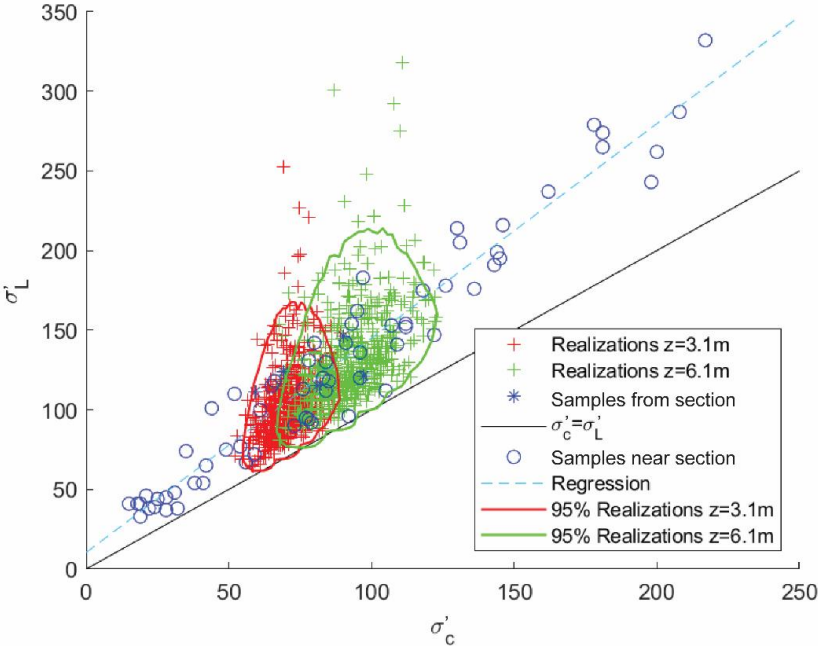
The limit pressure σ'_L must by definition be larger than σ'_c . To avoid violation of this restriction when sampling from the distributions, σ'_L was modelled by an additive random variable to the σ'_c , such that $\sigma'_L = \sigma'_c + \sigma'_\Delta$. The σ'_Δ was assessed from a larger number of CRS test samples taken in the vicinity of the evaluated cross section. The details of this model are presented in Escher (2022). Random samples generated using this model are shown in **Figure 68**.

The c_v was assumed to be log-normally distributed with mean $0.43 \cdot 10^{-8} \text{ m}^2/\text{s}$ and coefficient of variation of 50%. The mean was taken as the harmonic mean of the available test results and the coefficient of variation was assessed from the literature (see Escher (2022) for details on these assessments). It is assumed to not vary with depth.

The modulus M_0 was assessed based on an empirical transformation of the undrained shear strength: $M_0 = T_{M0} \tau_{ru}$, where the transformation factor T_{M0} was assumed to be normally distributed with mean 375 and coefficient of variation 20%, to account for the uncertainty in this transformation reported by Larsson & Sällfors (1986).

Additionally, the unit weight of the dry clay crust was assigned a constant value of 17 kN/m^3 (Larsson 2008).

Figure 68. Generation of σ'_c and σ'_L satisfying the restriction $\sigma'_c < \sigma'_L$.



Source: Escher 2022

Loads, groundwater and pore pressure

The load for the preloading is the weight of the road embankment, which is determined by the unit weight of the road embankment γ_{emb} . Note that this variable must be assessed from literature considering the planned construction material and compaction method, as the embankment is built after the design analyses have been completed. A log-normal distribution was assumed with a mean of 20.5 kN/m^3 and a coefficient of variation of 5%. Note also that the load to some extent depends on the compressibility of the clay: if the induced settlement is large, a large compensation in added material is required to meet the embankment target height at the end of the preloading

time (s_x in **Figure 64**); for such cases, an iterative approach may be needed to find a suitable initial embankment height, similar to that applied by Spross & Larsson (2021).

The groundwater level is assumed stable.

Model uncertainties

Based on Larsson (1986), the parameter M_0 was assessed through the empirical transformation $M_0 = T_{M0} \tau_{fu}$. In this case, T_{M0} was modelled as a normally distributed variable with mean = 375 and coefficient of variation = 20% to describe the uncertainty in the transformation. Thus, the M_0 parameter becomes a product of the lognormally distributed τ_{fu} and the normally distributed T_{M0} .

Reliability analysis

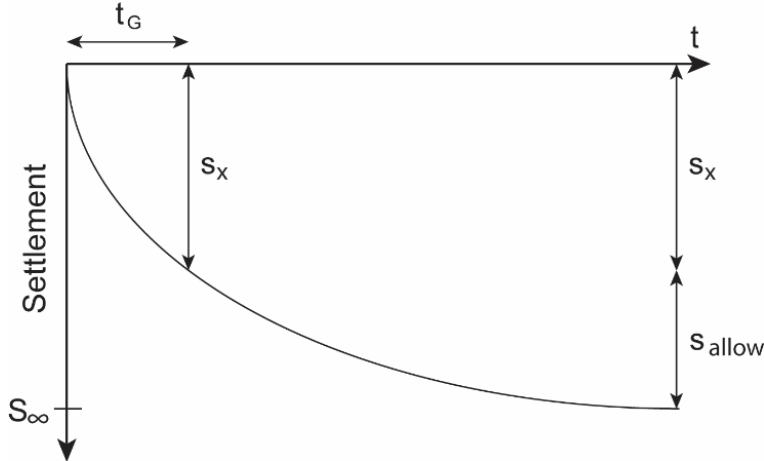
Selection of reliability method(s)

Crude Monte Carlo simulation was used, as the limit state is based on analytical equations which gives a short calculation time even for many evaluations of the limit state. Monte Carlo simulation can also easily generate random samples from the product $M_0 = T_{M0} \tau_{fu}$.

Estimation of probability of failure

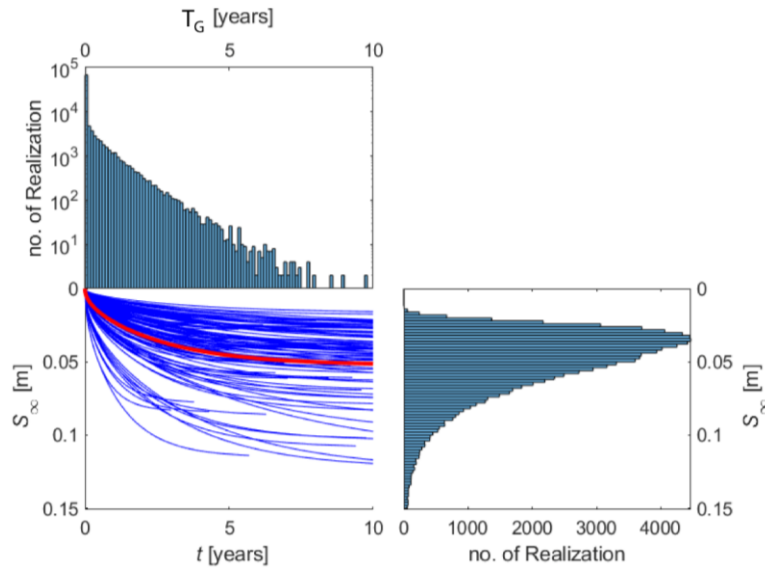
Generating 100.000 samples of the random variables, equally many time–settlement curves (“trajectories”) $s = U(t) \cdot s_\infty$ can be created, where the lower case letters for the settlement symbols indicate single trajectories (rather than random variables). For each trajectory, the needed settlement from preloading is calculated as $s_x = s_\infty - s_{allow}$. Inverting the relationship between settlement and time for s_x gives the time t_G for the trajectory (i.e. one outcome of TG in **Equation 90**). The procedure is illustrated for one trajectory in **Figure 69**. The result of the Monte Carlo simulation is shown in **Figure 70**.

Figure 69. Visualization of calculation procedure to find $t_{preload}$ for each generated settlement trajectory.



Source: Escher 2022

Figure 70. Monte Carlo simulation of settlement trajectories, histogram of their S_∞ , and required preloading time T_G to satisfy the allowable residual settlement.

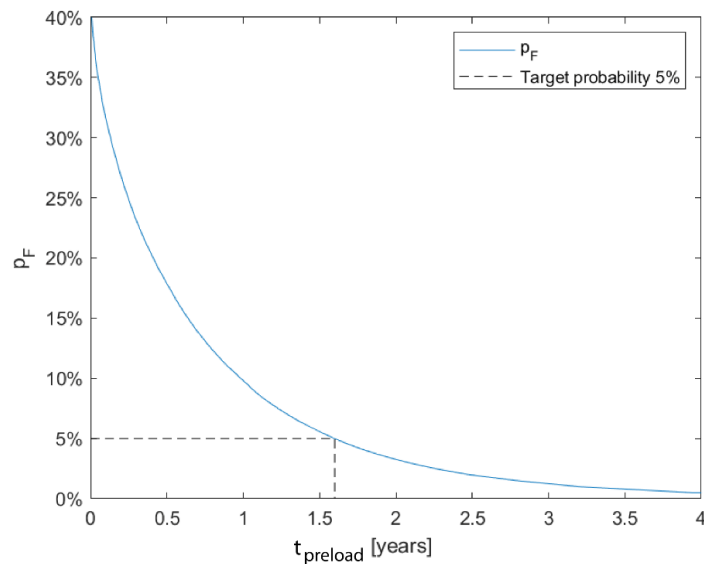


Source: Escher 2022

Interpretation of results and convergence

Having achieved 100 000 samples of T_G , the probability of violating the allowable residual settlement $s_{allow} = 5 \text{ cm}$ for different chosen preloading times $t_{preload}$ can be analysed. This is shown in **Figure 71**, from which it is clear that in this case, a preloading time of at least 1.6 years (580 days) is needed to satisfy the target failure probability $p_{f,T} = 5\%$.

Figure 71. Probability of violating the serviceability limit state (**Equation 90**) for different chosen $t_{preload}$.



Source: Escher 2022

Software Tools

A custom MATLAB code was used.

References

- Escher, K. (2022). Probabilistic settlement analysis for embankments using preloading without surcharge. MSc thesis. TRITA-ABE-MBT 22636. Stockholm: KTH Royal Institute of Technology. URL: <http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-317375>.
- Larsson, R. (1986). Consolidation of soft soils, Report 29. Swedish Geotechnical Institute, Linköping
- Larsson, R. (2008). Jords egenskaper [Properties of soil]. In Swedish. Swedish Geotechnical Institute, Linköping
- Larsson, R., Sällfors, G. (1986). Automatic continuous consolidation testing in Sweden. In: Consolidation of soils: testing and evaluation (eds R. Yong and F. Townsend). ASTM International, West Conshohocken, PA, USA, pp. 299–328.
- Müller, R., Larsson, S., Spross, J., (2016). Multivariate stability assessment during staged construction. Canadian Geotechnical Journal Vol. 53 No 4, pp. 603–618.
- Spross, J., Larsson, S. (2021). Probabilistic observational method for design of surcharges on vertical drains. Géotechnique Vol. 71 No 3, pp. 226–238.

B.3. Spread foundations

B.3.1 Vertically loaded spread foundation

Authors: Stéphane Commend

Reference: -

Problem definition

Description

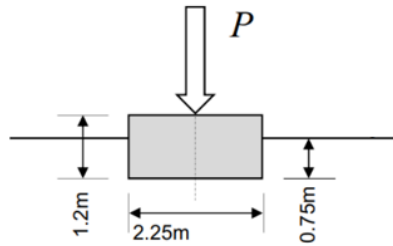
This spread foundation problem is inspired from the example “SAND 1” in *Probabilistic solutions for survey questions in “Are we overdesigning? – a survey of international practice”*, Oct. 14 2020, prepared by ISSMGE TC304 (TC304, 2020). In this study, a strip instead of a square footing (as in the TC304 example) is considered and a 2D FE analysis with ZSOIL conducted in order to show the benefits of a Bayesian analysis on the results of two reliability analyses: one for the bearing capacity of the strip footing, and the other one showing the probability of exceeding a given settlement.

Available Data

A sketch of the spread footing (see **Figure 72**

), as well as information on the soil’s friction angle ϕ and shear modulus G are provided in the TC 304 example.

Figure 72. Geometry and loading of strip footing.



Source: TC 304 2020

Limit State function(s)

The following limit state functions are considered:

$$g_1(X) = SF - 1 < 0$$

Equation 95.

where $SF = P_{ult}/P$, with P_{ult} the ultimate load computed by the FE model and $P = 1000$ kN/m the expected load on the footing.

$$g_2(X) = 4 \text{ mm} - u_y < 0$$

Equation 96.

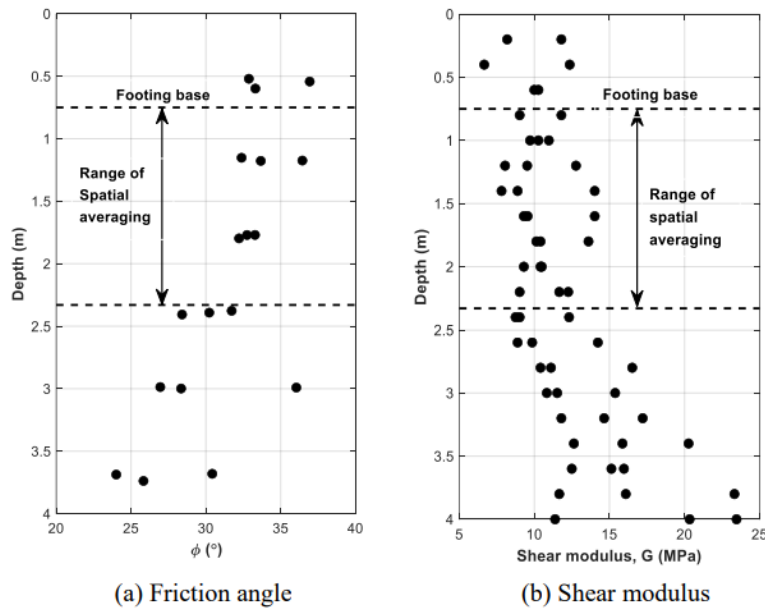
where u_y is the vertical settlement of the footing, computed by the FE model. A tolerable settlement of 4 mm is assumed in this example.

Target reliability

No target reliability is actually selected and/or determined.

Ground Model

Figure 73. Ground model - for friction angle (a) and shear modulus (b), taken from TC 304 (2020).



Source: TC 304 2020

Input Parameters (random and deterministic variables)

Table 35. Overview of deterministic and random variables within this example.

Deterministic Variable	Unit	Description
P	kN/m	Load on the strip footing
Random Variable	Unit	Description
E	kN/m ²	Sand's elastic modulus
φ	°	Sand's friction angle

Uncertainty characterization

Geotechnical units and parameters

- E: lognormal, mean = 29.750 kN/m², stdev = 6.190 kN/m²
- φ: lognormal, mean = 31.43°, stdev = 2.77°

The mean values and standard deviations are derived from the ground model described in TC 304 (2020).

Loads, groundwater and pore pressure

In this example, the load on the footing is considered to be deterministic, and equal to 1000 kN/m. There is no water table to be considered.

Model uncertainties

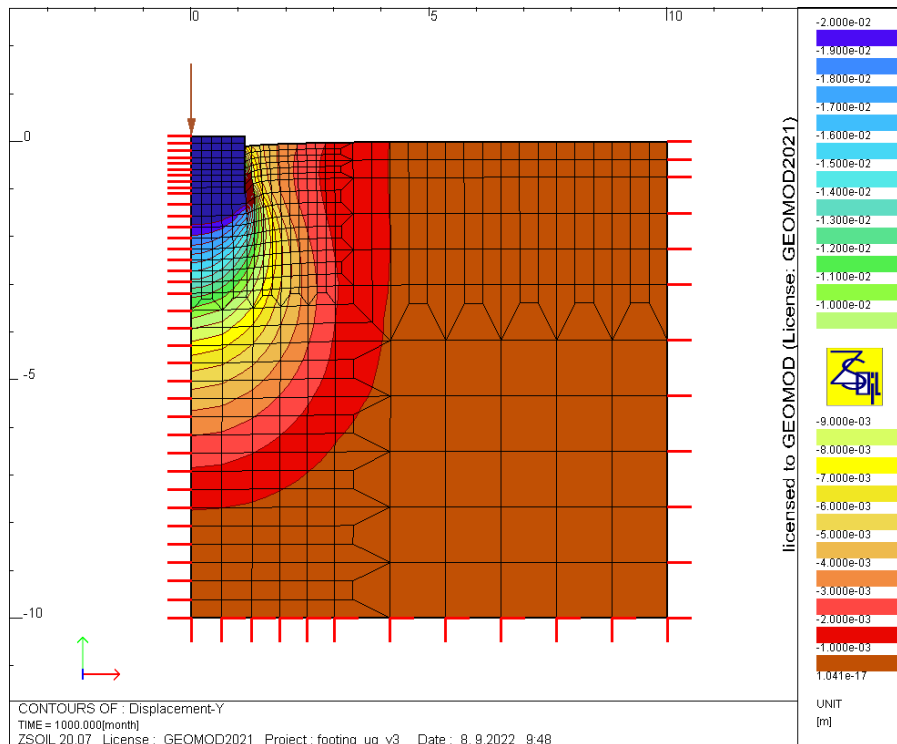
As described in TC 304 (2020), a multiplicative error is taken into account: $\epsilon = N(1, 0.28)$, where N indicates a Gaussian distribution.

Reliability analysis

Selection of reliability method(s)

A 2D FE model is constructed with ZSOIL (settlement results are shown in **Figure 74**). Polynomial Chaos Expansions (PCE, see Annex O) surrogates are built on 100 realizations of ZSOIL and then, on the PCE, Monte Carlo simulations are used to perform the reliability analyses. Note that, in order to select the input values (E_i , ϕ_i) for the 100 ZSOIL realizations, a Latin Hypercube Sampling (LHS) is used.

Figure 74. 2D mesh (with a vertical axis of symmetry on the left) and settlements for one of the ZSOIL realizations.



Source: Authors' own work

Estimation of probability of failure

Prior:

$$P_{f1}(SF < 1.0) = 0.9\%$$

Equation 97.

$$P_{f2}(u_y > 4.0 \text{ mm}) = 4.1\%$$

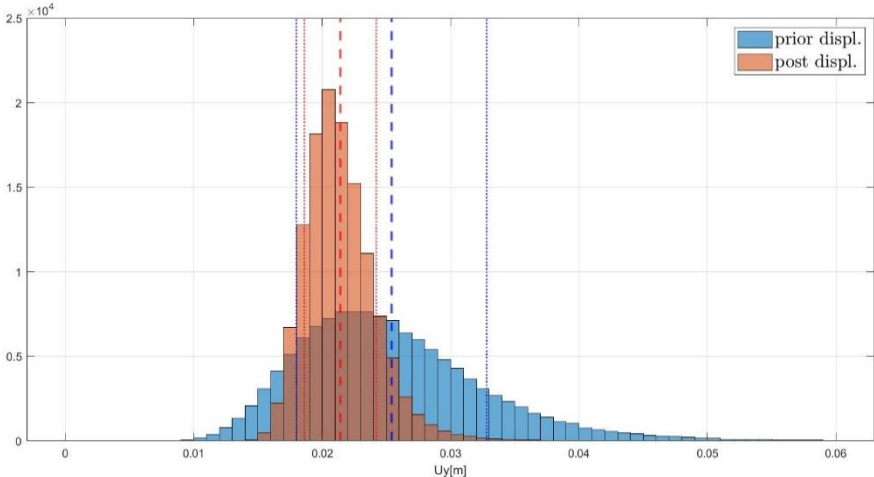
Equation 98.

Posterior:

Say now we have a new information on a building located close to the future one (hypothesis: same type of foundation, same geotechnical conditions): settlement u_y is measured = 15 mm for $P = 800$

kN/m. This new information allows us to update the inputs' probability density functions, as well as reliability analyses, both on safety factor $SF = P_{ult} / P$ and vertical settlement u_y (see **Figure 75**). This update is performed using Bayesian inverse analysis, as explained in the guidelines, chapter 7.

Figure 75. Prior and Posterior distributions of vertical settlement.



Source: Authors' own work

As a result, there is a tremendous change on P_{f2} , as expected:

$$P_{f1}(SF < 1.0) = 0.6\%$$

Equation 99.

$$P_{f2}(u_y > 4.0 \text{ mm}) = 0.06\%$$

Equation 100.

Interpretation of results and convergence

This example shows the potential of Bayesian inverse analysis on updating the probabilities of failure.

Software Tools

For the finite element calculations, ZSOIL (<https://www.zsoil.com/>) and for the uncertainty quantification, sensitivity and reliability analyses, UQLab (<https://www.uqlab.com/>) was used.

References

TC 304 (2020). Probabilistic solutions for survey questions in “Are we overdesigning? – a survey of international practice”, Oct. 14 2020, prepared by TC304.

B.4. Piled foundations

B.4.1 Comparison of reliability levels achieved by the pile design method of EN 1997-3 with the reliability targets of EN 1990

Authors: Antonis Mavritsakis

Reference: Hehenkamp, M. (2022). Reliability Analysis of Foundation Pile Designs from Eurocode 7. Master Thesis, TU Delft

Problem definition

Description

This pile design example investigates the comparison of the reliability levels achieved upon designing with the method suggested by EN 1997-3 and the reliability targets indicated by EN 1990-1. The analysis involves the utilization of real CPT data collected at a site in Amsterdam. The pile design depends on the Koppejan model for assessing the resistance of a pile based on CPT data (van Mierlo, J., & Koppejan, A., 1952). A sample of Koppejan parameters is derived by assessing the parameters per CPT profile. This sample allows for the generation of probability density functions (PDFs) for the Koppejan parameters and the derivation of their representative values according to EN 1997-1. These representative values are used to optimize the pile design. Following, the PDFs of the Koppejan parameters are employed in a probabilistic analysis for the designed pile, for the purpose of evaluating its reliability. The example performs the pile design and reliability evaluation for different numbers of CPTs in the site for the purpose of investigating the effect of additional site investigation data.

Note: The work performed in this example is based on a draft version of EN 1997-1 and EN 1997-3 valid December 2021.

Available Data

The example is based on CPT data from a site investigation campaign performed in Amsterdam, consisting of 36 CPTs. The dataset was provided by Deltares (2020).

Limit State function(s)

The limit state function (Z) for pile design is defined by **Equation 101**. Specifically, failure is met when the load applied at the top of the pile (F) surpasses the total resistance (R) of the pile, which can be decomposed into the base resistance (R_b) and the shaft resistance (R_s). When $Z \geq 0$ the pile does not fail.

$$Z = R_b + R_s - F$$

Equation 101.

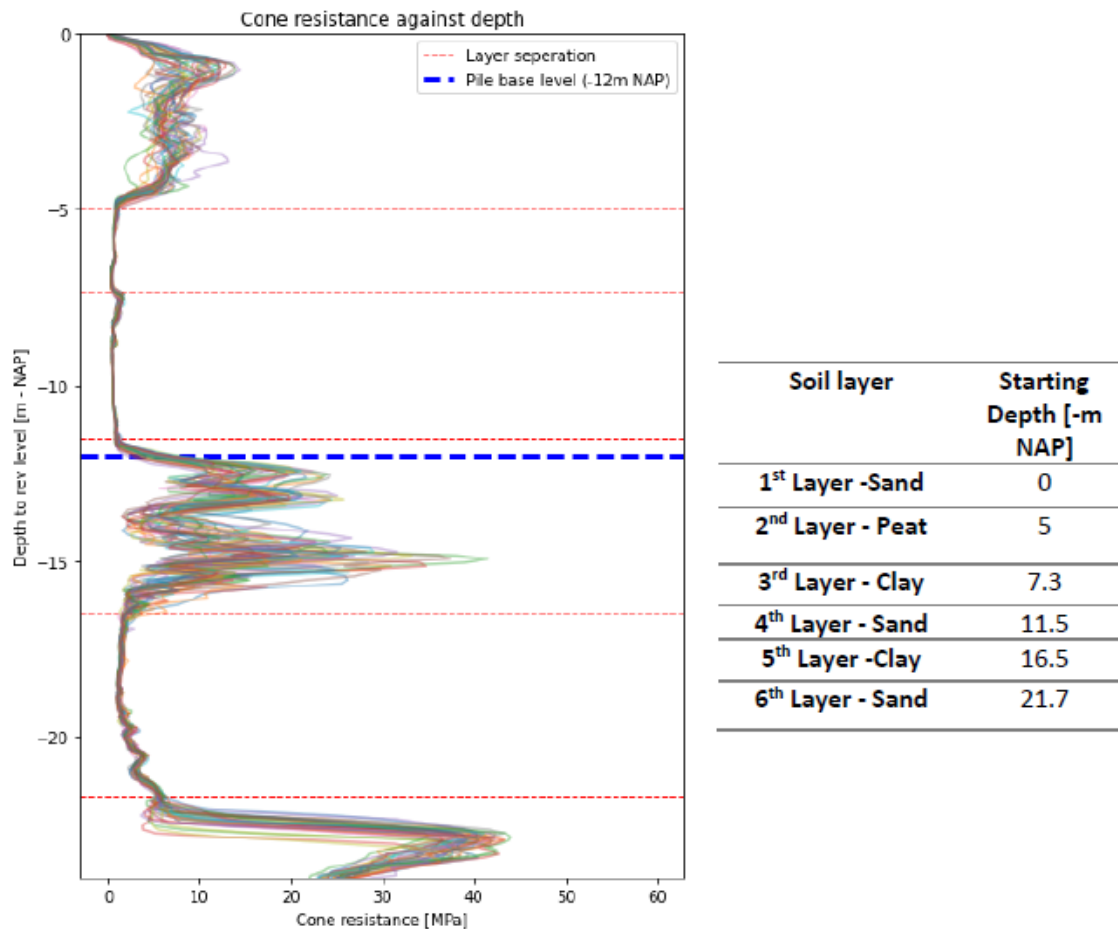
Target reliability

As specified by EN 1990-1, the target failure probability for the design is $P_{f,T} = 7.2 \cdot 10^{-5}$ in a lifespan of 50 years, which is associated to a target reliability index of $\beta_T = 3.80$.

Ground Model

The example uses a 1D ground model, assumed at the location where the designed pile will be installed. The identification of soil layers is based on the interpretation of the CPT data. Four of the recognised layers are useful for the analysis (from top to bottom): Holocene sand, peat, clay and Pleistocene sand (**Figure 76**). The latter acts as the foundation for the pile base. The layers are parametrized by the Koppejan parameters, which are considered stochastic and whose distribution varies according to the examined scenario of CPT sample size.

Figure 76. Plot of cone resistance per CPT over depth and soil stratigraphy.



Source: Hehenkamp 2022

Input Parameters (random and deterministic variables)

Table 36 shows the parameters used in the example. The pile dimensions are deterministic parameters of the example and are derived by pile designing according to EN 1997-3. The pile has a rectangular cross section. Pile load and resistance, and by extension the soil strength parameters, are considered as stochastic. All soil layers are parameterized by the shaft friction (q_s). The Pleistocene sand layer, where the piles are founded, is also parametrised by the Koppejan resistance (q_k).

Table 36. Input parameters.

Deterministic Variable	Unit	Description
L	m	Pile length
d	m	Pile width (rectangular cross section)
n	-	Number of CPT observations used in each scenario
Random Variable	Unit	Description
q_s	kPa	Shaft friction - $\sim N(\widehat{\mu}_{q_s}, \widehat{\sigma}_{q_s})$, or $\sim T(\widehat{\mu}_{q_s}, \widehat{\sigma}_{q_s})$
q_k	kPa	Koppejan resistance at base - $\sim N(\widehat{\mu}_{q_k}, \widehat{\sigma}_{q_k})$, or $\sim T(\widehat{\mu}_{q_k}, \widehat{\sigma}_{q_k})$,
m	-	Model factor $\sim N(1, 0.1)$
G	MN	Permanent pile load
Q	MN	Variable pile load

Uncertainty characterization

Geotechnical units and parameters

The uncertainty of the pile resistance R is directly connected to the resistance of each components R_b and R_s , and so, the uncertainty of the soil strength parameters. At this point, three different probabilistic models are composed for the soil strength. In the baseline model (model X), soil parameter distributions are Normal and the best estimates of their parameters are derived by using all CPT observations (**Equation 102**). In the second model (model A), the soil parameter distributions are Student T, in order to account for the sample size, and best estimates of their parameters are derived according to the specific CPT sample used in each examined scenario (**Equation 103**). The third model (model B) takes a Bayesian perspective into the derivation of parameter distributions. q_s and q_k are still normally distributed, but the parameters of the normal distributions are also described on a probabilistic level (**Equation 104**). Specifically, a prior distribution is assigned to the mean and standard deviations of q_k and q_s (**Equation 105** and **Equation 106**). The posterior distributions are derived by Bayesian updating with the conjugate prior model using the CPT observations considered in each scenario.

$$q_s \sim N(\widehat{\mu}_{q_s}, \widehat{\sigma}_{q_s}) \text{ and } q_k \sim N(\widehat{\mu}_{q_k}, \widehat{\sigma}_{q_k})$$

Equation 102. Model X

$$q_s \sim T(\widehat{\mu}_{q_s}, \widehat{\sigma}_{q_s}) \text{ and } q_k \sim T(\widehat{\mu}_{q_k}, \widehat{\sigma}_{q_k})$$

Equation 103. Model A

$$q_s \sim N(\mu_{q_s}, \sigma_{q_s}) \text{ and } q_k \sim T(\mu_{q_k}, \sigma_{q_k})$$

Equation 104. Model B

$$\mu_{q_s} \sim N(5, 1) \text{ and } \mu_{q_k} \sim N(5, 1)$$

Equation 105. Model B μ priors

$$\sigma_{q_s} \sim \text{Gamma}(1, 1) \text{ and } \sigma_{q_k} \sim \text{Gamma}(1, 1)$$

Equation 106. Model B σ priors

Loads, groundwater and pore pressure

The pile load (F) is composed of the permanent load (G) and the variable load (Q) (**Equation 107**). G is normally distributed, while Q follows a Gumbel distribution.

According to EN 1990-1, the ratio of permanent to variable representative load (χ) is estimated. With the assumption of EN 1990-1, the representative values for the permanent and variable load with a known design load ($F_d = 0.56$ MN) can be estimated by **Equation 109** and **Equation 110**. The associated partial safety factors are taken from EN 1990-1.

$$F = G + Q$$

Equation 107.

$$\chi = \frac{Q_k}{G_k + Q_k} = \frac{1}{3}$$

Equation 108.

$$G_{rep} = \frac{F_d}{\gamma_G + \frac{\chi}{1-\chi} \cdot \gamma_Q}$$

Equation 109.

$$Q_{rep} = G_k \cdot \frac{\chi}{1-\chi}$$

Equation 110.

Given the representative values of the permanent and variable loads, their distribution can be parameterized. **Equation 111** and **Equation 112** derive the mean and standard deviation of the permanent load. The representative value of the variable load is the 2% quantile and its CoV takes a value of 0.3. Given this information, the parameters of the Gumbel distribution are estimated.

$$P(G > G_k) = 0.5 \rightarrow \mu_G = G_k = 0.333 \text{ MPa}$$

Equation 111.

$$CoV_G = 0.05 \rightarrow \sigma_G = 0.016 \text{ MPa}$$

Equation 112.

The eventual distributions of G and Q are given by **Equation 113** and **Equation 114**.

$$G \sim N(0.333, 0.016)$$

Equation 113.

$$Q \sim \text{Gumbel}(0.937, 0.028)$$

Equation 114.

Model uncertainties

The pile resistance model includes a multiplicative model factor (m), which follows a normal distribution (**Equation 115**) (Deltares, 2020).

$$m \sim N(1, 0.1)$$

Equation 115.

Reliability analysis

Selection of reliability method(s)

The reliability of the pile design is assessed using the First Order Reliability Method (FORM).

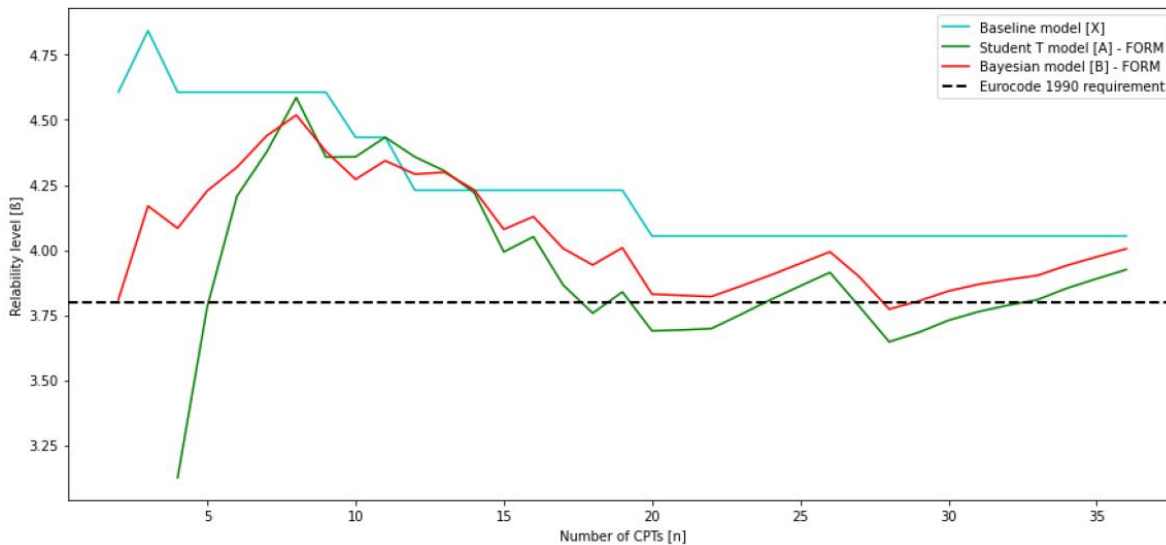
Estimation of probability of failure

The example analyses multiple scenarios, in each of which a CPT sample of a different size is used. Per scenario, the CPT sample is used to assess the parameters of the three probabilistic models for soil strength (Models X, A and B). Then, the pile is designed according to EN 1997-3 in a deterministic fashion using design values. It is mentioned that the design values are a function of the CPT sample size, and hence the pile dimensioning can vary between scenarios. The derived pile dimensioning is then used in probabilistic calculations per probabilistic soil strength model. This leads to the evaluation of the reliability index for each probabilistic model. Lastly, the baseline model (Model X) provides the same parameter distributions for all scenarios, since it is always based on all 36 CPTs performed at the site. However, its reliability results can vary per scenario, due to the different pile design resulting according to EN 1997-3.

Interpretation of results and convergence

Figure 77 presents the results of the analysis. The Student T and Bayesian models show good agreement and their reliability indices converge as the CPT sample grows. Exceptions can be spotted due to some localized fluctuations of the reliability index. Additionally, the baseline model perceives the design as more reliable than the other two probabilistic models for all sample sizes except from the range of $10 < n < 14$. The difference between the baseline model (Model X) and the other probabilistic quantification models (Models A and B) can be explained by local trends in the CPT data (see Appendix B-6 in Hehenkamp (2022)). Lastly, it is evident that all probabilistic models estimate that the reliability index of the pile design is greater than the target of $\beta_T = 3.8$ according to EN 1990-1.

Figure 77. Estimated reliability index per probabilistic model for different sizes of the CPT sample and comparison to the EN 1990 requirement.



Source: Hehenkamp 2022

Software Tools

The Probabilistic Toolkit of Deltares (Brinkmann, 2021) was the main software employed for this analysis. Also, for the work of Hehenkamp (2022), a Python script for adjusting the input to the Probabilistic Toolkit, as well as a Python script for statistical inference on the CPT data have also been developed.

References

- Brinkmann, R. (2021). Probabilistic toolkit- User Manual (version 2021). Deltares, Delft. URL: <https://www.deltares.nl/en/software/probabilistic-toolkit-ptk/>
- EN 1990-1:2023. Eurocode: Basis of structural and geotechnical design – Part 1: Design of new structures. (Draft Version) CEN/TC 250/SC 10.
- EN 1997-1:2023. Eurocode 7: Geotechnical design – Part 1: General rules. (Draft for Formal Vote) CEN/TC 250/SC 7/WG 1 “Evolution of 1997 series”.
- EN 1997-3:2024. Eurocode 7: Geotechnical design – Part 3: Geotechnical structures. CEN/TC 250/SC 10.
- Deltares (2020). Advies geotechnische draagkracht houten palen Amsterdam. Deltares, Delft/Amsterdam.
- Hehenkamp, M. (2022). Reliability Analysis of Foundation Pile Designs from Eurocode 7. Masterthesis TU Delft.
- van Mierlo, W.C., & Koppejan, A. (1952). Lengte en draagvermogen van heipalen. Bouw: Centraal weekblad voor het bouwwezen, 19-1-1952, no. 3, 1–11 (in Dutch), 1-11

B.4.2 Reliability updating of driven pile using pile tests and Bayesian inference

Authors: P. Arnold , A. Mavritsakis, S. Wilhelm

Reference: Arnold P., S. Wilhelm, B. Schädlich und T. Schweckendiek (2022). Zur Anwendung zuverlässigkeitsbasierter Methoden in der Bemessung von Pfahlgründungen auf Grundlage von Pfahlprobelastungen (in German). 28. Darmstädter Geotechnik-Kolloquium, S. 171-188.

Problem definition

Description

The following example demonstrates the application of reliability-based methods in combination with Bayesian inference to derive the load-bearing capacity of axially loaded driven piles.

Available Data

The data available for this example are results of a site investigation given by five Cone Penetration Tests (CPTs) and results from pile load tests.

Limit State function(s)

According to EN 1997-3, the representative pile resistance R_{rep} is:

$$R_{rep} = R_{s,rep} + R_{b,rep} = \sum_i A_{s,i} \cdot q_{s,i,rep} + A_b \cdot q_{b,rep}$$

Equation 116.

where $R_{s,rep}$ and $R_{b,rep}$ are the representative values of pile shaft and pile base resistance, $q_{s,i,rep}$ is the representative value of pile shaft friction in layer i , $q_{b,rep}$ is the representative value of the pile end bearing pressure, $A_{s,i}$ is the pile shaft area in layer i and A_b is the pile base area. The limit state function is $Z = R_s + R_b - N$ with N being the axial load.

Target reliability

EN 1990-1 defines different target reliabilities depending on the reference period for three reliability classes in relation to three failure consequence classes (also damage consequence classes). The partial safety factors specified in EN 1990-1 refer to reliability class RC 2, for which a reliability or the corresponding probability of failure, respectively, for a reference period of 1 and 50 years is specified ($P_f^1 \approx 10^{-6}$ and $P_f^{50} \approx 10^{-4}$).

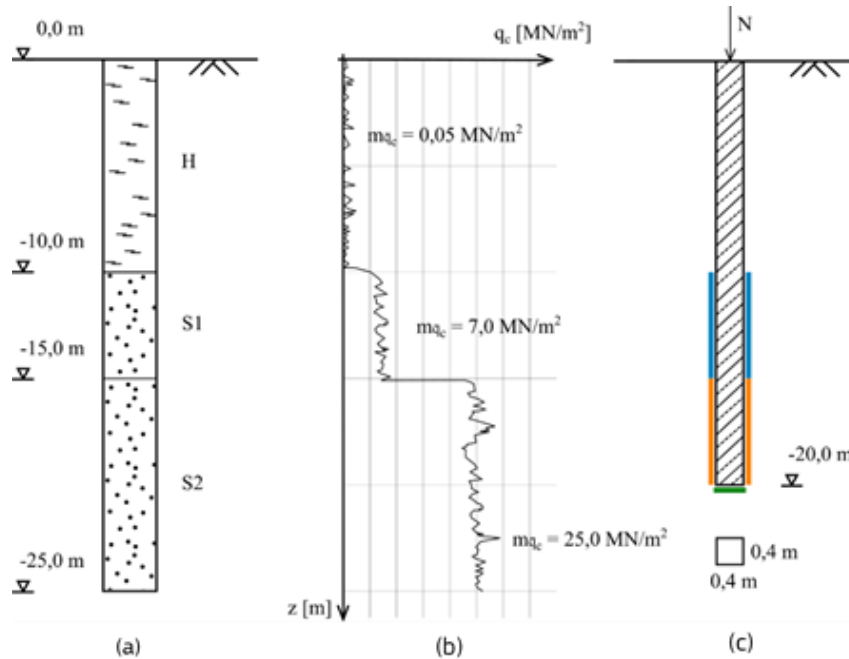
Ground Model

Figure 78 shows the design profile and the results of a cone penetration test, which is assumed to be representative for the construction site. Reinforced concrete piles ($L = 20$ m, $a = b = 0.4$ m) are to bridge the soft layer near the surface.

Five CPTs ($n = 5$) act as the basis for the design, which have been carried out on the construction site. For the sake of simplification, the heterogeneity and depth dependency of the tip resistance in the sand layers and the position of the CPTs relative to the pile location are neglected in the present example, i.e. a value \bar{q}_c averaged over the layer height describes the layer. **Table 37** shows the mean values $m_{\bar{q}_c}$ and the standard deviations $s_{\bar{q}_c}$ for the sand layers S1 and S2. For simplicity,

neither the strength nor negative skin friction effects of the peat layer are accounted for in this example.

Figure 78. Example set-up: (a) soil profile, (b) CPT-tip resistance profile, (c) bearing-capacity design.



Source: Arnold et al. 2022

Input Parameters (random and deterministic variables)

In this example, the uncertainty in the estimation of the pile bearing capacity (a) results in the variation of the soil parameters, represented by the CPT tip resistance $X_p = \{\bar{q}_c\}^T$ and (b) in the transformation model of the EA-Pfähle (2013) $X_M = \{m_{qs}; m_{qb}\}^T$.

Uncertainty characterization

Geotechnical units and parameters

Pile resistance from experience values.

On the right side of **Table 37**, the representative values of the pile shaft friction $q_{s,rep}$ and end bearing pressure $q_{b,rep}$ are derived based on the empirical relationship provided by the EA-Pfähle (2013). Since there is only a limited number of samples, the representative value of the tip resistance is determined for a one-sided 95% confidence interval based on EN 1997-1; $q_{c,rep} = m_{\bar{q}_c} - t_{n-1}(0,95) \cdot s_{\bar{q}_c} / \sqrt{n}$ where $t_{n-1}(0,95) = 2,132$ the 95th % quantile of the t-distribution for $n - 1 = 4$.

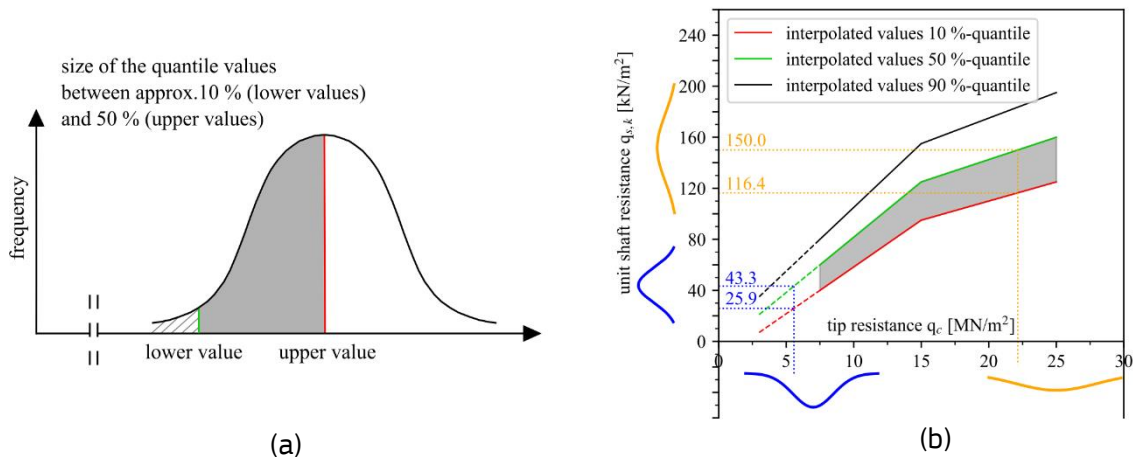
Table 37. Summary of input parameters for the example.

Layer	Depth	Point statistics of the CPT tip resistance			Derivation of pile shaft and base resistance based on the EA-Pfähle (2012) for $s/D_{eq} = 0,01$				
		$m_{\bar{q}_c}$	$s_{\bar{q}_c}$	$V_{\bar{q}_c}$	$q_{c,rep}$ [MN/m ²]	$q_{s,rep}$ [kN/m ²]		$q_{b,rep}$ [kN/m ²]	
		[MN/m ²]	[MN/m ²]	[-]		10%	50%	10%	50%
Peat	0 – 10 m	-	-	-	-	-	-	-	-
Sand S1	10 – 15 m	7.0	1.5	0.215	5.57	25.9	43.3	-	-
Sand S2	15 – 25 m	25.0	3.0	0.120	22.14	116.4	150	8 421	11 128

The empirical values provided by the EA-Pfähle (2013) are based on the results of many pile load tests and given in a range, where the lower and upper limit values are the 10% and 50% quantile of the distribution of the pile shaft friction and the end bearing pressure (see **Figure 79** (a) and Kempfert and Becker 2007). The derivation of empirical values for $q_{c,rep} < 7.5$ MN/m² is subject to limited applicability, but should be assumed here as a linear extrapolation for simplification.

Figure 79 (b) shows that the uncertainty in the estimated pile resistances is a product of the uncertainty in the EA-Pfähle transformation model and the uncertainty in the subsoil (represented here by the influence of the q_c -variation on the shaft friction, i.e. by the blue and yellow density functions corresponding with **Figure 78** and values in **Table 37**).

Figure 79. Quantile ranges of the density function for pile resistances from empirical values according to EA-Pfähle (2013): (a) based on the EA-Pfähle (2013), (b) representation of the variance of the pile skin friction $q_{s,rep}$ with q_c variance of the layers S1 and S2.



Source: EA-Pfähle 2013

Pile resistance from testing

Testing of installed piles aims to assess whether the piles can sustain the testing load. Naturally, since the resistance of all piles within a site are inter-dependent to some extent, the testing results at some piles can be used to update the resistance distribution of the all piles, and hence their probability of failure. To that end, Bayesian statistics can be used.

In order to check the influence of results from the pile load tests on the probability of failure, two cases are examined (see **Table 38**). In case A, two static pile load tests are carried out up to the defined failure criterion (e.g. $0.1 \cdot D_{eq}$). Accordingly, the bearing capacity of the tested piles is known $R = P_p$ (equality information). In case B it is assumed that four structural piles are tested in the SLS after production (e.g. to determine/confirm the load-deformation curve). Thus, in case B, the exact bearing capacity of the tested pile is not known after the test, but it is known that it is higher than the applied test load $R \geq P_p$ (inequality information).

The influence of the measurement error should also be determined for each independent static load test. For this it is assumed that the measurement error follows a normal distribution with $\mathcal{N}(0, \sigma_{meas})$.

Table 38. Summary of pile loading test.

Case	Test type	Tests n_s	Loading	σ_{meas} [MN]	Test loads P_p [MN]
A	static, test pile	2	ULS	[0; 0.05; 0.1]	4.15; 4.45
B	static, construction pile	4	SLS	[0; 0.05; 0.1]	1.70; 1.87; 2.04; 2.21

Loads, groundwater and pore pressure

The representative persistent axial static pile load is $N_{rep} = 1.7$ MN. Variable loads are not applied. The uncertainty in the permanent load is assumed to be normally distributed with a coefficient of variation of 5%, i.e. $f_N \sim \mathcal{N}(\mu_N; \sigma_N) = \mathcal{N}(1,7; 0,085)$. A permanent design situation is assumed. Thus, the design value of the actions is $N_d = \gamma_G \cdot N_{rep} = 1.35 \cdot 1.7$ MN = 2.3 MN.

Model uncertainties

The model factors for considering the uncertainty in the parameter transformation can be described as the deviation of a sample $q_s(x_i)$ and $q_b(x_i)$ relative to the mean of the distribution $\bar{q}_s = q_s(\bar{X}_p)$ and $\bar{q}_b = q_b(\bar{X}_p)$ as follows:

$$m_{qs} = q_s - \bar{q}_s \text{ and } m_{qb} = q_b - \bar{q}_b$$

Equation 117.

The pile resistance is then:

$$R = R_s(X_p; X_M) + R_b(X_p; X_M)$$

$$= \sum_i A_{s,i} (q_{s,i}(X_p) + m_{qs}) + A_b (q_b(X_p) + m_{qb})$$

Equation 118.

Since all input variables follow a normal distribution and the derivation of the pile bearing capacity is a linear function, the resulting pile resistance also follows a normal distribution $R \sim \mathcal{N}(\mu_R; \sigma_R)$ with $\mu_R = 3\,485$ kN and $\sigma_R = 478$ kN.

In a theoretically perfect transformation model, $m_{qs} = 0$ and $m_{qb} = 0$. Accordingly, the uncertainty in the pile resistance results solely from the inaccuracy in the description of the subsoil, described by

$X_p = \{\bar{q}_c^{S1}; \bar{q}_c^{S2}\}$ and $f_{\bar{q}_c} \sim \mathcal{N}(\mu_{\bar{q}_c}; \sigma_{\bar{q}_c})$, to $R(X_p) \sim \mathcal{N}(\mu_{R_p}; \sigma_{R_p})$ with $\mu_{R_p} = \mu_R = 3485 \text{ kN}$ and $\sigma_{R_p} = 387 \text{ kN}$. According to the simplifying assumptions made for this example in Section 1, this uncertainty is constant over the construction site.

Since all variables are normally distributed in this simplified example, the uncertainty in the transformation model can be derived directly:

$$\sigma_{R_m} = \sqrt{\sigma_R^2 - \sigma_{R_p}^2} = 280 \text{ kN}$$

Equation 119.

Reliability analysis

Selection of reliability method(s)

Within this example Bayesian inference is used to update the reliability of driven monopiles by use of closed analytical formulations as well as numerical modelling using Monte Carlo simulations.

Estimation of probability of failure

Pile design based on partial factor approach.

For this example, the deterministic design value of the pile resistance based on the experience values in EA-Pfähle (2013) is:

$$\begin{aligned} R_d^{EAP} &= \frac{R_{rep}^{EAP}}{\gamma_F} = \frac{\sum R_{s,rep}^{EAP} + R_{b,rep}^{EAP}}{\gamma_F} = \frac{1\,138 \text{ kN} + 1\,347 \text{ kN}}{1.4} = \frac{2\,485 \text{ kN}}{1.4} \\ &= 1\,775 \text{ kN} \end{aligned}$$

Equation 120.

where, in this example, the pile shaft and base resistances were derived at the 10%-quantile according to **Table 37**.

The verification of the external load-bearing capacity cannot be provided as $N_d/R_{d,EAP} = 1.296 \notin 1,0$. The minimum pile length required to fulfil the design criteria is 13.94 m.

Reliability based pile design.

Using the experience values from the EA-Pfähle (2013) and assuming that all soil parameters and pile resistance components are uncorrelated, the probability of failure (using a first-order second moment approximation) is:

$$\begin{aligned} P_f &= P(R - N < 0) = \Phi\left(-\frac{\mu_R - \mu_N}{\sqrt{\sigma_R^2 + \sigma_N^2}}\right) \\ &= \Phi(-3.677) = 1.17 \cdot 10^{-4} \end{aligned}$$

Equation 121.

The probability of failure does not state that out of 10,000 piles produced on the construction site, approx. 1.17 piles will fail, but that if 10,000 construction sites are generated under these

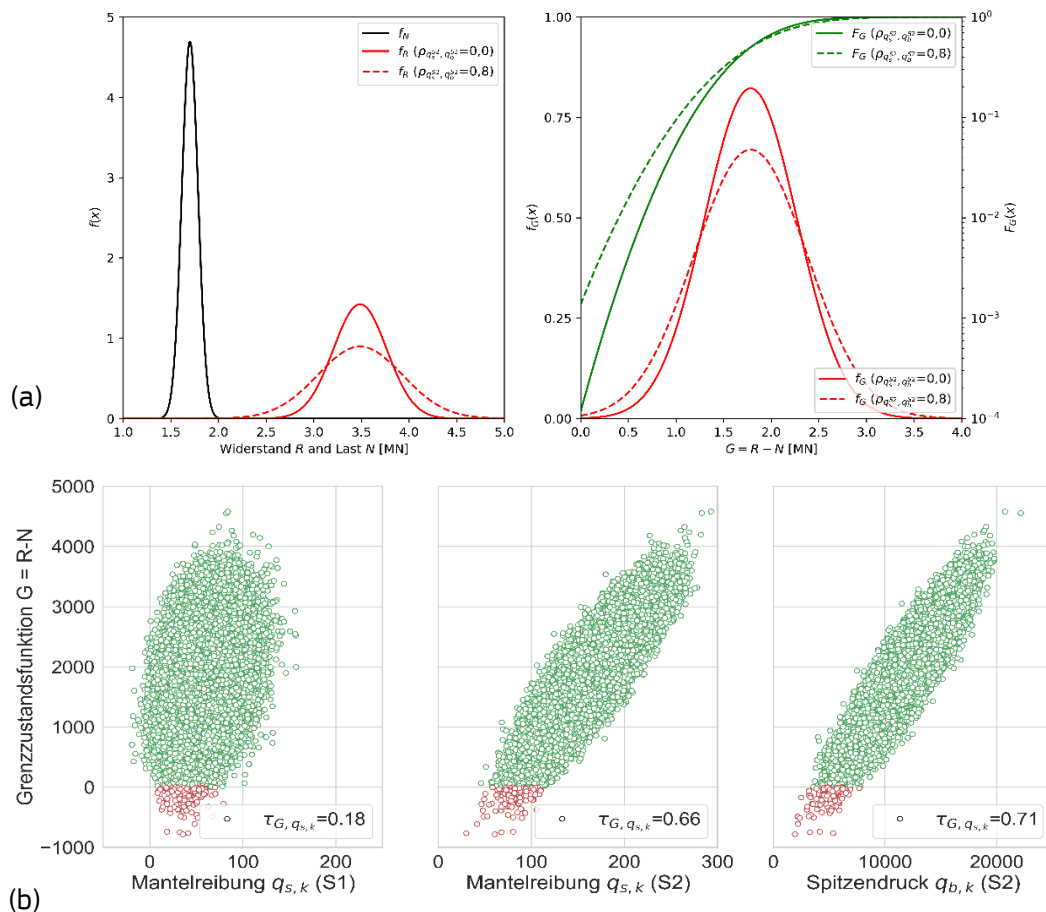
assumptions (soil layers homogeneous, pile resistances vary according to **Table 37** etc.), in approx. 1.17 cases the pile design criterion is not met.

However, it can be assumed that in situ a higher pile shaft resistance in the S2 layer is associated with a higher CPT tip resistance in the same layer. Accounting for this and assuming a positive correlation coefficient of $\rho_{q_{s,rep}, q_{b,rep}} = 0.8$ (simplified in the present case, neglecting the spatial correlation), the standard deviation of the resistance increases to $\sigma_R = 590$ kN and accordingly the failure probability to $P_f = 1.37 \cdot 10^{-3}$.

Figure 80 (a) shows the a-priori density functions for the uncorrelated and correlated case. **Figure 80** (b) shows an example of the parameter correlation between the shaft friction q_s and the base resistance pressure q_b and the performance function G . The strong positive correlation shows that the pile bearing capacity is highly sensitive to the variation in the parameters of layer S2. The few negative values of the skin friction in layer S1 result from the variance in q_c and are neglected in this example.

In any case, the target reliability levels defined in section B.4.2-1 have not been met.

Figure 80. (a) A priori density and distribution functions. (b) Sensitivity of the power function G vs. the variation of skin friction q_s or tip pressure q_b .



Source: Arnold et al. 2022

Reliability updating using results from pile testing.

Case A: Update of the load-bearing capacity according to EN 1997-3 and EA-Pfähle

Firstly, the effect of the test results on the probability of failure is quantified according to the relevant standards. The derivation of the representative pile bearing capacity $R_{c,rep}$ from pile load tests is summarized in EA-Pfähle (2013). The following applies to static test loads:

$$R_{c,rep} = \min \left[(R_{c,m})_{mean} / \xi_1 ; (R_{c,m})_{min} / \xi_2 \right]$$

Equation 122.

with $(R_{c,m})_{mean}$ mean value of the ultimate loads in the load tests, $(R_{c,m})_{min}$ smallest ultimate load in the load tests and ξ_1, ξ_2 scatter factors depending on the number of piles subjected to the test loads. For case A, $\xi_1 = 1.25$ and $\xi_2 = 1.15$ ($n = 2$, "soft" piles, i.e. no load-distributing head plate) which results in:

$$R_{c,rep} = \min \left(\frac{0.5 \cdot (4.15 \text{ MN} + 4.45 \text{ MN})}{1.25} ; \frac{4.15 \text{ MN}}{1.15} \right) = 3.44 \text{ MN}$$

Equation 123.

The design value of the pile resistance R_d is determined for the case of point loads with a reduced partial safety factor $\gamma_b = \gamma_s = \gamma_t = 1.1$ (compression piles) and results in $R_d = 3.13 \text{ MN}$. This gives the load-bearing capacity of the pile with a design utilization of $2.3 \text{ MN} / 3.13 \text{ MN} = 0.73$.

Case B: Indirect updating of the failure probability using equality tests

Following, Bayesian inference is used to update the pile probability of failure for the testing case with equality information. In this example only the mean of the pile resistance μ_R is updated (i.e. the model parameter to be inferred) and as such works as a random variable. The other parameters stay fixed to their deterministic values.

The a-priori information can be described with two density distributions:

Mean of pile resistance:

$$\mu_R \sim \mathcal{N}(\mu_{R_m}, \sigma_{R_m}) \rightarrow f'_{\mu_R} = f_N(\mu_R | \mu_{R_m}, \sigma_{R_m})$$

Equation 124.

Pile resistance:

$$R \sim \mathcal{N}(\mu_R, \sigma_{R_p}) \rightarrow f'_R = f_N(R | \mu_R, \sigma_{R_p})$$

Equation 125.

This Bayesian analysis accepts the mean value of pile resistance as its parameter for estimation. The uncertainty in this parameter is epistemic, as it reflects that the uncertainty in the

understanding of the model can be reduced with additional observations. This epistemic uncertainty is described by the parameter σ_{R_m} . Additionally, the model is used to make predictions for the actual pile resistance, which eventually determines the failure of a pile. This prediction incorporates a component of aleatoric uncertainty, which is reflected by the irreducible parameter σ_{R_p} of the predictive model.

In case A, $n_s = 2$ test piles are tested on the construction site before construction begins. Upon making prediction only for piles that have participated in the testing, the uncertainty is described by $\sigma_s = \sigma_{meas}$. On the other hand, if the information gained through pile testing is transferred to the entire construction site, the update is influenced by both the measurement error and the variability of the subsoil, i.e. $\sigma_s = \sqrt{(\sigma_{R_p})^2 + \sigma_{meas}^2}$. The mean value is not affected, since the Bayesian model built in this example does not include a spatial component (does not accounts for the spatial variability).

$$S_i \sim \mathcal{N}(\mu_R, \sigma_s) \rightarrow f_{S_i} = f_N(S_i | \mu_R, \sigma_s)$$

Equation 126.

Assuming that the density distributions are normally distributed, the posterior distribution of the mean value of the pile resistance can be determined using the conjugate priors model for implementing Bayes' theorem (Ang and Tang 1975):

$$\mu_R \sim \mathcal{N}(\mu''_{R_m}, \sigma''_{R_m}) \rightarrow f''_{\mu_R} = f_N(\mu_R | \mu''_{R_m}, \sigma''_{R_m})$$

Equation 127.

$$\text{with } \mu''_{R_m} = \frac{\bar{s} \cdot (\sigma_{R_m})^2 + \mu_{R_m} \cdot (\sigma_s^2 / n_s)}{(\sigma_{R_m})^2 + (\sigma_s^2 / n_s)} \text{ and } \sigma''_{R_m} = \sqrt{\frac{(\sigma_{R_m})^2 \cdot (\sigma_s^2 / n_s)}{(\sigma_{R_m})^2 + (\sigma_s^2 / n_s)}}$$

The posterior distribution of the pile bearing capacity is then:

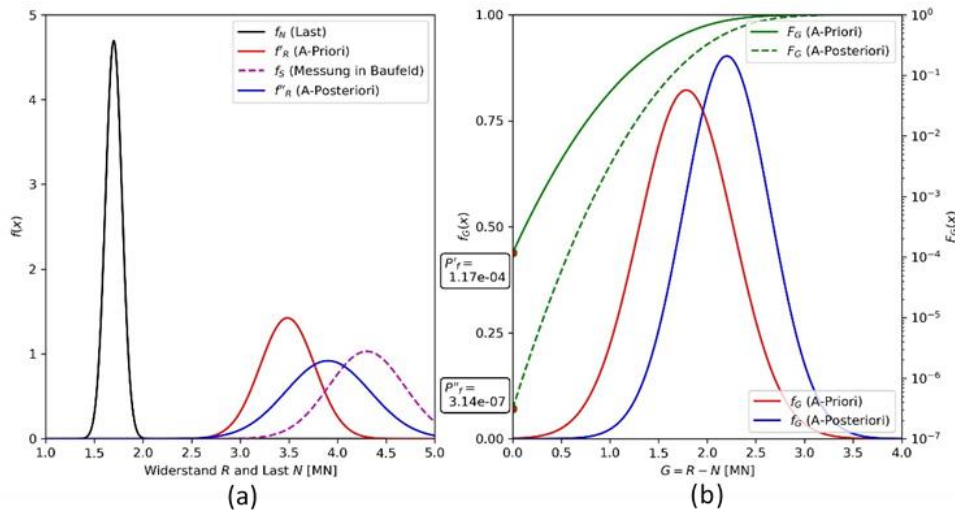
$$R \sim \mathcal{N}(\mu''_R, \sigma''_R)$$

Equation 128.

$$\text{with } \mu''_R = \mu''_{R_m} \text{ and } \sigma''_R = \sqrt{(\sigma''_{R_m})^2 + \sigma_{R_p}^2}.$$

Figure 81 shows the update of the probability of failure as an example for case A, assuming uncorrelated input parameters ($\rho_{q_s, q_b} = 0$) and a measurement without measurement errors ($\sigma_{meas} = 0$ kN). It can be seen that the update increases the estimated mean of the pile resistance, i.e. the posterior density function is to the right of the a priori density function, but the variance of the pile resistance increases somewhat. This can be justified by the fact that the variability in the construction site in the present example was assumed to be irreducible. The failure probability is nevertheless reduced from prior $P'_f = 1.17 \cdot 10^{-4}$ to posterior $P''_f = 3.14 \cdot 10^{-7}$ and would thus meet the requirements for a building with consequence class CC 2 according to EN 1990-1. This applies to all variations of case A, regardless of the correlation and the measurement error (see **Table 39**).

Figure 81. (a) Density functions of the variables, (b) density and distribution functions for performance function G .



Source: Arnold et al. 2022

If the representative load is increased to $N_{rep} = R_d/1.35 = 3.13 \text{ MN}/1.35 = 2.77 \text{ MN}$ (degree of utilization = 1.0), this in turn results in higher failure probabilities, which are below the reliability indices required by DIN EN 1990. In order to reduce the probability of failure, further pile load tests or subsoil investigations would be necessary.

Table 39. Summary of failure probabilities for case A.

Load	Corr. $\rho_{q_s,rep,q_b,rep}$	prior P'_f	posterior - P''_f		
			$\sigma_{meas} = 0.0$ MN	$\sigma_{meas} = 0.05$ MN	$\sigma_{meas} = 0.1$ MN
$N_{rep} = 1.7 \text{ MN}$	0.0	$1.17 \cdot 10^{-4}$	$3.14 \cdot 10^{-7}$	$3.92 \cdot 10^{-7}$	$7.34 \cdot 10^{-7}$
	0.8	$1.37 \cdot 10^{-3}$	$1.17 \cdot 10^{-7}$	$1.49 \cdot 10^{-7}$	$2.87 \cdot 10^{-7}$
$N_{rep} = R_d / 1.35$ $= 2.77 \text{ MN}$	0.0	---	$6.46 \cdot 10^{-3}$	$6.90 \cdot 10^{-3}$	$8.35 \cdot 10^{-3}$
	0.8	---	$2.85 \cdot 10^{-3}$	$3.08 \cdot 10^{-3}$	$3.82 \cdot 10^{-3}$

Case C: Indirect updating of the failure probability using inequality tests (Indirect updating of stress tests)

According to EN 1997-1 and EA-Pfähle, test loads up to the service load level (case B) cannot be used to derive pile resistances in the failure state. It is also not possible to reduce the partial safety factor on the resistance side on the basis of test loads up to the service load level (SLS). However, according to EN 1997-3, the model factor can be reduced when serviceability control tests are carried out, thereby increasing the design value of the pile resistance.

Using Bayesian inference, based on the inequality tests (test load $P_p = 1.0 \dots 1.3 N_{rep}$, see **Table 38**), the likelihood function based on a single measured value can be set up as follows:

$$L_S(\mu_R) = P(R \geq s | \mu_R, \sigma_{R_p}) = \int_s^{\infty} \frac{1}{\sigma_{R_p} \cdot \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{r-\mu_R}{\sigma_{R_p}}\right)^2\right\} dr$$

$$= 1 - \Phi\left[\frac{s-\mu_R}{\sigma_{R_p}}\right]$$

Equation 129.

where for several test load results n_s it is $L_{S_{n_s}}(\mu_R) = \prod_{i=1}^{n_s} P_1(R > s_i | \mu_R, \sigma_{R_p})$. The main parameter of the Bayesian model is again the mean of the pile resistance.

The posteriori distribution of the mean value of the pile resistance is proportional to:

$$f_{\mu_R | R \geq s}(\mu_R) \propto L_{S_{n_s}}(\mu_R) \cdot f_{\mu_R}(\mu_{R_m}, \sigma_{R_m})$$

$$= \left\{ \Phi\left[-\frac{s-\mu_R}{\sigma_{R_p}}\right] \right\} \cdot \frac{1}{\sigma_{R_m} \cdot \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\mu_R - \mu_{R_m}}{\sigma_{R_m}}\right)^2\right\}$$

Equation 130.

The posterior density function of the pile resistance R is:

$$f_{R | R \geq s}(r) = \int_{-\infty}^{\infty} f_R(r | \mu, \sigma_{R_p}) \cdot f_{\mu_R | R \geq s}(\mu) d\mu$$

Equation 131.

In this indirect update, the probability of failure then results in:

$$P_f'' = P(F | R \geq s) = \int_{-\infty}^{\infty} F_{R | R \geq s}(R \geq n) \cdot f_{N(load)}(n) dn$$

Equation 132.

where $F_{R | R \geq s}$ is the posterior cumulative density function of the pile resistance. The solution to the updating problem can be achieved via numerical methods.

Case D: Direct updating of the failure probability using inequality tests

The example above is also examined in a different perspective. According to Straub (2011, 2014), the a-posteriori failure probability can also be directly updated, taking into account the uncertainty in the measured values:

$$P_{F|s} = \frac{P(g(X) < 0 \cap h(X) > 0)}{P(h(X) > 0)}$$

Equation 133.

Where $g(\cdot)$ is the limit state function and $h(\cdot)$ is the function of the inequality observations (measurements). **Equation 133** can be rewritten as:

$$P_f'' = P(F|s) = \frac{P(R < N \cap R > s)}{P(R > s)} = \frac{P(R - N < 0 \cap R - s > 0)}{P(R - s > 0)}$$

Equation 134.

The probability of failure can, for example, be approximated numerically using Monte Carlo simulations that generate realizations of R, S and N:

$$P_f'' = P(F|s) = \frac{\sum_{i=1}^N 1_A[R_i < N_i] \cdot 1_A[R_i > S_i]}{\sum_{i=1}^N 1_A[R_i > S_i]}$$

Equation 135.

Where n_{mc} is the number of samples and 1_A is the indicator function:

$$1_A[x < y] = \begin{cases} 1, & \text{when } x < y \\ 0, & \text{when } x > y \end{cases}$$

Equation 136.

In our example, the indirect way is comparatively simple and can be described analytically to a good extent, since the density functions are normally distributed. In a more complex setting the direct route with a numerical approximation is the best choice.

Table 40 shows the posterior failure probabilities for case B, whereby only the results for the uncorrelated input parameters in layer S2 are given here. By carrying out the SLS tests on four structural piles, the probability of failure is significantly reduced.

If the empirical values according to EA-Pfähle are used as a priori information and the two ULS test loads from case A are not taken into account, the probability of failure is reduced by a factor of 10. However, the execution of the two test loads from case A is more effective because the a priori probability of failure is reduced by a factor of 1,000 when the breaking load is determined (see **Table 39**).

Furthermore, if the posterior density functions of case A are now used as priori information for case B, i.e. the prior mean value of the pile resistance increases, then the uncertainty is again minimized and the probability of failure again reduced by a factor of 10 (see **Table 40**).

The execution of SLS tests on building piles can thus help to confirm or increase the safety level set in the design and thus enable potential for use or later conversion.

Table 40. Summary of failure probabilities for case B.

Load	Prior information	prior - P_f'		posterior - P_f''	
		$\sigma_{meas} = 0.0 \text{ MN}$	$\sigma_{meas} = 0.1 \text{ MN}$	$\sigma_{meas} = 0.0 \text{ MN}$	$\sigma_{meas} = 0.1 \text{ MN}$
$N_{rep} = 1.7 \text{ MN}$	EA-Pfähle	$1.17 \cdot 10^{-4}$	---	$1.03 \cdot 10^{-5}$	$1.09 \cdot 10^{-5}$
	Case A	$3.14 \cdot 10^{-7}$	$7.34 \cdot 10^{-7}$	$4.06 \cdot 10^{-8}$	$6.50 \cdot 10^{-8}$

Interpretation of results and convergence

Using the design of an axially loaded pile as an example, it is shown that Bayesian inference is well suited to quantify uncertainties of various origins (variability in the foundation soil, uncertainty in

the transformation model, limited number and different types of test loads, measurement uncertainties, etc.) and thus to make them available to a reliability-based design or risk assessment.

Software Tools

A custom Python code was developed for the example. For parts of the analysis the library PyMC (<https://www.pymc.io/>) was used.

References

- Ang, A. H.-S., Tang, W. H. (1975). Probability Concepts in Engineering Planning and Design, Volume 1 – Basic Principles, 1st Edition, John Wiley and Sons, New York
- EN 1990-1:2023. Eurocode: Basis of structural and geotechnical design – Part 1: Design of new structures. (Draft Version) CEN/TC 250/SC 10.
- EN 1997-1:2023. Eurocode 7: Geotechnical design – Part 1: General rules. (Draft for Formal Vote) CEN/TC 250/SC 7/WG 1 “Evolution of 1997 series”.
- EN 1997-3:2024. Eurocode 7: Geotechnical design – Part 3: Geotechnical structures. CEN/TC 250/SC 10.
- EA-Pfähle (2013). Recommendations on Piling (EA Pfähle), Deutschen Gesellschaft für Geotechnik e. V. (Ed.), Ernst und Sohn.
- Kempfert, H.-G. und Becker, P. (2007). Grundlagen und Ergebnisse der Ableitung von axialen Pfahlwiderständen aus Erfahrungswerten für die EA-Pfähle, Bautechnik Vol. 84 No 7, pp. 441-449.
- Straub, D. (2011). Reliability updating with equality information. Probabilistic Engineering Mechanics, Vol. 26 No 2, pp. 254–258.
- Straub, D. (2014). Value of information analysis with structural reliability methods. Structural Safety Vol. 49, pp. 75-86.

B.5. Retaining structure

B.5.1 Probabilistic FEM analysis of a retaining wall

Authors: Alexandra Ene & Timo Schweckendiek

Reference: A. Ene, T. Schweckendiek and H. Popa (2022). Full Probabilistic Analysis with FEM for the Retaining Wall of a Deep Excavation. Proceedings of the 8th International Symposium on Reliability Engineering and Risk Management 4–7 September 2022, Hannover, Germany, pp. 439-446

Problem definition

Description

This example presents a full probabilistic analysis performed for a real case temporary retaining system of a deep excavation. The probabilistic analysis was combined with FEM using advanced constitutive models. The objective is to present the procedure followed for the probabilistic analysis coupled with FEM software in order to facilitate the implementation of reliability-based design, and

to assess the reliability obtained for such structure compared to the reliability produced by partial factor design.

The main uncertainties consisting of main geotechnical parameters and model factor were modelled as random variables and the limit state verification was expressed in terms of reliability index for SLS and ULS verification.

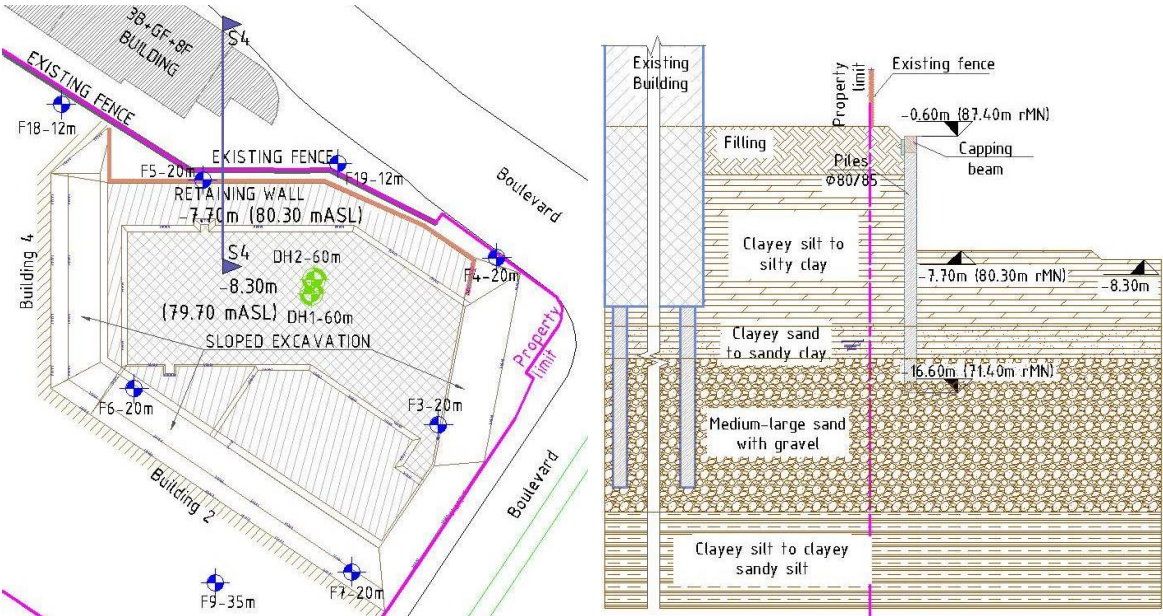
The reliability of the retaining wall at ULS was evaluated probabilistically just for the load (effects) part of the general verification, while working with design values for the resistances (see sub-section 4.3 in the guideline document).

Available data

The case study presented in this report refers to a temporary retaining wall for a deep excavation in Bucharest, Romania. The excavation pit for the two basement levels was 7.7 m deep for the marginal area and 8.3 m deep in the central area. On most of the area, a sloped excavation was considered, and on a side where the excavation was led near the property limit, a self-supporting embedded wall was provided. The retaining wall consisted of drilled reinforced concrete piles, 80 cm diameter at 85 cm inter-axes distance, 16 m length. The excavation pit layout, including adjacent constructions and site investigations, and a characteristic section of the retaining system are shown in **Figure 82**.

The ground investigations consisted of seven geotechnical boreholes with sampling of soil specimens and Standard Penetration Tests in cohesionless soil layers. The results of the investigations are described in more detail below.

Figure 82. Excavation and retaining system layout (left) and characteristic section (right).



Source: Ene and Schweckendiek 2022

Limit State function(s)

Verification of Serviceability Limit State for the retaining wall.

The limit state function for SLS was set as follows:

$$Z_{wall}^{SLS} = d_{wall}^{SLS} - \max U_{x,wall} \cdot f_m^{FEM}$$

Equation 137.

d_{wall}^{SLS} – target (limit) design value for the horizontal displacement of the retaining wall, set at 3.5 cm; $\max U_{x,wall}$ – maximum horizontal displacement of the retaining wall taken from the FEM analysis through automated commands; f_m^{FEM} – model factor considered for FEM analysis.

Verification of geotechnical Ultimate Limit State.

The limit state function for geotechnical ULS was set as follows:

$$\text{if } \sum M_{stage} > 0.995 : Z_{GEO}^{ULS} = \sum M_{sf} \cdot f_m^{FEM} - 1$$

$$\text{else: } Z_{GEO}^{ULS} = \sum M_{stage} \cdot f_m^{FEM} - 1$$

Equation 138.

if $\sum M_{stage} > 0.995$: $Z_{GEO}^{ULS} = \sum M_{sf} \cdot f_m^{FEM} - 1$ $\sum M_{stage}$ – ratio between the load for which failure occurs in FEM and the desired ultimate load taken from the FEM analysis through automated commands; $\sum M_{sf}$ – ratio between the initial soil shear resistance and the resistance for which failure occurs taken from the FEM analysis through automated commands; f_m^{FEM} – model factor considered for FEM analysis.

Verification of structural Ultimate Limit State

The limit state functions for structural ULS were set as follows:

$$Z_M^{ULS} = M_{Rd} - \max M_d \cdot f_m^{FEM}$$

Equation 139.

$$Z_S^{ULS} = S_{Rd} - \max S_d \cdot f_m^{FEM}$$

Equation 140.

M_{Rd} and S_{Rd} – design value of the bending moment and shear force, respectively, resistance/ capacity; $\max M_d$ and $\max S_d$ – design value for the bending moment and shear force, respectively, in the retaining wall taken from the FEM analysis through automated commands; f_m^{FEM} – model factor considered for FEM analysis.

The resistance (bending and shear capacity) was established based on deterministic analysis using representative values for the geotechnical parameters and applying the corresponding partial factors, according to Eurocode 7 (i.e. by partial factors for the permanent and unfavourable variable loads $\gamma_G = 1.35$ and $\gamma_Q = 1.5$, respectively).

It was assumed that the reinforcement of the retaining wall led to a capacity exactly equal to design load effect, although in practice this is rarely the case, and the reinforcement will often be slightly over-designed.

The retaining system of the excavation was modelled using the 2D Finite Element model for plane strain state in Plaxis 2019 software, by drained analysis but using total stresses shear resistance

parameters (Ene, Schweckendiek, & Popa, 2022). The Hardening Soil model with small stiffness behaviour was used with the main parameters given in **Table 41** and **Table 44**.

Target reliability

For Serviceability Limit State, reference is given for RC2 in EN 1990 (CEN, 2019) for the target reliability index: 1.5 for 50-years reference period and 2.9 for 1-year reference period, respectively.

The available target reliability index for ULS, as given in EN 1990 was considered with the following assumptions:

- Reference target reliability index was established for a design life of the structure of 50 years (i.e. 3.8 target value for ULS and 1.5 for SLS for RC2);
- A value of 4.0 for the reliability index should be enough to ensure minimum required reliability for lower reference period (e.g. 2 years);
- The implications of lowering the reliability index by one class (i.e. 3.3 target value for RC1) and by half a class (i.e. 3.55 target value) for temporary structures – as suggested in the Probabilistic Model Code (Joint Committee on Structural Safety, 2001) – is analysed separately.

For the structural ULS verification, only the load side of the problem is analysed probabilistically and the target beta value is taken as $\alpha_E \cdot \beta$. Thus, the target reliability index for ULS would be:

- $0.7 \cdot 4 = 2.8$ as for permanent structures, but at lower reference period;
- $0.7 \cdot 3.5 = 2.45$ as for temporary structures for which the reliability index is lowered by one class, and at lower reference period.

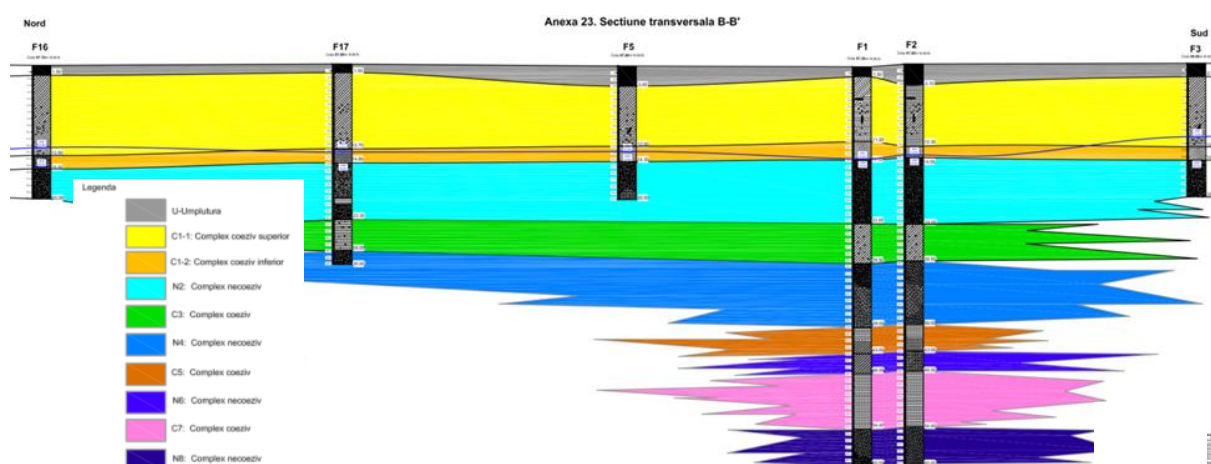
Ground Model

The main layers representative for the retaining wall model were:

5. Filling (with thickness varying between 1.50 m and 3.10 m), consisting of construction waste material, cemented or incorporated in a stiff clay;
6. Cohesive Layer C1-1 of clayey silt and silty clay, stiff and with increased consistency in some areas (thickness varying between 9.00 m and 10.80 m);
7. This was assimilated with the typical “Bucharest Clay” layer (yellow colour in Figure B.5.1-2).
8. Cohesive Layer C1-2 of sandy silty clay, sandy clayey silt, stiff (thickness varies between 1.70 m and 2.80 m);
9. Cohesionless layer 2 of sand and silty sand with little gravel, submerged, dense (thickness varies between 9.50 m and 10.00 m).

A geological section with the layers encountered within the site investigations is given in **Figure 83**.

Figure 83. Geological section on the eastern side of Building 1 (taken from the Geotechnical Report).



Source: Authors' own work

The groundwater level was found at depths varying between 13.20 and 14.50 m, which is below the final excavation level and did not involve any uplift concerns.

The results of the site and laboratory tests were statistically described to determine the geotechnical parameters for the design as follows: mean and standard deviation for spatial (layer) average properties of the "Bucharest Clay" (layer 1, silty clay), which is dominant for the retaining system in this project. The ground properties of the other soil layers were taken as representative values according to EN 1997-1 and *Romanian norm on determining the characteristic and design values for the geotechnical parameters* (NP 122:2010), assuming normal distribution and spatial averaging, as given in **Table 41**.

Table 41. Representative values for the geotechnical parameters used in the design.

Layers and levels	γ_{rep}^{sup} [kN/m ³]	φ'_{rep}^{inf} [°]	c'_{rep}^{inf} [kPa]	$E_{50} = E_{oed}$ [MPa]	E_{ur} [MPa]	G_0 [MPa]	$\gamma_{0.7}$ [-]	p_{ref} [kPa]
1. Filling (88÷85 m ASL)	19	15	5	3	9.7	40	1.00E-04	40
2. Silty clay (85÷75.5 m ASL)	See Table 44							
3. Sandy clay (75.5÷73.5 m ASL)	19.8	12.8	29.3	19.2	57.6	240	1.80E-04	280
4. Sand with gravel (73.5÷64 m ASL)	20.2	39	-	40	120	500	1.10E-04	470

γ_{rep}^{sup} – superior representative values for unit weight; φ'_{rep}^{inf} – inferior representative values for the effective angle of friction; c'_{rep}^{inf} – inferior representative values for the effective cohesion; E_{50} – the triaxial loading stiffness considered equal to E_{oed} – oedometric deformation modulus at the reference pressure; E_{ur} – unloading-reloading deformation modulus at the reference pressure; G_0 – small strain shear modulus at the reference pressure; $\gamma_{0.7}$ – strain level at 70% of shear modulus; p_{ref} – reference pressure; R_{inter} – strength reduction factor for interface.

Input Parameters (random and deterministic variables)

Since the 9 to 11 m thick silty clay layer, called “Bucharest Clay” is dominant for the current case study, the approach was to consider the statistical description for this layer and keep the other parameters as deterministic with representative values (see **Table 41**). An analysis of different statistical distributions for the geotechnical parameters was performed prior to the reliability analysis, using different assumptions (see sub-section B.5.1.2). Other relevant uncertainties related to the soil model parameters, together with a model factor (see B.5.1.3) were considered through a sensitivity analysis (Ene et al., 2022).

Based on the sensitivity analysis, the significant random variables chosen for the reliability analysis are the shear resistance parameters of the “Bucharest Clay” layer (c_{ref} and φ), and the model factor (f_m^{FEM}) as given in **Table 42**.

Table 42. Statistics of the stochastic variables used in the probabilistic analysis.

Variable	$tg(\varphi)$	c_{ref}	f_m^{FEM}
Mean value, μ	0.4251	39.5	1
	$tg(23.9^\circ)$		
Standard Deviation, σ_x	0,0369	5.95	0.1
Coefficient of variation, v_x	9%	15%	10%
Correlation factor, ρ	-0.8741		-
Distribution	Normal		

Uncertainty characterization

Geotechnical units and parameters

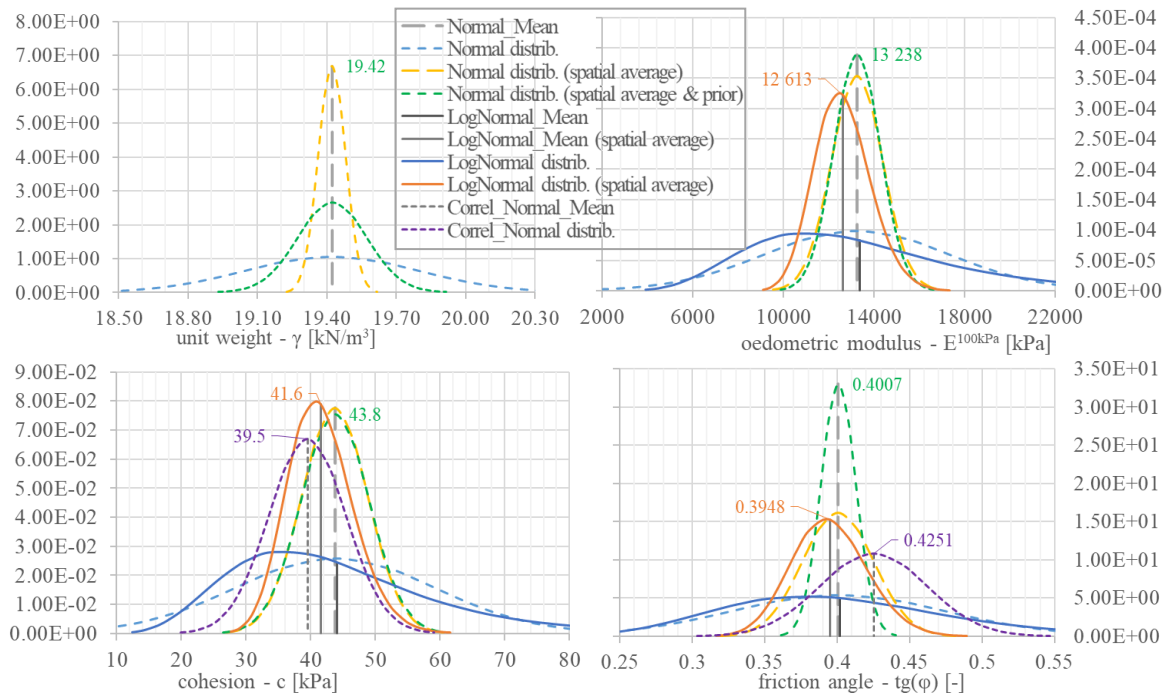
The probability distributions of the soil parameters considered as random variables are represented in **Figure 84** for the sample distribution and for spatial averaging (i.e. estimation of the mean value with 95% confidence level), considering statistical uncertainty through Student-t factor at 95% confidence level.

The following statistical distributions resulted as reasonable for the present case study, based on reliability analysis and on physical and mathematical justifications (Ene, 2021):

- Normal distribution for unit weight of the soil (at natural moisture content);
- Lognormal distribution for the deformation modulus (determined based on correlations with the oedometric modulus) – this was chosen because of the higher coefficient of variation of the sample data;
- Normal distribution resulted from linear regression analyses for the shear resistance parameters (tangent of the internal friction angle and cohesion) – resulted as rational from reliability analysis and justified by the high correlation of the two parameters that are determined from the same test.

The statistics of the geotechnical parameters considered as random variables for the present case study are given in **Table 43**, with the spatial averaging assumption and statistical uncertainty by Student-t factor, as mentioned before.

Figure 84. Probability distributions of the main geotechnical parameters.



Source: Authors' own work

Table 43. Statistics of the soil parameters considered as random variables for the “Bucharest Clay” layer.

Variable	$\gamma_{unsat} = \gamma_{sat}$	E_{50}^{ref}	$tg(\varphi)$	c_{ref}
	[kN/m ³]	[MPa]	[-]	[kPa]
Sample size, n (number of observations)	42	15	33	
Mean value, μ	19.42	12 552	0.4251 tg(23.9°)	39.5
Standard Deviation, σ_x	0.06	1 240	0.0369	5.95
Coefficient of variation, v_x	3%	10%	9%	15%
Correlation factor, ρ	-	-	-0.8741	
Distribution	Normal	Lognormal	Normal	

Loads, groundwater and pore pressure

The surcharge load outside the excavation was considered as deterministic in the probabilistic analysis, with the representative value of 5 kPa and with a partial factor of $\gamma_Q/\gamma_G = 1.11$.

The groundwater was not considered to influence the results.

Model uncertainties

A model factor for FEM analysis of the retaining structure (f_m^{FEM}) was considered as a random variable. This was taken as normally distributed with mean 1 and standard deviation 0.1 as per Section 3.7, table 3.7.5.1 for “Stability of retaining (sheetpiled) walls” in Probabilistic Model Code (Joint Committee on Structural Safety, 2001).

Reliability analysis

Selection of reliability method(s)

FORM analysis was attempted, but most of the times this could not reach convergence. This was noticed as the output from Plaxis 2D presented some randomness in the results for the same input values. This led to significant “noise” in the evaluation of the limit state function due to small calculation errors that exact methods such as FORM cannot handle.

Then, since crude Monte Carlo requires a very high number of calculations, Importance Sampling (IS) around the (estimated) design point was adopted. However, in order to obtain a good estimate of the design point, it is necessary to have some prior knowledge of the failure area. This was solved by performing the analysis in two steps: one IS run around the mean values, and a second IS run around the design point obtained from the first run in order to improve the precision. For the case study analysed within this example, it was necessary to perform about 1000 iterations around mean values and about 500 to 1000 iterations around the design point from the first calculation. Hence, in total about 1500–2000 Plaxis 2D model runs were required to reach a precision of 0.1 for the probability of failure (i.e. coefficient of variation).

At last, most of the reliability analyses were performed using Erraga metamodel developed at Delft University of Technology and Deltares (van den Eijnden, Schweckendiek, & Hicks, 2019) which led to very good results in less than 50 iterations most of times. The main results were verified by Importance Sampling performed in two steps, as described above. Also, some supplementary verifications were also performed by Directional Sampling.

Estimation of probability of failure

For the SLS verifications of the retaining wall, two models have been analysed:

1. Considering for the soil properties only the shear resistance parameters as stochastic variables;
2. Considering for the soil properties the unit weight of the soil, the linear deformation modulus and the shear resistance parameters as stochastic variables.

The reliability index β resulting from the analyses varied between 1.17 for the first scenario and 2.38 for the second scenario.

Verification of geotechnical Ultimate Limit State

For the verification of the geotechnical ULS, neither Erraga metamodel found any failure point in more than 120 iterations, nor did DS in 200 iterations (725 realizations).

An attempt was to use IS and, to reduce the calculation time for this, this was performed in one step around a design point taken from the Erraga realizations, as a close estimation of the limit state function. This analysis provided a reliability index of 4.86.

Verification of structural Ultimate Limit State

The reliability analysis for the structural ULS was performed using different values for the bending moment and shear force resistance, corresponding to different values of partial factor for the permanent loads.

Starting with the partial factors given in EN 1990 – $\gamma_G = 1.35$ for permanent loads, and $\gamma_Q = 1.5$ for variable loads – a reliability index β of about 3.5 was obtained. Lower values for the partial factors

were used afterwards to calculate the design resistance (bending moment and shear force capacity of the retaining wall), while keeping the ratio between the permanent load and the variable load partial factors constant with $\gamma_Q / \gamma_G = 1.11$.

The target values for the design resistances and the corresponding partial factors are given in **Table 44**, together with the main results obtained in terms of reliability.

Table 44. Reliability results for the verification of structural Ultimate Limit State for the normal design situation (Design Case 1).

LSF	Design resistance	γ_G	γ_Q / γ_G	β	P_f
Z_M^{ULS}	$M_{Rd} = 539.1$ [kNm/m]	1.35	1.11	3.45	2.80E-04
	$M_{Rd} = 499.1$ [kNm/m]	1.25	1.11	2.74	3.10E-03
Z_S^{ULS}	$S_{Rd} = 155.4$ [kN/m]	1.35	1.11	3.54	2.00E-04
	$S_{Rd} = 143.9$ [kN/m]	1.25	1.11	2.73	3.20E-03

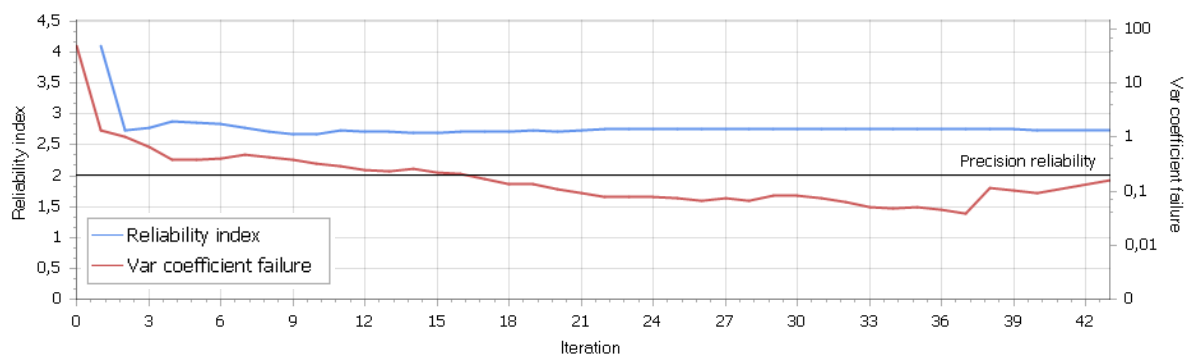
An example of the results obtained from the structural ULS reliability analysis is given below (**Table 45** to **Table 48** and **Figure 85** to **Figure 88**).

Table 45. Probabilistic results for limit state function $Z_M (M_{Rd} = 1.25 \cdot M_k)$ using Erraga metamodel.

Results	Variable			β	P_f
	$tg(\phi)$	C_{ref}	f_m^{FEM}		
Alpha	0.53	0.235	-0.815	2.74	3.10E-03
Influence factor	28.00%	5.50%	66.40%		
Correlated alpha – PTK	0.544	-0.358	-0.837		
Design point value – PTK	0.37009	50.09	1.23		
Design point value*	0.37009	45.34	1.23		

(*) Erraga estimated failure point

Figure 85. Probabilistic results for limit state function $Z_M (M_{Rd} = 1.25 \cdot M_k)$ using Erraga metamodel (screenshot Probabilistic Toolkit).

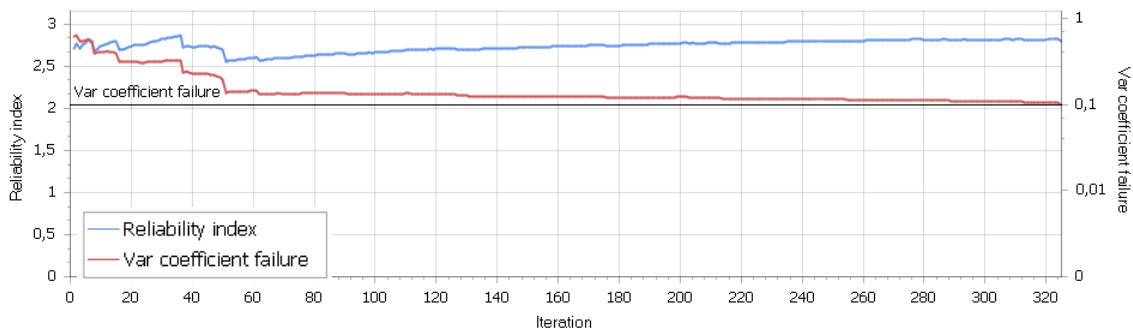


Source: Authors' own work

Table 46. Probabilistic results for limit state function $Z_M (M_{Rd} = 1.25 \cdot M_k)$ using importance sampling around design point from Erraga metamodel.

Results	Variable			β	P_f
	$tg(\varphi)$	C_{ref}	f_m^{FEM}		
Alpha	0.624	0.189	-0.759	2.79	2.60E-03
Influence factor	38.90%	3.60%	57.50%		
Correlated alpha	0.656	-0.477	-0.799		
Design point value	0.35748	52.88	1.22		

Figure 86. Probabilistic results for limit state function $Z_M (M_{Rd} = 1.25 \cdot M_k)$ using Importance Sampling around design point from Erraga metamodel (screenshot Probabilistic Toolkit).



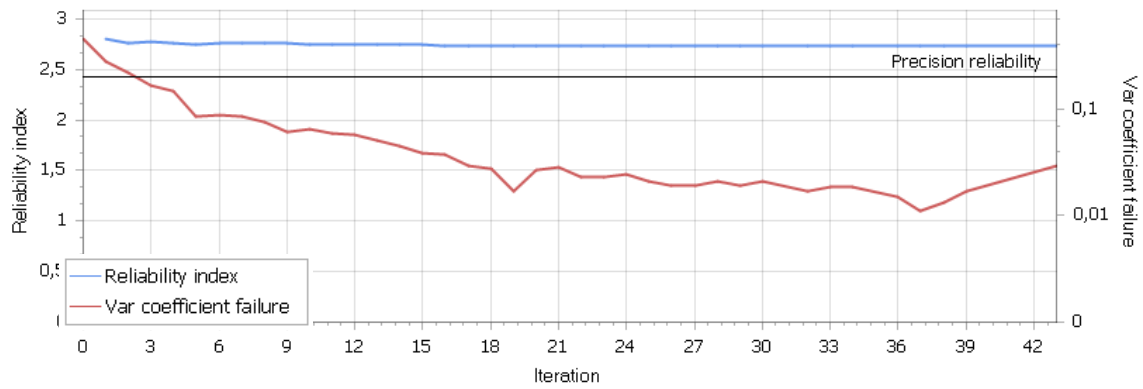
Source: Authors' own work

Table 47. Probabilistic results for limit state function $Z_S (S_{Rd} = 1.25 \cdot S_k)$ using Erraga metamodel.

Results	Variable			β	P_f
	$tg(\varphi)$	C_{ref}	f_m^{FEM}		
Alpha	0.63	0.219	-0.745	2.73	3.20E-01
Influence factor	39.7%	4.8%	55.5%		
Correlated alpha – PTK	0.654	-0.461	-0.773		
Design point value – PTK	0.35919	52.43	1.21		
Design point value*	0.35919	47.00	1.21		

(*) Erraga estimated failure point

Figure 87. Probabilistic results for limit state function $Z_S (S_{Rd} = 1.25 \cdot S_k)$ using Erraga metamodel (screenshot Probabilistic Toolkit).

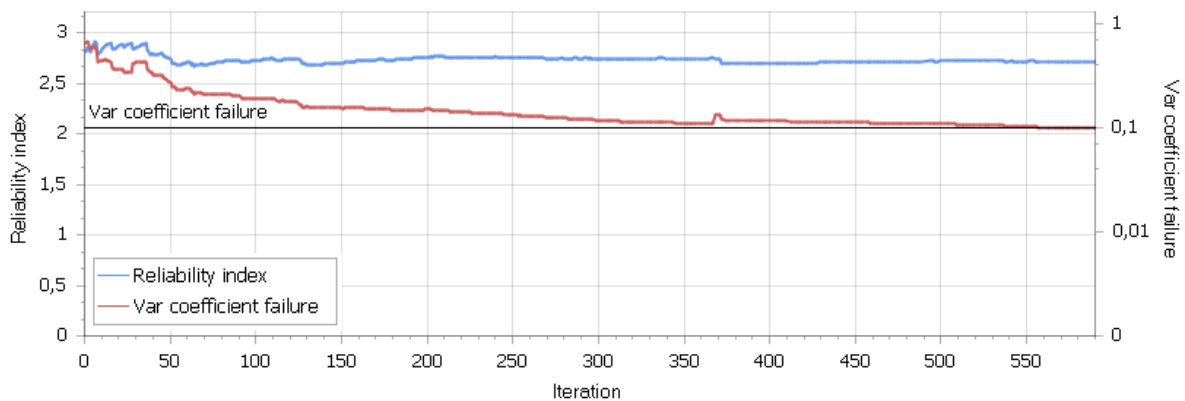


Source: Authors' own work

Table 48. Probabilistic results for limit state function $Z_S (S_{Rd} = 1.25 \cdot S_k)$ using Importance Sampling around design point from Erraga metamodel.

Results	Variable			β	P_f
	$tg(\varphi)$	C_{ref}	f_m^{FEM}		
Alpha	0.507	0.131	-0.852	2.72	3.30E-03
Influence factor	25.7%	1.7%	72.6%		
Correlated alpha	0.549	-0.411	-0.922		
Design point value	0.37001	50.5	1.25		

Figure 88. Probabilistic results for limit state function $Z_S (S_{Rd} = 1.25 \cdot S_k)$ using Importance Sampling around design point from Erraga metamodel (screenshot Probabilistic Toolkit).



Source: Authors' own work

Interpretation of results and convergence

First attempts to use FORM analysis encountered convergence problems due to significant “noise” in the evaluation of the limit state function. The noise appeared to be inherent to using the Hardening Soil Small Strain constitutive model. Attempts to adjust the FORM algorithm parameters to better deal with the noise improved the performance, but did not remove the convergence issues entirely.

Importance Sampling (IS) performed in two steps was found as a feasible (i.e. acceptable number of model evaluations) and accurate alternative. This procedure was verified for one of the analysis of the present case study through Directional Sampling, which required between 1200 and 2300 iterations, which means between 4500 and 8100 Plaxis 2D realizations for the same precision.

Erraga metamodel, used as an external model in Probabilistic Toolkit software, seems a promising new reliability method for complex analyses. This has shown good estimation of the reliability index, validated by classical methods such as Importance Sampling, in very few iterations. Also, in terms of estimation of the design point, Erraga provides trustworthy results.

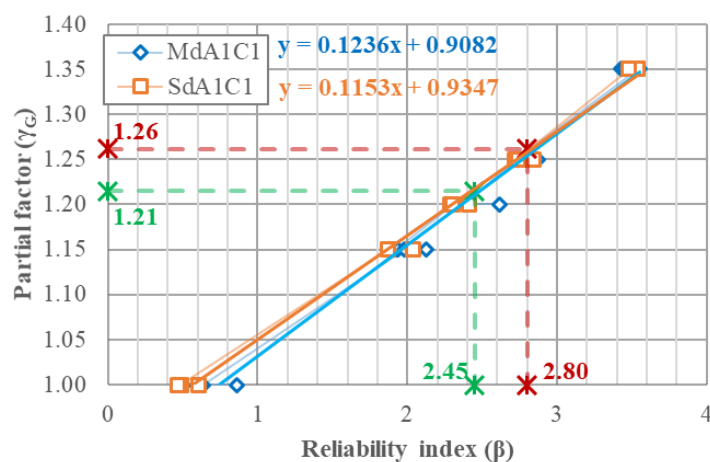
The reliability index resulted for SLS varies between 1.17 (which would be unacceptable according to the target value of 1.5 for permanent structures and 50-years reference period provided by Eurocode) and 2.38 (which should be acceptable for lower reference period and for mostly time invariant problems).

The preliminary results for the geotechnical ULS lead to significant margin compared to a target beta of 4, as considered for permanent structures with lower reference period for mostly time invariant problem. Although this limit state is not expected to dictate the design at least with the given geometrical details of the structure, validating the model can confirm the results obtained so far.

Using the partial factors given in Eurocode for loads, the reliability index obtained is significantly higher than the reference values chosen for the load verification: $\beta = 0.7 \cdot 4 = 2.8$ for permanent structures or $\beta = 0.7 \cdot 3.5 = 2.45$ for temporary structures, respectively (see sub-section “Target reliability”).

A relation was set between different values of partial factors for loads and reliability index was established, following a liner regression found as suitable (**Figure 89**).

Figure 89. Reliability results for the structural ULS verification.



Source: Ene and Schweckendiek 2022

Thus, it can be determined that – for this specific case study analysed – a set of partial factors $\gamma_G = 1.26$ and $\gamma_Q = 1.4$ would correspond to a reliability index of $\beta = 2.8$ for permanent structures, and a set of partial factors $\gamma_G = 1.21$ and $\gamma_Q = 1.34$ would correspond to a reliability index of $\beta = 2.45$ for temporary structures.

Both structural ULS verifications, Z_M^{ULS} and Z_S^{ULS} led to about the same values of the reliability index and corresponding partial factors.

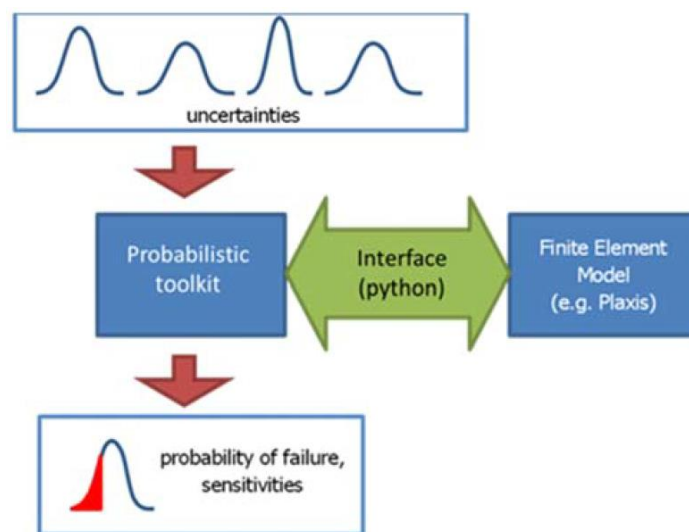
Software Tools

The coupling of a reliability tool with the FEM software represents setting the communication necessary to exchange the data (input and output parameters) between the two. Each of the models (FEM and reliability) is set independently and the calculations are performed in sequences.

Probabilistic Toolkit reliability software is used and coupled with FEM Plaxis 2D commercial software for geotechnical analysis and a new surrogate model is also used as a reliability method (Erraga metamodel).

The procedure of coupling Probabilistic Toolkit – PTK (Deltares, 2019a) with FEM software Plaxis 2D is represented in **Figure 90** (Deltares, 2019b).

Figure 90. Principle of the coupling between PTK and PLAXIS (Deltares, 2019b).



Source: Deltares 2019b

References

Deltares (2019a). Probabilistic toolkit- User Manual (version 2019). Deltares, Delft. URL: <https://www.deltares.nl/en/software/probabilistic-toolkit-ptk/>

Deltares (2019b). *Probabilistic Tools: Reliability Based Soil-Structure Analysis using FE*.

EN 1990-1:2023. Eurocode: Basis of structural and geotechnical design – Part 1: Design of new structures. (Draft Version) CEN/TC 250/SC 10.

EN 1997-1:2023. Eurocode 7: Geotechnical design – Part 1: General rules. (Draft for Formal Vote) CEN/TC 250/SC 7/WG 1 “Evolution of 1997 series”.

EN 1997-3:2024. Eurocode 7: Geotechnical design – Part 3: Geotechnical structures. CEN/TC 250/SC 10.

Ene, A. (2021). Determinarea valorilor caracteristice ale parametrilor geotehnici prin diferite metode statistice pentru calculul peretelui de susținere a unei excavații din București. *Lucrarile celei de-a XIV-a Conferinta Nationala de Geotehnica si Fundatii Bucuresti* (pp. 169-176). Bucharest: SRGF.

Ene, A., Schweckendiek, T., & Popa, H. (2022). Full Probabilistic Analysis with FEM for the Retaining Wall of a Deep Excavation. In E. Z.-K. Michael Beer (Ed.), *Proc. of The International Symposium on Reliability Engineering and Risk Management (ISRERM 2022)* (pp. 439-446). Hannover: Research Publishing, Singapore. doi:10.3850/978-981-18-5184-1_MS-13-104-cd

Joint Committee on Structural Safety (2001). *Probabilistic Model Code*. Technical University of Denmark.

Ministerul Dezvoltării Regionale și Turismului (2011, March 4). NP 122:2010 Romanian norm on selecting the characteristic and design values of geotechnical parameters. *Romanian Norm on Determining the Characteristic and Design Values for the Geotechnical Parameters*. România: Monitorul Oficial al României.

van den Eijnden, A. P., Schweckendiek, T., & Hicks, M. A. (2019). Metamodelling for geotechnical reliability analysis

B.5.2 Soldier pile wall

Author: Stéphane Commend

Reference: -

Problem definition

Description

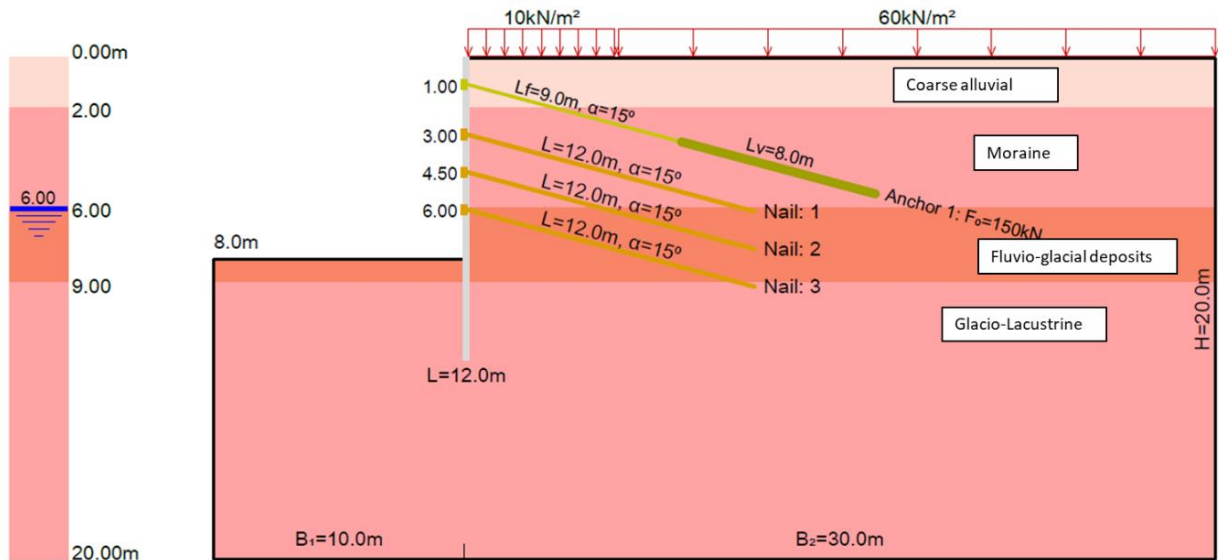
This example deals with the design of a soldier pile wall in Chardonne, Switzerland, conducted with a 2D finite element model. A deterministic analysis with representative values given in the geotechnical report was performed first, yielding results in terms of maximal bending moment in the steel profile, maximal horizontal displacement of the wall, and maximal settlement behind the wall.

As soil parameters and position of the water table are not known exactly, a sensitivity analysis is then performed on the main geotechnical parameters for the four soil layers (E_v , c' , φ') and on the water level, in order to show the influence of all these parameters on the quantities of interest (here: bending moment and settlements). Then, identifying that the water level and the friction angle of the weak “fluvio-glacial deposits” soil layer are the two most influent parameters, two reliability analyses are executed.

Available Data

A geotechnical report, with ranges indicated for apparent unit weight γ , effective friction angle φ' , effective cohesion c' and deformation modulus E_v for the four soil layers was available. It was not specified in the report if these values were representative (characteristic) or mean values. The pre-design of the retaining system was also available at the start of the analysis. The system geometry and the ground conditions are presented in **Figure 91**.

Figure 91. Investigated system of soldier pile wall.



Source: Authors' own work

Limit State function(s)

According to the quantities identified as critical for the design the limit state equations are defined as follows:

$$g_1(X) = 65 \text{ kNm/m} - M_{max} < 0$$

Equation 141.

$$g_2(X) = 20 \text{ mm} - s_{max} < 0$$

Equation 142.

where M_{max} is the absolute value of the maximal bending moment in the soldier pile wall, and s_{max} is the maximal settlement behind the wall, at the terrain level.

The limit for $g_1(X)$ has been set as the plastic bending moment of the pre-designed wall profile, and the limit for $g_2(X)$ has been set as a tolerable settlement defined by discussions with the client's neighbour.

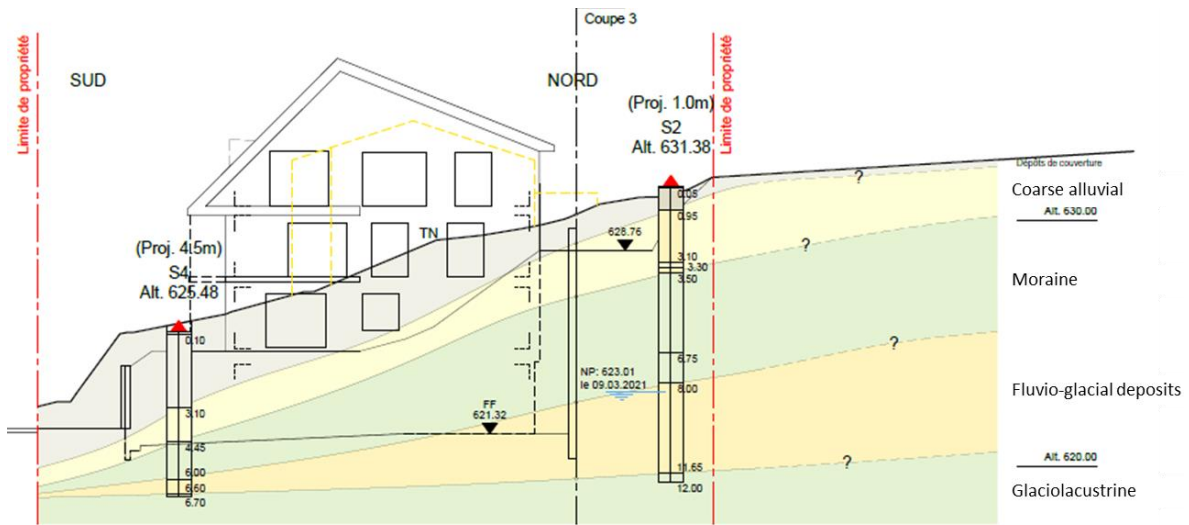
Target reliability

No target reliability is actually selected and/or determined.

Ground Model

The ground model used in the reliability analyses has been adopted from the geotechnical report and is depicted in **Figure 92**.

Figure 92. Ground model used for reliability analysis.



Source: Authors' own work

Ground property ranges given in the geotechnical report are summarized in **Table 49**.

Table 49. Range of ground property values of the present soil layers.

	γ [kN/m ³]	E [kN/m ²]	c' [kN/m ²]	ϕ' [°]
Coarse alluvial	19–21	10–15	0	32–35
Moraine	21–22	20–30	15–20	30–32
Fluvio-glacial deposits	20–21	20–25	0–2	32–35
Glaciolacustrine	21–22	20–30	15–20	26–28

Input Parameters (random and deterministic variables)

In this analysis the deformation modulus E_v , effective friction angle ϕ' and effective cohesion c' as well as the groundwater level h_{water} are considered to be of major influence on the system behavior and, hence, are treated as random variables. In opposite to this the unit weight of the soil layers is assumed to be of minor effect and, thus can be treated as deterministic input variable in the reliability analyses (see **Table 50**). Besides, also the surface loads are assumed to be deterministic.

Table 50. Overview of deterministic and random variable within this example.

Deterministic Variable	Unit	Description
γ	kN/m ³	Unit weight
q	kN/m	Surface load
Random Variable	Unit	Description
E	MN/m ²	Deformation modulus
c'	kN/m ²	Effective cohesion
ϕ'	°	Effective friction angle
h_{water}	m	Groundwater level

Uncertainty characterization

Geotechnical units and parameters

In this project, the probability density functions of the random input parameters (see section B.5.2) are defined based on own experience as well as discussions with the client, according to **Table 51**.

Table 51. Probability density functions of random input parameters.

	parameter	distribution	unit	mean	standard deviation
Coarse alluvial	E	Gaussian	MN/m ²	12.5	1.5
	c'	Lognormal	kN/m ²	0.4	0.3
	φ'	Gaussian	°	33.5	0.9
Moraine	E	Gaussian	MN/m ²	25	3
	c'	Gaussian	kN/m ²	17.5	1.5
	φ'	Gaussian	°	31	0.6
Fluvio-glacial deposits	E	Gaussian	MN/m ²	22.5	1
	c'	Lognormal	kN/m ²	1	0.5
	φ'	Gaussian	°	33.5	0.9
Glaciolacustrine	E	Gaussian	MN/m ²	25	3
	c'	Gaussian	kN/m ²	17.5	1.5
	φ'	Gaussian	°	27	0.6

Loads, groundwater and pore pressure

Accordingly, for the ground water level h_{water} [m] as random variable, a Gaussian distribution with a mean value of 6,0 m and a standard deviation of 0,5 m is defined.

Model uncertainties

Not considered in this example.

Reliability analysis

Selection of reliability method(s)

Propagating uncertainties usually requires many model evaluations, typically when the simulation runs through the Monte-Carlo algorithm. When such an approach is chosen, the number of required runs depends on the probability of failure, and the size of the input set must be large for enough accuracy. This makes the Monte-Carlo method particularly inappropriate for time-consuming finite elements models. Surrogate models (or meta-models) are therefore used for replacing the “true” finite element model by an easy-to-evaluate model in the form of a polynomial analytical expression, like the Polynomial Chaos Expansion (or PCE), (see Marelli and Sudret, 2021).

In order to link the quantities of interest Y with the model inputs X , the truncated PCE which represents the approximation of the “true” finite element model $Y = M(X)$ is given by the following expression:

$$Y^{PCE} = \sum (y_a \Psi_a(X))$$

Equation 143.

where y_α are coefficients of Ψ_α , a polynomial orthogonal basis.

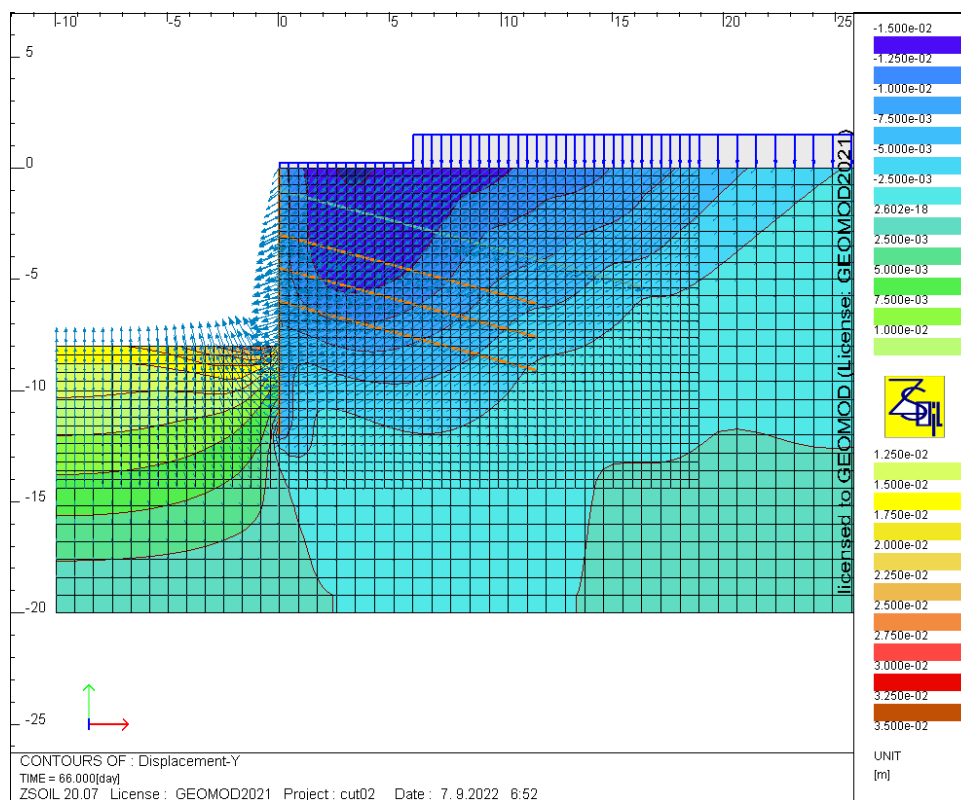
In this example, two Polynomial Chaos Expansions (PCE) surrogate models are built on 30 realisations of ZSWalls (a template of ZSOIL finite element computations), one for each quantity of interest, i.e., the maximal bending moment and the maximal settlement.

Figure 93 shows the vertical displacement field as well as displacement vectors for one of these 30 finite element realisations.

A criterion on an acceptable LOO (Leave One Out) error (see Allen, 1971) is selected in order to fix the number of samples computed with ZSOIL on which the PCE was built (here: 30).

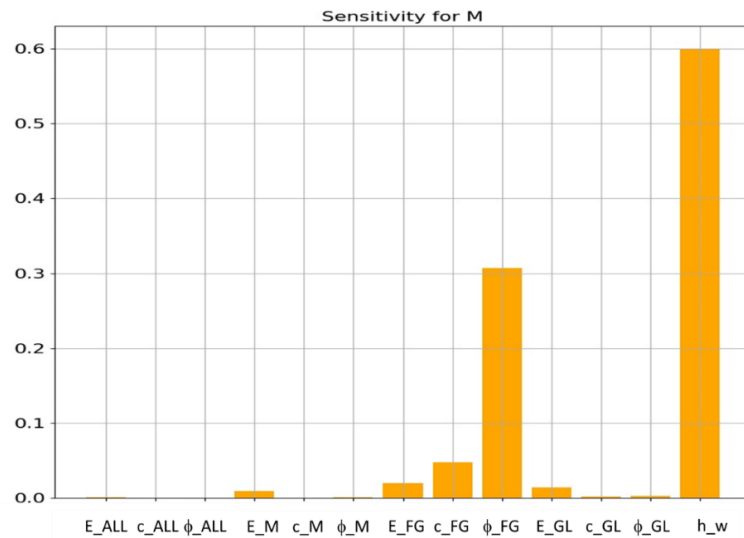
Then a sensitivity analysis, namely Sobol indices (see Marelli et al, 2021), is performed to evaluate the influence the random variables have on the limit state functions of **Equation 141** and **Equation 142**. **Figure 94** presents the results for the first quantity of interest (maximal bending moment). It indicates that the two most influent input parameters are the water table height and the glacio-lacustrine friction angle, respectively. It turns out that the same two input parameters are also most influencing on the second quantity of interest, which is the maximal settlement behind the wall.

Figure 93. Vertical displacement field and displacement vectors corresponding to one of the 30 finite element computations with ZSWalls.



Source: Authors' own work

Figure 94. Sobol indices illustrating the degree of sensitivity of the random input variables on the bending moment of the wall.



Source: Authors' own work

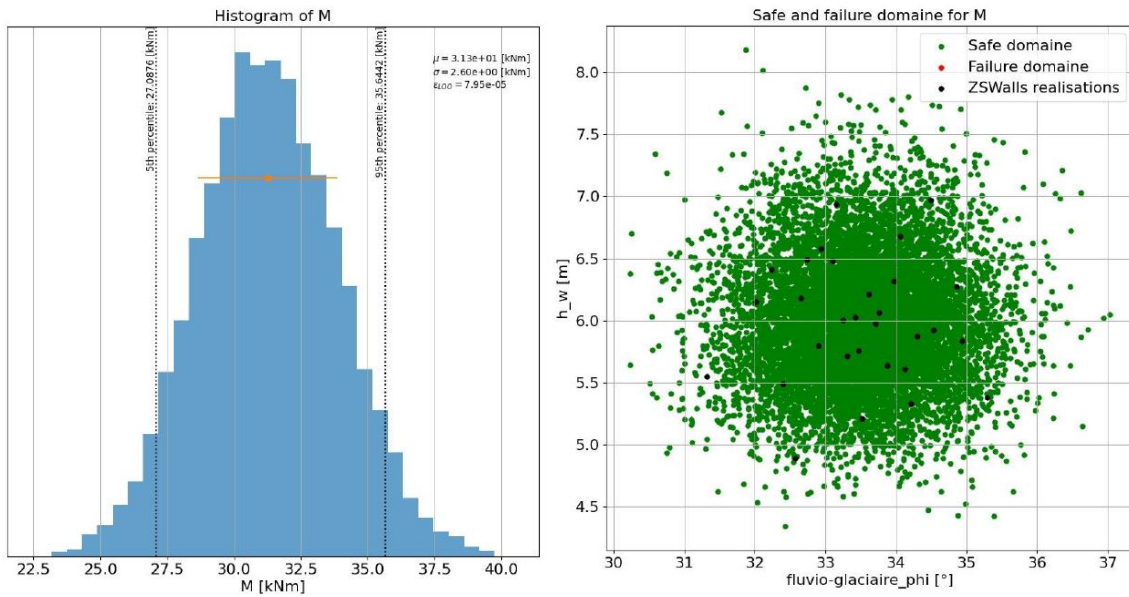
Finally, Monte-Carlo simulations are used on the PCE approximation in order to perform the reliability analyses described below, considering only the two most influential parameters. All other parameters, i.e. also the other random variables, are considered deterministic with their mean values.

Estimation of probability of failure

Figure 95 and **Figure 96** below summarize the results of the reliability analysis for both limit state functions $g_1(X)$ (**Equation 141**) and $g_2(X)$ (**Equation 142**). In each figure the histogram on the left represents the probability density function of the quantity of interest (maximum bending moment for $g_1(X)$ and maximum settlement for $g_2(X)$ computed on the 10.000 PCE points).

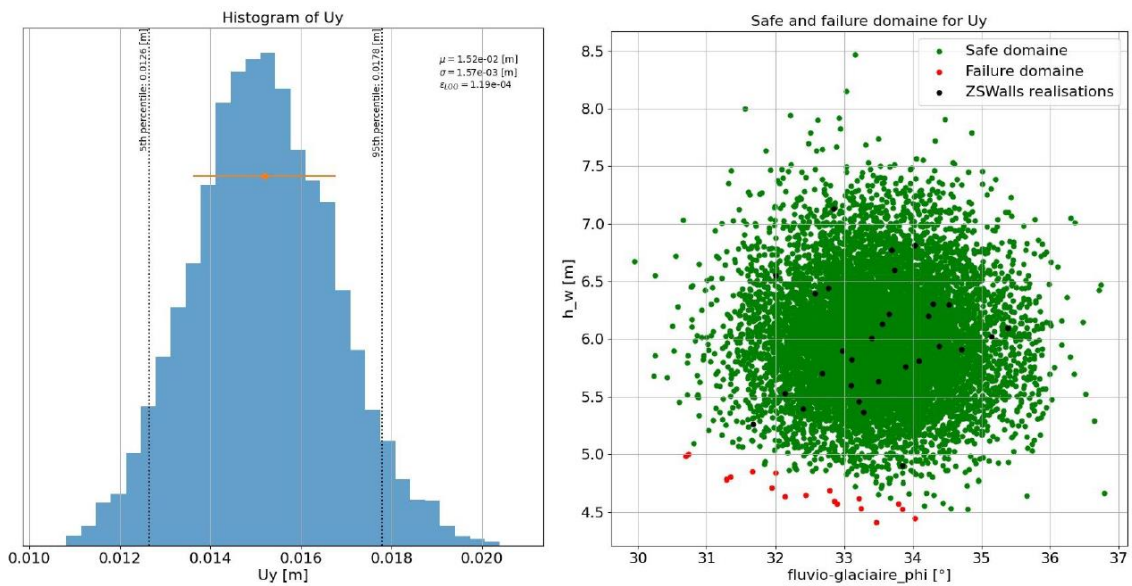
The diagram on the right in **Figure 95** and **Figure 96** show the 30 finite element samples, computed with ZSWalls in the water table height – glacio-lacustrine friction angle space, and the 10.000 PCE points in green (when $g_i(X) > 0$) or red (when $g_i(X) < 0$).

Figure 95. Results of the reliability analysis of limit state equation $g_1(X)$ – pdf of the maximum bending moment (left); PCE realisations based on FE samples depending on combinations of h_{water} and ϕ'_{FG} (right).



Source: Authors' own work

Figure 96. Results of the reliability analysis of limit state equation $g_2(X)$ – pdf of settlement (left); PCE realisations based on FE samples depending on combinations of h_{water} and ϕ'_{FG} (right)



Source: Authors' own work

Interpretation of results and convergence

It turns out that, under the chosen assumptions, the probability of exceeding the resistance of the sheet-pile wall is below 10^{-4} (= 1/number of Monte Carlo realisations), and the probability of exceeding a 20 mm settlement behind the wall is $1.8 \cdot 10^{-3}$.

This analysis gave the contractor confidence on ultimate limit state fulfilment, and helped him convince the upper neighbors that the risk of reaching significant settlements on their plot was very small.

Software Tools

For the finite element calculations ZSWalls (<https://www.zsoil.com/zswalls/>), a pre-processor for the finite element software ZSOIL, is used. The uncertainty quantification, sensitivity and reliability analyses is performed with UQLab (<https://www.uqlab.com>).

References

Allen, D. M. (1971) The prediction sum of squares as a criterion for selecting predictor variables. Technical Report Number 23, Department of Statistics, University of Kentucky, 1971 .

Marelli, S., Sudret, B. (2021) www.uqlab.com user manual – Polynomial chaos expansions, Tech. rep., Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, Switzerland, report # UQLab-V1.4-104.

Marelli, S., Lamas, C., Konakli, K., Mylonas, C., Wiederkehr, P. Sudret, B. (2012) www.uqlab.com user manual – Sensitivity analysis, Tech. rep., Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, Switzerland, report # UQLab-V1.4-106.

B.5.3 Propped embedded retaining wall

Authors: Stéphane Commend

Reference: -

Problem definition

Description

Geotechnique Suisse initiated a design challenge in 2021, based on problem „SAND 5“ in Probabilistic solutions for survey questions in “Are we overdesigning? – a survey of international practice”, Oct. 14 2020, prepared by TC304.

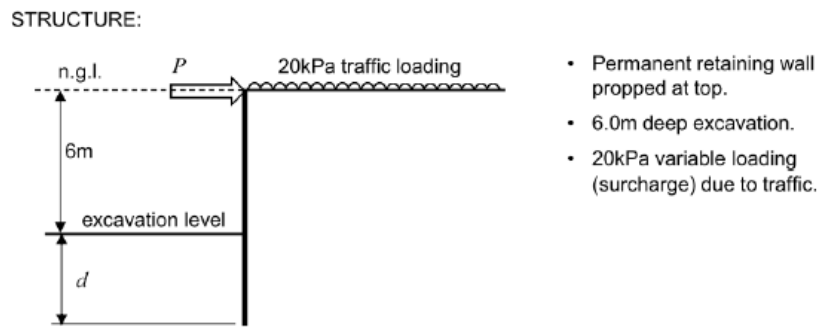
Homepage: <http://geotechnikschweiz.ch/?p=3400&lang=en>

The problem deals with a propped embedded retaining wall. This example aims at specifying the required depth of embedment d that satisfies that the probability of failure is lower than the target reliability (that has to be defined).

Available Data

A sketch of the retaining wall (see **Figure 97**) as well as information on grading, one SPT and five CPT tests are available.

Figure 97. Sketch of the retaining wall.



Source: TC 304 (2020)

Limit State function(s)

$$g(X) = SF - 1 < 0$$

Equation 144.

where SF is the global safety factor, obtained by dividing the cohesion c and the tangent of the friction angle ϕ' by an increasing value, until no more equilibrium can be achieved (c - ϕ reduction algorithm).

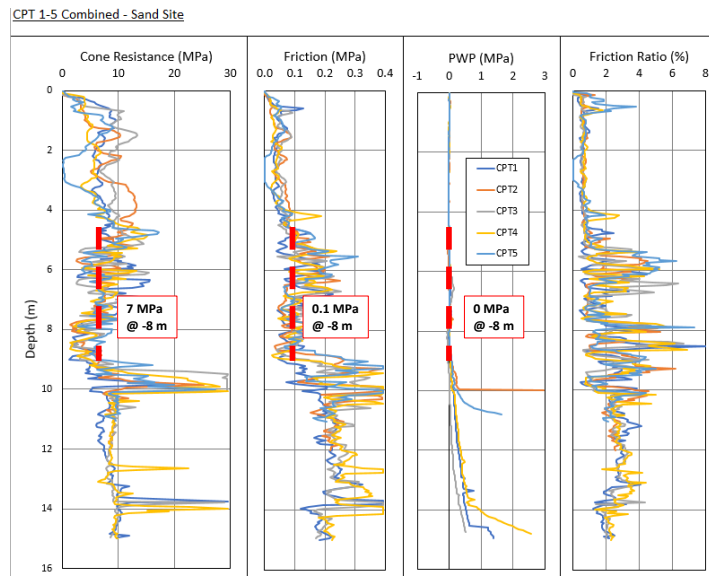
Target reliability

No target reliability was actually selected and/or determined.

Ground model

We use data provided in the 5 CPT tests, and selected a representative value for the cone resistance, the friction sleeve resistance and the pore water pressure (see **Figure 98**).

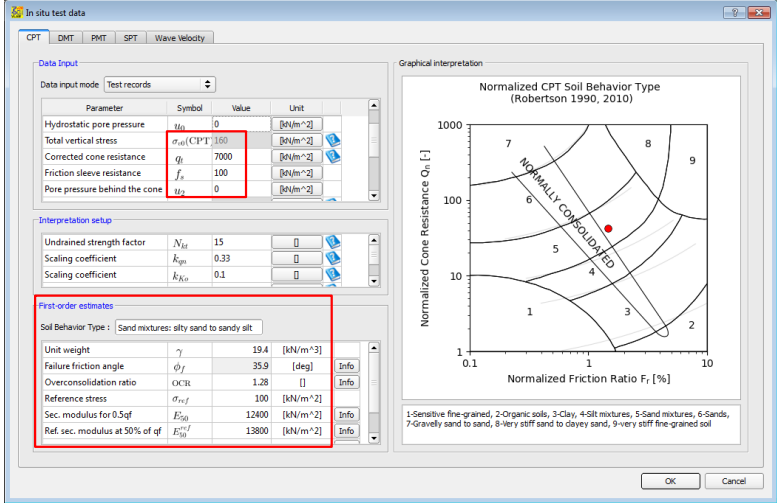
Figure 98. CPT 1-5 combined results and selected representative values.



Source: Authors' own work

Then, we use these values in order to determine the type of soil (sand mixture), select a constitutive model (Hardening Soil Model) and estimate its associated geotechnical parameters. For this we used ZSOIL v20.07 Virtual Lab, see **Figure 99**.

Figure 99. Investigated system of soldier pile wall.



Source: Authors' own work

Input Parameters (random and deterministic variables)

Table 52. Overview of deterministic and random variables within this example.

Deterministic Variable	Unit	Description
q	kN/m	Traffic load
Random Variable	Unit	Description
E	kN/m ²	Sand's elastic modulus
d	kN/m ²	Embedment depth (in the second approach)
φ	°	Sand's friction angle

Uncertainty characterization

Geotechnical units and parameters

- E: lognormal, mean = 14000 kN/m², stdev = 2800 kN/m²
- φ: lognormal, mean = 36°, stdev = between 3.6° and 7.2°

The mean values are directly taken from the ground model definition (see section B.5.3.1). The standard deviations derive from the assumption that COV(E) = 20% and COV(φ) = 10% to 20%, according to literature.

Loads, groundwater and pore pressure

In this example, the traffic load is considered to be deterministic and equal to 20 kPa.

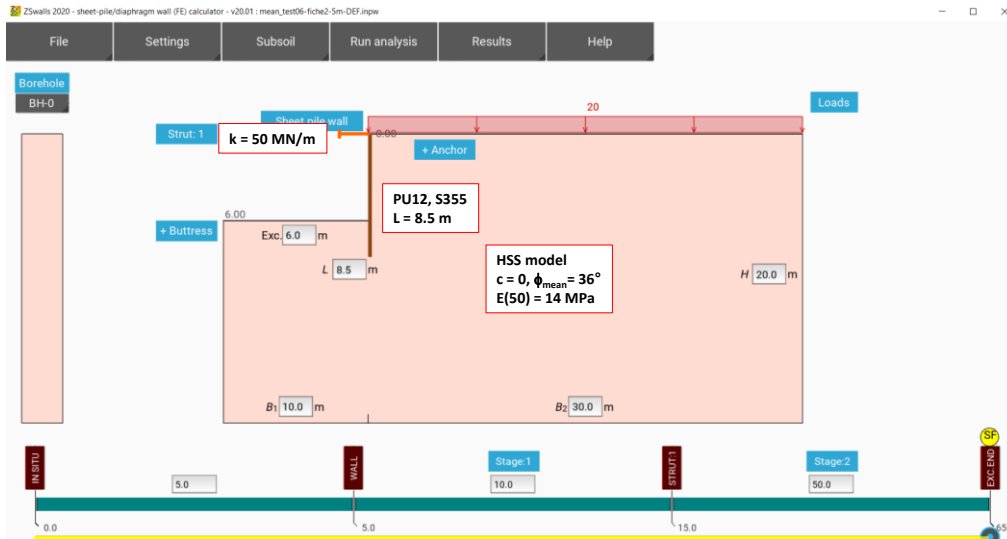
Model uncertainties

Not considered in this example.

Reliability analysis

We first construct a ZSWalls model, as shown in **Figure 100**. It has been admitted to use a PU12 sheet-pile wall and a strut with $k = 50 \text{ MN/m}$.

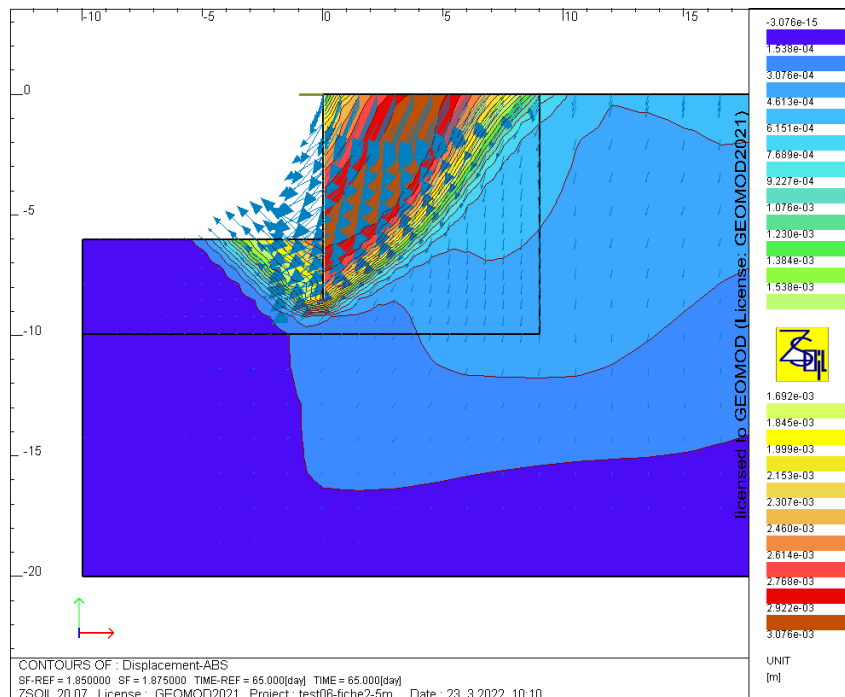
Figure 100. ZSWalls model.



Source: Authors' own work

The principal result of the finite element simulation is the global safety factor SF of the structure, determined by c - ϕ reduction, as well as the potential failure mechanism, see **Figure 101**.

Figure 101. Failure mechanism associated with the global safety factor, for one of the ZSWalls realisations.



Source: Authors' own work

First approach

We selected two different embedment depths ($d = 2.5$ m or 3.0 m), and chose $COV(\varphi)$ between 10% and 20%. We then built Polynomial Chaos Expansions (PCE, see Annex B.5.2) surrogates on 25 realisations of ZSWalls and, then, on 100'000 PCE points, we used Monte Carlo simulations to perform the reliability analyses.

A criterion on an acceptable LOO error (see Allen, 1971) was selected in order to fix the number of samples computed with ZSOIL on which the PCE was built.

Second approach

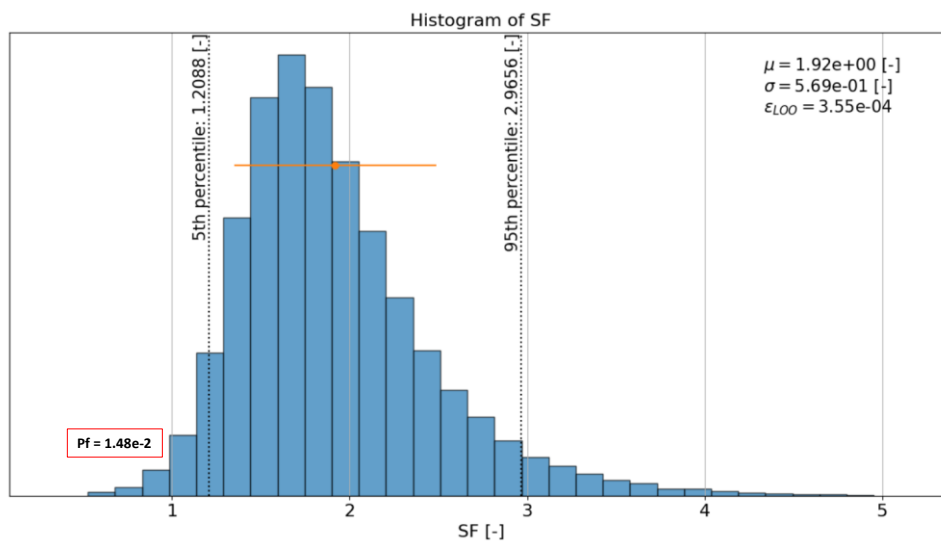
In addition to E and φ , we include the embedment depth d in our probabilistic input, and build our PCE on E , φ , and d . For this case, $COV(\varphi)$ is assumed to be equal to 20%, and the number of samples on which the PCE is built is equal to 200.

Estimation of probability of failure

First approach

Figure 102 illustrates the probability density function of the safety factor (SF) for $d = 2.5$ m and $COV(\varphi) = 20\%$. **Table 53** summarizes probabilities of failure and LOO errors computed with various $COV(\varphi)$ and embedment depths.

Figure 102. Histogram of safety factor and associated P_f for $d = 2.5$ m and $COV(\varphi) = 20\%$.



Source: Authors' own work

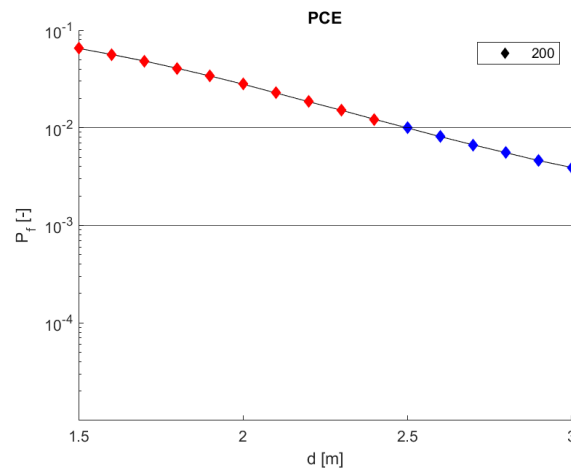
Table 53. Results for the first approach - probabilities of failure and LOO errors.

d [m]	COV(φ) [%]	P_f (SF < 1) [-]	Metamodel [-]	LOO error [-]
2.5	20	$1.48 \cdot 10^{-2}$	PCE(25)	$3.60 \cdot 10^{-4}$
2.5	15	$8.60 \cdot 10^{-4}$	PCE(25)	$8.30 \cdot 10^{-5}$
2.5	10	$1.00 \cdot 10^{-5}$	PCE(25)	$1.30 \cdot 10^{-4}$
3.0	20	$1.37 \cdot 10^{-3}$	PCE(25)	$1.35 \cdot 10^{-2}$

Second approach

As d is part of the PCE, it is now possible to plot $P_f = f(d)$ – see **Figure 103**, where for visualization purposes only the red dots stand for $P_f > 10^{-2}$, and the blue dots stand for $P_f < 10^{-2}$ – and then select d according to the target reliability (which is, at this point, unknown).

Figure 103. Probability of failure with respect to the embedment depth, for $\text{COV}(\varphi) = 20\%$.



Source: Authors' own work

Interpretation of results and convergence

For the first approach, for $d = 2.5$ m, P_f varies between 10^{-2} and 10^{-5} , depending on $\text{COV}(\varphi)$. $\text{COV}(\varphi)$ could/should therefore be updated using (direct or inverse) Bayesian techniques.

For the second approach, with $\text{COV}(\varphi) = 20\%$, the PCE built on 200 samples gives a probability of failure P_f in the 10^{-2} order of magnitude for $d = 2.5$ m, and in the 10^{-3} order of magnitude for $d = 3.0$ m, which confirms the results of the first approach.

Software Tools

For the finite element calculations ZSWalls (<https://www.zsoil.com/zswalls/>), a pre-processor for the finite element software ZSOIL, was used.

For the uncertainty quantification, sensitivity and reliability analyses, UQLab (<https://www.uqlab.com>) was used.

For the parameters estimation, based on CPTU tests ZSOIL v20.07 Virtual Lab (https://www.zsoil.com/zsoil_manual_2018/Rep-VirtualLab.pdf) was used.

References

Allen, D. M. (1971) The prediction sum of squares as a criterion for selecting predictor variables. Technical Report Number 23, Department of Statistics, University of Kentucky, 1971

TC 304 (2020). Probabilistic solutions for survey questions in “Are we overdesigning? – a survey of international practice”

B.6. Reinforced fill structures

B.6.1 Geogrid reinforced soil wall

Authors: Pietro Rimoldi

Reference: P. Rimoldi, P. Pezzano, E. Zannoni (2019). Reliability analysis of internal and external stability of geosynthetic reinforced soil retaining walls. Proc. GeoAmericas 2020, Rio de Janeiro, Brazil.

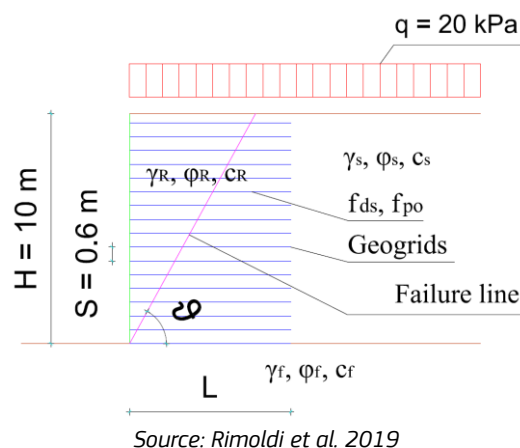
Problem definition

Description

This example shows a framework for the reliability design of a Geosynthetic Reinforced Soil Wall (GRSW) to explicitly address uncertainties in the design process and account for the actual safety and reliability level of a given design.

The following sketch (**Figure 104**) is an example of the system geometry of the design problem:

Figure 104. Example of the system geometry of a design problem.



Available Data

The data of the problem include wall geometry, soil parameters, interface parameters, geosynthetic reinforcement parameters.

The following **Table 54** lists all the design parameters with an example of selected values for a specific project. For GRSW the length L and the ultimate tensile strength of geogrids T_{ult} are the output of the design, since usually the vertical spacing is given by the facing or the geometry and it is not considered as stochastic. Hence the goal of the proposed procedure is to evaluate the values of L and T_{ult} producing 0.1% probability of failure (or any set probability of failure) for each of the five identified limit states (see next section).

The first task is to identify the limit state functions for five failure mechanisms of internal and external stability, which are used to calculate margins of safety in terms of probability of failure through Monte Carlo simulations, where all parameters can be set as either deterministic (with no associated variability) or probabilistic (with associated variability).

The final task is to perform probabilistic analyses, which can be carried out repeatedly by changing the deterministic and the stochastic parameters, and/or their associated variability. Results can provide a useful decision-making tool for preliminary design of GRSW based on target reliability levels.

Table 54. Input data for one run of Monte Carlo simulation with example of selected values for a specific project.

Deterministic Variable	Unit	Description	Value	Variability [%]
Wall geometry				
H	m	Height of wall	10.00	0.00
L	m	Length of georid	5.60	0.00
s	°	Vertical spacing of geogrids	0.60	0.00
N _{GG}	-	Number of geogrids	17.00	0.00
q	kPa	Uniform surcharge	20.00	0.00
Interface parameters				
f _{ds}	m	Direct shear factor	0.95	0.00
f _{po}	m	Pullout factor	0.95	0.00
Random Variable	Unit	Description	Value	Variability [%]
Soil parameters				
γ _R	kN/m ³	Saturated unit weight of the reinforced soil	18.00	5.00
φ _R	°	Friction angle of the reinforced soil	32.00	5.00
c' _R	kPa	Cohesion of the reinforced soil	0.00	20.00
γ _S	kN/m ³	Saturated unit weight of the backfill soil	18.00	20.00
φ _S	°	Friction angle of the backfill soil	30.00	20.00
c' _S	kPa	Cohesion of the backfill soil	0.00	20.00
γ _f	kN/m ³	Saturated unit weight of the foundation soil	18.00	20.00
φ _f	°	Friction angle of the foundation soil	28.00	20.00
c' _f	kPa	Cohesion of the foundation soil	0.00	20.00
Geosynthetic reinforcement				
Description of Geosynthetic Reinforcement: Bonded polyester geogrid				
T _{ult}	kN/m ³	T _{ult} of geogrids	70.00	1.00
RF _{cr}	-	Reduction Factor for creep	1.39	1.00
RF _{ID}	-	Reduction Factor for installation damage	1.10	1.00
RF _{ch}	-	Reduction Factor for chemical damage	1.20	1.00
RF _b	-	Reduction Factor for biological damage	1.00	0.00

Limit State function(s)

The limit state functions $G(x)$ which can be derived to evaluate the performance of GRSW against each failure mode for external and internal stability are represented by the equations providing the difference of resisting and active forces, stresses or moments for each of the considered failure modes: $G(x) = (R - A)$, where R represents the resisting forces, stresses or moments, and A represents the active forces, stresses or moments.

Sliding Failure

For GRSW the critical sliding usually occurs along the geosynthetic reinforcement at the base, when the friction between the fill and the geosynthetic is not enough to compensate for the external load, which causes the retaining wall to slide. With reference to the scheme in **Figure 105** (a), the limit state function for sliding failure is:

$$G_{ds} = L \cdot f_{ds} \cdot \tan \varphi_f \cdot (\gamma_R \cdot H + q) - (0.5 \cdot \gamma_s \cdot H^2 + q \cdot H) \cdot \tan^2(45 - \varphi_s/2)$$

Equation 145.

where:

φ_f, φ_s = friction angle of foundation soil and back soil, respectively (deg)

γ_R, γ_s = unit weight of reinforced soil and back soil, respectively (kN/m³)

f_{ds} = direct shear factor (-)

H = height of wall (m)

L = length of reinforcement (m)

q = uniformly distributed surcharge (kPa).

Note that $\tan \varphi_f$ is used in **Equation 145** assuming that the friction angle of the foundation soil, φ_f , is lower than the friction angle of the reinforced soil, φ_R .

Overturning Failure

Overturning occurs when the soil thrust behind the GRSW body is great enough to offset the retaining wall by rotation around the wall toe. With reference to the scheme in **Figure 105** (b), the limit state function for overturning is given by the difference of resisting and active overturning moments around the toe:

$$G_{ot} = 0.5 \cdot L \cdot (\gamma_R \cdot L \cdot H + q \cdot L) - (1/6 \cdot \gamma_s \cdot H^3 + 0.5 \cdot q \cdot H^2) \cdot \tan^2(45 - \varphi_s/2)$$

Equation 146.

Bearing Resistance Failure

Bearing resistance failure occurs when the subgrade soil beneath the reinforced soil wall fails under shear due to overloading or insufficiently constructed subgrades. With reference to the scheme in **Figure 105** (c) and using the Terzaghi formula for general shear failure of foundations, the limit state function for bearing resistance is given by the difference of the bearing capacity of foundation soil and the vertical stress on the foundation:

$$G_{bc} = 0.5 \cdot \gamma_f \cdot L \cdot N_\gamma - (\gamma_R \cdot H + q)$$

Equation 147.

with:

$$N_{\gamma} = 2 \cdot \tan[e^{\pi \cdot \tan \varphi_f} \cdot \tan^2(45 + \varphi_f/2) + 1]$$

Equation 148.

where:

φ_f = friction angle of foundation soil (deg)

γ_f = unit weight of foundation soil (kN/m³)

N_{γ} = bearing capacity factor (-)

Note that the Terzaghi formula can be replaced by more complex formulas (Meyerhoff, Brinch-Hansen, etc.) in the *limit* state function (**Equation 147**), while the framework remains valid.

Pullout Failure

Pullout failure occurs when the geogrid does not have sufficient length to resist the soil thrust in the influence area of each layer, thus causing failure by pullout. With reference to the scheme in **Figure 105** (d), the critical condition usually occurs for the top geogrid, which has the lowest vertical stress producing the pullout shear stresses needed for anchorage of the geogrid in the fill behind the potential failure surface.

The anchorage length L_a is defined as the total geogrid length L minus the length L_e from the face to the point where the potential failure surface intersects the geogrid. Assuming that the geogrid is an extensible reinforcement, the failure surface coincides with the Rankine failure surface, identified by a line passing from the toe and inclined at an angle ϑ . The limit state function for pullout is therefore given by the difference of the pullout resisting force developed along the anchorage length of the top geogrid and the soil thrust in the influence area of the top geogrid:

$$G_{po} = F_{po} - (F_{a-po} + F_{q-po})$$

Equation 149.

with:

$$\begin{aligned} F_{po} &= 2 \cdot \tau_{po} \cdot L_a = 2 \cdot \tau_{po} \cdot (L - L_e) \\ &= 2 \cdot f_{po} \cdot (\gamma_R \cdot z + q) \cdot (L - (H - z)) / \tan \vartheta \end{aligned}$$

Equation 150.

$$F_{a-po} = 0.5 \cdot \gamma_R \cdot \tan^2(45 - \varphi_R/2) \cdot (z + S/2)^2$$

Equation 151.

$$F_{q-po} = q \cdot (z + S/2) \cdot \tan^2(45 - \varphi_R/2)$$

Equation 152.

$$\vartheta = 45 + \varphi_R/2$$

Equation 153.

$$z = H - (N_{GG} - 1) \cdot S$$

Equation 154.

where:

F_{po} = pullout resisting force (kN/m)

F_{a-po} = active horizontal force produced by the self weight of the fill soil for pullout (kN/m)

F_{q-po} = active horizontal force produced by the surcharge for pullout (kN/m)

τ_{po} = pullout shear stresses on both sides (top and bottom) of the geogrid (kPa)

f_{po} = pullout factor (-)

ϕ_R = friction angle of reinforced soil (deg)

L = total length of the geogrid (m)

L_a = anchorage length of the geogrid (m)

L_e = length of geogrid between the face and the failure surface (m)

ϑ = inclination of the failure surface to the horizontal (deg)

z = depth of the geogrid from the top of the wall (m)

N_{GG} = total number of geogrid layers (-)

S = uniform vertical spacing of geogrids (m).

Note that the variable surcharge q is both favourable for safety, when calculating F_q , and unfavourable, when calculating F_{po} . Hence, strictly speaking, G_{po} should be calculated with both applying and not applying q . Nevertheless, the situation with q applied is usually the critical one.

Tensile Failure

Tensile failure occurs when the tensile strength of the geogrids is not enough to withstand the forces applied by the thrust of the soil. With reference to the scheme in **Figure 105** (e), for GRSW the critical condition usually occurs for the first geogrid above the toe, which has to withstand the highest horizontal stresses multiplied by its influence area. The limit state function for tensile failure is therefore given by the difference of the design strength T_D of this geogrid and the active horizontal force produced by the fill soil and the surcharge on the influence area of the geogrid:

$$G_{tf} = T_D - (F_{a-ts} + F_{q-ts})$$

Equation 155.

with:

$$T_D = \frac{T_{ult}}{RF_{cr} \cdot RF_{id} \cdot RF_c \cdot RF_b}$$

Equation 156.

$$F_{a-ts} = 0.5 \cdot \tan^2(45 - \varphi_R/2) \cdot [(H - S/2)^2 - (H - 3/2 \cdot S)^2]$$

Equation 157.

$$F_{a-ts} = 0.5 \cdot \tan^2(45 - \varphi_R/2) \cdot [(H - S/2)^2 - (H - 3/2 \cdot S)^2]$$

Equation 158.

where:

T_D = design tensile strength of the geogrid (kN/m)

F_{a-tf} = active horizontal force produced by the self-weight of the reinforced soil for tensile failure (kN/m)

F_{q-tf} = active horizontal force produced by the surcharge for tensile failure (kN/m)

RF_{cr} = Reduction Factor for tensile creep of geogrids (-)

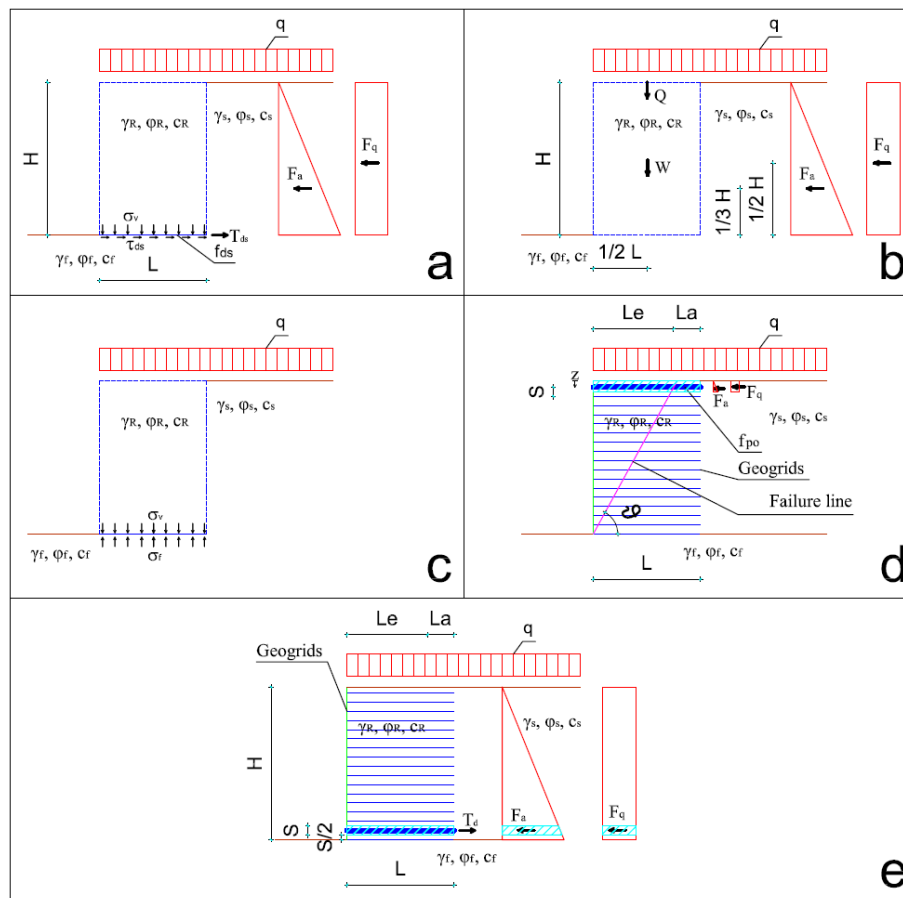
RF_{id} = Reduction Factor for installation damage of geogrids (-)

RF_c = Reduction Factor for chemical damage of geogrids (-)

RF_b = Reduction Factor for biological damage of geogrids (-)

Note that RF_b can always be assumed equal to 1.0 for geosynthetic reinforcement. Hence, this parameter can be considered as a deterministic value, while the other RFs can have a variability in respect to their nominal value and, therefore, these are considered as stochastic values.

Figure 105. Schemes of the limit states for internal and external stability of GRSW: a. sliding; b. overturning; c. bearing resistance; d. pullout; e. tensile failure.



Source: Rimoldi et al. 2019

Target reliability

Referring to the relationship between probability of failure P_f and reliability index β established in **Equation 46**, Bathurst (2018) argues that values of $\beta = 2.33$ and 3.09 correspond to probabilities of failure of $1/100$ and $1/1000$, respectively. The smaller β value is recommended as the target minimum reliability index for internal limit state design and LRFD calibration for GRSW. This value may appear small but GRSW walls are highly strength-redundant systems. In other words, if one reinforcement layer fails, other layers can compensate and thus system failure is unlikely. If a reinforcement layer is designed to just satisfy a target reliability of $\beta = 2.33$, the corresponding factor of safety can be as high as 1.70 . Anyway, in the referenced paper the value $\beta = 3.09$ is assumed in favour of safety, which corresponds to 0.1% probability of failure for each of the five limit states and is assumed as the target reliability.

Ground Model

The ground model includes three types of soil: the reinforced fill, the backfill (behind the reinforced soil block), and the foundation soil (see **Figure 104**).

All three soils can be considered as frictional only (effective cohesion $c' = 0$) or with both effective friction angle and cohesion. Note that under the present development of the method, in case of a

cohesive soil layer a unique value of cohesion for the whole layer and for any elapsed time shall be selected. The required parameters are listed in **Table 54**.

Input Parameters (random and deterministic variables)

Uncertainty in the true magnitude of load and resistance terms for the limit states introduced above is solely due to the estimation of the friction angle (ϕ) and unit weight (γ) of soils, and of the design tensile strength of geogrids T_D ; nevertheless, all parameters in the above defined limit state functions could be considered as stochastic variables.

Specifically, random values of $\phi_R, \phi_s, \phi_f, \gamma_R, \gamma_s, \gamma_f, T_{ult}, RF_{cr}, RF_{id}, RF_c$ are generated from probability distributions for these parameters within their set variabilities (in % of the nominal values) for a total of N times, and each set of values are used to compute values of the five limit state functions G.

An example of input values is shown in **Table 54**: if the variability is equal to 0 here, the parameter is considered as deterministic, while if the variability is > 0 , the parameter is considered as stochastic.

The following considerations apply in this example:

- The mean value of the Gaussian distribution is a conservative assumption; yet there is no better way to set the mean values at pre-design stage.
- The mean value has to be set by the designer, based on available data from geotechnical investigation, experience, and national standards.
- Variability is assumed as a percentage of the nominal value, since this is the easiest and the most common way to define variability for professional engineers.
- The probability distribution of all stochastic parameters is assumed to be normal, since at pre-design stage usually there are not enough data to assume more complex distributions.

Uncertainty characterization

Geotechnical units and parameters

The following considerations apply in this example:

- Unlike naturally deposited soils, the fill soil in a GRSW is an engineered material and therefore the variabilities in soil unit weight γ_R are low. According to Zannoni (2016) the variability of the friction angle ϕ_R can be as high as 10% according to a low coefficient of variation. For the back soil and the foundation soil the variabilities in soil unit weights (γ_s, γ_f) and friction angles (ϕ_s, ϕ_f) can be very large. Bond & Harris (2008) provide the indications reported in **Table 55**.

Table 55. Coefficient of variation (COV) of geotechnical and man-made-materials.

Material	Parameter		COV
Soil	coefficient of shearing resistance	$\tan(\phi)$	5-15%
	effective cohesion	c'	30-50%
	undrained strength	c_u	20-40%
	coefficient of compression	m_v	20-70%
	weight density	γ	1-10%

Source: Bond and Harris 2008

- The direct shear factor f_{ds} and the pullout factor f_{po} are usually obtained from direct shear and pullout laboratory tests on the specific geosynthetic and standard and/or site-specific soil. Their value can be set with no variability or with an associated variability, in which case also f_{ds} and f_{po} become stochastic parameters in Monte Carlo simulations. The decision whether to consider also f_{ds} and f_{po} as stochastic parameters, and in this case the variability to be set, is left to the designer. In any case the maximum value of f_{ds} and f_{po} shall be 1.0, since the interface cannot afford a higher friction angle than the adjacent soil.
- All the stochastic parameters are assumed to vary, according to the set variabilities, with higher or lower values than the nominal value, except for the friction angle of foundation soil φ_f , which (due to the assumed normal distribution) can take only lower values than the nominal value. Otherwise the resulting extremely large variability of N_v can make the standard deviation of the limit state function G_{bc} in **Equation 147** larger than its average value. Hence, for the purpose of this method, in the case of φ_f the nominal value shall be assumed as the maximum expected value.
- Variability is assumed as a percentage of the nominal value, since this is the easiest and the most common way to define variability for professional engineers.

The probability distribution of all stochastic parameters is assumed to be the normal distribution, since at pre-design stage usually there are not enough data to assume more complex distributions.

Loads, groundwater and pore pressure

The only considered external load is the uniformly distributed surcharge on top of the structure; it can be considered either as deterministic or stochastic. The surcharge and the self weight of soils produce horizontal thrusts on the reinforced block and vertical stresses on the base of the reinforced block. Neither groundwater nor pore pressure is considered.

Model uncertainties

The above defined limit state functions are based on the so-called Tie-back wedge method for the design of GRSW, where the failure line coincides with the Rankine failure line, with an inclination of $\vartheta = (45 + \varphi_R / 2)$ to the horizontal (see **Figure 104**).

This is an internationally accepted method for extensible reinforcement, like geogrids and geotextiles.

The model uncertainties are related to this Tie-back wedge method, and in particular to the value of ϑ . At present, model uncertainties cannot be evaluated due to a lack of data for this type of structures. Hence, they have not been considered in this example.

Reliability analysis

Selection of reliability method(s)

The probability of failure of each limit state can be computed using the above limit state functions, by producing N sets of values using Monte Carlo simulations.

Specifically, random values of φ_R , φ_s , φ_f , γ_R , γ_s , γ_f , T_{ult} , RF_{cr} , RF_{id} , RF_c , are generated from probability distributions for these parameters within their set variabilities (in % of the nominal values) for a

total of N times, with $N \geq 10,000$, and each set of values are used to compute values of the five limit state functions G.

Estimation of probability of failure

For GRSW the length L and the ultimate tensile strength of geogrids T_{ult} are the output of the design, since usually the vertical spacing is given by the facing or the geometry and it is not considered as stochastic.

The goal of the proposed method is to evaluate, through N Monte Carlo iterations, the values of L and T_{ult} producing 0.1% probability of failure (or any set probability of failure) for each of the five identified limit states. The procedure is the following:

- Set initial values of the length L and of the ultimate tensile strength of geogrids T_{ult} ;
- perform N Monte Carlo iterations, computing for each iteration the values of the five limit state functions G (see **Table 56**);
- as shown in **Table 57**, using the Excel function NORM.INV, calculate the value of each G function corresponding to 0.1% probability of failure (that is, corresponding to 0.1% probability that $G \leq 0$), which is assumed to be a reasonable engineering limit for GRSW, as discussed above;
- for each of the five limit state functions, modify the length L and/or the ultimate tensile strength of geogrids T_{ult} , and run the N Monte Carlo iterations; vary the values of L and/or T_{ult} by trial and error until all the G values with 0.1% probability become approx. equal to 0 (see **Table 57**).

Following this procedure, the values of L and T_{ult} producing 0.1% probability of failure for each of the five limit states can be evaluated. Obviously, the probability of failure can be set either larger or smaller than 0.1%.

Table 56. Example of Monte Carlo iterations.

Iteration Nr	Y_R	Φ_R	Y_S	Φ_S	Y_I	Φ_I	T_{ult}	RF _D	RF _{CH}	T_D	G_{ds} (kN/m)	G_{ot} (kN/m-m)	G_{bc} (kPa)	G_{po} (kN/m)	G_{is} (kN/m)
1	18.572	32.036	21.110	32.271	15.125	35.086	67.274	1.091	1.203	36.705	891.1	273.903	641.08	12.681	2.727
2	18.766	30.752	22.818	28.853	22.675	25.927	68.152	1.099	1.203	37.002	421.1	234.193	115.887	5.023	0.839
3	18.219	32.317	20.863	21.908	15.766	21.954	67.744	1.099	1.199	36.730	150.2	111.287	-73.791	14.27	3.491
4	18.354	31.066	16.393	35.819	17.039	23.234	68.110	1.106	1.207	37.033	502.8	300.999	-37.441	6.865	2.936
5	18.473	32.173	19.649	37.377	21.322	26.957	68.161	1.109	1.195	36.906	627.8	330.463	147.89	13.476	4.295
6	18.680	31.359	22.628	36.102	17.483	24.089	67.679	1.111	1.193	36.770	470.1	322.77	-14.578	8.66	1.981
7	17.184	33.484	17.525	36.837	19.281	24.794	67.204	1.103	1.194	36.397	511.0	288.427	42.366	20.78	6.663
8	17.639	32.543	19.277	25.676	16.837	26.086	67.014	1.098	1.190	36.576	386.6	156.169	49.372	15.464	4.885
9	17.860	30.462	14.447	21.671	18.125	20.674	67.119	1.095	1.193	36.634	295.1	97.14	-75.203	3.215	1.488
10	18.690	33.044	23.318	21.585	18.143	22.227	67.306	1.100	1.190	36.046	113.4	118.95	-53.97	18.673	4.303
11	18.634	33.181	22.498	36.047	15.860	25.536	66.983	1.102	1.196	36.884	524.8	292.404	7.764	19.462	4.379
12	17.753	31.484	17.387	34.026	19.305	21.283	67.950	1.098	1.205	37.272	375.9	272.345	-54.403	9.276	4.52
13	18.725	31.234	15.307	36.944	18.969	26.983	67.084	1.092	1.192	37.203	688.7	392.865	107.627	7.917	3.175
14	18.096	32.051	15.357	21.601	17.743	21.819	67.590	1.109	1.189	36.870	261.6	102.566	-50.168	12.681	2.688
15	17.815	32.664	20.662	30.259	21.815	22.862	68.004	1.102	1.201	36.776	320.0	226.621	3.65	16.214	2.833
16	18.317	33.251	12.901	23.091	15.943	24.228	67.354	1.099	1.199	36.564	496.2	194.007	24.41	19.78	4.233
17	17.611	30.558	19.971	25.403	22.694	27.264	67.188	1.097	1.195	36.866	411.1	151.168	196.153	3.772	2.54
18	18.097	33.305	14.752	23.637	17.232	19.612	67.064	1.107	1.199	36.527	220.7	136.79	-99.871	20.02	3.522
19	17.455	30.490	17.795	24.358	16.907	21.678	67.569	1.093	1.204	37.284	227.7	130.383	-62.078	3.361	5.072
20	18.353	32.917	22.906	24.095	13.010	24.139	67.545	1.109	1.198	36.807	238.0	151.421	-59.472	17.832	4.41

Source: Authors' own work

Table 57. Example of trial-and-error calculation.

Variable parameter	L_{ds}	L_{ot}	L_{bc}	L_{po}	T_{ult}
Value	7.00	2.08	7.53	5.68	88.00
Stochastic Parameter	G_{ds} [kN/m]	G_{ot} [(kN/m)·m]	G_{bc} [kPa]	G_{po} [kN/m]	G_{ts} [kN/m]
Average	178.115	111.992	244.501	13.070	5.202
Dev. Stand.	77.2373	48.1542	105.4172	5.6598	2.2356
Fractile	0.0100	0.0100	0.0100	0.0100	0.0100
Fractile Value	-1.566	-0.032	-0.737	-0.096	0.001

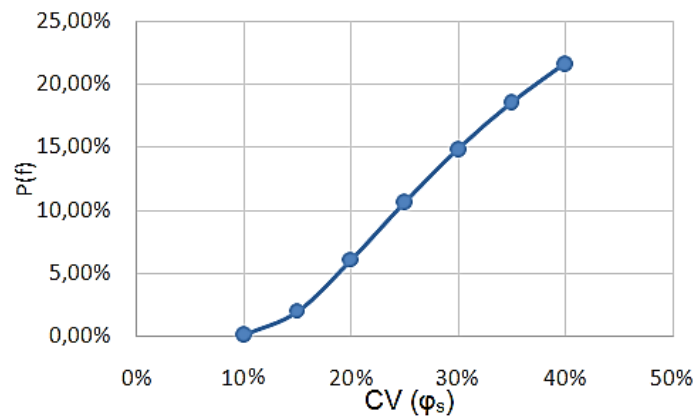
Source: Authors' own work

Interpretation of results and convergence

Results can be presented as charts of the variation of $P(f)$ for each failure mechanism with the CV of any stochastic parameter. Examples are shown in **Figure 106** and **Figure 107**.

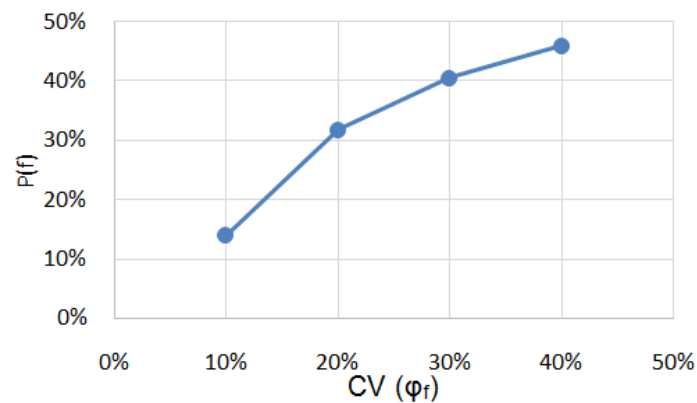
Since the output is the length and tensile strength of reinforcement, results can also be presented as charts of the variation of the tensile strength and the length with the CV of any stochastic parameter. Examples are shown in **Figure 108** and **Figure 109**.

Figure 106. Variation of $P(f)$ for sliding along the base with CV (φ_s)



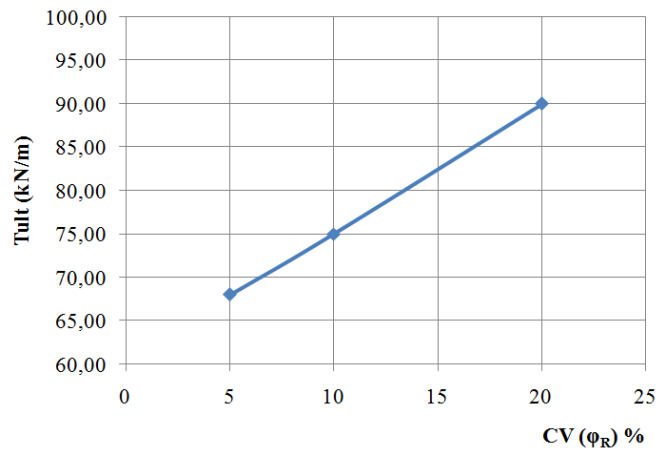
Source: Rimoldi et al. 2019

Figure 107. Variation of $P(f)$ for sliding along the base with CV (φ_f).



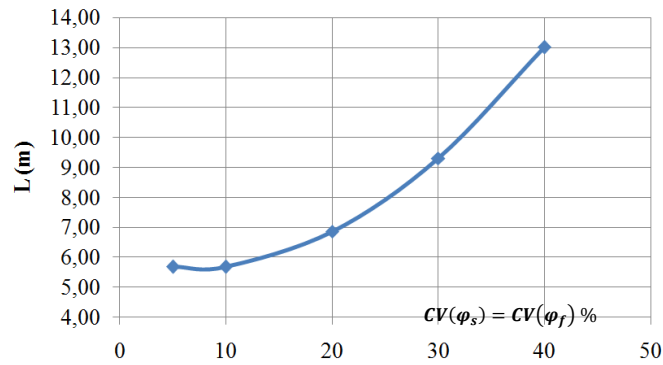
Source: Rimoldi et al. 2019

Figure 108. Variation of the tensile strength with $CV(\varphi_R)$ while $CV(\varphi_s) = CV(\varphi_f) = 20\%$.



Source: Rimoldi et al. 2019

Figure 109. Variation of the length L with $CV(\varphi_s) = CV(\varphi_f)$ when $CV(\varphi_R) = 5\%$.



Source: Rimoldi et al. 2019

Comparison of estimated reliability with target

An example calculation for a 10 m high GRSW is included in the referenced paper by Rimoldi et al (2020). A probabilistic analysis using a Monte Carlo simulation with 10.000 iterations was applied to the model considering the variation of the stochastic parameters as shown in **Table 54** for a specific run. The Monte Carlo simulation is used to compute the probability of failure and the values of length and tensile strength of geogrids associated with the probability of failure of 0.1%, for a given set of combination of the values of the stochastic parameters. An example of the first 20 Monte Carlo iterations is shown in **Table 56**. An example of results of a trial-and-error calculation is shown in **Table 57**.

The Monte Carlo simulations can be used for a sensitivity analysis with respect to design optimization, and/or for evaluating the probability of failure associated with a given variability of one stochastic parameter while all other parameters are set to given nominal values. For example, if only the friction angle of back soil φ_s is considered as a stochastic parameter and all other parameters are set with the nominal values in **Table 54**, the probability of failure P_f for sliding along the base varies from 0.17% to 21.68% when the variability, expressed as coefficient of variation CV (%), of the friction angle of back soil varies from 10% to 40% from its nominal value in **Table 54** (**Figure 106**). In terms of reliability β , the design moves from above average which

has a reliability index of 2.94 to an index of 0.79 which is a very hazardous design. These results would suggest to increase the number of geotechnical tests on the back soil for reducing the variability of the nominal value of φ_s to less than 10%.

Using the same procedure, if we consider only the friction angle of foundation soil φ_f as a stochastic parameter and we set all other parameters with the nominal values in **Table 54**, the probability of failure for sliding along the base varies with $CV(\varphi_f)$ as shown in **Figure 107**. An increase in the variability of the friction angle below the wall increases the probability of failure from 15% to more than 45%. These results suggest that with the set length L of 5.60 m it is impossible to reach a value of P_f of 0.1%. Hence the length would first have to be increased and this analysis would have to be repeated.

The Monte Carlo simulations can be used to evaluate the minimum value of one parameter to keep the probability of failure less than 0.1% when other stochastic parameters vary with given variabilities.

A further analysis varied the tensile strength of the geogrid in order to keep the probability of failure P_f at less than 0.1% when $CV(\varphi_R)$ varies from 5% to 20%, and while $CV(\varphi_s) = CV(\varphi_f) = 20\%$. As shown in **Figure 108**, the tensile strength increases with the increase of $CV(\varphi_R)$ from 68 kN/m to 90 kN/m. Such analysis suggests that increasing the controls on the fill in order to decrease $CV(\varphi_R)$, would allow to reduce the required tensile strength of the geogrid. Then, the costs of increased controls could be compared with the decreased costs of geogrids.

While the tensile strength only affects the stability in terms of tensile rupture, the length of the geosynthetics influences all the other limit state functions.

The same approach can then be followed considering the variation of the foundation properties and the effect on the length of the geogrid. The minimum length L to keep the probability of failure P_f at less than 0.1%, considering all limit state functions, when $CV(\varphi_s) = CV(\varphi_f)$ varies from 5% to 40%, while $CV(\varphi_R) = 5\%$ was investigated. The results are shown in **Figure 109**, where it can be seen that the minimum length varies from 5.87 m to 13.0 m. These results could be used to compare the increased cost of geotechnical investigations required to keep $CV(\varphi_s) = CV(\varphi_f) = 5\%$ with the reduced cost of the reinforced soil wall when the length L can be reduced to 5.87 m.

Software Tools

A specific Excel spreadsheet was developed. This spreadsheet allows to set the desired random variables with their associated variability (in terms of % of the nominal value of each parameter). The software allows to set also the required probability of failure and the required number of Monte Carlo iterations.

References

- Bathurst, R.J. (2018). The basics of probabilistic internal stability analysis and design of reinforced soil walls explained. Proc. 11ICG, 11th International Conference on Geosynthetics. Seoul, Korea.
- Bond, A. & Harris, A. (2008). Decoding Eurocode 7. Taylor & Francis, London, UK. 616 pp.
- Rimoldi, P., Pezzano, P., Zannoni E. (2019). Reliability analysis of internal and external stability of geosynthetics reinforced soil retaining walls. Proc. GeoAmericas 2020. Rio de Janeiro, Brazil.

Zannoni, E. (2016). Reliability based design in soil reinforcement applications. Proc. GeoAmericas 2016, 3rd PanAmerican Conf. on Geosynthetics. Miami Beach, USA.

B.7. Rock Engineering

B.7.1 Rock slope stability

Authors: Renato Pereira

References: Gasc-Barbier M, Ballion A, Virely D (2008). Design of large cuttings in jointed rock. Bulletin of Engineering Geology and the Environment 67, pp. 227-235. doi: 10.1007/s10064-008-0127-4;

Pereira R, Muralha J, Lamas L (2023). Stability analysis of a rock slope: Full-probabilistic approach. 15th International ISRM Congress 2023, October 09th - 14th, 2023, Salzburg, Austria.

Problem definition

Description

This example comes from a real project detailed in Gasc-Barbier et al (2008).

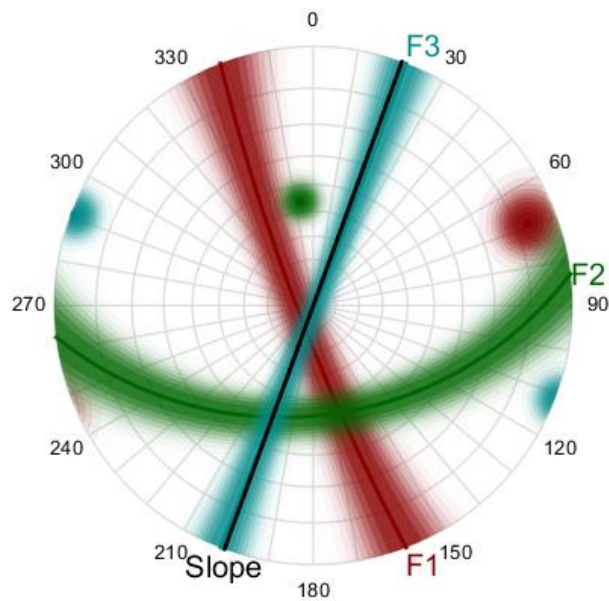
Rock cuttings were undertaken as part of the construction of a bypass in the French Pyrenees. The larger of those is 48 m high, 270 m long and requires the removal of more than 220,000 m³ of jointed gneiss with quartz and granite intrusions. Predicting the volume of rock that may fail after the cutting and consequently designing the ground reinforcement is then relevant.

This study consists of the probabilistic slope stability analysis, specifically concerning wedge failure modes.

Available Data

The slope is nearly vertical and oriented at N110°E. Statistical analysis of joint orientation data, obtained from outcrops and scanlines, allowed the distinction of three joint sets (F1, F2 and F3) and the estimation of the respective mean orientations (dip and dip direction), dispersion and spacing parameters. **Figure 110** shows the lower hemisphere projection of the joint sets, also representing their dispersion by means of 99% confidence interval (Fisher distribution).

Figure 110. Lower hemisphere equal-angle projection of the joint sets identified on field.



Source: Pereira et al. 2023

Limit State function(s)

Graphical analysis of the lower hemisphere projection of the joint planes (**Figure 110**) reveals the existence of a predominant failure mechanism characterized by rock wedges defined by the intersection of joint sets F1 and F2. Joint set F3, sub-vertical and sub-parallel to the slope face, might cut the rock slope at multiple distances from the slope face, also enabling the occurrence of toppling failure mechanisms. Without loss of generality, this study focus on the stability of the rock wedge defined by joint sets F1 and F2, but considering the effect of other joints from set F3 possibly intersecting it.

Wedge stability, either a portion of it or the whole wedge, is quantified through a factor of safety generically given by the quotient between stabilizing and destabilizing effects of actions. A closed-form solution for the factor of safety can be obtained from the block theory and the vector method, which encompasses the possibility of sliding along one or both joints, depending on the joint orientation and the direction of the total force.

The corresponding limit state function G can then be explicitly defined by,

$$G(x) = FS(x) - 1$$

Equation 159.

where FS is the factor of safety given in terms of the random variables x involved in the problem.

Target reliability

Target reliability (or probability of failure) must be set in order to have a minimum value to compare the calculated reliability or to design the minimum reinforcement needed. This value would depend, in accordance with EN 1997-1, on the geotechnical category of the project. In this case (major highway), a geotechnical category between GC2 and GC3 would fit, depending on the complexity of the ground model and the amount of information available. However, since no

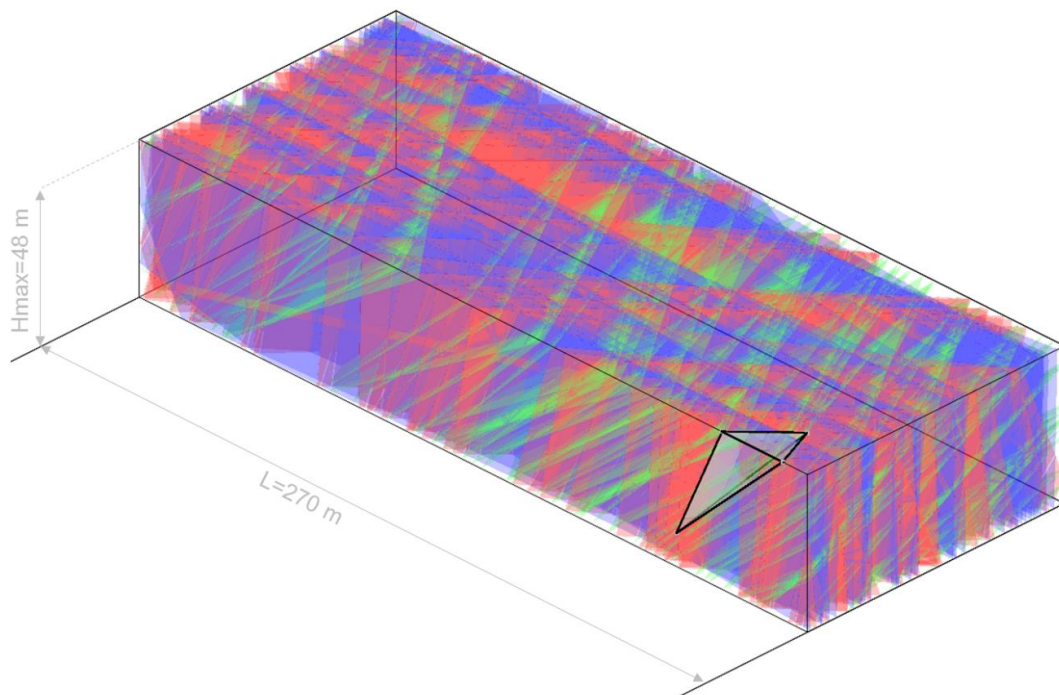
relevant information is directly given in the source documents, this decision would be purely speculative and therefore only the probability of failure is calculated in this example. No comparison to a target value is made.

Ground Model

The ground model is a description of the media (ground and groundwater), including its geometric, mechanical and other relevant properties. In this example, although not mentioned, it is plausible to admit the presence of groundwater and heterogeneity of the ground properties. However, given the geometric uncertainties on the orientation and spacing between rock joints, the jointed structure of the media is actually unknown. In such cases, the discontinuities within the geotechnical unit need to be modelled explicitly.

A realization of the ground model is illustrated in **Figure 111**, showing a discrete fracture network (DFN) derived from the aforementioned joint orientation data.

Figure 111. Realization of a discrete fracture network (DFN).



Source: Authors' own work

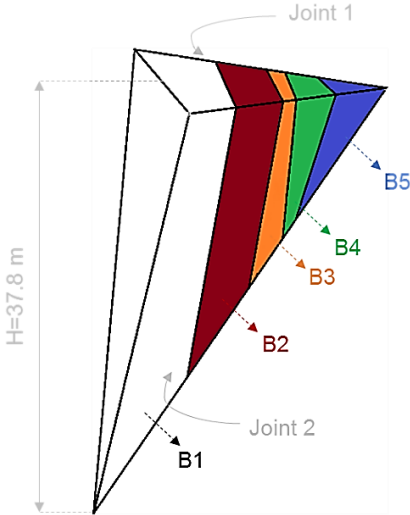
Geotechnical design model

A geotechnical design model is a conceptual representation of the media, derived from the ground model, for the verification of the limit state. It is then not uncommon to derive different geotechnical design models from the same ground model, for analyzing different failure mechanisms.

In this case, being the wedge instability the only failure mechanism being considered, a geotechnical design model can be an isolated wedge limited by any two joints from sets F1 and F2. Conservatively, these joints are modelled as planar and persistent. However, other joints intersecting the wedge can be accounted for and the wedge may be actually divided into several

blocks by other joints from set F3. **Figure 112** illustrates a realization of a wedge with 4 intersecting joints from set F3, taken from the generated DFN presented in **Figure 111**.

Figure 112. Rock wedge highlighted in **Figure 111**, intersected by 4 joints from joint set 3 forming 5 rock blocks.



Source: Authors' own work

Input Parameters (random and deterministic variables)

The input parameters are the deterministic and random variables described in **Table 58**.

Table 58. Input parameters.

Deterministic variables	Unit	Description (Constant value)
sdip /sdd	°/°	Slope face orientation, i.e. dip /dip direction (90/110)
bdip /bdd	°/°	Slope bench orientation (0/110)
Hmax	m	Slope height (48)
H	m	Wedge height (varied from 2 to Hmax)
ρ_w	kg/m ³	Water density (1000)
Random variables	Unit	Description
ρ_R	kg/m ³	Rock density
Hw	m	Groundwater level (in relation to the slope toe)
jdip _j /jdd _j	°/°	Joint j orientation (dip /dip direction)
θ_j	kPa	Shear strength model uncertainty (joint j)
kw _j	1	Rate of the water pressure at toe (joint j)
ΔL_3	m	Joint (set F3) spacing

Uncertainty characterization

Geotechnical units and parameters

Rock engineering problems may depend on the jointed structure of the ground. To analyze the stability of moderately weathered rock masses, when their jointed structure must be considered

explicitly, the boundaries of the geotechnical units are treated similarly to other discontinuities within the geotechnical unit. In this example, the slope is considered belonging to the same geotechnical unit (gneissic material) so the rock joints in it are actually the relevant discontinuities to be taken into account.

Rock joints are geometrically characterized by numerous unknown features. While some of them (persistency, waviness, roughness, aperture, etc) can be indirectly accounted by affecting the joint shear strength parameters, others, such as orientation and spacing, must be modelled as random variables. Field information is often available allowing to perform statistical inference of their distribution parameters. In this case, probabilistic distribution parameters of both joint orientation and spacing are given and summarized in **Table 59**.

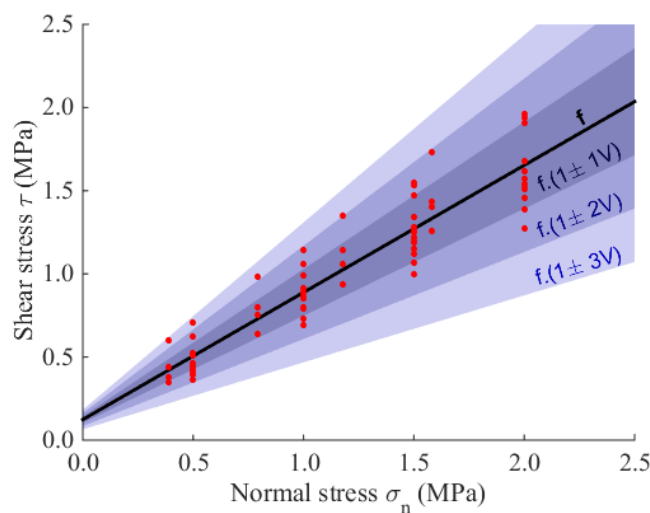
Table 59. Description of random variables related to geometric uncertainties.

Variables	Distribution	Parameters
$jdip_1 / jdd_1$	Fisher	{Mean orientation=83 /249}; k=86
$jdip_2 / jdd_2$	Fisher	{Mean orientation=44 /173}; k=105
$jdip_3 / jdd_3$	Fisher	{Mean orientation=89 /111}; k=196.5
ΔL_3	Exponential	$\lambda=3$

Joint shear strength

The actual joint shear strength is also unknown. Laboratory testing programs are usually performed on a limited number of specimens. In this example, no information regarding the joint shear testing is available. **Figure 113** shows the outcome of a rock joint shear testing program, from a different project.

Figure 113. Outcome of a rock joint shear testing program.



Source: Authors' own work

In practice, a model defining the failure criteria is adjusted to the results. The model can be previously selected, based on expert judgement, or derived after comparative statistical tests. In rock mechanics, nonlinear failure criteria, whose parameters are not always easy to be associated with physical properties, are often used to define the shear strength. This encourages to consider

the model itself as a random variable, as opposed to the model parameters. This option also facilitates test data analysis. Accordingly, the shear strength of joint j can be given by,

$$\tau_j = f(\sigma_{n,j}, \bar{p}_j) \cdot \theta_j$$

Equation 160.

Where $f(\cdot)$ is the adjusted shear strength model, $\sigma_{n,j}$ is the normal stress, \bar{p}_j is the model best estimate parameters and θ_j is the model uncertainty which also carries, in this formulation, the inherent variability of the shear strength (see section 5.3). Note that the multiplicative form of the model uncertainty, as opposed to the additive form, allows to use the coefficient of variation V directly as a standard deviation. In face of the limited number of specimens usually tested, statistical variability should also be accounted for (see section 5.3.5).

In this example, the joint shear strength is assumed to follow the Mohr-Coulomb failure criteria with cohesion $\bar{c}_j=40$ kPa and friction angle $\bar{\varphi}_j=35^\circ$ as best estimate parameters for joints from both sets F1 and F2. The respective model uncertainties, θ_1 and θ_2 , are assumed to follow a lognormal distribution with unit mean and coefficient of variation of 20%. Although no specific information is available, the model uncertainty associated with the shear strength along different joints is considered positively correlated with a correlation coefficient of 0.3, which seems to be a realistic hypothesis.

Loads, groundwater and pore pressure

Dead load

The dead load of each block is dependent on the rock density which is here considered as a random variable following a normal distribution with mean value 2600 kg/m^3 and standard deviation of 50 kg/m^3 .

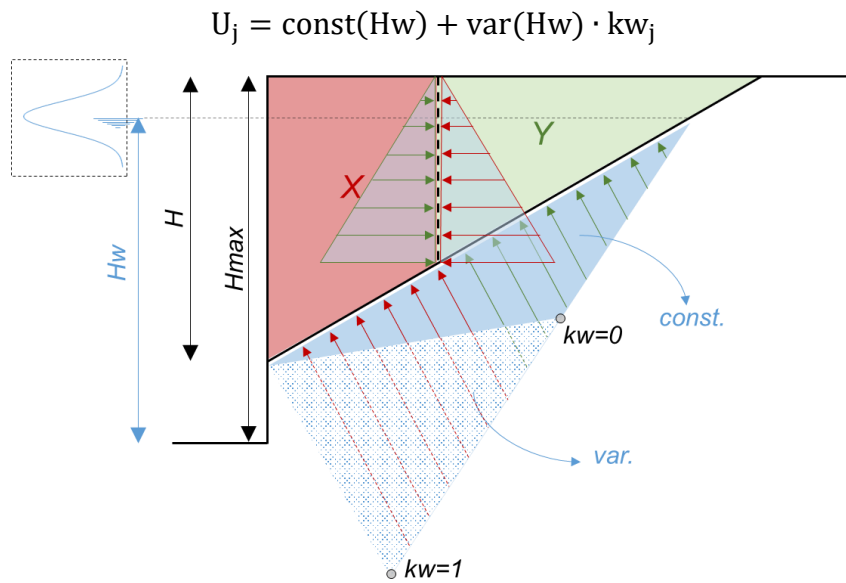
Water (variable) load

As mentioned, the presence of groundwater is plausible. The groundwater would exert a hydrostatic pressure normal to the submerged portions of the block faces. The point-in-time (quasi-permanent) groundwater level is assumed to follow a normal distribution with mean value of 40 m and standard deviation of 3 m (measured in relation to the slope toe). Furthermore, it is assumed that the water pressure along any joint increases linearly with depth. This is an idealized pressure distribution which may deviate from reality due to the spatial variability of the hydraulic conductivity along joints. This issue is specifically accounted for joints from sets F1 and F2. Joints from set F3, which serve as tensile cracks, are assumed to be fully opened at failure. Two independent random variables kw_1 and kw_2 , uniformly distributed between 0 and 1, are introduced to simulate that the peak water pressure may be located at the toe ($kw=1$) or at the mid-height of the submerged portion ($kw=0$) of the joints from sets F1 and F2, respectively. **Figure 114** illustrates the water load effects on joints. In short, there is then a constant part (const) of the total joint water pressure U_j , that solely depends on the groundwater level, and a variable part (var), affected also by kw , i.e.,

$$U_j = \text{const}(Hw) + \text{var}(Hw) \cdot kw_j$$

Equation 161.

Figure 114. Representation of the water load effects on joints (equivalent two-dimensional problem).



Source: Authors' own work

Model uncertainties

Not considered.

Reliability analysis

System reliability formulation

Accepting the existence of other sub-vertical joints that intersect it and may behave as tension cracks, the rock wedge consists in an assembly or a system of blocks. The state of the system can then be inferred from the state (stability or instability) of each block. However, since an inner block cannot be unstable without the instability of all the outward blocks, this problem, that would originally require testing the stability of each block individually, can be simplified. In turn, the system is considered to fail whether any set of blocks, from the slope face to any joint serving as tension crack, is unstable. Therefore, a rock wedge intersected by N other joints can be idealized as a series system with $N + 1$ alternative failure modes ($i = 1, \dots, N + 1$), corresponding to the instability of the key set of blocks X (from block B_1 up to block B_i) and the stability of the remaining set of blocks Y (from block B_{i+1} to the block B_{N+1}), i.e.,

$$\begin{aligned}
 p_{f,\text{sys}} &= P\left(\left[G_{B_1} < 0 \quad \cap \quad G_{B_2+\dots+B_{N+1}} > 0 \right] \cup \right. \\
 &\quad \left[G_{B_1+B_2} < 0 \quad \cap \quad G_{B_3+\dots+B_{N+1}} > 0 \right] \cup \dots \cup \\
 &\quad \left. \left[G_{B_1+\dots+B_N} < 0 \quad \cap \quad G_{B_{N+1}} > 0 \right] \cup \right. \\
 &\quad \left. G_{B_1+\dots+B_{N+1}} < 0 \right) \\
 &= P\left(\bigcup_{i=1}^{N+1} \left[G_{X(i)} < 0 \quad \cap \quad G_{Y(i)} > 0 \right] \right)
 \end{aligned}$$

Equation 162.

Note that the failure mode corresponding to $i = N + 1$ accounts for the instability of the whole wedge.

Selection of reliability method(s)

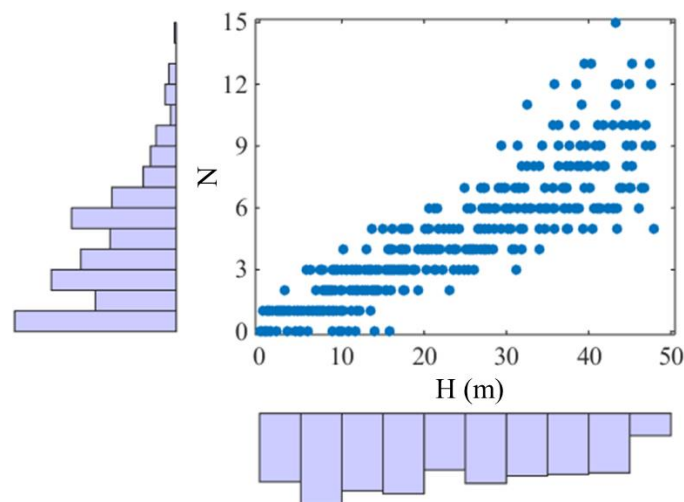
The system probability of failure can be computed using direct simulation techniques, i.e. the ‘crude’ Monte Carlo method (see section 6.2.4). The system is considered to have failed whenever any failure mode occurs. For each realization of the random variables, all failure modes must be tested (or at least, until the system fails for the first time). The use of the crude Monte Carlo method in this case is only practical because the failure mode verification relies on solving limit states with explicit (analytical) formulations. For a wedge with height H , the following algorithm is used:

- Step 1. Generate values of the random variables involved in all failure modes ($\rho_R, H_w, kw_1, kw_2, \theta_1, \theta_2, jdip_1 / jdd_1$ and $jdip_2 / jdd_2$).
- Step 2. Compute the length of rock wedge (L) from $H, jdip_1 / jdd_1$ and $jdip_2 / jdd_2$.
- Step 3. Generate m joints (orientation $jdip_3 / jdd_3$ and relative location ΔL_3) from set F3. Note that only the joints intersecting the wedge are relevant, i.e. $\sum_{i=1}^N \Delta L_{3,i} < L$.
- Step 4. Check the occurrence of failure for each $N+1$ failure mode.
- Step 5. Estimate the system probability of failure and the corresponding coefficient of variation.
- Step 6. Repeat the previous steps until a small coefficient of variation (here set to 2%) of the estimated probability of failure is achieved.

Estimation of probability of failure

The procedure described above was first performed for $H = H_{max} = 48$ m and then repeated considering gradually smaller heights of the rock wedge, since it is expected that several wedges with variable dimensions exist throughout the rock slope. Actually, from the DFN shown in **Figure 111**, several wedges are formed with variable heights. **Figure 115** shows the relation between the wedge height (H) and the number of joints (N) from set F3 intersecting them. The histogram of the wedges heights is also shown illustrating its randomness.

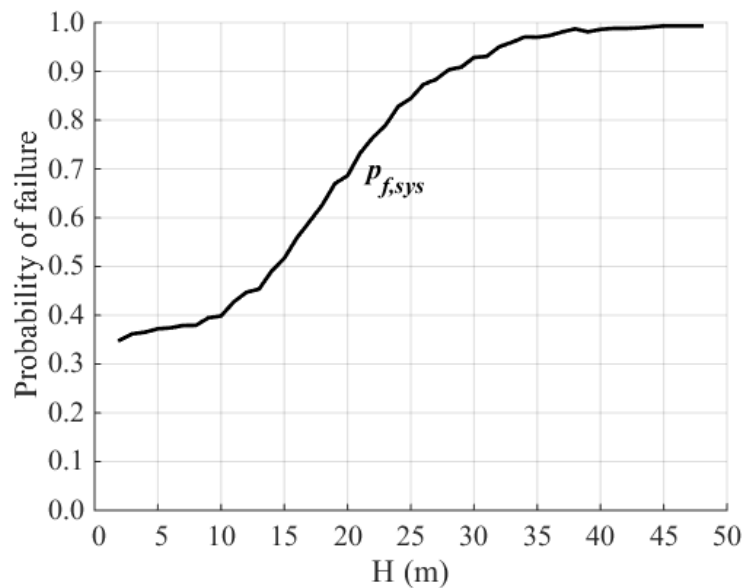
Figure 115. Wedge height vs number of joints from set F3 intersecting them, for the DFN in **Figure 111**.



Source: Authors' own work

Figure 116 shows the system probability of failure versus the rock wedge height.

Figure 116. Output from reliability analysis.



Source: Pereira et al. 2023

Interpretation of results and convergence

Given, on one hand, the unfavorable geometry of the wedge and, on the other hand, the statistical properties considered for the other random variables, especially the shear strength and the groundwater level, the system probability of failure of higher wedges is very high. Although no comparison to target values is made here, the probability of failure is obviously excessive and reinforcement measures are certainly required. The extent of those is yet dependent on the targeted probability of failure.

Since the target probability of failure (or reliability) is closely related to the consequences of failure, it is not uncommon to have this requirement varying with the height (or volume) of the unstable blocks. That is possible, even in the context of EN 1997-1, by assigning different geotechnical categories within the same project as a function of the severity of the failure event.

Final remarks

The groundwater level, as the only variable load taken into account, conditions the time reference period related to the estimated probability of failure. As the point-in-time distribution of the groundwater level was assumed, the estimated value is in fact a measure of the instantaneous probability of failure. For other time reference periods, the distribution of the groundwater level must be corrected, prior to the performance of reliability analysis, using extreme value theory.

The joint shear strength is, as described, derived from test data analysis. These tests are performed on small specimens taken from different locations. Scale effects and spatial variability are probably present. This situation would require an approach similar to what is detailed in section 5.3.2. However, in rock engineering, this is still an open discussion and research must be undertaken. Nonetheless, the strategy followed in this example is on the conservative side since it does not perform the averaging of the properties that would be required in that case.

Software Tools

Probabilistic studies similar to this can be performed with the SWedge software from [®]Rocscience.

References

Gasc-Barbier M, Ballion A, Virely D (2008). Design of large cuttings in jointed rock. Bulletin of Engineering Geology and the Environment Vol. 67, pp. 227-235.

Pereira R, Muralha J, Lamas L (2023). Stability analysis of a rock slope: Full-probabilistic approach. 15th International ISRM Congress 2023, October 09th - 14th, 2023, Salzburg, Austria.

B.8. Further readings

The following references include selected examples of RBD for further study. The references are structured according to the main topic of the publication. In addition, the topics of ground properties and rock engineering are also listed.

Table 60. Further literature on probabilistic design and assessment of various geotechnical structures and topics with a brief description of the contents.

Source	Short summary and main contents
1 Slopes, cuttings, and embankments	
Müller, R.; Larsson, S. & Spross, J. (2016): Multivariate stability assessment during staged construction; DOI: 10.1139/cgj-2015-0037	For staging the construction of embankments on soft clay, an important aspect in deterministic or probabilistic stability analyses is the assessment of the representative average values and associated uncertainties for the undrained shear strength as the height of the embankment is sequentially increased. Assessments made prior to construction can be verified by performing observations during the construction phase. All relevant available information should be incorporated into an analysis to increase the level of confidence and the objectivity of the assessment. An extended multivariate approach is applied to assess the undrained shear strength using different indirect measurement methods during the staged construction of the Veda embankment (Sweden).
Lanzafame, R. & Sitar, N. (2018): Reliability Analysis of the Influence of Woody Vegetation on Levee Performance; DOI: 10.13140/RG.2.2.34674.20163	In this paper, the effect of vegetation on the performance of California levees was quantified using a probabilistic/stochastic analysis. The First Order Reliability Method (FORM) was used to evaluate the incremental contribution to structural deficiency posed by vegetation by calculating slope stability and seepage through and under the model levees in the presence and absence of vegetation. The principal strength of the reliability analysis approach is that it can directly account for the inherent variability, i.e. stochastic nature, of the input variables and the associated degree of uncertainty about their actual values.
Jongejan, R.; Drosos, V.; Giannakou, A.; et al. (2018): Probabilistic assessments of flood defence performance subject to induced seismicity; DOI: 10.1007/s10518-018-0521-7	Gas extraction in the Netherlands has caused seismicity. A method was needed for probabilistic assessment of the seismic performance of the levees that protect low-lying polders against flooding. By combining the First Order Reliability Method (FORM) with response surfaces it was proved possible to strongly reduce the required number of simulations with advanced numerical models to obtain reliable failure probability

	estimates. To illustrate the workings of the method, an application to a levee cross-section with a sheet pile wall is presented.
Wolebo, A. P. (2016): Advanced Probabilistic Slope Stability on Rissa Slope	This thesis focuses on the evaluation of the effect of uncertainty and soil variability on stability analysis within the framework of probabilistic methods and contributes to the application of advanced probabilistic methods in slope stability analysis. Conditional random finite element method (CRFEM) creates a computational model able to estimate the probability of failure of a slope while fully accounting for spatial variability of the soil.
2 Spread foundations	
3 Piled foundations	
4 Retaining structures	
5 Anchors	
6 Reinforced fill structures	
7 Ground reinforcing elements	
8 Ground improvement	
Spross, J.; Bergman, N. & Larsson, S. (2021): Reliability-Based Verification of Serviceability Limit States of Dry Deep Mixing Columns. DOI: 10.1061/(ASCE)GT.1943-5606.0002458.	The structural behavior of the soil volume improved with columns is difficult to predict due to the existence of considerable uncertainties in the mixing process and the structural interaction between the columns and the untreated soil. This paper probabilistically investigates two serviceability limit states of deep mixing columns from a system reliability perspective. A design framework employing the observational method is proposed that considers allowable residual settlements, excessive settlement from column yielding, and the curing time of the columns. The design framework facilitates an effective reduction of the geotechnical uncertainty during construction and promotes risk-aware decision-making during both design and construction of the embankment.
9 Groundwater control	
10 Ground properties	
Müller, R.; Larsson, S. & Spross, J. (2014): Extended multivariate approach for uncertainty reduction in the assessment of undrained shear strength in clay; DOI: 10.1139/cgj-2012-0176	Important features of the multivariate approach are discussed, and an extension to this approach is proposed whereby the total uncertainty in site investigation methods due to spatial averaging is assessed prior to its adoption. Results from a site investigation of spatially averaged values of undrained shear strength and the corresponding coefficient of variation in Veda sulphide clay were used as a practical illustration of the extended multivariate approach and provide a basis for discussion. The inherent variability and scales of fluctuation for different methods are presented. The study shows the usefulness of the extended multivariate approach for the

	evaluation of representative values of and based on results from different methods.
11 Rock	
Bozorgzadeh, N. & Harrison, J. P. (2019): Reliability-based design in rock engineering: Application of Bayesian regression methods to rock strength data; DOI: 10.1016/j.jrmge.2019.02.002	This paper examines the Hoek-Brown (H-B) strength criterion within the RBD framework, and presents three distinct analyses using a Bayesian approach. Firstly, a compilation of intact compressive strength test data for six rock types is used to examine uncertainty and variability in the estimated H-B parameters m and σ_c , and corresponding predicted axial strength. The second analysis uses an extensive set of compressive and tensile (both direct and indirect) strength data for a granodiorite, together with a new Bayesian regression model, to develop joint probability distributions of m and σ_c suitable for use in RBD. The third analysis uses the granodiorite data to investigate the important matter of developing characteristic strength criteria.
Poisel, R.; Hoedlmoser, N.; Kolenprat, B. & Hofmann, R. (2017): The influence of joint orientation uncertainties on the stability of rock structures – do we need alternatives to partial factors?	The stability of rock structures is massively influenced by joints especially by their orientation uncertainties. Three examples of a slope, of a foundation and of a tunnel in rock are to show the importance of taking uncertainties of joint orientation into account.

Getting in touch with the EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696,
- via the following form: european-union.europa.eu/contact-eu/write-us_en.

Finding information about the EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

Science for policy

The Joint Research Centre (JRC) provides independent, evidence-based knowledge and science, supporting EU policies to positively impact society



EU Science Hub

[Joint-research-centre.ec.europa.eu](https://joint-research-centre.ec.europa.eu)



Publications Office
of the European Union