RUHR-UNIVERSITÄT BOCHUM

SFB 837
Interaction Modeling in Mechanized Tunneling

Raoul Hölter

Optimal Experimental Design in Geotechnical Engineering

Bochum 2021

Heft 71

Schriftenreihe des Lehrstuhls für
Bodenmechanik, Grundbau und Umweltgeotechnik

Herausgeber: Torsten Wichtmann

ISSN 2699-1020

Raoul Hölter Optimal Experimental Design in Geotechnical Engineering

71

ISSN 2699-1020

Ruhr-Universität Bochum
Schriftenreihe Bodenmechanik, Grundbau und Umweltgeotechnik
Heft 71

Herausgeber:
Prof. Dr.-Ing. habil. Torsten Wichtmann
Ruhr-Universität Bochum
Fakultät für Bau- und Umweltingenieurwissenschaften
Lehrstuhl für Bodenmechanik, Grundbau und Umweltgeotechnik
44801 Bochum
Telefon: 0234/ 3226135
Telefax: 0234/ 3214236
Internet: www.bgu.ruhr-uni-bochum.de

ISSN 2699-1020

Dissertation
as a requirement for the degree of
Doktor-Ingenieur (Dr.-Ing.)

Optimal Experimental Design in Geotechnical Engineering

submitted by
Raoul Hölter

Bochum, December 2021

Reviewers
Prof. Dr.-Ing. habil. Torsten Wichtmann
Prof. Dr. rer. nat. Tom Lahmer
Prof. Dr. Maria Datcheva

Preface of the editor

When realizing complex geotechnical structures like deep excavations or tunnels the application of the observational method is common practice nowadays. Thereby a numerical model of the geotechnical problem is built, while the structure and its vicinity are equipped with sensors. The results of the measurements during construction are used to improve the numerical model, particularly the parameters of the constitutive models applied, by back analysis. In that way the model is continuously enhanced, allowing more accurate predictions of deformations and limit states. If the model predicts the loss of serviceability or even failure suitable countermeasures must be adopted.

Previous research works concentrated on the back analysis of the model parameters based on measurements at given positions. The dissertation of Raoul Hölter goes one step beyond and focusses on the question where to place the sensors, how many of them and which types, in order to enable the most accurate back analysis. The thesis thus discusses the application of methods of optimum experimental design (OED), known from other research disciplines, in geotechnical engineering.

Based on an extensive literature review, mainly covering studies from research disciplines beside civil engineering, Mr. Hölter has identified three OED approaches being promising for an adoption in geotechnical engineering, namely the global sensitivity analysis (GSA), the Bootstrap method and the Bayesian OED. The thesis summarizes the fundamentals of these methods along with the involved mathematical basics and shows their implementation and application to different geotechnical problems, namely a model test with consolidation of clay under step-wise loading, ground settlements caused by mechanized tunneling and a dike subjected to a rapid drawdown. Based on the studied examples Mr. Hölter discusses the advantages and drawbacks of the different OED methods and gives recommendations for their practical application taking into account the large computational effort of some numerical tools. Amongst others, one interesting outcome of the investigated examples is that the sensors should be placed in zones with high gradient of the measured quantity, rather than choosing positions with high absolute values.

The present thesis may be seen as the first systematic research work focussing on the application of OED in geotechnical engineering. Such application could lead to an optimized arrangement of sensors with regard to the collected information and costs, and to a more reliable determination of the parameters of numerical models by back analysis in the framework of the observational method.

This research has been done in the framework of the project C2 of Collaborate Research Center (SFB) 837 "Interaction Modeling in Mechanized Tunneling" at Ruhr-Universität Bochum. The funding of German Research Council (DFG) is gratefully acknowledged.

Torsten Wichtmann

Acknowledgement

This research work was performed at the Chair of Soil Mechanics, Foundation Engineering and Environmental Geotechnics at Ruhr-Universität Bochum. At first, I want to thank Prof. Dr.-Ing. habil. Tom Schanz who gave me the opportunity to join his team at this chair and who generated my enthusiasm for the investigated research topic, but who could sadly not attend its finalisation. In this regard, I would like to thank Prof. Dr.-Ing. habil. Torsten Wichtmann who enabled me to finalise this doctoral thesis and who accepted to take over the supervision of this work.

The thesis was funded by the German Research Foundation DFG as a part of the collaborated research center SFB 837 "Interaction Modelling in Mechanized Tunneling" within the Subproject C2 "System and Parameter Identification Methods for Ground Models in Mechanized Tunneling". I want to express my gratitude for being part of this SFB 837 that provided an inspiring and productive research environment.

Dr. Maria Datcheva from Bulgarian Academy of Sciences from the first day on supervised and supported this work, helped me in particular in terms of the mathematical aspects of this work and contributed as reviewer of this thesis to its successful finalisation, wherefore I would like to especially thank her.

In this context I would also like to express my gratitude to Prof. Dr.-Ing. habil. Achim Hettler for giving me the opportunity to work as research assistant at the Chair of Soil Mechanics and Foundation Engineering at TU Dortmund since October 2015, allowing me to obtain a new point of view on the field of geotechnical engineering.

I also would like to thank Prof. Dr. rer. nat. Tom Lahmer from Bauhaus-Universität Weimar who approved being reviewer of this thesis and who introduced me to his research group.

My thanks goes to the colleagues at the chair of Computing in Engineering at Ruhr-Universität Bochum who welcomed me with open arms when joined their institute in 2016, but of course also to my colleagues at the chair of Soil Mechanics, Foundation Engineering and Environmental Geotechnics who were a great team and who will always be

remembered in that way. A special thank goes to my student co-workers for their commitment in performing computer experiments: Maximilian Schoen, M.Sc., Armel Meda, M.Sc., Mohammad Sanayei, M.Sc., and Otabek Mirzakhmedov, M.Sc..

Finally, I want to thank my wife Sarah for being always on my side, her continuous support, and keeping me laughing also on the tough days of this research journey.

Bochum, December 2021

Raoul Hölter

Abstract

Performing measurements during construction and maintenance of major geotechnical projects like tunnels, ship locks, or dikes is state of the art in geotechnical engineering and usually obligatory. Doing so allows a comparison of the obtained measurement data with preliminary made predictions and therefore, to identify possible discrepancies. This procedure known as observational method always requires on the one hand a validated model that provides reliable system responses and on the other hand corresponding measurement data. While a lot of effort has been dedicated to develop highly sophisticated models and measurement devices, little research work is available on how the measurements should be conceptually set up, i.e. which measurement data should be gathered.

However, the problem of obtaining most informative data to validate a model that reflects a specific system is not limited to geotechnical engineering. Indeed, it is encountered in almost any scientific research field. To address this type of problems, the concept of optimal experimental design has been developed. This concept aspires to identify and develop measurement set-ups, the so-called experimental designs, that provide reliable data to identify the models' most relevant parameters. Hereby, the aspect of reliability plays a major role in all considerations of optimal experimental design as the identified experimental design should be robust towards uncertainties of the model and of the measurement data itself. Therefore, the approaches of experimental design mostly consist of statistical evaluations.

The present work considers several statistical methods employed in different research fields in view of their benefit in context of geotechnical problems where the soil's uncertainty is predominant and measurements are often sparsely distributed over large model/system domains. The focused on methods of optimal experimental design are the global sensitivity analysis, the Bootstrap evaluation, and the Bayesian updating. These methods are introduced in detail and applied to representative problems like a geotechnical laboratory device, the water drawdown acting on a dike, and mechanised tunnel construction. To systematically evaluate the models of these problems that are generated using the

finite element method, further mathematical tools like metamodelling and optimisation algorithms are employed and introduced in an additional chapter.

In the several considered applications, it is shown how the different methods allow an improvement of the parameter identification results, i.e. a reduction of the uncertainty range. However, also the limitations of the different methods are demonstrated which are mostly related to the required computational effort. Accordingly, recommendations are given how the different methods might be employed depending on the problem at hand and how combining the methods can increase their efficiency. Using these recommendations, efficient measurement set-ups can be designed and the observational method is improved by being placed on a rational basis for decision-making.

Contents

List of symbols

$\boldsymbol{A}$	Amplitude matrix
$\boldsymbol{b}$	vector of polynomial basis terms
$\boldsymbol{B}$	Size of Bootstrap or other statistical samples
$\mathfrak{B}$	Matrix of interpolation coefficients
c	Shape parameter
$\boldsymbol{c}$	Vector of approximation coefficients
$\boldsymbol{C_y}$	Covariance matrix of model responses $\boldsymbol{y}$
$\boldsymbol{C_\theta}$	Covariance matrix of identified parameters $\boldsymbol{\theta}$
$\boldsymbol{C}_{\boldsymbol{\theta},\mathrm{Boot}}$	Covariance matrix of identified parameters $\boldsymbol{\theta}$ using the Bootstrap approach
CV	Coefficient of variation
CR	Crossover probability
d	Number of dimensions of the design parameter space Π
$\boldsymbol{d}_\mathrm{H}$	Haussdorf distance
$\boldsymbol{\delta}$	Experimental design
$\boldsymbol{\delta}_0$	Initial experimental design
$\boldsymbol{y}_\mathrm{c}$	Experimental design selected without optimisation
$\boldsymbol{y}_\mathrm{o}$	Experimental design identified as optimum for the given purpose
e	Ratio of artifically added noise
E	Expectancy
G	Generation/Iteration step of DE algorithm
$\boldsymbol{\gamma}_{\mathrm{E}ig}$	Matrix of eigenvectors
F	Weighting/Scaling Factor of DE Algorithm
FIM	Fisher information matrix
J	Number of dimensions of the output space
J_min	Minimum of the objective function
K	Amount of candidate points to be investigated in a design space
κ	Stiffness derived from swelling line in MCC model
κ_s	Scaling factor to define number of sigma points $\boldsymbol{\sigma}$
λ	Stiffness derived from virgin compression line in MCC model
$\boldsymbol{\lambda}_{\mathrm{E}ig}$	Matrix of eigenvalues multiplied by unity matrix
M	Slope of critical state line in MCC model
n	Number of model responses/outputs
N	Number of sample points in the parameter space $\boldsymbol{\Theta}$
NP	Population size/Number of sample points in DE algorithm

P	Amount of employed sensors for one type of output
Φ	Optimality criterion
Φ_{F}	CDF of uncertain model parameters
$\boldsymbol{\Phi}_{\mathbf{POD}}$	matrix of proper orthogonal basis vectors
$\boldsymbol{\Pi}$	d-dimensional design parameter space
Q	number of evaluations to fit $\tilde{U}$
σ	Standard deviation
$\boldsymbol{\sigma}$	Matrix containing the $2 \cdot (s + \kappa_{\mathrm{s}})$ sigma points
s	Dimension of model parameter space $\boldsymbol{\Theta}$
S	Sensitivity matrix employed in FIM
SS_{ij}	Scaled Sensitivity index of j^{th} output to i^{th} parameter
$\boldsymbol{\theta}$	Model parameters
$\tilde{\boldsymbol{\theta}}$	Back calculated model parameters
$\bar{\boldsymbol{\theta}}$	Mean value of back-calculated model parameters
$\hat{\boldsymbol{\theta}}$	Optimal model parameters
$\tilde{\boldsymbol{\theta}}_r$	Vector of identified parameters extracted from $\tilde{\mathfrak{T}}$
$\tilde{\mathfrak{T}}$	Set/Matrix of back-calculated model parameter samples
$\mathfrak{T}$	Set/Matrix of model parameter samples $\boldsymbol{\theta}$
$\boldsymbol{\Theta}$	s-dimensional space of model parameters
u	Utility function
U	Utility of experimental design
$\tilde{U}$	Approximated Utility by one-dimensional metamodel
V	Variance
w	Weighting factor
$\boldsymbol{y}$	Vector of model responses
y_{x}	Horizontal displacements
y_{y}	Vertical displacements
y_{w}	(Excess) Pore water pressure
μ_i	Arithmetic mean of distributed value
$\tilde{\boldsymbol{y}}$	Measurement data affected by measurement errors

$\mathbf{Y}^*$	Bootstrap sample
$\mathfrak{Y}$	Matrix of model responses generated from different parameter samples
$\tilde{\mathfrak{Y}}$	Matrix of falsified model responses
ω	Randomly generated value
$\boldsymbol{\Omega}$	n-dimensional model output space

List of abbreviations

CPT	Cone Penetration Test
CDF	Cumulative Distribution Function
DE	Differential Evolution Algorithm
DOE	Design of Experiment
GA	Genetic Algorithm
GSA	Global Sensitivity Analysis
HS	Hardening Soil (constitutive model)
LHS	Latin Hypercube Sampling
LSA	Local Sensitivity Analysis
LSM	Least Square Method
MC	Monte Carlo
MCC	Modified Cam Clay (constitutive model)
MSE	Mean Square Error
NRMSE	Normalised root mean square error
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
PDF	Probability Density Function
POD RBF	Proper Orthogonal Decomposition with Radial Basis Functions
RBD	Reliability Based Design
SPT	Standard Penetration Test
TBM	Tunnel Boring Machine
VB	Variance Based sensitivity method
ACE	Approximate Coordinate Exchange

1. Introduction

1.1. Background

Employment of measurement data during the construction of large geotechnical structures like tunnels, excavation pits, or dikes is the common approach, not only because it helps to assess the safety of the structure, but because it is mandatory, e.g. according to Eurocode 7 (DIN EN 1997-1). However, there are two different kinds of measurements, or better to say, two kinds of data interpretation: either they are interpreted related to a threshold (e.g. maximum settlement), or as piece of information that is used to validate a model of the system at hand.
In case we think of thresholds, it is only possible to evaluate what happened up to the moment of the measurement and decide whether to react or not. This approach is commonly used and well described e.g. in Janin et al. (2012). The alternative of using measurement data for model validation was introduced as "observational method" in Peck (1969*b*). In such a case, a model of the considered situation is needed that can be validated. The validation takes place by comparing the model response with the measurement data. In case the model agrees with the measurement, it can be used to predict future situations, if not, different parameters of the model that describe e.g. the geometry or material behaviour might be adapted to fit the measurement.
A simple example from the field of geotechnical engineering is the loading of a shallow foundation. The more it is loaded, the more settlement will be measured. However, if a certain settlement is assumed as failure criterion, the maximum load capacity of the footing depends on the soil properties below. Due to intrinsic uncertainties of natural soils, these properties are not known exactly and either they can be assumed as very conservative values with large safety margins, or it is tried to match the correct value as accurate as possible. The first approach, called "conventional design" in Spross and Johansson (2017) needs observation only in case the settlements exceed the acceptable settlements. Using the observational method, as will be presented later, the settlement of the footing could be measured to validate a model of the system, e.g. the simple influence factor method

introduced in Kany (1959). Doing so, it can be identified if the initial assumptions of the soil properties were correct, or if they were over- or underestimated. Afterwards, the design parameter, that is the loading itself in the assumed case, can be adapted. This can lead to a safer design, and also to a more cost efficient design.
However, when using measurement data, we must be aware that large uncertainties are probably affecting the system. In Phoon and Kulhawy (1999*b*), model uncertainty, measurement uncertainty, and soil heterogeneity are named as main sources of inaccuracy in geotechnical engineering. Therefore, the obtained measurement data may be misleading. The aspect of model accuracy will be discussed later on in this thesis, but merely as secondary source of uncertainty, while the focus is set on the measurement uncertainty and soil inhomogeneity.
In the context of the above introductive example, soil inhomogeneity could mean that the stiffness, unlike expected, decreases with depth due to the geological history of the site. In such case, small loadings that impact only surface near areas would provoke explicitly small settlements. A model validation would identify a high stiffness of the soil and might tempt to assume a high allowable loading. In case the load is increased, the ratio of the additionally applied stress to the stress due to own weight will become larger in any depth below the footing. In deeper areas of the soil, increasing this ratio would now become relevant for settlements and cause them to be higher than expected, proving the previous assumption to be false.
Besides inhomogeneity, measurement errors must be considered that can be of epistemic or aleatoric origin, e.g. precision deficiency of the measurement device or wrong calibration to ambient temperature, respectively. These sources of uncertainty downgrade the advantage of the observational method. However, in Meier et al. (2009) and Meier et al. (2013) it is shown that for different "monitoring options", as various arrangement of the sensors are called by Spross and Johansson (2017), different qualities of the model validation are obtained in terms of model accuracy. Therefore, one should ask by which monitoring option the efficiency of the method could to maximised.
Of course, this aspect has never been neglected in geotechnical engineering, but usually a monitoring design is set up according to engineering judgement. Even though this "judgement" is a very powerful tool, it is not a general guideline that everybody approves in the same manner. Design codes such as the aforementioned DIN EN 1997-1 request a monitoring for complex geotechnical structures, but do not give any details on how this monitoring should be arranged.
Contributing to rationalise this aspect is the objective of this thesis. Every project is different and so are their needs for monitoring. Therefore, this work does not intend to

be a reference book where to look up which measurement device to place where in a certain case, but it is supposed to be a platform that provides different concepts that can be applied depending on dimension, complexity or level of detail of the considered case. Decision making how to apply and to interpret these methods is enabled and does not replace but improve the engineering judgement.

1.2. Objectives

The present work aims to investigate how the process of parameter identification, that is the core element of model validation, is influenced by different types of data and their inherent uncertainty. Ultimately, an awareness should be raised for the possible benefits of a thoughtful selection of measurement data that is employed for model calibration: A more adequate model allows more precise predictions of the system behaviour. Furthermore, less needed measurement instrumentation for the same accuracy level enables immediate economic savings.

To a certain extend, the work presented in this thesis is the continuation of the work of Miro (2016) which has been, like the one at hand, prepared in the framework of the collaborative research group SFB 837 "Interaction Modeling in Mechanized Tunneling" at Ruhr-Universität Bochum in the subproject C2, entitled "System and Parameter Identification Methods for Ground Models in Mechanized Tunneling". In Miro (2016) the process of model calibration, as is called therein the verification of the existing model, simply "uses" measurement data and assumes the position, accuracy, etc. of the corresponding sensors as given. The present work extends this previous one by adding an additional dimension to the problem: if we want to validate our model, which data should we use at all for this purpose?

To answer this question, it should be first described what is understood by model validation that is a complex topic, including several steps and mathematical methods. As this process is influenced by the uncertainties mentioned above, finding a solution, i.e. a set of parameters that allows a reproduction of the measured data, means solving an ill-posed problem as usually no solution will allow a perfect fit and many solutions may allow a good approximation. The main objective in this work of finding the optimal measurement arrangement is therefore a second optimisation problem that aims to provide optimal boundary values to the first one of identifying parameters. The objectives of the investigation are as follows:

- Convey the concept of optimal experimental design (OED) and discuss how this concept can be helpful to improve safety and efficiency in geotechnical engineering. Considering the characteristics of this research field, a literature review of existing methods of OED is performed.
- Provide a survey on different approaches to perform model parameter identification based on back analysis of measurement data with specific application to geotechnical problems and optimal experimental design. In this context, the mathematical tools that are needed for this purpose must be introduced.
- It is one key objective to adapt and develop different existing and new approaches for model parameter identification, employing geotechnical measurement data collected based on optimal experimental design. Hereby, explicit focus is set on the question which impact different sources of uncertainty may have on the result of parameter identification.
- The different considered approaches of optimal experimental design should be applied to selected examples that are representative for geotechnical engineering. Thereby, the different features, advantages, and drawbacks should be depicted.
- Illustrate which possibilities the concept of OED offers in geotechnical engineering. It shall be indicated for which types of problems it is beneficial to invest this additional effort and which gain in reliability can be obtained. The aspect of efficiency is included in the evaluation, as the computational effort becomes a crucial aspect in some of the approaches.

1.3. Layout of the thesis

After this introduction chapter, the second out of eight chapters introduces the topic of OED and presents the state of the art in this research field. As most research work is not from the field of civil and much less of geotechnical engineering, a thematic bridge is built to this topic, showing some works that meet the ideas of OED in the broadest sense.

In the third chapter, the mathematical and statistical tools are discussed that are employed through this thesis. Some of these methods have been introduced in the previous work of Miro (2016) such as metamodelling or global sensitivity analysis, but as they are applied in different manners in the present study, they are outlined again with special focus on the intended application. Other methods such as the Bootstrap or the sigma points

are new in the framework of this research project and are explained more fundamentally in the corresponding chapters.

In chapter four, the first candidate method is introduced that employs a spatial global sensitivity analysis to detect promising sensor placements. The concept is applied to several geotechnical problems considering real cases of a laboratory experiment, a tunnel construction, and the synthetic case of a dike that is subjected to a hydromechanical impact.

The fifth chapter focuses on the Bootstrap resampling method that allows to reproduce artificial noisy data. This noisy data is assigned to explicit sensor types and allows more accurate definition of sensor arrangements, i.e. experimental designs. A continuation of this approach is presented by employing the so-called sigma point method. Thereby, extensive distributions of noisy data are substituted by their first order moments.

In chapter six, the ideas of Bayesian OED are taken up. This method allows to define the objective of an experimental design to be the reduction of uncertainty that is inevitable when handling soil parameters.

After introducing three possible concepts of OED, a merged evaluation is performed in chapter seven. Hereby, the different methods are compared, their pros and cons are highlighted and recommendations are given on how to consider them in different applications.

The eighth chapter consists of a conclusion of the present work, combined with an outlook on possible extensions and further application examples.

2. Optimal Experimental Design

2.1. General ideas

This chapter illustrates the general concept of optimal experimental design (OED) to the reader. Its historical development is described as well as the different research fields where it is applied currently. As the concepts of OED are employed in different contexts, one important aspect is to determine to which type of applications it is suitable: can the information of interest be directly measured or does it need to be back-calculated as handled in this work? On the opposite, the differentiation is explained between the concept of OED for real experiments or in-situ measurements, and the design of computer experiments. Thereby, the focus of this work is stated and defined precisely, as no sharp differentiation is possible or reasonable between these methods.
In general, the following description given in Wolkenhauer et al. (2008) should be considered: "Performing experiments to obtain a rich enough set of experimental data is a costly and time-consuming activity. The purpose of optimal experimental design (OED) is to devise the necessary dynamic experiments in such a way that the parameters are estimated from the resulting experimental data with the best possible statistical quality, which is usually a measure of the accuracy and/or decorrelation of the estimated parameters. In other words, based on model candidates, we seek to design the best possible experiments in order to facilitate system identification." This relatively recent work that is actually from the field of biochemistry quite accurately describes what is intended within this thesis to be investigated and to be applied to geotechnical problems.

2.2. State of the art

2.2.1. Historical review

Of course, ever since people were systematically measuring an event, they reflected on how this measuring might be most efficient, so to say they looked for the optimal experimental design. However, the monograph Fisher (1935) is generally accepted as first systematic and scientific study on this topic. First, several examples related to evolution, psychology, or agriculture are provided therein to prove the relevance of testing (i.e. performing experiments) before performing the actual operation. The most employed example of investigating which treatment of seeds allows the largest harvest of grains well demonstrates the complexity of the problem of experimental design: testing the same type of seeds, they can be sowed on different types of soils and treated with different types of manures. By varying one by one these factors, one can identify their individual impact on the harvest. However, this might be quite insufficient, as hereby the correlation among the factors is not considered. If more than two of these factors are investigated, the higher order factor correlations should also be considered, e.g. the combination of two different manures on different soils. Testing every possible combination to identify the most valuable set-up is called a full factorial design. It is obvious that investigating all of these possible combinations quickly leads to an unbearable number of tests and unbearable experimental costs, especially in case these tests are as extensive as in agriculture where one test will take several months.
The problem becomes even more complex if the factors (or system variables as called later on) do not behave linearly. In case e.g. the amount of potassium sulphate added to the ground is doubled, will the positive effect on the harvest be doubled? Presumably not. Therefore, one should also vary the experiments considering the amounts of the different factors, what leads to an even larger number of tests. It is indicated how testing certain combinations of factors allows to gain the same information content as others what allows to leave out certain parts of the full factorial design. It is pointed out that certain factors will have more impact on the experiments' results and that it is important to identify these factors. Thereby, the concept of sensitivity analysis, that is described in detail later on, is anticipated without explicitly naming it. Finally, also the aspect of randomness and error is discussed, explaining that these may falsify the outcome of any experimental series. To overcome such random alternation of test subjects, the possibility of statistical evaluation is introduced. Hereby, relative changes are suggested as measurable criteria instead of absolute values to make treatments applied to different original states compar-

able. However, this uncertainty inevitably leads to an additional number of required tests. It is indeed pointed out that a large number of tests is often not possible in real world experiments. For this reason, in the present thesis, the numerical models and their more time-efficient substitute, the metamodel, are employed; an option that R.A. Fisher did not have in 1935.

The merits of the described monograph are to well describe the problem of the "design of experiments", to show how extensive such design can become, and what attempts can be employed to reduce the extend of such designs while maintaining their provided information content. Fundamental ideas are formulated such as having the variance of experimental results introduced as criterion of interest that allows to estimate the quality of an experimental design or the question how to relate monetary costs to the information gain of an experiment. Making this information gain measurable is the essential first step to be able to compare different experimental designs. Within Fisher (1935), it is stated "that methods of estimation which proceed without reference to the possibility of evaluating the quantity of information actually contained in the data, are liable to be defective in the quantity that they utilise". A formulation that might be applicable to any branch of research, but especially to the field of geotechnical engineering where information are often derived from sparse data. To relate data to information, Fisher deduces that the variance of a distribution obtained as results from an experimental design allows to describe the obtained information content. Accordingly, if some parts of the experimental design are left out, the obtained distribution will be different and its information content less. However, one can now compare which experimental design allows to be as close as possible to the full factorial design (i.e. to the most reliable distribution) for least experimental effort.

To describe the information content I of the experimental data $\tilde{\boldsymbol{y}}$ with respect to the parameter θ_i, the following formulation can be employed:

$$I = \frac{\partial^2 \tilde{\boldsymbol{y}}}{\partial^2 \theta_i} \ , \ i = 1 \dots s, \tag{2.1}$$

where s is the dimension of the considered parameter space. The matrix that accrues when considering all parameters and their correlations is often called the Fisher-information matrix FIM whereby several similar formulations exist.

Conceptually, the FIM turns out to be the variances' expectancy of the function's second derivatives to each of its parameters. It should be mentioned indeed that in this thesis, such functions are approximated by distributions obtained by random sampling. To allow consideration of uncertainties arising not only from the input but also from the measurement's uncertainty, the covariance matrix of the outputs $\boldsymbol{C}_y$, or $\boldsymbol{C}_{\tilde{y}}$ in case experimental

data is employed, is included in the formulation that is employed for OED purposes:

$$FIM = \boldsymbol{S}_l^{\mathrm{T}} \cdot \boldsymbol{C}_y^{-1} \cdot \boldsymbol{S}_l \tag{2.2}$$

wherein, according to 2.1, the matrix $\boldsymbol{S}_l$ is:

$$\boldsymbol{S}_l = \begin{pmatrix} \frac{\partial y_{1,l}}{\partial \theta_1} & \frac{\partial y_{1,l}}{\partial \theta_2} & \cdots & \frac{\partial y_{1,l}}{\partial \theta_s} \\ \frac{\partial y_{2,l}}{\partial \theta_1} & \frac{\partial y_{2,l}}{\partial \theta_2} & \cdots & \vdots \\ & \vdots & & \\ \frac{\partial y_{N,l}}{\partial \theta_1} & \cdots & \frac{\partial y_{N,l}}{\partial \theta_{s-1}} & \frac{\partial y_{N,l}}{\partial \theta_s} \end{pmatrix} \tag{2.3}$$

The index l refers to the variation of the considered experimental design $\boldsymbol{\delta}_l$ whereby each design leads to a different FIM that allows comparison among them.
Up to date, using the FIM remains a valuable approach for many types of applications described in the following. In Bandemer and Bellmann (1979), the state of the art up to the 1970th is presented and the concepts of Fisher (1935) are transferred to more systematic and fundamental approaches. The different optimality criteria, described in detail in Sec. 2.4.3 are hereby considered. Identifying an OED is performed based on the formal application of analytical functions, and the concept of the least square method (LSM) and close-by alternatives is presented. The algebraic properties of the FIM are investigated in a far more fundamental mathematical description than needed and intended for the present application-oriented thesis. Hereby, it should be remembered that at that time extensive FE simulations and their statistic evaluation was not possible due to lack of computational power. None of the applications considered in this thesis could have been performed using the techniques of the 1970s, even though the objectives and underlying concepts are the same.
Another aspect in the framework of OED is that one should differentiate between problems where the parameters of interest need to be back-calculated as it is the case in the present thesis and those where they can directly be measured. An example for the latter case is given in Mukherjee et al. (2017) where it is described how to find optimal sensor placements to detect contaminants in a water network. As the water demand in such a network is uncertain, finding these optimal placements is still a complex optimisation problem, but it does not require the additional step of parameter identification. Using a stochastic simulation of the water demand is combined with a reweighting scheme to simulate the impact of this varying demand on the contaminants' propagation. With different deterministic and stochastic optimisation approaches, including different uncertainty scenarios, the optimal sensor locations are identified exhibiting quite different results and showing the need to consider effects of uncertainty in OED applications.

Another concept that is related to the term of OED is the generation of samples for computer experiments. Within this thesis, such methods are employed to generate data for metamodels that substitute the more time-consuming FE models. This is introduced briefly in Sec. 3.2.2 while a detailed overview on such approaches such as the Latin hypercube sampling, the Box-Behnken-design, and the central composite design is provided in Giunta et al. (2003).

2.2.2. Current developments

The employment of the before introduced FIM is still very common as shown by several examples in Uciński (2005) and Patan (2012) from the same research group. There, like in those problems considered in this thesis, the problem of where to place sensors in a, as it is called therein, "distributed system" is investigated. In the considered application, the time-dependent pollutant transport over an urban area is studied among others. To avoid clustering, the sensor density is employed as design variable in an optimisation loop to identify the experimental design that provides the least value of the assigned cost function. The problem whether to consider discrete candidate points or a continuous domain is addressed, indicating the higher complexity of the second case, as discussed later on in chapter 6. Thereby, the aspect of reducing the computational effort as side aspect in the task of experimental design optimisation is underlined, but also the limitation of the FIM to problems that are linear or of "mild nonlinearity". To rank different experimental designs, the mean square error (MSE) is employed and, among others, the Φ_{A}- and Φ_{D} optimality criteria are applied to the FIM. These allow to evaluate candidate designs under aspects of optimality whereby these different aspects are described in more detail in Sec. 2.4.3 and to which an overview is provided in Nishii (1993).
In Lahmer et al. (2008), an example from material science is provided considering the properties of piezoelectric discs. The behaviour of these discs can be well described by partial differential equations (PDE), but the parameters of these PDEs are uncertain due to their production process. To identify these parameters, an accentuation of the piezometers within a possible bandwidth of frequencies is simulated under consideration of possible noise in the case of real data. The solution to this problem is found by generating artificial noisy data and afterwards minimising the MSE between these noisy measurements and the model responses by varying the system parameters. Besides identifying adequate measurement frequencies, it is investigated which number of frequencies should be employed to allow the most efficient parameter identification. This shows how complex the OED problem can be in any application, as the design space $\boldsymbol{\Pi}$ in which the possible

experimental design $\boldsymbol{\delta}_l$ is defined can be arbitrarily extended. At the same time, it can be seen that the fundamental concepts, as in this case the reduction of the MSE, are universally applicable. However, the specific application always needs individual adaptation, wherefore the employed methods cannot be transferred directly to problems of geotechnical engineering, what is one major motivation of the thesis at hand.
From the same author and using a similar approach, in Lahmer (2011) it is intended to place different types of sensors in a gravity dam. Similar to the problems studied in this thesis, the sensors (pore water pressure transducers and mechanical measurements) can vary their position over the whole domain of the model. Hereby, the problem of clustering appears, as the sensors tend to regroup in the same position; a circumstance that is mathematically logic, but often not applicable and not desired as locally arising technical problems might affect all sensors at once instead of only few of them.
In Schenkendorf et al. (2009) different methods of OED are compared in an application of system biology. To identify the parameters that control the growth of bacteria in a bioreactor, described by a series of PDEs, a back calculation of these parameters is necessary. Therefore, the aforementioned FIM is applied as well as the Bootstrap and the sigma-point method as described more in detail in Schenkendorf (2014). The comparison of these methods shows that using Bootstrap and sigma-point method allows to identify experimental designs of higher quality in case of the considered non-linear models. However, using the Bootstrap method has its limitations as it requires numerous calls of the time-consuming back analysis algorithm.
To increase the efficiency of OED approaches, and not try all possibilities within a design space, the concept of Bayesian OED can be employed as done among others in Vanlier et al. (2012) and Huan and Marzouk (2013) which consider applications from system biology and combustion kinetics, respectively. In the latter one e.g., it is intended to identify the parameters of ordinary differential equations (ODE) that describe the chemical reactions that take place in a shock tube. Even though only two design parameters, the initial temperature and the equivalence ratio, are considered as design variables, the response surface of the utility function (as inverse equivalent to the cost function in other applications) is non linear and requires a detailed investigation of the design space. Using the Bayes theorem, new candidate designs are selected based on investigations of prior tested ones, and thereby transferring the OED process to an active optimisation problem.
In Reichert et al. (2019), the impact of systematic and random noise with and without bias is investigated while the present thesis focuses on Gaußian zero mean noise. In that publication, a dynamic excitation is applied to a beam structure represented by a differential equation that provides the acceleration of the system as model response to which the

different types of noise are applied. The evaluation of this noisy data is performed using the FIM in combination with the Φ_D-criterion, but also using the MSE. The different types of errors lead to varying results, but not to fundamental changes of the experimental design. More relevant is the selection of the design criterion. In this regard, the results could be interpreted that the Φ_D-criterion aims by its formulation to identify all parameters that are included in $\boldsymbol{\theta}$. Therein, the relevance of these parameters for the model response is not taken into consideration. The MSE-criterion indeed exclusively cares for reducing the discrepancy between model response and noisy data. Hereby, a less-relevant parameter might be identified less accurately than a relevant one, but this does not matter as this is not the objective. One could say that the attention paid to the parameters is weighted according to their relevance to the model response. This is principally the idea of sensitivity analysis as introduced later on in Sec. 4.1. To enable using the Φ_D criterion while still considering the relevance of the individual parameters, in the applications shown in this thesis, a sensitivity analysis is performed preliminary. The obtained sensitivity indices can be employed as weighting factors of the objective function or to exclude certain less-relevant parameters at all from the OED process.
The concept of global sensitivity indices for the purpose of OED as suggested in Hölter et al. (2015) and described in Sec. 4.1, was adopted in Lo and Leung (2018) for the purpose of investigating a slope, demonstrating that the concepts introduced within this thesis can be transferred to other geotechnical problems. The examples described in this overview show on the one hand the diversity of application possibilities and on the other hand that various approaches exist. However, in the field of geotechnical engineering almost no applications of OED for parameter identification, apart Schanz and Meier (2008) from the same research group, had been published to the best of the author's knowledge. This motivates the present thesis to investigate which of the previously mentioned methods can be applied to relevant problems and how successful these can be performed.

2.3. Motivation in geotechnical engineering

2.3.1. Numerical modelling in geotechnical engineering

It should be mentioned that the concepts of OED presented within this thesis are supposed to be generally valid and can be applied to any type of model, ranging from conventional analytical functions to most sophisticated numerical models, as shown in the examples mentioned in Sec. 2.2. However, as in recent decades numerical simulations have been

increasingly employed for prediction and design in geotechnical engineering, the present thesis focuses on this type of models. This is of special relevance as such models can have various input details and can provide detailed results that include the whole geometric and time model domain. However, as the quality and reliability of such results is only given in case of validated models, the benefit of using concepts of OED is especially given in case of numerical models. Within this thesis, the finite element method (FEM) is utilised and if not stated elsewise, the commercial software packages Plaxis 2D (Plaxis bv, 2019) and Plaxis 3D (Plaxis bv, 2018) are employed.
When using the FEM, the soil medium is assumed to be a continuum that is discretised into a number of two- or three-dimensional elements. For each of these elements, stresses and strains are calculated and brought to equilibrium with external forces and deformations in a system of equations that includes all other elements. Afterwards, the material responses as stresses are available for all so-called Gauß-points in which the element properties are defined, while further responses as displacements are obtained in nodes or can be interpolated in between. The interpolation is convenient for the present work on OED, as thereby different measurement designs, corresponding to results from different positions of the model domain can be compared without further computational effort caused by additional Gauß-points or nodes. The employment of the FEM in geotechnical engineering has been established in science and practice in the last decades and is broadly presented in Potts (1999) and Potts and Zdravković (2001). More specific applications of the FEM in the context of validation of tunnel simulation models can be found e.g. in Miro et al. (2015) and Zhao et al. (2015, 2019), while in Müthing et al. (2018) and Mahmoudi et al. (2020) the settlement of an embankment and the stability of a slope are investigated, respectively.

2.3.2. Model validation and verification

Setting up an FE model might be an easy task sometimes, but one should consider that any new model needs to be approved what is often the more complex part of the modelling. "Approving" a numerical model, or any other kind of model, is effectuated by verification and validation. These terms can be found even in the titles of numerous publications from the field of geotechnical engineering e.g. (Gupta et al., 2008; Tsinidis et al., 2015; Zhao et al., 2015) and they should be clarified in this section to understand how they are related to the topic of OED.
In several publications, e.g. Balci (1998); Carson (2002); Thacker et al. (2004), approaches are presented that describe model verification and validation and how to differentiate these

two terms. Key statements of these authors are given in the following.
The part of verification should be performed first, but might be reconsidered after validation in the sense of an iterative process. Hereby, it should be controlled that all predefined steps that the model should consider are included and executed as agreed on by all stakeholders. The internal logic of the model must be ensured and reasonable results generated. Results should still be reasonable and continuous in case the input parameters of the model are varied within realistic ranges.
Validation indeed includes the aspect of comparison of the model predictions to certain data that is accepted as reference, like in-situ measurements or simulation data obtained from other approved models. By variation of the model parameters, the model outputs change accordingly and can be adapted to this data. A perfect fit of model outputs and measurements will not be possible due to unavoidable model limitations, but they should coincide with satisfying accuracy. That means, a comparison of confidence intervals is effectuated to find out whether the model outputs are within a certain range of acceptable discrepancy from the measurement data. If this is the case, the model can be accepted as valid. If not, the origin of this discrepancy should be identified, by deeper investigating the model behaviour. Beside, in case additional data is obtained, the validated model should be able to allow an interpretation of the changed conditions or to predict future events.
Thacker et al. (2004) summarises the description as follows:

- "Verification is the process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model."
- "Validation is the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model."

The explanations given above originate from publications in the context of production logistics and aerospace engineering. However, they seem to be generally valid to any research or application as they can be found similar in Brinkgreve (2013) where this topic is discussed with focus on geotechnical engineering. In this field, model validation is of special importance as the parameters of the considered soil materials are highly uncertain and represent a complex constitutive behaviour. Often, the model parameters are initially simply assumed as bandwidth with uniform distribution and they obtain specific values only by the process of validation. Besides, due to the complex material behaviour, the selection of an adequate constitutive model becomes an important task and part of the process of model verification and validation. To ensure an adequate representation

of the constitutive behaviour, element tests can be simulated for qualitative assessment of the soil behaviour. This approach corresponds to the statements of Brinkgreve (2013) and Balci (1998) that a model can only be valid if its submodels are all valid. As it is possible that different types of errors compensate for each other, a simple comparison of model response and reference data is not sufficient for model validation, but a deeper understanding of the model behaviour should be the objective, for example by sensitivity analyses or convergence tests of the FE discretisation.

An example of model verification is given in Hölter et al. (2016) where different constitutive models are compared with respect to their ability to reproduce the behaviour of Ballina clay on which a road embankment was supposed to be constructed. Selecting the correct constitutive model was part of the verification of the FE model, as it could be controlled if initial assumptions such as creep behaviour are realised in the model. However, as stated in Brinkgreve (2013), one should note that using a commercially available software package, such as Plaxis 2D (Plaxis bv, 2019) or Plaxis 3D (Plaxis bv, 2019) like in the case discussed here as well as in the other examples considered in this thesis, simplifies the process of verification as it avoids most errors due to coding mistakes and similar.

The validation of the aforementioned embankment model took place within a second step, described in Müthing et al. (2018). Here, specific measurement data from the considered site was employed and the model could be validated to identify the relevant soil parameters of the selected constitutive model. To this, statistical evaluations are performed, to investigate whether the assumed bandwidths of parameters produce results in reasonable ranges and whether the measured data is appropriately located within these ranges. Additionally, a global sensitivity analysis is performed to understand which parameters are most influencing the model results and should therefore be in focus of the investigations. Further examples of model validation corresponding to the concepts described herein can be found e.g. in (Schädler et al., 2015; Meier et al., 2013).

Another demonstrative example of model validation in geotechnical engineering is given in Zhao et al. (2015). Therein, it is described in detail how a model of mechanised tunnel construction is set up and then validated using in situ measurement data. However, it is evident that data from certain measurement points contribute more to the validation process than from others. Using an OED concept might have allowed to find a sensor arrangement that provides data for more reliable model validation to the involved researchers.

Within this context, the question arises where OED should be located in the context of verification and validation as this is the objective of the present thesis. As mentioned

above, validation requires reference data to improve a model and increase its accuracy, or decrease its uncertainty, respectively. The given descriptions indicate that validation should consist of several steps and, depending on the quality of the employed data, the validation process can be more or less successful, i.e. the uncertainty can be more or less reduced.
As stated above, data is needed to perform model validation, but due to the uncertainties described in Sec. 2.3.3, data is always to be questioned when employed for model validation. Vice versa, it means that to perform model validation one should always aspire to have valuable data. Finding out which type of data is most valuable and therefore allows the most reliable model validation is exactly the objective of OED and key issue of this thesis. However, as the OED procedure is model based, at least a verified model is required. That means that OED should be considered in between verification and validation, or as first step of validation.

2.3.3. Uncertainty in geotechnical engineering

The previous sections introduced the aspects of model validation to reduce the uncertainty of model parameters and explained that the objective of OED is to reduce this uncertainty. Therefore, it should be now discussed how uncertainty is assessed in geotechnical engineering and how it can be categorised. Within the framework of reliability-based design that will be described more in detail in Sec. 3.4, the question arises how a value can be assigned to soil properties that characterises their uncertainty so that they could be employed as design values.
It should be differentiated between input and output uncertainty. In the framework of this thesis, it is intended to transform measured data (the input) to constitutive parameter values (the output). The parameter values should be identified as accurate as possible to allow their precise prediction. The issue about uncertainty of this prediction is that the uncertainty in context of geotechnical engineering is not limited to one source, but according to Phoon and Kulhawy (1999*b*) to the three categories of inherent variability (described as soil heterogeneity in the following), measurement error, and transformation uncertainty named here as model uncertainty. The interaction of these aspects is rudimentary shown in Fig. 2.1. The figure describes that from the actual soil properties and those that are estimated to perform a simulation or to design a structure, the different sources of uncertainty are introduced. However, it is not apparent how to differ among the impact of these uncertainty sources, wherefore the following overview is given.

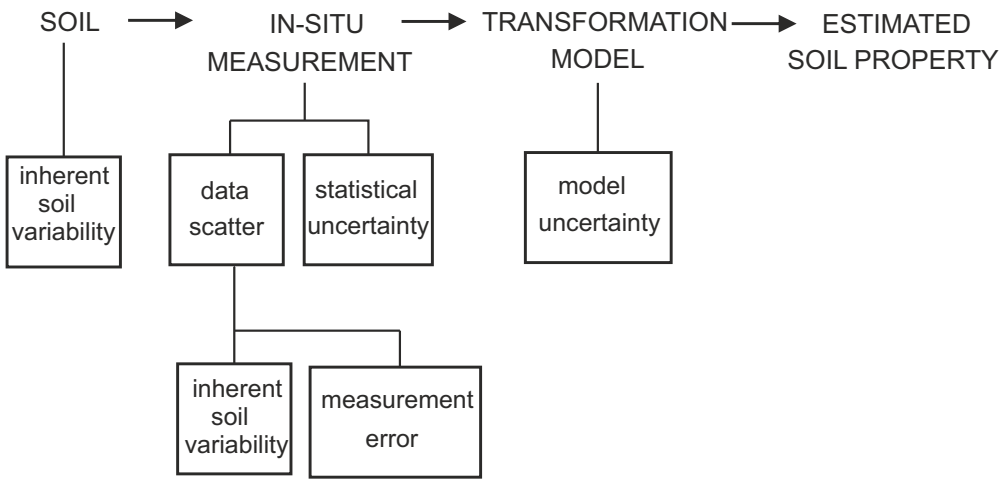


Figure 2.1.: Uncertainty in soil property estimates, after Kulhawy et al. (1992).

2.3.3.1. Soil heterogeneity

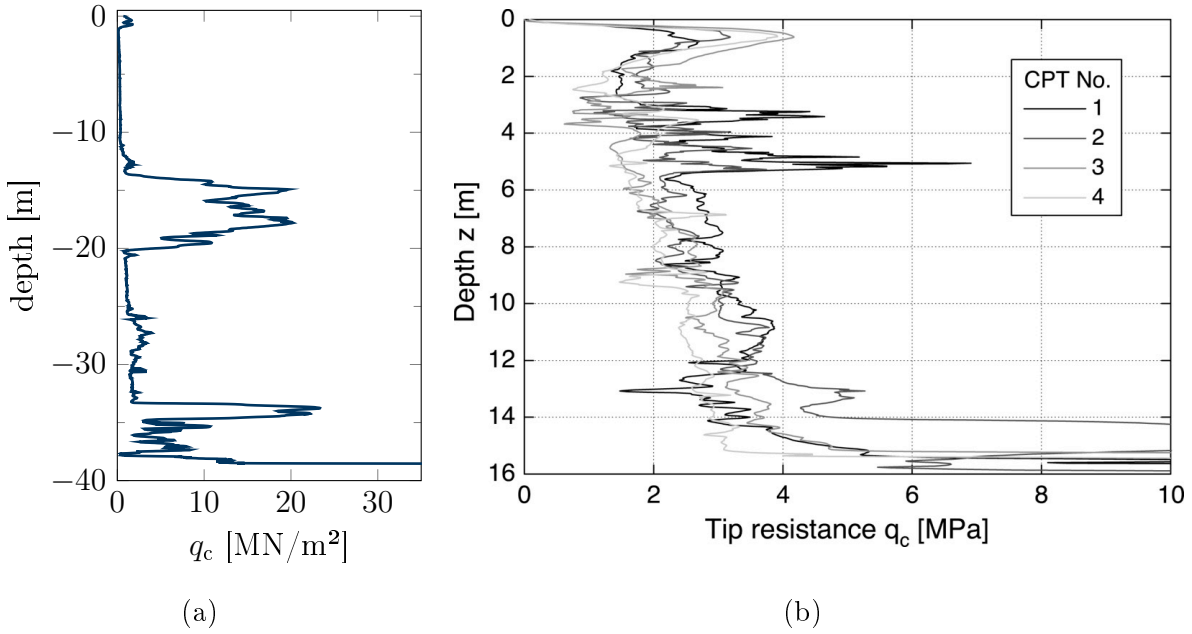
Naturally "grown" soils as they are encountered within many geotechnical problems are not homogeneous materials especially compared to other media encountered in civil engineering such as steel or concrete. They are the outcome of long lasting geological processes such as diagenesis or sedimentation. These processes overlap and are influenced by further effects such as overconsolidation and leaving behind of erratic boulders in case of ice ages, changing river courses after flood events, and any type of man-made impacts. The final outcome can be anything and needs to be characterised using in-situ or laboratory experiments. Also "artificially" built-up soil bodies as employed for dikes, dams, or embankments are not perfectly homogeneous. Due to irregular compaction during construction and production errors, but also through weather effects, a certain variation appears in such structures, requiring to take uncertainty into account.
Typical field tests to determine the soil properties and to find out about the stratigraphy are cone penetration tests or standard penetration tests. More detailed knowledge can be obtained by laboratory testing. Such tests can vary from simple classification tests according e.g. to DIN 18196 (2011-05) to sophisticated tests performed in oedometric or triaxial devices. However, these test are always local insights at discrete points of the considered (model) domain as they are performed on a marginal share of the whole soil body. Conventional sample sizes in triaxial devices are a diameter of 5 or 10 cm and a height of 10 or 20 cm, respectively. Corresponding values of oedometer devices are a diameter 7 cm and a height of 2 cm. This aspect of proportionally small samples becomes more relevant the larger the construction site is. The properties of the soil between the points in which the tests were performed are formally unknown. Instead of simple linear interpolation of

the properties, several approaches to identify properties of the soil between the known points have been developed in the last decades such as Kriging (Dubourg et al. (2011); Ragkousis et al. (2016)). However, still a large share of uncertainty remains on the actual values.

But also in case of a single location at the ground surface, the soil properties are not constant over depth as shown in Fig. 2.2. The data shown here was obtained using cone penetration tests (CPT) in Australia (Fig. 2.2a) and Germany (Fig. 2.2b). In Fig. 2.2a, it can be seen that at least 5 different soil layers are encountered (0 m to 12 m, 12 m to 20 m, 20 m to 33 m, 33 m to 38 m, and below 38 m). However, one should ask what parameter values should be applied for each of these layers as there is apparently strong fluctuation within the layers. Another question might be if five layers are a correct assumption, or if one should additionally subdivide some of them. In Fig. 2.2b, the results of four different tests, all performed on the same site of a dumped mining deposit, are shown. It can be seen that the four curves follow a similar trend that can probably be divided into four layers, but even more than in the previous figure, the problem becomes evident that any description of such soil would be affected by large uncertainty.

One approach to account for the soil's heterogeneity is the so called random field approach as applied among others in Mahmoudi et al. (2020); Mahmoudi (2017); Kasama and Whittle (2016). Hereby, the varying soil properties are reproduced using e.g. Gaußian distributions. The so-called correlation length defines in which distance the soil properties may vary. Reference correlation lengths can be found in Phoon and Kulhawy (1999*a*), however the lengths provided therein are quite vast and may not be representative for any application. The other drawback of the random field approach is the considerable computational effort that is required to perform a sufficient number of simulations until results converge to a reliable value, while they cannot be substituted by a metamodel as described in Sec. 3.2.

Within this thesis, the soil body is divided into different soil layers if discrepancies among different soil areas justify it, but within these layers, the soil properties are assumed homogeneous. The uncertainty arising from soil heterogeneity is considered by assuming a bandwidth for each constitutive parameter that is of relevance for the considered model response. Using random parameter samples within these bandwidths, it is intended to obtain model responses to any possible combination of soil parameters.



(a) (b)

Figure 2.2.: Tip resistance q_c of CPT investigations on two different sites a: from Ballina, Australia, obtained in the framework of data published in Pineda et al. (2016) b: from a dump in the Rhenish lignite-mining area (Wichtmann et al., 2019).

2.3.3.2. Measurement uncertainty

When generating measurement data, one should be aware that they include some technically unavoidable misfit between the physical value and the recorded one. Describing this error is very complex. The strain gauges for example produced by the company Glötzl that are commonly employed in geotechnical applications as the one described in Detert et al. (2020) are provided with specific ranges that indicate the measurement accuracy (Glöztl GmbH, 2020). The indication of such devices is generated by varying the resistance to an electrical circle and measuring the voltage. Even though the voltage metering is highly accurate, several aspects may falsify the device's output. The resistance tolerance e.g. describes within which range the actual resistance may be related to the indicated one. Besides this epistemic error, further aleatoric measurement errors may arise from inaccuracies in calibration or installation. Such highly sensitive devices can be influenced by further external impacts like temperature, rain, or fluctuating external effects that are not part of the foreseen loading. Depending on the type of sensor, different error types may result that can be described by means of statistical moments (mean, standard deviation, bias, skewness, etc.). Within the present thesis, only Gaußian white noise is considered as type of measurement error, as extending the performed investigations also to higher order

moments would be out of the scope of the work. In Reichert et al. (2019) andSchenkendorf (2014) indeed, the impact of different types of noise on the results of OED investigations is analysed. To account for the unbiased or white noise, artificial error is generated using random sampling methods that generate distributions of values according to the employed Gaußian distributions. Hereby, the magnitude of error is defined by the standard deviation of the employed distribution. In case such falsified or noisy data is employed to back-calculate model parameters in a process of model validation, this probably leads to an ill-posed problem. When identifying a solution that leads to the least discrepancy between model response and measured data, it should therefore always be considered how reliable the employed data is. Accordingly, repeated back analyses on different data points are desirable, leading probably to a certain range of results by which the measurement error induced parameter uncertainty can be described. This procedure is applied e.g. in Alkam and Lahmer (2019), where concrete parameters of prestressed railway poles are identified using observed data. It is shown indeed that using different approaches of parameter identification, the obtained uncertainty on these parameters varies. However, as in geotechnical engineering practice experiments can often not be repeated, it is advisable to employ different types of measurements simultaneously or over a long time period as shown in Knabe et al. (2012) and Müthing et al. (2018). In both cases, it is intended to reproduce and predict the time-dependent behaviour of an embankment on soft ground. As such "experiment" is not repeatable for reasons of time and money, displacement and pore water pressure data is employed to identify the constitutive parameters of the underlying soil and to reduce the uncertainty of the model parameters. As still considerable discrepancy between model prediction and measurements remains, probably a more systematic error is present that is hard to be identified if additionally the aspect of model uncertainty is relevant as described in the following.

2.3.3.3. Model uncertainty

Model uncertainty can be described as epistemic uncertainty, as it originates from the fact that it is not possible or not known how to perfectly adapt a model to the real phenomenon that is intended to be reproduced. It can have many origins, as the term "model" itself is somehow uncertain. In case of the geometrical model, uncertainty might arise from wrong dimensions: If one intends to calculate the bearing capacity of a shallow foundation of 1 m width, but a width of 1.5 m is employed, the model is obviously wrong and will overestimate the bearing capacity. Further uncertainty arises from the employed constitutive model that should be able to reproduce the soil properties accurately. Simulating a structured

clay with the Hardening soil small strain model will not be successful although it is a quite sophisticated model. In Hölter et al. (2016) e.g., it is demonstrated how different constitutive models perform in reproducing data from a constant rate of strain test. As none of them allows a "perfect match", model uncertainty might be propagated to the subsequent calculations by an inaccurate constitutive model.
Finally the calculation model must be adequate. E.g. in Motra et al. (2016), 60 different models to calculate the bearing capacity of shallow foundations are investigated. As they provide 60 different results, at least 59 of them are not correct. But does it mean they are wrong? Better not to say so. Depending on the specific application, each of them might have its advantages, wherefore a considerate selection of the model is always necessary as it is not possible to say what is the true model. An evaluation of model uncertainty in comparison to measurement data is given in Tang et al. (2017), while a more general overview is provided in Phoon and Tang (2019). Identifying a model-based OED cannot be successful if the model is not accurate. Even though the detailed assessment of model uncertainty is out of the scope of this thesis, and it will be assumed that the employed models are "correct", an awareness of this issue should be risen. For applications based on real measurement data, the aspect of model uncertainty still becomes relevant as the reduction of discrepancy between model response and measurements cannot be more reduced than to the error arising from the inaccurate model itself.

2.4. Methodology

2.4.1. Parameter uncertainty

The aspects of uncertainty described in the previous sections must be reflected in the employed constitutive parameter values when applying them to a simulation model. In both cases, using purely synthetic examples or real scenario based, a very low initial knowledge level is assumed in this thesis, corresponding to highly uncertain soil parameters. In Miro (2016), this is called "pessimistic scenario" in contrast to situations in which more detailed knowledge about the actual parameter values is given. The employed data should correspond to the knowledge an engineer would have if he or she gets the information "this is a normally consolidated clay" or "this is a coarse grained sand". Such descriptions are normatively defined e.g. in DIN 18196 (2011-05) and DIN EN ISO 14688 (2018-05). In generally accepted textbooks like Das (2006), Mitchell and Soga (2005), and Lambe and Whitman (1969), adequate ranges of parameters can be found for the different types of

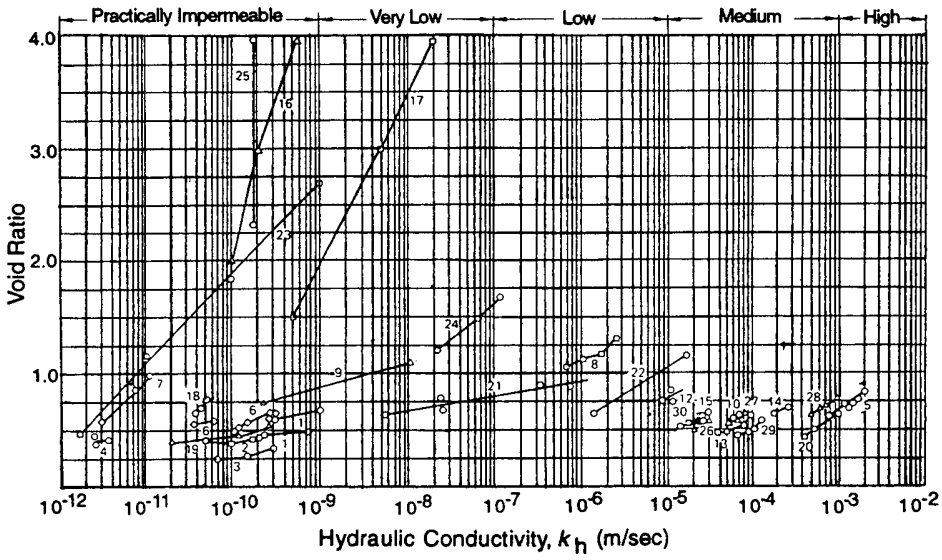


Figure 2.3.: Hydraulic conductivity values for several soils from Lambe and Whitman (1969)

Table 2.1.: Maximum and minimum void ratios, porosities, and unit weights for several granular soils, modified from Lambe and Whitman (1969).

	Void ratio [-]		Porosity [%]		Dry unit weight [kN/m^3]	
	e_{max}	e_{min}	n_{max}	n_{min}	$\gamma_{d,min}$	$\gamma_{d,max}$
Uniform spheres	0.91	0.35	47.6	26	-	-
Standard Ottawa sand	0.80	0.50	44	33	14.5	17.3
Clean uniform sand	1.0	0.40	50	29	13.0	18.5
Uniform inorganic silt	1.1	0.40	52	29	12.6	18.5
Silty sand	0.90	0.30	47	23	13.7	20.0
Fine to coarse sand	0.95	0.20	49	17	13.4	21.7
Micaceous sand	1.2	0.40	55	29	11.9	18.9
Silty sand and gravel	0.85	0.14	46	12	14.0	22.9

soils. Exemplary datasets are given in Fig. 2.3 and Tab. 2.1.
The diagram shown in Fig. 2.3 displays the relationship of void ratio and hydraulic conductivity (that is equivalent to the term of "coefficient of permeability" that is employed further on in this thesis) of 30 different types of soil. Tab. 2.1, from the same textbook, provides in tabular manner bandwidths of void ratio, porosity, and soil weights. Such empirical reference data sets allow to asses parameter values, depending on the considered type of soil. The investigations performed in this thesis employ constitutive parameter ranges originating from such data sets. In terms of probability distributions, these ranges should be assumed as uniformly distributed. By means of back analysis of either artificial or in-situ measurement data, more precise knowledge on the parameters is gained and their bandwidths can be adapted by confining them or by modification of the distribution type.

2.4.2. Covariance matrix $C_{\boldsymbol{\theta}}$

The assumption of parameter ranges and the uncertainty that comes along with them, mentioned in the previous Sec. 2.4.1, refers to the initial assumptions that are necessary to set up a model where the exact material parameter values are unknown. One objective of model validation is now to gain knowledge on these parameters by back calculating measured data. If this is performed for a set of data points $\tilde{\boldsymbol{y}}$, as many parameter samples $\tilde{\boldsymbol{\theta}}$ are obtained. The parameters should be somehow distributed within the initially assumed range what should be seen as a reduction of the parameter uncertainty.
This uncertainty can be described by the variance $\sigma^2(\tilde{\boldsymbol{\theta}})$. In contrast to the initial assumptions mentioned afore, the variance or standard deviation σ does not depend on assumptions, but on the data $\tilde{\boldsymbol{y}}$ that again depends on the employed experimental design $\boldsymbol{\delta}$. Therefore, the variance must be calculated exactly to be able to compare the efficiency or quality of different experimental designs. The variance of one single variable, or parameter, θ describes how strong the sample of size N deviates by its individuals $\tilde{\theta}_i$ in average from its mean value $\bar{\theta}$ as described in Eq. 2.4:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} \left(\tilde{\theta}_i - \bar{\theta} \right)^2 \tag{2.4}$$

However, in case of several parameters as encountered in the applications in this thesis, one should also consider the so called covariance that accounts for the correlation among the different individual parameters. To include the correlation of all considered parameters,

the covariances are described in the covariance matrix $\boldsymbol{C_\theta}$, of dimension s x s:

$$\boldsymbol{C_\theta} = \frac{1}{N}\sum_{i=1}^{N}(\tilde{\boldsymbol{\theta}}_i - \overline{\boldsymbol{\theta}})(\tilde{\boldsymbol{\theta}}_i - \overline{\boldsymbol{\theta}})^T = \begin{pmatrix} \sigma^2(\theta_1) & \sigma(\theta_1)\sigma(\theta_2) & \dots & \sigma(\theta_1)\sigma(\theta_s) \\ & \vdots & & \\ \sigma(\theta_s)\sigma(\theta_1) & \sigma(\theta_s)\sigma(\theta_2) & \dots & \sigma^2(\theta_s) \end{pmatrix} \quad (2.5)$$

This matrix describes the correlation among all considered parameters of the model. A more visual representation is provided in Fig. 2.4.

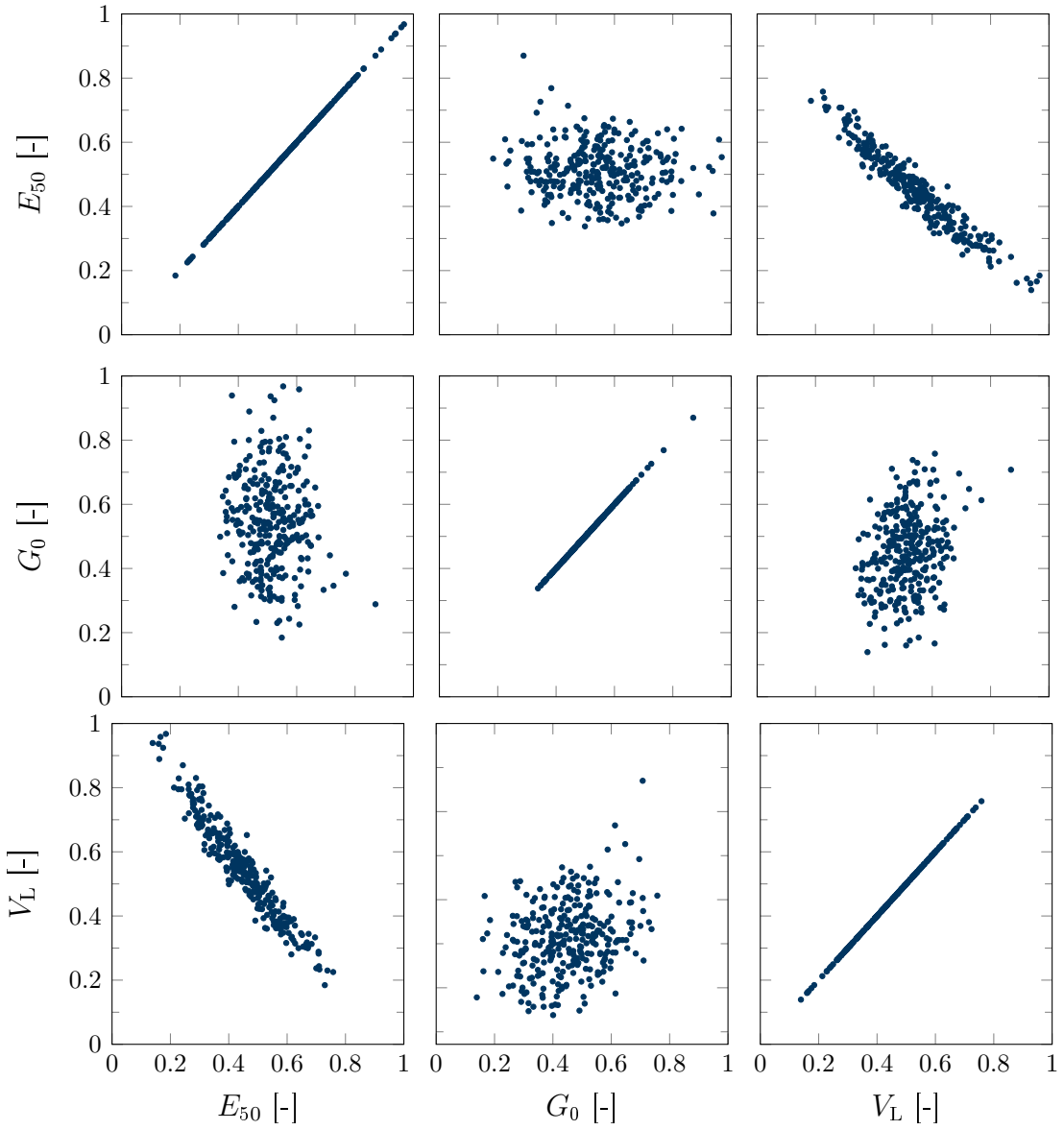


Figure 2.4.: Correlation plots of an exemplary case with three parameters E_{50}, G_0, and V_L with 300 back-calculated samples showing different amount of correlation among the different parameters.

The three parameters E_{50} (pressure-referenced Young's modulus), G_0 (pressure-referenced small-strain shear modulus), and V_L (volume loss) are herein back-calculated 300 times from vertical settlements to identify the relevant parameters of a simulation of tunnel construction that is described more in detail in Sec. 6. The back-calculated settlement values diverge by a normally distributed random error that was added to the original value. The nine plots in Fig. 2.4, in which the parameter combinations are plotted in a normalised way within the allowable ranges of the metamodel, reflect the nine elements that the corresponding covariance matrix would have. The elements on the symmetry axis are represented by straight lines as they describe the correlation of one parameter with itself that will always be one. In contrast, the correlations between G_0 and E_{50} as well as G_0 and V_L seem to be very weak as these plots show dispersed data clouds. The more the scattering is circular, the less the parameters are correlated and the wider the scattering is, the larger the value of the covariance is. In case of E_{50} and V_L, a strong negative correlation can be observed. These observations make sense as G_0 is relevant for settlements in the range of small strains, while E_{50} and V_L have similar impacts in case of larger strains. Therefore, it is plausible that in case one parameter is high, the other one must be smaller.

2.4.3. Optimality, Cost, and Utility functions

Within the process of OED, it is intended to identify the optimal experimental design $\boldsymbol{\delta}^*$. "Optimal" means that experimental design that allows to identify the parameters of interest with least uncertainty. In case of a two-dimensional parameter space, Fig. 2.5 outlines the scheme of parameter identification using OED. Corresponding to the correlation plots shown in Fig. 2.4, the different experiments are represented as ellipses that reflect the bandwidth of probability of the two parameters. The least efficient experiment "1" allows the least accurate parameter identification. Further improved experiments allow a more accurate parameter identification and a reduction of the parameter ellipse.

This uncertainty is numerically expressed in the covariance matrix $\boldsymbol{C_\theta}$ as described in the previous section or by another information matrix as the FIM described in Sec. 2.2.1. However, as a matrix has several elements, the question arises how to rate them with respect to the specific requirements of OED. Therefore, so-called criterion functions Φ are employed that are also known as optimality function, cost function, or utility function, whereby the latter one is to be maximised while the others are to be minimised. As the constitutive parameters do not depend on the experimental design, the following relation

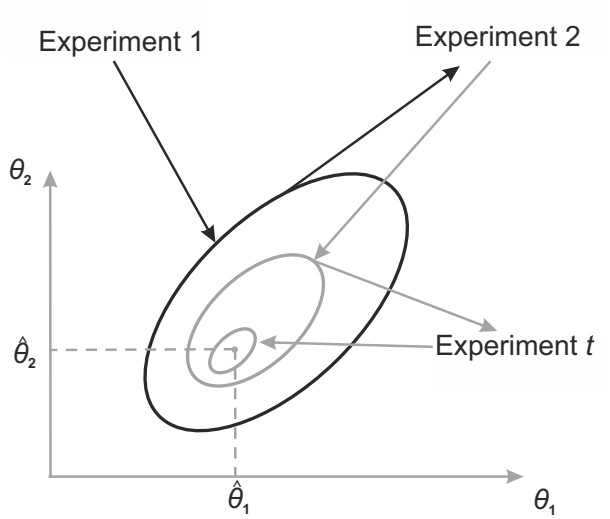


Figure 2.5.: Concept of OED-process for a two-dimensional parameter space (modified after Schenkendorf et al. (2009)).

is given.

$$\Phi_{i,j} = f\left(\boldsymbol{C}_{\boldsymbol{\theta}} | \boldsymbol{\delta}_j\right) \ , j \in 1, \ldots, t \tag{2.6}$$

Herein, the index i indicates that different variants of the criterion function Φ_i are possible and t is the total number of considered experimental designs. In Nishii (1993), an overview is given on the different possible criteria that can be employed, whereby any further is conceivable. In the literature, several criteria are considered as follow:

$$\Phi_{\mathrm{D}} = \det(\boldsymbol{C}_{\boldsymbol{\theta}}) \tag{2.7}$$

$$\Phi_{\mathrm{A}} = \mathrm{tr}(\boldsymbol{C}_{\boldsymbol{\theta}}) \tag{2.8}$$

$$\Phi_{\mathrm{M}} = \max_i \sqrt{\boldsymbol{C}_{\boldsymbol{\theta},ii}} \tag{2.9}$$

$$\Phi_{\mathrm{E}} = \min \lambda(\boldsymbol{C}_{\boldsymbol{\theta}}) \tag{2.10}$$

$$\Phi_{\mathrm{E}}^{*} = \frac{\max \lambda(\boldsymbol{C}_{\boldsymbol{\theta}})}{\min \lambda(\boldsymbol{C}_{\boldsymbol{\theta}})} \tag{2.11}$$

By applying these functions to $\boldsymbol{C}_{\boldsymbol{\theta}}$, a discrete value is assigned to each considered experimental design, allowing to rank the designs or to employ optimisation algorithms. It is indicated that depending on the application, the different criteria have different suitability. In Nishii (1993) e.g., it is described that the Φ_{A} criterion does not consider any correlation among the parameters, wherefore it should be used in case all parameters have the same importance. However, different researches and investigations of the author show that no

generally valid rules for the selection of the optimality criterion can be provided. It might be rather recommended to employ several of the criteria at the same time and to compare the results. A certain experience of the user is required to interpret the results adequately and to select the most proper one for the current application. In Sec. 5, this procedure is followed by applying both Φ_{D} and Φ_{A} to the obtained covariance matrices. The resulting rankings are interpreted according to engineering judgement to identify which criterion provides more reliable results.

3. Fundamental mathematical methods

3.1. General

In principle, the whole topic of OED can be described as a conventional optimisation problem: for a given function $\Phi(\boldsymbol{\delta}, \boldsymbol{\theta})$, find that $\boldsymbol{\delta}$ that provides the minimum response of Φ, whereby, $\boldsymbol{\delta}$ denotes a possible experimental design and $\boldsymbol{\theta}$ is a vector of the employed model parameters. Φ is the criterion that is applied to the result of the optimisation algorithm to evaluate how good the uncertain parameters $\boldsymbol{\theta}$ can be identified. However, solving this problem implies several non-trivial steps. To identify soil parameters based on observed values, an inverse analysis is necessary, whereby the term of back analysis can also be found for the considered applications.
Depending on the procedure how the back-analysis is performed, it will still demand numerous calls of $\Phi(\boldsymbol{\delta}, \boldsymbol{\theta})$ to allow the process of back-analysis to converge, as such optimisation problems are mostly solved iteratively. This is feasible in case an explicit function is at hand, but in this thesis numerical solutions of Finite-Element models (FE) instead of analytical ones are considered whose calculations may need several hours. Therefore, so called meta- or surrogate models are employed that substitute the computationally expensive FE-model by relating the input parameters $\boldsymbol{\theta}$ to the model outputs $\boldsymbol{y}(\boldsymbol{\theta})$ by means of a mathematical regression $\hat{f}(\boldsymbol{\theta})$. Using such metamodels, it is possible to perform numerous evaluations of the problem at hand within short time. By employing input values that are sampled according to specific probability density functions (PDF), probabilistic calculations become possible that reflect the system response not by a single value, but again as distribution of values. The use of PDFs as model inputs, or their approximation by Monte-Carlo sampling, is one background of the topic of Bayesian probability presented in chapter 6. Here, PDFs are not just employed to describe parameter and model response distributions, but also to evaluate how good a certain OED works to retrace an initial distribution of uncertain input parameters.

These four methods: metamodelling, back analysis, probabilistic analysis, and Bayesian inference, shall be introduced in the following sections to provide the essential basics to the methods employed in the later chapters of this thesis.

3.2. Metamodelling

3.2.1. Introduction

Employing the approaches of OED that are described in this thesis needs a lot of effort and expert knowledge and most probably it will benefit and find application in complex problems of geotechnical engineering like dikes, tunnels, or deep excavation pits. As described in Sec. 2.3.1, such cases are nowadays evaluated using numerical methods such as the Finite-Element method rather than with analytical approaches.
The drawback of this method is the considerable computational effort that is necessary to obtain the numerical solution. Simulations with simple 2D models can be performed within few minutes, but complex 3D-simulations of tunnel excavations, as considered in the following chapters, might need several hours, depending on the level of details of the model. In case of one single model run, this might be affordable, but for the intended operations in this thesis such as parameter identification (Sec. 3.3), probabilistic evaluations (Sec. 3.4), and global sensitivity analysis (Chapter 4), numerous model calls are necessary, but would be too time-consuming.
To avoid this issue, so-called metamodels can be employed. Such models are also referred to as surrogate models as e.g. in Cao et al. (2016) and Matott and Rabideau (2008) but this basically means the same: an approach to approximate the mechanisms that translates input to output data of a certain model by an analytical function. This analytical function is computationally cheap to call and hence well suited for the aforementioned operations. However, one should keep in mind that for calibrating the metamodel still a certain amount of model runs are necessary, and this might still be highly time consuming. The details and challenges of generating a metamodel in the context of geotechnical engineering can be found in Miro (2016), Khaledi (2017), and Zhao (2018). To refine the focus of these aspects on the topic of *OED*, the descriptions of the following sections are provided. Of course, numerous further options are available to generate metamodels or similar approaches to substitute computationally cost-intensive models such as neural networks in e.g. Freitag et al. (2018), or polynomial chaos expansion (Blatman and Sudret, 2010), and Kriging (Dubourg et al. (2011); Ragkousis et al. (2016)), but the following

review refers to those employed in this thesis that can be related to the group of response surface surrogates.

3.2.2. Metamodel data generation

Consider a given model $f : \boldsymbol{\Theta} \rightarrow \mathfrak{Y}$ that depends on the parameters $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, whereby $\boldsymbol{\Theta}$ is a predefined and constrained parameter space of dimension s. In the context of this thesis, $f(\boldsymbol{\theta})$ are approximations to initial and boundary value problems that are formulated in a FE model that delivers the outputs $\boldsymbol{y}$. To set up a reliable metamodel $\hat{f}$, that is able to approximate the model f, it is necessary to generate a sufficiently large database in the input but as well as in the output space $\mathfrak{Y}$. A set $\mathfrak{T} = \{\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_N\}$ of input samples is generated in the input space and it is run in the model to obtain the corresponding responses $\boldsymbol{y} = \{\boldsymbol{y}_1, ..., \boldsymbol{y}_N\}$. Input sample, as mentioned above called $\boldsymbol{\theta}$, is a combination of model parameters, that is selected e.g. randomly or in another way among the members of the model parameter values space $\boldsymbol{\Theta}$. An individual set of samples $\mathfrak{T}$ indeed assembles a large amount of input samples $\boldsymbol{\theta}$.
Hereby, several aspects should be considered. The generated set $\mathfrak{T}$ of samples should cover the input space in a way that all relevant features of the input-output relationship of the model can be captured. Finding a way to efficiently generate meaningful samples in a defined parameter space is a similar problem as the problem of OED. While the overall objective of this thesis is to provide approaches to optimise in-situ or laboratory experiments, in this section, optimal sampling designs in computer experiments are described, often called as design of experiments (DOE). In Giunta et al. (2003), two methods are referred to as "classical DOE" and "modern DOE", stating that the essential difference consists in the fact that the "classical" DOE refers to experiments in which random error exists, while this is not the case for computer experiments. That strict discrimination should be employed very well-considered as this terminology is not generally accepted. Both names and concepts are close to each other and an explicit differentiation might not be possible as they might even overlap in certain applications.
In Pronzato (2008), this is pointed out by the given description of DOE that well fits to both problems as it is stated that "DOE, which can be apprehended as a technique for extracting the most useful information from data to be collected, is thus a central (and sometimes hidden) methodology in every occasion where unknown quantities must be estimated and the choice of a method for this estimation is open. DOE may therefore serve different purposes and happens to be a suitable vehicle for establishing links between problems like optimization, estimation, prediction and control."

As in the present thesis it is intended to present optimal experimental designs by using data obtained from computer simulations, a short overview on the concepts for design of computer experiments is provided.
The most often encountered way of generating parameter samples for a metamodel generation is to use "random" sampling, known as Monte Carlo (MC) sampling introduced by Metropolis and Ulam (1949). The samples are not really random, but generated according to a defined algorithm as described e.g. in von Neumann (1951), wherefore this method might also be described as pseudo-random sampling. Due to the nature of randomness, certain areas of the parameter space may stay empty while in other areas several clusters of samples may arise. Samples generated by MC are not correlated with each other such that it is possible to exclude parts of the set, e.g to obtain a test set or to readjust the ranges of the parameter boundaries.
However, with increasing number of parameters, equivalent to the dimension of the input space, the density of parameter samples in this space is decreasing. Therefore, it might be recommended to use other sampling methods such as the Latin Hypercube Sampling (LHS), introduced by McKay et al. (1979). Hereby, the s-dimensional design space is divided into a multidimensional grid of N^s cells, where N describes the number of selected samples. Each of the N samples is placed randomly within one cell, but with the restriction that for each one-dimensional projection of the cells only one sample is admitted, comparable to the popular brain twister called Sudoku, while it has been employed for simulations of geotechnical problems e.g. in Zhao et al. (2015) and Miro et al. (2014).
The MC and LHS sampling methods are introduced here, as they are employed in the following applications, while numerous others exist, such as Orthogonal Array sampling or Quasi-Monte Carlo sampling that can be found e.g. in Giunta et al. (2003). An overview on previously developed sampling strategies for comparable computer experiments is indeed given in Sacks et al. (1989).
Further relevant aspects that must be considered before generating a set of parameter samples are adequate boundaries for each of the considered parameters. In some cases it might be obvious, e.g. the parameter range for cohesion should not include negative values, but often a well-considered decision is necessary. Parameters should represent a physically meaningful behaviour that allows a realistic representation of natural conditions in the considered engineering application. In case it is known e.g. that in a hydro-mechanically coupled problem the soil consists of sand, it is required to limit the range of the coefficient of permeability between 10^{-6} to 10^{-3} m/s. However, parameters must be consistent with the soil type. For a given soil, correlations exist that relate density and permeability. In case the coefficient of permeability is very high, it is most probable that the density is low.

Therefore, it would be beneficial to add restrictions in the sampling to exclude unrealistic parameter combinations. Employing such restrictions reduces the number of samples to be applied to the model calculated or increases the significance of the same amount of employed samples in the admitted area, respectively.
The crucial question how many samples are necessary to generate a good metamodel has not been considered up to now and there might be no satisfying answer. In case the response surface of the model is perfectly linear, it might be sufficient to have a minimum number of samples, but linearity is often not the case. The necessary set of samples will increase with the non-linearity and dimensionality of the problem, but also depends on the selected approximation function. In Zhao and Xue (2010) and Khaledi et al. (2014), detailed studies are presented on how the aforementioned aspects are related to the sample size and sampling type, whereby the latter publication explicitly refers to applications in geotechnical engineering. In these publications, the impact of sample size, dimension, and non-linearity on the model accuracy was proven, but no generally valid solution can be provided. Finding the adequate sample size demands experience and knowledge about the considered model and its approximation function. For most examples presented in this thesis sample sizes between 100 and 250 samples are selected, a range that has been succesfully employed in comparable applications e.g. in Miro et al. (2014) and Zhao et al. (2015). Finally, the only way to understand how accurate a metamodel is, is to perform an accuracy test as described in Sec 3.2.4.

3.2.3. Approximation function

To create a metamodel that is able to reproduce the characteristics of the original model, it is necessary that the approximation functions fit the model responses of the generated samples as accurately as possible. To this, numerous approximation approaches have been developed in the past decades of which those employed in this work are described in the following.

3.2.3.1. Least square method

A well-proven and efficient method is the so-called least-square regression where multivariate polynomial basis functions are employed to approximate model outputs or any other type of data. In the present work, polynomials of degree two are employed. The Least Square Method (LSM) is accessibly described in Press et al. (2007), but was developed 200 years ago in parallel by C.F. Gauß and A.M. Legendre and published by the latter

one (Legendre, 1805). It refers to the optimisation problem shown in Eq. 3.1 that intends to find the minimum $J_{\min}$ of the sum of squares of distances between the model responses $\boldsymbol{y}$ and the approximation $\hat{\boldsymbol{f}}(\boldsymbol{\theta})$.

$$J_{\min}(\boldsymbol{c}) = \min \sum_{i}^{N} ||\boldsymbol{y}_i - \hat{\boldsymbol{f}}\left(\boldsymbol{\theta}_i, \boldsymbol{c}\right)||^2 \tag{3.1}$$

N refers to the number of $i \in [1...N]$ sampling points that are located in a s-dimensional parameter space $\boldsymbol{\Theta}$. To modify $\hat{f}$ in a manner that it solves the optimisation problem in Eq. 3.1, the vector $\boldsymbol{c}$ of regression coefficients is applied to the matrix of polynomial basis terms $\boldsymbol{b}\left(\boldsymbol{\theta}\right)$ that allow efficient differentiability, as shown in Eq. 3.2, as being polynomials.

$$\hat{\boldsymbol{f}}(\mathfrak{T}) = \hat{\boldsymbol{f}}(\boldsymbol{\theta}|c_0 \dots c_k) = \boldsymbol{b}\left(\mathfrak{T}\right)\boldsymbol{c}, \tag{3.2}$$

whereby the number of coefficients k depends on the dimension s of the parameter space $\boldsymbol{\Theta}$ and the order of the employed polynomials m:

$$k = \frac{(s+m)!}{m!s!} \tag{3.3}$$

Accordingly, the vectors $\boldsymbol{b}_i\left(\boldsymbol{\theta}\right)$ of which the matrix of polynomial terms is composed, will have the same length of k elements. In the exemplary case of $m = 2$, the following matrix results:

$$\boldsymbol{b}_i = [\theta_{i,1}^2, \theta_{i,2}^2, ..., \theta_{i,s}^2, \theta_{i,1}\theta_{i,2}, \theta_{i,1}\theta_{i,1}, ..., \theta_{i,s-1}\theta_{i,s}, \theta_{i,1}, \theta_{i,1}, ..., \theta_{i,s}] \tag{3.4}$$

To find the minimum of the objective formulated in Eq. 3.1, a system of linear equations is generated and solved according to Eq. 3.5

$$\boldsymbol{c} = \left[\sum_{i=1}^{N} \boldsymbol{b}_i \boldsymbol{b}_i^{\mathrm{T}}\right]^{-1} \sum_{i=1}^{N} \boldsymbol{b}_i \boldsymbol{y}_i \tag{3.5}$$

Extensions of the least square concept are the Weighted Least Square where an additional weighting term is introduced to penalise sample points that cause model responses with large discrepancies and the Moving Least Square (MLS) approach, introduced in Lancaster and Salkauskas (1981) where the least square approximation is performed individually for each available sample point.

3.2.3.2. POD RBF

The other approximation approach employed herein is the Proper Orthogonal Decomposition with Radial Basis Functions (POD RBF). Using the concept introduced by Buljak (2012), the part of POD starts by the N x n snapshot matrix $\boldsymbol{Y}_{\mathrm{POD}}$ of outputs whose number of rows and columns is corresponding to the number of parameter samples N and model responses n, respectively. By means of the eigenvalues λ and eigenvectors $\boldsymbol{\nu}$ of its Gramian, both having the length N, the snapshot matrix can be decomposed into the matrix of proper orthogonal basis vectors $\boldsymbol{\Phi}_{\mathrm{POD}} = [\boldsymbol{\varphi}_{\mathrm{POD},1} \dots \boldsymbol{\varphi}_{\mathrm{POD},n}]$ (3.6) and the amplitude matrix $\boldsymbol{A}$, as shown in Eq. 3.7. The basis vectors $\boldsymbol{\varphi}_{\mathrm{POD}}$ have a length of n such that $\boldsymbol{\Phi}_{\mathrm{POD}}$ is of dimension N x n.

$$\boldsymbol{\varphi}_{\mathrm{POD},i} = \boldsymbol{Y}_{\mathrm{POD},i} \cdot \boldsymbol{\nu}_i \cdot \lambda_i^{-1/2} \ , \ i = 1, \dots N \tag{3.6}$$

$$\boldsymbol{Y}_{\mathrm{POD}} = \boldsymbol{\Phi}_{\mathrm{POD}} \boldsymbol{A}, \qquad \boldsymbol{Y}_{\mathrm{POD}} \boldsymbol{A} = \boldsymbol{\Phi}_{\mathrm{POD}}^{\mathrm{T}} \tag{3.7}$$

The matrix $\boldsymbol{\Phi}_{\mathrm{POD}}$ is sorted in decreasing order of the eigenvalues λ_i. Thereby, it is possible to define a threshold value from which on all smaller values are neglected, defining the residual part as "truncated POD basis" $\hat{\boldsymbol{\Phi}}_{\mathrm{POD}}$. Accordingly, the amplitude matrix $\boldsymbol{A}$ is reduced to $\hat{\boldsymbol{A}}$.
Subsequently, in the framework of RBF, radial functions are employed to fit the reduced amplitude matrix $\hat{\boldsymbol{A}}$. To this, several types of radial functions, such as linear, cubic, or Gaussian can be found in literature, e.g. in Krishnamurthy (2005), that characterise the discrepancy between two parameter samples. In the present work, the multiquadratic approach described in Eq. 3.8 is employed.

$$g_i(\boldsymbol{\theta}) = \left(||\boldsymbol{\theta} - \boldsymbol{\theta}_i||^2 + c^2 \right)^{-0.5} \ , \ i = 1, \dots N \tag{3.8}$$

The smoothness of the functions is controlled by the shape parameter c that should be adapted by the user to improve the fit of the metamodel. For each of the N parameter samples, the matrix $\boldsymbol{G}$ is computed:

$$\boldsymbol{G} = \begin{pmatrix} g_1(\boldsymbol{\theta}_1) & g_1(\boldsymbol{\theta}_2) & \dots & g_1(\boldsymbol{\theta}_N) \\ & \vdots & & \\ g_N(\boldsymbol{\theta}_1) & g_N(\boldsymbol{\theta}_2) & \dots & g_N(\boldsymbol{\theta}_N) \end{pmatrix} \tag{3.9}$$

Next, the matrix of interpolation coefficients $\mathfrak{B}$ with dimension N x s must be determined, using Eq. 3.10.

$$\mathfrak{B} = \hat{\boldsymbol{A}} \, \boldsymbol{G}^{-1} \tag{3.10}$$

Having finally these components, any parameter sample $\boldsymbol{\theta}$ within the defined boundaries of $\boldsymbol{\Theta}$ can be employed to generate the vector of metamodel responses $\hat{\boldsymbol{f}}(\boldsymbol{\theta})$ of length n using the following formulation:

$$\hat{\boldsymbol{f}}(\boldsymbol{\theta}) = \hat{\boldsymbol{\Phi}}_{\mathrm{POD}} \; \mathfrak{B} \; \boldsymbol{g}(\boldsymbol{\theta}) \tag{3.11}$$

A possible extension to this method is to use the so-called Extended Radial Basis Functions (POD ERBF) as described in Khaledi et al. (2014). To describe the distance between a model response and the corresponding response of the metamodel for same input data, that extension employs, besides the radial basis functions, non-radial functions, too.

3.2.4. Goodness of metamodel

As described above, the most important feature of a metamodel is its agreement with the original data, or prediction goodness. Based on experience gained by means of several applications of the techniques described in this chapter, one might become aware of the correlations between model complexity, sample size, and approximation algorithm, but it is impossible to know the metamodel's prediction goodness in advance. Therefore, to ensure this goodness, every metamodel has to be tested before being employed. In the present work, an independent test sample of input parameters is generated and run in the numerical model and the metamodel to compare both results. The comparison is performed using the normalised root mean square error $NRMSE$ (Eq. 3.12) or the coefficient of determination R^2, shown in Eq. 3.13.

$$NRMSE = \sqrt{\frac{\sum_i^n (y_i - \hat{f}(\boldsymbol{\theta}_i))^2}{n \cdot (y_{max} - y_{min})^2}} \tag{3.12}$$

$$R^2 = 1 - \frac{\sum_i^n (y_i - \hat{f}(\boldsymbol{\theta}_i))^2}{\sum_i^n (y_i - \bar{y})^2} \tag{3.13}$$

Here, $\hat{f}(\boldsymbol{\theta}_i)$ represents the n metamodel responses to the i^{th} sample and $\bar{y}$ is the mean of the numerical model responses obtained to each of the N input samples. y_{max} and y_{min} are the maximum and minimum value of all considered responses of the initial value. As they directly influence the value of $NRMSE$, special attention should be paid to the data in case of outliers. Such outliers can often indicate problems that require an adjustment

either of the numerical model or of the defined parameter space. Good metamodels will have an $NRMSE$ and R^2 close to 0 and 1, respectively. Which type of error measure should be used, again depends on the considered situation as shown in the evaluation of the following Fig. 3.1. In these three plots, the blue continuous line represents the function of a metamodel that fits both, the green and the grey dots, that are assumed to be results $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ of some numerical model. Those outputs are artificially generated to demonstrate the error measures. To generate the green and the grey dots, Eq. 3.14 and Eq. 3.15 are employed, respectively.

$$y_i = \hat{f}(x_i) + \omega \cdot \hat{f}(x_i) \ , \ i = 1, \dots 100 \tag{3.14}$$

$$y_i = \hat{f}(x_i) + \omega \cdot \bar{y} \ , \ i = 1, \dots, 100 \tag{3.15}$$

Here, the coefficient ω denotes a random value, generated according to a uniform distribution in the range between -1 and 1. The simple functions that represent the course of $\hat{f}$ are given in Tab. 3.1. Below in that table, the values of the $NRMSE$ and R^2 error functions are given that describe how good the model is fitting the random data. Usually, and especially in those publications referred to in this section, metamodel accuracy is considered to be good enough in case of $R^2 \geq 0.95$ and NRMSE ≤ 0.05. Accordingly, in case (a), for increasing and constant error, a metamodel with such fitting behaviour would be assumed to be good enough in case of R^2 error criterion but not in terms of $NRMSE$. In case (b), the result is somehow reversed. For constant and increasing error,

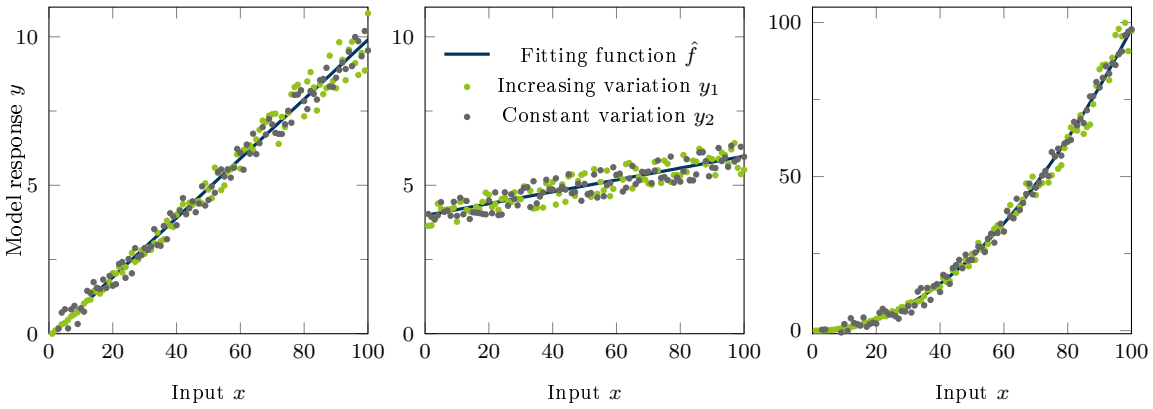


Figure 3.1.: Demonstration example of different artificial error applied to a (a) steep linear function, (b) shallow linear function, and (c) quadratic function.

Table 3.1.: Evaluation of error measures.

Error type	Plot	a $\hat{f} = 0.1 \cdot x$	b $\hat{f} = 4.5 + 0.01 \cdot x$	c $\hat{f} = x^2$
Constant				
	$NRMSE$	0.089171	0.156149	0.01892437
	R^2	0.990738	0.715809	0.99609393
Increase				
	$NRMSE$	0.10986	0.149886	0.02337788
	R^2	0.986761	0.738314	0.99413571

the $NRMSE$ indicates higher accuracy than for R^2, whereas it is not good enough in both cases. This might be unexpected as the same type of function and error is employed in both cases (a) and (b). The causal difference is the deviation σ of the results y_i that is five times larger in (a) than in (b) (2.930 to 0.586). In case (c), for both error types a high accuracy is indicated, whereby it should be mentioned that $\sigma(\boldsymbol{y})$ reaches a value of 30.279 in this case.

This simple example shows that a large coefficient of variation, that relates variation to mean value ($CV = \sigma(\boldsymbol{y})/\bar{y}$), indicates a better accuracy of the metamodel, even if the relative error is the same. As in case (b), this can lead to rejection of the metamodel. It is therefore recommended to always inspect by error criteria but also by plots of residuals or else whether the metamodel is providing accurate and reasonable results. Besides, one should always normalise both, the input and the output data, to have an equal impact of all types of values on the criteria described in Eqs. 3.12 & 3.13.

In literature, further approaches can be found to investigate the accuracy of the metamodel. For example, in Most and Will (2008) the coefficient of prognosis is determined that is a model-independent measure to assess the model quality. It is employed to create a so-called metamodel of optimal prognosis, that is an approximation to the model that is not considering unimportant factors. In turn, Atamturktur et al. (2015) suggest a so-called coverage metrics, where the coverage of the design space with respect to suitability for metamodel creation is considered and where additional samples in the parameter space should be placed to improve the metamodel's accuracy most.

3.3. Back analysis

3.3.1. Concept

The numerical models employed in this work allow to transfer input parameters, usually soil properties, to system responses like displacements or stresses. However, in many situations it might be more interesting to go the opposite way. The whole concept of the observational methods described in chapter 1 is based on performing back analyses. Observations or measurements $\tilde{\boldsymbol{y}}$ are compared to model outputs $\boldsymbol{y}$ obtained from initial assumptions on model parameters $\boldsymbol{\theta}$. To reduce the discrepancy between measurement and model response, the model parameters are varied. In the end this corresponds to the formulation of the objective function, described in Sec. 3.3.2. The complexity of this problem emerges from the circumstance that there are usually several model parameters that are correlated with each other and that the system behaviour is non-linear, explicitly in the considered geotechnical applications. Therefore, finding that optimal parameter combination $\hat{\boldsymbol{\theta}}$ that causes the minimum discrepancy between measurement and model response requires so-called optimisation algorithms as described in Sec. 3.3.3.

3.3.2. Objective function

Defining the distance between two points can be done in different ways. Several possible approaches to define discrepancy with application to geotechnical problems can be found in Meier (2008). The aforementioned discrepancy between measurements $\tilde{\boldsymbol{y}}$ and model responses $\boldsymbol{y}(\boldsymbol{\theta})$ is usually described by the objective function J as follows:

$$J_{\min}(\boldsymbol{\theta}) = ||\tilde{\boldsymbol{y}} - \boldsymbol{y}(\boldsymbol{\theta})|| \tag{3.16}$$

If several results are considered simultaneously, it might be favourable to employ a normalised objective function as shown in Eq. 3.17.

$$J_{\min} = \sum_{i=1}^{N} \left| \frac{\tilde{y}_i - y_i(\boldsymbol{\theta})}{\tilde{y}_i} \right| \tag{3.17}$$

In case some of the measurements are of less relevance for the model or less trustworthy, the objective function can be modified by including a vector of weighting factors $\boldsymbol{w}$ that reflects the relevance of the individual measurements $\tilde{y}_i$:

$$J_{\min} = \sum_{i=1}^{N} w_i \left| \frac{\tilde{y}_i - y_i(\boldsymbol{\theta})}{\tilde{y}_i} \right| \tag{3.18}$$

The criterion described in Eqs. 3.16 ff. uses the Euclidean distance that defines the minimum distance between two points by the square root of the sum of squares of its coordinates' distances as described by the theorem of Pythagoras in the two-dimensional case. It is the most commonly employed criteria in optimisation problems, as e.g. in (Miro, 2016; Zhao et al., 2015; Müthing et al., 2018). Another criterion is the so-called Haussdorf-distance d_{H}, described e.g. in Rote (1991) and successfully applied in Fidanova et al. (2013) to the problem of biological fed-batch cultivation process. Hereby, instead of comparing two individual outputs like in the Euclidean distance, all points of two data sets are compared in common as described by Eq. 3.19:

$$\boldsymbol{d}_{\mathrm{H}}(\boldsymbol{A},\boldsymbol{B}) = \max\left\{\max_{a\in\boldsymbol{A}}\left\{\min_{b\in\boldsymbol{B}}\left\{\boldsymbol{d}(a,b)\right\}\right\}, \max_{b\in\boldsymbol{B}}\left\{\min_{a\in\boldsymbol{A}}\left\{\boldsymbol{d}(a,b)\right\}\right\}\right\} \tag{3.19}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are two sets of data, e.g. measurement data and model responses. This approach makes sense in case these data sets are obtained as series, e.g. measurement at different time steps or positions.

To demonstrate how the objective function may influence simulation results using back-analysis, an example is studied in which the Haussdorf and Euclidean distance are applied to the same problem of time-dependent pore water pressure dissipation in an experimental device introduced in Nishimura et al. (2014). This example is considered more in detail in Sec. 4.2.1. The outcome is shown in Fig. 3.2. Measurement data from three pore water pressure transducers (in Points 1, 2, and 3) is displayed here over time by the blue, green, and grey dots. Using back analysis of this data, it is intended to fit a FE model by adapting the soil parameters. The back analysis is performed once using the Euclidean distance and once the Hausdorff metric. The identified soil parameters obtained with both criteria are run in the model and the model responses are plotted in Fig. 3.2 as straight (Euclidean distance) and dashed line (Hausdorff metric). At first sight, the results are quite similar, whereby the "Euclidean curve" has in general less discrepancy to the measurement data. Only close to the peaks of pore water pressure after approximately 60, 80, and 150 minutes, the "Hausdorff curve" is closer to the measurements (with exception of Point 2). These results are comprehensible, as the Hausdorff metric becomes relevant when elements of the two data sets $\boldsymbol{A}$ and $\boldsymbol{B}$ are close to each other (in an Euclidean meaning) even though they do not correspond to each other (e.g. correspond to different time steps as in the present example). However, the objective function reaches a slightly smaller value using the Euclidean distance as criterion (0.300) compared to the Hausdorff metric (0.306). Therefore, the Euclidean distance is employed in the work described in the following sections as it additionally demands less computational effort.

The special relevance of the topic of objective function for the present thesis comes from

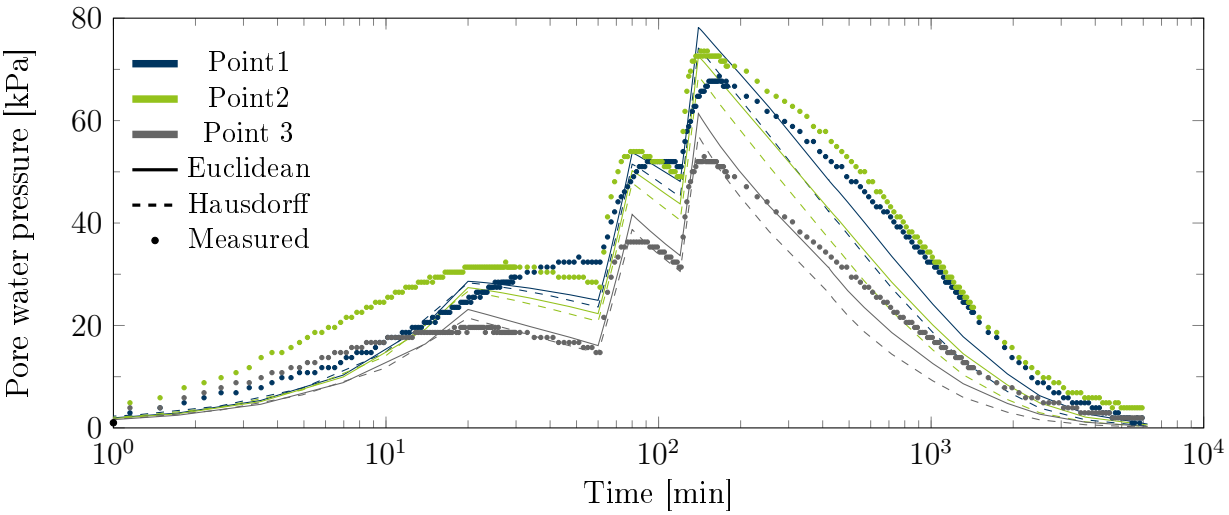


Figure 3.2.: Measurement data fitted to FE model results using Hausdorff metric and Euclidean distance.

the aspect that as obviously any model response, and measurement data as well, depend not only on the model parameters but also on the employed time and space coordinates. The objective function will be different depending on where and when the disagreement between model prediction and observation is compared. These additional features are denominated as experimental design parameters $\boldsymbol{\delta}$ in contrast to the model parameters $\boldsymbol{\theta}$ as they are user-defined, while the model parameters are supposed to enable a reflection of some natural situation. Accordingly, the objective function Eq. 3.16 is changed to Eq. 3.20

$$J_{\min}(\boldsymbol{\theta}, \boldsymbol{\delta}) = ||\tilde{\boldsymbol{y}}(\boldsymbol{\delta}) - \boldsymbol{y}(\boldsymbol{\theta}, \boldsymbol{\delta})|| \tag{3.20}$$

Hereby, the OED problem becomes a nested optimisation loop in which in the inner loop the model parameters are optimised, while in the outer loop the experimental design parameters are optimised with respect to the reliability of the identified parameters in the inner loop. To obtain a measure of the reliability of the outer loop's parameters, a Monte-Carlo analysis is performed that consists of numerous repetitions of the back-analysis and will be explained in detail in Sec. 3.4.
This method requires large computational cost, therefore, one essential aspect of the OED approaches presented herein, is to find time efficient solutions to the problem of OED. Accordingly, the first step is to select an efficient optimisation algorithm that finds the minimum of Eq. 3.16, or 3.20, respectively, and that is described in the following.

3.3.3. Optimisation algorithm

The objecitve function $J_{\min}$ introduced in the previous subsection describes the discrepancy between measured and calculated data. As in terms of model validation this discrepancy should be as small as possible, one aims to find that set of model parameters that provides the minimum of Eq. 3.20, such that the optimisation problem is formulated as described in Eqs. 3.21. The optimum set of parameters $\boldsymbol{\theta}^*$ are identified using an experimental design, here the optimal experimental design $\boldsymbol{\delta}^*$. However, as said before, finding this optimal design is also an optimisation problem. In that case, Eq. 3.22 is employed, where the optimal experimental design $\boldsymbol{\delta}^*$ is searched within the design space $\boldsymbol{\Pi}$, given the constitutive parameters $\boldsymbol{\theta}$. One should be aware that for different ranges of $\boldsymbol{\Theta}$, the obtained optimal experimental design $\boldsymbol{\delta}^*$ will also change.

$$J_{\min}(\boldsymbol{\theta}^*) = \min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} J(\boldsymbol{\theta}|\boldsymbol{\delta}^*) \tag{3.21}$$

$$J_{\min}(\boldsymbol{\delta}^*) = \min_{\boldsymbol{\delta}\in\boldsymbol{\Pi}} J(\boldsymbol{\delta}|\boldsymbol{\theta}^*) \tag{3.22}$$

As Eqs. 3.21 & 3.22 are nested, one fundamental problem of OED becomes clear: the optimal experimental design can only be identified, having the exact parameters. However, in case the parameters are known, there is no need for an experiment to identify them. Therefore, the approach should be rather understood as an iterative process, than as a unique matter.

Such objective functions as described in Sec. 3.3.2 are encountered in many scientific fields as e.g. economy (Lwin et al., 2017), aircraft engineering (Steiner et al., 2014), material sciences (Gomes et al., 2011), and wherever one specific solution is needed in a problem that has an infinite number of potential solutions. Optimisation problems can be distinguished between linear and non-linear optimisation. For linear optimisation, the Simplex-algorithm, as described in Dantzig (1987), can be employed where the vertices that define the space of possible solutions are investigated until the optimal solution is found. Another linear optimisation algorithm is the Karmarkar's algorithm (Karmarkar, 1984) that does not anymore follow vertices but can find solutions by intersecting the parametric space.

For non-linear models, the optimisation may lead to local and global solutions. A global solution means that the overal optimum of the objective function is encountered. A local solution indeed means that there is a zero point of the first derivative of the model's objective function, i.e. all solutions in vicinity of this local solution will provide larger

values of the objective function. In such cases, more complex optimisation algorithms are required whereby a distinction can be made between local and global optimisation methods.

The local optimisation can be used in case it is known that there is only one minimum in the objective function or if it can be ensured that no local solution is located between the starting point of the algorithm and the global solution. Hereby, the assigned boundaries of the optimisation area must be taken into consideration. It is possible that by employing a model, e.g. a constitutive soil model, the objective function has a global minimum according to its mathematical formulation, but that it is outside of its physical range of validity. In such case, one should be aware that the solution of interest is a local minimum within the limited search space. Besides, one should ensure that the optimisation does not converge to the boundaries of the search space that often consist of constrained local boundaries. Well-known examples of this group of local optimisation algorithms are among others the Newton method, where the quotient of first and second derivative of a function is used to iteratively find the minimum of a function, the Newton-Gauß algorithm that performs a local linearisation of the function in concern to stepwise reach a minimum, and the Nelder-Mead algorithm (Nelder and Mead, 1965) also known as Downhill-Simplex method. As its name indicates, it employs a simplex of $n + 1$ vertices that is iteratively regenerated based on that point of the simplex with the smallest response values. These local search algorithms have been further developed like the algorithm of total inversion introduced by Tarantola and Valette (1982) to conduct complex operations as the full waveform inversion.

However, these algorithms mentioned above are susceptible to converge possibly to the nearest minimum that might not be the global minimum of the search space, but a local one or to face an ill-posed problem if the estimation of the derivatives is not possible. Besides, these approaches require a continuously differentiable equation. In the cases considered in the present thesis, the so-called "model $f(\boldsymbol{\theta})$" is an FE-implementation that is not numerically differentiable. Its substitution, the metamodel $\hat{f}(\boldsymbol{\theta})$, could indeed be differentiable. However, it remains the problem of convergence to local minima.

Therefore, the group of nature-inspired algorithms that are mostly population-based algorithms is considered as they can be efficiently employed to find global minima. Numerous different algorithms of this category have been introduced in the last decades that intend to find optimal solutions by imitation of behaviours that can be observed in nature like Particle Swarm Optimization (Kennedy and Eberhart, 1995), Ant Colony Optimisation (Gambardella and Dorigo, 1996), Cuckoo Optimization (Rajabioun, 2011), or Spotted Hyena Optimizer (Dhiman and Kumar, 2017). Further algorithms might be inspired by

social behaviour or physical reactions like Simulated Annealing (Kirkpatrick et al., 1983) or League Championship Algorithm (Kashan, 2014). Overviews on the innumerable methods can be found in Behera et al. (2015) and Amaran et al. (2016).
In the present work, the Differential Evolutionary Algorithm (DE) (Storn and Price, 1997) and the Genetic Algorithm (GA) (Holland, 1975) are employed that both origin from the aforementioned group of population-based algorithms and are explained in detail in Secs. 3.3.3.1 and 3.3.3.2.

3.3.3.1. Differential Evolutionary Algorithm

The objective of using the DE algorithm, or any other one, is to find that optimal parameter combination that reduces most the discrepancy between model response and measurement data, i.e. for which the minimum of the objective function is obtained. To this, the DE algorithm aims to imitate the evolution of a population during several generations, whereby the different generations $1, 2, 3, ..., G, G+1, ...$ correspond to the iteration loops of the algorithms. The first generation consists of a "population" of NP vectors $\boldsymbol{\theta}_i$ that are sampled uniform randomly within the parameter space $\boldsymbol{\Theta}$. To generate a new generation of samples $\boldsymbol{v}_i$, the weighted difference of two vectors $\boldsymbol{\theta}_{i_2}$ and $\boldsymbol{\theta}_{i_3}$ is added to a third vector $\boldsymbol{\theta}_{i_1}$, what is referred to as "differential mutation". As shown in Eq. 3.23, this weighting is effectuated by the parameter F that is usually set between zero and two, or zero and one (Vincenzi and Savoia, 2015).

$$\boldsymbol{\theta}_{i,G+1} = \boldsymbol{\theta}_{r_1,G} + F \cdot (\boldsymbol{\theta}_{r_2,G} - \boldsymbol{\theta}_{r_3,G}) \ , \ F \in [0, 2] \tag{3.23}$$

where the indices $r_i, i = 1, 2, 3, ..., NP$ are employed to differentiate them from i. Each of the obtained new "population members" $\boldsymbol{\theta}_{i,G+1}$ are applied to the objective function and its results are compared to those corresponding to the previous generation $\boldsymbol{\theta}_{i,G}$ what is usually referred to as "crossover". This process of "evolutionary" generation of new samples is visualised in Fig. 3.3. In case the new population member yields a smaller value of the objective function than the old one, it is accepted, i.e. becomes member of the next generation: $f(\boldsymbol{\theta}_{i,G+1}) < f(\boldsymbol{\theta}_{i,G})$. This is repeated until the number of obtained differential mutations that have crossover probability CR between zero and one corresponds to the dimensions s of the parameter space. In case one population member $\boldsymbol{\theta}_i$ of the next generation $G+1$ is violating the bounds of $\boldsymbol{\Theta}$, it will be rejected. To perform the statistical operations described in Secs. 5 & 6, the coding language **R** is employed (Tippmann, 2015) that is implemented using the integrated development environment RStudio (RStudio Team, 2015). The employment of DE algorithm in **R** is enabled using

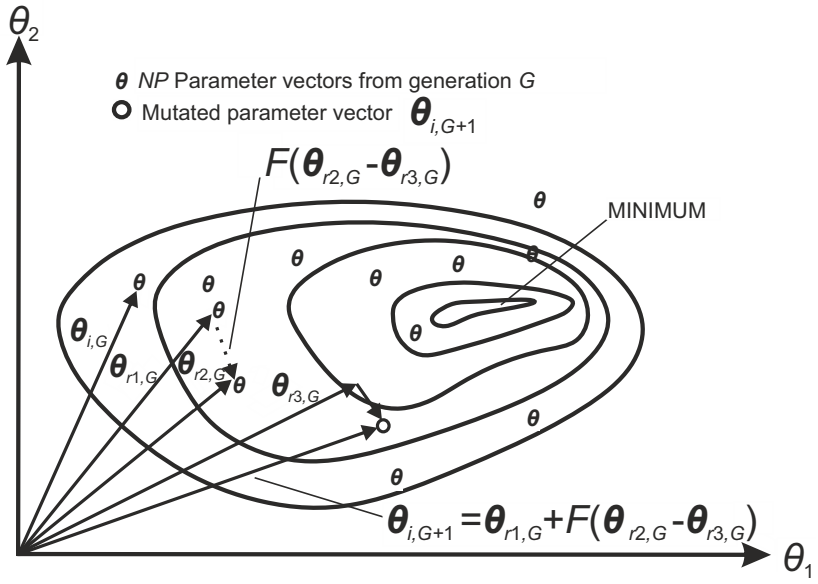


Figure 3.3.: An example of a two-dimensional objective function showing its contour lines and the process for generating $\boldsymbol{\theta}_{i,G+1}$, modified after Storn and Price (1997).

the package **DEoptim**. As the parameter identifications are not performed a single or few times, but numerous times to derive comparable statistical measures, the computational efficiency is of high relevance. Beside the efficient concept described afore, the coding was performed "using floating-point instead of bit-string encoding of population members, and arithmetic operations instead of logical operations in mutation, in contrast to classic GAs"(Ardia et al., 2011) to allow particular fast results. Within this **R**-package, the aforementioned variables of population size, mutation factor or maximum iteration steps can be selected individually according to the current problem, as will be demonstrated in following sections.

3.3.3.2. Genetic Algorithm

The genetic algorithm was introduced in Holland (1975). As mentioned above, it is from the same type of algorithms as the DE algorithms and its fundamental concepts are quite similar wherefore only a short description should be given here. Again, an initial set of random samples, called population, is generated. These parent generation samples are applied to the model of interest to obtain the corresponding responses. Using the concepts of selection, mutation, and crossover, a new generation of samples is generated that is supposed to outperform the parents. Selection means that those members of the parent generation that best fulfil the criteria of the objective function will automatically

be part of the next generation. For mutation, samples from the parent generation are randomly varied from their previous generation state, essential to avoid local minima. Within crossover sampling, two samples from the parent generation are taken and merged to a new sample like the segments of two chromosomes in genetics. Besides the population size, crossover and mutation chance can be selected by the user to influence the behaviour of the algorithm. The GA algorithm was applied for problems of geotechnical engineering in Khaledi et al. (2014), Khaledi et al. (2016), and Meier (2008). The comparisons in the latter one indicate that the GA algorithm shows high efficiency, even compared to DE or particle swarm optimisation but being more time-consuming. Within the framework of this thesis, the GA is therefore employed for specific optimisation when few accurate runs are required while the DE algorithm is utilised for numerous repetitions in the framework of statistical evaluations.

3.4. Statistical evaluation of geotechnical problems

Global safety or partial safety factor concepts as described e.g. in DIN EN 1997-1 (2014-03) follow deterministic approaches. Input parameters are employed as characteristic and discrete values and resulting effects and resistances are modified by partial safety factors (or the input parameters themselves in case of limit state GEO-3). However, using probabilistic approaches are an upcoming tendency in engineering design called reliability based design (RBD).
To a certain extent, the concept of OED requires the "probabilistic model thinking": as described previously, the objective of OED is to reduce uncertainty of model responses and predictions. Using deterministic models, single values are obtained that do not allow assessment how reliable such model responses are. The responses of probabilistic models are rather distributions that can be characterised by mean and variance. Here, using efficient model validation techniques including OED, the uncertainty that is described by the variance can be reduced, leading to more efficient experimental designs and predictions. To further illustrate these concepts, the RBD approach shall be briefly introduced in the following.

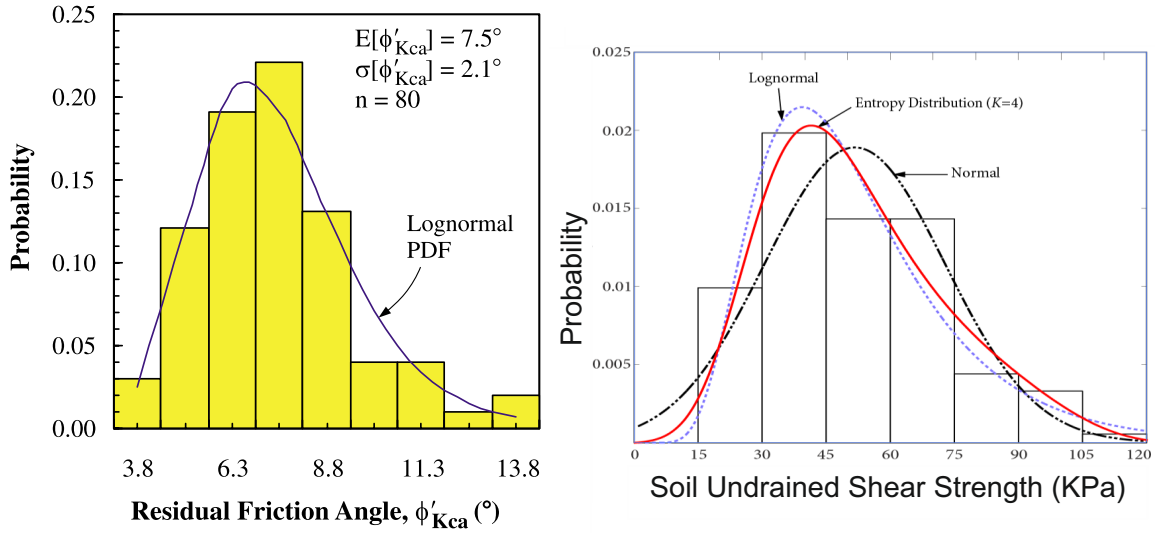
3.4.1. Types of distributions of input parameters

Up to now, parameter ranges have been mentioned several times in this thesis without precise details on their appearance. However, one should be aware that in a certain loc-

ation, soil parameters can vary within such a range, but it is very probable that values closer to the mean value occur more often than values close the limits of the range. To express how probable it is that a certain value might occur, probability distributions are employed. The question which type of distribution to employ is in a certain disagreement with the aspect of available data to fit the distribution. To be able to differentiate whether a set of data is $\mathcal{X}^2$, Weibull, or lognormally distributed, a lot of test data is required. As in most practical applications this is not the case, geotechnical engineers mostly refer to the normal, lognormal, and uniform distribution. This procedure has been shown practicable in several applications and for different soil properties. In Fig. 3.4 (a), this is shown exemplary for the friction angle of a clayey shale described in El-Ramly et al. (2003). Based on Fig. 3.4 (b), the problem of insufficient data is demonstrated. 121 vane shear tests were performed on a clayey-silt layer in the same area, what can be regarded as very high number. Using the Kolmogorov-Smirnov test, Kanwar and Deng (2019) are able to prove the entropy distribution to have the best fit, but few tests more or less may have provided another result as shown by the two other curves (normal and lognormal distribution) plotted in that diagram.
In general, the employed type of parameter distribution function should reflect the actual knowledge on the parameters. Therefore, in case only the soil type is known and no information about the actual parameter values are available, employing the uniform distribution is recommended as performed among others in this thesis. In case of more information normal or lognormal PDFs should be applied, assigning the expectancy as mean value and variance according to the actual state of knowledge. The lognormal distribution has the advantage that values beyond a certain threshold value can be avoided. This is of relevance, as most soil parameters for example cannot become negative. Using Bayesian inference, as in Miro et al. (2015) and in chapter 6 of this thesis, the variance term is iteratively reduced according to the current state of knowledge. More details on this topic can be found in Uzielli et al. (2006), where a detailed overview is provided on the possible ranges and distribution types that can be expected depending on the considered soil type.

3.4.2. Monte-Carlo sampling method

A direct transfer from the inputs PDFs to output PDFs would require to calculate integrals what is not possible in case of the FE models that are employed in the present thesis. To overcome this issue, the so-called Monte-Carlo simulation can be employed to approximate the output distribution. It was introduced previously in Sec. 3.2.2 in the



(a) (b)

Figure 3.4.: a: Probability histogram of the residual friction angle of the Kca clay–shale (El-Ramly et al., 2003), b: Histogram for soil undrained shear strength and normal, lognormal, and entropy distributions (Kanwar and Deng, 2019).

context of generating data to set up a metamodel. In the context of probabilistic simulations, the MC-simulation will behave quite similar. Instead of generating the samples uniformly random within the whole parameter space, they are still generated randomly, but according to the distribution of each of the input parameters. That means the uniform distribution is replaced by e.g. a normal or log normal distribution. In the next step, all samples are applied to the employed model to obtain the outputs of interest. These outputs converge to a new distribution whereby it is not necessarily of the same type as the input distribution. The more samples are generated, the more accurate the output distribution will be, but the computational cost rises in equivalent. From this distribution, statistics can be derived that allow to evaluate the system reliability. In case the probability of specific failure event is of interest, this might become quite extensive as in civil engineering applications required failure probabilities are often very small.

3.4.3. Use in modern engineering design

Based on the scientific concepts mentioned above, the probabilistic concepts are transferred to engineering practice e.g. in DIN EN 1990 (2010-12) where the reliability index β is found nowadays as design criterion. In Low and Phoon (2015) and Low (2017), application examples are given where RBD is applied to typical examples of geotechnical

engineering. The reliability index β can be calculated using Eq. 3.24 according to Hasofer and Lind (1974).

$$\beta = \min_{\boldsymbol{\theta} \in \boldsymbol{F}} \sqrt{(\boldsymbol{\theta} - \boldsymbol{\mu})^T \boldsymbol{C}_{\boldsymbol{\theta}}^{-1} (\boldsymbol{\theta} - \boldsymbol{\mu})} \tag{3.24}$$

Here again, $\boldsymbol{\theta}$ are the uncertain parameters, $\boldsymbol{\mu}$ is the expected mean, and $\boldsymbol{C_\theta}$ is the covariance matrix of these parameters. $\boldsymbol{F}$ is the failure domain of the parameter space in which a certain threshold, e.g. a maximum settlement or a limit load, is exceeded as described by function g in Eq. 3.25.

$$\boldsymbol{F} = g(\boldsymbol{\theta}) < 0 \tag{3.25}$$

This index β is related to failure probability according to Eq. 3.26:

$$P_\mathrm{f} \approx 1 - \Phi_\mathrm{F}(\beta) = \Phi_\mathrm{F}(-\beta) \tag{3.26}$$

Here, Φ_F is the cumulative distribution function (CDF) of the uncertain model input parameters $\boldsymbol{\theta}$. That combination of $\boldsymbol{\theta}$ that allows to minimise the square root in Eq. 3.24 for the prescribed value of β should accordingly be employed as design parameter set as it allows the most reliable design for the required level of safety. This concept of RBD is illustrated in Fig. 3.5. Within the exemplary parameter space of friction angle φ and cohesion c, as it would be relevant e.g. for a problem of slope stability, the limit state surface separates safe and unsafe (failure) areas. The parameter means μ_c and μ_φ are

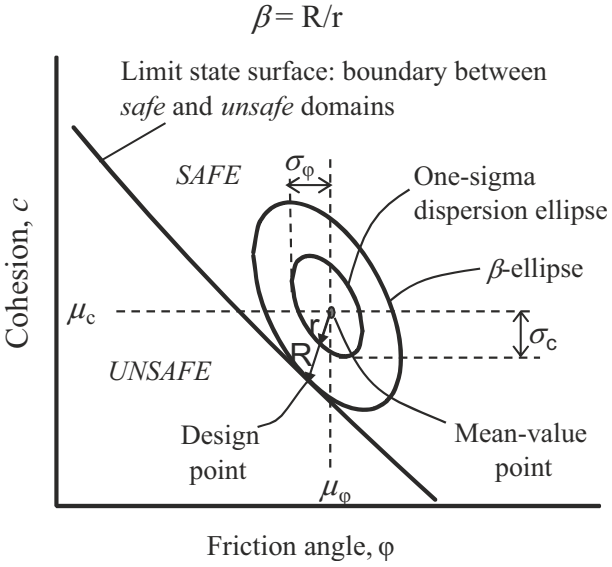


Figure 3.5.: Illustration of the reliability index β in the plane when c and φ are negatively correlated, modified after Low and Phoon (2015).

obviously on the safe side and the parameter combination for which the design should be performed is graphically obtained by expanding the ellipse of standard deviations until it is tangent to the limit state surface. The thereby obtained ratio of the radii of outer and inner ellipse corresponds to the reliability index β.
As mentioned above, identifying properly the limit state surface or P_f is often not directly possible and is usually approximated by using MC simulations. Especially in case of small failure probabilities, the crude MC can become inefficient as it covers evenly the whole parameter space with samples, while only a small area is of interest. In such cases, applying the so-called subset simulation as described in Mahmoudi et al. (2017) and Mahmoudi et al. (2020) can be employed that gradually retraces the parameter space to the relevant areas.

3.4.4. Application example

The employment of probabilistic design and prediction methods is not part of the main objectives of this thesis. However, it is employed in the context of Bayesian OED that is described in chapter 6. Besides, the concepts of OED should be considered in the overall context of system and parameter identification with uncertain soil parameters and their impact on model validation. This was studied in detail in Müthing et al. (2018) by investigating and predicting the settlement and excess pore water pressure of an embankment placed on soft clay at several time steps after construction whereby the final settlement was not yet reached. The accordingly generated 2D FE model is shown in Fig. 3.6. The interpretation of constant rate of strain (CRS) oedometer tests enabled to set up an initial assessment of the most adequate constitutive model that is the soft soil creep model and its required parameters for the four soft layers. Under consideration of appropriate parameter boundaries and correlations, input parameter distributions are defined using normal distributions with exception of friction angle and permeability that are assumed to be log normal. 10000 MC samples of these input parameters are applied to a metamodel of the time-consuming 2D FE simulation. In Fig. 3.7, the output distributions of settlement in the centreline and pore water pressure 2 m below ground surface of the embankment (point 0 and point VWP1a, respectively in Fig. 3.6) are shown by the green bars. In Fig. 3.7b, a normal distribution can be observed, while in Fig. 3.7a, a more log normal distribution appears.
In a second step, in-situ measurement data is employed to perform a back analysis of the relevant soil parameters that are identified by GSA to be the overconsolidation ratio OCR, the modified compression index λ^*, the modified swelling index κ^*, the modified

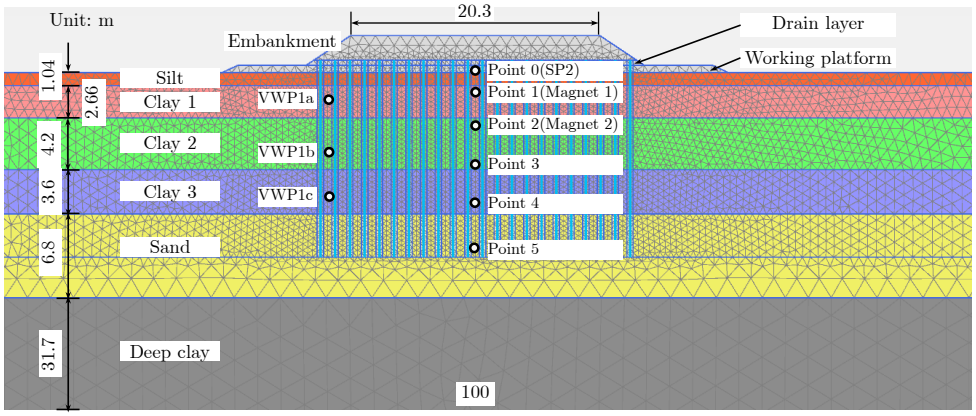


Figure 3.6.: Geometry and FE discretisation of the 2D model of the considered embankment and its subsoil (Müthing et al., 2018).

creep index μ^*, and the permeability coefficient k. Using the identified parameters, the MC simulation is repeated, leading to the blue bars in Fig. 3.7. Within this specific application, the *COV* of the parameters is not changed after back calculation. Thereby, it is intended to consider that the actual uncertainty of the subsoil regarding inhomogeneity was not reduced. Accordingly, it is comprehensible that the output bandwidths do not reduce. The actually measured values for the two given locations are indicated by the blue dashed line. In case of the settlements, initial and final prediction are close to each other and within the margin of standard deviation of the first prediction, but still an improvement can be observed.

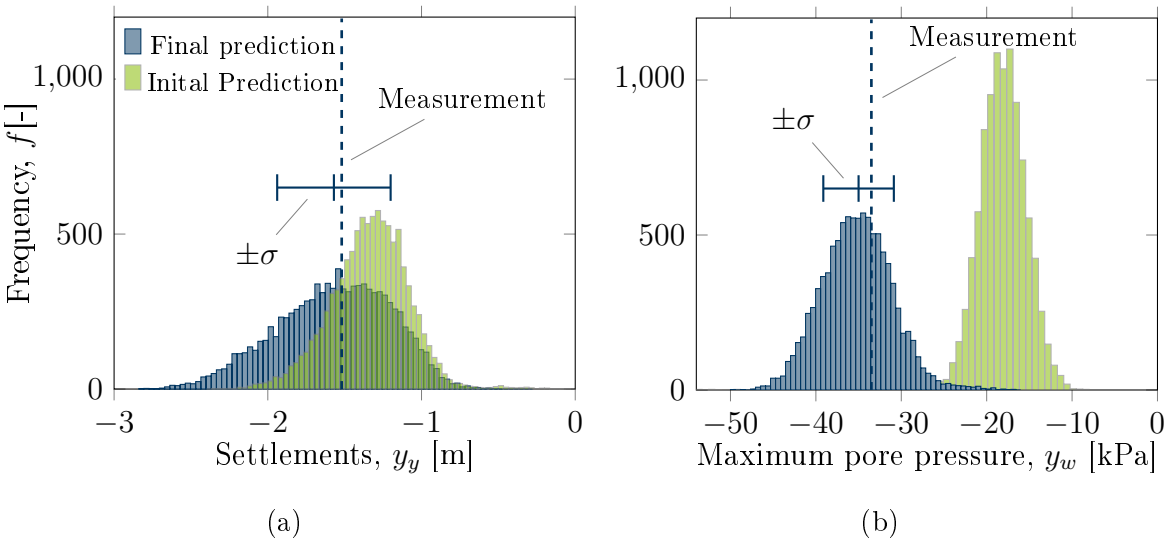


Figure 3.7.: Histograms of probability distribution of a: settlements and b: pore water pressure.

Table 3.2.: Values of identified parameters from initial and back calculated parameter sets.

Soil layer	OCR [-]		λ^* [-]		κ^* [-]		μ^*[-]		k[m/day]	
	+	*	+	*	+	*	+	*	+	*
Silt	2.86	3.99	0.158	0.115	0.053	0.048	6.2 E-3	5.0 E-3	2 E-3	9 E-4
Clay 1	2.71	2.09	0.177	0.170	0.010	0.064	4.3 E-3	4.4 E-3	2 E-4	7 E-4
Clay 2	1.73	1.60	0.277	0.320	0.010	0.048	4.3 E-3	4.0 E-3	6 E-4	1 E-4
Clay 3	1.81	1.58	0.239	0.279	0.012	0.023	4.9 E-3	6.3 E-3	6 E-4	6 E-5

$^+$ parameter values from initial parameter set (blind prediction)

* parameter values from optimised parameter set (back calculated)

In Tab. 3.2, the initially determined parameter values and those identified by back calculation are listed for the four controlling layers. In case of the pore water pressure, the final prediction shifted considerably. In both cases, the mean values of the distributions are now close to the measured data. However, they do not perfectly match and therefore, this probabilistic prediction is a helpful tool to indicate how realistic the prediction is in comparison with in-situ data. Besides, it shows how the prediction accuracy improves with a larger database.

A specific failure case is not defined in this application as the only objective is to predict the long-term settlement behaviour. However, such consideration would be simple to add, once the data shown in Fig. 3.7 is at hand. More details on this application can be found in Hölter et al. (2016) and Müthing et al. (2018), especially with respect to the initial test interpretation, the FE model set up and the long-term investigation. Further detailed applications of RBD to problems of geotechnical engineering can be found in Mahmoudi et al. (2017) or Miro et al. (2015). There, the convergence behaviour of a rock cavern and the settlement induced by a tunnel excavation are considered as synthetic examples, respectively.

3.5. Bayesian analysis

3.5.1. Concept

The previously described methods are employed to account for the uncertainty of the model parameters $\boldsymbol{\theta}$. However, the described process of back analysis is still following a deterministic concept, that means that if a value is back calculated as described in Sec. 3.3.2, the probability of the available measurement $\tilde{\boldsymbol{y}}$ and of the back calculated parameter set $\boldsymbol{\theta}$ is not considered, i.e. a uniform distribution is assumed. Assuming both values as probability density functions (PDF) would allow to evaluate how probable a certain measurement $\tilde{\boldsymbol{y}}$ is with respect to the current knowledge on the input data. As the measurement data used for back calculation is at the same time an extension of the current data knowledge, one might be interested to "update" the previous knowledge. This concept was first formulated in Bayes and Price (1763) for probability values while it is still valid for random variables as described by Eq. 3.27.

$$p(\theta|y) = \frac{p(y|\theta) \cdot p(\theta)}{p(y)} \tag{3.27}$$

Herein, $p(y)$ describes the PDF of model output that is used to normalise the result. In the framework of this thesis, this would typically be noisy measurement data $\tilde{\boldsymbol{y}}$ (see Sec. 2.3.3.2) that is somehow distributed around its mean value $\bar{y}$. $p(\theta)$ accordingly is the PDF of the model parameters. These would usually be soil parameters that are often very uncertain in geotechnical engineering and that are usually approximated by normal or lognormal distributions as described in Sec. 2.4.1. Within the Bayesian concept, this distribution $p(\theta)$ is called as prior probability function or "prior knowledge", as it describes the state of information before it has been updated using Eq. 3.27. Accordingly, $p(\theta|y)$ is called the posterior probability function or posterior knowledge as it includes the new information y. The link between prior and posterior distribution is the conditional probability $p(y|\theta)$ or likelihood that describes how probable the new data y is, given the known probability distribution of θ.

Following this procedure of Bayesian updating, a parameter identification becomes initially much more complex than as described in the previous sections, but this probabilistic manner allows to adapt the parameter space. Hereby, a faster identification is enabled by the retracing or it allows to expand or shift it in case the model conditions have basically changed. For a more detailed introduction to the topic, one is referred to Miro (2016) and the references therein, where the concept of Bayesian statistics is employed considering an application of mechanised tunnel simulation.

3.5.2. Review on applications of Bayesian inference

For several decades, applications of Bayesian updating have been applied in various fields as e.g. in economics as shown in the publication of Kooiman et al. (1985) where the productivity of floor spaces of shops is investigated. Herein, the human behaviour as large source of uncertain behaviour is modelled and this uncertainty shall be reduced by using observations of purchasing value of sales, occupancy costs per square meter, and the relative share of sales of fresh products. As changing prices cause shifting of supply and demand of products, the model state is continuously changing in this example. By using MC simulation to approximate the prior and posterior PDFs, the authors achieve to reduce the uncertainty of their prediction of supply and demand to 20% of the initial value. An example from structural mechanics is given in Alkam and Lahmer (2019). Here, the vibration behaviour of train track poles exposed to cyclic loading in a 3-point bending test and dynamic loading in a vibration test is investigated in common. In-situ observations are considered to validate a model that has uncertain parameters. As different test set ups are used, different FE models are generated that include the same six uncertain constitutive parameters to be identified. Each observation is employed for one Bayesian inference step after which the prior knowledge is updated. By comparing the results with a deterministic approach, it is shown that in particular the variances of the parameters can be identified more accurately using the Bayesian approach, but at the cost of larger computational efforts.
Miro (2016) and Nguyen (2017) both employed Bayesian inference to update 3D simulation models of tunnel excavations. The uncertainty of the prior estimation of the soil parameters is therein reduced stepwise using measurement data that is obtained after each new excavation step of the tunnel, but each time affected with artificial random noise. Miro (2016) is using ground surface responses (horizontal and vertical displacements) induced by the tunnel excavation while in Nguyen (2017) seismic wave propagation is considered. In both cases indeed, the measurement positions are fixed and not considered as parameters that could be optimised. Within this thesis, accordingly the objective is to identify an experimental design by which the prior probability distributions of soil parameters can be narrowed most efficiently.
These examples show how Bayesian inference is widely spread over numerous fields of research. Within the present thesis, its ability to adapt the process of parameter identification in case of changing model conditions is utilised, whereby not the model itself changes, but the experimental design to obtain data from it.

3.6. Summary

The methods introduced in this chapter are employed in the further parts of this thesis without detailed explanation but rather reference to this chapter will be given. The discussed topics of metamodelling, back analysis, statistical analysis and Bayesian inference are not originating in the field of geotechnics but were applied in various other fields before as shown by the cited literature. However, due to the large uncertainties encountered in geotechnical engineering, these methods are considered here in a special focus. Working with uncertain data instead of deterministic values is undeniably a trend that will gain relevance in future engineering practice as it already has in research. For this reason, the given overview provides the relevant "tools" to understand and handle the methods encountered in this thesis. They are described more in detail and more fundamental in the cited references, while they are here described with special focus on the intended applications.

4. Global sensitivity analysis in the framework of OED

4.1. Method

4.1.1. Fundamental approaches in sensitivity analysis

The description given in Sec. 2.3.2 on model validation clarifies that it is not sufficient to prove that a model generates correct results $\boldsymbol{y}$ for a certain situation, represented by a specific set of input parameters $\boldsymbol{\theta}$. A reliable model validation implies to obtain a holistic understanding of the model behaviour. In case of mathematical models, either analytical or numerical, this means essentially understanding how the different input parameters θ_i influence the model responses y_i and which correlation effects in between the model parameters might occur. This type of investigation is nowadays known as sensitivity analysis and got rising attention in the last decades. In case of linear analytical models, this task corresponds to the partial derivatives of the model variables, but for more complex models, especially in case of FE-models like those considered in this thesis, more advanced techniques are necessary that are explained in the following.

4.1.1.1. Local sensitivity analysis

The main classification of sensitivity analysis methods is the differentiation into local and global methods. Hereby, the terms of local and global do not refer to geometric coordinates, but to the investigated area in the parameter space. While the global sensitivity analysis (GSA) is supposed to investigate the whole defined parameter space and any correlation effects of combinations therein, the local sensitivity analysis (LSA) only investigates variations of individual parameters in an area close to a defined point in the parameter space. As aforementioned, in case the model of interest is represented by a function that is linear with respect to the employed variables, the sensitivity of a parameter corresponds to the

derivative of the model equation. However, if the model is not linear, a certain knowledge on acceptable ranges of parameter values is necessary. In case of FE models that consider initial boundary value problems and do not have a functional relationship between input and output parameters, not an analytical but a numerical solution is obtained, and this procedure of determining derivatives is not possible. Therefore, the approach described by Eq. 4.1 is employed to calculated the so-called scaled sensitivity SS_{ij} where i refers to the different parameters and j to the output of interest as introduced by Hill (1998) and employed in Zhang et al. (2003):

$$SS_{ij} = \left(\frac{\Delta y_j}{\Delta \theta_i}\right) \theta_i = \left(\frac{y_j\left(\theta_i + \Delta\theta_i\right) - y_j\left(\theta_i\right)}{\Delta\theta_i}\right) \theta_i \tag{4.1}$$

$\Delta\theta_i$ describes the increment of the i^{th} input parameter that is deviated from its so-called base point. Finding the optimal increment size is a complex procedure which strongly influences the reliability and significance of the obtained values of SS_{ij}. If the increment is too large, non-linearities may falsify the result, but if it is too small, response variation might be smaller than the inevitable numerical error and provide also misleading results. In Zarev (2014), the procedure of finding optimal values for $\Delta\theta_i$ is described in detail, taking into consideration truncation and condition error. As this procedure is time-consuming and the LSA is sometimes employed as "preliminary sensitivity analysis" as in Zhao et al. (2014), empirical values such as 5 % or 10 % of θ_i are often employed as $\Delta\theta_i$. In Zhao et al. (2014) it is intended to identify those soil parameters that influence most the settlement behaviour induced by a mechanised tunnel construction. As several soil layers are involved, a total number of more than 100 parameters would be required to be investigated. In this publication, LSA is applied to those numerous parameters as a kind of filter to find out which of them could be excluded from the model calibration procedure in the particular case study. On the ten parameters identified to be relevant, the investigation is focused in Zhao et al. (2015) performing a parameter identification using field observations.
To combine the information of model outputs from different locations or time steps of the simulation history, counted by the index j in Eq. 4.1, the composite scaled sensitivity CSS can be employed:

$$CSS_i = \sqrt{\frac{1}{N}\sum_{j=0}^{N}\left(SS_{ij}\right)^2} \tag{4.2}$$

This formulation can be found in Hill and Tiedeman (2005) and is employed in the publications mentioned above to obtain a more reliable sensitivity measure that is based on more information than just that of one single observation point. However, as in the

framework of OED it is explicitly intended to differentiate between individual points to decide which of them will provide more reliable information, this measure seems not to be suitable for the purpose of interest. All in all, the concept of LSA persuades by its simplicity and efficiency and is therefore often applied and considered adequate for types of models in which the aforementioned limitations do not restrain its application. Even in case of non-linear models, an LSA may be appropriate if the considered parameter space is that narrow that the model behaviour can be assumed to be linear. However, these aspects should be well-considered before employment.

4.1.1.2. Global sensitivity analysis

Several different approaches of GSA have been developed in the last decades that all have their varying advantages, such as the Elementary Effect method (Morris, 1991), the random balance design method (Cukier et al., 1973; Schaibly and Shuler, 1973) or the Variance Based (VB) method (Saltelli et al., 2008). Their general concept is to "study how the uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model input factors", Saltelli et al. (2010). The term of "uncertainty" plays a major role in this formulation. It indicates that the sensitivity of a parameter will not only depend on the model itself but also on the assumed bandwidth and distribution type of the considered parameter. As the GSA considers parameter samples from the whole defined parameter space, expanding this space might cause very different model responses, especially when the model is non linear. Therefore, the selection of parameter bandwidth must be well-considered and accurate. When considering e.g. the friction angle of sand as an input parameter, assuming a range from 30 to 45° should be reasonable if there is no additional information. The typical reaction of an engineer might be: "Let us increase the range to 25 to 50° to be on the safe side". In case of GSA this is not recommended as such a large bandwidth would not describe the same material any more and the legitimate question arises whether the employed model is still valid. Additionally, the relevance of other parameters would be decreased, leading to a misinterpretation of the model behaviour.
A detailed overview on the different methods in context of geotechnical engineering, with recommendations which one to use in case of specific applications, mostly depending on the number of parameters and complexity of the model, can be found in Mahmoudi (2017) and Mahmoudi et al. (2019). For the requirements that are encountered in the present thesis, it makes sense to employ one single method to not make the procedure more complicated. Therefore, the VB method is employed that is regarded as most robust

considering different possible circumstances like parameter number or model complexity. The drawback of growing computational costs by using VB-GSA is overcome by the use of adequate metamodels (see Sec. 3.2) that enable to considerably reduce the time of individual model runs as described in Sec. 3.2. Accordingly, this method is described more in detail in the following.

The VB method was inspired by Cukier et al. (1973) and found its first formulation in Sobol' (1993), wherefore this method is also known as Sobol' method, and the obtained sensitivity measures as Sobol' indices. In the following decades, several publications summarised in Saltelli et al. (2008) improved and extended that approach to the state that is considered in the present work. Further general works on this topic can be found e.g. in Ferretti et al. (2016) and Lo Piano et al. (2021).

The aforementioned sensitivity indices refer to the different orders of effects among the parameters. The variance V_j of a certain model response y_j due to a parameter θ_i, called the "effect", can be described as:

$$V_j = V_{\boldsymbol{\theta}_i}\left(E_{\boldsymbol{\theta}_{\sim i}}\left(\boldsymbol{y}|\theta_i\right)\right) \ , \tag{4.3}$$

where E describes the expectation and $\boldsymbol{\theta}_{\sim i}$ is the vector of all input parameters with exception of θ_i. Accordingly, $E_{\boldsymbol{\theta}_{\sim i}}\left(\boldsymbol{y}|\theta_i\right)$ corresponds to the expected value of $\boldsymbol{y}$ that is obtained when all parameters except θ_i are varied. If now this individual variance V_i is normalised by the overall variance $V(\boldsymbol{y})$ of the outputs of $\boldsymbol{y}(\boldsymbol{\theta})$, the so-called first order index S_i is obtained:

$$S_i = \frac{V_{\theta_i}\left(E_{\boldsymbol{\theta}_{\sim i}}\left(\boldsymbol{y}|\theta_i\right)\right)}{V(\boldsymbol{y})} \tag{4.4}$$

Because the indices S_i are normalised by $V(\boldsymbol{y})$, that includes all contributions to the model response variance, the sum of all S_i equals one.

However, as the sensitivity index in Eq. 4.7 refers to the impact of each parameter variation $V_{\boldsymbol{\theta}_i}$ individually, variations that occur due to correlated effects of two parameters will not be captured. Therefore, it is often recommended to consider the second or higher order effect terms expressed by Eq. 4.5:

$$V_{ij} = V_{\theta_i\theta_j}\left(E_{\boldsymbol{\theta}_{\sim ij}}\left(\boldsymbol{y}|\theta_i,\theta_j\right)\right) - V_{\theta_i}\left(E_{\boldsymbol{\theta}_{\sim i}}\left(\boldsymbol{y}|\theta_i\right)\right) - V_{\theta_j}\left(E_{\boldsymbol{\theta}_{\sim j}}\left(\boldsymbol{y}|\theta_j\right)\right), \tag{4.5}$$

such that the sum of variance terms equals to:

$$V(\boldsymbol{y}) = \sum_i V_i + \sum_i \sum_{i>j} V_{ij} + \cdots + V_{12\ldots d}, \tag{4.6}$$

where d corresponds to the length of the parameter vector $\boldsymbol{\theta}$. Accordingly, the total sensitivity index S_{T_i} is obtained:

$$S_{Ti} = \frac{E_{\boldsymbol{\theta}_{\sim i}}\left(V_{\theta_i}\left(\boldsymbol{y}|\boldsymbol{\theta}_{\sim i}\right)\right)}{V(\boldsymbol{y})} = 1 - \frac{V_{\theta_{\sim i}}\left(E_{\theta_i}\left(\boldsymbol{y}|\boldsymbol{\theta}_{\sim i}\right)\right)}{V(\boldsymbol{y})} \tag{4.7}$$

As the higher order sensitivity measures are included in the S_{Ti}-indices, their sum can become larger than one, indicating stronger correlation among the parameters the higher $\sum_i S_{Ti}$ becomes. In case the parameters are not correlated, i.e. independent, S_{Ti} becomes equal to S_i. As in the present work constitutive models are employed that actually include correlation effects to adequately reproduce the natural soil behaviour, the total effect sensitivity index S_{Ti} is mostly employed.

The VB method allows an accurate estimation of all parameter contributions on the model responses, but has the major drawback of causing higher computational costs due to calculating the different variance terms. Decisive for an employment of the VB sensitivity method is an efficient method of sampling and using the model responses to estimate the sensitivity indices. In the present thesis, the approach introduced in Jansen (1999) that is described in Miro et al. (2014) for the employment in geotechnical applications is applied, using Latin Hypercube Sampling (LHS) to generate the required parameter samples. In Dimov and Georgieva (2010) a comparison is provided on different methods to calculate the herein employed Sobol' indices. However, as it is shown in Saltelli et al. (2010) and specifically for geotechnical applications in Mahmoudi et al. (2019), the required number of samples easily reaches several thousands. In case calculations of little complexity are executed, such as element tests, also in geotechnical engineering, GSA can be applied directly to a numerical model as shown in Barciaga (2022). However, in case of full-scale FE-simulations that often require several hours of calculation time, this would be infeasible. In particular, it can not be known in advance how many samples will be necessary to reach convergence in the estimation of the variance terms. Therefore, the VB-GSA is not applied to the results of a numerical model, but to a metamodel as described in Sec. 3.2. Doing so, many thousands of parameter samples can be considered to calculate the different variance terms, probably more than necessary, but ensuring reliable results. This procedure is introduced in Miro et al. (2014) and further employed in Miro (2016) and Zhao et al. (2015) for the purpose of simulation of mechanised tunnelling, but was also successfully applied to further geotechnical problems as embankments or rock caverns, respectively (Müthing et al., 2018; Mahmoudi et al., 2017).

4.1.1.3. GSA under the point of view of OED

The limitations of LSA mentioned in Sec. 4.1.1.1 regarding parameter space exploration and correlation effects are still relevant when considering the purposes of OED. In this regard, it is to be noted that the FIM includes, besides the covariance matrix of output variation, the matrix of derivatives of the different model responses with respect to all considered model parameters. Therefore, the FIM as employed e.g. in (Uciński, 2005; Patan, 2012) is very closely related to calculating the LSA index SS_{ij}. The numerous former applications that employed LSA should not be disparaged and are definitely valid in case of linear models. Besides, the extensive computational efforts required when performing GSA might have prevented its employment for OED.
The application examples considered in this thesis consist of FE models where a differentiation of the solution is unkown. Besides, the employed constitutive models describe non linear behaviour with correlation effects among their parameters. Therefore, the employment of GSA instead of LSA is considered to be preferred for the model validation process in general and explicitly for OED.
With respect to the objectives of this thesis formulated in Sec. 1.2, one can reconsider which benefits the GSA can contribute to the ideas of OED. The profit of sensitivity analyses with respect to the topic of OED is initially to know which parameters are most relevant at all. The experimental design should focus on these parameters as identifying a parameter that has no influence would be thriftless with respect to the usually limited resources. In certain publications as e.g. in Rodriguez-Fernandez et al. (2007) and Schenkendorf et al. (2018) from the field of chemical engineering, sensitivity measures are employed to perform an entire process of OED.
The conceptual link between GSA and OED is that those parameters to which the model response is most sensitive, consequently cause a larger bandwidth of results if they are varied. The larger the output variation is, the easier the influence of the individual parameters can be identified and discriminated from uncertainty related errors, an idea that is formulated similar previously in Kay (1993). Hence, it is possible to identify the most sensitive parameters most accurately by means of a back analysis as described before in Sec. 3.3. To know which parameters are most relevant for the considered system allows to focus the investigations on these parameters and to neglect the less relevant ones by fixing them to roughly estimated values.
The approach presented in this chapter exclusively employs sensitivity measures for performing OED, using two different steps of GSA described in the following. In chapters 5 and 6, further methods for OED are applied. However, GSA is still performed as first

essential step to identify the most relevant model parameters on which the specific OED approach is applied afterwards.

4.1.2. Concept of modified sensitivity index $S^*_{T_i}$

First ideas to the concept of sensitivity analysis in a spatially distributed manner for geotechnical purposes can be found in Schanz and Meier (2008) where indeed the composite scaled sensitivity index CSS (Eq. 4.2) is employed with the drawbacks described in Sec. 4.1.1.1. In that work, the interpretation of results was performed more conceptually and qualitatively compared to the systematic manner of description in the following sections. Using the variance-based GSA that is described in Sec 4.1.1.2 allows to identify in a more reliable manner how much the different constitutive parameters influence a certain model response. In the context of this thesis, the first order sensitivity index S_{T_i} is employed that also considers the impact of correlation effects which play larger roles in constitutive models of soils. As described in the previous section, this implies also that those parameters with high sensitivity can be identified most accurately. Knowing on which parameters the experimental design should focus at all is already an essential part of the OED.

However, one should consider that a FE model response depends not only on the values of the employed constitutive parameters $\boldsymbol{\theta}$. Indeed, also the time step and position at which responses are obtained, summarised in the experimental design parameter vector $\boldsymbol{\delta}$, influence these responses. Besides, one should consider that one single model can provide different types of model responses. This seems to be very obvious, but under the point of view of OED it is of relevance to mention. One could consider the common case of a retaining wall for which initial model assumption should be validated to ensure small wall deflection at the end of construction. Which model data $\boldsymbol{y}$ should be used for model validation, i.e. which field measurements $\tilde{\boldsymbol{y}}$ should be gathered? Displacements at the top of the wall, or better horizontal earth pressure in greater depth? No generally valid answer is possible as measurement uncertainty, and hence the potential model uncertainty reduction, of both will be different. However, to find the best measurement arrangement, these circumstances should be considered and be incorporated in the decision making.

The next point that is crucial in context of validation of FE models in civil and especially geotechnical engineering applications is the position of a measurement to collect the required information. Modern FE applications allow to provide model responses for any positions of the FE model geometry. For each of these positions, a GSA can be performed and each time it will show different results, depending on properties of the current specific

position in the context of the model. Like in the consideration regarding time or model response, here again the approach is to use that position with highest sensitivity to a certain parameter to place a sensor to identify this specific parameter. However, one should be aware that the sensitivity index S_{T_i}, that is obtained when using the variance-based method, provides normalised values. This means that positions in which almost no model response is obtained, like on the outer boundaries of the model, can still exhibit a high sensitivity to a certain input parameter. This problem can be solved by including the relevance of an output by means of its variance. To consider in combination the different aforementioned aspects, the modified sensitivity index $S^*_{T_i}$ is proposed:

$$S^*_{T_{i,j,k}} = \frac{S_i \cdot \sigma_{j,k}}{\max_k \sigma_{j,k}} \tag{4.8}$$

In Eq. 4.8, $S^*_{T_{i,j,k}}$ denotes the well-known sensitivity index as it was introduced in Sec. 4.1.1.2 for a certain parameter θ_i, but specifically for the j^{th} model response obtained at the k^{th} position. $\sigma_{j,k}$ denotes the standard deviation of the j^{th} model response obtained at the k^{th} position and $\max_k \sigma_{j,k}$ corresponds to the maximum of the standard deviation of the j^{th} model response obtained from all of the considered positions. It must be clearly differentiated between different types of model responses like pore water pressure and vertical displacement and same type of model responses obtained at different positions of the model, even though they are all merged in the vector of model responses $\boldsymbol{y}$.
Multiplying S_{T_i}, obtained at the k^{th} position and for the j^{th} model response, with the corresponding $\sigma_{j,k}$ allows to consider the relevance of a certain model response. In case in-situ measurement data is available, it can be included in this value to account e.g. for less reliable devices. Placing $\sigma_{j,\max}$ into the denominator allows to normalise the term and to compare thereby sensitivity indices that refer to different output types regarding their relevance to identify a certain input parameter.
The modified sensitivity index $S^*_{T_{i,j,k}}$ should be calculated for any output and position that is considered as candidate for a sensor in an experimental design, i.e. for any considered combination of $\boldsymbol{\delta}$. In case no specific candidate points are defined in advance, it is recommended to generate a grid of points in the candidate area as will be demonstrated in the following sections. After obtaining $S^*_{T_{i,j,k}}$ in each of the grid points, contour plots can be generated by interpolation that allow a better visualisation and interpretation of the results. These contour plots should be used first of all for consistency to ensure that the preceding calculations were correct. Beside, these plots provide new insights into the model behaviour such that, even without using them for OED, they can contribute to the validation of the model.
With regards to questions of designing an experimental set-up, those areas of the contour

plots with highest values of $S^*_{T_{i,j,k}}$ are the best locations to place sensors of the corresponding response type j. The different values of $S^*_{T_{i,j,k}}$ should be used as weighting factors w_i of the objective function (see Eq. 3.17) in a subsequent parameter identification.
In case different time steps should be considered, the approach is analogue to having different model response types. For each time step, $S^*_{T_{i,j,k}}$ values are calculated considering now model parameters, type of model response, position, and time. This procedure is presented exemplary in Sec. 4.2.2 for the case of tunnel construction, where the tunnel advancement changes with the considered time step and accordingly the suggested measurement arrangement, and also in Sec. 4.2.1 where the progression of a laboratory experiment is investigated accordingly. Having that many different dimensions for which $S^*_{T_{i,j,k}}$ must be calculated individually shows how beneficial it is to perform a general GSA in advance that considers all (or at least all uncertain) model parameters and their influence on the model response of interest. Afterwards, the calculation of $S^*_{T_{i,j,k}}$ and generating the contour plots is effectuated only for the relevant parameters, saving a lot of time and computational power.
The advantage of this method is that it is quite time efficient (at least compared to those methods introduced in Chapters 5 & 6) and that it delivers good additional insights to the model structure. However, as this method does not test specific sensor arrangements, it might be quite inaccurate. It does not allow to quantitatively compare the utility of two specific experimental designs with different locations and types of sensors without additional calculations. To do so, these experimental designs must be tested, using artificial measurement data as real data would be usually not affordable (or not available, as the experimental design should be identified before its installation). As an infinite number of feasible experimental designs is possible, finding the best experimental design in this way is impossible.
However, in many practical applications, the task of OED is not that much sophisticated, but more of the kind of "Three sensors should be placed. In which area will they be most meaningful?" For such cases, the introduced method is a powerful instrument that provides comprehensible results.
In case one of the more sophisticated methods that are presented in the subsequent chapters are employed, the introduced approach of using spatial GSA can indeed strongly contribute as preliminary step. Doing so, relevant areas in time and space are identified in advance and the experimental design parameter space of the detailed investigation can be retraced accordingly.

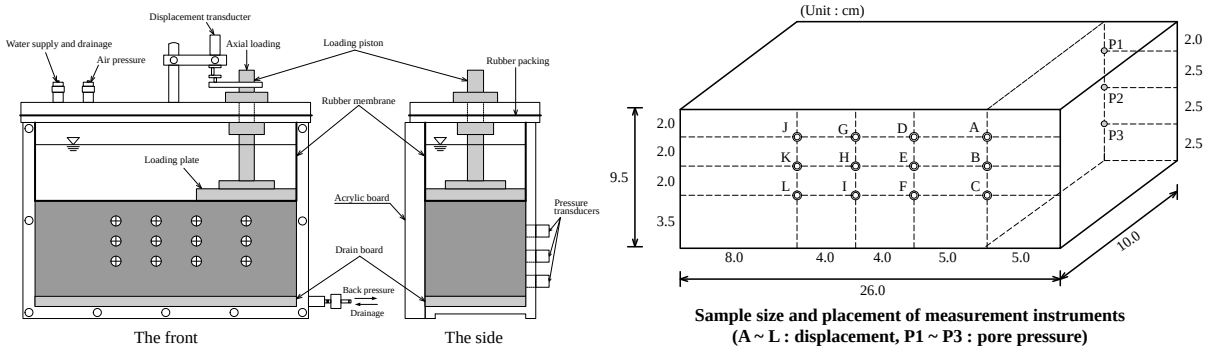
4.2. Application examples

4.2.1. Geotechnical testing device

To demonstrate the applicability of the method of spatial GSA introduced in Sec. 4.1.2, an example is required that should be very simple to understand at first view. The set-up of the testing apparatus introduced in Shuku et al. (2012) seemed to match this requirement. It was developed to investigate the behaviour of soft clay under anisotropic loading and to validate constitutive models.

4.2.1.1. Experimental design and numerical model

A rectangular shaped reconstituted soil sample is placed in the device that is displayed in Fig. 4.1a and uniformly loaded by a constraining pressure applied by air pressure on a water surface. Besides, it is eccentrically loaded with a loading plate on the border of the device to obtain anisotropic loading conditions. Drainage of the sample is possible at the bottom of the device by a filter plate, while the top of the sample is sealed with an impermeable membrane avoiding contact between the sample and the water loading. The sides consist of an impermeable and inflexible frame. Besides, at the boundaries at the top (between soil and membrane) and on the left side (between soil and rigid device wall), a sheet of filter paper is placed whose influence on the test results is not fully interpreted. Measurements are performed in several spots shown in Fig. 4.1b. Herein, the positions A to L refer to positions where horizontal and vertical displacements are recorded using particle image velocimetry, while at positions P1 to P3 pore water pressures were measured. Besides the original intention of investigation of soil properties and calibration of constitutive models as performed in Shuku et al. (2012) and Nishimura et al. (2014), with respect to the present thesis, the question of OED can be posed to this experimental set-up. The precisely defined experimental set-up allows its accurate reproduction as an FE model and in case different sets of measurement results are provided by the experimenters, these data can be considered with respect to model validation purposes. An FE model of the experimental device is generated as plain strain model as shown in Fig. 4.2 with fixed side and bottom boundaries and drainage allowed at the bottom. On top, a distributed load is applied across the entire width of the sample to consider the water load of 49 kN/m^2 on the membrane that is employed to consolidate the sample. On the right-hand side of the sample, an additional distributed load is applied to simulate the impact of the loading plate and that is, in contrast to the water load, modelled using a stiff



(a) (b)

Figure 4.1.: Testing device (Shuku et al., 2012) a: Set-up of model test equipment b: Block sample size and placement of measurement instruments.

loading plate. During the six hours lasting experiment, three load increments of 30 kPa are applied via the load plate to reach a total load of 90 kPa additionally to the water load. To account for the circumstance that the load is not immediately transferred from the load piston to the soil body, the load is applied as a ramp function over short time in the FE model as displayed in Fig. 4.2b. The impact of this ramping is shown in Fig.4.2c where instead of a sudden jump with unrealistically high values, the pore water pressure in the three considered points increases slightly time shifted with lower amplitude and qualitatively closer to the measured values.

The caused displacements are strongly time dependent due to the low permeability of the Kasaoka clay (liquid limit $w_{\mathrm{L}} = 102$ %, plastic limit $w_{\mathrm{P}} = 30.6$ %, particle density $\rho_{\mathrm{s}} = 2.694$ g/cm^3, clay fraction = 55 %, silt fraction = 45 %). To simulate the constitutive behaviour of the clay, the Modified Cam-Clay model is employed (Roscoe and Burland, 1968), that is one of the first models that implemented the concept of critical state soil mechanics and that is often seen as one of the ancestors of the Cam-Clay family of constitutive models. Herein, the stress dependent stiffness is described by the parameters λ for virgin compression and κ for swelling (un- and reloading). These two parameters can be obtained from oedometer test results plotted in a σ-e diagram where the two stiffness parameters are tangent to the experimental data that are assumed to follow a linear trend when the stress is plotted in logarithmic scale. This is approximately the case for normally consolidated soils as encountered in the considered experiment. Further aspects that are often encountered in natural soils as anisotropy, structure, or creep are not considered in this model but have been subject of several later improvements of the model as described in Hölter et al. (2016). The hydro-mechanical coupling of the model is effectuated by considering that the soil's permeability coefficient k decreases with decreasing void ratio

as described by Eq. 4.9

$$k = k_0 \cdot 10^{\frac{e_0 - e}{c_k}}, \tag{4.9}$$

where k_0 denotes the initial permeability of the soil and e_0 is the initial void ratio. c_k is the correlation coefficient that defines how strong the permeability is depending on changes in the void ratio and that should be defined between 0 and 1.

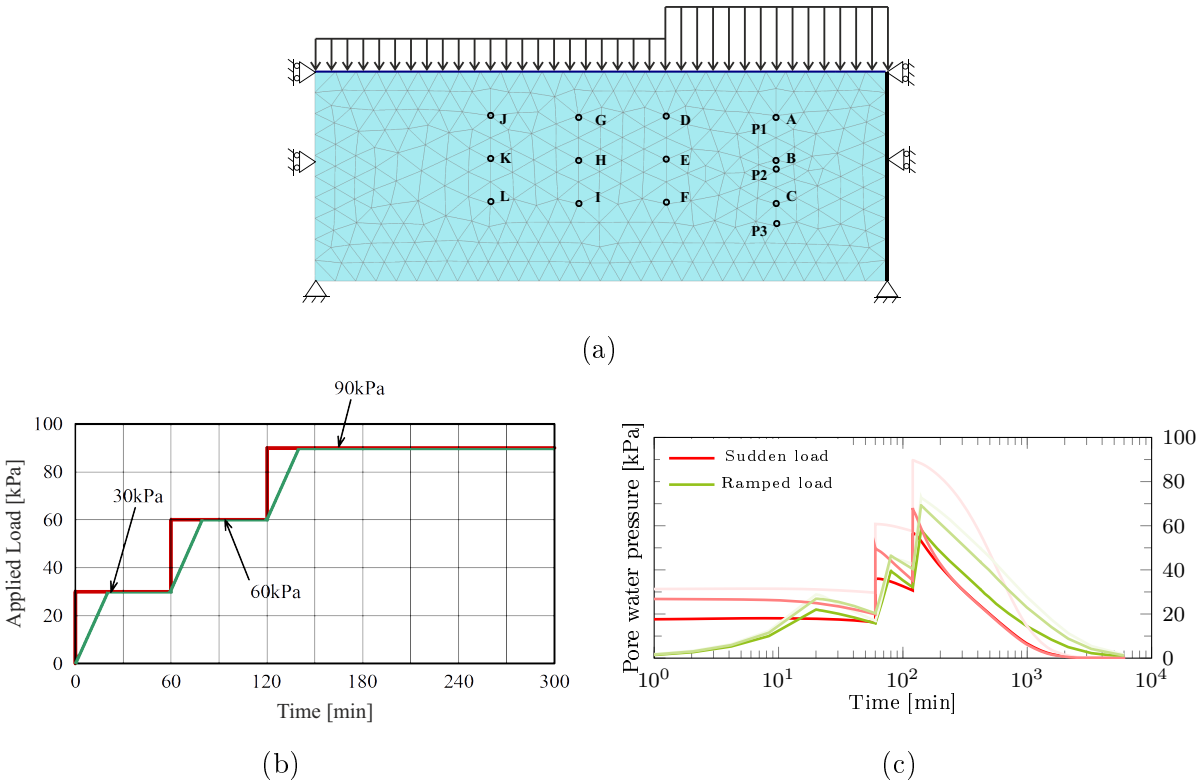


Figure 4.2.: Numerical implementation a: FE-model of considered experiment b: Visualisation of experimental (red) and numerical (green) loading pattern c: Pore water pressure over time in simulations with sudden (red) and ramp-like (green) loading and the three curves corresponding to the sensor spots P1, P2, and P3.

4.2.1.2. Parameter identification

Before applying the introduced OED techniques to the considered application, an initial parameter optimisation is performed using the available measurement data but no further information about the soil properties. The parameters of the model that are assumed to be unknown are κ, λ, the inclination of the critical state line M, the Poisson's ratio ν,

Table 4.1.: Considered MCC parameter bandwidths

Parameter	lower bound	upper bound	identified value	unit
λ	0.07	0.300	0.300	[-]
κ	0.0158	0.122	0.089	[-]
M	0.772	1.420	1.25	[-]
ν	0.025	0.49	0.36	[-]
c_k	0.2	1.0	0.23	[-]
k_0	$1 \cdot 10^{-10}$	$1 \cdot 10^{-8}$	$5 \cdot 10^{-9}$	[m/s]

the correlation coefficient c_k, and the initial permeability coefficient k_0. Hereby, it is assumed that the permeability is the same in horizontal and vertical direction ($k_x = k_y$). The bandwidths of the parameters that must be defined as candidate search space to which the optimisation algorithm is applied is obtained from accepted reference literature (Mitchell and Soga, 2005; Taylor, 1948) and given in Tab. 4.1. Within the ranges of Tab. 4.1, two sets of parameter samples are pseudo-randomly generated using LHS and run in the model. With the results of the first set of 100 samples, a metamodel using the POD RBF approach is generated as described in Sec. 3.2.3, while the second set of 25 samples is employed to test its accuracy. As results of the FE-model, the vertical and horizontal displacements of the points A to L and the excess pore water pressure in the points P1 to P3 are recorded.

Using the genetic algorithm, that set of parameters is searched that allows to reduce most the discrepancy between the response of the metamodel and the measured data, whereby the cost function is formulated according to Eq. 3.18 to consider the vertical displacements in the points A, B, and C and the pore water pressure in the points P1, P2, and P3. The non-employed measurement data can be employed afterwards for verification. For these six measurement sets, the outputs are considered at the beginning and at the end of each loading stage and for that moment when all measurements converged to a final value, corresponding to 42 values. As all measurements are considered equivalent, the weighting factor w_i in Eq. 3.18 is set to the same value for all 42 values. The identified parameter set is displayed in Tab. 4.1. The value of 0.3 for λ is located on the edge of the defined boundaries what should be avoided as it might indicate an artificial minimum of the objective function, avoidable with wider boundaries. Therefore, an extension of the parameter space was manually performed by testing further parameter samples with larger values of λ, but the identified set turned out to still provide the best results.

For a visual comparison, the results are displayed in Figs. 4.3a & 4.3b. Herein, it can be seen that the final settlements correspond in model and simulation. The pore water pressure is fully dissipated at the time of final settlements, indicating that no further settlements can be expected. However, the displacement behaviour over time does not fit well, as in the simulation larger settlements take place earlier than in the experiment. Correspondingly, one could expect a decrease of the pore water pressure in the simulation that takes place earlier than in the experiment, but this is not the case. The curves in Fig. 4.3b show a high quantitative conformity between experiment and simulation, but exhibit the qualitative discrepancy that in the simulation the pore water pressure in point 1 is larger than in point 2 while the experimental results are vice versa.
To verify if the process was successful, it is recommended to consider not only data included in the process of inverse analysis, but also independent data. Therefore, the vertical displacements over time in the proximate points D to F are considered in Fig. 4.4a.

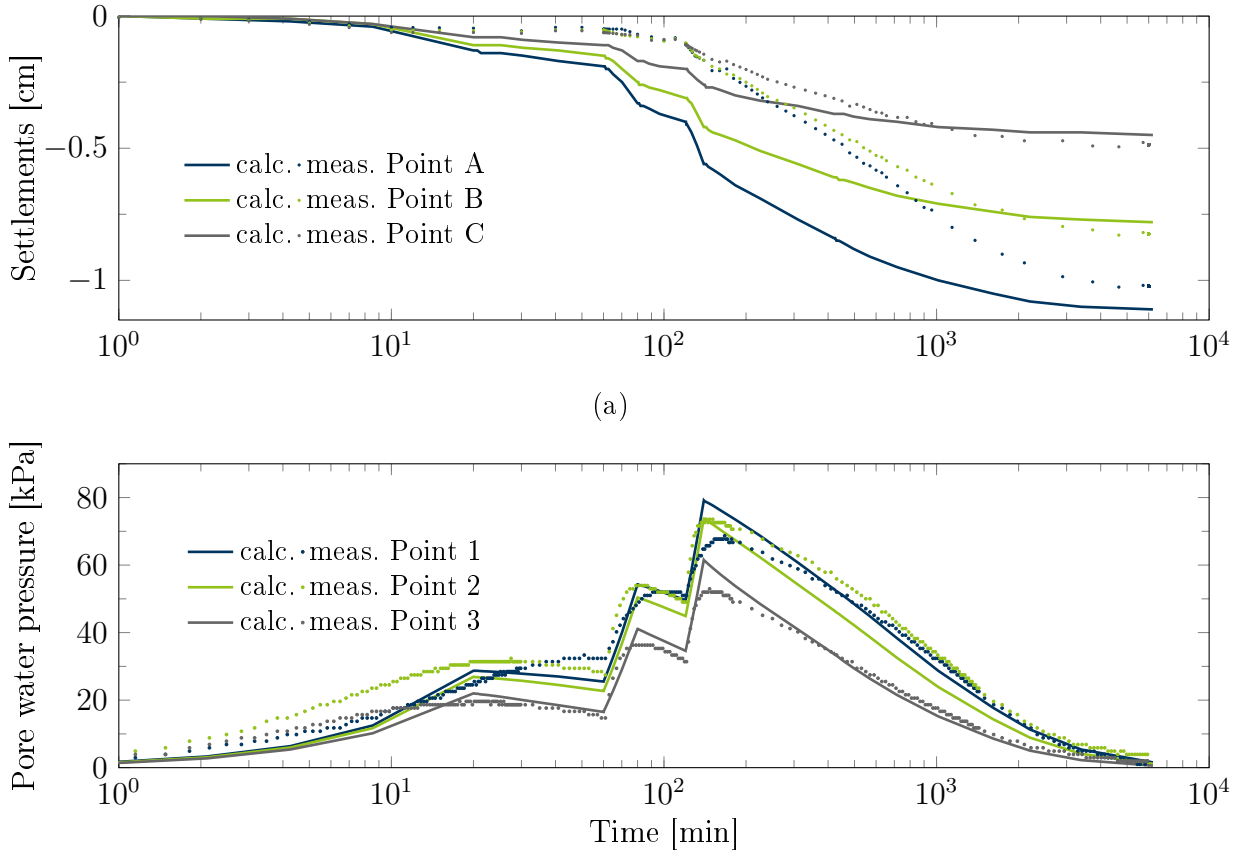


Figure 4.3.: Comparison of measurement data and FE outputs obtained for identified optimal parameter set in case of a: settlements and b: pore water pressure.

It can be seen that the simulation results match the experimental data well, even better in case of the results at intermediate time values. A comparison of horizontal displacements is shown in Fig. 4.4b. Here, neither the qualitative nor the quantitative agreement are satisfying as the displacements in Point A are supposed to be smaller than in Point B and overall, the displacements are overestimated by approximately 100 %.
The qualitative misfit in the Figs. 4.3b & 4.4b discussed above, can most probably not be solved by an improved adaptation of the soil parameters. Especially the circumstance that in Fig. 4.3b the measured pore water pressure is higher in Point 2 than in Point 1 can only be explained by systematic measurement errors or, more probable, by an additional drainage path induced by the filter paper that is wrapped around the soil sample.
To proceed the validation process, a global sensitivity analysis is performed to understand which of the employed parameters are most influencing the model responses. To do so, the procedure of VB-GSA as described in Sec. 4.1.1.2 is employed using the same metamodel and the same parameter bandwidth as for the parameter identification. Hereby, different times are considered: after 60 and 120 minutes, immediately after the second and third loading are fully applied, and after 6000 minutes, when the consolidation process seems to have converged.
The sensitivity results for the vertical and the horizontal displacement in point A as well as the pore water pressure in point 1 are displayed in Fig. 4.5. In case of vertical settlements, y_{y} becomes increasingly sensitive to the primary stiffness parameter λ while the opposite happens with the permeability coefficient k_0 and the remaining parameters are of low relevance. Such behaviour coincides with expectations as λ governs the final settlements in the MCC-model. However, the settlement does not take place immediately but is delayed by the consolidation process as can be seen in Fig. 4.3a. As the consolidation is mostly governed by the permeability coefficient k, this parameter has initially a high influence on the settlements that decreases as the excess pore water pressure dissipates as displayed in Fig. 4.3b. The influence of the parameters on the pore water pressure y_{w} is quite constant over time, whereby one should consider that the y_w value is close to zero at the third evaluation stage after 6000 min.
The results obtained in case of the horizontal displacements in Fig. 4.5b are more irregular and complex to interpret. At the final stage, the influence of λ and M is comprehensible as horizontal displacements will probably be induced by shear stresses. The decrease of relevance of the Poisson's ratio ν might be explained by the pore water pressure that avoids high effective stresses in the soil and therefore plastic deformations. The elastic deformation again is related to ν, wherefore this parameter is only relevant in the early

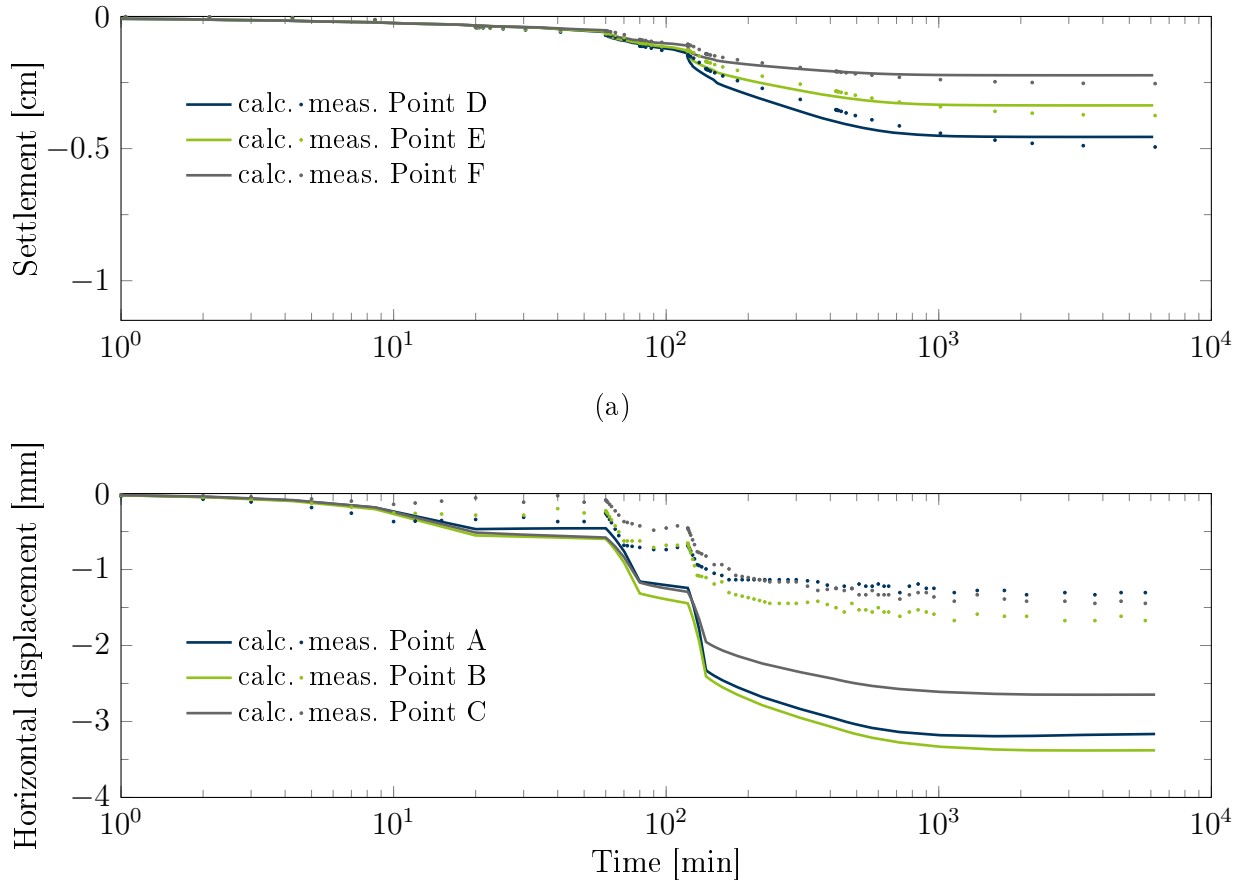


Figure 4.4.: Comparison of model and measurement data not employed in inverse analysis in case of a: vertical displacements in Points D to F and b: horizontal displacements in Points A to C.

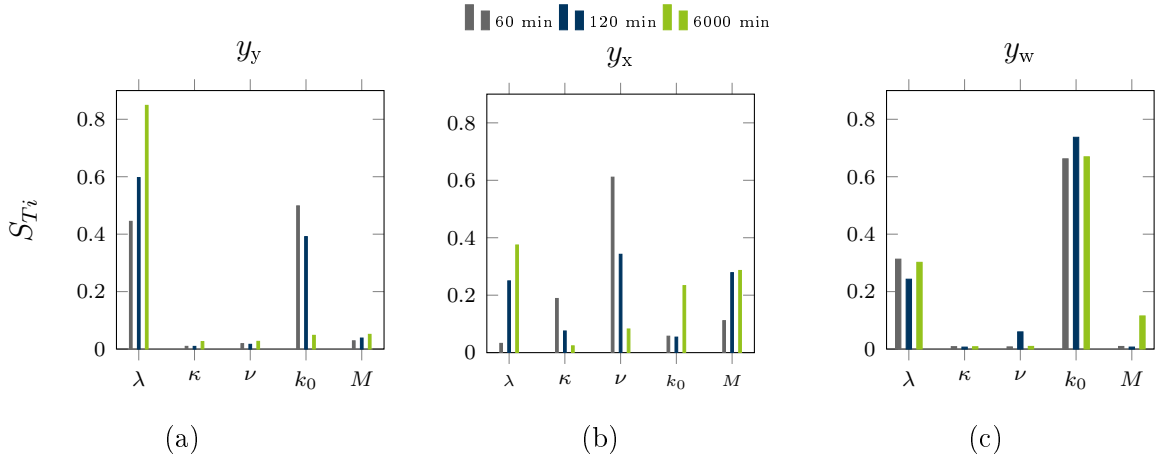


Figure 4.5.: Global sensitivity analysis of model outputs for a: vertical displacements y_y and b: horizontal displacements y_x in Point A, and c: pore water pressure y_w in Point 1.

phases of the simulation. Together with the outcome of Fig. 4.4a, the impression arises that an investigation of horizontal displacements is not reliable when using the MCC model for this problem.
The dependency of the results on the considered time step in the case of displacements (Fig. 4.5b) indicates how relevant a well-considered selection of the time step is. One should be aware that the decision from which time step data should be employed for GSA is already part of OED considerations. If the question was which parameter is most influencing settlement of the soil, using data obtained after 60 minutes would be totally misleading as almost no settlements occurred up to that point.

4.2.1.3. Application of spatial global sensitivity analysis

Besides the question when to obtain meaningful data, also the consideration where to measure is relevant in this example. Therefore, the concept of spatial GSA, as introduced in Sec. 4.1.2, is applied to the considered experiment. The name is selected as the GSA that is often applied pointwise or for global model behaviour is now adopted to a whole model geometry (or at least a part of it). To do so, the parameter space is reduced to those parameters that are identified as most relevant according to the results of the GSA shown in Fig. 4.5, i.e. the stiffness λ and the initial permeability coefficient k_0. Additionally, the considered ranges of these two parameters are reduced to $k = 1.3 \cdot 10^{-9}$ to $1.3 \cdot 10^{-8}$ m/s and $\lambda = 0.251$ to 0.361. While the values given in Tab. 4.1 are supposed to represent possible ranges according to the considered type of soil, these new ranges should correspond to a possible identification accuracy while having the identified value in its middle. Following the approach described in Sec. 4.1.2, a grid of 152 points is generated over the whole model domain in which the outputs of interest are obtained, namely y_y and y_w while y_x is considered to be less reliable. The grid is shown in Fig. 4.6 where each x marks one spot in which model responses are obtained. It can be seen that the arrangement is densified in the upper right corner below the loading plate as it is expected that more relevant differences are obtained in the results in that area. Again, 100 parameter samples are generated to set up the metamodel as well as a second set of 25 samples for the verification of the metamodel. It is noteworthy that the approach to employ the same metamodel as described above in Sec. 4.2.1.2, i.e. POD RBF with same shape coefficient c, was not successful. Even if c is varied over its whole reasonable range, no satisfying agreement between metamodel results and FE results can be achieved. The demanded accuracy higher than 95 % can indeed be obtained using an LSM metamodel that is based on less sophisticated quadratic polynomials. This shows the aforementioned reflections

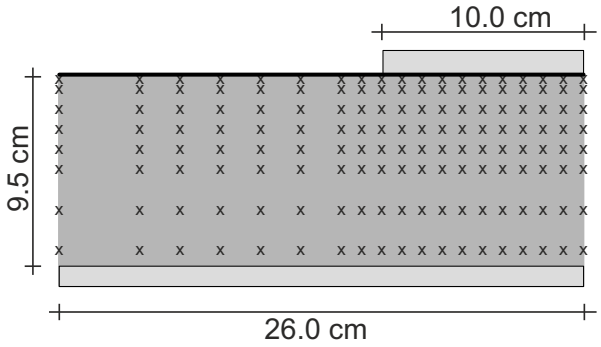


Figure 4.6.: Arrangement of points on side of experimental device in which model responses are obtained to perform spatial GSA.

that a more complex metamodel does not necessarily provide more accurate results, and that any new situation requires an individually assigned metamodel.
Using this metamodel, Eq. 4.8 is employed to calculate the modified sensitivity indices $S^*_{T_{i,j,k,l}}$, whereby i refers to the two parameters of interest k_0 and λ and j to the outputs y_{y} and y_{w}, respectively. By k, the 152 locations are considered. As mentioned in Sec. 4.1.2, the application range of the modified sensitivity index S^*_T can be extended to further dimensions what is performed here by means of time. As shown by the figures in Sec. 4.2.1.2 before, the model behaviour and responses considerably change during the course of the experiment. Therefore, the modified sensitivity indices are obtained at $l = 1, \ldots, 7$ time steps, namely the same as employed for the previous back-analysis and at mid-term of the final consolidation phase (20, 60, 80, 120, 140, 420, 6000 minutes). For a better comprehension, the S^*_T results are smoothed using quadratic interpolation and visualised using contour plots that are displayed in Fig. 4.7 for the first considered phase, while the results of all phases are shown in Figs. A.1 to A.7 of Appendix A. In each contour plot the scaling ranges from dark blue, corresponding to zero to dark red for values close to one. Above the contour plots, two diagrams with the course of y_{y} and y_{w} over time are displayed showing which time step is currently considered. In each time step, the two contour plots on the left-hand side display S^*_T of pore water pressure towards k_0 (middle figure) and λ (bottom), while the plots on the right-hand side contain analogous presentations for the settlements.
In the first considered time step (Fig. A.1), the plots are well defined: displacement data is most sensitive in the upper right corner, close to the loading plate, whereby it is more sensitive to λ than to k_0. Pore water pressure indeed is most sensitive to areas at the bottom of the device, here, k_0 being more relevant than λ. Initially, it might be confusing

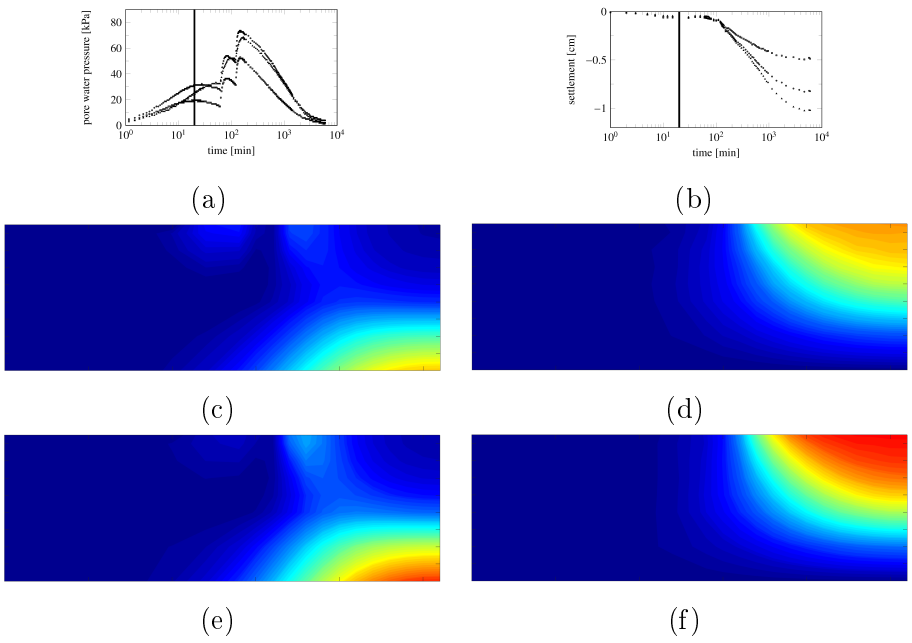


Figure 4.7.: Vertical bar pointing out the considered time step after 20 minutes with respect to a pore water pressure y_w and b settlement y_y evolution; sensitivity distribution for y_w towards c: k and e: λ, and for y_y towards d: k and f: λ

why the soil properties in the area at the device's bottom should be relevant for the pore water pressure, but if remembering the essentials of chapter 2 on OED, it is clear that not the maximum value is decisive, but the maximum gradient. At the top of the model, close to the loading plate, there is no possibility for the excessive water to drain easily. Accordingly, almost no dissipation of pore water pressure can take place up to the considered time step, so that the pore water pressure will stay the same, independently of the soil parameters. At the bottom indeed, due to the drain board installed at the bottom of the model container, drainage paths are short and pore water pressure dissipation can take place, the faster the larger the value of k_0 is. As this direct relationship exists between soil property and measurement outcome, it makes sense to perform measurements of the pore water pressure at this location of the device in this early phase of the experiment.
In the next phase e.g. Fig. A.2, consolidation has been ongoing for 40 minutes and excess pore water pressure seems to be mostly dissipated in the area close to the bottom. Therefore, the areas of high sensitivities of y_w are "moving" upwards in direction to the loading plate. After 80 minutes (Fig. A.3), the second loading occurred and sensitivity patterns overlap to a certain extend, leading to the contours shown in Figs. A.3c & A.3e. With ongoing time of the experiment, where loading and consolidation phases alternate but also overlap, the results' interpretation becomes more complex. As the S_T^*-values are

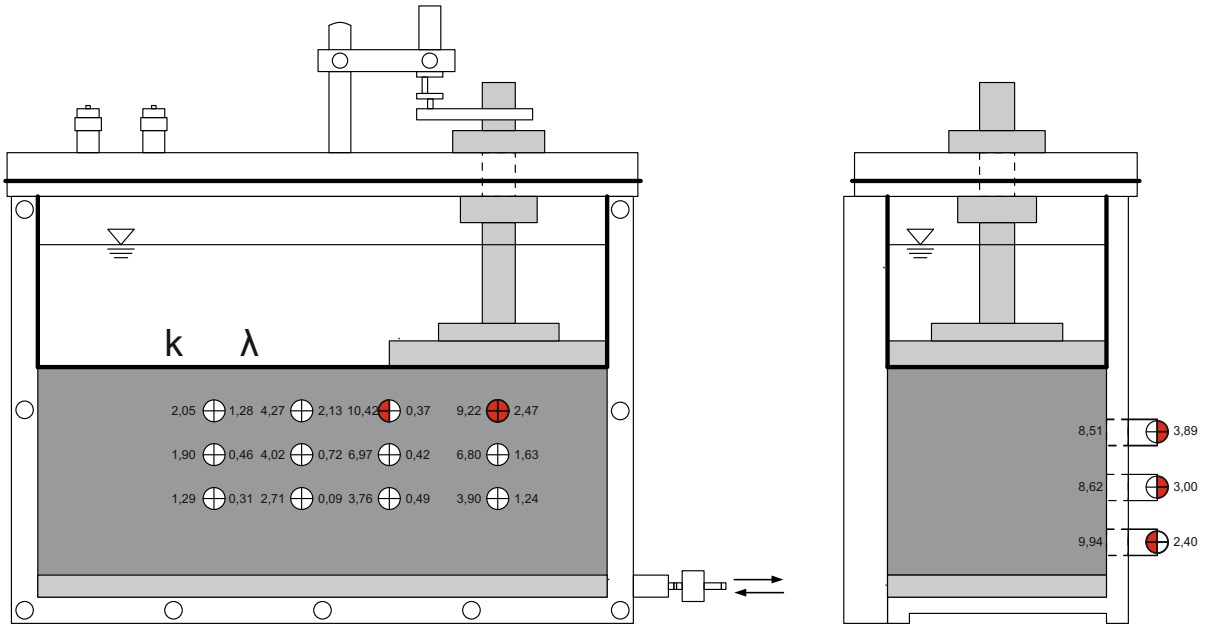


Figure 4.8.: Weighting factors obtained according to GSA for each of the existing measurement positions for time step $t = 140$ min.

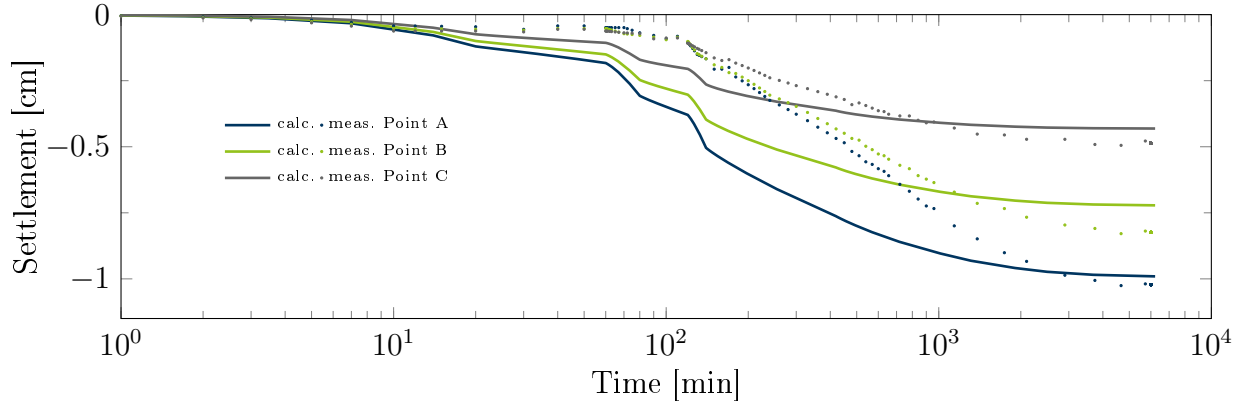
normalised by the maximum deviation value of the specific phase, the contours indicate most relevant areas in each time step individually, but do not rank among the different time steps.

The next step in sense of OED is to employ this information to suggest an improved arrangement of the sensor configuration. As this is obviously not possible in the considered steps as the experiment is completed, the gained data can be employed to suggest an improved weighting of the measurement data for parameter identification. To do so, the values of S_T^* are calculated in the points A to L for the displacements and in the points P1 to P3 for the pore water pressure, separately for each considered time step. These values can by employed as weighting factors in the objective function to be employed as shown in Eq. 3.18. This allows to consider some of the measurement data as more important than others. The measurement data with higher weighting is obtained in positions where the parameter to be identified is more sensitive to the specific output.

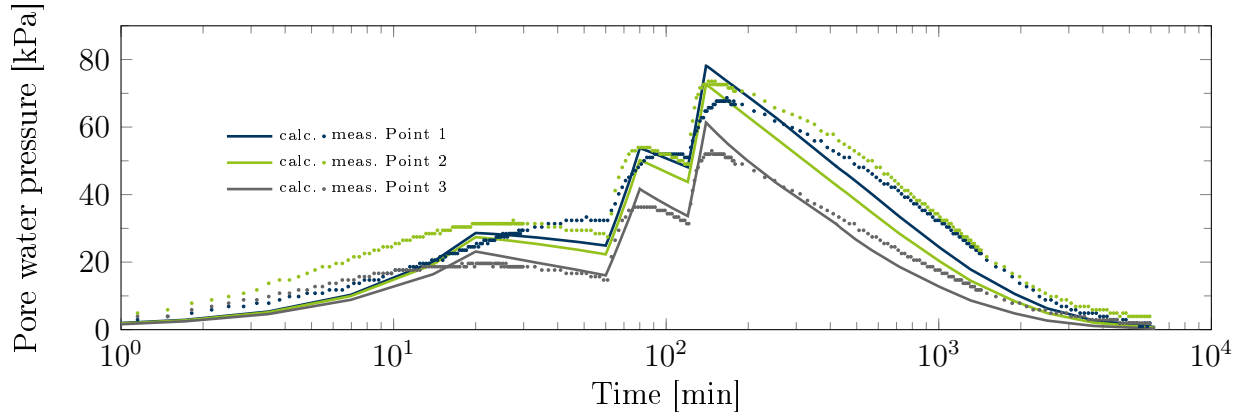
The outcome is displayed in Fig. 4.8 where on the right-hand side of each sensor position the modified sensitivity index with respect to λ is displayed while the corresponding value for k_0 is provided on the left-hand side. Exemplary, the results are displayed for time step $t = 140$ min, but the procedure is performed for all considered time steps. Doing so, 42 weighting factors $w_{k,l}$ for the specific time step and position are obtained and included in the subsequent parameter identification, where only the two relevant parameters are

considered while the remaining four are fixed to the values determined in the initial parameter identification.

As result again a value 0.3 is obtained for λ while $5.53 \cdot 10^{-9}$ m/s is identified as optimal solution for k_0. The model responses with this parameter set are shown in Fig. 4.9. In comparison to Fig. 4.3, it can be seen that the final settlement of Point A is matched better, but besides, no fundamental improvement can be observed, especially with respect to the qualitative misfits discussed above. Further modifications of the objective function are considered by additional weighting factors proportional to the time between the reference points or logarithmic scaling of these time spans, by formulating independent objective functions for y_w and y_y, or, as introduced in Sec. 3.3.2, by employing the Hausdorff metric as an alternative definition of distance in the objective function. However, none of these methods provides substantial improvement in reducing the discrepancy between measured



(a)



(b)

Figure 4.9.: Comparison of measurements and simulation data after optimisation for a: vertical displacements y_y and b: pore water pressure y_w.

and simulated data. Nevertheless, employing these contour plots significantly contributes to the validation of the model. The inaccurate simulation of the drainage paths could be identified and deeper understanding of the model parameters' interaction is obtained. In context of improving the experimental setup, it could be shown that the same quality of model validation could be obtained using much less measurement data. This would allow to perform and validate future experiments using the same device more time and cost efficient.
A reduced part of the results shown in the present Sec. 4.2.1 were published previously in Hölter et al. (2015). There, the focus is set on generating the contour plots shown in Appdenix A by calculating the modified sensitivity index S_T^* for the whole geometry of the model test in one exemplary time step. Thereby, it could be identified how the following parameter identification could be effectuated most efficiently.

4.2.2. Mechanised tunnel construction

The application example shown in the previous section is easily understandable, because of the simplicity of the problem, but might be of low relevance for engineering practice due to its specific hydraulic and mechanical boundary conditions and its measurement instrumentation. Therefore, the two following sections consider the topics of mechanised tunneling and water-retaining dikes where extensive monitoring programs are common and offer a large optimisation potential.
Mechanised tunnel constructions are in terms of size, cost, construction time, risk potential, and conflict potential among the involved stakeholders, some of the largest and most difficult projects encountered in civil engineering. Especially urban tunnel construction projects must be planned and executed very carefully as the advancement of the tunnel may cause settlements at the ground surface and thereby induce damage to adjacent buildings. The complex processes of mechanised tunnel construction and approaches to implement them in FE-models are described in the following. Subsequently, it is described how the concept of OED can be considered in this context.

4.2.2.1. FE-simulation of mechanised tunnel construction

The construction of tunnels inevitably leads to a redistribution of stresses around the tunnel, as a consequence of excavating material and accordingly reducing soil weight, but also due to the forces applied to the soil by the tunnel boring machine (TBM) during the construction. Depending on the host material of the tunnel this often leads to strains and

displacements that propagate in form of settlements (or upheave) to the ground surface. In case of urban tunnels with small overburden depth, this can cause costly damages to the surface infrastructure. Compared to conventional tunnel excavation with shotcrete stabilisation, mechanised tunneling induces less disturbance as the stress level can be maintained to a certain extend by the TBM. The set-up of a modern TBM is displayed in Fig. 4.10 for the case of a slurry- or mixshield. The complexity of this machine can hardly be captured entirely in a FEM simulation that is primarily focused on soil behaviour. Therefore, certain simplifications are undertaken in the simulations considered in this thesis: the advancement of the TBM and hence the progress of the tunnel excavation is accompanied by the installation of ring segments consisting of usually six to ten concrete lining elements, placed one by one by the errector (Fig. 4.10-9). The installation of these elements might be simulated individually as performed in Gall et al. (2018) or using a continuous plate element as in the present work. This simplification introduced in Blom (2002) is valid in case that the focus is not laid on the lining forces and stresses, but on the behaviour of the surrounding soil. To account for the reduced bending stiffness due to the joints of the elements, the stiffness of the plate elements is hereby reduced accordingly. The interaction between the tunnel and the surrounding soil is ensured by the grout that is injected at the tailskin of the TBM (Fig. 4.10-10) all around the lining ring. The simulation of the grout and its pressure is a complex hydromechanical process and content of numerous publications (e.g. Masini et al. (2014), Meschke et al. (1996), Lavasan et al. (2018)). The approach employed in the models considered here consists of applying a radial surface pressure between the lining elements and the shield elements which amount should equal the applied grout pressure. As in each excavation phase, the TBM advances by the width of one segment, the "grouting pressure segment" follows, assuming that in the previous segment the final stiffness of the grout is reached.
The technical need for the grout backfill results from the overcut of the cutting wheel and the TBM conicity, necessary for the TBM to advance, and the thickness of the shield itself. These details are reflected in the FE simulation by applying a contraction factor to the lining elements, corresponding to a volumetric strain. The second pressure transferred by the TBM to the soil during excavation is the so-called face pressure that is usually applied either by bentonite suspension or by foam injected earth, depending on the surrounding ground conditions, whereby Fig. 4.10 illustrates a Mixshield or Slurryshield TBM, employed in coarse grained soils. The face pressure is necessary to enable the rotation of the cutting wheel (Fig. 4.10-1) and to prevent a collapse of the soil as described e.g. in Mollon et al. (2013). As on the other hand a blow-out is possible in case of a too high face pressure, the steering of this value needs to be well considered and continuously adjusted

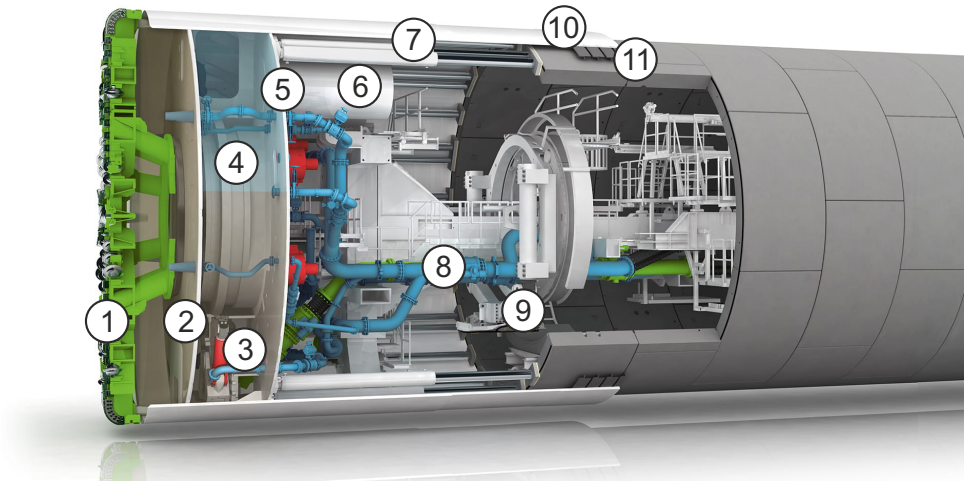


Figure 4.10.: Usual set-up of mixshield TBM: 1) cutting wheel, 2) submerged wall, 3) jaw crusher, 4) air cushion, 5) bulkhead, 6) air lock, 7) thrust cylinders, 8) slurry circuit, 9) erector, 10) tailskin, 11) backfilling, modified after Herrenknecht AG (2020).

just like the grout pressure. In the employed models, the face pressure is simulated by a distributed load that linearly increases with depth and that corresponds to the local horizontal earth pressure. Detailed descriptions on FE simulations of tunnels, especially with respect to the model validation techniques employed herein can be found in Zhao (2018). The accurate simulation of the tunnelling process can play a major role in the early prediction of the ground reaction, what allows to adapt the steering parameters of the TBM or to initiate external counter-measures. However, this requires a careful model validation based on in-situ data. As the TBM continuously advances, it would also be necessary to repeat the model validation, depending on the current location of the TBM. These circumstances show the relevance and potential of employing OED in the framework of mechanised tunnel constructions.

4.2.2.2. Considered example of tunnel excavation

To apply the method of spatial sensitivity analysis, a model is generated of a synthetic example of a tunnel construction in vicinity of a building at the ground surface. The employed material and process parameters as well as the geometries correspond to usually encountered properties in such applications. The diameter D of the tunnel is 8.5 m with a width of the lining segment of 1.5 m that corresponds to the advancement within each simulation step. The length of the TBM is assumed to be 9.0 m, equal to the width of 6 lining segments and advancement steps, respectively. The overburden depth is 1 D and the

contraction factor applied to the plate elements of the TBM is 1 %. To consider the tunnel-building interaction, at the ground surface a quadratic building is placed with an offset from the building centreline to the tunnel centreline of 1.75 D and a width of 1.5 D. Only the foundation slab is simulated, but according to Franzius et al. (2004) with properties that correspond to the weight and stiffness of a 10-storey building. The first simulation steps of the model are as follows. First, the soil material in the geometry model is activated by the method of $K0$-procedure to generate an initial stress distribution, defined only by the weight of the employed soil. Herein, according to Jaky $K0$ is assumed to be $1 - \sin\varphi$. In the next phase, the building foundation is applied with its assigned load. To consider only the incremental impact of the tunnel, the displacements caused by the self-weight and the building are reset to zero. After this reset, the actual tunnel excavation starts by removing the first soil cluster in the tunnel course and activating the plate elements of the TBM-shield. By repeating this process, the whole width of the model geometry is crossed within 48 consecutive calculation phases whereby the plate elements of the shield are replaced after six phases first by the "grout segment" and then by the elements of the lining.

To the foundation, the lining, and the TBM shield, linear-elastic material properties are assigned that are provided in Tab. 4.2. The contact between the soil and the structural elements, that are the foundation as well as the lining segments and the TBM-shield, is simulated using interface elements that allow relative displacement of adjacent mesh elements along the interface. The interface friction can be controlled by means of the reduction factor $R_{\mathbf{int}}$ that is set to 0.6 in the current application. This value can vary between zero and one, whereby one corresponds to a full frictional contact ($= 1 \cdot \tan\varphi'$). The excavation length of the tunnel, corresponding to the model length accounts to 72 m. The further model dimensions are as shown in Fig. 4.11: 40 m depth and 102 m width to ensure that boundary effects are avoided. However, the model bottom boundary is geometrically fixed while the surface is free.

Table 4.2.: Considered parameters of linear elastic constitutive model for tunnel lining, TBM shield and building.

Parameter	Lining	TBM-shield	Building	Unit
Bending stiffness (EI)	$2.18 \cdot 10^5$	$7.50 \cdot 10^5$	$4.39 \cdot 10^9$	[kNm^2/m]
Axial stiffness (EA)	$1.05 \cdot 10^7$	$7.35 \cdot 10^7$	$3.80 \cdot 10^7$	[kN/m]
Weight	25	90	-	[kN/m^3]
Load	-	-	100	[kN/m^2]

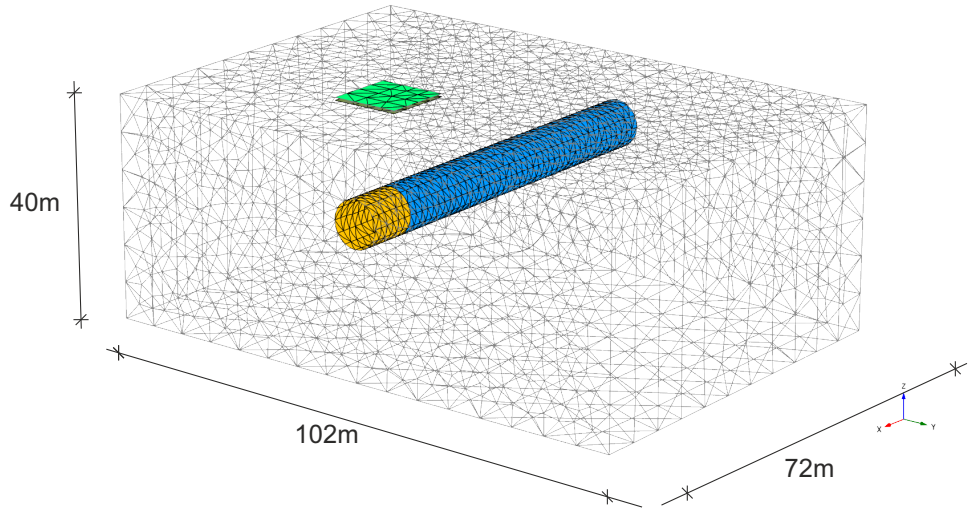
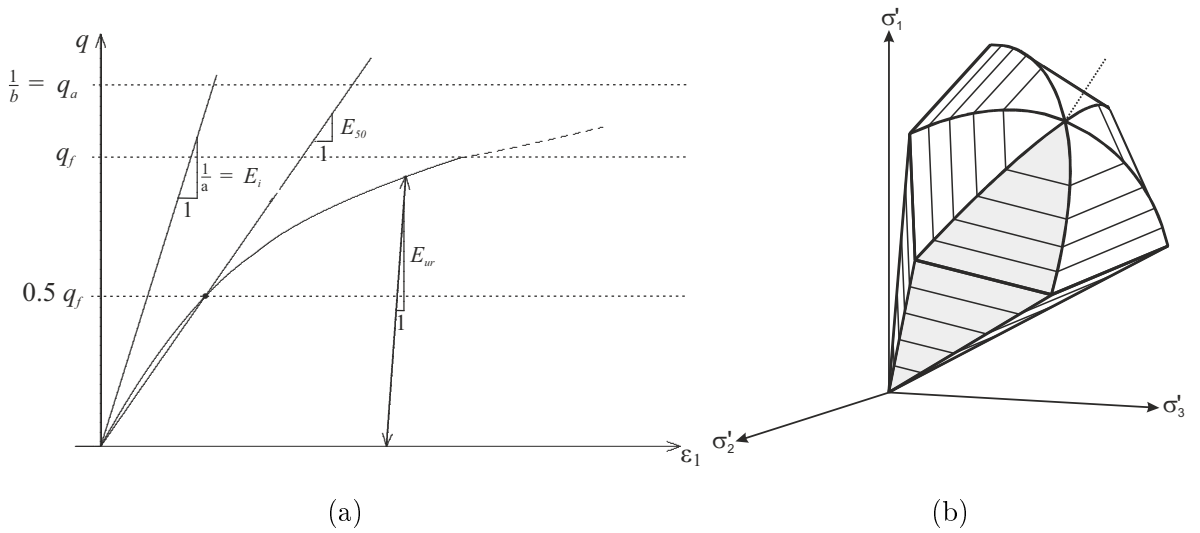


Figure 4.11.: Employed FE model discretisation and model dimensions of 3D tunnel simulation (yellow: TBM shield, blue: lining segments, green: building foundation).

On the four sides, the model boundaries are fixed in horizontal direction but free in vertical direction. The whole soil domain is simulated using the Hardening Soil (HS) constitutive model. This model, introduced in Schanz (1998) and further described in Schanz et al. (1999) employs the same failure criterion as the Mohr-Coulomb model, but extends it with stress-depending hardening formulations for isotropic and deviatoric loading. Thereby, the simulated stress-strain behaviour is closer to that of real soils as in case of the Mohr-Coulomb (MC) model. The HS model is of the same model hierarchy as the MC model and became quite popular in the geotechnical community. It was extended in Benz (2007) to account for the increased stiffness at small strains. Depending on the type of loading, different stiffness moduli are considered as illustrated in Fig. 4.12a. In case of confined boundary conditions, the stiffness modulus $E_{\text{oed}}^{\text{ref}}$ is considered while for deviatoric loading the secant stiffness E_{50}^{ref} is employed. Hereby, the index 50 refers to the 50% of maximum shear stress q_f that is used to construct the secant in the $\varepsilon_1 - q$ diagram. In case of un- and reloading, the deformation behaviour is defined by the stiffness $E_{\text{ur}}^{\text{ref}}$. The stress dependency of these three moduli is defined as displayed in Eq. 4.10 for E_{50}, while it is analogue for the two other stiffnesses.

$$E_{50} = E_{50}^{\text{ref}} \left(\frac{c' \cos\varphi' - \sigma_3' \sin\varphi'}{c' \cos\varphi' + p^{\text{ref}} \sin\varphi'} \right)^m \tag{4.10}$$

Herein, c' and φ' are the well known effective shear strength parameters cohesion and friction angle. p^{ref} reflects the reference effective stress at which the reference stiffness E_{50}^{ref} was identified, i.e. usually the constant effective lateral stress in a triaxial test, and σ_3' is



(a) (b)

Figure 4.12.: Hardening Soil (HS) model considered for the FE model a: Deviatoric strain - axial strain relationship with definition of different stiffnesses, b: yield surface in the principal stress space, modified after Schanz (1998).

the ambient minor principal stress. The exponent m is a parameter without direct physical meaning that must be identified on appropriate test data. It defines how strong the stress dependency of the respective stiffness is. In case of $m = 0$ the stiffness remains constant like in the Mohr-Coulomb model. Usually, for sands values of 0.5 to 0.8 are recommended while for soft clays the value should be around 1.0.
The yield surface of the HS model is displayed in Fig. 4.12b, whereby one should differentiate the deviatoric yield function $f_{\mathbf{dev}}$:

$$f_{\mathbf{dev}} = \frac{q_a}{E_{50}} \frac{(\sigma'_1 - \sigma'_3)}{q_a - (\sigma'_1 - \sigma'_3)} - \frac{2(\sigma'_1 - \sigma'_2)}{E_{\mathrm{ur}}} - \gamma^{\mathrm{p}}, \tag{4.11}$$

where q_a is the equivalent stress at failure and γ^{p} is the plastic shear strain, and the cap yield surface for isotropic loading $f_{\mathbf{cap}}$:

$$f_{\mathbf{cap}} = \frac{3J_2}{\alpha^2} + p'^2 - p'^2_p. \tag{4.12}$$

Here, J_2 is the second principal invariant of the stress deviator, α is a model parameter, p' is the effective mean stress and p'_p is the effective preconsolidation pressure. The parameter values employed in the current model are displayed in Tab. 4.3, corresponding to a quite soft normally consolidated soil. As shown in Zhao et al. (2015) and Miro et al. (2014), the parameters with highest impact on ground settlements induced by mechanised tunnel advancements are the friction angle φ' and the stiffness parameters E^{ref}_{oed}, E^{ref}_{50}, and $E^{\mathrm{ref}}_{\mathrm{ur}}$. All parameters are assumed to be approximately known, but as these four parameters

are assumed to be of high relevance, it is of interest to identify them accurately and continuously to possibly react in the construction process. Accordingly, it is assumed that the other parameters are known sufficiently well and do not require a further validation as long as the initial general assumptions on soil type etc. are valid. However, according to Schanz (1998) and the experiences made in Zhao et al. (2015), E_{50}^{ref} is defined as being equal to E_{oed}^{ref}, reducing the amount of independent constitutive parameters of interest to be identified to three.

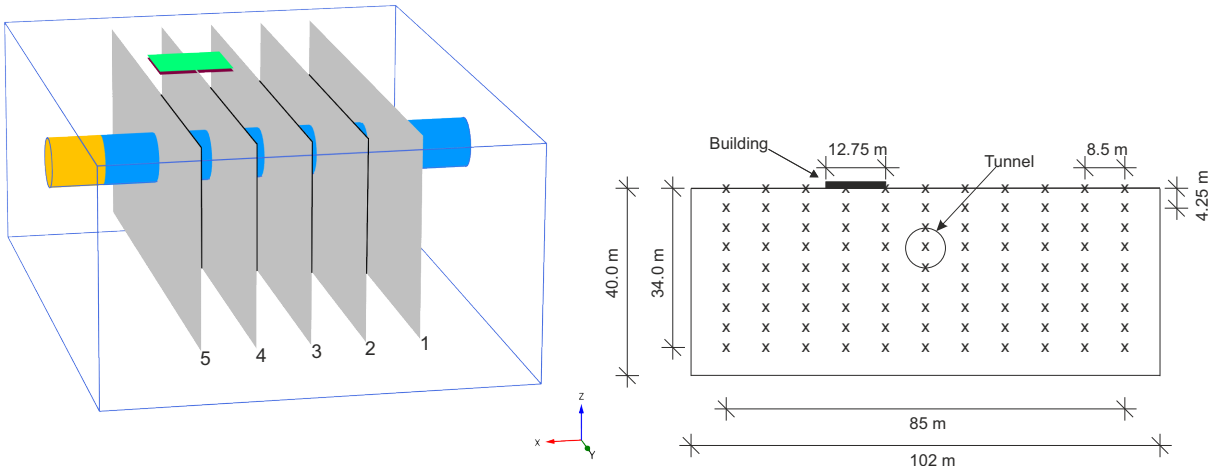
Table 4.3.: Employed HS parameters of the soil in the simulation model of TBM advancement.

Parameter	Description	Value / Range	Unit
φ'	Friction angle	[20; 25]	[°]
ψ'	Dilatancy angle	0	[°]
c'	Cohesion	0	[kN/m^2]
E_{50}^{ref}	Secant stiffness in triaxial test	[5,000 ; 15,000]	[kN/m^2]
E_{oed}^{ref}	Tangent stiffness in oedometer test	[5,000 ; 15,000]	[kN/m^2]
E_{ur}^{ref}	Unloading-reloading stiffness	[15,000 ; 45,000]	[kN/m^2]
p^{ref}	Reference stress	100	[kN/m^2]
m	Exponent	1	[-]
ν_{ur}	Poisson's ratio	0.2	[-]
γ	Soil unit weight	17	[kN/m^3]
K_0	Coefficient of lateral earth pressure	0.57	[-]

4.2.2.3. Spatial sensitivity analysis application

In the considered case of a tunnel that is excavated, any measurable value that could provide information about the soil properties will be caused by the excavation itself. In case no building at the ground surface is considered ("green field" conditions), but only the TBM and its continuous advancement, the optimal experimental design would be constant by its relative position to the TBM. Therefore, the impact of the building and its interaction with the TBM is of high interest in this application example. In Zhao et al. (2015) and Miro (2016) for example, constitutive model parameters were identified based on measurement data of vertical displacements obtained in field and artificially generated, respectively. However, the location of the measurements was not subject of a preliminary OED investigation. The same accounts for the time-dependency of the results as they

were all weighted the same in the objective function to determine the parameters' values. In the present case, this is explicitly intended: finding optimal locations for measuring vertical displacements during the advancement and bypassing of the tunnel.
The concept of spatial GSA, introduced in Sec. 4.1.2 is employed therefore. To do so, some initial assumptions on the parameters of the experimental design should be made. As the considered FE model has a 3D geometry, some simplification is done by not performing a consecutive three-dimensional GSA, but in five vertical cross sections that are passed centrally by the tunnel. The performed measurement, i.e. the FE model response employed in the GSA is the vertical displacement. Measurements shall be performed once in each of the sections each time the TBM reaches one of them (excavation steps 13, 16, 25, 31, and 37), leading to a total of 25 measurement informations to be obtained. It is assumed that the location of the sensors can be changed within the sections in between the excavation steps. The locations of the five sections are shown in Fig. 4.13a. The horizontal distance between the sections accounts to 9 m, what is exactly the length of the TBM. One can see that the first section is located 20.5 m in front of the footing edge and 18 m behind the geometrical model boundary. Thereby it is intended to obtain "green field" conditions for this section without the possibly falsifying boundary effects. Having the equal distance of 9 m between the sections, the fourth section is exactly below the centreline of the building and the fifth is 2.25 behind it.
For each of the sections, a grid of points is generated similar to that shown in Fig. 4.6 in which S_T^* should be obtained. The distribution of the 99 points is shown in Fig. 4.13b.
To enable performing GSA, the FE simulation is run 150 times using different input para-



(a) (b)

Figure 4.13.: a: Position of the considered vertical monitoring sections within the FE model and b: position of GSA points in each of these sections.

meter samples that are generated using LHS within the boundaries given in Tab. 4.3. As results, the vertical displacements in the 99 positions shown in Fig. 4.13b of the five sections shown in Fig. 4.13a in the five considered excavation steps are calculated. For each of these 2475 outputs, the modified sensitivity index $S^*_{\mathrm{T}_i}$ is calculated, whereby i refers to the three parameters of interest: φ', $E^{\mathrm{ref}}_{\mathrm{ur}}$, and $E^{\mathrm{ref}}_{\mathrm{oed}}$ (that equals E^{ref}_{50}). The visualisation of the results is performed again by means of contour plots displaying $S^*_{\mathrm{T}_{i,k}}$ that are shown in Appendix B. As showing all 75 contour plots would be beyond the scope, exemplary the plots obtained in the third (green field close to the building) and fourth (below the building, s. Fig. 4.13a) section are shown for the three considered parameters. In Figs. B.1 to B.3, the third section is considered. That means in the first two considered excavation steps 13 and 19 (e.g. B.1a and B.1b in case of parameter E_{50}) the TBM has not yet reached the considered section. In excavation step 25 (i.e. Figs. B.1c, B.2c, and B.3c) the TBM face reached the considered section and in the last two excavation steps 31 and 37, the TBM face has passed the section. Therefore, no model response is obtained in position marked in Fig. 4.13b in the middle of the tunnel, causing gap in the interpolated contour plot.

In Figs. B.4 to B.6 the same plots are obtained but, 9 m ahead in section four. Accordingly, the "tunnel gap" appears only in the last plot (Figs. B.4e, B.5e, and B.6e). The contents of the plots are not simply shifted by one slide, as the TBM-building interaction is different as discussed in the following

A holistic description of the obtained results is complex due to the vast amount of information. It can be observed that depending on the current position of the TBM and the position of the different sections, large differences in the patterns of S^*_T are on hand, but also the absolute values strongly differ, as they are all normalised by the maximum standard deviation of the current phase (not section!). For a better visualisation of the results, Fig. 4.14 exemplary displays the maximum value of $S^*_{\mathrm{T}_i}$ in the third section (highlighted in red in the figure) over the whole trajectory of the tunnel. The three curves in the figure show the maximum values in each of the five considered excavation steps for the different considered parameters, irrespective in which of the 99 points they occur. In this figure, it can be seen that in the early phases of the model evaluation the unloading-reloading stiffness is the most relevant parameter, while its relevance decreases as the TBM passes the considered section. Conversely, the friction angle becomes more and more relevant, while E^{ref}_{50} is of average relevance. By generating such a plot for each of the five considered measurement sections indicated in Fig. 4.13b, an optimal experimental design is derived. Therefore, the maximum $S^*_{\mathrm{T}_i}$ values are summed up in Tabs. 4.4 - 4.6, for each measurement section and the corresponding excavation phase. The five considered excavation

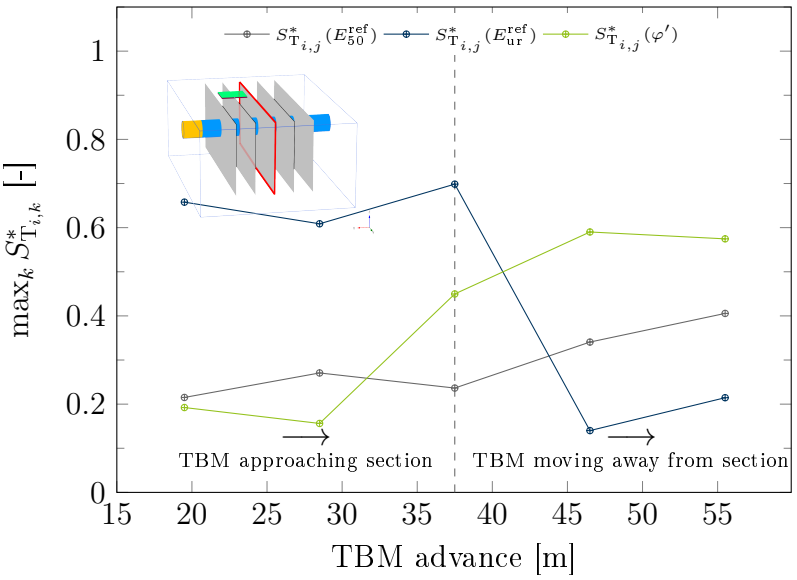


Figure 4.14.: Variation of maximum sensitivity index $S^*_{\mathrm{T}_{i,k}}$ in considered third monitoring section with tunnel advancement.

phases, i.e. 13, 19, 25, 31, 37 correspond to a TBM advancement of 19.5 m, 28.5 m, 37.5 m, 46.5 m, 55.5 m. As mentioned above, it is assumed that in each excavation step one measurement can be taken in each of the five sections, resulting in 25 measurements. The colours of the values in Tabs. 4.4 - 4.6 indicate the ranking among the three different parameters for each individual measurement value: green (most sensitive), blue (second), and black (least sensitive). It can be seen, that in none of the considered cases E^{ref}_{50} and accordingly $E^{\mathrm{ref}}_{\mathrm{oed}}$ is the most relevant parameter while $E^{\mathrm{ref}}_{\mathrm{ur}}$ shows the tendency to be most relevant in early phases and sections reached later by the TBM and hence, φ' is more relevant in the late phases and sections reached earlier by the TBM.
The established experimental design is based on the $S^*_{\mathrm{T}_i}$ values obtained in each of the five sections at the five considered excavation steps. Those locations within the grid of points shown in Fig. 4.13b that exhibit the highest $S^*_{\mathrm{T}_i}$ values are selected as sensor locations.
The obtained optimised experimental design is visualised exemplary for the first considered measurement time step that is excavation phase 13 in Fig. 4.15a, whereby the blue dots show the arrangement of measurement points after optimisation. The red dots correspond to an assumed "custom" experimental design where the sensors are all placed in the centreline over the tunnel at the ground surface and that is considered to compare the efficiency of the optimised experimental design. It is selected without any optimisation procedure, but just by assumption that measuring at the ground surface above the tunnel might be a reasonable and comprehensible approach. The positions in the custom experi-

Table 4.4.: Maximum $S^*_{T_i}$ values for E^{ref}_{50}.

	Excav. Phase				
Meas.Section	13	19	25	31	37
1st	0.17439	0.22959	0.45879	0.36235	0.44176
2nd	0.25779	0.19577	0.35357	0.32885	0.44189
3rd	0.21516	0.27069	0.23645	0.34069	0.40560
4th	0.20122	0.22714	0.28131	0.34455	0.33319
5th	0.20452	0.22690	0.24550	0.34436	0.30386

Table 4.5.: Maximum $S^*_{T_i}$ values for E^{ref}_{ur}.

	Excav. Phase				
Meas.Section	13	19	25	31	37
1st	0.57539	0.45610	0.27651	0.14093	0.19115
2nd	0.58753	0.59081	0.40197	0.13623	0.19667
3rd	0.65780	0.60872	0.69859	0.14006	0.21458
4th	0.67594	0.67491	0.58747	0.15788	0.44992
5th	0.72834	0.69258	0.67221	0.26799	0.64527

Table 4.6.: Maximum $S^*_{T_i}$ values for φ'.

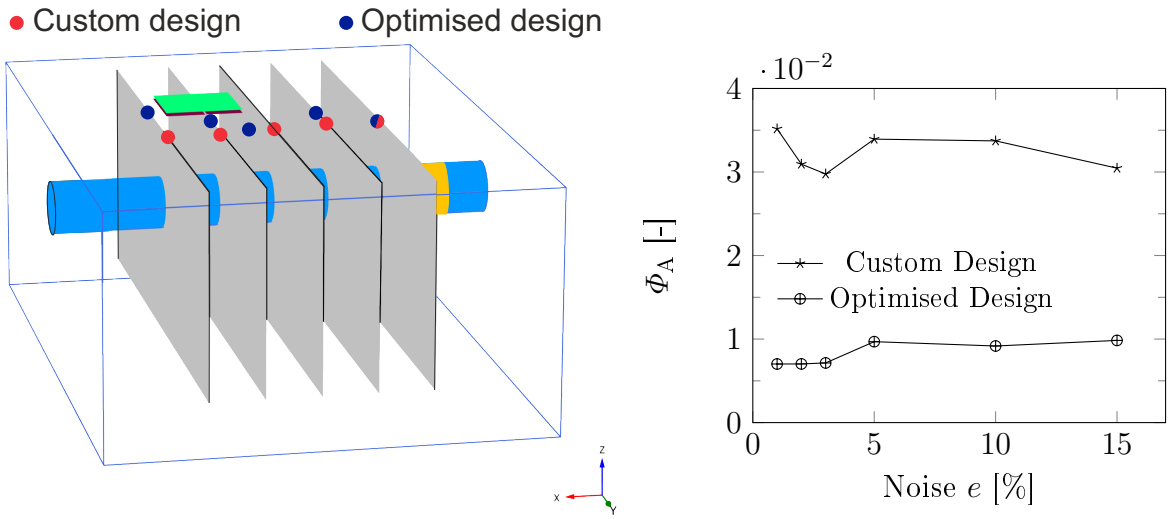
	Excavation	Phase			
Meas.Section	13	19	25	31	37
1st	0.24363	0.45795	0.58568	0.61499	0.59550
2nd	0.18577	0.24424	0.55092	0.60325	0.59387
3rd	0.19232	0.15625	0.44970	0.59017	0.57459
4th	0.18999	0.16974	0.17257	0.53811	0.45593
5th	0.19039	0.16216	0.16283	0.47334	0.22820

mental design are the same among all five considered excavation steps, while they change in each step in the optimised experimental design. In the first excavation step shown in Fig. 4.15a, the location of the sensor in the first section coincides with the custom experimental design, i.e. it is located at the ground surface in the tunnel centreline. In the second section, the optimal sensor location is still at the ground surface, but located with an offset of 8.5 m to the right-hand side (in direction of tunnel excavation). In the third section, the sensor should be located 17 m on right-hand side and 4.25 m below ground surface. The fourth sensor is in the same location as the second one with an offset of 8.5 m to the right-hand side, on the building's boundary. In the fifth section the sensor is located again at the ground surface, but this time 17.5 m right of the tunnel centreline.
As relevant as the position of the measurements are the weighting factors that are related to the data obtained in these points. In case of the blue optimised experimental design, the weighting factors w_i that are employed in Eq. 3.18, are assigned according to the values from Tabs. 4.4 to 4.6. Within the custom experimental design, as no more knowledge is available, each measurement information is weighted equally with 0.04 (= 1/25).
To evaluate the efficiency of the chosen experimental design, parameter identification is performed using both experimental designs. Hereby, artificial noisy data is generated that is supposed to reproduce in situ measurements that are affected with measurement noise, a procedure that is successfully employed in Lahmer (2011) and Schenkendorf et al. (2009) to test the efficiency of an experimental design. Doing so, the model results that are obtained using the mean values $\bar{\boldsymbol{\theta}}$ of the constitutive parameters given in Tab. 4.3 are employed. Once, the results are obtained for the custom experimental design, $\boldsymbol{y}_\mathrm{c} = f(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\delta}})$ and once for the optimised experimental design $\boldsymbol{y}_\mathrm{o} = f(\bar{\boldsymbol{\theta}}, \boldsymbol{\delta}^*)$. Both result vectors consisting of 25 elements are then falsified with artificial Gaußian white noise. As various types of noise can be considered, as discussed in Reichert et al. (2019), several of them are employed within this thesis. In the present case the noise is added according to Eq. 4.13.

$$\tilde{\boldsymbol{y}} = \boldsymbol{y} + \boldsymbol{y} \cdot \omega \cdot e, \; \omega \sim \mathcal{N}(0,1) \tag{4.13}$$

Herein, ω is a normally distributed zero-mean random number with standard deviation of one. e is a user-defined value that reflects the accuracy of the employed sensors. In the current example, values of $e = 0.01, 0.02, 0.03, 0.05, 0.10$, and 0.15 are selected. For each value of e, matrices of 100 x 25 elements of ω are generated and employed in Eq. 4.13 to obtain a reliable amount of back-calculation results. In the next step, each of the vectors of $\tilde{\boldsymbol{y}}$ with its 25 elements is back-calculated using the genetic algorithm described in Sec. 3.3.3.2. The algorithm is selected because of its generally accepted good convergence results. The higher calculation costs are accepted, as much less optimisation runs are re-

quired, compared to the applications considered in the following chapters.
Within each run of the optimisation algorithm, the defined objective is to minimise the discrepancy between model results $\boldsymbol{y}_{\mathrm{c}}(\boldsymbol{\theta})$ and $\boldsymbol{y}_{\mathrm{o}}(\boldsymbol{\theta})$, respectively, and the corresponding falsified values $\tilde{\boldsymbol{y}}$, by exploring the constitutive parameter space $\boldsymbol{\Theta}$, defined by the ranges given in Tab. 4.3. The parameter values are normalised by the boundary values of the given ranges to avoid undesired weighting of the optimisation results by having strongly different absolute values of the parameters. The hereby employed objective function is the one given in Eq. 3.18, using the weighting factors w_i as described above. By repeating this 100 times, two covariance matrices $\boldsymbol{C}_{\boldsymbol{\theta}}$ can be generated according to Sec. 2.4.2 that reflect the uncertainty by which the parameters are identified using the two different experimental designs. To evaluate $\boldsymbol{C}_{\boldsymbol{\theta}}$, the optimality criterion Φ_{A} is employed, i.e. the trace of each matrix is obtained. The results of this investigation are shown in Fig. 4.15b, where Φ_{A} is plotted over the applied noise ratio e for both experimental designs. It can be seen that over all considered values of e, the optimised experimental design produces much smaller values of Φ_{A} what indicates that employing the OED concept is highly beneficial to validate the constitutive parameters of the model. The non-consistent course, especially in case of the custom experimental design, is probably related to the small number of 100 samples of $\tilde{\boldsymbol{y}}$: using the custom experimental design, the response surface of the objective function is expected to be much flatter than in case of the optimised experimental design, wherefore outliers are more likely to happen and distorting the value of Φ_{A}.
When considering the presented results, one should be aware that the approach has its drawbacks. First, the suggested experimental design is technically feasible, but much more costly than the custom experimental design as its sensors are changing in positions and even go underground. Second, the suggested experimental design seems to be better than the custom one, but it should be tested by applying to some measurement data (real or artificial) to prove its efficiency compared to the custom experimental design or any other experimental design. Therefore, the interaction among the sensors and the impact of the measurement error should be considered what is done in the following. Nevertheless, when using the optimal experimental design and the corresponding weighting factors identified with this method, it is anyway beneficial to know where to place sensors, when to perform measurements, and to which objectives. A reduced review on this example can also be found in Hölter et al. (2018a). Finally, one should consider that in this academic example only vertical displacements in discrete points are considered for the sake of simplicity. As in the other examples shown in this thesis, considering displacement fields, lateral displacements or pore water pressure would probably allow much more efficient experimental designs, especially in case of combining the different type of measurement data.



(a) (b)

Figure 4.15.: a: Location of sensors in custom (red) and optimised (blue) experimental design at first considered excavation step and b: reliability of parameter identification for different designs and noise scenarios.

In Zhao et al. (2018), the concept of spatial GSA is employed in a similar manner to a 2D case of mechanised tunnelling. The employed model is based on data of a centrifuge test, introduced in Shahin et al. (2011, 2016) where a tunnel excavation close to a shallow foundation at the ground surface is simulated. The two considered experimental designs in this investigation are derived like those presented in this section: one based on GSA-plots and one without investigation by distribution of sensors at the ground surface as shown in Fig. 4.16. The comparison of the two experimental designs is done without adding further noise but by repeated back-calculation of the same data. Accordingly, only the uncertainty of the optimisation algorithms and the metamodel are influencing the variation of the identified parameters. Still, it could be shown that using the GSA-based experimental design, a more reliable identification of parameters is provided than using the intuitive sensor arrangement.

4.2.3. Dike under rapid water drawdown

Another common problem in geotechnical engineering is the safety of dams that are employed for water maintenance or as flood barriers in emergency cases. Failure of such structures can have dramatic impacts wherefore they are in focus of several research works (Alonso and Pinyol, 2016; Viratjandr and Michalowski, 2006). In these works it is

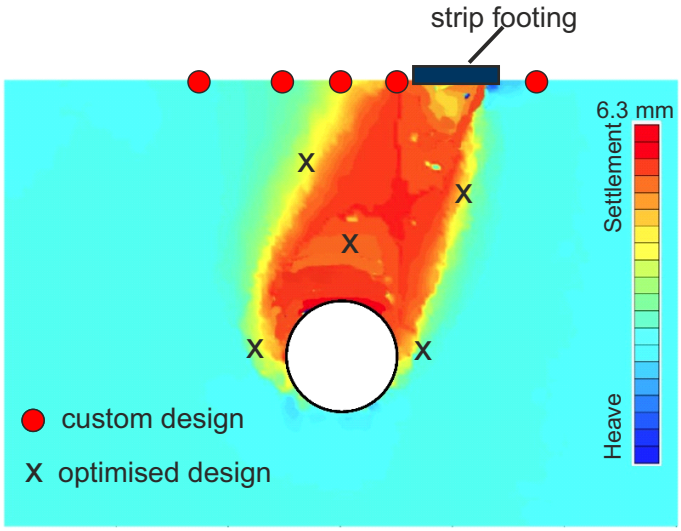


Figure 4.16.: Custom and optimised experimental designs for a centrifuge model test simulating a shallow tunnel excavation with a strip footing at the ground surface, in comparison to recorded field of vertical displacement (Zhao et al., 2018).

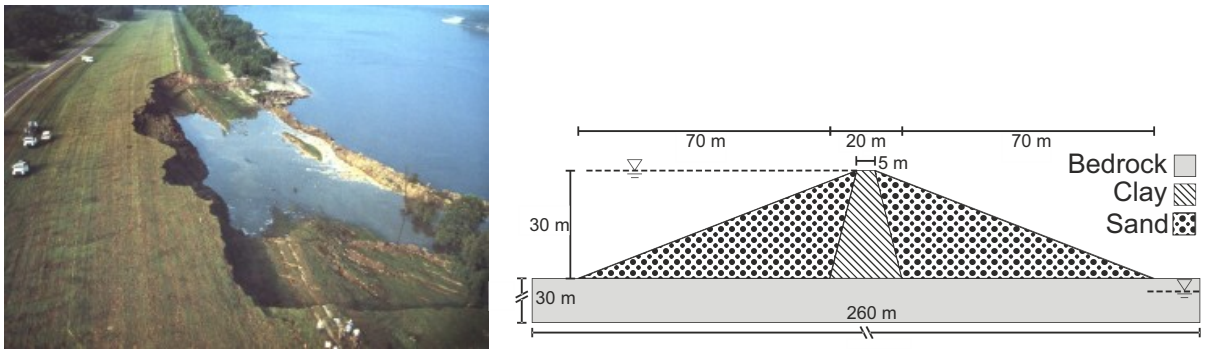
explained that the most critical state is not when the water level has reached its maximum, but when it rapidly decreases. This is due to the excess pore water pressure distribution within the soil body that will lead to an additional driving force on the water side and can lead to a slope failure as shown in Fig. 4.17a. To early predict a possible failure and to detect weak dike areas, extensive monitoring is often performed to ensure the stability of these structures. As discussed in chapter 1, considering just threshold values might not be as efficient as performing model validation using frequently updated measurement data. Here, the aspect of OED becomes relevant as many different options are possible w.r.t. position, type, and amount of sensors. Therefore, the problem of a dike or dam exposed to a rapid drawdown of the water level is selected as application example in this thesis.

4.2.3.1. Numerical simulation

To reproduce the behaviour induced by the decrease of the water level in a FE-model, a hydro-mechanically coupled analysis is required. Therefore, hydraulic boundary conditions must be provided additionally to the mechanical ones. The geometry of the generated 2D model is shown in Fig. 4.17b. The ground water head on the right-hand side of the model is assumed below ground surface. On the left-hand side, the water head is initially set to the height of the dam at 30 m, but is decreased in the following calculation steps to provoke the effects encountered in reality when the water level in rivers or reservoirs

rapidly decreases.

The geometry of the dam is given by a clay core with a width of 20 m at the ground level that narrows to 5 m at the dike crown. This clay body is surrounded by two inclined sand shoulders of 70 m length at the ground surface. As often encountered in dam constructions the clay core serves as hydraulic barrier, while the sand ensures stability of the construction. For both materials, the HS constitutive model is employed, using typical material properties. Among the parameters of this constitutive model that have been previously described in Sec. 4.2.2.2, the permeability coefficient k, the friction angle φ', the stiffness E_{50}^{ref}, and the cohesion c' are assumed to be most relevant in the intended simulation. Hereby, the cohesion is only applied to the clay and the stiffness E_{50}^{ref} is correlated to the other stiffnesses of the HS model E_{oed}^{ref} and E_{ur}^{ref}. For these seven parameters, reference values and bandwidths for the investigations performed in the following are given in Tab. 4.7, whereby the indices 1 and 2 refer to the sand and the clay properties, respectively. The bandwidths are selected quite narrow to reflect that in artificial structures the uncertainty should be lower than in case of natural soils. The underlying bedrock is modelled using a linear elastic constitutive model in manner that it has low influence on the behaviour of the dam, i.e. with low permeability and high stiffness. In the initial phase of the simulation with this model, the stress distribution induced by the own weight of the dam and the underlying soil is calculated, using the gravity loading method suggested by Plaxis2D in case of non-horizontal soil layers (the construction of the dam is not simulated). Next, the hydraulic boundary conditions are applied as mentioned before such that a constant seepage line of the water level is obtained. As third phase, drawdown of the water level is initiated. Therefore, the hydraulic boundary condition on the left-hand side is changed, corresponding to a reduction of the water level by a certain height, e.g. 5 m within five



(a) (b)

Figure 4.17.: Water retention structure a: structural failure after water drawdown (MEC International, LLC, 2020) and b: dimensions of generated model.

days. Starting from phase two, the simulation of water drawdown is repeated for 10, 15, 20, 25, and 30 m, each within 5 days. This phase is defined as "fully-coupled flow deformation analysis", i.e. a simultaneous calculation of deformations and pore water pressures is performed. Subsequent to the phases of water drawdown, a so-called safety analysis is performed. Hereby, the friction parameters φ' and c' (and dilatancy angle ψ' in case it would become larger than φ') of the soil are reduced proportionally until a failure of the structure is obtained. This method, also known as φ'-c reduction is available as own calculation type in the employed software. One should be aware that within the safety analysis, any constitutive model is reduced to the MC-model and the stress distribution is maintained constant. That means that the hydro-mechanical coupling is suspended wherefore employing safety analysis in combination with water flow is only recommended to a certain extend, but should be acceptable for this academic example. The obtained values of F_{SF} are plotted in Fig. 4.18 over the applied water drawdown. It can be seen that the more the water level is decreased, the smaller the safety factor becomes. It drops below one at approximately 20 m water drawdown, i.e. the failure occurs. From zero to five meter water drawdown, one can observe an increase in F_{SF}. An explanation can be given based on the field plots of deviatoric strain in Fig. 4.19. In the upper Fig. (4.19a), the water level is decreased by 5 m and the resulting failure mechnism results from the pressure of the free water of the reservoir pushing against the dike, causing failure to the "air" side. In case the water level decreases more, like in the case shown in Fig. (4.19b), the water pressure of the reservoir becomes less relevant and the seepage forces that cause a failure towards the "water" side, i.e. a landslide like the one shown in Fig. 4.17a, are decisive for the type of failure mechanism. The ratio between input parameter values and

Table 4.7.: Reference values and assumed ranges of variation of the uncertain HS–model parameters. The indices 1 and 2 refer to the sand and clay parts of the dike, respectively.

Parameter	Initial Value	Range	Unit
k_1	7.06 e-6	[7.06e-7, 7.06 e-5]	[m/s]
k_2	1.16 e-9	[1.16 e-10, 1.16 e-8]	[m/s]
φ'_1	35	[30, 40]	[°]
φ'_2	20	[16, 24]	[°]
$E^{\mathrm{ref}}_{50,1}$	20000	[16000, 24000]	[kPa]
$E^{\mathrm{ref}}_{50,2}$	15000	[12000, 18000]	[kPa]
c'_2	10	[8, 12]	[kPa]

their values at failure defines the so-called factor of safety F_{SF}:

$$F_{\mathrm{SF}} = \frac{\tan \varphi'_{\mathrm{input}}}{\tan \varphi'_{\mathrm{failure}}} = \frac{c'_{\mathrm{input}}}{c'_{\mathrm{failure}}} \tag{4.14}$$

This means that two different failure mechanisms take place depending on the amount of water drawdown. However, the mechanism shown in Fig. 4.19a only becomes controlling in case of 0 or 5 m water drawdown leading to a F_{SF} much higher than one and therefore being less relevant for the current study.

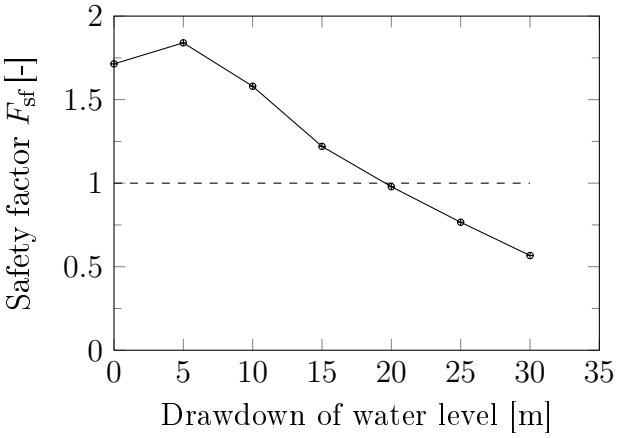


Figure 4.18.: Safety factor F_{SF} of dike, versus drawdown of water level.

4.2.3.2. Application of GSA-based OED

According to Fig. 4.18, it could be concluded that the dike is stable until the water level does not decrease more than 20 m within 5 days (In case the timespan is reduced, the allowable water drawdown height would accordingly decrease). However, these results rely on the deterministic model parameter values given in the second column of Tab. 4.7. As described in Sec. 2.3.3, one should be aware that the soil properties encountered in field might deviate from these values. Therefore, a measurement based model validation helps to assess the current state of the structure. To identify which type of measurement data would be most suitable for such validation and where it should be recorded, the OED concept of spatial GSA introduced in this chapter is applied to this example. It is defined to use sensors that record displacements in vertical and horizontal directions and also sensors to monitor pore water pressure that can be placed independently of each other.
As first step, a global sensitivity analysis is performed to identify which parameters are most relevant for the stability ot the dam under the considered impact. To do so, the

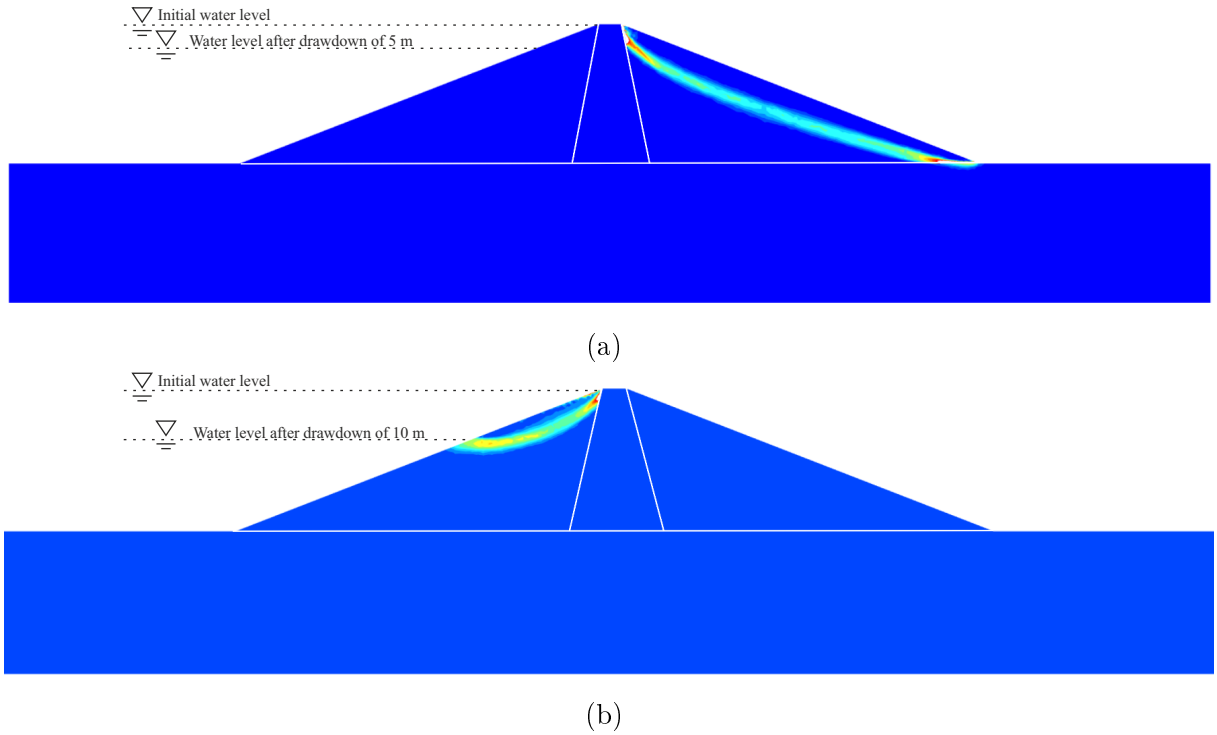


Figure 4.19.: Obtained failure mechanism visualised by the deviatoric strains in case of a: a 5 m water drawdown and b: a 10 m water drawdown.

procedure described in the previous sections of generating a metamodel and testing its accuracy is followed. The set of model parameter samples $\mathfrak{T}$ is generated within the ranges given in Tab. 4.7 and consists of 100 samples, complemented by a second set of 50 samples for testing the accuracy, whereby the considered model response is the safety factor of the dike. As was shown in Fig. 4.18, water drawdowns of 20 m or more are most critical for the stability of the dike, wherefore 20 m, 25 m and 30 m water drawdown are considered in the GSA. The GSA results are plotted in Fig. 4.20.

It can be seen that in general the permeability coefficient k_1 and the friction angle φ'_1 of the sand are most relevant, whereby the relevance is varying for different water drawdown values. The larger the water drawdown values get, the more the friction angles φ'_1 and φ'_2 gain in relevance. At the same time the permeabilities k_1 and k_2 decrease in relevance while stiffness and cohesion are generally of low relevance. Due to their comparably high impact, the two parameters φ'_1 and k_1 are placed in focus of the ongoing investigation.

One crucial aspect in this example is that the decisive model response, the factor of safety F_{SF}, is not a quantity that is measurable in-situ. It can only be obtained by running the numerical model. In contrast, the aforementioned displacements and the pore water pressure can be measured in both: simulation and in-situ. One should be aware about this

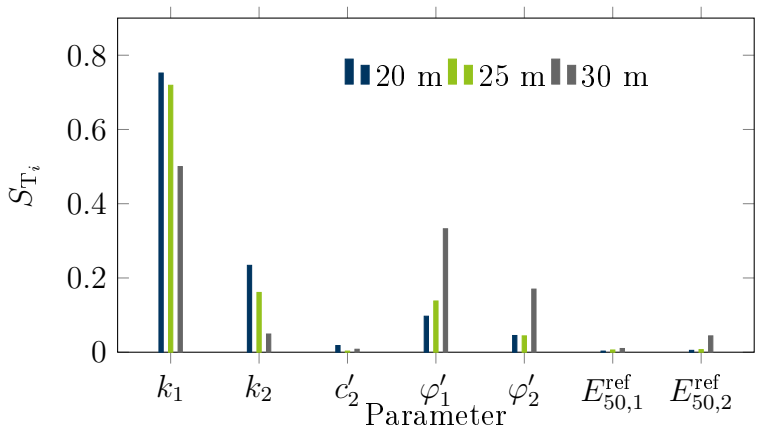


Figure 4.20.: Results of GSA that indicate the impact of the considered parameters on the safety factor of the dike in case of different water drawdown scenarios.

change in the considered model response for the further considerations. The model is run again in total 150 times to obtain a database to generate and test a metamodel, but this time the constitutive parameter space is limited to the two relevant ones, while the obtained model responses are now the horizontal displacement y_{x}, the vertical displacement y_{y} and the pore water pressure y_{w} instead of the safety factor. As a water drawdown of 30 m can be considered as the worst case, the corresponding model phase is selected as the one in which the data is extracted. However, if one would evaluate the data after the end of the water drawdown this would be too late, as failure would already have occurred. Therefore, the data is extracted after 2.5 days. That means half of the water drawdown has taken place, and possible countermeasures could be initiated in case a critical situation is identified based on the measurements.

To perform the spatial sensitivity analysis and calculate the modified sensitivity index $S^*_{\mathrm{T}_i}$, these model responses are obtained in 263 locations of the model geometry shown in Fig. 4.21. The marked points are distributed over the whole model domain, but densified on the left lower area as it is expected that this area might be most relevant. The results of $S^*_{\mathrm{T}i}$ for the three considered types of model outputs for the two parameters φ'_1 and k_1 are displayed as six contour plots in Tab. 4.8. As in the previous sections, dark blue areas correspond to a value of zero while the dark red zones corresponds to the highest relative value of one, whereby the relevant areas are encircled by a thin white line to indicate them. At first sight, one can conclude that the displacements have very little sensitivity to the permeability coefficient. The same applies to the pore water pressure's sensitivity to the friction angle. While this could have been expected, the locations of the sensitive areas are interesting. In case of the displacements, they are both located close to the left

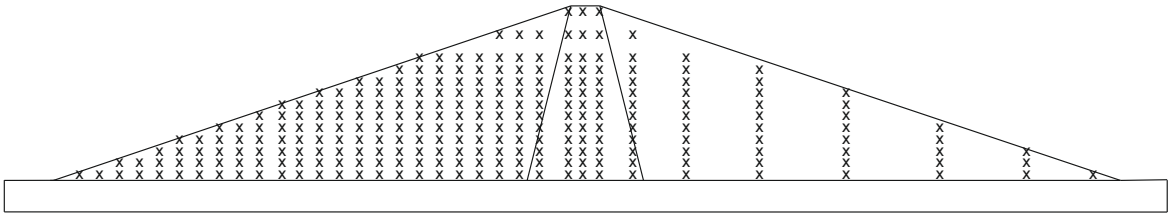


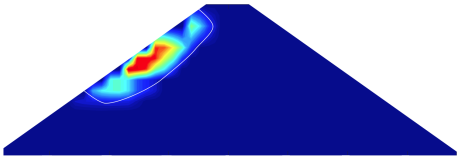
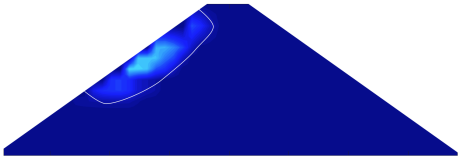
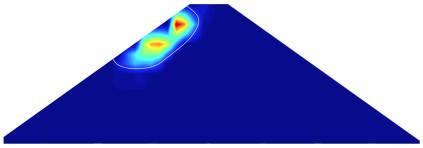
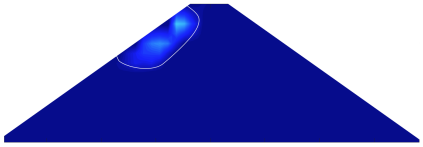
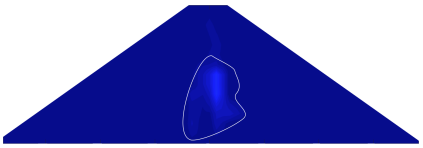
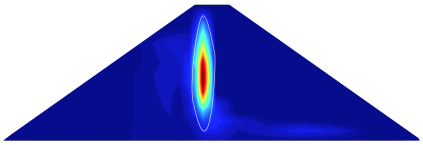
Figure 4.21.: Reference points for GSA distributed over the geometry of the dike model.

surface of the dike, in that area where the gradient of possible displacements is largest, but not the absolute displacement. In case of the pore water pressure measurements, the highest sensitivity is indicated on the left-hand side of the core of the dam. This might be surprising, as the preceding GSA has shown that not the permeability coefficient of the core but that of the surrounding sand is of interest. As any pore water pressure in the clay part should dissipate into the sand, the value of y_{w} in the clay will be a more accurate indicator for the permeability coefficient of the sand than pore water pressure measurements in the sand itself. Besides, as the permeability of the sand is very high, it might not be possible to perform accurate pore water pressure measurements there for practical reasons.
To sum up, it can be recommended to perform displacement measurements in the upper left side of the dike to identify φ_1' and to combine this with the measurement of pore water pressures in the left part of the clay core to identify k_1.
Combining both information provides the most reliable data for model validation. This exemplary application can be found in Hölter et al. (2019a), where the approach is employed for a preliminary experimental design that helps to identify areas where to place sensors to measure displacements and pore water pressures.

4.3. Summary

The variety of the three examples given in this chapter shows that the introduced approach of spatial global sensitivity analysis is not restricted to a single type of application. In the first example of an experimental device, the individual mathematical steps of the procedure are explained in detail, that is followed over the whole chapter. It is shown how the two most relevant parameters are identified and how both types of measurements are employed to identify these parameters most accurately. The limitations of the method are indeed reached when model error becomes larger than the parametric error as seems to be the case in that application. In the second example of 3D tunnel construction, only one

Table 4.8.: $S^*_{\mathrm{T}i}$ distributions for parameters φ'_1 and k_1 w.r.t y_{x}, y_{y}, and y_{w}.

Considered output	Sensitivity w.r.t. φ'_1	Sensitivity w.r.t. k_1
y_{x}		
y_{y}		
y_{w}		

type of measurements is considered, wherefore the specific location and timestep are of relevance in that application. The complexity of OED investigations that arises from the three dimensions of the problem and the five different considered time steps is pointed out. Here, an approach to validate the identified experimental design is applied using artificial noisy data and the optimality criterion Φ_{A} applied to the individual covariance matrix. The third example considers a dam subjected to a rapid drawdown of the water level. It shows how in-situ measurable model responses are influenced by different constitutive parameters that are most relevant for the stability of the structure what is the focus of interest. However, no knowledge is obtained regarding which sensor type provides more valuable information or how many sensors might be useful. To investigate these aspects, the following chapter provides a possible approach in which the number of sensors and their accuracy can be included.

5. The Bootstrap method in the concept of OED

5.1. Methodology

The approach to OED in geotechnical engineering presented in the previous sections reaches its limits when it comes to evaluation and comparison of specific experimental design scenarios and the aspect of including measurement error. In the present chapter, the concept of Bootstrapping is explained that is initially introduced as sampling method and that was applied to OED purposes in Schenkendorf et al. (2009). The method is applied exemplary to the case of the dike that was introduced in the previous chapter. As computational requirements are increasing when using this approach, the sigma-point method is applied to improve the calculation efficiency.

5.1.1. Introduction

The concept of Bootstrap is initially introduced in Efron (1979), where it is employed to allow performing statistical evaluations on data that would originally not be sufficiently large. Its denomination originates from the phrase "To pull yourself up by your bootstraps" that means to cope with a difficult situation without any external help as was demonstrated by Baron Munchausen who pulled himself and his horse out of the mud by his hair (Raspe, 1785). To a certain extend the same is done here: A sample of model responses $\boldsymbol{y} = (y_1, y_2, \ldots, y_N)$ is assumed to follow the random distribution $\hat{F}$. From this distribution, a Bootstrap sample $\mathbf{Y}^* = (Y_1^*, Y_2^*, \ldots, Y_N^*)$ is drawn with replacement with N indicating the size of the sample. The Bootstrap concept now consists of first generating a large number B of such samples $\mathbf{Y}^{*1}, \mathbf{Y}^{*2}, \ldots, \mathbf{Y}^{*B}$. From each of the B samples any statistics like the standard deviation σ, mean $\bar{\mathbf{Y}}$ or median $M(\mathbf{Y})$ can be obtained, but which might not reflect the correct properties of $\hat{F}$ due to the insufficient size of N. The

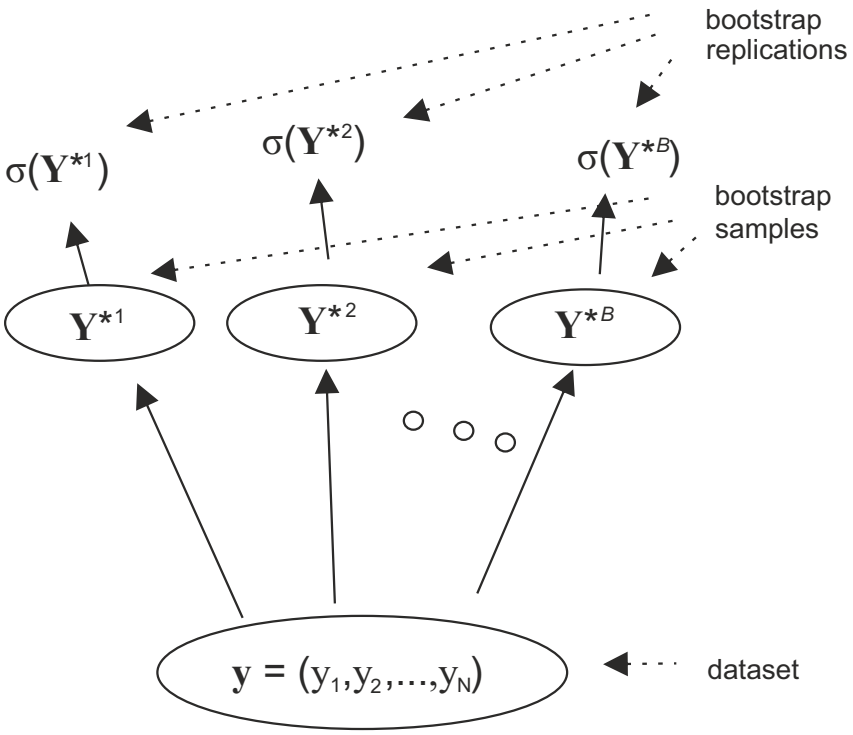


Figure 5.1.: Schematic of the bootstrap process for estimating the standard deviation of a statistic $\sigma(\mathbf{y})$, after Efron and Tibshirani (1993).

procedure is illustrated in Fig. 5.1. To obtain the standard deviation, Eq. 5.1 considers the median value of the B Bootstrap samples $M\left(\mathbf{Y}^{*b}\right)$.

$$\sigma_{boot} = \left\{ \sum_{b=1}^{B} \left[M\left(\mathbf{Y}^{*b}\right) - \bar{\mathbf{Y}}^{*} \right]^2 / (B-1) \right\}^{\frac{1}{2}} \tag{5.1}$$

Herein, $\bar{\mathbf{Y}}^{*} = \sum_{b=1}^{B} \mathbf{Y}^{*b}/B$ is the mean value of median values. In case B is sufficiently large, Eq. 5.1 will converge to the generally accepted formulation of the standard deviation σ defined by Eq. 5.2.

$$\sigma = \left\{ \sum_{i=1}^{n} \left[M\left(y_i\right) - \bar{\mathbf{y}} \right]^2 / (N-1) \right\}^{\frac{1}{2}} \tag{5.2}$$

whereby N has to be adequately large to ensure the representativeness of the sample for the given population.

By reusing initially sparse datasets, it is finally possible to let the statistics of the Bootstrap samples converge to the statistics of the unknown distribution $\hat{F}$ and to obtain meaningful and reliable results. Geotechnical engineering seems to be predestined for the employment of the Bootstrap method as here large uncertainties are encountered (see

Sec. 2.3.3) and coincidentally few measurement or sampling data is available due to high costs. However, in Bourdeau and Amundaray (2005) the authors had a large database of soil properties originating from standard penetration (SPT) and cone penetration tests (CPT) available and employed the Bootstrap method on parts of the data to prove its operability. Necessary sample size N and number of resamples B are investigated to identify which values are required for convergence of the results, but also the possible falsifying impact of outliers is illustrated. An application of the approach to a specific geotechnical problem is given in Most and Knabe (2010) where a reliability analysis of a shallow foundation is performed considering homogeneous and inhomogeneous model parameter uncertainties. A similar problem can be found in Li et al. (2015) where the problem of slope stability is considered. The Bootstrap approach is employed therein to identify which type of probability distribution function is able to best fit the available data. The large database allows to investigate the influence of both, uncertain constitutive parameters and uncertain distribution parameters, on the factor of safety F_{sf}. In Luo et al. (2013), a further typical problem from geotechnical engineering is considered, i.e. the wall deflection and surface settlements in case of a braced excavation. As performed in the present thesis, from the numerous existing parameters, the attention is concentrated on those few that are identified as being most relevant for the model responses of interest. The authors intend to provide a confidence interval of the probability of exceedance of one of the model responses. As typical in geotechnical engineering, they indicate that the amount of available measurement data is too small to derive accurate statistics. In the considered case of the Taipei National Enterprise Center Excavation, the tangent modulus and undrained shear strength obtained from triaxial tests on ten disturbed and 7 undisturbed clay samples are employed as dataset. According to Eq. 5.1, the mean and standard deviation of the assumed normal distribution are obtained using $B = 10000$ Bootstrap samples. With these statistical characteristics, the reliability index of the possible failure is calculated using a semi-empirical analytical model for estimation of displacements in braced excavations named KJHH model.

Of course, the Bootstrap approach can also be applied in other types of research as shown in Joshi et al. (2006) for a problem from the field of system biology. Here, kinetic parameters of ordinary differential equations that describe processes of metabolisms are in focus. As only few measurements are available in the described research, the Bootstrap resampling enabled to obtain sufficiently large datasets to precisely define the confidence intervals of the identified parameters. The question of confidence intervals of parameters identified from measured data now inevitably leads to the aspect of OED and how these intervals can be narrowed as will be considered in the following.

5.1.2. Bootstrap in the context of OED

The problem of parameter identification in biological systems discussed in Joshi et al. (2006) is reconsidered again in Schenkendorf et al. (2009), but with the focus on OED applications. The Bootstrap approach is employed therein in particular to determine the covariance matrix $\boldsymbol{C_\theta}$ that is introduced in Sec. 2.4.2 and to which one of the cost functions or optimality criteria $\Phi_i\left(\boldsymbol{C_\theta}\right)$ is applied that are introduced in Sec. 2.4.3. The authors generate at first artificial noisy measurement data. The measurable experimental result and therefore, the considered model response is the concentration of biomass in a growth reactor that will depend on the two uncertain parameters of maximum growth rate and the substrate affinity. In the next step, these two parameters are back calculated using the falsified values of the biomass concentration. The experimental design parameter of the experiment that is supposed to be optimised is the time dependent inlet flow of the reactor. Therefore, the three variables that define the linear equation of the inlet flow are varied to find the minimum of $\Phi_{E^*}\left(\boldsymbol{C_\theta}\right)$. As performing the repeated runs of the Bootstrap method would be too time consuming, the authors settle to describe it conceptually and show the result of applying the FIM (s. Sec. 2.2.1) and the Sigma Point Method that will be introduced later on in this chapter.

5.2. Application to geotechnical engineering problem

In the following, the concept of Bootstrap is applied to a problem of geotechnical engineering. As described in Schenkendorf et al. (2009), the computational requirements are a limiting factor of this method as a large number of parameter identification runs are needed. Therefore, it is conceivable that this method should not be employed as a single option, but probably as refinement tool in an already restricted design parameter space. This approach is followed herein where the results of Sec. 4.2.3 are employed as starting point for applying the Bootstrap method. In case the design space is smaller, fewer candidates points must be considered in the possible experimental design set ups.
The number of sensors and of candidate points significantly influences the computational requirements as shown by Eq. 5.3

$$t = \binom{P}{K}^J \quad , \tag{5.3}$$

where t is the number of possible sensor combinations, P is the number of sensors employed to measure each of the J types of measurement in the corresponding K candidate locations. To each of the selected candidate points, a set of N output values is required. This leads to a matrix $\tilde{\mathfrak{Y}}$ of dimensions N x $J \cdot K$. The elements of this matrix can be artificially generated by adding random noise to model results as performed in the example to be presented, but for real case application, the Bootstrap concept should be employed to enrich measurement data that will most probably be quite sparse and not available for all positions of interest. In the next step, back calculation of the model parameters of interest is undertaken. To include the aspect of sensor position interaction, always two sensors, one of each kind of measurements (one for displacement and one for pore water pressure), are employed to back calculate one of the parameters of interest. This laborious procedure allows a reduction of computational effort in the following steps. Its outcome is the matrix of back calculated parameters $\tilde{\mathfrak{T}}$ given in Eq. 5.4.

$$\tilde{\mathfrak{T}} = \begin{pmatrix} \tilde{\theta}_{1,1,1} & \cdots & \tilde{\theta}_{1,K^J,1} & \tilde{\theta}_{2,1,1} & \cdots & \tilde{\theta}_{s,K^J,1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \tilde{\theta}_{1,1,N} & \cdots & \tilde{\theta}_{1,K^J,N} & \tilde{\theta}_{2,1,N} & \cdots & \tilde{\theta}_{s,K^J,N} \end{pmatrix} \tag{5.4}$$

Calculating all the $N \cdot K^J$ elements of this matrix is the crucial because most time consuming part of the evaluation. However, afterwards no more back calculation runs are necessary. From the matrix $\tilde{\mathfrak{T}}$, the two column matrix $\tilde{\boldsymbol{\theta}}_r$ of identified parameters is extracted for each possible experimental design. Hereby, the $N \cdot P^J$ elements of each column of $\tilde{\boldsymbol{\theta}}_r$ are constituted of the columns of $\tilde{\mathfrak{T}}$ that correspond to the current r^{th} sensor arrangements.

$$\tilde{\boldsymbol{\theta}}_r = \begin{pmatrix} \tilde{\boldsymbol{\theta}}_{r,1} \\ \vdots \\ \tilde{\boldsymbol{\theta}}_{r,B} \end{pmatrix}, \quad r \in \{1, ..., t\} \tag{5.5}$$

As the error terms that were added to $\tilde{\mathfrak{Y}}$ are randomly generated, calculating the mean values $\overline{\boldsymbol{\theta}}$ of the parameters of interest according to Eq. 5.6 provides a reliable estimate of this value.

$$\overline{\boldsymbol{\theta}} = \frac{1}{B} \sum_{i=1}^{B} \tilde{\boldsymbol{\theta}}_{r,i} \tag{5.6}$$

Knowing the parameter values $\tilde{\boldsymbol{\theta}}_r$ that correspond to the back calculation of the different experimental designs and their means $\overline{\boldsymbol{\theta}}$, the covariance matrices originating from Bootstrap data $\boldsymbol{C}_{\boldsymbol{\theta}_i,\text{Boot}}$ can be calculated for every considered experimental design as described

in Eq. 5.7.

$$\boldsymbol{C}_{\boldsymbol{\theta}_i,\mathrm{Boot}} = \frac{1}{B-1}\sum_{i=1}^{B}(\tilde{\boldsymbol{\theta}}_{r,i} - \overline{\boldsymbol{\theta}})(\tilde{\boldsymbol{\theta}}_{r,i} - \overline{\boldsymbol{\theta}})^T \tag{5.7}$$

To these matrices, an optimality criterion Φ_i is applied and finally, all experimental designs are ranked according to this criteria, indicating the optimal experimental design by the smallest value.

5.2.1. Example introduction

The same example of a dam under rapid water drawdown as in Sec. 4.2.3 shall be employed to illustrate the application of the Bootstrap method. Using a GSA, it was identified that the friction angle φ'_1 and the permeability coefficient k_1 of the sand parts are most influencing the stability of the considered dam, wherefore these are in focus of the OED application. The objective is now to find a solution for the problem where to place sensors for measuring displacement and pore water pressure. This was also investigated in Sec. 4.2.3 and the distributions of $S^*_{\mathrm{T}i}$ obtained from the spatially distributed GSAs indicate two relevant areas. Tab. 4.8 shows that displacements should be measured on the left side of the dam close to the ground surface, while pore water pressures should be measured on the left side of the clay core. These results shall be employed in the current investigation to reduce the required calculation runs. Accordingly, the design space $\boldsymbol{\Pi}$ that would have been the whole model geometry is restricted to those areas with highest value of $S^*_{\mathrm{T}i}$. Here, three sensors of each type shall be placed in a manner to reduce as much as possible the constitutive parameter uncertainty. The sensors can be placed in discrete positions what defines the number of possible experimental designs and prevents a possible clustering of the sensors. This problem might occur in case one single area is identified as most promising for a certain measurement. An unconstraint search algorithm might then converge to a solution where all sensors should be placed at the same position as observed and discussed in Lahmer (2011). Such clustering should be avoided as it is hardly applicable in praxis and could overrate impacts from local inhomogeneities.

5.2.2. Application

Within each of the high sensitivity areas, eleven positions are defined as displayed in Fig. 5.2. Here, the points 1 to 11 are candidate points for the displacement sensors and

12 to 22 are possible positions for the pore water pressure sensors. It is assumed that horizontal and vertical displacements are recorded by one sensor at a time. The sensor positions 11 and 17 are control points to test whether the obtained results are reasonable. As these points are outside the indicated sensitivity areas, they should not be identified as candidates of an experimental design. In a first investigation, it is assumed that three sensors are employed for each type of measurements. Within the nomenclature introduced in Sec. 5.2, there is in the present example $K = 11$, $P = 3$, and $J = 2$. According to Eq. 5.3, this implies that there are $t = 27225$ possible combinations to arrange the alltogether six scheduled sensors. The OED task is accordingly to find out which combination of the sensors allows the most reliable identification of the constitutive parameters of interest φ_1' and k_1.

As the considered candidate points shown in Fig. 5.2 are among those shown in Fig. 4.21, the data generated in Sec. 4.2.3.2 can be reused and no additional forward calculations of the FE model are necessary. However, immediate reuse of the output data is not possible for the simple reason that by adding noise to the model response, one might obtain values of $\tilde{y}_{j_k}$ that can not be back calculated within the given boundaries of the parameter space $\boldsymbol{\Theta}$. Instead, the parameter identification would converge to values on the edge of this space, leading to unusable data.

The overall concept introduced in the previous section, that is illustrated in Fig. 5.3, consists of first generating artificial noisy data for each of the K candidate points. Like in Schenkendorf et al. (2009), the employed data is artificially generated as a synthetic example is considered. 150 new random input samples are generated within a restrained input space that are transferred to output samples for each of the 22 positions using the previously generated metamodel. These results are falsified using Gaußian white noise as described by Eq. 5.8

$$\tilde{y}_{j_k} = y_{j_k} + y_{j_k} \cdot e \cdot \omega \ , \ \omega \sim \mathcal{N}(0,1), \tag{5.8}$$

whereby j refers to the two different considered output types and k to the eleven different positions. The amount of noise is defined by e that is set to 0.1. To visualise the impact of adding this noise, Fig. 5.4 displays exemplary FE model results and noisy data of different exemplary positions and output types. The green dots correspond to the 150 model results that are obtained for randomly generated input parameters, while the blue dots were generated by applying Eq. 5.8. These blue dots shown in Figs. 5.4a & 5.4b represent four out of the 33 columns of the matrix $\tilde{\mathfrak{Y}}$. To obtain the 242-column matrix $\tilde{\mathfrak{T}}$, the noisy data is back calculated as described above using one displacement and one pore water pressure

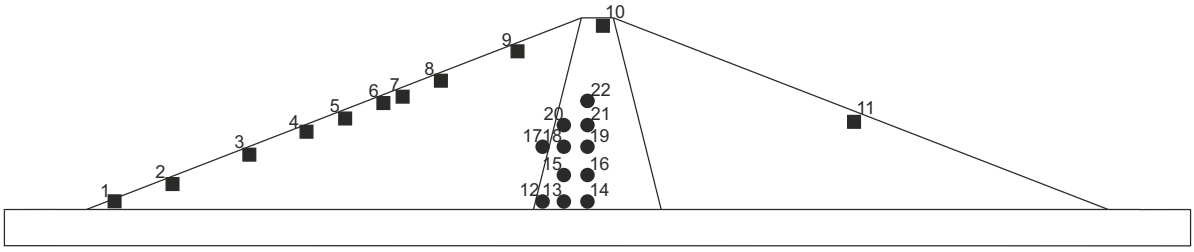


Figure 5.2.: Location of the 22 points considered in OED process using the Bootstrap approach (Squares for displacements & circles for pore water pressure).

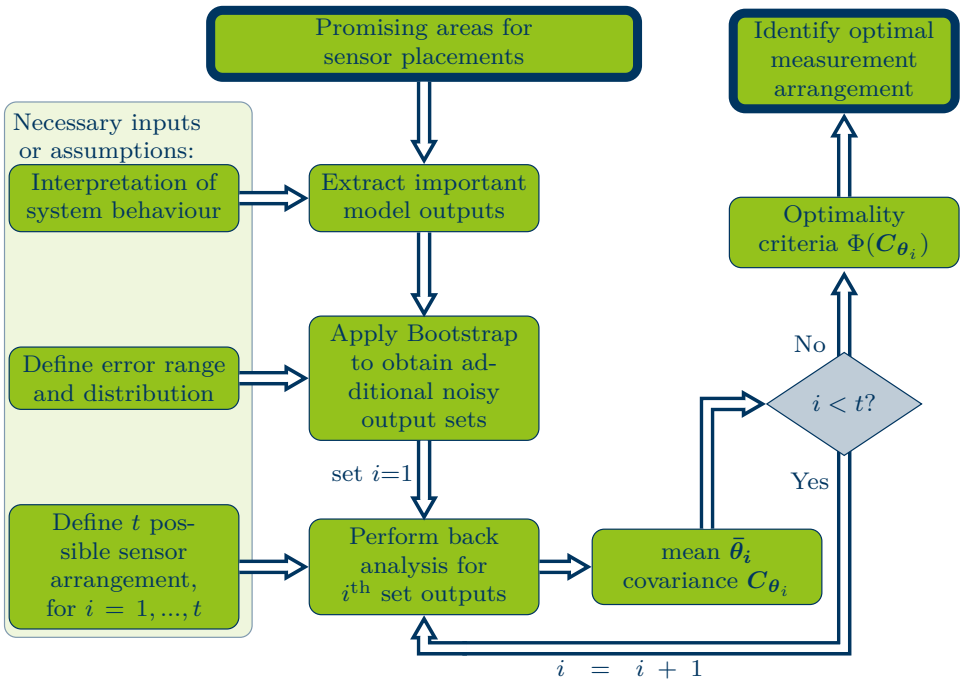


Figure 5.3.: Flowchart of OED procedure using Bootstrap approach in the intended context of geotechnical applications.

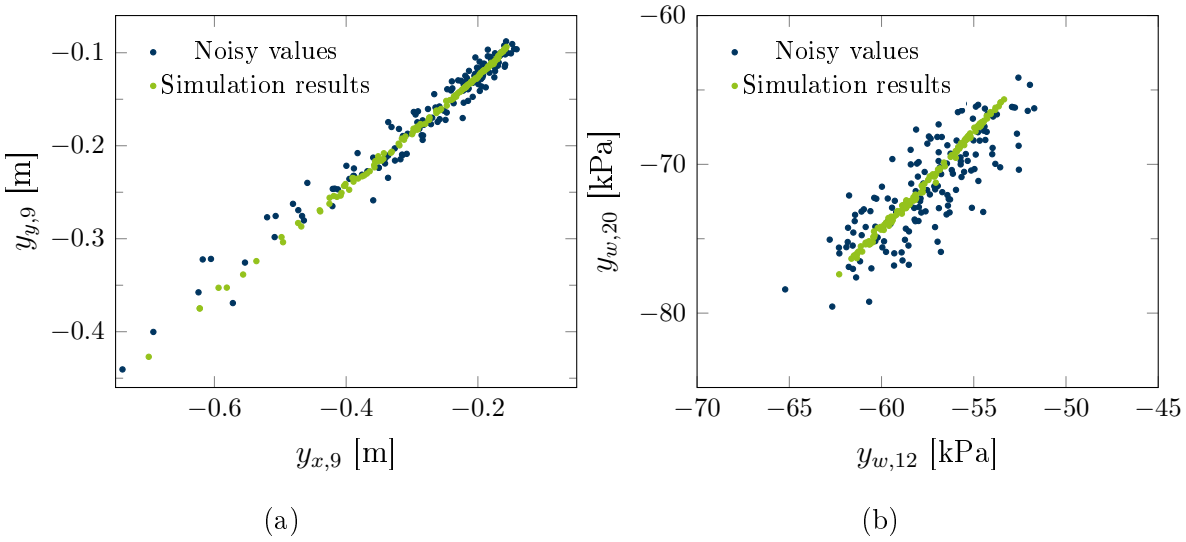


Figure 5.4.: Output values from simulation compared with systematically falsified values a: horizontal and vertical displacement of same position ($\boldsymbol{y}_{x,9}$ & $\boldsymbol{y}_{y,9}$) and ($\tilde{\boldsymbol{y}}_{x,9}$ & $\tilde{\boldsymbol{y}}_{y,9}$) b: pore water pressure of two different positions ($\boldsymbol{y}_{w,12}$ & $\boldsymbol{y}_{w,20}$) and ($\tilde{\boldsymbol{y}}_{w,12}$ & $\tilde{\boldsymbol{y}}_{w,20}$).

sensor at a time. As hereby 18150 ($= N \cdot K^J$) processes of back calculation are necessary to identify each entry of $\tilde{\mathfrak{T}}$, time efficiency is of high relevance. To account for this aspect, the DE-algorithm is employed that is implemented in the package **DEoptim** Mullen et al. (2011) for the statistical computing language **R** (e.g. Tippmann, 2015) and that is described in Sec. 3.3.3.1. As optimisation parameters, the population of one generation is assigned to 1000, a maximum of 200 iteration steps is defined, and a crossover probability of 0.5 is selected. Performing all parameter identifications took around five days using a PC with Intel i7-4790 3.6 GHz processor.

To each experimental design, starting with the positions (1, 2, 3, 12, 13, 14) and ending with (9, 10, 11, 20, 21, 22), the corresponding 18 ($=P^J \cdot s$) columns of $\tilde{\mathfrak{T}}$ are merged to build the vector $\tilde{\boldsymbol{\theta}}_r$. Using Eq. 5.6 and Eq. 5.7 to identify the statistics of interest that are mean and covariance matrix of the parameters is effectuated as next step. As B equals to 1350 for each experimental design, a reliable estimation of both statistics of interest is obtained. To visualise the convergence behaviour, Figs. 5.5a & 5.5b are displayed. It can be seen that in case of the 1350 parameter values, a result is obtained that is quite close to the model mean values of $\varphi' = 35°$ and $k = 7.06 \cdot 10^{-6}$ m/s given in Tab. 4.7. Besides, in the employed logarithmic scale it is clarified that using only 150 runs of back calculation per individual sensor data would not have been sufficient.

Doing so, for each of the t possible combinations one covariance matrix $\boldsymbol{C}_{\boldsymbol{\theta}_i,\mathrm{Boot}}$ is obtained to which the optimality criteria Φ_A and Φ_D are both applied that are the trace and

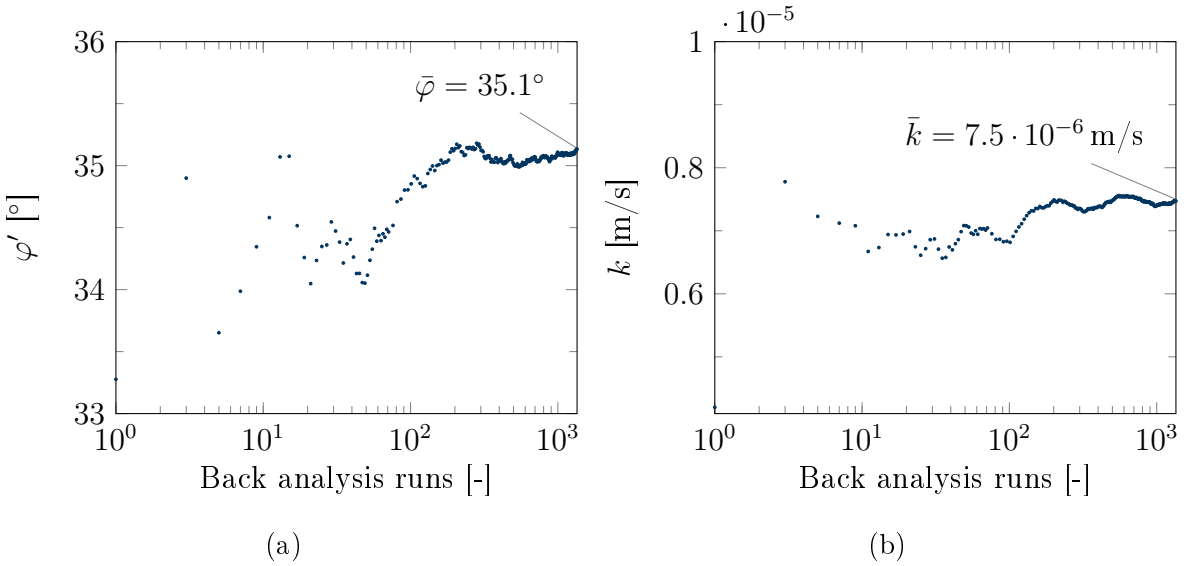


Figure 5.5.: Convergence behaviour of the identified parameters φ' (a) and k (b) towards the actual values over the 1350 back-calculated values.

the determinant, respectively. As now for each combination of sensors (i.e. experimental design) one specific value is available, a ranking can be performed that is shown for the best and worst two experimental designs in Tabs. 5.1 & 5.2. As discussed in Sec. 2.4.3 it is difficult to define in advance which optimality criterion is most suitable, but comparing both tables brings additional insights. First, one can see that the experimental design ranked best is in both cases the same, using the positions 4, 5, 6, 19, 21, and 22 what confirms that the method leads to consistent results. That fits to the circumstance that the second best experimental design is not a totally different one, but differs only by one position. However, the second best experimental design is not the same using both criteria: compared to the best-ranked experimental design, sensor position 6 is replaced by 7 in case of A and by 9 in case of D. More interesting might be a view on the opposite end of the ranking. Here, quite different sensor arrangements are suggested by both criteria. However, the two "control points" 11 and 17 can be considered here, as they are expected to be part of the worst experimental designs. In contrast to criterion A, in case of criterion D both points can be found in the worst experimental designs. This would indicate that overall the Φ_{D}-criterion is the better choice in case of this application.

5.2.3. Variation on sensor types

Unlike the concept of spatial sensitivity distribution introduced in Chapter 4, using the Bootstrap based approach allows to explicitly compare specific experimental designs and

Table 5.1.: Ranking the individual sensor positions – optimality criterion Φ_{A}.

Sens. Pos. y_y/y_x	Sens. Pos. y_w	Φ_{A}	Rank
[4 5 6]	[19 21 22]	2.404 e-2	1
[4 5 7]	[19 21 22]	2.406 e-2	2
[2 3 6]	[12 13 15]	4.607 e-2	27224
[1 2 3]	[12 13 15]	4.608 e-2	27225

Table 5.2.: Ranking the individual sensor positions – optimality criterion Φ_{D}.

Sens. Pos. y_y/y_x	Sens. Pos. y_w	Φ_{D}	Rank
[4 5 6]	[19 21 22]	2.493 e-5	1
[4 5 9]	[19 21 22]	2.535 e-5	2
[1 2 11]	[14 17 20]	1.4192 e-4	27224
[1 9 11]	[14 16 17]	1.4286 e-4	27225

thereby find the optimum one. To do so, the experimental design parameters like amount, accuracy, or type of sensors can be varied to see what the impact would be. As consequently, infinite possibilities of variation exist, in the following, the case of varying the amount of sensors is exemplary presented. In the previous section, it is stated that three sensors should be employed for each output type, the pore water pressure and the displacements. However, the number of three is actually an arbitrary value. One should consider that every sensor causes costs, in monetary way but also in the sense that it will generate data causing additional computational costs in their evaluation. Therefore, the evaluation performed in the previous section is repeated by changing the number of sensors, starting by using only one of each kind and then stepwise increasing to three. The obtained results are displayed in Fig. 5.6 by means of box plots, for each of the nine considered sensor combinations. To visualise the complexity of the OED task, the number of the corresponding possible arrangements t is shown by the green line. The box plot visualisation is selected to provide an impression on the reliability of the different scenarios. Therefore, the whiskers of each box plot represent the bandwidth of the 50 best results. The value of 50 is arbitrary chosen, but allows to comprehend what happens in case not the best scenario is selected, but a close by one, what could happen e.g. if the installation process is not carefully performed or the model has some inaccuracies. As could be expected, using only one of each sensor type results in the highest value of the optimality criterion

and in the widest spread of the box plot, i.e. the least reliable parameter set is obtained, while the opposite is the case for using three sensors of each type. The combinations in between provide more interesting information. It can be seen that adding one pore water sensor or one displacements sensor (scenarios 2 & 3) results in the almost same optimal (minimum) value of Φ_D, but adding one more pore water pressure sensor results in much larger box plot bandwidth, wherefore adding one displacement sensor would be preferable. This effect is even more pronounced in the two next scenarios (i.e. having 1-3 and 3-1 sensors). Here, it can be seen that having three displacement and one pore water pressure sensor results into a smaller box plot bandwidth but also a smaller value of Φ_D, even compared to the following scenario 2-2. Subsequently, the more sensors are employed, the more Φ_D and its bandwidth are reduced. However, as between the eighth and the ninth scenario (3-2 and 3-3), only marginal improvement is obtained, no further scenarios are investigated. This demonstrates that adding more and more sensors cannot continuously improve the process of model validation. Finding out how extensive a monitoring program needs to be at maximum could accordingly save a lot of costs by reducing the number of devices.

In Hölter et al. (2019a), a reduced description of this example is given. There, it is presented in combination with the method that was described in Sec. 4.2.3 where a preliminary investigation was performed. One should be aware that to identify a unique solution as optimal sensor arrangement, the method employed here required extensive computational

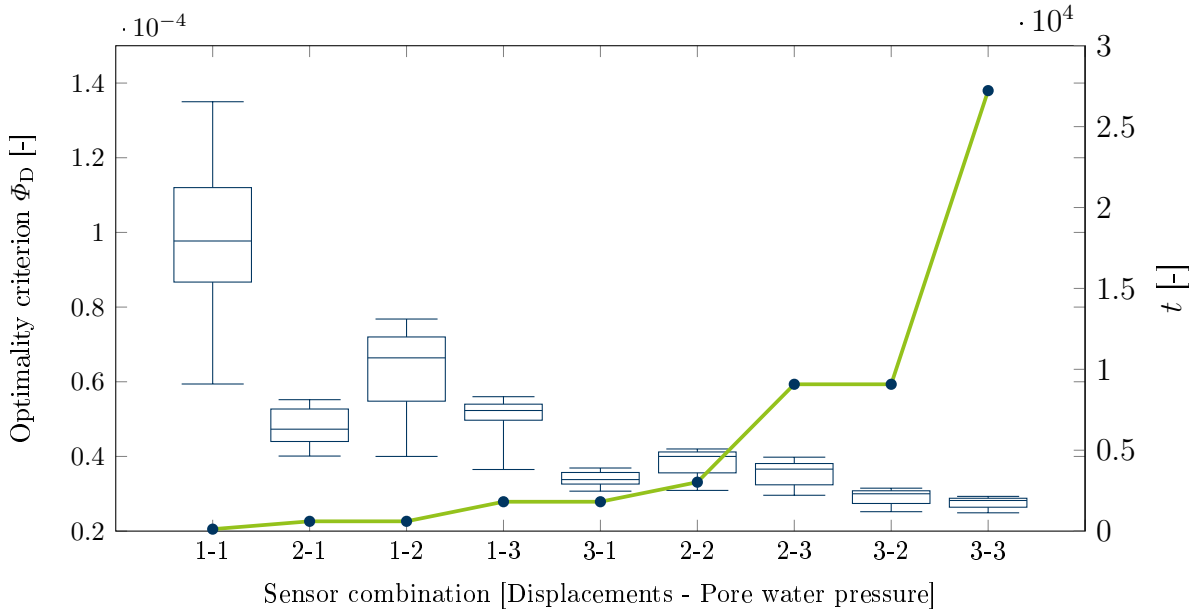


Figure 5.6.: Comparison of the number of sensors (boxplots), using D-optimality criterion compared to possible number t of arrangement combinations (blue dots with green line).

effort. In case the design space would have included the whole model domain and the candidate points would have been distributed with the same density as in the performed evaluation, it would not have been possible to perform the evaluation as the computational requirements would not have been bearable.
This is the case even though the presented method simplifies the process of back analysis by obtaining the results $\tilde{\boldsymbol{\theta}}_r$ of the current experimental design from the matrix $\tilde{\mathfrak{T}}$ instead of repeating the calculation each time. Using this approach as refinement of the previous GSA results makes this approach practical. To further reduce the computational costs, in the following section an approach is presented to substitute the numerous metamodel calls caused by MC sampling.

5.3. Extension with sigma points

5.3.1. Concept

The previously introduced Bootstrap approach obtains precise results by extensive resampling of data until an accurate distribution of that data is reached from which the statistics of interest can be derived. This is not a problem in the references given in Sec. 5.1.1 as generating new samples can be performed almost instantly. In the context of OED indeed, the advantage of obtaining a large database becomes also the drawback, because the more output data is generated, the more time-intensive back analysis runs should be performed. The sigma point method introduced in Julier and Uhlmann (1996) provides an approach to solve this problem. The general idea of transferring the input distribution by performing the model processes to an output distribution is maintained similar, but instead of approximating the inputs by a random MC process, deterministic samples are generated. The intention of the authors is to provide an advanced filtering algorithm, especially more suitable for non-linear problems than the Kalman filter, as instead of linearising the transformation process, i.e. the simulation model, a limited number of model runs is required.
In the original publication, it is employed to predict the future state of a system based on current observations under consideration of both, uncertain model inputs and outputs. The method can be employed in case the distribution of inputs is known, the transformation is known, but the distribution of the model outputs is unknown, what is exactly the case in the problems described in this thesis. The terminology must be employed carefully in this context, as the "transformation" as called in Julier and Uhlmann (1996) is in the

present application the back calculation of model parameters. That means the terms of input and output are employed here opposite to the initial publication.
The data contained in the vector of noisy measurement data $\tilde{\boldsymbol{y}}$, or matrix $\tilde{\mathfrak{Y}}$, can be described by its mean and standard deviation $\bar{y}$ and σ_y, or $\bar{\mathbf{y}}$ and $\boldsymbol{\sigma}_y$, respectively and the minimum required number of information to describe a normal distribution in an J-dimensional output space accounts to $2J+1$.
Accordingly, the simple concept of the sigma point method is to employ exactly $2J+1$ deterministic values to describe the variance of the considered distribution instead of approximating the distribution by random sampling. These values are the so called sigma points for which the vector of disturbances $\boldsymbol{\sigma}$ is defined according to Eq. 5.9 and that are added to the mean value to obtain the distribution of interest.

$$\boldsymbol{\sigma} = \pm\sqrt{(J+\kappa_\mathrm{s})\,\boldsymbol{C_y}} \tag{5.9}$$

Herein, $\boldsymbol{C_y}$ is the covariance matrix of the J different considered output types and κ_s is a factor to enhance the accuracy of the approach. The higher the value of κ_s, the higher the order of the error terms that can be considered using the sigma points approach, but also the more sigma points are required. As in the present work only Gaußian white noise is considered, the value of κ_s is set to zero. In Schenkendorf (2014) indeed, application examples are given using sigma points with several different types of noises and distributions, including increasing values of κ_s and according weighting factors. Within the framework of sigma points, any of the possibilities to extract a square root of the matrix in Eq. 5.9 is permitted. Afterwards, the matrix of $\boldsymbol{\sigma}$ is added to the vector of mean values. To obtain the unknown statistics of the distribution of interest, the "transformation process" (i.e. in the present application the parameter identification) is performed and as many response values are obtained as sigma point samples were generated. In Julier and Uhlmann (1996), several theoretical application examples are given in which it is proven that the obtained mean values using sigma points are equal to the correct means and that the results are independent of κ_s. This process is schematically illustrated in Fig. 5.7. From the obtained response values, that are symbolised by the ellipse on the left side, the mean and variance can be calculated according to Eq. 5.10 that are same as from a whole distribution due to the considered selection in the sampling process.

$$\boldsymbol{C}_{\boldsymbol{\theta},SP} = \frac{1}{2(J+\kappa_\mathrm{s})}\sum_{i=1}^{2J} w_i(\tilde{\boldsymbol{\theta}}_i - \overline{\boldsymbol{\theta}})(\tilde{\boldsymbol{\theta}}_i - \overline{\boldsymbol{\theta}})^T \tag{5.10}$$

The weighting factor w_i depends on J, but essentially also on the type of error and should be adjusted depending on the current application as explained in Julier and Uhlmann

(1996).
Using the sigma points approach, several applications were performed in recent years that show the utility of this method. In Azam et al. (2012), the sigma point method is employed for updating a dynamic system of layered composites using uncertainty affected observations. Also in Nguyen and Nestorović (2016), it is used for an iterative identification of soil parameters in the context of tunnel excavation and its performance is compared to the extended Kalman filter. As no metamodel is employed, the computational efficiency than can be improved using this method is of high interest in both publications.
However, instead of using the sigma points method for updating a changing system, one could also employ it to compare distributions of different responses obtained from the same system. This approach to OED is introduced in Schenkendorf et al. (2009) with application to the same system of bioreactor growth to which is referred to in Sec. 5.1.2. For the one-time process of parameter identification, it is shown that using seven sigma points leads to almost same results as using the 10000 Bootstrap samples. In contrast to employing the Bootstrap approach, using the sigma point approach for OED was computationally feasible at all and the identified experimental design exhibits smaller variation of the identified parameters compared to the optimal experimental design identified using the FIM.

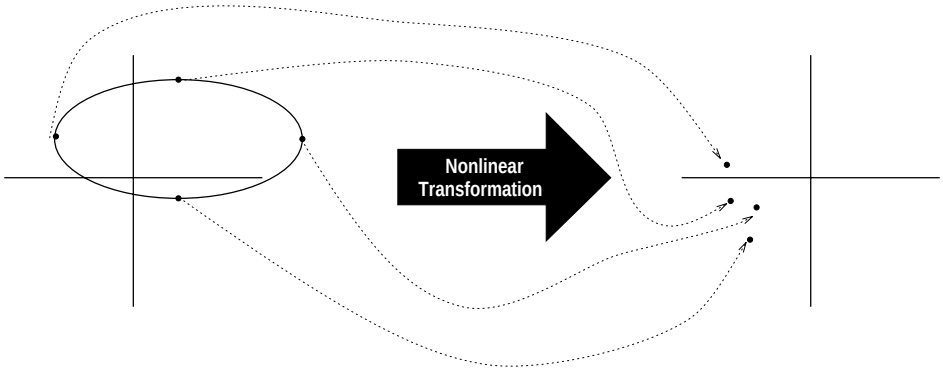


Figure 5.7.: Sigma points capturing the mean and covariance of the distribution are transformed independently. The mean and covariance of the transformed sigma points define the predicted state, after Julier and Uhlmann (1996).

5.3.2. Reconsidering the Dike example

The application example of the dike subjected to a rapid water drawdown introduced in Sec. 4.2.3 and investigated using the Bootstrap method in Sec. 5.2 is now considered

again. In general, the procedure is the same as described for the application of the Bootstrap approach. The difference consists in not generating random samples to obtain the matrix of identified parameters $\tilde{\mathfrak{T}}$ and subsequently $\boldsymbol{C}_{\boldsymbol{\theta}_i,\mathrm{Boot}}$, but to create the matrix of sigma point disturbances $\boldsymbol{\sigma}$.
To do so, the matrix $\tilde{\mathfrak{Y}}$ is reused that contains 150 normally distributed noisy output values for each of the 22 considered candidate locations. In the locations 1 to 11 horizontal and vertical displacement can be measured wherefore the matrix $\tilde{\mathfrak{Y}}$ contains 150 rows and 33 columns. To make the procedure consistent with the previous Bootstrap approach, three output columns out of matrix $\tilde{\mathfrak{Y}}$, one for each measurement type, are combined to generate the covariance matrix $\boldsymbol{C}_{\boldsymbol{y}}$. As the locations for horizontal and vertical displacement measurements are coupled, 121 realisations of $\boldsymbol{C}_{\boldsymbol{y}}$ are obtained according to Eq. 5.11

$$\boldsymbol{C}_{\boldsymbol{y}_{i,j}} = COV(\boldsymbol{y}_i, \boldsymbol{y}_{i+11}, \boldsymbol{y}_j)\ ,\ i = 1,\ldots,11, j = 1,\ldots,11 \tag{5.11}$$

As Gaußian white noise is assumed in $\tilde{\mathfrak{T}}$, κ_s can be set to zero. To calculate the square root of the covariance matrix of outputs $\boldsymbol{C}_{\boldsymbol{y}}$ the following formulation is employed:

$$\sqrt{\boldsymbol{C}_{\boldsymbol{y}}} = \boldsymbol{\gamma}_{\mathrm{E}ig} \cdot \sqrt{\boldsymbol{\lambda}_{\mathrm{E}ig}} \cdot \gamma_{\mathrm{E}ig}^{-1}\ , \tag{5.12}$$

whereby, $\boldsymbol{\gamma}_{\mathrm{E}ig}$ is the matrix of eigenvectors of $\boldsymbol{C}_{\boldsymbol{y}}$ and $\boldsymbol{\lambda}_{\mathrm{E}ig}$ is the matrix obtained by multiplying the vector of eigenvalues of $\boldsymbol{C}_{\boldsymbol{y}}$ by the identity matrix. To each model response y_i obtained using the mean values of input parameters, the sigma point disturbances calculated according to Eq. 5.9 are added to generate the matrix $\tilde{\mathfrak{Y}}_{SP}$ (Eq. 5.13).

$$\tilde{\mathfrak{Y}}_{SP} = \boldsymbol{y} + \boldsymbol{\sigma} \tag{5.13}$$

As for each model response $2J+1$ sigma points are obtained (including the undisturbed model response itself), the matrix $\tilde{\mathfrak{Y}}_{SP}$ has 7 x 121 elements. As the next steps are analogue to those described in the Bootstrap approach $\tilde{\mathfrak{Y}}_{SP}$ is back-calculated one by one to obtain the matrix $\tilde{\mathfrak{T}}_{SP}$ that accordingly consists of seven lines instead of 150 in the previous approach. Using this matrix of identified parameters, Eq. 5.10 is employed to back calculate the corresponding parameters. The therein employed weighting factors w_i are generated as follows:

$$w_i = \frac{1}{2 \cdot (J+1)}\ ,\ i = 1, ..., 2J \tag{5.14}$$

$$w_{2n+1} = \frac{1}{(J+1)}, \tag{5.15}$$

whereby $2J+1$ refers to the unmodified model response $\boldsymbol{y}$. Assuming again three sensors for each type of output, $t = 27225$ possible realisations of $\boldsymbol{C}_{\boldsymbol{\theta},SP}$ are obtained to which the

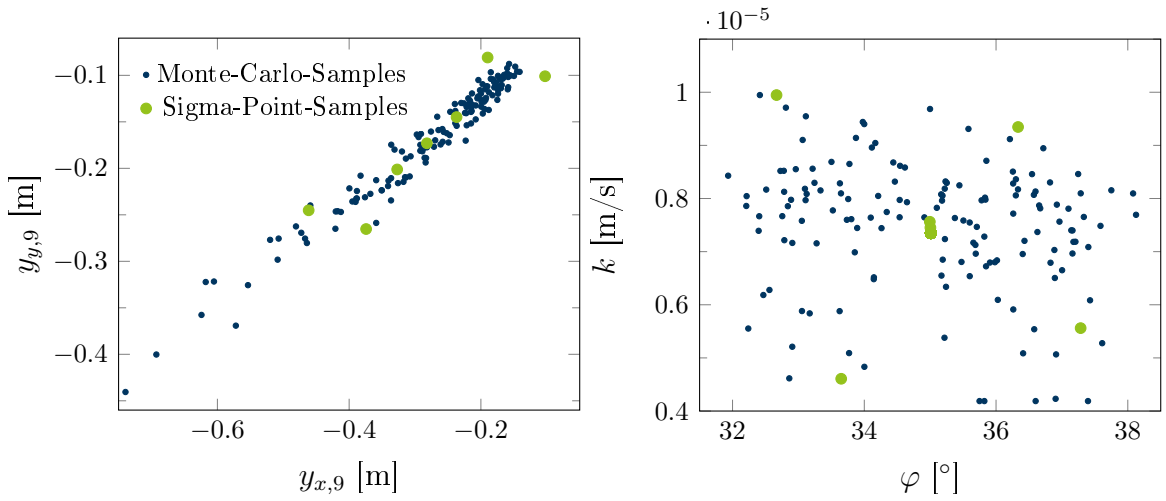
optimality criterion Φ_D is applied. According to the value of Φ_D, a ranking is generated that indicates which experimental design is most favourable. The best and worst three cases are listed in Tab. 5.3.

The comparison of the results with those obtained in Sec. 5.2.2 using the Bootstrap method shows good agreement in case of the displacement sensors, but quite significant differences for the pore water pressure sensors. With respect to the displacement sensors, the best experimental designs include in both methods the positions 4 and 6 and in the worst experimental designs the positions 1 and 11. Within the locations of the pore water pressure sensors, position 21 is among the best experimental designs in both methods, but the further positions are different. Indeed, position 19 that is among the best experimental designs using the Bootstrap approach is now found as part of the worst experimental designs. Besides, location 17 that was selected as control point is not found in the worst experimental designs. However, as the 11 candidate points for pore water pressure sensors are located close to each other, their model responses are quite similar. It would have been more favourable to select them in larger distance to each other to enable an a priori qualitative statement and compare it with the obtained results. The obvious differences in the results does not necessarily origin from a too inaccurate representation of the output distribution by using the sigma points. One should be aware that using the Bootstrap method, falsified output samples were randomly generated. Even though the input space was restrained as described in Sec. 5.2.2, it can not be excluded that some of the error values were strong outliers that did not allow the DE algorithm to find the global minimum within the defined parameter boundaries. In such cases, the algorithm converges to the best possible results, possibly lying on the edge of the parameter space. Therefore, using the sigma points might even be more reliable, as here the output samples are generated in a more controlled manner and their back calculation results are sure to be within the defined parameter boundaries.

The decisive difference between the Bootstrap and the sigma point approach is illustrated in Figs. 5.8a & 5.8b. The random samples known from Fig 5.4a are shown again in blue, but this time with the according green sigma points in contrast. In Fig. 5.8b identified parameter combinations of φ and k are shown. As each dot is representative for one run of back-calculation, it is clear that the sigma point approach is much more time efficient. In the present case, where the previous 150 random samples are substituted by seven sigma points this corresponds to a reduction of 95%. In Hölter et al. (2018b), some of the presented findings are illustrated in a condensed manner.

Table 5.3.: Ranking of the individual sensor positions using sigma-points approach – optimality criterion Φ_{D}.

Sens. Pos. y_y/y_x	Sens. Pos. y_w	Φ_{D}	Rank
[4 5 6]	[15 18 21]	3.58 e -6	1
[4 5 7]	[13 15 21]	3.76 e -6	2
[4 5 7]	[13 18 21]	3.79 e -6	3
[1 10 11]	[14 18 19]	4.65 e -2	27223
[1 10 11]	[15 18 19]	4.73 e -2	27224
[1 10 11]	[14 15 18]	5.16 e -2	27225



(a) (b)

Figure 5.8.: a: Randomly generated output data in comparison to deterministically generated sigma points in case of displacement output at point 9 b: parameter values obtained by back calculating both types of noisy data.

5.4. Summary

The approaches presented in this chapter intend to identify an optimal experimental design by back-calculating noisy data assigned to all candidate positions in which a possible measurement is considered. The variance, respectively the covariance, of the identified constitutive soil parameters indicates which experimental design allows the most reliable parameter identification. The two presented methods differ by the approach of obtaining these parameter statistics. Using the Bootstrap approach, MC simulations are performed that require numerous samples. The sigma point approach employs only few representative points which amount is related to the dimension of the output data (number of considered model responses). The configuration of the results is the same in both methods: a ranking in which each possible experimental design is listed according to its optimality criterion value (value of cost function Φ_{A} or Φ_{D}, respectively). The drawback of using the Bootstrap approach is the major calculation effort that is required to back calculate the noisy data, even though the number of calculations is reduced by employing the Bootstrap resampling method. However, the OED results obtained in the considered example correspond to the expectations made by engineering judgement. The best identified experimental designs are consistent within both methods. In case of the worst experimental design, the sigma point method shows partly inconsistent results, especially in case of pore water pressure measurements. To investigate the origin of this discrepancy, it would be of interest to perform similar calculations with a wider spread of the locations for pore water pressure measurements.

Besides, one should be aware that testing any possible combination might be not an efficient method. The described method is well suited in case the search area is limited, a small finite number of discrete candidate points is considered, or a restriction by previous investigations is possible as in the present case. In case no previous knowledge is available and a free search in design space is required or desired, a type of optimisation algorithm-based search would be more favourable. This idea is taken on in the following chapter where the concept of Bayesian updating is employed for OED.

6. Bayesian OED

The methods introduced in the previous chapter intend to identify an OED by defining a priori a set of candidate locations in which sensors may be located and by testing subsequently each possible experimental design combination. This allows a very accurate identification of the OED, but one might have the impression that this is not an efficient method. Point 11 for example in Fig. 5.2 is part of the poorest experimental designs; a finding that should rise quite early in testing the different designs. An efficient approach would be to use this knowledge and to exclude any experimental design that contains this point in the further search.
Using knowledge obtained prior to improve the posterior steps is now the idea of the Bayesian updating or inference that is introduced in Sec. 3.5. The present chapter describes how this concept can be considered in the framework of OED. After referring to several applications from other research fields, an approach for the employment in geotechnical engineering is introduced as well as an application of this approach to an example of mechanised tunnelling simulation.

6.1. Concepts

As the objective of OED for parameter identification is in general to reduce parameter uncertainties based on observed values, the Bayes' theorem introduced in Sec. 3.5 should be considered to improve this process. Looking at the Bayes' theorem as shown in Eq. 3.27, the uncertainty of the parameters of interest $\boldsymbol{\theta}$ are considered in the form of probability distributions $P(\boldsymbol{\theta})$ that are introduced in Sec. 3.5. Measurement data $\tilde{\boldsymbol{y}}$ contains the "true" system responses $\boldsymbol{y}$, but is subjected to a certain ratio of measurement error e that is assumed to correspond to a known probability distribution. Therefore, the obtained data is regarded as an event of probability $P(\tilde{\boldsymbol{y}})$. The conditional probability $P(\tilde{\boldsymbol{y}}|\boldsymbol{\theta})$ describes the probability that a certain output results from a parameter combination, what is assumed to be known, while $P(\boldsymbol{\theta}|\tilde{\boldsymbol{y}})$ is the value of interest that describes the probability of a set of parameters to be correct given a measured output value as formulated in Eq. 3.27.

As it is demonstrated in the previous chapters, the measurement data depends on the experimental design $\boldsymbol{\delta}$ that is applied to the considered problem. Transferring this idea to the Bayes' theorem, Eq. 3.27 is extended to Eq. 6.1.

$$P(\boldsymbol{\theta}|\tilde{\boldsymbol{y}}, \boldsymbol{\delta}) = \frac{P(\tilde{\boldsymbol{y}}|\boldsymbol{\theta}, \boldsymbol{\delta}) \cdot P(\boldsymbol{\theta}|\boldsymbol{\delta})}{P(\tilde{\boldsymbol{y}}|\boldsymbol{\delta})} \tag{6.1}$$

Herein, $P(\boldsymbol{\theta}|\tilde{\boldsymbol{y}}, \boldsymbol{\delta})$ is the posterior probability of the parameters $\boldsymbol{\theta}$ given a set of measurements $\tilde{\boldsymbol{y}}$ that was obtained using the experimental design $\boldsymbol{\delta}$. Within the process of Bayesian OED, it is investigated which experimental design $\boldsymbol{\delta}$ provides measurement data that has the highest probability of reducing the initial uncertainty of the model parameters.

6.1.1. Applications in OED

In the last decades, the concept of Bayesian inference was employed in several applications of OED. Chaloner and Verdinelli (1995) provides a detailed introduction to the topic. Therein, a number of simple examples from biomedicine demonstrate the concept of Bayesian OED. Unlike the approaches introduced in the previous chapters where a cost or optimality function was optimised by minimisation, the scientific community has agreed on maximising the utility U in the framework of Bayesian OED to find the optimal design. Accordingly, the experimental design $\boldsymbol{\delta}$ should be selected in a way that its data increases most the utility U from the prior to the posterior state. In Vanlier et al. (2012), another example from system biology is presented where the interactions of different proteins in a cell are simulated using a system of ordinary differential equations (ODEs). The objective of the therein performed Bayesian OED investigation is to identify at which location in time measurements should be performed. By identifying optimal sampling times, the authors are able to reduce the necessary size of employed samples that they identify as time consuming restriction.

In particular Huan and Marzouk (2013) should be named to which the procedure in this thesis refers. Therein, beside an introductory synthetic example, the problem of combustion kinetics in a shock tube is investigated. The chemical reactions that take place in a shock tube after ignition are elementary chemical reactions and assumed to take place spatially homogeneous (i.e., well-mixed) at constant pressure, called therefore "zero-dimensional". However, describing which processes take place among the different hydrocarbons is described by a system of ODE with uncertain kinetic parameters that depend on temperature and molar or mass fractions as state variables. The intention of

Huan and Marzouk (2013) is therefore to identify at which temperature and equivalence ratio the uncertain kinetic parameters can be identified most reliable from the following observable values: the peak value of the heat release rate, the peak concentrations of various intermediate chemical species, and the times at which these peak values occur. These form accordingly the experimental design $\boldsymbol{\delta}$ and the parameter vector $\boldsymbol{\theta}$, respectively. The vector of model responses $\boldsymbol{y}$ is falsified with Gaußian white noise with standard deviation of 10% of the output value.
The general formulation to the utility U in that publication that is also employed within this thesis is described in Eq. 6.2:

$$U(\boldsymbol{\delta}) = \int_{\boldsymbol{\Omega}} \int_{\boldsymbol{\Theta}} u(\boldsymbol{\delta}, \tilde{\boldsymbol{y}}, \boldsymbol{\theta}) p\left(\boldsymbol{\theta}, \tilde{\boldsymbol{y}} | \boldsymbol{\delta}\right) d\boldsymbol{\theta} d\boldsymbol{y} \tag{6.2}$$

As in the previous chapters of this thesis, $\boldsymbol{\delta}$ is the considered experimental design, $\boldsymbol{\theta}$ is the employed set of constitutive parameters from the parameter space $\boldsymbol{\Theta}$ and $\tilde{\boldsymbol{y}}$ is the obtained measurement data that is subjected to measurement errors and $\boldsymbol{\Omega}$ is the output space of dimension n. $u\left(\boldsymbol{\delta}, \tilde{\boldsymbol{y}}, \boldsymbol{\theta}\right)$ is the utility function to which the integral $U(\boldsymbol{\delta})$ is the expected utility. Using the Kullback–Leibler divergence, the utility function is defined as follows:

$$u(\boldsymbol{\delta}, \tilde{\boldsymbol{y}}, \boldsymbol{\theta}) = \int_{\boldsymbol{\Theta}} p\left(\boldsymbol{\theta} | \tilde{\boldsymbol{y}}, \boldsymbol{\delta}\right) \ln \left[\frac{p\left(\boldsymbol{\theta} | \tilde{\boldsymbol{y}}, \boldsymbol{\delta}\right)}{p(\boldsymbol{\theta})} \right] d\boldsymbol{\theta} = u(\boldsymbol{\delta}, \tilde{\boldsymbol{y}}) \tag{6.3}$$

As the system parameters are independent of the employed design, the utility does also not depend on them. In Huan and Marzouk (2013), Eqs. 6.2 and 6.3 are merged to the formulation of Eq. 6.4.

$$U(\boldsymbol{\delta}) = \int_{\boldsymbol{\Omega}} \int_{\boldsymbol{\Theta}} p\left(\boldsymbol{\theta} | \tilde{\boldsymbol{y}}, \boldsymbol{\delta}\right) \ln \left[\frac{p\left(\boldsymbol{\theta} | \tilde{\boldsymbol{y}}, \boldsymbol{\delta}\right)}{p(\boldsymbol{\theta})} \right] d\boldsymbol{\theta}\ p(\tilde{\boldsymbol{y}} | \boldsymbol{\delta}) d\boldsymbol{y} \tag{6.4}$$

The optimal experimental design $\boldsymbol{\delta}^*$ is accordingly defined as follows:

$$\boldsymbol{\delta}^* = \arg\max_{\boldsymbol{\delta} \in \boldsymbol{\Pi}} U(\boldsymbol{\delta}) \tag{6.5}$$

However, solving the integral in Eq. 6.4 is hardly possible if the output data $\boldsymbol{y}$ is obtained from FE models as it is the case in the present thesis. Also in the referred application of Huan and Marzouk (2013), these integrals are numerically approximated. In that publication, the approach of using polynomial chaos surrogate is employed to generate noisy output samples and two different optimisation algorithms are employed to approximate Eq. 6.4. The authors perform the design optimisation for a single experiment as well as for a series of experiments, demonstrating thereby, which impact this repetition can have on the reliability of obtained results as well as on the experimental designs themselves. As

an application, the advancement of a tunnel excavation is considered in this thesis, which is described in detail in Sec. 6.3.1. As such "experiment" can obviously not be repeated, this corresponds to the case of a single experiment.

6.1.2. Aproximate coordinate exchange algorithm

The aforementioned problem that the integral in Eq. 6.4 can hardly be solved arises, beside the aspect of employing FE-models, from the aspect of high dimensionality of $\boldsymbol{\delta}$. In combination with the number s of constitutive input parameters $\boldsymbol{\theta}$ of the model, the curse of dimensionality rapidly obstructs solving this optimisation problem.
Therefore, in the present work the Approximate Coordinate Exchange (ACE) algorithm is employed that is introduced in Meyer and Nachtsheim (1995) and adapted for Bayesian OED in Overstall and Woods (2017). The basic idea of this optimisation algorithm is to investigate each of the d elements of the design vector $\boldsymbol{\delta}$ one by one and to repeat this operation iteratively until convergence of the utility is obtained. To accurately capture the utility U of a certain design, a series of one-dimensional metamodels $\tilde{U}$, or emulators as called in Overstall and Woods (2017), are generated to perform the single-parameter optimisations:

$$\tilde{U} = \tilde{U}(\delta_i|\boldsymbol{\delta}_{(i)}) \ , \ i = 1, \ldots, d \tag{6.6}$$

The utility U of each individual design is evaluated using a Monte-Carlo Simulation:

$$U(\boldsymbol{\delta}) \approx \tilde{U}(\boldsymbol{\delta}) = \tilde{U}(\delta_i|\boldsymbol{\delta}_{(i)}) = \sum_{l=1}^{B} u\left(\boldsymbol{\delta}, \tilde{\boldsymbol{y}}_l, \boldsymbol{\theta}_l\right)/B \tag{6.7}$$

where B corresponds to the number of Monte-Carlo samples that are employed to approximate the utility of the i^{th} design and that are drawn from the probability distribution $p\left(\boldsymbol{\theta}, \tilde{\boldsymbol{y}}|\boldsymbol{\delta}\right)$. To start the optimisation, an initial design $\boldsymbol{\delta}_0$ is either randomly sampled in $\boldsymbol{\Pi}$ or prescribed according to a possible a priori knowledge of the user. For each of the d dimensions of the design, the optimisation is run to maximise $\tilde{U}$ and identify the corresponding best design $\boldsymbol{\delta}^{\dagger}$ that should outperform the current design $\boldsymbol{\delta}^{\text{c}}$. After identifying the best design among all dimensions d of $\boldsymbol{\delta}$, the distribution is updated in the sence of Bayesian inference according to the knowledge obtained in the current iteration step. To do so, a new probability distribution of parameters $\boldsymbol{\theta}$ is calculated as follows:

$$p_l^{\dagger} = 1 - T_{2B-2}\left(-\frac{B\tilde{U}\left(\boldsymbol{\delta}^{c\dagger}\right) - B\tilde{U}\left(\boldsymbol{\delta}^{c}\right)}{\sqrt{2B\hat{\nu}_l}}\right) \tag{6.8}$$

with T_{2B-2} being a probability distribution function for a Student's t-distribution with $2B-2$ degrees of freedom and the coefficient $\tilde{\nu}_l$ as follows:

$$\hat{\nu}_l = \frac{\sum_{l=1}^{B}\left[u\left(\boldsymbol{\delta}^{c\dagger}, \boldsymbol{y}_l^{\dagger}, \boldsymbol{\theta}_l^{\dagger}\right) - \tilde{U}(\boldsymbol{\delta}^{c\dagger})\right]^2 + \sum_{l=1}^{B}\left[u\left(\boldsymbol{\delta}^{c}, \boldsymbol{y}_l^{c}, \boldsymbol{\theta}_l^{c}\right) - \tilde{U}(\boldsymbol{\delta}^{c})\right]^2}{2B-2} \tag{6.9}$$

Using this PDF, the samples are generated that are employed in the next iteration to calculate the utility $\tilde{U}(\boldsymbol{\delta})$. The number of iteration steps can be selected by the user and should allow convergence.

In Overstall and Woods (2017), a second phase to the algorithm is suggested in which it is tested if the utility of a certain design can even be improved by clustering the individuals of the identified design. As in applications of geotechnical engineering this phenomenon is especially undesired, this step is not considered in the present thesis. The authors apply the algorithms to synthetic examples to make it comprehensible to the reader, but also to an example of pharmacokinetics where once again the sampling time is decisive to obtain as much information as possible from an experiment and should therefore be optimised.

The ACE algorithm is implemented in the **R** package **acebayes**. The main attributes of this package are the utility function $u(\boldsymbol{\delta}, \tilde{\boldsymbol{y}})$ that is to be maximised, the initial experimental design $\boldsymbol{\delta}_\mathbf{o}$, the limits of the design space $\boldsymbol{\Pi}$, the size of B, the number Q of support points to generate the one-dimensional metamodel $\tilde{U}$, and the maximum number of iteration runs. Besides, a function to sample from the prior parameter distribution is required that is called within the utility function.

6.2. Employment in geotechnical engineering

In the framework of this thesis, the introduced Bayesian OED shall be applied to a problem of geotechnical engineering using the ACE algorithm for optimisation. The objective of the application would be similar to those in the problems discussed in the previous chapters 4 & 5: it should be identified where to place sensors in a defined area in a way to identify most reliably the relevant parameters of the soil, i.e. to reduce most the uncertainty of the elements of $\boldsymbol{\theta}$. The uncertainty of the soil parameters must be defined initially in terms of PDFs, whereby the selection of the type of PDF should reflect the actual knowledge of the soil properties.

As for the intended application, like for the other examples shown in this thesis, a time consuming FE-model is employed and not an analytical model, it would not be possible to call this model repeatedly during the optimisation. Therefore, it is necessary to create

a sufficiently large database in advance that enables to generate adequate metamodels as described similarly in Sec. 3.2. These models should be of the following type:

$$\boldsymbol{y} = (y_1, y_2, \ldots, y_n) = f(\boldsymbol{\theta}), \boldsymbol{y} \in \mathbb{R}^n, \boldsymbol{\theta} \in \mathbb{R}^s \tag{6.10}$$

The n elements of $\boldsymbol{y}$ should enable to cover as dense as possible the experimental design space in which the optimal experimental design $\boldsymbol{\delta}^*$ is identified afterwards. The complexity of this identification as well as the required number N of samples $\boldsymbol{\theta}$ depends on the number s of parameters that are considered in the model. Therefore, it is strongly recommended to perform a preliminary GSA that is introduced in Sec. 4.1.1.2. Doing so, it is possible to identify and reduce the amount of parameters considered in the model to those that have a relevant impact on the system response of interest. Hereby, it should be mentioned that these are not necessarily consistent with those parameters that are most influencing the model responses in $\boldsymbol{y}$.
After obtaining the corresponding snapshot matrix $\mathfrak{Y}$ of model outputs, a metamodel is generated that substitutes the model described by Eq. 6.10. This could be done either as in Eq. 6.11 or as in Eq. 6.12.

$$y_i = \hat{f}(\boldsymbol{\theta}, i), i = 1, \ldots, n \tag{6.11}$$

$$y = \hat{f}(\boldsymbol{\theta}, u_\mathrm{x}, u_\mathrm{z}) \tag{6.12}$$

Eq. 6.11 provides output data in any of the n points considered in Eq. 6.10 and nowhere else. Using Eq. 6.12, an additional interpolation between these points is performed and model responses for any coordinate combination $(u_\mathrm{x}, u_\mathrm{z})$ within the defined output space are obtained. This could be extended to additionally obtain u_y in case of a 3-dimensional design space or to any further considered dimension. The first option prohibits clustering due to the distance between the selected output positions. The latter one indeed allows a more detailed optimisation as one is not restricted to consider points located on the grid of model output locations (or any other selected arrangement), but this requires a higher level of abstraction. Along with the generation of such more complex metamodel, a potentially higher approximation error is obtained, contributing to the model uncertainty described in Sec. 2.3.3.3. Within this thesis, the grid-free version is employed for the sake of a larger difference to the OED approaches considered in the previous chapters.
As first step of performing Bayesian OED, an initial design $\boldsymbol{\delta}_0$ is generated within that part of the model geometry in which the experimental design to be optimised should be located, corresponding to the design space $\boldsymbol{\Pi}$. This can be performed either by generating a random sample of dimension d within the boundaries of $\boldsymbol{\Pi}$ or by defining a specific one. The latter approach can be useful if the model responses in some areas of the design

space are uniformly insensitive to the varied parameters of the experimental design. In such cases, no higher utility can be identified and no update of the parameter distribution as described in Eq. 6.8 can be obtained.
Similar to the OED concepts described in the previous chapter, artificial noisy data is generated that corresponds to the locations of the experimental design $\boldsymbol{\delta}_0$. In this application, a Gaußian white noise is applied that is independent of the specific location of the sensor, and added to a sample of $l = 1, \ldots, B$ output values of $\boldsymbol{y}$ as described by Eq. 6.13:

$$\tilde{\boldsymbol{y}}_l(\boldsymbol{\theta}, \boldsymbol{\delta}_0) = \boldsymbol{y}(\boldsymbol{\theta}, \boldsymbol{\delta}_0) + e \cdot \omega,\ \omega \sim \mathcal{N}[0, 1] \tag{6.13}$$

where e is again a user-defined value to control the noise level. In the next step, the parameters of interest $\boldsymbol{\theta}$ are back-calculated from each of the noisy samples $\tilde{\boldsymbol{y}}_l$:

$$\tilde{\boldsymbol{\theta}}_l = f^{-1}(\tilde{\boldsymbol{y}}_l | \boldsymbol{\delta}^c)\ l = 1, \ldots, B\ , \tag{6.14}$$

whereby $\boldsymbol{\delta}^c$ refers to the currently investigated experimental design. The corresponding utility U ($\tilde{U}$, respectively) of a certain experimental design $\boldsymbol{\delta}^c$ should be highest if the uncertainty of the identified parameters is smallest, wherefore the following utility function is proposed:

$$\tilde{U}(\boldsymbol{\delta}^c) = \sum_{l=1}^{B} \left(\frac{-|\tilde{\boldsymbol{\theta}}_{\boldsymbol{l}}(\boldsymbol{\delta}^{\boldsymbol{c}}) - \bar{\boldsymbol{\theta}}|}{\bar{\boldsymbol{\theta}}} + \frac{1}{\det \boldsymbol{C}_{\boldsymbol{\theta}}(\boldsymbol{\delta}^c)} \right) / B \tag{6.15}$$

$\bar{\boldsymbol{\theta}}$ describes the mean value of the identified parameter values using the current experimental design such that the term $-|\tilde{\boldsymbol{\theta}}_{\boldsymbol{l}} - \bar{\boldsymbol{\theta}}|$ maximises when the discrepancy is small. Doing so allows to capture the spread of the distribution of u while $\det \boldsymbol{C}_{\boldsymbol{\theta}}(\boldsymbol{\delta}^c)^{-1}$ controls the mean value of this distribution. The covariance matrix $\boldsymbol{C}_{\boldsymbol{\theta}}$ is obtained as described in Eq. 2.5. Employing Eq. 6.15 as objective function in the ACE algorithm introduced in the previous section shall allow to identify the optimal arrangement $\boldsymbol{\delta}^*$ that leads to the most reliable experimental design. Its application is demonstrated in the following.

6.3. Application

6.3.1. Considered example

The considered example is that of a twin tunnel construction in an urban area, that partly underpasses a nine-storey building as illustrated in Fig. 6.1a. The position of the building

and how it is related to the two tunnel tubes is displayed in Fig. 6.1b. Within this example, the simulation of the construction of line 5 of the Milan Metro in a section between the stations Lotto and Portello is considered. The soil stratigraphy, tunnel dimensions, and TBM details are accurately reproduced in the FE model that has been introduced in Fargnoli et al. (2015a) and Fargnoli et al. (2015b) and that is employed for the present work, too. A visualisation of the model and its dimensions is provided in Fig. 6.2. The three soil layers that are two gravelly sand layers with a silty sand layer in between are simulated using the HSsmall model, introduced by Benz (2007).

(a) (b)

Figure 6.1.: a: Front view and b: top view of the considered building and location of applied measurement points, modified, after Fargnoli et al. (2015b).

6.3.2. Motivation and objectives of application

The HSsmall model corresponds mostly to the HS model introduced in Sec. 4.2.2.2, but extends it by two further parameters that allow to consider an increased stiffness in case of small strains in the soil. The building that is specifically considered in this model is selected because an extensive monitoring program has been applied to it. This monitoring program consists of 13 sensors with the positions indicated in Fig. 6.1b. The sensors record the local vertical displacements during the consecutive construction of the two tunnel tubes. In parallel, the position of the TBM, the applied face pressure and grouting pressure are recorded. This enables to validate the numerical model and to some extend the results of an OED process.

Mechanised tunnel constructions are challenging projects not only because of their overall

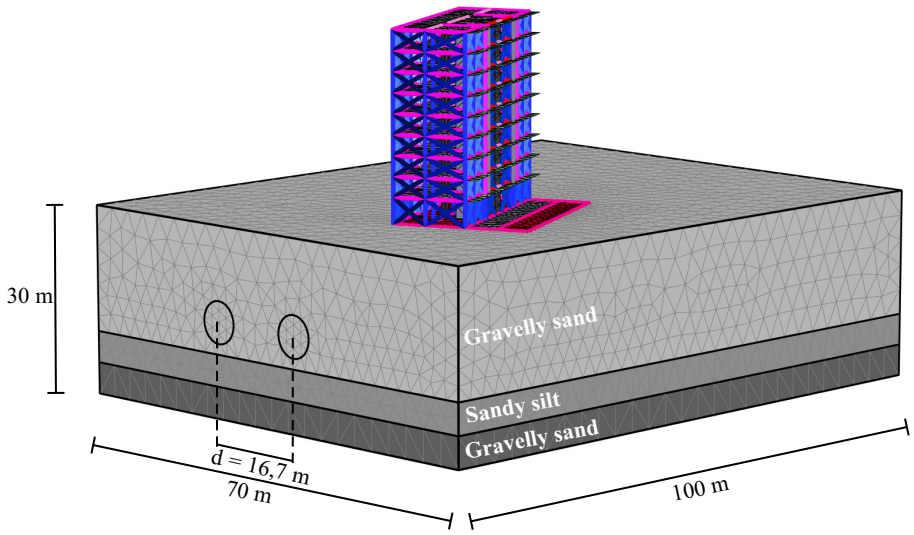


Figure 6.2.: FE model of considered interaction of twin tunnel bypassing the nine-storey building, modified, after Schoen (2018*a*).

size, but also because the structure is installed in a large soil area. Obtaining holistic detailed soil properties that allow an accurate prediction of the tunnelling induced impacts is hardly possible.

As mentioned above and described in Fargnoli et al. (2015a) and Fargnoli et al. (2015b), in the present case thirteen measurement sensors have been installed on the edge of the considered building. The overall objective of performing measurements is to ensure the safety of the endangered structures. Simply considering measurements as threshold values that should not be exceeded, as described in chapter 1 as "conventional design" is one way, but the more sustainable way followed in this thesis is using measurement data for parameter identification and model validation. If an adequate model is obtained that can very precisely predict the building behaviour, countermeasures can be more efficiently assessed that lead to improve the building's safety. Therefore, the objective becomes to obtain a reliable model by identifying the relevant soil properties using the method of back analysis, described in Sec. 3.3, which will benefit from considering the ideas of OED. However, considering the current example one may ask 'what is the benefit of identifying soil parameters in an area after the tunnel has been constructed there?' and this is of course a justified question, as this procedure should not end in itself, performing parameter identification for the sake of parameter identification.

Several aspects must be considered here. First, the building simulated in this model is one out of numerous in the city. Any knowledge gained here is valuable for the next building

on the ongoing path of the tunnel as well, as soil properties do vary over distances, but they do not jump suddenly. Second, the presented results are obtained for that moment after the first TBM has passed the building. However, the measurement data is available for each excavation step, i.e. as well for phases where the TBM has not reached the building. When bringing this approach to application, the optimisation of the measurement arrangement should be performed in a continuous and adaptive manner. Hereby, the suggested positions will presumably be varying by time and therefore enable earlier data gathering and processing. Information would be obtained before the TBM reaches the building. Over time, the closer the TBM approaches the more relevant information are obtained, but with less options to react by countermeasures.
Third, similar to the second aspect mentioned but probably more apparent, the evaluation is performed after the first tunnel has been constructed. As the second tunnel follows in parallel and causes the final impact on the building, all technical countermeasures or adaptations of the TBM steering can be undertaken based on the current result but with respect to the final situation.
For these reasons, it makes sense to not only take into consideration those sensors located on the simulated building itself, but also in the far field around it. The considered design space is set in a way that any position in the range of the settlement trough of the tunnel construction is included. To be able to consider also these areas, artificial measurement data is generated.

6.3.3. Preliminary work

In Fargnoli et al. (2015a) and Fargnoli et al. (2015b) the focus of investigation is set on the soil-structure interaction behaviour that is induced by the construction of the twin-tunnel. The measurement data is employed in combination with the numerical results to qualitatively validate the therein introduced analytical approaches to describe this interaction and not to identify the soil properties. Studies performed preliminary to the topic of Bayesian OED investigated the aspects of parameter identification, GSA, and GSA-based OED: In Schoen (2018*a*), the in-situ measurement data is employed to back-calculate the constitutive parameters of the surrounding soil, following the concept described in Sec. 3.3. Besides geometry and types of soil layers, any further prior knowledge is explicitly ignored, only relying on the capacities of the developed mathematical methods. Based on engineering judgement, it was assumed that the parameters that are most relevant for possible settlements or tilting of the building are the grouting pressure p_v and face pressure p_s of the TBM, the site-related volume loss factor V_L, the friction angle φ_1, the small-strain

stiffness $G_{0,1}$, and the secant stiffness $E_{50,1}^{\mathrm{ref}}$ of the first soil layer and the secant stiffness $E_{50,2}^{\mathrm{ref}}$ of the second soil layer as well as the stiffness of the concrete-made building foundations E_{F}. It should be mentioned that for both layers the secant stiffness E_{50}^{ref} is assumed to be correlated with the tangent stiffness $E_{\mathrm{oed}}^{\mathrm{ref}}$ and the unloading-reloading stiffness $E_{\mathrm{ur}}^{\mathrm{ref}}$ by the factor of three (i.e. $E_{\mathrm{ur}}^{\mathrm{ref}} = 3 \cdot E_{\mathrm{oed}}^{\mathrm{ref}} = 3 \cdot E_{50}^{\mathrm{ref}}$) as performed previously in Sec. 4.2.3.1. When in the following the term "stiffness" is employed, it refers to those three parameters combined.

For model validation, only the measurement data obtained until the first tunnel tube is constructed is employed such that a prediction can be performed for the excavation of the second tunnel tube and be compared to the corresponding measurement data. Doing so, the results displayed in Fig. 6.3a are obtained. Here, the measured settlements obtained after excavation of the first tube in the thirteen measurement points shown in Fig. 6.1b are displayed by the blue triangles. By green circles and grey cubes, the settlements are displayed that correspond to the identified parameters using the actual FE-model and the corresponding metamodel, respectively. It can be seen that in some cases the discrepancy between metamodel and FE-model is larger than between metamodel and measurements. In such cases, one should consider that a better agreement between metamodel and measurements might not be possible without further model improvement. As still an overall good agreement is obtained between measurements and simulation, the identified set of parameters is employed for further investigations and for the prediction of the settlements caused by the construction of the second tube as shown in Fig. 6.3b. The measurements after finalising the second tube are compared with the corresponding results of the FE-model. Still a good agreement can be observed with exception of the points R4, T1, T2, and T3. Here, one should consider that as shown in Fig. 6.1a, next to the considered building, another one is located. This second building (and all further located in the row of buildings) will induce a pre-stressing of the subsoil and thus a stiffness increase that causes the lower settlements observed in reality. Further investigations with continuous buildings would be meaningful with respect to this point. To obtain a deeper understanding of the system behaviour, in Schoen (2018*b*) several GSAs are performed, relating the aforementioned soil and system parameters to the average settlements and tilting in different axis of the building at final stage of the construction. In Figs. 6.4 & 6.5, the results of the GSAs performed after excavation of the first and the second tunnel tube are displayed, relating the seven investigated input parameters to the vertical settlements of each of the thirteen measurement points shown in Fig. 6.1b. Looking a the displayed results, it can be seen that they vary over time and space, but that generally the parameters $E_{50,1}^{\mathrm{ref}}$, $G_{0,1}$, and V_{L}

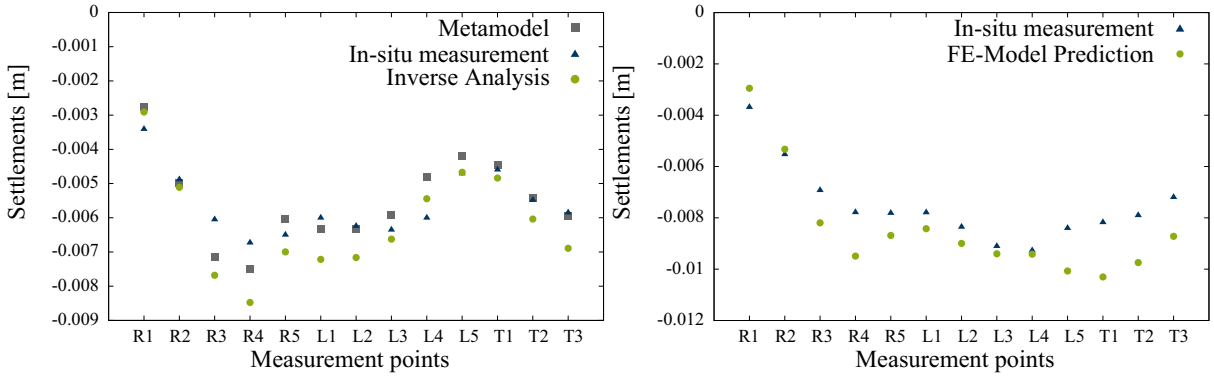


Figure 6.3.: Comparison of measured and a: back-calculated data of first tunnel excavation and b: measured and predicted data of second tunnel excavation.

have most influence on the investigated results. For this reason, all further investigations, in the preliminary and the present work, put focus only on these three parameters.

6.3.4. Optimal locations of sensors for vertical displacements

In the following, it is illustrated how the concept of Bayesian OED using the ACE algorithm as introduced in the Secs. 6.1.1 & 6.1.2 is applied to the example described in the previous section.

At first, the experimental design optimisation problem should be defined. After completing the first tunnel tube, the vertical settlements should be obtained at any place at the ground surface within the limits of the employed FE-model shown in Fig. 6.2. Therefore, a grid of points with a distance of 5 m in between each of them is defined over the model range that is shown in Fig. 6.6a. The position of the building is indicated and in its area no points are defined, leading to a total of 259 points. A sample set of 120 combinations of the three parameters of interest is generated using LHS and applied to the FE-model. The bandwidth of these parameters is narrowed and retraced according to the findings of the preliminary performed parameter identification while the remaining four parameters are fixed to the identified values. These ranges are given in Tab. 6.1 together with the identified ranges of the other parameters.

To substitute the FE-model, a metamodel using quadratic polynomial regression is employed that enables to receive the settlement output in the 259 locations for any parameter combination within the ranges given in Tab. 6.1. To be able to receive settlement data in between these points, the polynomial interpolation method that is introduced in Akima

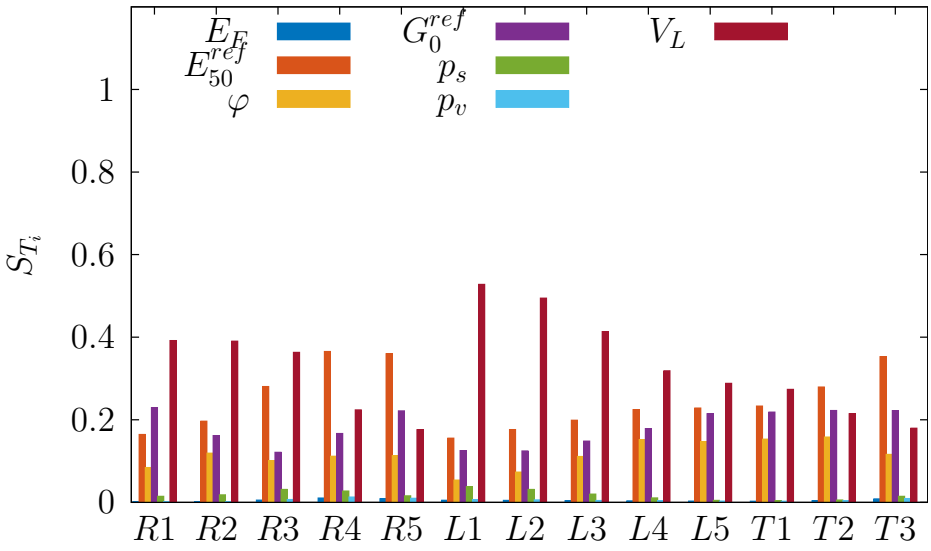


Figure 6.4.: Results of GSA applied in the thirteen measurement points considering the impact of the seven parameters of interest on the settlements after excavation of the first tunnel tube.

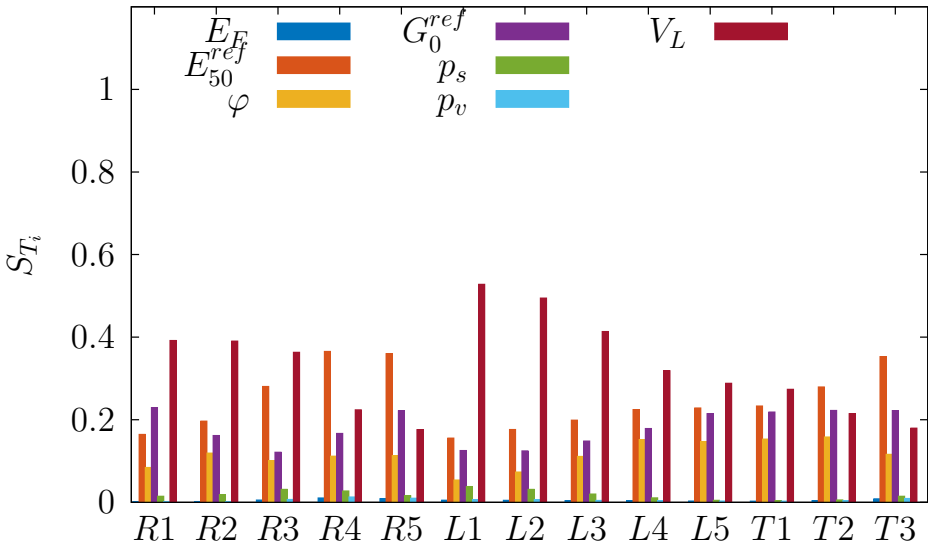
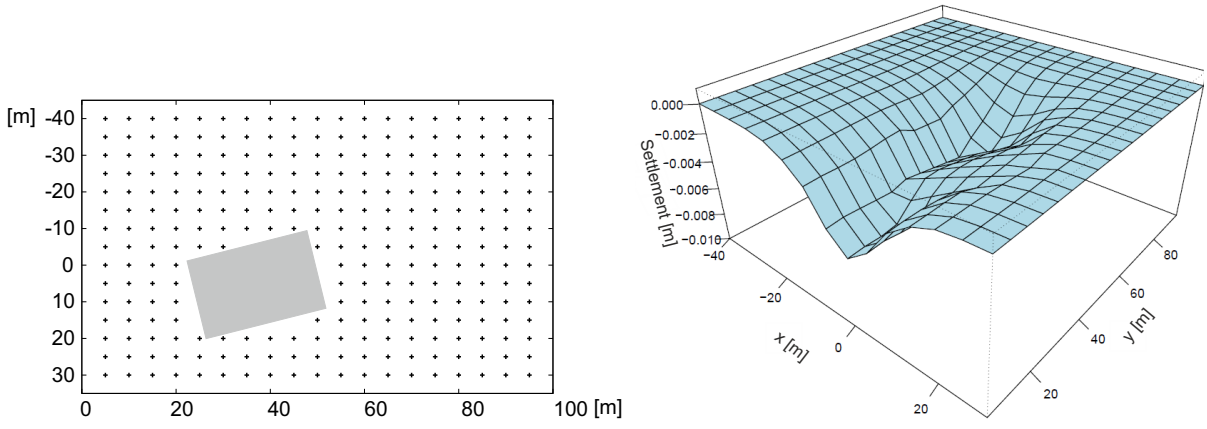


Figure 6.5.: Results of GSA applied in the thirteen measurement points considering the impact of the seven parameters of interest on the settlements after excavation of the second tunnel tube.

(1978) and implemented in the **akima** package in **R** is employed.
The limits of the grid defined in Fig. 6.6a with exception of the surface of the building define the limits of the area in which the sensors might be placed to identify the relevant parameters. It is defined that 6 sensors should be placed in this area. As each sensor can freely move into both dimensions of the surface, finding the optimal experimental design $\boldsymbol{\delta}^*$ becomes a twelve-dimensional optimisation problem. To set-up an initial experimental design, it is sufficient to generate a random 6 x 2 matrix within the given geometric boundaries. However, this can lead to severe problems in certain cases. When looking at Fig. 6.6b where the expected settlement trough obtained for an arbitrary parameter combination is shown, one can see that some areas on the boundaries of the model are not influenced by the building or the tunnel excavation at all and remain perfectly flat. If the ACE algorithm starts its search for experimental designs with high utilities in such an area, the model response will always be the same and the utility of the experimental design, too. Accordingly, none of the investigated current experimental designs is accepted as new best experimental design $\boldsymbol{\delta}^\dagger$ and the initial experimental design $\boldsymbol{\delta}_0$ is always kept. In the present application, it is therefore appropriate to place the initial locations of the six sensors anywhere in the range of influence of the tunnel such that the different back analysis runs lead to varying responses of the utility function u.
The initial distributions, or prior according to ACE denotation, of the parameters are generated as normally distributed samples, as described in Sec. 3.4.1, with means equal to the means of bandwidth defined in Tab. 6.1 and a COV of 0.29. Doing so, it is almost ensured that no samples are generated outside of $\boldsymbol{\Theta}$. These parameter samples are trans-



(a) (b)

Figure 6.6.: a: locations of points in the FE-model where the settlements are recorded for metamodel generation (Schoen, 2018*a*) b: 3D view of settlement trough after first tunnel tube underpassing the considered building.

Table 6.1.: Identified parameter values and assumed parameter ranges for setting up the metamodel of tunnel simulation.

Parameter	Range / Value	Unit
$E_{50,1}^{\mathrm{ref}}$	[25000 - 40000]	[kN/m^2]
$G_{0,1}$	[125000 - 180000]	[kN/m^2]
V_{L}	[0.3 - 0.4]	[-]
φ_1	28.31	[°]
p_{v}	126	[kN/m^2]
p_{s}	155	[kN/m^2]
E_{F}	[23846.28]	[kN/m^2]

ferred to model responses $\mathfrak{Y}$ in the six locations of the initial experimental design. In the current application, the number of samples is set to $B = 100$ and the model responses are falsified according to Eq. 6.13 to obtain $\tilde{\mathfrak{Y}}$. Within the ACE-algorithm, these B noisy samples are back-calculated using the DE-algorithm. The obtained parameter values $\tilde{\boldsymbol{\theta}}$ are evaluated using Eq. 6.15 to identify the utility $\tilde{U}$ of the current experimental design $\boldsymbol{\delta}^c$. This is repeated Q times in each of the d dimensions of the design parameter space $\boldsymbol{\Pi}$. In the current example, Q is kept at its default value of 20, while the maximal number of iteration runs is set to eight. As shown in Fig 6.7, this number seems to be sufficient as the obtainted value of $\tilde{U}$ is converging. Limiting the number of iterations becomes a relevant aspect because of the calculation effort. In the present quite simple example, with $B = 100$, $Q = 20$, $d = 12$, and eight iterations this multiplies to 192,000 runs of back analysis requiring several days of calculation on a conventional PC or 8 h for one iteration step. The obtained results are displayed in Fig. 6.8. In red and green, the settlement at the ground surface in the whole model domain is shown for one exemplary parameter combination, with the location of the building marked as black block. When comparing to Fig. 6.2, one can identify the settlement pattern at the ground surface that is caused by the tunnel excavation after approximately 60 m of TBM advancement. By white cubes, the locations of the sensors in the initial experimental design $\boldsymbol{\delta}_0$ are displayed. One can see that they are placed in areas where large settlements take place for the reasons mention above. The black crossmarks symbolise the final experimental design to which the algorithm converges, i.e. for which the highest utility could be obtained. It can be seen, that the black marks move away from the initial experimental design to the edge of the settlement trough. Here, it should be mentioned that the settlement distribution corresponds to one arbitrary realisation of the input parameters $\boldsymbol{\theta}$. For other combinations, the

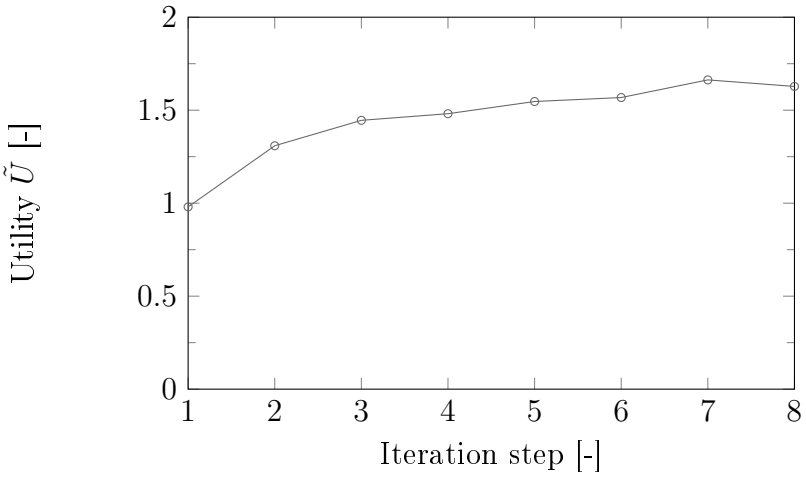


Figure 6.7.: Convergence of Utility $\tilde{U}$ of best current experimental design during steps of iteration progress.

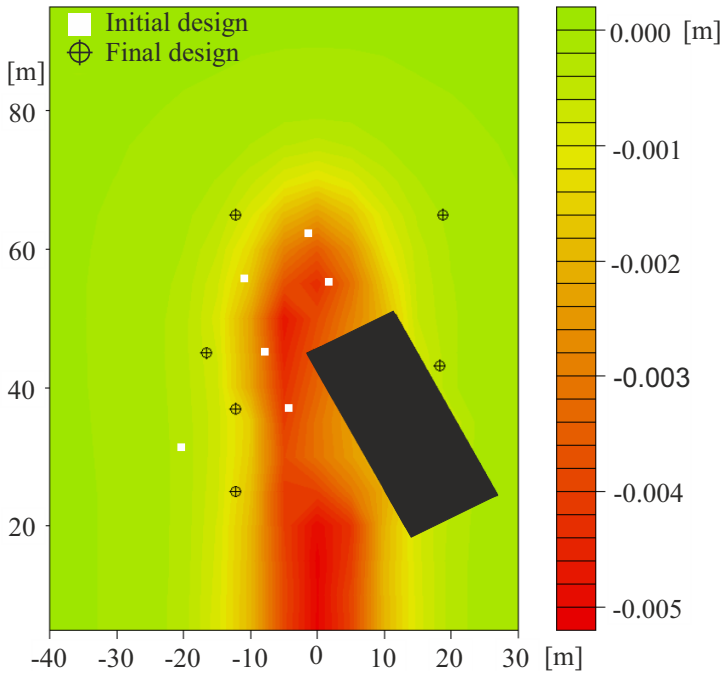


Figure 6.8.: Initial experimental design $\boldsymbol{\delta}_0$ and final experimental design $\boldsymbol{\delta}^*$ of the Bayesian OED mapped over displacement field of the ground surface. The position of the building is shown by the black colour.

settlement distribution might be wider or narrower. The obtained arrangement of sensors does not correspond to a certain set of parameters, but is supposed to be most reliable over the considered range of parameters. The locations closer to the edge of the settlement trough are in accordance with the observations made in the previous chapters that measurements might be most informative, if they are performed where the gradient of a model response is largest.

This corresponds also to results of a study performed preliminary to the one presented here. Thereby, instead of using a FE-model to calculate the surface settlements induced by tunnel excavation, the commonly known Peck equation, introduced in Peck (1969*a*) is employed that approximates the tunnelling induced settlement curve by a Gaußian curve given in Eq. 6.16:

$$S_y(x) = \frac{V_\mathrm{l}}{\sqrt{2\pi i}} \exp\left(-\frac{x^2}{2i_\mathrm{x}^2}\right) \tag{6.16}$$

where V_l describes the volume loss and i_x is the distance in the transversal direction, called x-direction herein, to the settlement trough's inflection point. This equation allows to relate shape and maximum of settlement with the volume loss. To calculate the settlements in longitudinal direction, an analogue equation is available, depending accordingly on i_z instead of i_x. Combining both equations, a three-dimensional settlement trough is obtained as displayed in Fig. 6.9c that depends on the location of the tunnel. In these equations, the width and height of the settlements trough directly correlate with V_l and i and these two parameters are assumed to be unknown and objective of the OED. The procedure is exactly the same as described above considering the FE model but using the analytical model instead. Three sensors are scheduled to provide measurements to identify V_l and three more sensors are used to identify i, whereby i_x is assumed to be equal to i_y. The results of this operation are shown in Fig. 6.10. In this greenfield example the red coloured area indicates maximum calculated settlement using the Peck equation, while the dark blue zones correspond to zero settlement. The white dots describe the initial experimental design provided to avoid non-convergence as described above. The red dots indeed are the points that are obtained as experimental design with highest utility. The three dots closer to the zone of largest settlement are those designated to identify V_l, while the three on the edge of the settlement trough are intended to be used to identify i. This result might be trivial and agrees with what could be expected. However, it confirms that the results obtained for the case employing the FE model and its metamodel are plausible. The FE model based application of Bayesian OED is described in a quite rudimentary and non-conceived manner in Hölter et al. (2019).

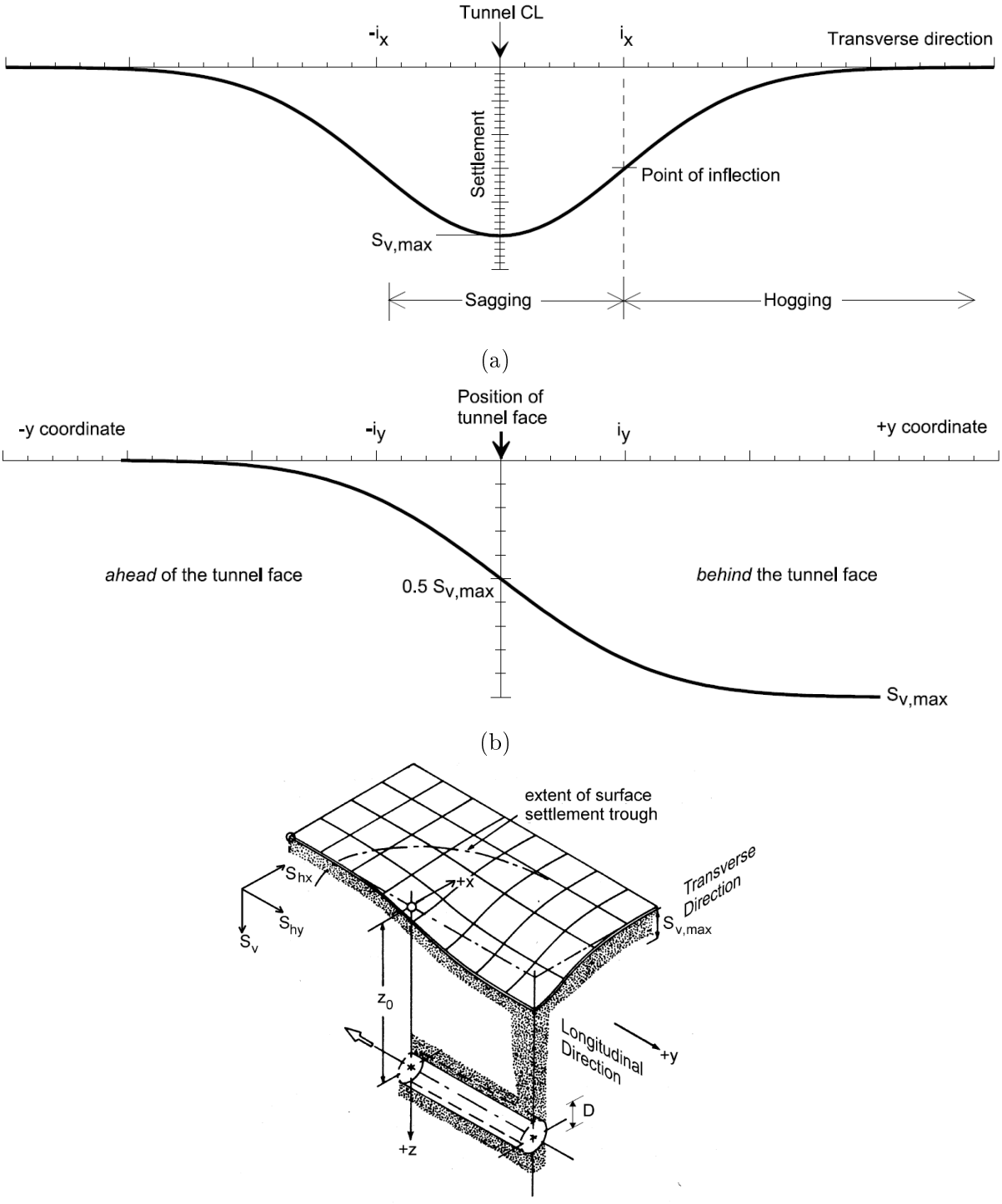


Figure 6.9.: Description of Bayesian OED in analytical application, a: transversal settlement curve, b: longitudinal settlement curve, c: resulting settlement trough in relation to tunnel position.

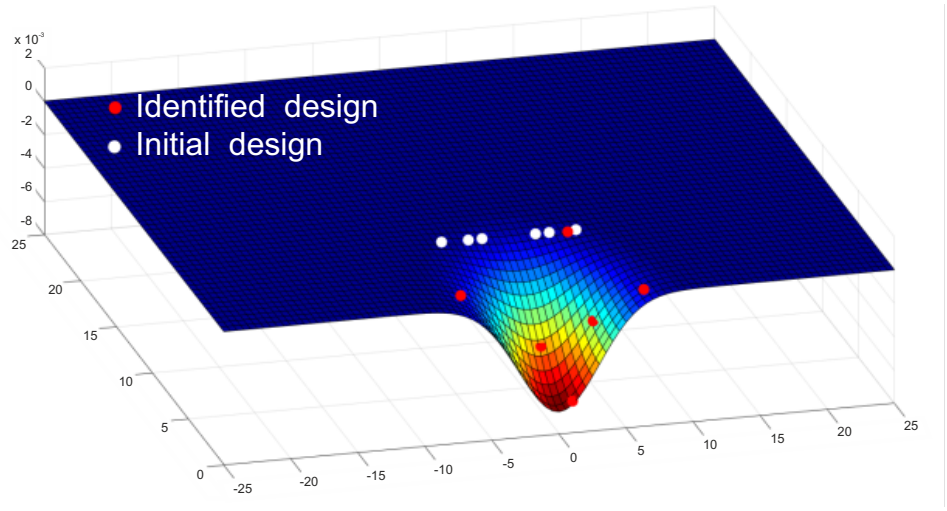


Figure 6.10.: Results of Bayesian OED application to analytical problem.

6.4. Summary

The concept of Bayesian OED introduced in this chapter shows to be a valuable alternative to the previously considered approaches and has been proven to be efficient in other scientific disciplines. Within this thesis, it is intended to adapt this approach for geotechnical applications. As one main aspect of OED in this field is the question where to locate measurement spots, the iterative and adaptive procedure of this method is convenient.
The fundamental concept of Bayesian OED is introduced that consist of maximising the utility U of a certain experimental design $\boldsymbol{\delta}$ to identify uncertain parameters $\boldsymbol{\theta}$ by means of noisy data $\tilde{\boldsymbol{y}}$. As the conditional probability of the constitutive parameters and the model responses are influenced by the employed experimental design, this experimental design can be optimised to maximise the utility of the Bayesian inference. As this utility is described by an integral that can hardly be solved (s.Eq. 6.2), the ACE algorithm is introduced to approximate this integral.
This concept is adapted for problems of geotechnical engineering by proposing how to handle the parameter distribution, the artificial noisy data, and the formulation of the utility. It is applied to the simulation of a real world case of a twin tunnel bypassing a nine-storey building. The objective of the investigation is finding the optimum position of sensors that record ground surface settlements in a manner that the parameters that influence most the building behaviour can be back-calculated most reliably. The results obtained in this example are comprehensible and in accordance with those from other approaches, especially with respect to the aspect of obtaining locations where the gradient of model responses is largest instead of such with large absolute values. As the calculation

time is acceptable, the Bayesian OED seems to be an at least equivalent alternative to the methods introduced previously.
However, several aspects are identified that might reduce the profit of this method: having an initial experimental design that allows the algorithm to proceed is crucial and might not always be as obvious to identify as in the present case. The high level of abstraction that is caused by using several metamodels and the approximation of the utility integral can induce a considerable model error. Furthermore, the results may depend on the initial assumptions of the parameter distributions. In the present case, the parameters are assumed to be normally distributed what is legitimate in case of the stiffnesses. With respect to the distribution of the volume loss, much less information is available. Detailed investigations on identifying the correct value of this parameter are performed as for example in Marshall et al. (2012) and databases are established like in Dimmock and Mair (2007), but to date by the best of the author's knowledge no recommendations are available on what probability distribution should be employed. However, the volume loss is not a conventional soil parameter. It results from the interaction of soil and TBM. As practitioners aspire to reduce this value as much as possible to avoid surface damage, this parameter is the objective of many research and development studies and therefore changing with the improvement of the TBMs. Therefore, the assumption of a bandwidth from 0.3 to 0.4 might anyway be inappropriate in near future.
The procedure shown in this example refers to the assumptions made, especially the considered number of and type of sensors, design space, and type of added noise. The obtained results are therefore only valid for the current application and may strongly vary if these variables are changed, even though the performed procedure would be the same.

7. Comparison of the employed methods

Within the three last chapters, different OED approaches are introduced and their application to geotechnical problems is demonstrated. The different approaches that are first OED by means of spatial GSA, second, OED using bootstrapping, and third, Bayesian OED, have been applied to different geotechnical problems, wherefore a qualitative comparison should be further undertaken. However, a quantitative comparison using the same experimental setup is not meaningful as the applicability of the particular OED approach is case sensitive.

7.1. Role of the measurement data diversity

Using only the vertical displacements as in case of the 3D tunnel simulation in Sec. 4.2.2.3 shows how problematic this can be in case of several parameters that have to be identified. The $S^*_{\mathrm{T}_{i,k}}$-values never indicate the parameter E_{50}^{ref} to be most relevant. Therefore, its identification is performed using data related to $S^*_{\mathrm{T}_{i,k}}$-values smaller than in case of the two other parameters $E_{\mathrm{ur}}^{\mathrm{ref}}$ and φ' and hence probably less reliable. In case of the dike in Sec. 4.2.3 this is even more drastic as the relevant parameter k_1 has almost no impact on the displacement data that is employed for parameter identification. On the other hand, in this example a similarly low sensitivity of the pore water pressure to the parameter φ'_1 is observed. Therefore, identifying in a reliable manner both parameters k_1 and φ'_1 is only possible using measurements related to different types of physical processes.
How many measurement points and which of these measurement points are necessary cannot be identified using OED by means of spatial GSA. To this, OED using bootstrapping provides the necessary additional information when performing the analysis for different sensor combinations as shown in Fig. 5.6. Here, one can see that the data obtained from the displacement sensors contribute more to the reliable overall identification of both

constitutive parameters than the data from the pore water pressure sensors. Moreover, installing more than five sensors does not have further positive impact.
Different insight, indeed, is obtained from the application example with the experimental device discussed in Sec. 4.2.1. Here, preliminary studies showed that it is possible to identify both parameters of interest, the permeability coefficient and the stiffness, by only one type of measurements. However, this is because both parameters are related to the same state parameter that is the void ratio e of the soil as indicated by Eq. 4.9. This shows that in case of coupled field problems, depending on the type of application and of employed constitutive model, using different measurement types might be essential or "only" an improvement of the reliability.
Independent of the investigated problem, the following statement should be considered as generally valid regarding where to place sensors: It is not most promising to place sensors where the model response of interest is the largest, but where its gradient is the largest. This is most comprehensive in case of the experimental testing device described in Sec. 4.2.1 when considering the problem where the pore water pressure measurements should be performed. The measurement data are not most informative directly below the plate where the pore water pressure has larger but constant values, but in a certain distance where most of the excess pore water pressure is dissipated and therefore, the gradient of the pore water pressure is largest. Moreover, similar observations are made in the dike example (Sec. 4.2.3) and in the application to mechanised tunnelling (Sec. 6.3.4). Identifying areas (in time as well as in space) where model responses strongly vary over short distances is therefore of importance in the context of OED.
It can be concluded that based on the results in this thesis, any further general suggestion would be too speculative and should depend on the specific application, wherefore the examples shown in this thesis are valuable references.

7.2. Computational efficiency

As mentioned several times, the calculation time strongly influences the applicability of the different approaches to OED. In the approaches to OED considered in this thesis, three features mainly influence the calculation time to complete the OED task: the model complexity, the dimensions of the parameter space, and the dimensions of the experimental design. However, these three features are not equally important for the different OED strategies. Therefore, it should be considered how the properties of the parameter

and design space as well as the model complexity will impact the different available OED approaches before selecting one of them. In certain cases, as shown in chapter 5, the preliminary evaluation can significantly reduce the computational effort. If the experimental design space only contains one parameter, e.g. time, a more complex OED approach might be computationally affordable than in the case of a complex problem of e.g. locating several sensors in a multidimensional design space.
In general, and this has been done in all OED applications presented here, performing a GSA is an important first step as it allows to identify the relevant parameters and to shrink the constitutive parameter space by excluding the less significant parameters. For performing afterwards the OED by means of spatial GSA, it is necessary to consider only the impact of these most relevant parameters on the model responses of interest. Therefore, a second set of parameter samples, containing only those relevant ones, must be generated and run in the model. Depending on the type of simulation, running the new set of samples in the simulation model can be highly time consuming even though the sample size could be smaller, as the parameter space has less dimensions. However, once all samples have been applied to the simulation model, only few further computational effort is required. Performing the spatial GSA, even on hundreds of grid points is a task of a few minutes. In case the grid should be refined, in FE-models additional points can be defined, and the output data is extracted in these points (assuming that the model results are stored) to re-perform the "densified" spatial GSA. In case few candidate experimental designs are scheduled from results of OED by means of spatial GSA and tested for comparison using artificial noisy data, this procedure is also operable within reasonable time limits. Hereby only these scheduled specific experimental designs are tested including a few hundreds runs of back analysis as in the example shown in Fig. 4.15b, instead of testing hundreds or thousands experimental designs as done in case of OED using bootstrapping or Bayesian OED.
In contrast, OED using bootstrapping is highly time consuming, but partially independent of the complexity of the employed FE-model as its computational effort mostly results from the required numerous back analysis runs. This already applies to the quite simple model considered in Sec. 5. For more complex models, their calculation time will become proportionally more relevant. In the problem studied by Schenkendorf et al. (2009), this approach could not be employed at all because of the required computation time, even though a computationally cheap analytical model was considered as the simulation model. To enable OED using bootstrapping at all, in this thesis an application of this approach is conducted whereby the sensor data is evaluated individually instead of cluster-wise. This simplification reduced considerably the calculation time to several hours instead of

weeks for the example considered herein. The generated matrix of identified parameters $\tilde{\mathfrak{T}}$ (see Eq. 5.4) allows to consider other configurations of the number of sensors without a new parameter identification process. This is performed in Sec. 5.2.3, where the number of sensors for displacements and pore water pressure is varied from one to three. Accordingly, OED using bootstrapping is well suited if such questions are relevant in a specific OED problem.

However, in the application example of the dike exposed to a rapid drawdown of the water level studied in this thesis, only eleven candidate points are included in the analysis for both, displacements and pore water pressure. The computational effort would grow exponentially in case the number of these points is increased. Therefore, employing the OED using bootstrapping is only reasonable in case only few candidate points are to be studied or if the design space can be reduced in advance. The application example of the dike shown in chapter 5 follows the latter approach by performing OED by means of spatial GSA in advance to OED using bootstrapping. Although a preliminary OED by means of spatial GSA requires a considerable calculation time, it is still more efficient than having a Bootstrap design space (number of candidate points) of double or triple size.

In the same chapter 5, the sigma-point method is considered that actually follows a similar concept as the OED using bootstrapping, but is way more efficient in generating the matrix $\tilde{\mathfrak{T}}$. Instead of generating a distribution of noisy measurement data that is back-calculated, this distribution is substituted by a few points whose number increases linearly with the dimension of the model parameter space. Accordingly, it is necessary to precisely know the characteristics of this distribution what can be an obstruction in case in-situ measurement data is employed. The sigma-point method can be seen as an efficient improvement of the OED using bootstrapping, but requires probably a deeper understanding and an even more careful handling.

The employed approach of Bayesian OED requires a reliable metamodel of good quality, wherefore numerous, possibly highly time consuming FE-model runs are necessary. For spatial interpolation, as performed in Sec. 6.3.4, no additional model runs are required, but additional model results data. This can be obtained by data extraction on a finer grid of locations in the already executed simulation runs. Like in the OED using bootstrapping, the actual procedure of identifying the OED by performing numerous runs of back-analysis causes large computational costs. Depending on the selected parameters of the employed ACE-algorithm (and of the DE-algorithm in the inner loop), one optimisation run for the problem investigated in Sec. 6.3.4 needs several days to converge. The considered design space herein is considerably larger than in the dike example. However, no spatial GSA is performed in advance to reduce the design space. The non-promising

parts of the design space that would have been excluded thereby, are though identified in the iterative process of the ACE-algorithm. Thus, one can conclude that performing Bayesian OED is more suitable in case large computational capacities are available but few human capacities for pre-processing data and interpreting results and vice-verse in case of combining OED with spatial GSA and bootstrapping, as here additional intermediate steps are required.
Besides, one should estimate which variation of the experimental design boundary conditions should be performed. Any variation like the number of sensors, the amount of added noise, or the candidate space requires an entire new optimisation run. In case of varying the number of sensors, the OED using bootstrapping (and the sigma point method, respectively) is advantageous for the reasons described above.

7.3. Informativity and reliability

An elementary aspect of the comparison of the different approaches is what type of information or results are provided and how, and for what further use, the obtained output can be employed.
The results that the OED by means of spatial GSA provides differ from those obtained by OED using bootstrapping and the Bayesian OED as employing spatial GSA indicates which data is most informative, while the latter two tell which specific design allows most accurate parameter identification. This apparently small difference should not be ignored. In case an experimental design is selected based on spatial GSA, only the model parameter uncertainty is considered and not the experimental design parameter variation. Accordingly, the influence of sensor number or accuracy is not part of the experimental design decision. Still, it is very probable that performing solely OED by means of spatial GSA, one selects a very good experimental design even though it might not be the best one. The resulting gain or loss of uncertainty might by smaller than the uncertainty resulting anyway from soil heterogeneity, model uncertainty, and measurement error that were described in Sec. 2.3.3. That means using OED by means of spatial GSA is suitable in case the experimental design components (e.g. number, type, and accuracy of measurement devices) are fixed a priori. If one is interested to make a decision on these aspects, this method is not recommended, while the method is well suitable in case one is interested in appropriate locations in time or space.
The advantage of OED using bootstrapping (and sigma-points) and of the Bayesian OED

is that specific experimental designs are applied to a considered OED problem and their results depend on both, the problem and the applied experimental design as well. Both approaches employ optimality criteria, that were introduced in Sec. 2.4.3, to evaluate the obtained parameter values with respect to the quality of the considered experimental design. However, as was demonstrated in Sec. 5.2.2 and according to the literature referred therein, applying these optimality criteria is not trivial but needs certain experience on the topic. Beside, one should apply different optimality criteria on the same specific application to compare the results and consider how they correlate with expectations based on engineering judgement.
The meticulous investigation and documentation of the considered experimental designs in case of OED using bootstrapping (and sigma points) method is rather time consuming, but it is of advantage when it comes to deeper interpretation of the results. The results for each specific experimental design that is tested are documented and can be evaluated by different optimality criteria, allowing a further detailed evaluation like the box plot diagrams shown in Fig. 5.6. Using the Bayesian OED and ACE algorithm, only the final experimental design and the utility of the best experimental design of each iteration $\boldsymbol{\delta}^{\dagger}$ are stored, making the results to a kind of black box. As outlined in the previous section, the Bayesian OED is therefore again more useful for one single evaluation that provides precise results to be employed without extensive additional interpretation. To perform detailed scientific analysis, the OED using bootstrapping (and sigma points) seems to be more suitable.

7.4. Summary

The comparisons and evaluations performed in this chapter intend to merge the knowledge gained in the previous chapters. Considering the aspects of different types of measurements, computational efficiency, and accuracy and reliability of results, it is demonstrated that depending on the specific problem and application, the introduced approaches to OED strongly differ in their performance. Obviously there is no universal optimum approach to OED and each application will require a preliminary study to make the decision what OED concept would be most appropriate. In addition, the introduced OED methods and the approaches to their assessment can be a useful platform to rely on in future investigations.

8. Conclusion

8.1. Main findings of present work

Introducing concepts of optimal experimental design to the field of geotechnical engineering is the overall purpose of this thesis. The methods of measurement based model validation were improved in the last years. However, it is identified that the question which data to use for such validation is rarely considered and hence the topic of OED comes up. Within this thesis, a methodology is developed to perform OED fully adapted to the circumstances encountered in geotechnical engineering.
Previous studies within the research project "System and Parameter Identification Methods for Ground Models in Mechanized Tunneling" in whose framework this thesis is embedded provide a methodological background to this research. In this thesis, the concepts of sensitivity analysis, metamodelling, and back analysis are adopted and extended by the aspect of not considering measurement data as given, but as an arbitrary variable that can and should be selected in an optimal way to improve the efficiency of the geotechnical model identification and validation.
Three approaches, namely the spatial GSA (i.e. to perform GSA in spatially distributed manner), Bootstrap analysis, and Bayesian OED, are considered to identify an OED, whereby operating these three methods cause an increasing computational effort. Accordingly, reducing the dimension of the parameter space and of the design space is undertaken as the first step of successfully identifying the OED. In connection to this, a detailed analysis is presented in the thesis on the model parameter sensitivity analysis and especially how the global sensitivity analysis can be employed to identify the relevant (constitutive) model parameters on which should be focused. A modified sensitivity index $S^*_{T_i}$ is proposed that allows to compare the parameter sensitivity independently of the different model output dimensions. By applying the modified sensitivity index $S^*_{T_i}$ to the sensitivity analysis of three different simulation examples, several key finding are achieved:

- The sensitivity index $S^*_{T_i}$ allows to identify in which time intervals and in which special areas in the considered FE model a certain model parameter can be identified most reliably.
- By performing an extensive model calibration, obtaining a higher agreement between experimental and simulation results is enabled. The residual misfit between model and measurement data that remains after model calibration is identified to result from the insufficiently adequate model. Therefore, the optimal experimental design cannot compensate for a poor model.
- It is shown that specific optimal experimental designs can be determined employing $S^*_{T_i}$. By using the experimental design established based on sensitivity plots, more reliable results of the parameter identification procedure are obtained. However, it is indicated that such experimental designs may not account sufficiently and appropriately for the impact of measurement error.
- As the computational effort to calculate $S^*_{T_i}$ depends only on the number of simulation runs and not on the number of considered observation points in the geometrical model, it is also employed as preliminary step to reduce the range of the parameter space $\boldsymbol{\Pi}$ in which the optimal experimental design $\boldsymbol{\delta}^*$ has to be identified.

As second approach, the Bootstrap and the sigma-point methods are employed to perform statistical evaluation, whereby detailed investigation of individual experimental designs is enabled. Hereby, any possible combination of sensors on a predefined set of candidate locations is considered. The quality of a measurement design program is identified by applying an optimality criterion to the covariance matrix that reflects the uncertainty of the identified parameters. This method is very versatile as the procedure could be repeated for different types and measures of error and different sensor types. Within this thesis, the variation of number and types of sensors is undertaken, leading to relevant information about how additional sensors contribute to a further improvement of the model validation. Using the Bootstrap concept allows to artificially increase this database, what enables to obtain probability distributions in case of insufficient in-situ observations. However, the Bootstrap approach demands large computational efforts as each time one experimental design is considered, a whole distribution of back-calculated results needs to be generated. By the performed studies, the following insights are gained :

- Like in the referred to publications in Sec. 2.4.3, it is observed that different optimality criteria may lead to different optimal experimental designs. It is therefore suggested to actually employ several of them and scrutinise the obtained results.

- The sigma point method is a valuable approach to boost the efficiency of the Bootstrap approach. In the considered application, using the sigma point method or not leads to identified experimental designs of same quality. However, due to the necessity to know the actual distribution of the considered measurement data, this approach is restricted to artificial data (or very extensive measurement data sets).
- The Bootstrap approach is presumed to be the most accurate of the considered OED methods. However, it is highly time consuming, wherefore it is applied to an example with a small design space. To reduce the computational effort, performing the spatial GSA approach in advance to reduce the design space is successfully applied.

To consider even large model geometries, i.e. large experimental design spaces, the concept of Bayesian OED is applied as third approach to OED in this thesis. At first glance, several positive aspects of this approach are identified, as it allows a grid independent solution, the optimisation is more efficient as it does not test all over the design space, and it allows to account for the uncertainty type of the model parameters. However, the aspect of model uncertainty is identified as potentially affecting the optimisation process: first, a metamodel is generated to substitute the time-consuming FE-calculation and second, an interpolation among the metamodel's support points is done to enable the finding of solutions also between the support points. As the accumulation of the potential model errors resulting from the different regressions is unseizable, the obtained results are considered as potentially uncertain. Therefore, one should employ this approach very carefully and keep in mind the following key aspects:

- An experimental design that is comprehensive according to the findings of the spatial GSA and Bootstrap approaches is identified using the suggested Bayesian OED approach.
- The required computational effort is smaller in the considered example than e.g. in the Bootstrap application while the design space is considerably larger.
- A rough knowledge on suitable experimental designs is necessary to provide an adequate initial guess for the experimental design. In case this "initial design" is set up completely random and located in non-sensitive areas, the ACE algorithm is not converging to an optimal solution.
- The nature of the geotechnical problems requires to take into account a high level of model uncertainty. For this reason, the employment of Bayesian OED in context of FE models is critically discussed and the results scrutinised. The combination with

other methods like GSA, or spatial sensitivity allows an independent comparison of the results or to limit the boundaries of the optimisation problem.

The insights obtained in this thesis allow to develop concepts for solving the OED problem in geotechnical projects. Depending on the specific application as well as the available time and budget, different combinations or individual concepts introduced in this thesis might be most promising. However, even though fundamental ideas are presented herein, further research should be performed on this topic whose major objectives are formulated in the following.

8.2. Outlook

Two major aspects should be considered in future works on the considered topic of OED in geotechnical engineering. First, the extension of the presented methods to further and more complex applications and second, the practical validation, whereby both aspects might by considered in same projects. There are further aspects that have played a minor or negligible role within this thesis such as to be mentioned the soil variability, the higher order moments of the measurement error, and variation of the amount of error. Deeper research on each of these topics would be desirable. These could be individual extensions or continuations of the studies presented herein, where the investigations are repeated with varying types of error, or by substituting the homogeneous soil layers in one of the FE models by random field clusters as performed in Mahmoudi et al. (2020).
Regarding the two main topics for future work, it should be mentioned that the examples considered within this thesis, especially those in chapters 4 and 5 are rather simple ones. The dike simulation e.g. only considered one case of water drawdown wherefore the identified experimental design is optimal for that case, but not necessarily for other drawdown scenarios. To consider this aspect, the decrease of the water level could have been considered as additional random variable in the simulation. However, this would have made the simulation more complex and, even more problematic, the results less comprehensible. In the current application, the relationships of water drawdown, the resulting failure mechanism, and the identified positions of sensors make sense and enable the reader to understand the OED approach.
Considering more complex applications would allow more meaningful results. Within the considered applications, the discrepancies of the optimality criteria showed only few variation within the considered experimental design parameter space. Investigating the case e.g. of a large slope, like the one described in Li et al. (2019), including several soil lay-

ers and an alternating ground surface that is subjected to fluctuating meteorological and hydraulic impacts would require extensive modelling effort, but differences between good and poor experimental results would become more demonstrative. To perform such investigations, the experiences gained in the framework of this thesis allow the selection of an adequate methodology to ensure an efficient handling even of such a complex application. To transfer the introduced concepts to engineering practice, real case applications are necessary to prove their functionality and the measurable cost savings that can be achieved considering them. Project costs will be higher with a detailed planning of an experimental design than without. Besides, these techniques only make sense if during the construction phase (or even more the operation phase for e.g. a dike) a continuous validation is practised. This is only affordable in context of large projects. In such applications (like tunnels, dams, deep excavation pits, etc.) the potential of saving costs would be high and the expenses for the "numerical assistance" relatively low. However, such applications are not feasible in context of pure research projects, but require interaction of science and practice. Therefore, persuading practitioners of the advantages of implementing OED to their projects is a major next step to take.

A. Sensitivity contour plots of sample testing experiment

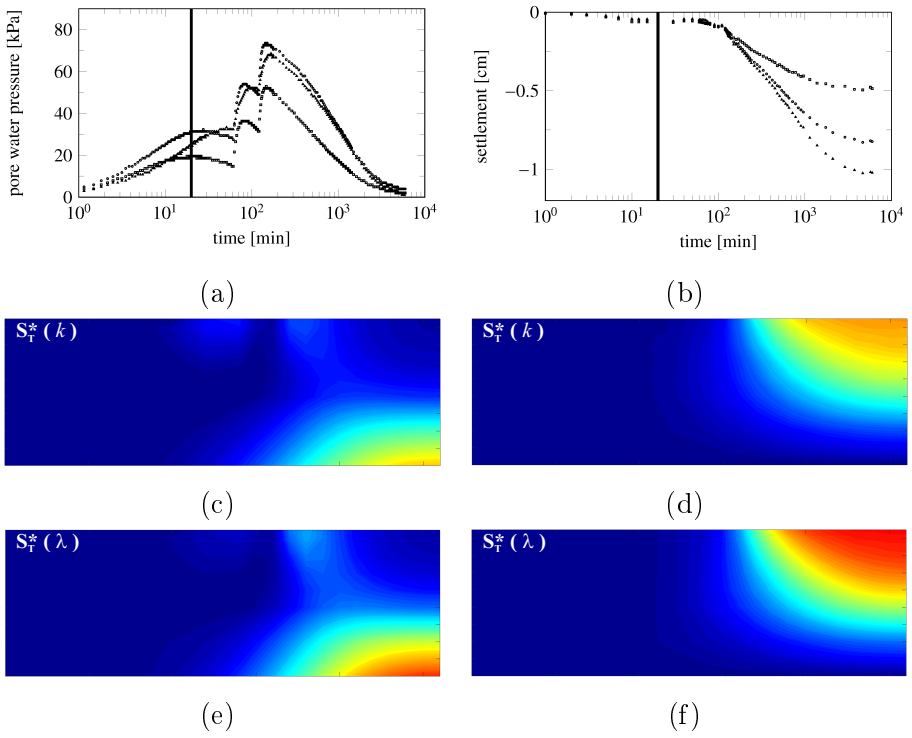


Figure A.1.: Vertical bar pointing out the considered time step after 20 minutes with respect to a pore water pressure y_w and b settlement y_y evolution; sensitivity distribution for y_w towards c: k and e: λ, and for y_y towards d: k and f: λ

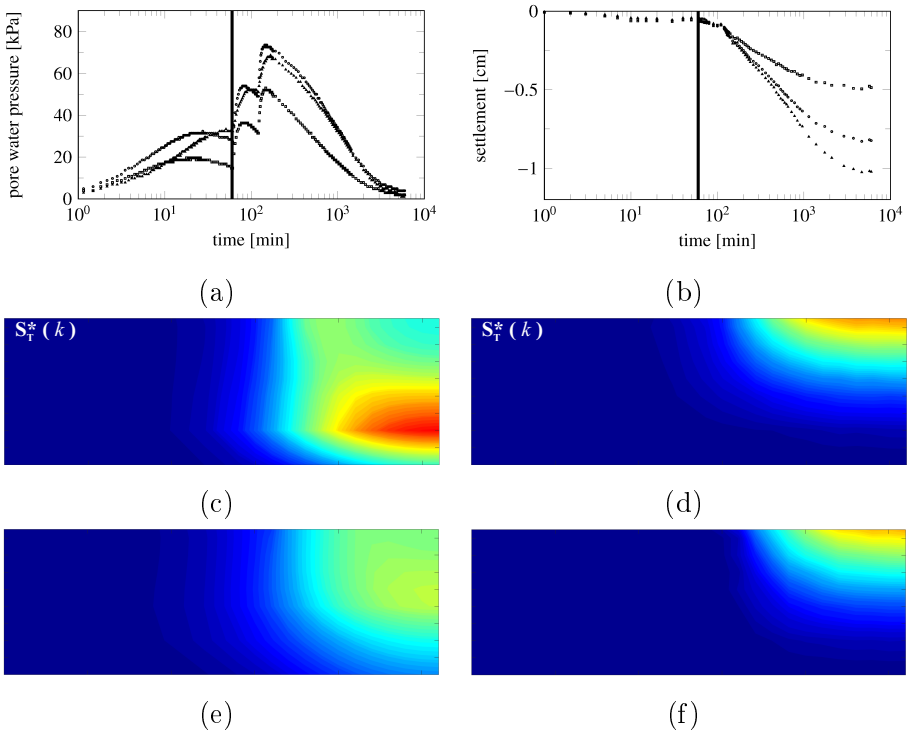


Figure A.2.: Vertical bar pointing out the considered time step after 60 minutes with respect to a pore water pressure y_w and b settlement y_y evolution; sensitivity distribution for y_w towards c: k and e: λ, and for y_y towards d: k and f: λ

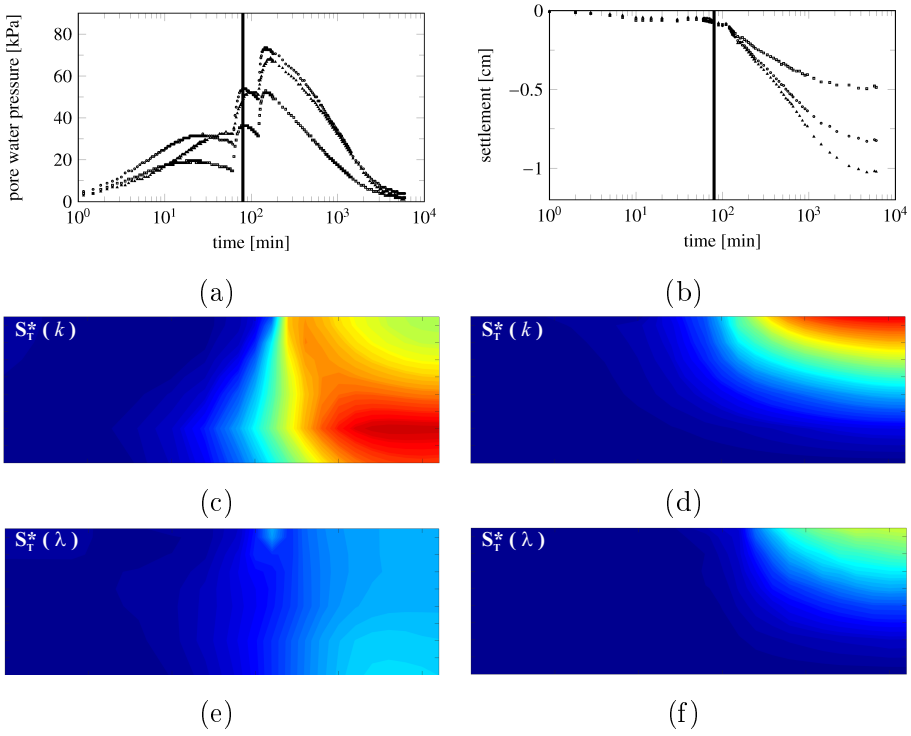


Figure A.3.: Vertical bar pointing out the considered time step after 80 minutes with respect to a pore water pressure y_w and b settlement y_y evolution; sensitivity distribution for y_w towards c: k and e: λ, and for y_y towards d: k and f: λ

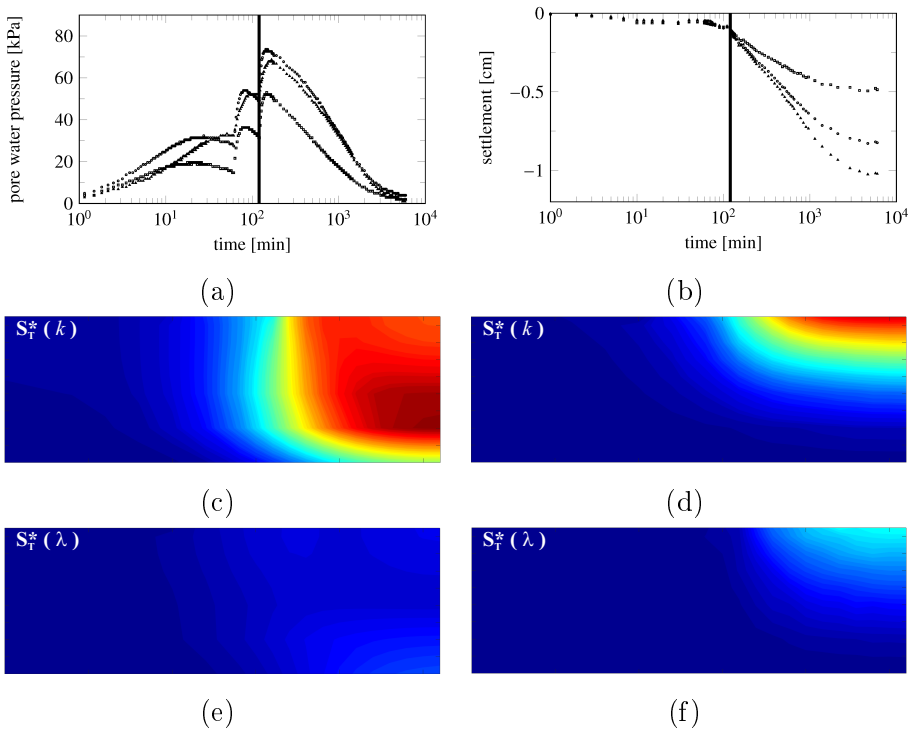


Figure A.4.: Vertical bar pointing out the considered time step after 120 minutes with respect to a pore water pressure y_w and b settlement y_y evolution; sensitivity distribution for y_w towards c: k and e: λ, and for y_y towards d: k and f: λ

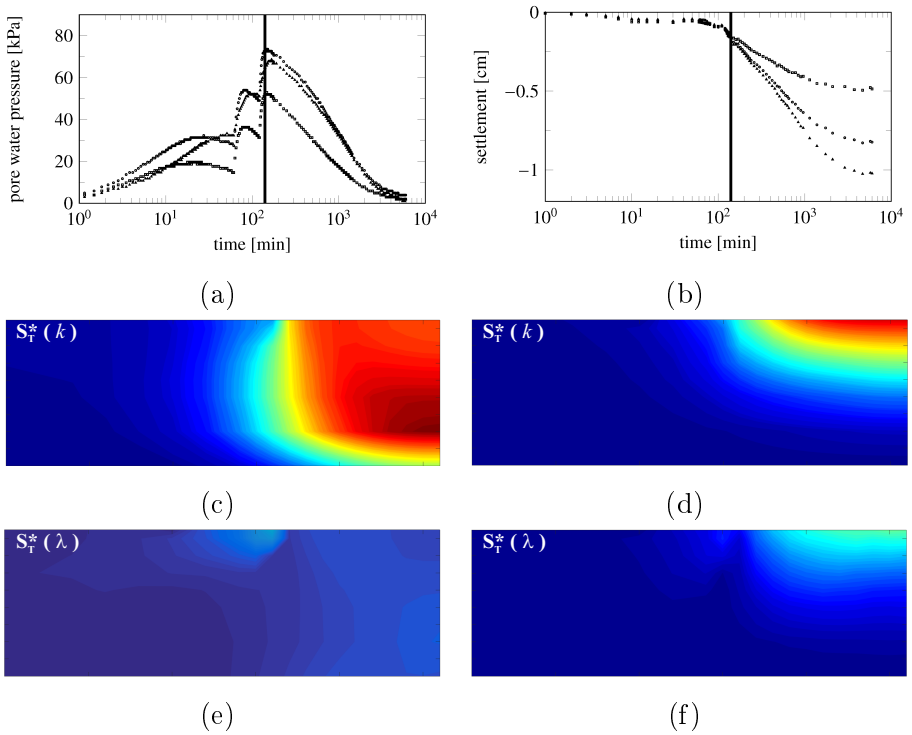


Figure A.5.: Vertical bar pointing out the considered time step after 140 minutes with respect to a pore water pressure y_w and b settlement y_y evolution; sensitivity distribution for y_w towards c: k and e: λ, and for y_y towards d: k and f: λ

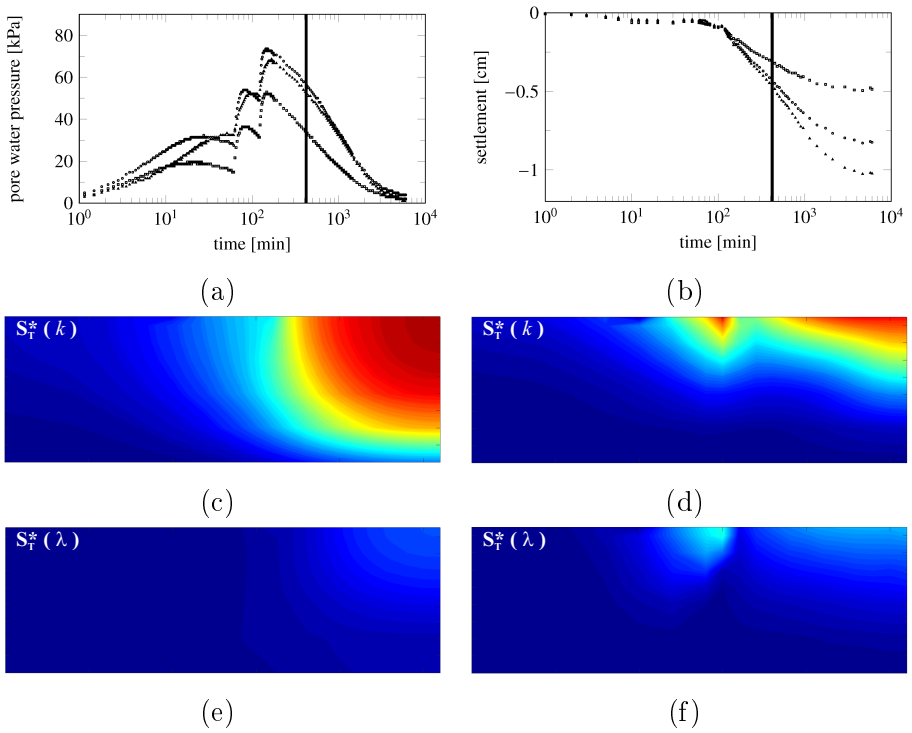


Figure A.6.: Vertical bar pointing out the considered time step after 420 minutes with respect to a pore water pressure y_w and b settlement y_y evolution; sensitivity distribution for y_w towards c: k and e: λ, and for y_y towards d: k and f: λ

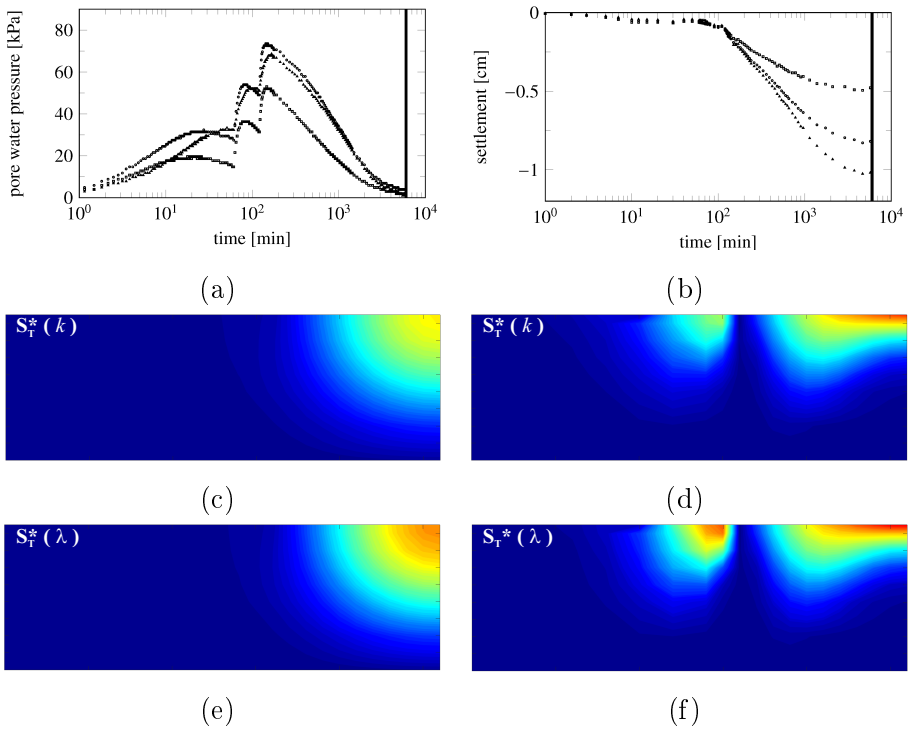


Figure A.7.: Vertical bar pointing out the considered time step after 6000 minutes with respect to a pore water pressure y_w and b settlement y_y evolution; sensitivity distribution for y_w towards c: k and e: λ, and for y_y towards d: k and f: λ

B. Sensitivity contour plots of TBM advancement

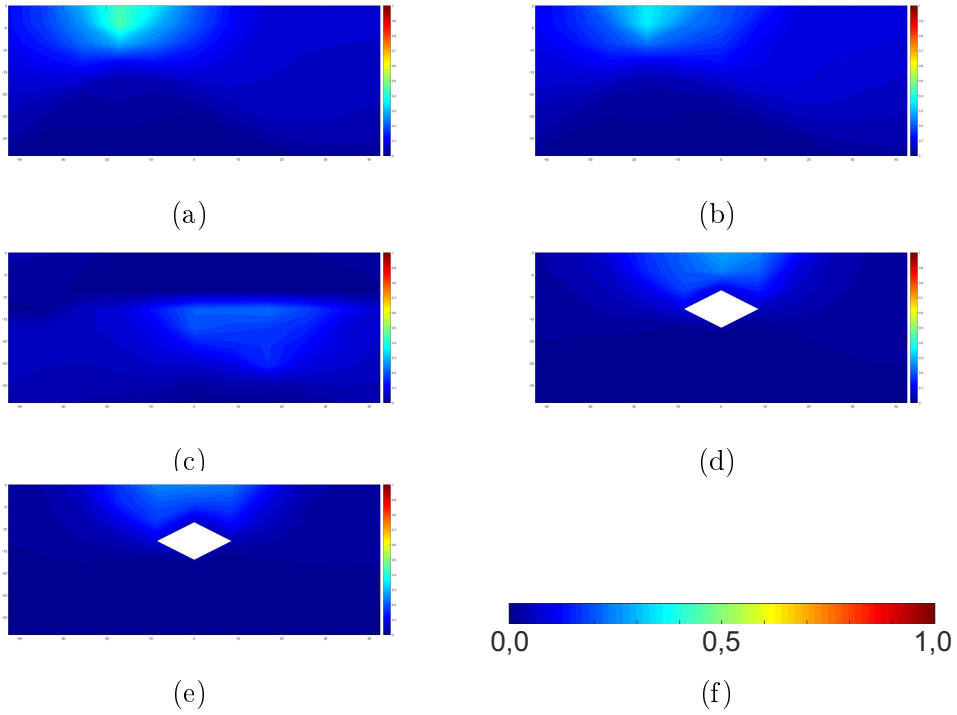
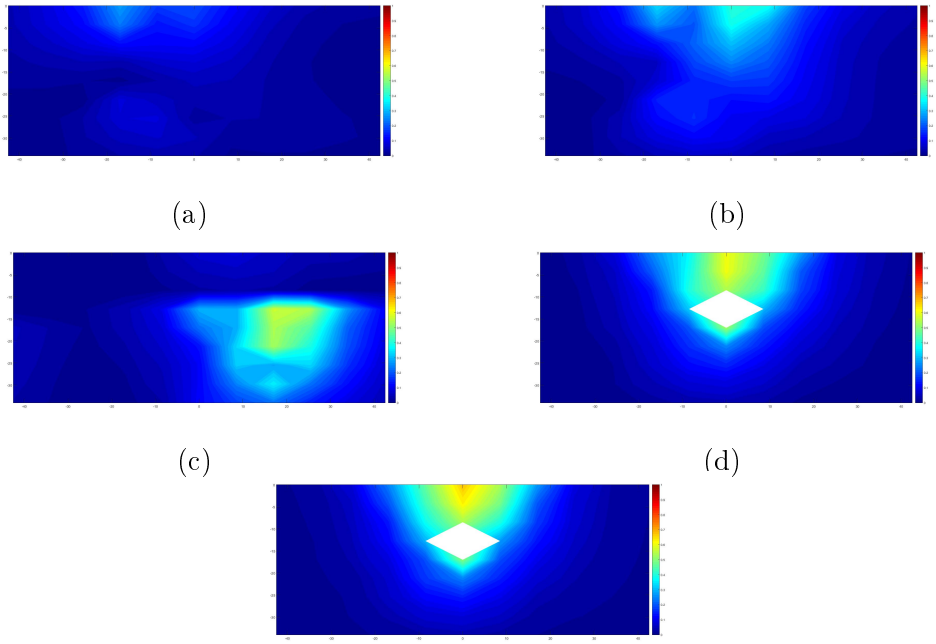
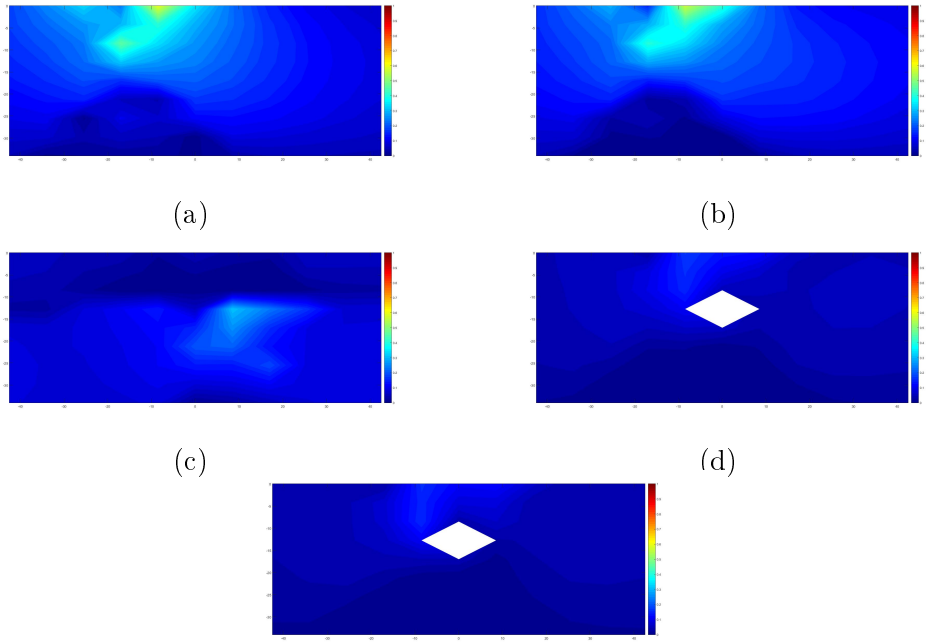


Figure B.1.: Contourplots of GSA with respect to E_{50} in section three in excavation phase (a) 13, (b) 19, (c) 25, (d) 31, and (e) 37 and scale bar of $S^*_{\mathrm{T}_{i,k}}$ (f)



(a) (b) (c) (d) (e)

Figure B.2.: Contourplots of GSA with respect to E_{ur} in section three in excavation phase (a) 13, (b) 19, (c) 25, (d) 31, and (e) 37



(a) (b) (c) (d) (e)

Figure B.3.: Contourplots of GSA with respect to φ' in section three in excavation phase (a) 13, (b) 19, (c) 25, (d) 31, and (e) 37

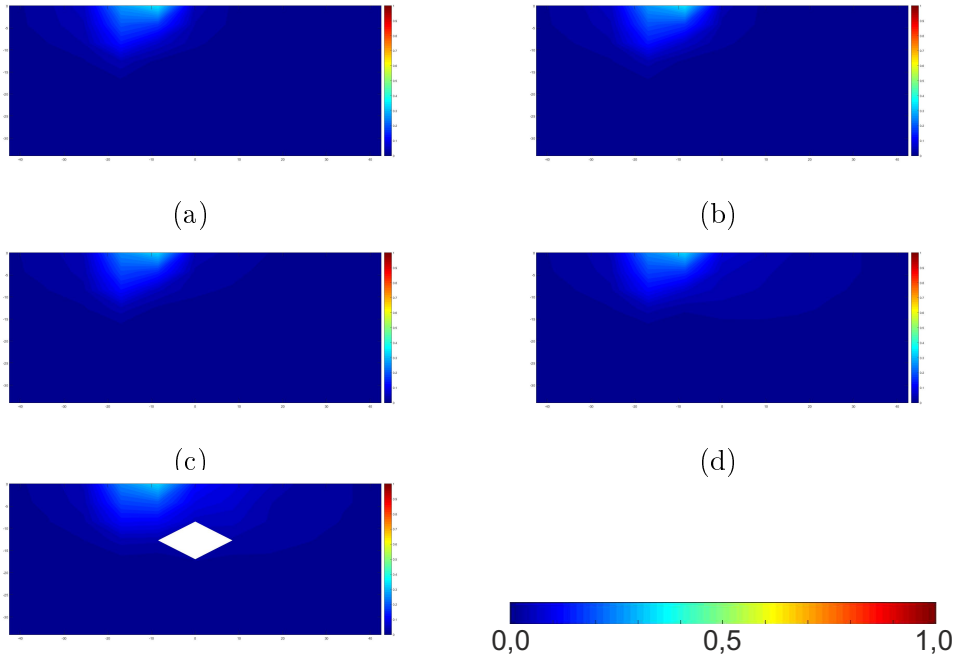


Figure B.4.: Contourplots of GSA with respect to E_{50} in section four in excavation phase (a) 13, (b) 19, (c) 25, (d) 31, and (e) 37 and scale bar of $S^*_{\mathrm{T}_{i,k}}$ (f)

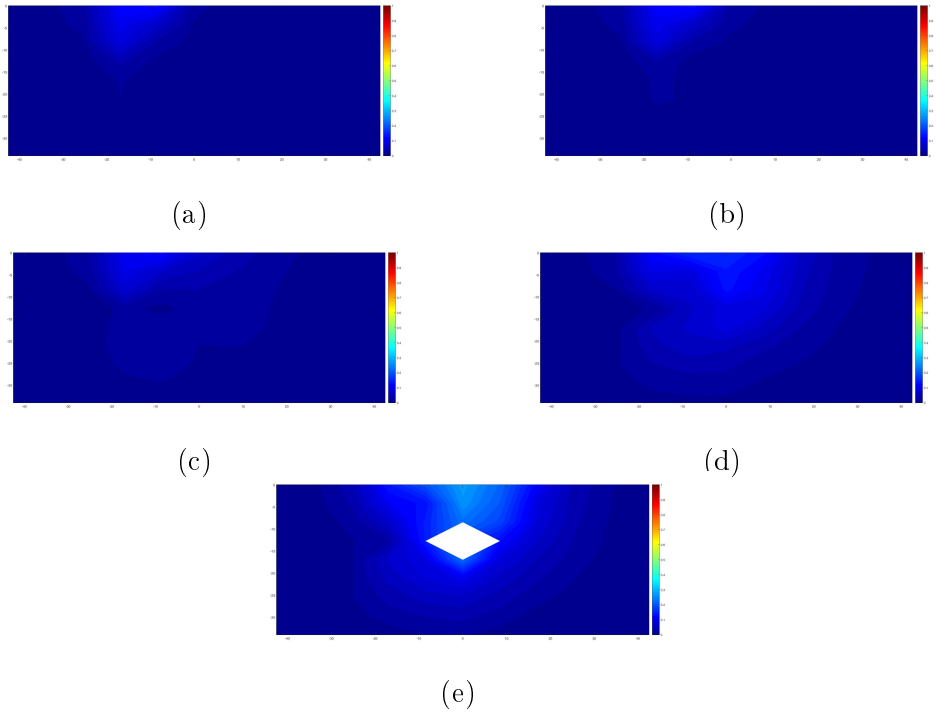
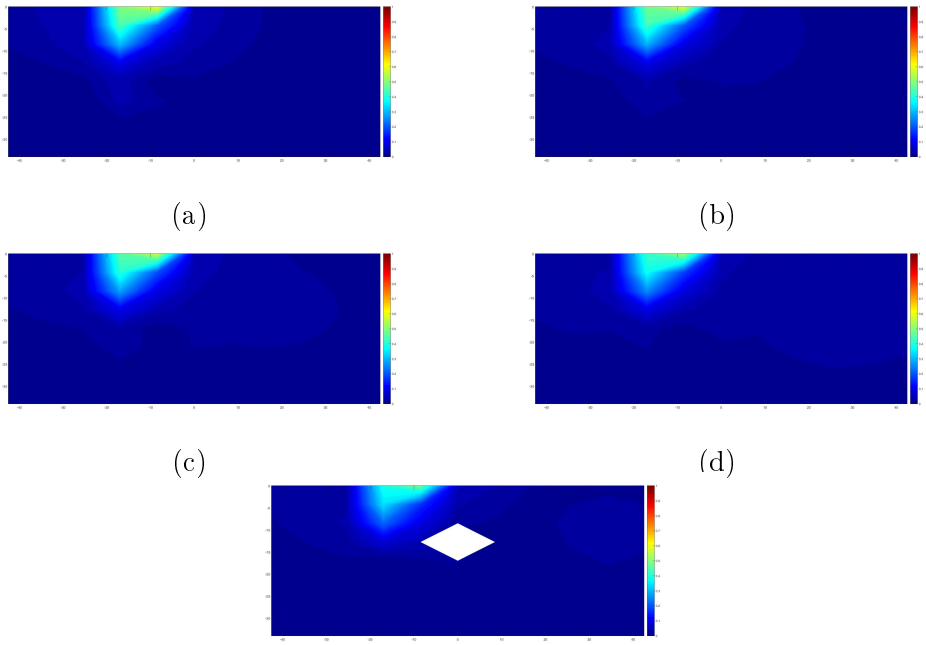


Figure B.5.: Contourplots of GSA with respect to E_{ur} in section four in excavation phase (a) 13, (b) 19, (c) 25, (d) 31, and (e) 37



(a) (b)

(c) (d)

(e)

Figure B.6.: Contourplots of GSA with respect to φ' in section four in excavation phase (a) 13, (b) 19, (c) 25, (d) 31, and (e) 37

Bibliography

Akima, H. (1978), 'A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points', *ACM Transactions on Mathematical. Software* **4**(2), 148–159.

Alkam, F. and Lahmer, T. (2019), 'Quantifying the uncertainty of identified parameters of prestressed concrete poles using the experimental measurements and different optimization methods', *Engineering and Applied Sciences* **4**(4), 84–92.

Alonso, E. E. and Pinyol, N. M. (2016), 'Numerical analysis of rapid drawdown: Applications in real cases', *Water Science and Engineering* **9**(3), 175 – 182.

Amaran, S., Sahinidis, N. V., Sharda, B. and Bury, S. J. (2016), 'Simulation optimization: a review of algorithms and applications', *Annals of Operations Research* **240**(1), 351–380.

Ardia, D., Boudt, K., Carl, P., Mullenl, K. and Peterson, B. G. (2011), 'Differential evolution with DEoptim: An application to non-convex portfolio optimization', *The R journal* **3**(1), 27–34.

Atamturktur, S., Egeberg, M. C., Hemez, F. M. and Stevens, G. N. (2015), 'Defining coverage of an operational domain using a modified nearest-neighbor metric', *Mechanical Systems and Signal Processing* **50-51**, 349 – 361.

Azam, S. E., Ghisi, A. and Mariani, S. (2012), 'Parallelized sigma-point Kalman filtering for structural dynamics', *Computers and Structures* **92-93**, 193 – 205.

Balci, O. (1998), Verification, validation, and accreditation, *in* 'Proceedings of the 30th Conference on Winter Simulation', WSC '98, IEEE Computer Society Press, Los Alamitos, CA, USA, pp. 41–4.

Bandemer, H. and Bellmann, A. (1979), *Statistische Versuchsplanung*, Vol. 19.

Barciaga, T. (2022), Adaptive Constitutive Modeling of Soil with Special Consideration of Destructuration [in preparation], PhD thesis, Ruhr-Universität Bochum, Bochum, Germany.

Bayes, T. and Price, R. (1763), 'An essay towards solving a problem in the doctrine of chances', *Philosophical Transactions of the Royal Society of London* **53**, 370–418.

Behera, S., Sahoo, S. and Pati, B. (2015), 'A review on optimization algorithms and application to wind energy integration to grid', *Renewable and Sustainable Energy Reviews* **48**, 214 – 227.

Benz, T. (2007), Small-Strain Stiffness of Soils and its Numerical Consequences, PhD thesis, Universität Stuttgart, Stuttgart, Germany.

Blatman, G. and Sudret, B. (2010), 'An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis', *Probabilistic Engineering Mechanics* **25**(2), 183 – 197.

Blom, C. (2002), Design philosophy of concrete linings for tunnels in soft soils, PhD thesis, Delft University of Technology, Delft, The Netherlands.

Bourdeau, P. L. and Amundaray, J. I. (2005), 'Non-parametric simulation of geotechnical variability', *Géotechnique* **55**(2), 95–108.

Brinkgreve, R. (2013), Validating geotechnical finite element models, *in* S. Pietruszczak and G. Pande, eds, 'COMGEO III: proceedings of the 3rd International Symposium on Computational Geomechanics', pp. 292–304.

Buljak, V. (2012), *Inverse Analysis with Model Reduction: Proper Orthogonal Decomposition in Structural Mechanics*, Springer Verlag, Berlin, Heidelberg, Germany.

Cao, B.-T., Freitag, S. and Meschke, G. (2016), 'A hybrid RNN-GPOD surrogate model for real-time settlement predictions in mechanised tunnelling', *Advanced Modeling and Simulation in Engineering Sciences* **3**(5).

Carson, II, J. S. (2002), Verification validation: Model verification and validation, *in* 'Proceedings of the 34th Conference on Winter Simulation: Exploring New Frontiers', WSC '02, Winter Simulation Conference, pp. 52–58.

Chaloner, K. and Verdinelli, I. (1995), 'Bayesian experimental design: A review', *Statistical Science* **10**(3), 273–304.

Cukier, R. I., Fortuin, C. M., Shuler, K. E., Petschek, A. G. and Schaibly, J. H. (1973), 'Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory', *The Journal of Chemical Physics* **59**(8), 3873–3878.

Dantzig, G. B. (1987), Origins of the simplex method, Technical report, Stanford University, Stanford, CA, Technical report.

Das, B. M. (2006), *Principles of geotechnical Engineering*, 7th edn, Cengage learning.

Detert, O., König, D., van Duijnen, P., Lavasan, A. A., Hölter, R. and van Eekelen, S. (2020), 'Zum Tragverhalten von mit Geogittern gehaltenen Spundwänden', *Bautechnik* (9).

Dhiman, G. and Kumar, V. (2017), 'Spotted hyena optimizer: A novel bio-inspired based metaheuristic technique for engineering applications', *Advances in Engineering Software* **114**, 48 – 70.

Dimmock, P. S. and Mair, R. J. (2007), 'Estimating volume loss for open-face tunnels in london clay', *Proceedings of the Institution of Civil Engineers - Geotechnical Engineering* **160**(1), 13–22.

Dimov, I. and Georgieva, R. (2010), 'Monte Carlo algorithms for evaluating Sobol' sensitivity indices', *Mathematics and Computers in Simulation* **81**(3), 506 – 514.

DIN 18196 (2011-05), 'Earthworks and foundations - Soil classification for civil engineering purposes'.

DIN EN 1990 (2010-12), 'Eurocode: Basis of structural design'.

DIN EN 1997-1 (2014-03), 'Eurocode 7: Geotechnical design - Part 1: General rules'.

DIN EN ISO 14688 (2018-05), 'Geotechnical investigation and testing – Identification and classification of soil'.

Dubourg, V., Sudret, B. and Bourinet, J.-M. (2011), 'Reliability-based design optimization using kriging surrogates and subset simulation', *Structural and Multidisciplinary Optimization* **44**(5), 673–690.

Efron, B. (1979), 'Bootstrap methods: Another Look at the Jackknife', *The Annals of Statistics* **7**(1), 1–26.

Efron, B. and Tibshirani, R. (1993), *An Introduction to the Bootstrap*, Chapman and Hall, New York, NY.

El-Ramly, H., Morgenstern, N. R. and Cruden, D. M. (2003), 'Probabilistic stability analysis of a tailings dyke on presheared clay-shale', *Canadian Geotechnical Journal* **40**(1), 192–208.

Fargnoli, V., Boldini, D. and Amorosi, A. (2015b), 'Twin tunnel excavation in coarse grained soils: Observations and numerical back-predictions under free field conditions and in presence of a surface structure', *Tunnelling and Underground Space Technology* **49**, 454 – 469.

Fargnoli, V., Gragnano, C. G., Boldini, D. and Amorosi, A. (2015a), '3D numerical modelling of soil structure interaction during EPB tunnelling', *Géotechnique* **65**(1), 23–37.

Ferretti, F., Saltelli, A. and Tarantola, S. (2016), 'Trends in sensitivity analysis practice in the last decade', *Science of The Total Environment* **568**, 666–670.

Fidanova, S., Roeva, O. and Ganzha, M. (2013), *ACO and GA for Parameter Settings of E. coli Fed-Batch Cultivation Model*, Springer International Publishing, Heidelberg, pp. 51–71.

Fisher, R. (1935), *The Design of Experiments*, Oliver and Boyd, Edinburgh.

Franzius, J., Potts, D., Addenbrooke, T. and Burland, J. (2004), 'The influence of building weight on tunnelling-induced ground and building deformation', *Soils and Foundations* **44**(1), 25–38.

Freitag, S., Cao, B., Ninić, J. and Meschke, G. (2018), 'Recurrent neural networks and proper orthogonal decomposition with interval data for real-time predictions of mechanised tunnelling processes', *Computers & Structures* **207**, 258 – 273.

Gall, V. E., Marwan, A., Smarslik, M., Obel, M., Mark, P. and Meschke, G. (2018), 'A holistic approach for the investigation of lining response to mechanized tunneling induced construction loadings', *Underground Space* **3**(1), 45 – 60. Computational Methods in Mechanized Tunneling.

Gambardella, L. and Dorigo, M. (1996), Solving Symmetric and Asymmetric TSPs by Ant Colonies, *in* 'IEEE Conference on Evolutionary Computation, Nagoya, Japan', pp. 622–627.

Giunta, A., Wojtkiewicz, S. and Eldred, M. (2003), Overview of modern design of experiments methods for computational simulations, *in* 'Proceedings of the 41st AIAA Aerospace Sciences Meeting and Exhibit,', pp. AIAA–2003–0649.

Glöztl GmbH (2020), 'Data sheet GWD 20 G'.
URL: *http://www.gloetzl.de/produkte/messwertaufnehmer/setzung-und-hebung.htm*

Gomes, H. M., Awruch, A. M. and Lopes, P. A. M. (2011), 'Reliability based optimization of laminated composite structures using genetic algorithms and artificial neural networks', *Structural Safety* **33**(3), 186 – 195.

Gupta, S., Degrande, G. and Lombaert, G. (2008), Experimental validation of a numerical model for subway induced vibrations, *in* B. Schulte-Werning, D. Thompson, P.-E. Gautier, C. Hanson, B. Hemsworth, J. Nelson, T. Maeda and P. de Vos, eds, 'Noise and Vibration Mitigation for Rail Transportation Systems', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 108–114.

Hasofer, A. M. and Lind, N. C. (1974), 'Exact and invariant second moment code forma', *Journal of the Engineering Mechanics Division* **100**(1), 111–121.

Herrenknecht AG (2020), 'Mixshield tunnel boring machine'.
URL: *https://www.herrenknecht.com/de/produkte/productdetail/mixschild/*

Hill, M. (1998), Methods and guidelines for effecitve model calibration, Technical report, U.S. Geological Survey Water-Resources Investigations Report.

Hill, M. C. and Tiedeman, C. R. (2005), *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty*, John Wiley and Sons, Ltd.

Holland, J. (1975), *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, University of Michigan Press, Ann Arbor.

Hölter, R., Lieske, W., Müthing, N., Schmüdderich, C., Zhao, C. and Schanz, T. (2016), Settlement predicion for an embankment on soft, structured clay using smeared soil-drain replacement technique and probabillistic analysis, *in* S. W. Sloan and R. B. Kelly, eds, 'Embankment and Footing Prediction Symposium', ARC Centre of Excellence for Geotechnical Science and Engineering, pp. CGSE EFPS–009.

Hölter, R., Mahmoudi, E., Datcheva, M. and Schanz, T. (2018b), 'Using sigma-points to identify optimal experimental design for dike monitoring', *IFAC-PapersOnLine* **51**(2), 759 – 764. 9th Vienna International Conference on Mathematical Modelling.

Hölter, R., Mahmoudi, E., Rose, S., König, M., Datcheva, M. and Schanz, T. (2019a), 'Employment of the bootstrap method for optimal sensor location considering uncertainties in a coupled hydro-mechanical application', *Applied Soft Computing* **75**, 298 – 309.

Hölter, R., Mahmoudi, E. and Schanz, T. (2015), Optimal sensor location for parameter identification in soft clay, *in* 'Application of Mathematics in Technical and Natural Sciences, Albena, Bulgaria', p. 030005.

Hölter, R., Schoen, M., Lavasan, A. A. and Mahmoudi, E. (2019), Model validation using Bayesian optimal experimental design in urban mechanised tunnelling, *in* '3rd UNCECOMP 2019 ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering'.

Hölter, R., Zhao, C., Mahmoudi, E., Lavasan, A. A., Datcheva, M., König, M. and Schanz, T. (2018a), 'Optimal measurement design for parameter identification in mechanized tunneling', *Underground Space* **3**(1), 34 – 44.

Huan, X. and Marzouk, Y. M. (2013), 'Simulation-based optimal Bayesian experimental design for nonlinear systems', *Journal of Computational Physics* **232**(1), 288–317.

Janin, J., Daniel, D., Emirault, F., Guilloux, A. and Lebissonnais, H. (2012), Settlement monitoring and tunnelling process adaptation - case of South Toulon tunnel, *in* G. Viggiani, ed., 'Geotechnical Aspects of Underground Construction in Soft Ground', pp. 205–212.

Jansen, M. J. (1999), 'Analysis of variance designs for model output', *Computer Physics Communications* **117**(1), 35 – 43.

Joshi, M., Seidel-Morgenstern, A. and Kremling, A. (2006), 'Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems', *Metabolic Engineering* **8**(5), 447 – 455.

Julier, S. and Uhlmann, J. (1996), A general method for approximating nonlinear transformation of probability distribution, Technical report, University of Oxford, Oxford, UK, Technical report.

Kanwar, N. and Deng, J. (2019), 'A comparison of probabilistic distributions of undrained shear strength of soils in nipigon river, canada', *IOP Conference Series: Earth and Environmental Science* **351**, 012024.

Kany, M. (1959), *Beitrag zur Berechnung von Flächengründungen*, Wilhelm Ernst und Sohn, Berlin.

Karmarkar, N. (1984), 'A new polynomial-time algorithm for linear programming', *Combinatorica* **4**(4), 373–395.

Kasama, K. and Whittle, A. J. (2016), 'Effect of spatial variability on the slope stability using random field numerical limit analyses', *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* **10**(1), 42–54.

Kashan, A. H. (2014), 'League championship algorithm (LCA): An algorithm for global optimization inspired by sport championships', *Applied Soft Computing* **16**, 171 – 200.

Kay, S. (1993), *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall PTR, Upper Saddle River, NJ.

Kennedy, J. and Eberhart, R. (1995), Particle Swarm Optimization, *in* 'Proceedings of IV. IEEE International Conference on Neural Networks', pp. 1942–1948.

Khaledi, K. (2017), Constitutive Modeling of Rock Salt with Application to Energy Storage Caverns, PhD thesis, Ruhr-Universität Bochum, Bochum, Germany.

Khaledi, K., Mahmoudi, E., Datcheva, M., König, D. and Schanz, T. (2016), 'Sensitivity analysis and parameter identification of a time dependent constitutive model for rock salt', *Journal of Computational and Applied Mathematics* **293**, 128 – 138.

Khaledi, K., Miro, S., König, M. and Schanz, T. (2014), 'Robust and reliable metamodels for mechanized tunnel simulations', *Computers and Geotechnics* **61**, 1–12.

Kirkpatrick, S., Gelatt, C. D. and Vecchi, M. P. (1983), 'Optimization by simulated annealing', *Science* **220**(4598), 671–680.

Knabe, T., Schweiger, H. F. and Schanz, T. (2012), 'Calibration of constitutive parameters by inverse analysis for a geotechnical boundary problem', *Canadian Geotechnical Journal* **49**(2), 170–183.

Kooiman, P., Dijk, H. K. V. and Thurik, A. (1985), 'Likelihood diagnostics and Bayesian analysis of a micro-economic disequilibrium model for retail services', *Journal of Econometrics* **29**(1), 121 – 148.

Krishnamurthy, T. (2005), Comparison of response surface construction methods for derivative estimation using moving least squares, kriging and radial basis functions, *in* 'Proceedings of the 46st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics & Materials Conference', pp. AIAA 2005–1821.

Kulhawy, F., Birgisson, B. and Grigoriu, M. (1992), Reliability based foundation design for transmission line structures: transformation models for in situ tests, Technical report, Electric Power Research Institute, Palo Alto, CA.

Lahmer, T. (2011), 'Optimal experimental design for nonlinear ill-posed problems applied to gravity dams', *Inverse Problems* **27**(12), 125005.

Lahmer, T., Kaltenbacher, B. and Schulz, V. (2008), 'Optimal measurement selection for piezoelectric material tensor identification', *Inverse Problems in Science and Engineering* **16**(3), 369–387.

Lambe, T. and Whitman, R. (1969), *Soil Mechanics*, John Wiley, New York.

Lancaster, P. and Salkauskas, K. (1981), 'Surfaces generated by moving least squares methods', *Mathematics of Computation* **37**(15), 141 –158.

Lavasan, A. A., Zhao, C., Barciaga, T., Schaufler, A., Steeb, H. and Schanz, T. (2018), 'Numerical investigation of tunneling in saturated soil: the role of construction and operation periods', *Acta Geotechnican* **13**, 671–291.

Legendre, A.-M. (1805), *Nouvelles méthodes pour la détermination des orbites des comètes*, Firmin Didiot, Paris, France.

Li, D.-Q., Tang, X.-S. and Phoon, K.-K. (2015), 'Bootstrap method for characterizing the effect of uncertainty in shear strength parameters on slope reliability', *Reliability Engineering & System Safety* **140**, 99 – 106.

Li, X., Zhao, C., Hölter, R., Datcheva, M. and Lavasan, A. A. (2019), 'Modelling of a large landslide problem under water level fluctuation—model calibration and verification', *Geosciences* **9**(2), 89.

Lo, M. K. and Leung, Y. F. (2018), 'Reliability Assessment of Slopes Considering Sampling Influence and Spatial Variability by Sobol' Sensitivity index', *Journal of Geotechnical and Geoenvironmental Engineering* **144**(4), 04018010.

Lo Piano, S., Ferretti, F., Puy, A., Albrecht, D. and Saltelli, A. (2021), 'Variance-based sensitivity analysis: The quest for better estimators and designs between explorativity and economy', *Reliability Engineering and System Safety* **206**, 107300.

Low, B. K. (2017), 'Insights from reliability-based design to complement load and resistance factor design approach', *Journal of Geotechnical and Geoenvironmental Engineering* **143**(11), 04017089.

Low, B. K. and Phoon, K.-K. (2015), 'Reliability-based design and its complementary role to Eurocode 7 design approach', *Computers and Geotechnics* **65**, 30 – 44.

Luo, Z., Atamturktur, S. and Juang, C. H. (2013), 'Bootstrapping for characterizing the effect of uncertainty in sample statistics for braced excavations', *Journal of Geotechnical and Geoenvironmental Engineering* **139**(1), 13–23.

Lwin, K. T., Qu, R. and MacCarthy, B. L. (2017), 'Mean-var portfolio optimization: A nonparametric approach', *European Journal of Operational Research* **260**(2), 751 – 766.

Mahmoudi, E. (2017), Probabilistic analysis of a rock salt cavern with application to energy storage systems, PhD thesis, Ruhr-Universität Bochum, Bochum, Germany.

Mahmoudi, E., Hölter, R., Georgieva, R., König, M. and Schanz, T. (2019), 'On the Global Sensitivity Analysis Methods in Geotechnical Engineering: A Comparative Study on a Rock Salt Energy Storage', *International Journal of Civil Engineering* **17**, 131–143.

Mahmoudi, E., Khaledi, K., Miro, S., König, D. and Schanz, T. (2017), 'Probabilistic Analysis of a Rock Salt Cavern with Application to Energy Storage Systems', *Rock Mechanics and Rock Engineering* **50**(1), 139–157.

Mahmoudi, E., Schmüdderich, C., Hölter, R., Zhao, C., Wichtmann, T. and König, M. (2020), 'Stochastic field simulation of slope stability problems: Improvement and reduction of computational effort', *Computer Methods in Applied Mechanics and Engineering* **369**, 113167.

Marshall, A., Farrell, R., Klar, A. and Mair, R. (2012), 'Tunnels in sands: the effect of size, depth and volume loss on greenfield displacements', *Géotechnique* **62**(5), 385–399.

Masini, L., Rampello, S. and Soga, K. (2014), 'An approach to evaluate the efficiency of compensation grouting', *Journal of Geotechnical and Geoenvironmental Engineering* **140**(12), 04014073.

Matott, L. S. and Rabideau, A. J. (2008), 'Calibration of complex subsurface reaction models using a surrogate-model approach', *Advances in Water Resources* **31**(12), 1697 – 1707.

McKay, M., J.Beckman, R. and Conover, W. (1979), 'A comparison of three methods for selecting values of input variables in the analysis of output from a computer code', *Technometrics* **21**, 239 – 245.

MEC International, LLC (2020), 'Levee failure'.
URL: *http://www.mecinternational.us/wp-content/uploads/2014/02/Levee-failure-300x213.jpg/*

Meier, J. (2008), Parameterbestimmung mittels inverser Verfahren für geotechnische Problemstellungen, PhD thesis, Bauhaus-Universität Weimar, Weimar, Germany.

Meier, J., Datcheva, M., Moser, M. and Schanz, T. (2009), 'Identification of constitutive and geometrical parameters for slope instability modelling application to mountain-splitting area Reutte/ Tyrol', *Austrian Journal of Earth Sciences* **102**, 81–89.

Meier, J., Moser, M., Datcheva, M. and Schanz, T. (2013), 'Numerical modeling and inverse parameter estimation of the large-scale mass movement Gradenbach in Carinthia (Austria)', *Acta Geotechnica* **8**(4), 355–371.

Meschke, G., Kropik, C. and Mang, H. A. (1996), 'Numerical analyses of tunnel linings by means of a viscoplastic material model for shotcrete', *International Journal for Numerical Methods in Engineering* **39**(18), 3145–3162.

Metropolis, N. and Ulam, S. (1949), 'The Monte Carlo method', *Journal of the American statistical Association* **44**(247), 335–341.

Meyer, R. K. and Nachtsheim, C. J. (1995), 'The coordinate-exchange algorithm for constructing exact optimal experimental designs', *Technometrics* **37**(1), 60–69.

Miro, S. (2016), Calibration of Numerical Models Considering Uncertainties - Application to Mechanized Tunnel Simulations, PhD thesis, Ruhr-Universität Bochum, Bochum, Germany.

Miro, S., Hartmann, D. and Schanz, T. (2014), 'Global sensitivity analysis for subsoil parameter estimation in mechanized tunneling', *Computers and Geotechnics* **56**, 80–88.

Miro, S., König, M., Hartmann, D. and Schanz, T. (2015), 'A probabilistic analysis of subsoil parameters uncertainty impacts on tunnel-induced ground movements with a back-analysis study', *Computers and Geotechnics* **68**, 38 – 53.

Mitchell, J. and Soga, K. (2005), *Fundamentals of Soil Behavior*, 3rd edn, John Wiley, New York.

Mollon, G., Dias, D. and Soubra, A.-H. (2013), 'Continuous velocity fields for collapse and blowout of a pressurized tunnel face in purely cohesive soil', *International Journal for Numerical and Analytical Methods in Geomechanics* **37**(13), 2061–2083.

Morris, M. D. (1991), 'Factorial sampling plans for preliminary computational experiments', *Technometrics* **33**(2), 161–174.

Most, T. and Knabe, T. (2010), 'Reliability analysis of the bearing failure problem considering uncertain stochastic parameters', *Computers and Geotechnics* **37**(3), 299 – 310.

Most, T. and Will, J. (2008), Meta-model of optimal prognosis - an automatic approach for variable reduction and optimal meta-model selection, *in* 'Proceedings of the Weimar Optimization and Stochastic Days 2008'.

Motra, H., Stutz, H. and Wuttke, F. (2016), 'Quality assessment of soil bearing capacity factor models of shallow foundations', *Soils and Foundations* **56**(2), 265 – 276.

Mukherjee, R., Diwekar, U. M. and Vaseashta, A. (2017), 'Optimal sensor placement with mitigation strategy for water network systems under uncertainty', *Computers and Chemical Engineering* **103**, 91 – 102.

Mullen, K., Ardia, D., Gil, D., Windover, D. and Cline, J. (2011), 'DEoptim: An R package for global optimization by differential evolution', *Journal of Statistical Software, Articles* **40**(6), 1–26.

Müthing, N., Zhao, C., Hölter, R. and Schanz, T. (2018), 'Settlement prediction for an embankment on soft clay', *Computers and Geotechnics* **93**, 87 – 103.

Nelder, J. A. and Mead, R. (1965), 'A Simplex Method for Function Minimization', *The Computer Journal* **7**(4), 308–313.

Nguyen, L. T. (2017), Inference of ground condition in mechanized tunneling via inverse analysis using sequential Bayesian filtering, PhD thesis, Ruhr-Universität Bochum, Bochum, Germany.

Nguyen, L. T. and Nestorović, T. (2016), 'Nonlinear Kalman filters for model calibration of soil parameters for geomechanical modeling in mechanized tunneling', *Journal of Computing in Civil Engineering* **30**(2), 04015025.

Nishii, R. (1993), 'Optimality of experimental designs', *Discrete Mathematics* **116**(1), 209 – 225.

Nishimura, S., Shuku, T. and Fujisawa, K. (2014), 'Prediction of multidimensional deformation behavior based on observed values', *International Journal of Geomechanics* **14**(3), 04014011.

Overstall, A. M. and Woods, D. C. (2017), 'Bayesian design of experiments using approximate coordinate exchange', *Technometrics* **59**(4), 458–470.

Patan, M. (2012), 'Distributed scheduling of sensor networks for identification of spatio-temporal processes', *International Journal of Applied Mathematics and Computer Science* **22**(2), 299 – 311.

Peck, R. (1969*a*), Deep excavations and tunneling in soft ground, *in* 'State-of-the Art Volume, Seventh International Conference on Soil Mechanics and Foundations, Mexico City', p. 225–290.

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

Phoon, K.-K. and Kulhawy, F. (1999*a*), 'Characterization of geotechnical variability', *Canadian geotechnical journal* **36**, 612 – 624.

Phoon, K.-K. and Kulhawy, F. (1999*b*), 'Evaluation of geotechnical property variability', *Canadian Geotechnical Journal* **36**, 625 – 639.

Phoon, K.-K. and Tang, C. (2019), 'Characterisation of geotechnical model uncertainty', *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* **13**(2), 101–130.

Pineda, J., Suwal, L., Kelly, R., Bates, L. and Sloan, S. (2016), 'Characterisation of Ballina clay', *Géotechnique* **66**(27), 556–577.

Plaxis bv (2018), 'PLAXIS 3D Reference Manual'.

Plaxis bv (2019), 'PLAXIS 2D Reference Manual'.

Potts, D. M. (1999), *Finite element analysis in geotechnical engineering*, Thomas Telford Publishing.

Potts, D. M. and Zdravković, L. (2001), *Finite Element Analysis in Geotechnical Engineering: Volume two - Application*, Thomas Telford Publishing.

Press, W. P., Teukolsky, Saul A. amd Vetterling, W. T. and Flannery, B. P. (2007), *Numerical Recipes - The Art of Scientific Computing*, 3. edn, Cambridge University Press.

Pronzato, L. (2008), 'Optimal experimental design and some related control problems', *Automatica* **44**(2), 303 – 325.

Ragkousis, G. E., Curzen, N. and Bressloff, N. W. (2016), 'Multi-objective optimisation of stent dilation strategy in a patient-specific coronary artery via computational and surrogate modelling', *Journal of Biomechanics* **49**(2), 205 – 215.

Rajabioun, R. (2011), 'Cuckoo optimization algorithm', *Applied Soft Computing* **11**(8), 5508 – 5518.

Raspe, R. E. (1785), *Baron Munchausen's Narrative of his Marvellous Travels and Campaigns in Russia*, England.

Reichert, I., Olney, P. and Lahmer, T. (2019), 'Influence of the error description on model-based design of experiments', *Engineering Structures* **193**, 100 – 109.

Rodriguez-Fernandez, M., Kucherenko, S., Pantelides, C. and Shah, N. (2007), Optimal experimental design based on global sensitivity analysis, *in* V. Pleşu and P. Şerban Agachi, eds, '17th European Symposium on Computer Aided Process Engineering', Vol. 24 of *Computer Aided Chemical Engineering*, Elsevier, pp. 63 – 68.

Roscoe, K. and Burland, J. (1968), *On the generalised stress-strain behaviour of 'wet' clay*, Cambridge University Press.

Rote, G. (1991), 'Computing the minimum Hausdorff distance between two point sets on a line under translation', *Information Processing Letters* **38**(3), 123 – 127.

RStudio Team (2015), *RStudio: Integrated Development Environment for R*, RStudio, Inc., Boston, MA.
URL: *http://www.rstudio.com/*

Sacks, J., Welch, W. J., Mitchell, T. J. and Wynn, H. P. (1989), 'Design and analysis of computer experiments', *Statist. Sci.* **4**(4), 409–423.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. and Tarantola, S. (2010), 'Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index', *Computer Physics Communications* **181**(2), 259 – 270.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J. and Gatelli, D. (2008), *Global sensitivity analysis. The Primer*, John Wiley and Sons.

Schädler, W., Borgatti, L., Corsini, A., Meier, J., Ronchetti, F. and Schanz, T. (2015), 'Geomechanical assessment of the Corvara earthflow through numerical modelling and inverse analysis', *Landslides* **12**(3), 495–510.

Schaibly, J. H. and Shuler, K. E. (1973), 'Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. II Applications', *The Journal of Chemical Physics* **59**(8), 3879–3888.

Schanz, T. (1998), Zur Modellierung des mechanischen Verhaltens von Reibungsmaterialien, habilitation, Institut für Geotechnik, Universität Stuttgart, Stuttgart, Germany.

Schanz, T. and Meier, J. (2008), 'Gestaltung, Validierung und Optimierung von Messprogrammen für geotechnische Aufgabenstellungen', *Bautechnik* **85**(5), 307–316.

Schanz, T., Vermeer, P. and Bonnier, P. (1999), The hardening soil model: Formulation and verification, *in* 'Proceedings of 1st International PLAXIS Symposium on Beyond 2000 in Computational Geotechnics', Balkema, pp. 281–296.

Schenkendorf, R. (2014), Optimal Experimental Design for Parameter Identification and Model Selection, PhD thesis, Otto-von-Guericke-Universität, Magdeburg, Germany.

Schenkendorf, R., Kremling, A. and Mangold, M. (2009), 'Optimal experimental design with the sigma point method', *IET Systems Biology* **3**(1), 10–23.

Schenkendorf, R., Xie, X., Rehbein, M., Scholl, S. and Krewer, U. (2018), 'The impact of global sensitivities and design measures in model-based optimal experimental design', *Processes* **6**(4).

Schoen, M. (2018*a*), 'Inverse Analyse von Bodenparameter auf Grundlage von Messdaten einer Tunnelunterfahrung eines mehrstöckigen Gebäudes'. Project work, Ruhr-Universität Bochum, Bochum, Germany.

Schoen, M. (2018*b*), 'Optimierte Modellvalidierung mittels Messdaten am Beispiel des Tunnelvortriebs der Mailänder Metrolinie 5', Master's thesis, Ruhr-Universität Bochum, Bochum, Germany.

Shahin, H. M., Nakai, T., Ishii, K., Iwata, T. and Kuroi, S. (2016), 'Investigation of influence of tunneling on existing building and tunnel: model tests and numerical simulations', *Acta Geotechnica* **11**(3), 679–692.

Shahin, H. M., Nakai, T., Zhang, F., Kikumoto, M. and Nakahara, E. (2011), 'Behavior of ground and response of existing foundation due to tunneling', *Soils and Foundations* **51**(3), 395–409.

Shuku, T., Murakami, A., ichi Nishimura, S., Fujisawa, K. and Nakamura, K. (2012), 'Parameter identification for Cam-Clay model in partial loading model tests using the particle filter', *Soils and Foundations* **52**(2), 279 – 298.

Sobol', I. (1993), 'Sensitivity estimates for nonlinear mathematical models', *Mathematical Models in Civil Engineering* **1**(4), 407–414.

Spross, J. and Johansson, F. (2017), 'When is the observational method in geotechnical engineering favourable?', *Structural Safety* **66**, 17–26.

Steiner, H.-J., Vratny, P. C., Gologan, C., Wieczorek, K., Isikveren, A. T. and Hornung, M. (2014), 'Optimum number of engines for transport aircraft employing electrically powered distributed propulsion', *CEAS Aeronautical Journal* **5**(2), 157–170.

Storn, R. and Price, K. (1997), 'Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces', *Journal of Global Optimization* **11**(4), 341–359.

Tang, C., Phoon, K.-K., Zhang, L. and Li, D.-Q. (2017), 'Model uncertainty for predicting the bearing capacity of sand overlying clay', *International Journal of Geomechanics* **17**(7), 04017015.

Tarantola, A. and Valette, B. (1982), 'Generalized nonlinear inverse problems solved using the least squares criterion', *Reviews of Geophysics* **20**(2), 219–232.

Taylor, D. (1948), *Fundamentals of Soil Behavior*, John Wiley, New York.

Thacker, B. H., Doebling, S. W., Hemez, F. M., Anderson, M. C., Pepin, J. E. and Rodriguez, E. A. (2004), Concepts of model verification and validation, Technical report, Los Alamos National Laboratory, Los Alamos, NM, Technical report.

Tippmann, S. (2015), 'Programming tools: Adventures with R', *Nature* **517**(7532), 109–110.

Tsinidis, G., Pitilakis, K., Madabhushi, G. and Heron, C. (2015), 'Dynamic response of flexible square tunnels: centrifuge testing and validation of existing design methodologies', *Géotechnique* **65**(5), 401–417.

Uciński, D. (2005), *Optimal Measurement Methods for Distributed Parameter System Identification*, CRC Press, Boca Raton.

Uzielli, S., Lacasse, S., Nadim, F. and Phoon, K. (2006), Soil variability analysis for geotechnical practice, *in* 'Proceedings of the Second International Workshop on Characterisation and Engineering Properties of Natural Soils, Singapore'.

Vanlier, J., Tiemann, C. A., Hilbers, P. A. J. and van Riel, N. A. W. (2012), 'A Bayesian approach to targeted experiment design', *Bioinformatics* **28**(8), 1136–1142.

Vincenzi, L. and Savoia, M. (2015), 'Coupling response surface and differential evolution for parameter identification problems', *Computer-Aided Civil and Infrastructure Engineering* **30**(5), 376–393.

Viratjandr, C. and Michalowski, R. L. (2006), 'Limit analysis of submerged slopes subjected to water drawdown', *Canadian Geotechnical Journal* **43**(8), 802–814.

von Neumann, J. (1951), 'Various techniques used in connection with random digits', *Journal of Research of the National Bureau of Standards, Applied Mathematics* **3**, 36 – 38.

Wichtmann, T., Kimmig, I., Steller, K., Triantafyllidis, T., Back, M. and Dahmen, D. (2019), 'Correlations of the liquefaction resistance of sands in spreader dumps of lignite opencast mines with cpt tip resistance and shear wave velocity', *Soil Dynamics and Earthquake Engineering* **124**, 184–196.

Wolkenhauer, O., Wellstead, P., Cho, K.-H., Banga, J. R. and Balsa-Canto, E. (2008), 'Parameter estimation and optimal experimental design', *Essays in Biochemistry* **45**, 195–210.

Zarev, V. (2014), Model Identification for the Adaption of Numerical Simulation Models - Application to Mechanized Shield Tunneling, PhD thesis, Ruhr-Universität Bochum, Bochum, Germany.

Zhang, Z. F., Ward, A. L. and Gee, G. W. (2003), 'Estimating soil hydraulic parameters of a field drainage experiment using inverse techniques', *Vadose Zone Journal* **2**(2), 201–211.

Zhao, C. (2018), A Contribution to Modeling of Mechanized Tunnel Excavation, PhD thesis, Ruhr-Universität Bochum, Bochum, Germany.

Zhao, C., Hölter, R., König, M. and Lavasan, A. A. (2019), 'A hybrid model for estimation of ground movements due to mechanized tunnel excavation', *Computer-Aided Civil and Infrastructure Engineering* **34**(7), 586–601.

Zhao, C., Lavasan, A. A., Barciaga, T., Veselin, Datcheva, M. and Schanz, T. (2015), 'Model validation and calibration via back analysis for mechanized tunnel simulations - The Western Scheldt tunnel case', *Computers and Geotechnics* **69**, 601–614.

Zhao, C., Lavasan, A. A., Hölter, R. and Schanz, T. (2018), 'Mechanized tunneling induced building settlements and design of optimal monitoring strategies based on sensitivity field', *Computers and Geotechnics* **97**, 246 – 260.

Zhao, C., Lavasan, A. A. and Schanz, T. (2014), Sensitivity analysis of the model response in mechanized tunneling simulation -A case study assessment, *in* H. Rodrigues et al., eds, '4th International Conference on Engineering Optimization', pp. 491–496.

Zhao, D. and Xue, D. (2010), 'A comparative study of metamodeling methods considering sample quality merits', *Structural and Multidisciplinary Optimization* **42**(6), 923–938.

Schriftenreihe des Instituts für Grundbau, Wasserwesen und Verkehrswesen der Ruhr-Universität Bochum

Herausgeber: H.L. Jessberger

1 (1979) **Hans Ludwig Jessberger**
Grundbau und Bodenmechanik an der Ruhr-Universität Bochum

2 (1978) **Joachim Klein**
Nichtlineares Kriechen von künstlich gefrorenem Emschermergel

3 (1979) **Heinz-Joachim Gödecke**
Die Dynamische Intensivverdichtung wenig wasserdurchlässiger Böden

4 (1979) **Poul V. Lade**
Three Dimensional Stress-Strain Behaviour and Modeling of Soils

5 (1979) **Roland Pusch**
Creep of soils

6 (1979) **Norbert Diekmann**
Zeitabhängiges, nichtlineares Spannungs-Verformungsverhalten von gefrorenem Schluff unter triaxialer Belastung

7 (1979) **Rudolf Dörr**
Zeitabhängiges Setzungsverhalten von Gründungen in Schnee, Firn und Eis der Antarktis am Beispiel der deutschen Georg-von-Neumayer- und Filchner-Station

8 (1984) **Ulrich Güttler**
Beurteilung des Steifigkeits- und Nachverdichtungsverhaltens von ungebundenen Mineralstoffen

9 (1986) **Peter Jordan**
Einfluss der Belastungsfrequenz und der partiellen Entwässerungs-möglichkeiten auf die Verflüssigung von Feinsand

10 (1986) **Eugen Makowski**
Modellierung der künstlichen Bodenvereisung im grundwasserdurchströmten Untergrund mit der Methode der finiten Elemente

11 (1986) **Reinhard A. Beine**
Verdichtungswirkung der Fallmasse auf Lastausbreitung in nichtbindigem Boden bei der Dynamischen Intensivverdichtung

12 (1986) **Wolfgang Ebel**
Einfluss des Spannungspfades auf das Spannungs-Verformungsverhalten von gefrorenem Schluff im Hinblick auf die Berechnung von Gefrierschächten

13 (1987) **Uwe Stoffers**
Berechnungen und Zentrifugen-Modellversuche zur Verformungsabhängigkeit der Ausbaubeanspruchung von Tunnelausbauten in Lockergestein

14 (1988) **Gerhard Thiel**
Steifigkeit und Dämpfung von wassergesättigtem Feinsand unter Erdbebenbelastung

15 (1991) **Mahmud Thaher**
Tragverhalten von Pfahl-Platten-Gründungen im bindigen Baugrund,
Berechnungsmodelle und Zentrifugen-Modellversuche

Schriftenreihe des Instituts für Grundbau der Ruhr-Universität Bochum

Herausgeber: H.L. Jessberger

16 (1992) **Rainer Scherbeck**
Geotechnisches Verhalten mineralischer Deponieabdichtungsschichten
bei ungleichförmiger Verformungswirkung

17 (1992) **Martin M. Bizialiele**
Torsional Cyclic Loading Response of a Single Pile in Sand

18 (1993) **Michael Kotthaus**
Zum Tragverhalten von horizontal belasteten Pfahlreihen aus langen Pfählen in Sand

19 (1993) **Ulrich Mann**
Stofftransport durch mineralische Deponieabdichtungen:
Versuchsmethodik und Berechnungsverfahren

20 (1992) **Festschrift anlässlich des 60. Geburtstages von**
Prof. Dr.-Ing. H. L. Jessberger
20 Jahre Grundbau und Bodenmechanik an der Ruhr-Universität Bochum

21 (1993) **Stephan Demmert**
Analyse des Emissionsverhaltens einer Kombinationsabdichtung im Rahmen der
Risikobetrachtung von Abfalldeponien

22 (1994) **Diethard König**
Beanspruchung von Tunnel- und Schachtausbauten in kohäsionslosem Lockergestein
unter Berücksichtigung der Verformung im Boden

23 (1995) **Thomas Neteler**
Bewertungsmodell für die nutzungsbezogene Auswahl von Verfahren zur Altlastensanierung

24 (1995) **Ralph Kockel**
Scherfestigkeit von Mischabfall im Hinblick auf die Standsicherheit von Deponien

25 (1996) **Jan Laue**
Zur Setzung von Flachfundamenten auf Sand unter wiederholten Lastereignissen

26 (1996) **Gunnar Heibrock**
Zur Rissbildung durch Austrocknung in mineralischen Abdichtungsschichten
an der Basis von Deponien

27 (1996) **Thomas Siemer**
Zentrifugen-Modellversuche zur dynamischen Wechselwirkung zwischen Bauwerken
und Baugrund infolge stoßartiger Belastung

28 (1996) **Viswanadham V. S. Bhamidipati**
Geosynthetic Reinforced Mineral Sealing Layers of Landfills

29 (1997) **Frank Trappmann**
Abschätzung von technischem Risiko und Energiebedarf bei Sanierungsmaßnahmen für Altlasten

30 (1997) **André Schürmann**
Zum Erddruck auf unverankerte flexible Verbauwände

31 (1997) **Jessberger, H. L. (Herausgeber)**
Environment Geotechnics, Report of ISSMGE Technical Committee TC 5 on Environmental Geotechnics

Schriftenreihe des Institutes für Grundbau und Bodenmechanik der Ruhr-Universität Bochum

Herausgeber: Th. Triantafyllidis

32 (2000) **Triantafyllidis, Th. (Herausgeber)**
Boden unter fast zyklischer Belastung: Erfahrung und Forschungsergebnisse (Workshop)

33 (2002) **Christof Gehle**
Bruch- und Scherverhalten von Gesteinstrennflächen mit dazwischenliegenden Materialbrücken

34 (2003) **Andrzej Niemunis**
Extended hypoplastic models for soils

35 (2004) **Christiane Hof**
Über das Verpressankertragverhalten unter kalklösendem Kohlensäureangriff

36 (2004) **René Schäfer**
Einfluss der Herstellungsmethode auf das Verformungsverhalten von Schlitzwänden in weichen bindigen Böden

37 (2005) **Henning Wolf**
Zur Scherfugenbänderung granularer Materialien unter Extensionsbeanspruchung

38 (2005) **Torsten Wichtmann**
Explicit accumulation model for non-cohesive soils under cyclic loading

39 (2008) **Christoph M. Loreck**
Die Entwicklung des Frischbetondruckes bei der Herstellung von Schlitzwänden

40 (2008) **Igor Arsic**
Über die Bettung von Rohrleitungen in Flüssigböden

41 (2009) **Anna Arwanitaki**
Über das Kontaktverhalten zwischen einer Zweiphasenschlitzwand und nichtbindigen Böden

Schriftenreihe des Lehrstuhls für Grundbau, Boden- und Felsmechanik der Ruhr-Universität Bochum

Herausgeber: T. Schanz

42 (2009) **Yvonne Lins**
Hydro-Mechanical Properties of Partially Saturated Sand

43 (2010) **Tom Schanz (Herausgeber)**
Geotechnische Herausforderungen beim Umbau des Emscher-Systems,
Beiträge zum RuhrGeo Tag 2010

44 (2010) **Jamal Alabdullah**
Testing Unsaturated Soil for Plane Strain Conditions: A New Double-Wall Biaxial Device

45 (2011) **Lars Röchter**
Systeme paralleler Scherbänder unter Extension im ebenen Verformungszustand

46 (2011) **Yasir Al-Badran**
Volumetric Yielding Behavior of Unsaturated Fine-Grained Soils

47 (2011) **Usque ad finem**
Selected research papers

48 (2012) **Muhammad Ibrar Khan**
Hydraulic Conductivity of Moderate and Highly Dense Expansive Clays

49 (2014) **Long Nguyen-Tuan**
Coupled Thermo-Hydro-Mechanical Analysis: Experimental and Back Analysis

50 (2014) **Tom Schanz (Herausgeber)**
Ende des Steinkohlenbergbaus im Ruhrrevier: Realität und Perspektiven für die Geotechnik, Beiträge zum RuhrGeo Tag 2014

51 (2014) **Usque ad finem**
Selected research papers

52 (2014) **Houman Soleimani Fard**
Study on the Hydro-Mechanical Behaviour of Fiber Reinforced Fine Grained Soils with Application to the Preservation of Historical Monuments

53 (2014) **Wiebke Baille**
Hydro-Mechanical Behavior of Clays - Significance of Mineralogy

54 (2014) **Qasim Abdulkarem Jassim Al-Obaidi**
Hydro-Mechanical Behavior of Collapsible Soils

55 (2015) **Veselin Zarev**
Model Identification for the Adaption of Numerical Simulation Models - Application to Mechanized Shield Tunneling

56 (2015) **Meisam Goudarzy**
Micro and Macro Mechanical Assessment of Small and Intermediate Strain Properties of Granular Material

57 (2016) **Oliver Detert**
Analyse einer selbstregulierenden interaktiven Membrangründung für Schüttkörper auf geringtragfähigen Böden

58 (2016) **Yang Yang**
Analyses of Heat Transfer and Temperature-induced Behaviour in Geotechnics

59 (2016) **Alborz Pourzargar**
Application of suction stress concept to partially saturated compacted soils

60 (2017) **Hanna Haase**
Multiscale analysis of clay-polymer composites for Geoenvironmental applications

61 (2017) **Kavan Khaledi**
Constitutive modeling of rock salt with application to energy storage caverns

62 (2017) **Nina Silvia Müthing**
On the consolidation behavior of fine-grained soils under cyclic loading

63 (2017) **Elham Mahmoudi**
Probabilistic analysis of a rock salt cavern with application to energy storage systems

64 (2017) **Negar Rahemi**
Evaluation of liquefaction behavior of sandy soils using critical state soil mechanics and instability concept

65 (2018) **Chenyang Zhao**
A contribution to modeling of mechanized tunnel excavation

66 (2018) **Tom Schanz (Herausgeber)**
Innovationen im Spezialtiefbau und in der Umweltgeotechnik,
Beiträge zum RuhrGeo Tag 2018

67 (2019) **Linzhi Lang**
Hydro-Mechanical Behaviour of Bentonite-Based Materials Used for Disposal of Radioactive Wastes

68 (2019) **Usama Al-Anbaki**
Hydraulic Interaction of Soil and Nonwoven Geotextiles under Unsaturated Conditions

69 (2019) **Abhishek Rawat**
Coupled Hydro-mechanical Behavior of a Compacted Bentonite-Sand Mixture: Experimental and Numerical Investigations

Schriftenreihe des Lehrstuhls für Bodenmechanik, Grundbau und Umweltgeotechnik der Ruhr-Universität Bochum

Herausgeber: T. Wichtmann

70 (2019) **Mahmoud Qarmout**
Tunnel face stability using Kinematical Element Method (KEM)

71 (2021) **Raoul Hölter**
Optimal Experimental Design in Geotechnical Engineering