Evaluation of Inverse Analysis Methods with Numerical Simulation of Slope Excavation

A. Mavritsakis

Master of Science Thesis

TU Delft

Civil Engineering and Geosciences

Department of Geo-Engineering

Evaluation of inverse analysis methods with numerical simulation for slope excavation

By

A. Mavritsakis

in partial fulfilment of the requirements for the degree of

Master of Science

in Geo-Engineering

at the Delft University of Technology.

Thesis committee:

Prof. dr. M. A. Hicks

Dr. P. J. Vardon

Dr. A. P. van den Eijnden

Dr. ir. F. C. Vossepoel

PREFACE

This dissertation started in December 2016 in order to accomplish the requirements for a Geo-Engineering Master's degree at TU Delft. While it demanded expertise on several fields and a higher tier of programming skills, this thesis helped me improve as an engineer and a soon-to-be young professional. Ultimately, I can claim that working on it had been straining, yet rewarding and fulfilling for my research curiosity. Nevertheless, I would not have been able to deliver it on my own.

First of all I would like to thank Bram van Eijnden for his assistance on this dissertation. His advice on both theoretical and practical aspects of my work through every stage has been critical and accurate. Furthermore, his detailed-oriented attitude along with his scientific enthusiasm greatly enhanced the quality of my efforts.

Also, I would like to thank the rest of my committee members M. Hicks, P. Vardon, F. Vossepoel for their insightful guidance against any major problems encountered or decisions that had to be taken. Admittedly, their aid has been most valuable and rendered the completion of my thesis possible.

Moreover, I would like to thank my family for their support, as well as my friends, who stood by me throughout my years as a student both in the Netherlands and in Greece.

Finally, I thank TU Delft for its high quality of studies and effort put into assimilating foreign students. All in all, this institution has enabled me to pursue my personal goals and aspirations.

Antonios Mavritsakis, 2017

Contents

Abstract

Insufficient information on soil parameters and their spatial variability pose as a main factors of uncertainty in geotechnical design. When soil response differs from the expected, the notion of inverse analysis becomes relevant; back-calculating the parameter set able to reproduce the monitored observations. Accordingly, its application attempts to clarify the effective soil conditions and allows for an update of the design based on in-situ measurements. The main goal of this thesis is assessing the adaption of available inverse analysis concepts in geotechnical engineering, as well as evaluating the reliability of their outcome. Specifically, the methods are employed on a slope construction problem in order to exhibit their function, compare their efficiency and point out the obstacles encountered in their application.

After selecting the three most competitive approaches (namely: the Markov Chain Monte Carlo, the Genetic Algorithm and the Regularized Particle Filter), a synthetic case of RFEM model simulating a staged slope excavation is composed to test them, thus includes both soil heterogeneity and a pseudo-time scheme. Horizontal displacements at specified mesh points, mimicking the readings of inclinometers, act as the observations. Subsequently, the soil parameters to be identified are the undrained shear strength and the Young's modulus, as they yield the heaviest impact on displacement generation. Essentially, the objective of the inverse analysis is to minimize the error between observations and their RFEM-simulated counterpart. Furthermore, since all methods produce stochastically defined approximations of the parameter fields, a reliability analysis is suitable for their appraisal. Utilizing a Monte Carlo framework, the confidence curves of the results are plotted and judged. In an effort to comply with modern design standards, the reliability level of 95% acts as one of the evaluation criteria. Moreover, a practice-driven example employs the reliability concept in predicting the displacements of the next stage.

Inverse methods are evaluated according to their computational performance, their efficiency in identifying soil parameters, as well as their behavior in the reliability analysis. In this study, the Genetic Algorithm is proven to be the most competent inverse analysis approach. It easily adapts to the problem, requires fewer RFEM runs and provides accurate and reliable results. Additionally, in the course of the project two main hindrances where encountered: the approximation of the strength field when applying a Linearly Elastic – Perfectly Plastic constitutive model and the combination of the Regularized Particle Filter with the RFEM model. Both are thoroughly analyzed and the conditions required for their mitigation are presented with numerical examples. Ultimately the thesis delivers on its research goals by providing insight on inverse method applications. It includes an RFEM setting, that employs authentic elements, and also addresses some practice-related aspects. All in all, this is an effort towards the development of the numerical inverse analysis and hopefully its adoption as part of an integrated geotechnical design.

Nomenclature

CDF: Cumulative Density Function

CoV: Coefficient of Variation

c_u: Undrained shear strength of soil

E:Young's modulus of soil

EnKF: Ensemble Kalman Filter

FEM: Finite Element Method

GA: Genetic Algorithm

KF: Kalman Filter

MCMC: Markov Chain Monte Carlo

ML: Maximum Likelihood

PDF: Probability Density Function

PF: Particle filter

RFEM: Random Finite Element Method

RPF: Regularized Particle Filter

SD: Standard deviation for state proposal

V: Ensemble of eigenvectors scaled to the eigenvalues of the heterogeneity matrix

γ: Unit weight of soil

θ_h: Horizontal scale of fluctuation

θ_v: Vertical scale of fluctuation

μ: Mean

ν: Poisson's ratio of soil

σ: Standard deviation

ϕ: Friction angle of soil

ψ: Dilatancy angle of soil

1. Introduction

1.1. Thesis background

A critical issue of geotechnical engineering has always been parameter uncertainty. Even when advanced analysis concepts are employed in project design, actual soil response cannot be simulated perfectly and in-situ measurements differ from those expected. While this may happen due to the inability to fully grasp soil behaviour, it is mostly accredited to the inaccuracy of assigning soil parameters and charting the heterogeneity. At this point, the notion of inverse analysis is relevant.

Fundamentally, the inverse analysis is the process of calculating the set of parameters that is able to produce the measured observations when applied in the operator connecting them (Groetsch, 1999). In other words, the inverse analysis aims into allocating the input of a function, given its output. While this procedure can be easily employed on linear problems, where the inverse function can be explicitly stated, recent increase in computational power has driven it to more complex applications. Dealing with a non-linear operator means that the effect of every cause on the final product is ambiguous, thus hindering the back-calculation of the parameters. Therefore, probability theory is introduced into the inverse analysis. Instead of calculating the inverse function, such methods attempt to simulate the parameter set through series of guided sampling. As a result, inverse analysis is now applicable to Geo-engineering issues more than ever.

So far, inverse analysis has been present in geotechnical design, but usually in a more empirical fashion. For instance, the Observational Method, pioneered by Karl Terzaghi, has been the cornerstone on the New Austrian Tunnelling Method (NATM), allowing for the revaluation of design parameters based on measurements of tunnel wall convergence (Peck, 1969). The need for applying an inverse analysis is critical when measurements deviate from model predictions. Firstly, this can be attributed to the incompetency of deterministic models in capturing soil heterogeneity. Secondly, although soil testing is one of the pillars for standardized geotechnical engineering, its results are often alienated from reality. Laboratory testing inherently induces parameter errors, as specimens are disturbed when removed from actual project conditions. On the other hand, in-situ testing examines soil parameters in natural situations, but heavily relies on empirical correlations, which might not always be precise.

The objective of the inverse analysis is to identify the effective soil parameter field. By taking advantage of actual soil response, it efficiently delivers an unbiased estimation of heterogeneity patterns and properties of the subsoil. Besides, this knowledge is critical for geotechnical design. The implementation of inverse analysis in practice is founded on the presence of feedback loops. While construction advances based on the initial design, monitoring provides the observations to be employed in the inverse analysis. The resulting recognition of the underlying parameter field is then used to reassess the design for later stages of the project.

This thesis inspects the application of the inverse analysis on a synthetic case, constituted by the multi-staged construction of a slope. Specifically, displacement profiles act as the measurements of each phase, while the back-calculation attempts to approximate the user-defined source properties. The framework of this investigation is the utilization of the Random Finite Element Method (RFEM) (Griffiths & Fenton, 2004), which combines random field theory with typical Finite Element Method (FEM) calculations, hence enabling soil variability in the slope construction FEM simulation. Based on this, the inverse analysis provides fields of parameters that comply with the given heterogeneity pattern and reproduce the set measurements when supplied in the RFEM model. Accordingly, the outcome of most inverse methods is stochastic in nature, meaning that parameters are described by statistical distributions. Therefore, the reliability of these results has to be thoroughly appraised, here accomplished through a Monte Carlo analysis on the collected samples.

Extending this concept, the confidence of utilizing the inverse analysis in predicting the response of the construction has to be evaluated, before accepting it as part of an integrated geotechnical design.

Notable, this thesis is based on the work of Vardon, Liu, & Hicks (2016) and Noordam (2016), who have dealt with recognizing heterogeneous fields of hydraulic conductivity using water head measurements.

1.2. Thesis goals

This report attempts to expand the knowledge in the topic of inverse analysis by the application of three special features.

- The soil is treated as a heterogeneous material by employing the RFEM method. In this way soil properties vary in space and thus the inverse analysis solution is dependent on both the parameter values and their spatial distribution.
- The adoption of a pseudo-time scheme means that observations change over time, while soil properties and their spatial variability remain the same. Implementing several construction stages in a slope excavation means that the product of the inverse analysis will have to comply with the boundary conditions in a different concept.
- Practice-oriented interpretation of the results highlights the applicability of the inverse analysis, as well as its importance for an integrated design.

Specifically, the thesis attempts to recognize spatially variable soil parameters in a staged slope construction problem by employing displacement profiles. As a result, a stochastic approach on soil parameters is provided, and conclusions can be drawn on inverse analysis schemes and the utilization of the results in engineering practice.

1.3. Research questions

Research on the inverse analysis in Geo-engineering has grown significantly, especially over the latest years, as the increased computational power available recently is able to remove simulation hindrances. While, several schemes have been applied on a variety of geotechnical problems, there is still an abundance of persisting questions. The scope of this thesis is addressing the following queries:

- How do different inverse analysis schemes apply to the multi-staged RFEM model?
- What is the performance of these methods?
- Which hindrances can be met when applying an inverse analysis?
- How reliable are the results of the process?
- What are the benefits of applying inverse analysis as part of an integrated design?

1.4. Report structure

This report presents every step in the application of inverse analysis methods in the said RFEM model. The operation of the inverse analysis algorithms, their implementation in the RFEM framework and the calibration of model parameters are described in-depth. Moreover, the efficiency of the methods is reviewed, as well as their practice-oriented potential. Specifically, the chapters composing the report are:

- The *Theoretical Background*, where the philosophy and mathematical foundation of the examined methods are described. Also, the changes necessary, so that the methods fit the current investigation, are pointed out.
- The *Method comparison*, in which the performance of the presented inverse analysis schemes in a simple FEM model is assessed. Thus, the fittest methods can be selected.
- The *Inverse Analysis*, where the outcome of the analyses is discussed. Arguments are made regarding the operation of the algorithms on the multi-staged slope excavation model and the accuracy of the results .
- The *Reliability Analysis*. Here, the stochastic nature of the inverse analysis results is interpreted as an equivalent reliability.
- The *Conclusions*, where the comparison of the inverse analysis methods is resolved and their merits and drawbacks are outlined.
- The *Recommendations for future research*, in which some suggestions considering the extension of the thesis' topic.

2. Theoretical Background

This chapter focuses on the elaboration of the examined inverse analysis methods. After a brief introduction to the Bayesian theory basics, the five selected methods are presented. For each one of them, the general philosophy is elaborated, the advantages and drawbacks are listed and finally their application in the examined inverse problem is analyzed. Namely, the inverse analysis schemes are the: Markov Chain Monte Carlo, Maximum Likelihood, Genetic Algorithm, Kalman Filter and Particle filter. While mathematical research offers a variety of available inverse methods, the final selection is reduced to these five mainly due to performance criteria (robustness, problem suitability, computational tolls), time restrains and prior knowledge.

2.1. Brief introduction to the Bayesian approach

Baye's Theorem, as proposed in 1763, states that the probability of an event is proportional to prior knowledge of the parameters affecting it and their likelihood of happening. Bayesian analysis offers an alternative view on statistical inference problems, compared to the traditional frequentist thinking. Instead of regarding probability as a the frequency of event appearance after multiple trials, the Bayesian approach interprets it to a degree of belief on a statement about unknown quantities (Wikle & Berliner, 2007).

$$p(x \mid y) = \frac{p(y \mid x)\, p(x)}{p(y)} \propto p(y \mid x)\, p(x) \tag{2.1}$$

$$Posterior \propto Likelihood * Prior \tag{2.2}$$

According to (Glickman & Van Dyk, 2007) and (Wikle & Berliner, 2007):

- $p(y \mid x)$: This is the *likelihood function* (data distribution), or the process that connects parameters (*x*) to evidence (*y*). It often utilizes observed data which are treated as a function of model parameters. Essentially, the likelihood function expresses the probability of observing *y* given the parameter value *x*. In the examined inverse problem, the RFEM process plays an imperative role in forming the likelihood function. Also, most methods assume that the likelihood function is too complicated to treat as a known, and thus orchestrate schemes in order to bypass its direct utilization.

- $p(x)$: This term denotes the *prior distribution*, which is the knowledge of the parameter state a priori. Based on the approach followed, the analyst can select a prior distribution in a subjective or objective manner. In terms of this report, the prior is the initial estimate of soil parameter distributions.

- $p(y)$: The *marginal distribution*, which is taken as in formulation (Eq. 2.3). often, it is not possible to be solved analytically and is treated as a "normalizing constant" in Baye's Theorem (Wikle & Berliner, 2007).

- $p(x \mid y)$: The *posterior distribution*, which is an update on the knowledge of model parameters (*x*), in the light of evidence (*y*). The posterior is proportional to the probability of the parameters (prior) and the likelihood of encountering the observed evidence (likelihood). For the Bayesian approach, the posterior distribution is the foundation on any analysis; assessing it renders stochastic evaluation of hypotheses or calculation of event probabilities possible. Specifically, now the inverse analysis aims into providing the posterior distribution of soil parameters, allocating effective parameter values based on the displacement measurements (Glickman & Van Dyk, 2007).

$$p(y) = \int p(x)\, p(y \mid x) dx \tag{2.3}$$

Concluding, the essentials of Bayesian inference were presented in brief. While deeper understanding of the topic demands an extensive analysis, the basic tools of this method are elaborated so that the application and operation of the following inverse methods is comprehensible.

"*Baye's theorem is to the theory of probability what Pythgoras' theorem is to geometry*." (Jeffreys, 1931)

2.2. Markov Chain Monte Carlo

The Monte Carlo method (MC) is a stochastic analysis method that aims to optimize a function by relying on random sampling of input parameters. Essentially, the MC utilizes massive numbers of simulations (realizations) that generate draws of parameter values out of a given probability distribution (prior), producing different input combinations (Gilks, Richardson, & Spiegelhalter, 1996). The randomness involved in the procedure poses as an efficient way to bypass the complexity of a function, like an FEM model. Thus, the MC is a resourceful alternative for dealing with problems where parameter uncertainty is dominant and optimization is required (like inverse analysis, locating the worst case scenario, etc.) or for analysing random draws from a probability distribution (reliability analysis).

When applying the Monte Carlo method, it is essential that the sample is representative of the reality, or in other words, that the sampled values are picked with a frequency matching their probability of appearance. In order to achieve this while dropping the realization number needed, the theory of a Markov Chain process is applied, in a form of a component algorithm. A Markov Chain is a procedure in which the next step is only dependent on the current state and not the path followed until reaching it (Gilks et al., 1996).A new step is picked from a distribution centred around the current parameter values. Then, this procedure creates a "random walk" which attempts to approach the target distribution. Besides, to ensure that the path converges to the target, an algorithm-component is added, judging which values are accepted, using error minimization as a criterion. Then, method convergence is also guaranteed when lower numbers of trials are involved. The Markov Chain is implemented through the Modified Metropolis-Hastings algorithm, a procedure allowing for representative sampling using the target distribution as a basis (Chib & Greenberg, 1995).

Added together, a Markov Chain Monte Carlo (MCMC) is a method of competently simulating a complex probabilistic process while preserving the accurate portrayal of the parameter statistics. Applying the Markov Chain part enhances the speed of the analysis, assuring representative sampling even with fewer realizations.

Markov Chain Monte Carlo is quite popular in probabilistic design and analysis for a variety of reasons. Firstly, the mathematical background involved is quite basic, making it approachable to any engineer. Moreover, complex issues of the function can be easily solved, like parameter interdependency or coupled phenomena. In addition, the MCMC is capable of operating with a large number of parameters (interdependent or not), which can freely take any value in the domain.

On the other hand, MCMC holds some major disadvantages. The most crucial of all is the heavy toll on computer resources and the extended simulation time required. Although the Markov Chain algorithms can reduce the realizations needed, the number of iterations is still a problem. Also, the fragility of the representing ability of the sampling is always an issue to be kept in the user's mind. Finally, mathematically the MCMC suffers from two issues: dependency on starting values (burn in) and autocorrelation (the steps of the chain are dependent by definition) (Papaioannou, Betz, Zwirglmaier, & Straub, 2015).

Of course, Markov Chain Monte Carlo can be applied in the case of inverse analysis after specific modifications. In this thesis, the relationship between input and output of the RFEM code is deemed too complex to deal

with analytically, and so, MCMC is used to simulate it. The goal is to reach the given measurements, and so, the error of the procedure is defined as the deviation between them and the RFEM generated measurements. The point of applying a Markov Chain is that every new realization will produce an error at least lower than the previous run. Then, the new input set is accepted, and the new step of the path is simulated. With every iteration, another "ring" is added to the Markov Chain, connecting the initial position to the approximate area of the accepted solutions, which is never accurately calculated, but rather depicted as a cloud of realizations in a confined region.

In each iteration, values of the chosen parameters are drawn from their probability distributions by employing the Modified Metropolis-Hastings algorithm, which can be implemented in several variations. This algorithm (Soubra & Bastidas-Arteaga, 2014) includes the following steps, also illustrated in the respective flowchart (Fig. 1):

- The current variable state enters as an input.
- A new, proposal state is formulated, by randomly picking a value from a proposal distribution that has a mean equal to the current value and a user-defined standard distribution.
- Following, the acceptance of the new value is determined. In simple terms, if the new value has a larger probability than the current one, both calculated for the terms of the target distribution, then the acceptance is guaranteed. Else, the probability of acceptance is equal to the ratio of the prior probabilities for the proposal to that for the current state.
- The final variable state exits, with variables having changed where the proposal was accepted.

The new step is then used as an input in the RFEM and its divergence from the goal measurements is calculated. If the error is lower than that of the previous iteration, the new draw is accepted, otherwise, the realization is re-run and the algorithm is repeated. In this way, only realizations that are beneficial to the advance of the analysis are kept and the consistency between the posterior distribution and the total sample distribution is spoiled. Nevertheless, there is effect between random draws of variables to some extent, as the performance of the input set dictates the acceptance of the new step.

2.3. Maximum Likelihood

The Maximum Likelihood (ML) method is yet another approach of conducting an inverse analysis. It employs a series of iterations aiming in maximizing the likelihood of producing the observation measurements. Mathematically, likelihood is directly analogous to the conditional probability of generating the observations [x^*] under the hypothesis parameters [p] (Eq. 2.4) (Myung, 2003).

$$L = k^* f(x^* | p) \tag{2.4}$$

Setting an objective function that includes the error [x^*-x] as the main term, the problem can now be re-stated as the attempt to minimize this function (Eq. 2.5) (Ledesma, 2014):

$$J = (x^*-x)^T * C_x^{-1} * (x^*-x) \tag{2.5}$$

(Where C_x is the covariance matrix of the measurements)

The method allows for a consecutive alteration of the hypothesis in a manner that the objective function gradually drops. In order to change the parameters, ML heavily relies on the derivation of the objective function and correcting the input so that the error is reduced. The process of deriving the product of the simulation by the input parameters generates the sensitivity matrix [A], which expresses the correlation

between them for a specific iteration. After a succession of steps, the method is expected to converge to a single solution [p_{k+1}].

$$[A] = \frac{dp}{dx} = K^{-1}\left(\frac{df}{dp} - \frac{dK}{dp}x\right) \tag{2.6}$$

$$\Delta x = (x^{*} - x) \tag{2.7}$$

$$\Delta p = (A^{T} C_{x}^{-1} A)^{-1} A^{T} C_{x}^{-1} \Delta x \tag{2.8}$$

$$\begin{aligned} p_{k+1} &= p_{k} + \Delta p_{k} \\ (J(p_{k+1}) &\leq J(p_{k})) \end{aligned} \tag{2.9}$$

The Maximum Likelihood approach possesses some advantages when applied in an inverse analysis. Mainly, it is the only examined method that does not include a factor of randomness. Moreover, it can be quite fast comparatively to the other methods, when a relatively simple model is back-calculated. Lastly, a fact to its favour is that the method does not require a parameter distribution as an input, but rather a decent guess for the initial variable state.

On the other hand, major drawbacks take a toll on the Maximum Likelihood method. The greatest one is the constant need of handling the stiffness matrix. In a complex FEM run, this can mean a vast increase in computational effort and time, restricting or prohibiting the analysis. Also, it should be noted that each chain created can only lead to a single solution.

The Maximum Likelihood method is mostly competent when applied in an FEM model. An FEM simulation uses the stiffness matrix [*K*] as its cornerstone, connecting actions *f* to reactions [*x*]. Having discussed the basics of the ML, the sensitivity matrix can now be estimated easily in each iteration.

After this statement, the application of ML in the framework of an inverse analysis can be broken down according to the flowchart (Fig. 2):

- The covariance of the observations has to be calculated.
- The error function is formulated.
- An initial "guess" is given to the examined parameters.
- The RFEM model produces the first results.
- The error function is calculated. If the error is larger than the tolerance, then the sensitivity matrix is estimated, and following, the change of the input for the next iteration.
- A new set of input is tested.

In this method, no prior distribution is used for the parameters, whose change is tightly interconnected. The evolution of this method is structured on a succession of steps, creating a path from the initial guess to a single solution.

2.4. Genetic Algorithm

After his extensive research in the Galapagos Islands, Charles Darwin established his Theory of Evolution on the famous notion: “survival of the fittest”. Much like the evolutionary procedure of life, a genetic algorithm is an optimization algorithm that exploits the concept of this theory.

In general, the algorithm is organized in generations. Each generation is composed of individuals that represent a set of inputs for the examined function. The result of each set is judged by the user-defined "fitness function", and so the efficiency of each individual is assessed. As long as an adequate solution is not achieved, a new generation is to be produced, based on three main routines (Coley, 1999), (Mitchell, 1996):

- **Selection**: The individuals are ranked based on their fitness score. Following the Theory of Evolution, the fittest ones are selected for "breeding", as they are the most potent candidates for producing a solution in the next generation. The "parents" are randomly selected, but the fittest have a higher probability of selection.
- **Crossover**: The selected individuals are used to generate "offsprings", which act as the individuals of the succeeding generation. This is done through a user-defined function, designed to exploit the best parts of the "gene pool" and also to include a random factor.
- **Mutation**: This routine adds the factor of randomness, completing the optimization procedure. An offspring is created by randomly altering the parent, and not through a crossover with another individual. In this way the algorithm is potent to better explore the variable domain, a feature critical in case the process is "trapped" when following the original breeding routine.

There are certain benefits in applying a genetic algorithm in this thesis. First of all, a genetic algorithm poses as an ideal solution for stochastic analyses, as it bypasses the complex analytical solution of a back calculation (Carr, 2014). Also, it can deal with large numbers of variables and is potent to locate multiple optima; as the individuals of each generation search in parallel, the process is able avoid being trapped to a single local optimization point. Thus, a number of solutions can be provided.

On the other hand, genetic algorithms bear significant drawbacks. The greatest one is that the procedure might not locate the global optimum. As the starting generation is randomly selected, but also the crossover and mutation routines include a random factor, it is possible that the solutions miss the "best fit" (Carr, 2014). Hence, the algorithm requires an adequate initial sample to ensure competency. Furthermore, the GA can be time consuming and exhausting on computational resources.

In particular, the Genetic Algorithm in this thesis is applied in order to assess the set of parameters that offers the closest prediction of the measurements, when applied in the RFEM model (Fig. 3). The fitness function is simply the inverse of each individual's error, meaning that the closer the attempt, the higher the fitness score. Moreover, as the values of the input parameters are non-binary, the crossover is consisted of applying a weighted average between parents, using the fitness score, and then allowing for randomness through a draw from a normal distribution. Finally, the mutants are created by randomly picking a parameter from a normal distribution that has the parent as an average and a user-defined standard deviation. It should be noted that the routines are applied separately for each parameter examined, however, their fitness is calculated for the set. This means that the calculation of the new generation seems independent for each parameter, but their evolution is tightly connected to the performance of their set (Haupt & Haupt, 2004).

2.5. Kalman Filter

The Kalman Filter (KF) is a recursive Bayesian algorithm that reads noise-containing measurements at multiple time stages and estimates the parameter values that produce them. Also, it is potent to operate in a steady state, looping until the convergence criteria are met. The cornerstone of the Kalman Filter is the assumption that variables and measurement noise of a Gaussian distribution are processed through a linear function, and so, the final product will also be distributed in a Gaussian manner (Bishop & Welch, 2001). It heavily relies on the successive handling of the parameter-variable covariance matrix and the covariance matrix of the

measurement error (noise). Eventually, both matrices contribute to the "Kalman Gain", which expresses the measure of change that each variable has to undergo.

Many variations of the Kalman Filter exist, however, two of them must be highlighted in the margins of this thesis: the Extended Kalman Filter (EKF) and the Ensemble Kalman Filter (EnKF). The first one is a version of the original method that accounts for a nonlinear function. Its application is essential, considering the complexity of the FEM code, connecting the observations and the parameters in an ambiguous way. The main idea is that the nonlinear function can be approximated at points with its Jacobian, and so, the original Filter can be applied with slight modifications (Terejanu, 2008). Moreover, the Ensemble Kalman Filter is a combination with the Monte Carlo method. Since maintaining and advancing the variable covariance matrix can be computationally exhausting, the EnKF employs an ensemble of variable values that are used to calculate the covariance in each assimilation step (Evensen, 2009). Each ensemble member is composed by random values for each parameter, and thus, represents a different parameter state. The first ensemble is assembled by random Monte Carlo sampling from prior distributions and its size has to preserve a representative sample. Then, the change of each ensemble member is calculated based on the employed Kalman Filter version. Ultimately, the parameter posterior is the distribution of the ensemble.

According to literature (Evensen, 2003):

- Having an ensemble (A) of N states, then the process is mathematically described as follows:
- Ensemble perturbation matrix: $A' = A(I - 1_N), where: 1_N = \frac{1}{N} square\ matrix$
- Ensemble covariance matrix: $P_e = \frac{A' * (A')^T}{1-N}$
- Measurement covariance matrix: $R_e = \frac{Y' * (Y')^T}{1-N}$, $with\ Y\ containing\ the\ measurements\ noise$
- Kalman Gain: $K_g = P_e J^T (J\ P_e J^T)^{-1}$, $where\ J\ is\ the\ Jacobian\ for\ each\ state$
- Change in the ensemble: $DA_{1:N} = K_g\ (D - F(A_{1:N}))$
- New ensemble: $A_{new} = A + DA$

The Ensemble Kalman Filter yields some major advantages. Firstly, the use of an ensemble allows for a faster and easier calculation, bypassing the handling and transfer of the variable covariance matrix among iterations. Secondly, the product of this procedure is not a mere solution value, but rather a statistical distribution of the solution. Lastly, the distribution of the variables advances by individually evolving each member of the ensemble, allowing for a simpler and quicker process. However, the samples are not independent, as the method guarantees their interdependency.

So far, two main disadvantages of the KF can be noted. The first problem arises in dealing with observation noise. The user has to manually add a noise term in the measurement field, so that the observation covariance can be defined. This noise must have a mean of zero, but its standard deviation is vague and decided by the user. Also, the Kalman Filter is mostly fitted for linear problems, due to its mathematical foundation. Applying the Extended form means that the Jacobian of the FEM has to be calculated, taking a toll on computational performance, without guaranteeing proportionally adequate results.

In the context of this thesis, the Extended/Ensemble Kalman Filter approach offers some significant insight. The application procedure can be narrowed down to simple steps following the respective flowchart (Fig. 4):

- Generate the first ensemble using Monte Carlo sampling from the given parameter distributions.
- Set the observed measurements and add a noise component. The noise is chosen to be normally distributed, with a mean of zero and a standard deviations equal to one thousandth of the observed values.

- Estimate the ensemble and measurement covariance matrices.
- Calculate the Jacobian for each ensemble member, if the Extended form is applied.
- Calculate the Kalman Gain for each ensemble member.
- Calculate the change of each ensemble member.
- Create the new ensemble by adding the changes to each member.
- Iterate until the error given by the ensemble mean is lower than the tolerance.
- Finally produce a vector of means and standard deviations for every parameter in the final ensemble.

The visualization of each step in a variable space is a cloud of points. The final cloud of each stage will have a mean that is within tolerance from the solution. However, as every ensemble individual is directed towards convergence, the final cloud may include a number of points in tolerable proximity to the solution, thus forming a domain matching the analytical surface. Also, the path followed by the mean of each cloud can be a point of interest.

Last but not least, the Kalman method provides a distribution of the solution. This means that further statistical processing can take place, evaluating the accuracy of the prior distribution and assessing the representation of the data.

2.6. Particle Filter

The Particle Filter (PF) is a genetic type Monte Carlo scheme that can be utilized in inverse analysis. It is similar to the Ensemble Kalman Filter, in the sense that an ensemble of individuals representing variable states (particles) is the core of the procedure, which is initially sampled using the Monte Carlo method. PF is based on the idea of Importance Sampling; a target distribution is approached by sampling from a prior distribution, and weighting in order to approximate the target. Thus, the particles are weighted based on their performance and the final solution is the weighted average of the ensemble (Orhan, 2012).

The first version of this method is the Sequential Importance Sampling (SIS), where the weight of each particle is evaluated based on its value in the previous assimilation step and its current performance (Doucet & Johansen, 2009). In this approach, the particles remain constant and the solution is altered due to the change in weight. However, this process brings up the issue of degeneracy, where after several iterations of the SIS, only a few particles are left with a significant weight. This means that computational effort is spent on particles with extremely low importance and also the procedure deviates from the solution, as only a limited number of particles have an impact. The degeneracy problem is corrected with the use of resampling, which includes drawing a new group of particles according to the discrete distribution of the weights. Then, the new particles are all attributed with the same weight, but important particles will appear more often in the group. After resampling, the solution is the average of the new ensemble. Also, as resampling is heavily affected by the weight of each particle, the importance of its performance is clearly significant and so the PF comes close to the philosophy of a genetic algorithm.

This approach leads to the second version of Particle Filtering, the Sequential Importance Resampling (SIR) (Arulampalam, Maskell, Gordon, & Clapp, 2002). This approach resamples in the end of every assimilation step, and so the weights of the next run are only a matter of particle performance. However, the problem of sample impoverishment arises. As a computer can only simulate the cumulative distribution of the weightings in a discrete manner, then particles with larger weight are picked multiple times, whereas lower weight particles are left out. This leads to an ensemble of reduced variance, finally collapsing to the same particle in a few assimilation steps. The Regularized Particle Filter (RPF) employs a tactic dealing with this issue; it applies a distribution kernel, converting the discrete cumulative distribution into a continuous one.

As a part of the "Filter" family of methods, the particle filter approach yields mostly similar benefits and drawback as the Kalman Filter. Their most obvious differences can be easily summed up. In contrast to the EnKF, this method operates better in non-linear problems, posing as an ideal solution in RFEM analysis. However, in order to render it efficient and computationally affordable, as well as to overcome the problems discussed, the algorithm increases in complexity.

The Particle Filter/ SIR algorithm can be a competitive alternative for an efficient inverse analysis. As with the Kalman Filter, each particle represents a variable state, and as a group, the final product is going to be a solution distribution, with the mean being within error tolerance. The sequence for the RPF as applied in this thesis is listed according to the flowchart (Fig. 5) (Arulampalam et al., 2002):

- Generate the first particle group using Monte Carlo sampling from the prior parameter distributions [*X*].
- Calculate the error of every particle. The weight is assigned based on a Gaussian distribution (Eq. 2.10), thus the closer to the solution, the fitter the variable combination. Then, normalise the weights by dividing with their sum (Eq. 2.11).
- The resampling procedure takes place. The discrete cumulative distribution of the weights is used to pick the new particles. Afterwards, the weights are all set to equal.
- A distribution kernel is fitted on the new weights of the resampled ensemble. As suggested by the literature, the Epanechnikov kernel is the most valid choice for converting to a continuous distribution, since all weights are equal after the resampling. However, a slight variation also tested in this thesis includes the employment of a normal distribution kernel in the place of Epanechnikov.
- A random sample is taken from the normal distribution kernel as a measure against sample impoverishment [ε]. Multiplied by the kernel bandwidth, it offers the change of every particle [*DX*] (Eq. 2.12). This is added to the resampled ensemble (Eq. 2.13).
- Iterate until the error of the ensemble mean is lower than the tolerance.
- Finally produce a vector of means and standard deviations for every parameter in the final group.

In this case, the Cholesky decomposition is unnecessary.

$$w_i = \frac{1}{\sqrt{2\pi\sigma_{error}^2}} e^{-\frac{error_i^2}{2\sigma_{error}^2}} \tag{2.10}$$

$$w_i = \frac{w_i}{\sum_{i=1}^{n} w_i} \tag{2.11}$$

$$DX = h * \varepsilon \tag{2.12}$$

$$X* = X + DX \tag{2.13}$$

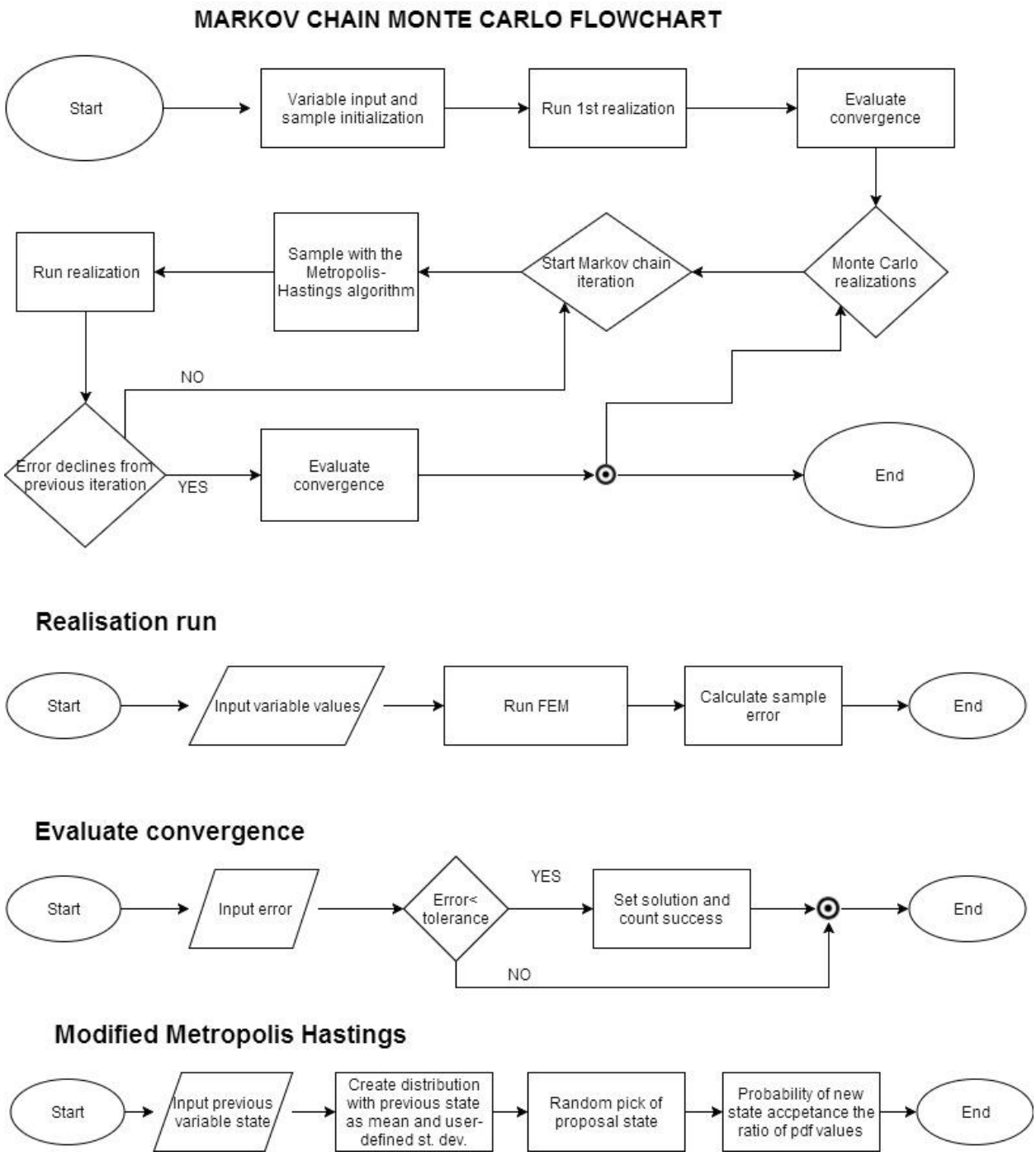

Fig. 1: Markov Chain Monte Carlo flowchart

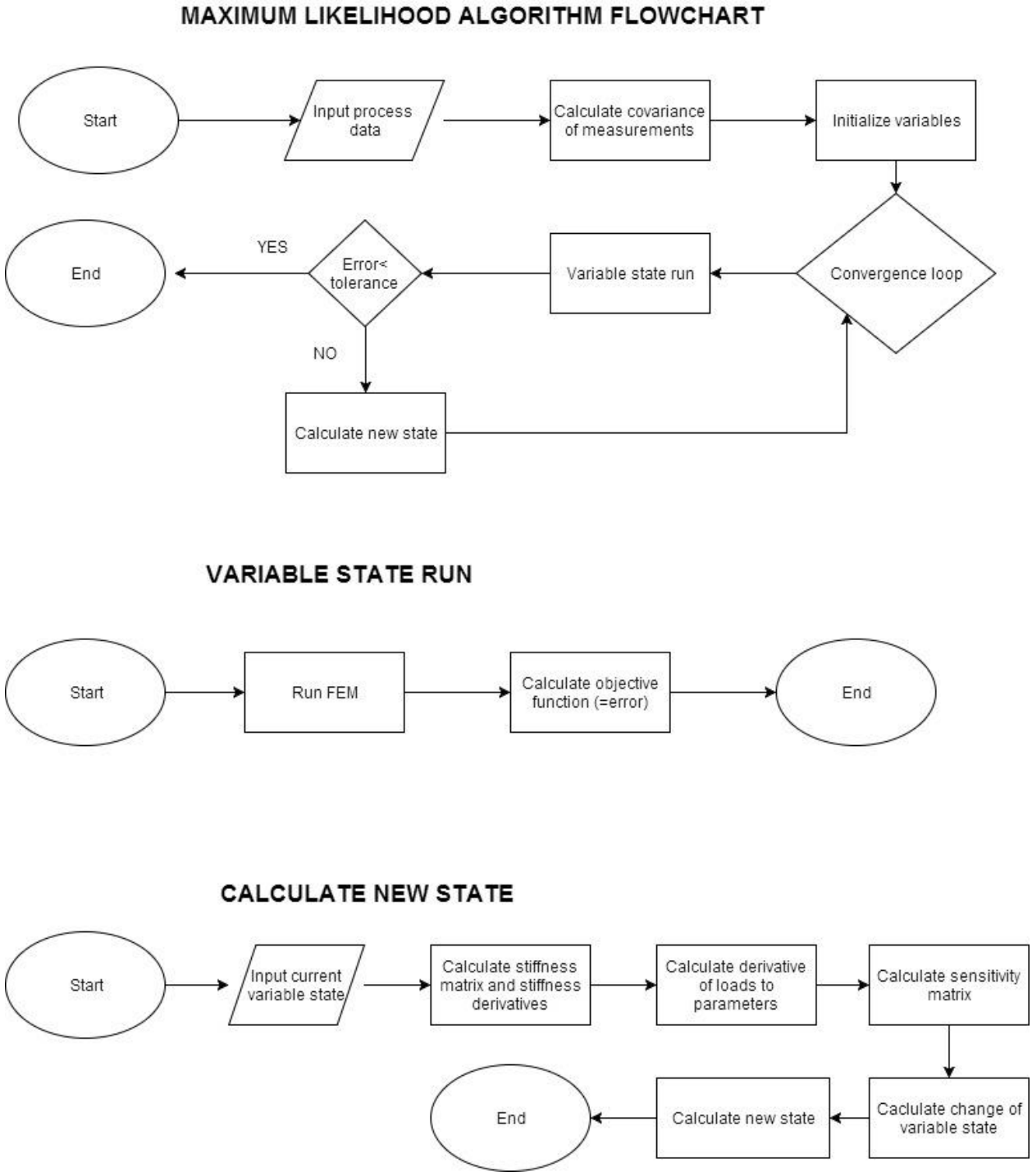

Fig. 2: Maximum Likelihood flowchart

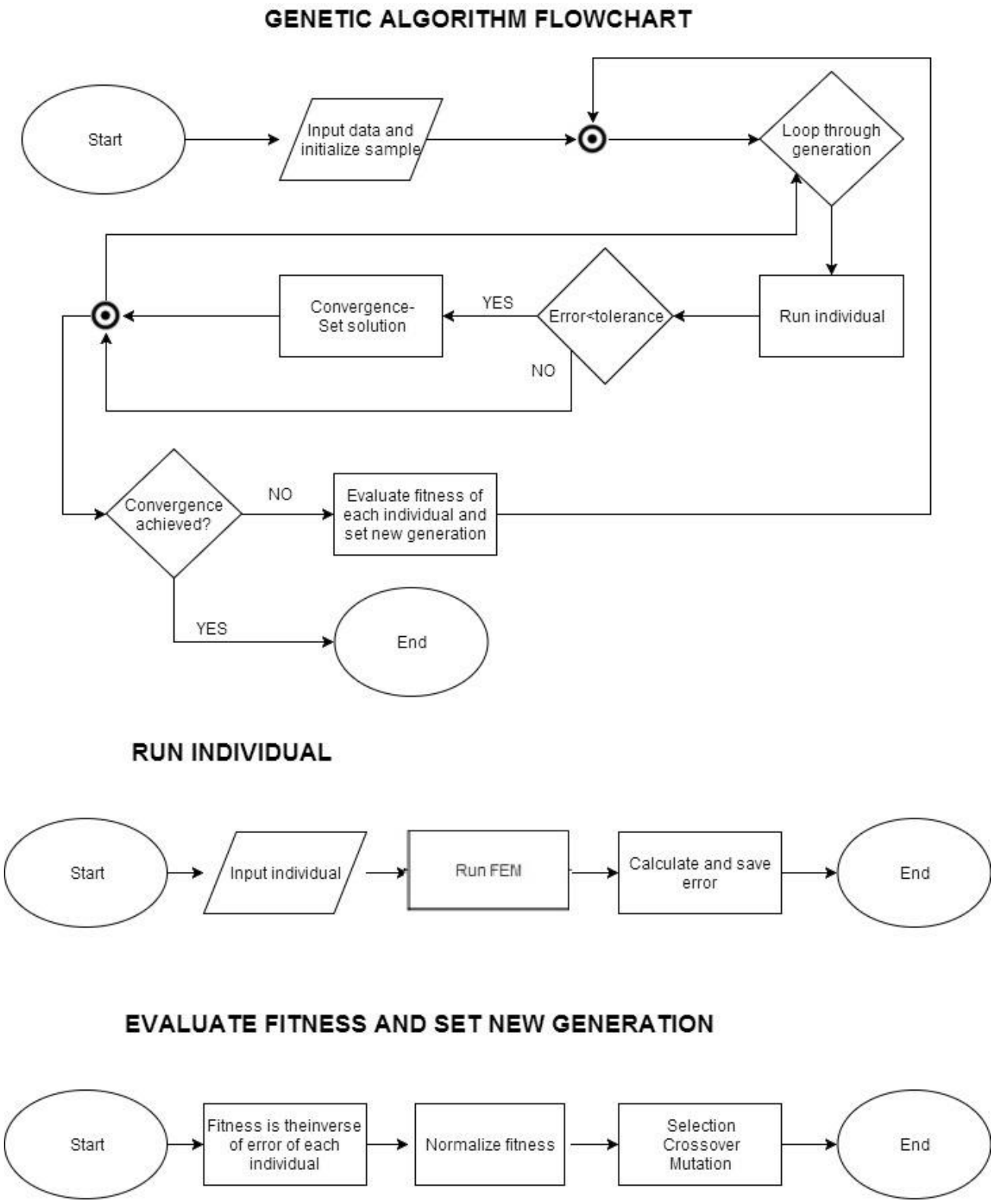

Fig. 3: Genetic Algorithm flowchart

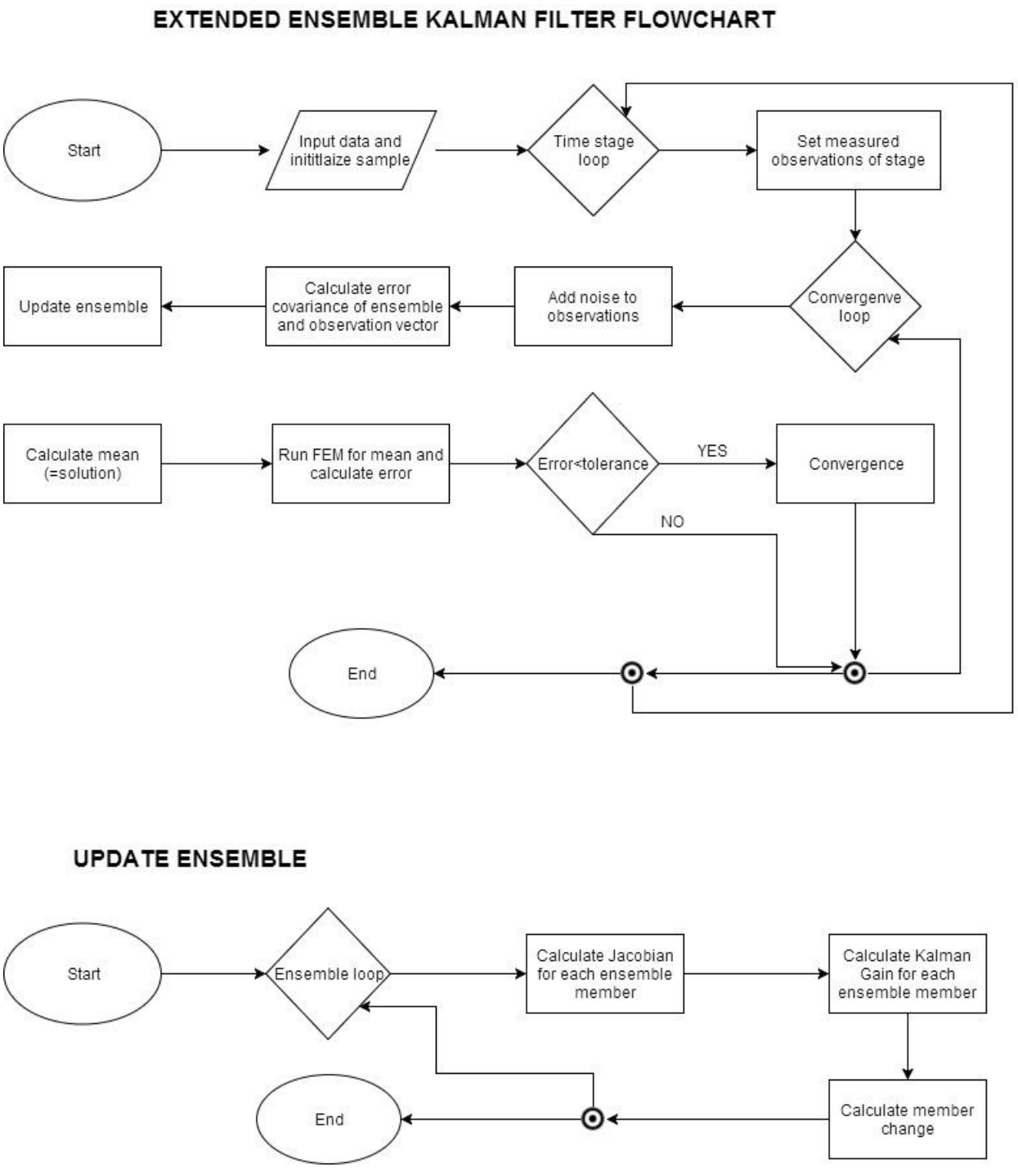

Fig. 4: Extended Ensemble Kalman Filter flowchart

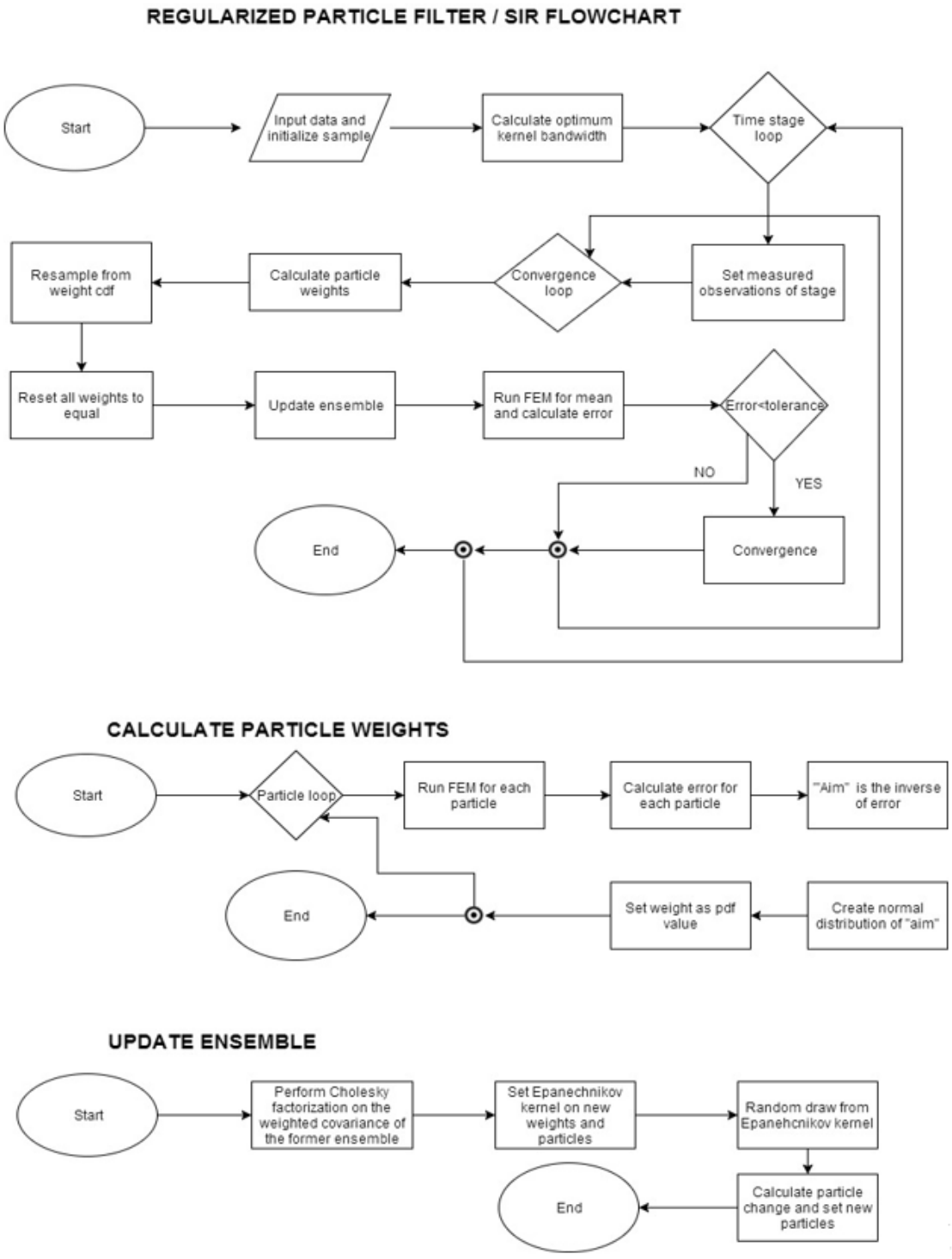

Fig. 5: Regularized Particle Filter flowchart

3. Method Comparison

After compiling the various inverse analysis methods, testing is required in order to evaluate, calibrate and rank their performance. Therefore, Finite Element codes are orchestrated as examples, simulating engineering problems.

3.1. Example: Waterflow in 1D

The first test is selected to be a simple one-dimensional waterflow problem. It is analogous to horizontal flow through a soil-containing pipe, with a fixed water head at both ends. The observation measure is the inflow/outflow of the system. Finally, the variables of the problem are: the permeability of the soil, the diameter of the pipe and the length of each pipe segment, which are homogenous properties of the domain.

It should be noted that this example is not optimum for calibrating the algorithms:

- It is unrealistic, as usually waterflow cannot be accurately measured, thus not fitting as an observation measure.
- It would be better to have a vector of observations and such would be the water head at the inner nodes of the model. However, these values are a function of the ratios of the parameters and not the parameters themselves, creating problems in the simple case where properties are uniform in the model.
- The problem is too simple for a FEM solution. It can be easily solved analytically and the properties of the simulation have no impact on the result.

Despite these drawbacks, the first example is employed in order to observe the behaviour of the inverse analysis mechanism for each method. Finite Elements are only used in order to check the consistency of a simple model with the compiled algorithms.

Table 1: Distribution of variables and outflow of mean values for the 1D waterflow inverse problem

VARIABLE	MEAN	ST. DEVIATION	FEM OUTPUT (m^3/s)
PERMEABILITY (m/s)	1.00	0.10	0.0105
DIAMETER (m)	0.20	0.02	
LENGTH (m)	1.00	0.10	

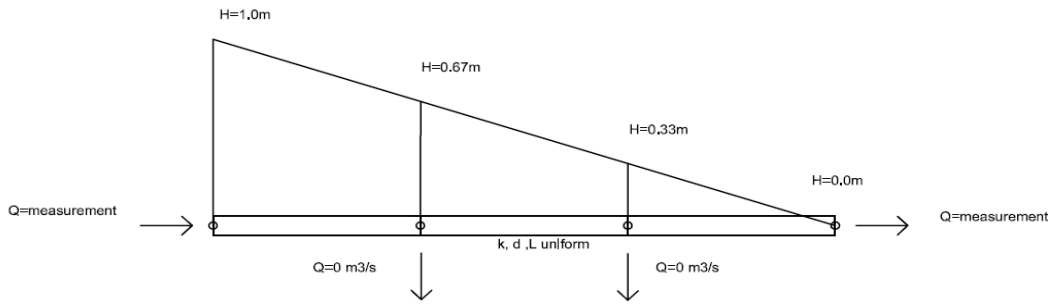

Fig. 6: Mesh and boundary conditions of simple 1D waterflow FEM model

A simple mathematical description of the effect of every variable on the FEM result clarifies that the product is linear, quadratic and inverse for the permeability, diameter and length respectively. The analytical solution is plotted, so that it can fit on the inverse analysis results.

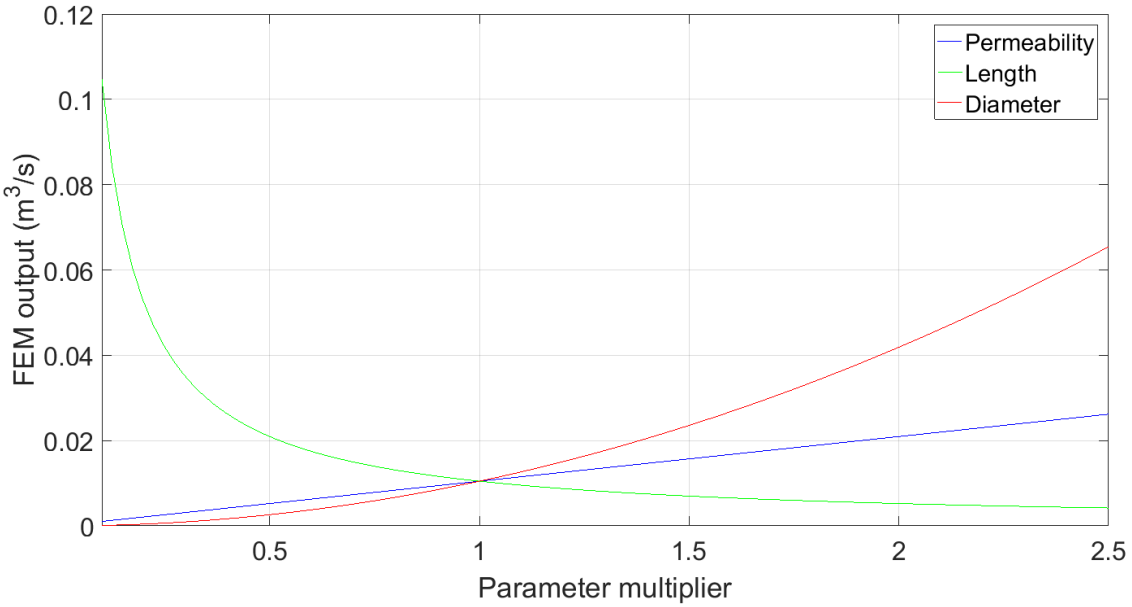

Fig. 7: Sensitivity analysis of parameters on outflow

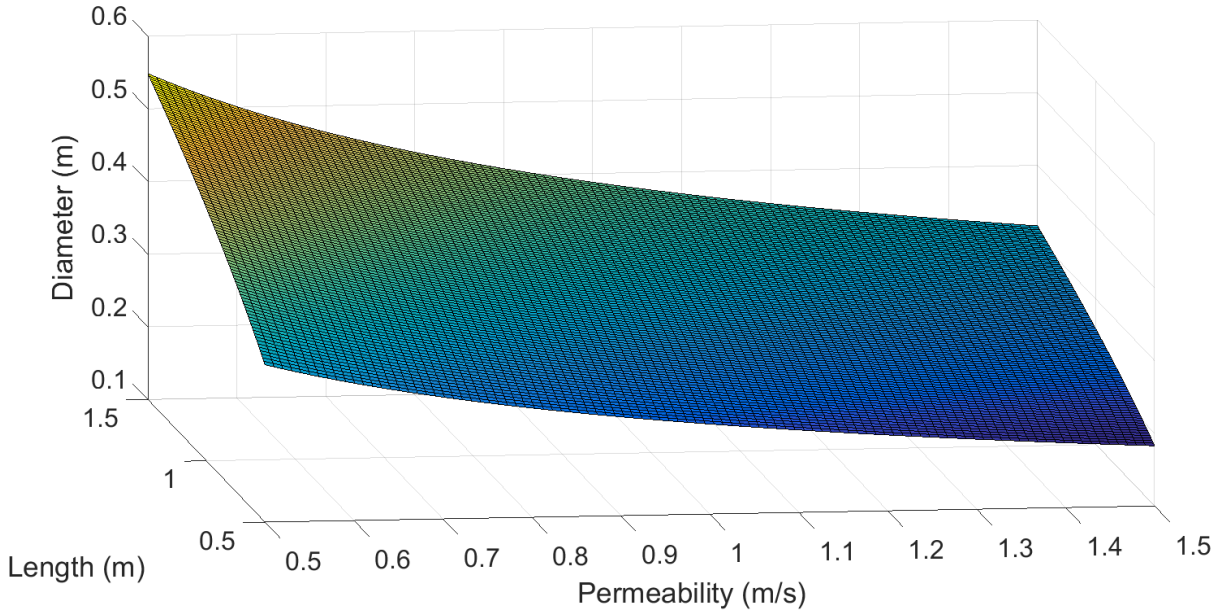

Fig. 8: Analytical solution of the inverse problem

Then, each method is tested on this simple FEM model. For the inverse analysis, the observation is set to be the flow of the mean values altered by a multiplier (0.5, 1.2, 1.8, 2.5 and 10.0) so that the analysis has to track the solution in a sub-domain different from the prior. All inverse analyses work with a 10^{-5} tolerance.

Method performance is evaluated based on the number of FEM repetitions needed (computational strain), the number of solutions provided, as this is the size of the output sample, and time required. Initially, since the FEM calculations are going to occupy most of the analysis time, it is essential that their repetitions are kept to the minimum (originally, the ratio of FEM runs to inverse analysis iterations should have been presented, but it is not as some methods do not operate in such terms). Following this, the size of successful iterations is significant, since it produces the posterior distribution of variables or parameters. Last but not least, time consumption is a measure of method readiness as opposed to problem stiffness.

Table 2: Performance of the five inverse analysis schemes in the described criteria

PERFORMANCE EVALUATION			
METHOD	**FEM Iterations**	**No. of Solutions**	**Time (sec)**
MCMC	157844	978	88
MAXIMUM LIKELIHOOD	117	1	3
GENETIC ALGORITHM	26400	100	4
EXT. EN. KALMAN FILTER	11173	Distribution	38
PARTICLE FILTER	12	Distribution	3

3.1.1. Markov Chain Monte Carlo results

The Markov Chain manages to track an area where the error stands within tolerance to the solution. Generally, the chain starts with larger leaps, but as it progresses, error decrease is disrupted and new points tend to be accepted closer to the initial ones. When the chain achieves the error tolerance, the other branch of the Markov criterion is activated; the error of the new variable state has to be within tolerable margins. Thus, the chain starts to "wander" around the solution. This is highly evident when the chain is plotted together with the analytical solution (Fig. 9).

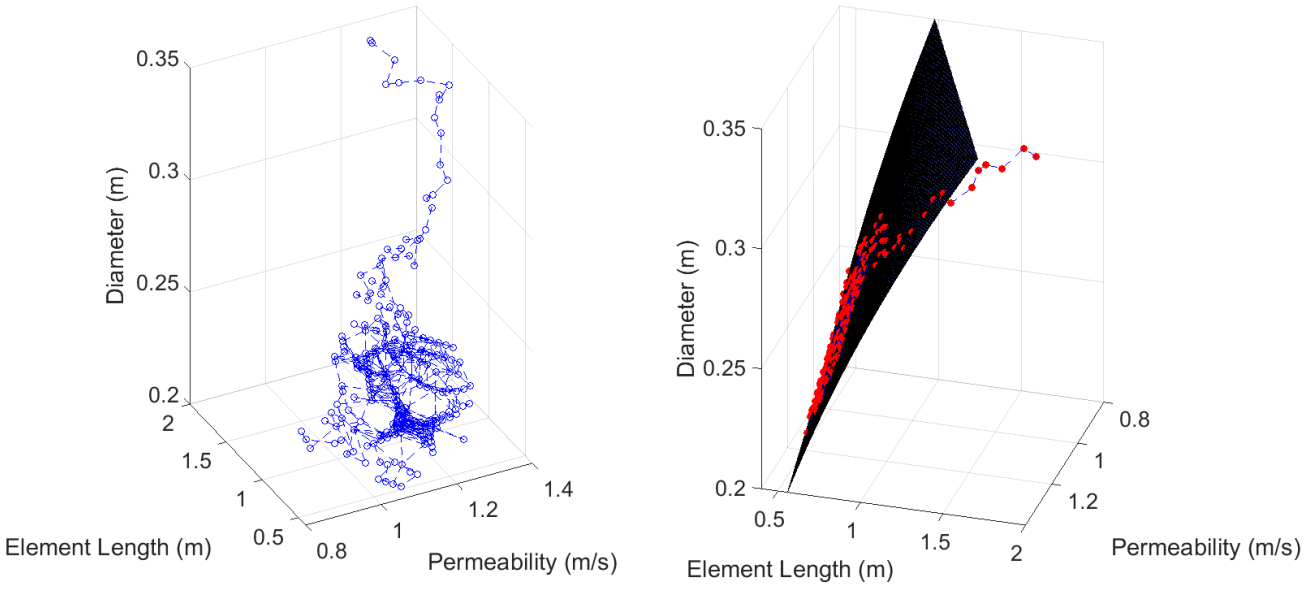

Fig. 9: Markov path for the 1D waterflow inverse problem

Notably, in the histogram plot (Fig. 10) the final distribution approaches a normal one, a property which might be problem-sensitive. Comparing the normal distributions, the sampled one (all values sampled) is close to the solution (only accepted values), but has to be at least wider, since the second one is its subset (Fig. 11).

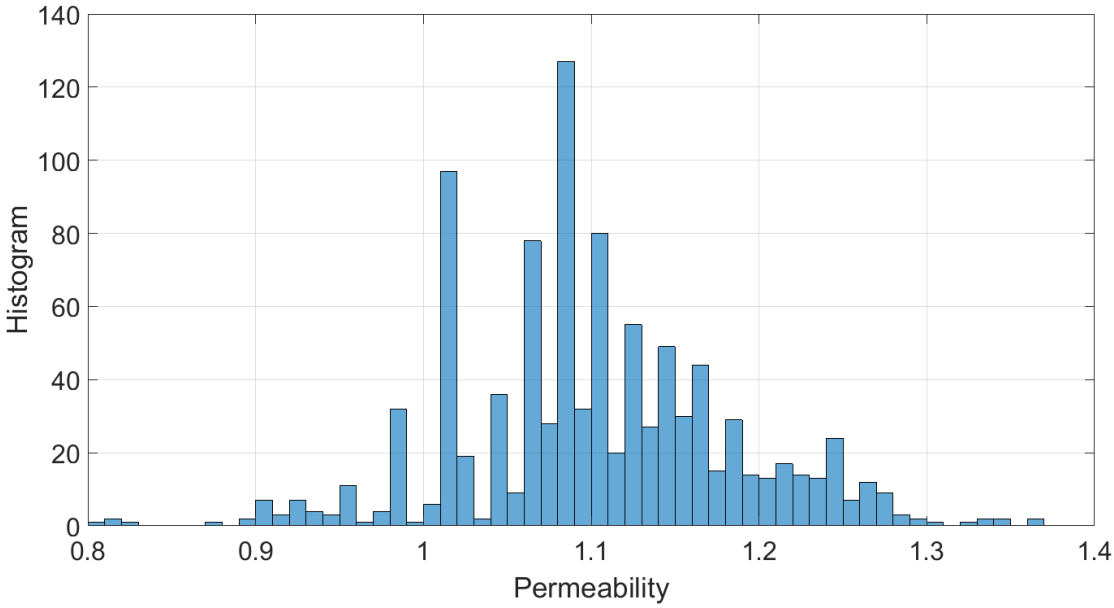

Fig. 10: Histogram of sampled permeability values in MCMC

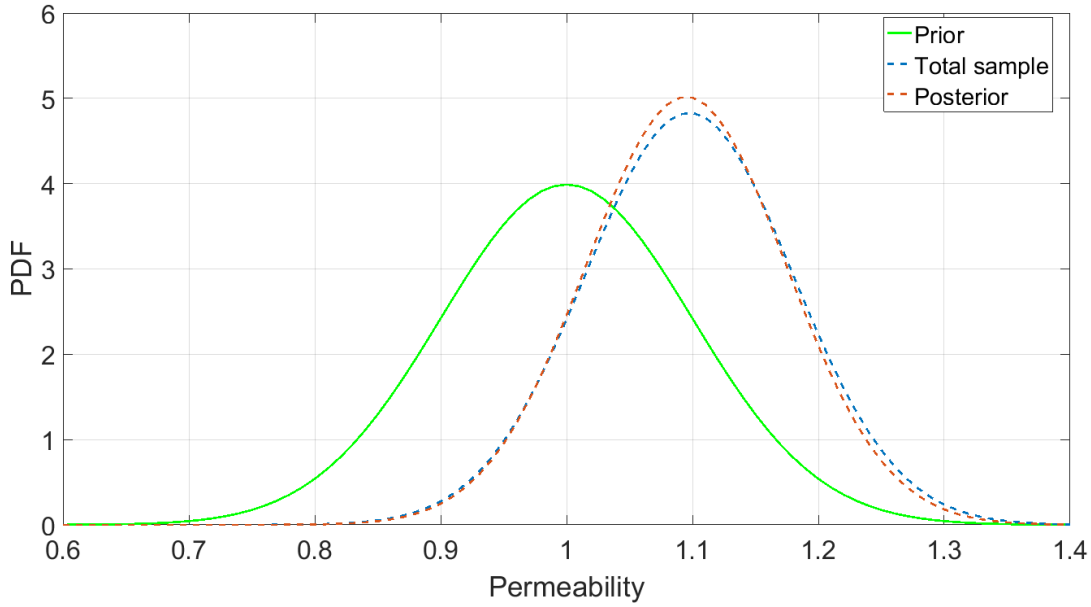

Fig. 11: Comparison of permeability PDF for the sampled, posterior and prior distributions

3.1.2. Maximum Likelihood results

The initial variable state of the Maximum Likelihood method is taken to be the mean. In the beginning, the method advances with larger leaps, which get smaller as the solution surface is approached. An ideal angle between the surface and the line defined by the initial point and the point of convergence would be a right

one, signalling that the algorithm has followed the shortest trail. Also, in this case the path is linear, a feature which can be safely accredited to model simplicity.

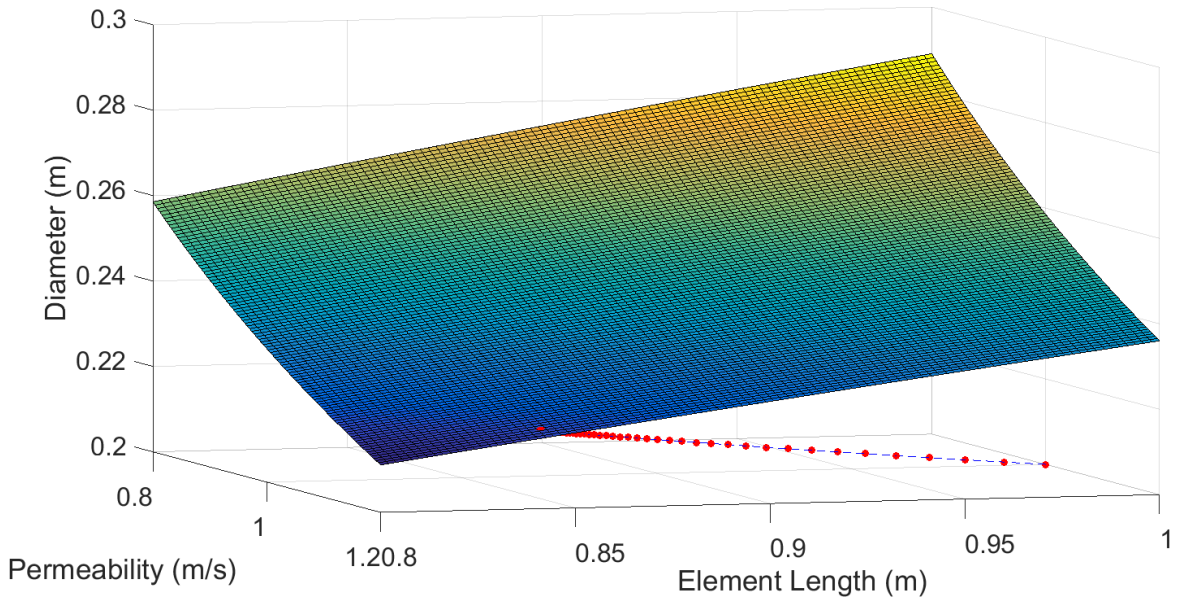

Fig. 12: Convergence of Maximum Likelihood method for the 1D waterflow inverse model in the parameter domain

3.1.3. Genetic Algorithm results

The Genetic Algorithm reproduces the population of each generation, illustrated with a cloud in the 3D space, until a point (for this version) within tolerance is achieved (Fig. 13). Hence, the final cloud is depicted, in relation to the analytical solution. Moreover, the Probability Density Function of each parameter along the generations exhibits the evolution of the variable distribution until convergence (Fig. 14).

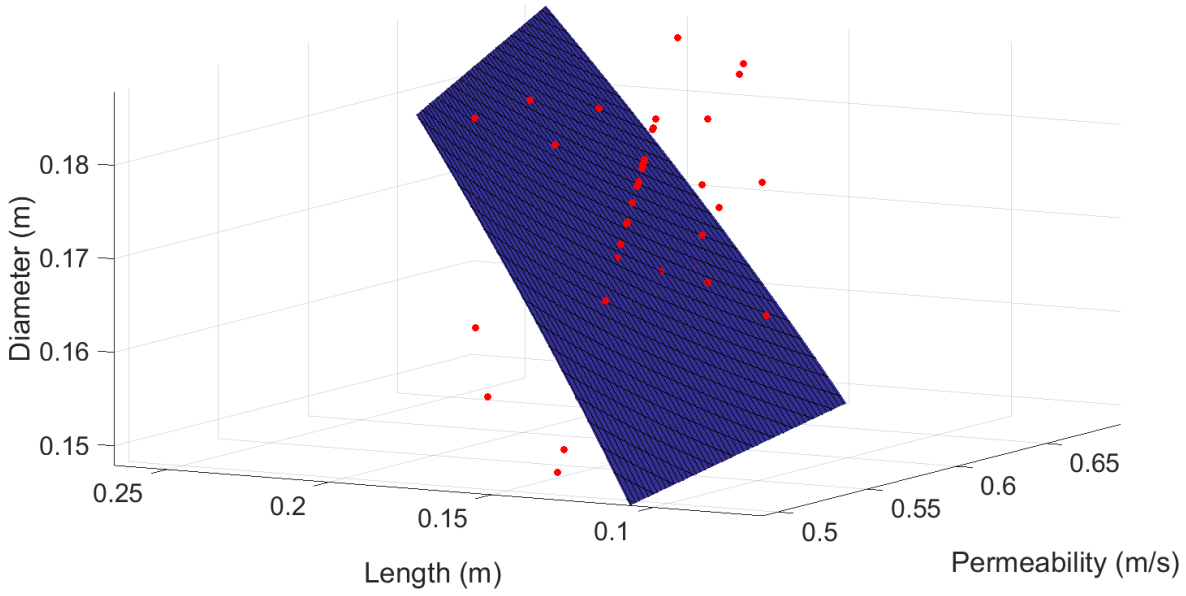

Fig. 13: Cloud of final Generation for the GA in the 1D waterflow inverse problem

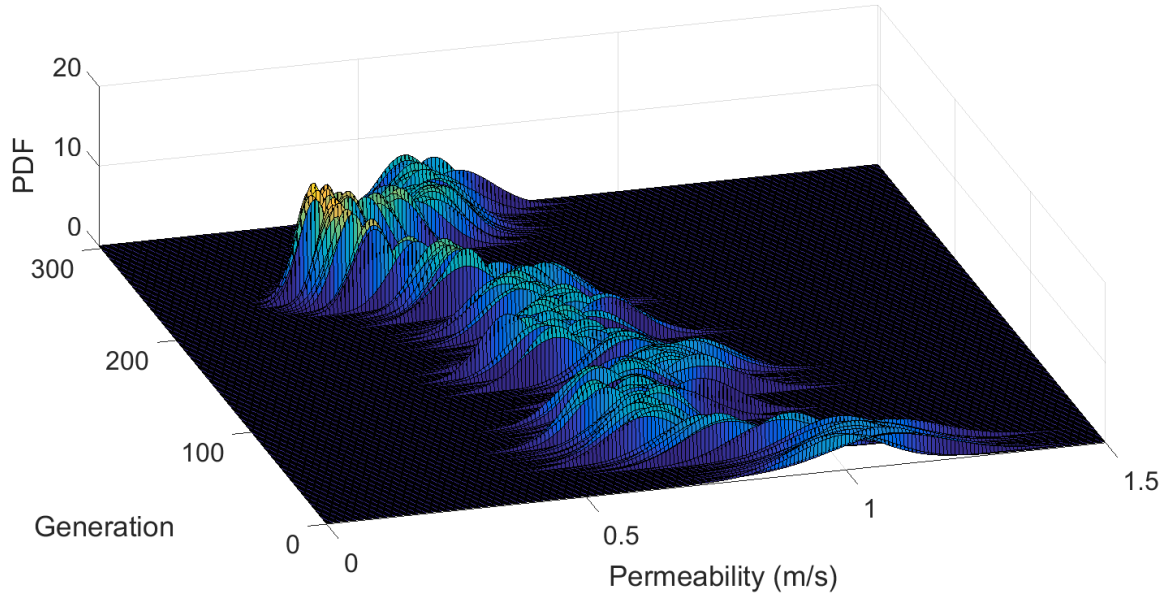

Fig. 14: Evolution of variable PDF to generations (permeability)

3.1.4. Kalman Filter results

This method guides the ensemble of variable sates until its mean produces a tolerable error. Each ensemble can be depicted with a 3D cloud. A significant result of this procedure is the fact that most of the points in the final cloud are also solutions (Fig. 15).

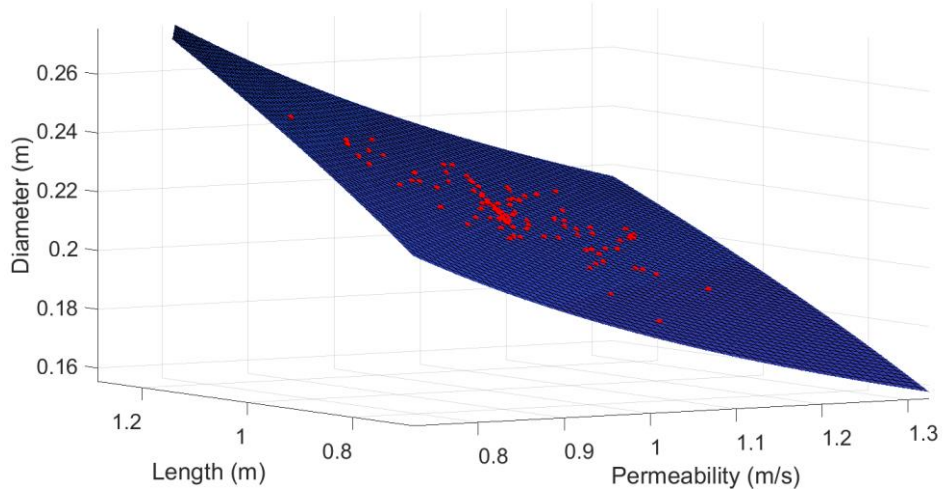

Fig. 15: Final cloud of the Kalman Filter method in the 1D waterflow inverse problem

3.1.5. Particle Filter results

As previously, in the RPF each particle ensemble is represented by a cloud (Fig. 16). Contrary to the previous method, most particles are not a solution while their mean is, implying that the sample has a higher standard deviation. Also, usual practice of the RPF includes plotting the particle values of a variable for every iteration in order to present the evolution of its mean and variance (Fig. 17).

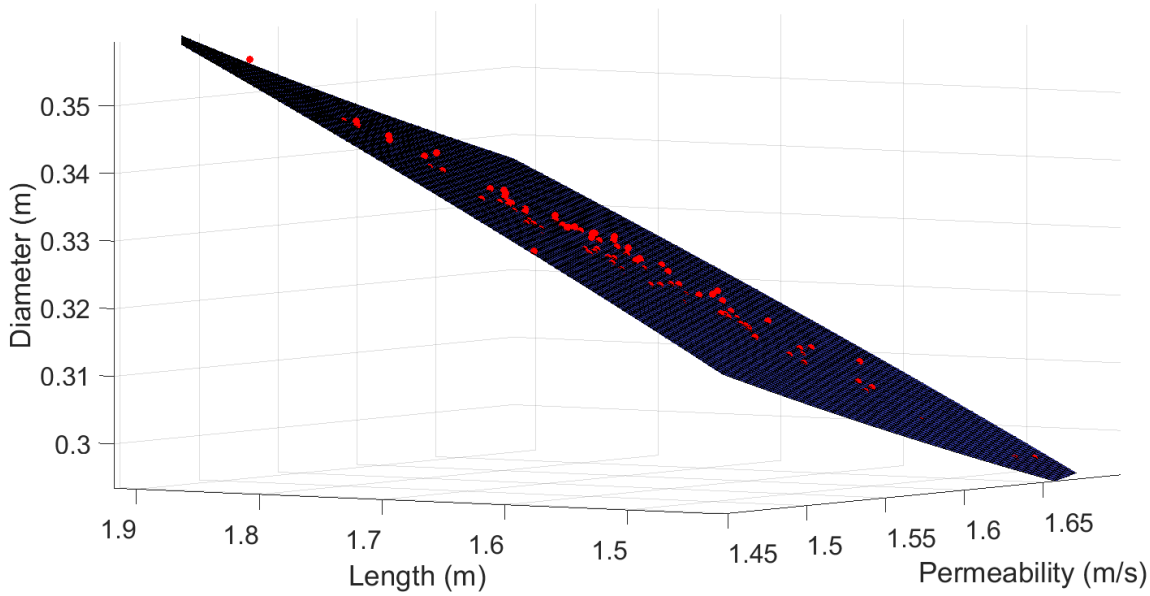

Fig. 16: Final cloud for the Particle Filter in the 1D waterflow inverse problem

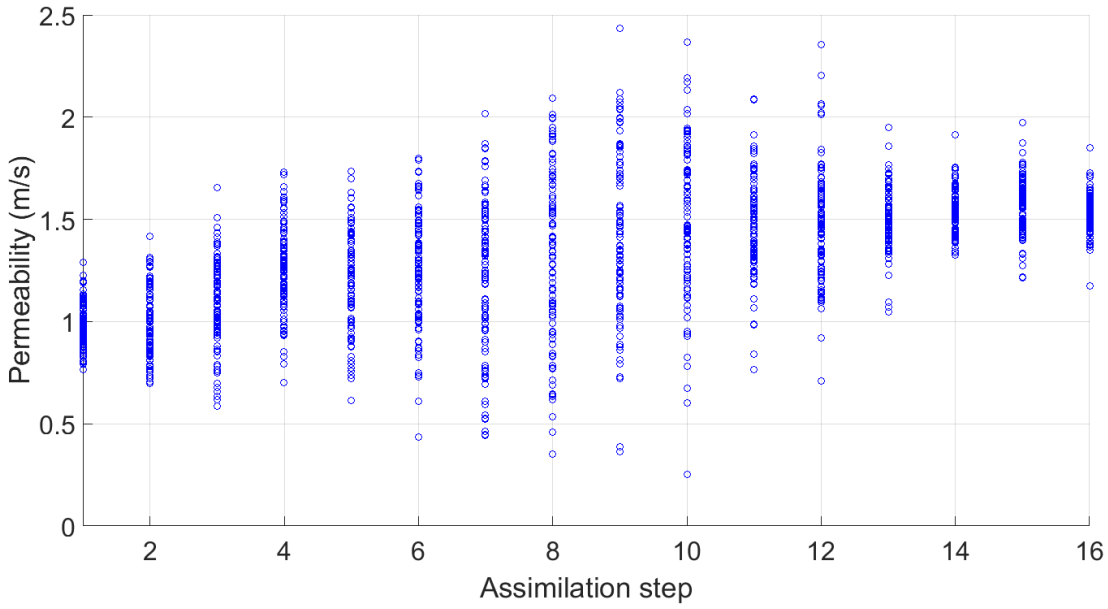

Fig. 17: Particle evolution along steady state iterations

3.2. Conclusions

These statistics are deemed as a basis for attempting to compare method conduct and efficiency. In the end, the Markov Chain Monte Carlo, the Genetic Algorithm and the Regularized Particle Filter are the methods chosen for further application. While having the most FEM iterations, MCMC is a solid foundation for the method development and comparison on the main research of the project. Furthermore, the other two approaches are promising, mostly because of their ability to explore the domain and allocate multiple optima. Besides the particle filter is able to provide a distribution as a solution, which establishes it as an ideal contender for method merging. On the other hand, the Maximum Likelihood and Extended Kalman Filter were

rejected, mainly because of their need of constantly deriving the stiffness matrix. For a large FEM mesh this might prove to be computationally fruitless and might also conflict with the nature of the calculations. For example, geotechnical knowledge verifies that a mesh with yielded points experiences a softer response to loading. However, the undrained shear strength of the soil does not take part in the formation of the stiffness matrix, hence being attributed a derivative of zero, seemingly being treated as irrelevant to the problem. The Ensemble Kalman Filter, on the other hand, did not pose as an attractive alternative as the Regularized Particle Filter.

This preparatory analysis poses as an excellent opportunity for method calibration and evaluation of inner mechanics. Developing the inverse analysis schemes in a simple 1D flow model allows for thorough supervision of method conduct and illustrative result interpretation in a three - variable space. Hence, a solid foundation has been established for expanding the inverse analysis into the main FEM model.

4. Methodology

After introducing the inverse analysis methods, employing them in simple cases and comparing the performances, their application on the main problem has to be resolved. The main RFEM model simulates a slope excavation taking place in four stages. A heterogeneous soil layer is considered, resting on firm bedrock. The excavation is fast enough to assume that the soil behaves in an undrained manner.

Firstly, an appropriate constitutive model has to be selected for the soil. Secondly, a scheme accounting for soil heterogeneity has to be introduced into the model. Then, the error formulation describing the convergence of the inverse analysis is presented, as well as the concepts of selecting the prior and target distributions of the inverse problem. Following, the main model of the boundary value problem is finally set up by compiling the mesh and defining the slope excavation stages. Also, the monitoring points, where the observations are taken, are selected. Lastly, methods of inverse analysis results visualization and processing are described.

4.1. Soil modelling

The FEM analysis adopts undrained soil conditions in the simulation. Thus, soil strength is only characterized by its undrained shear strength (c_u) and the calculations are executed on total stresses, since pore water pressure is unknown. This conditions appears to be realistic in a dike analysis, when waterflow and consolidation are not considered. Moreover, since the embankment material is usually normally consolidated clay, short-term, undrained behavior is critical.

Soil behavior is simulated by the Linearly Elastic - Perfectly Plastic (LEPP) Mohr-Coulomb constitutive model, adopting the simplified Tresca failure criterion. This is because in undrained conditions the friction angle is considered null and only undrained shear strength is utilized, then the Mohr-Coulomb failure surface is constant along the isotropic stress axis, taking the value of c_u for lode angles corresponding to principal axes and having a hexagonal crossection (Schofield & Wroth, 1968).

After examining the effect of the parameters included in the Tresca model on soil response, only undrained shear strength and Young's modulus will be included as variables in the inverse analysis. This is because these properties dictate the generation of displacements and the evolution of the failure mechanism, while parameters as the dilatancy angle, unit weight and Poisson's ratio provide a stiffer inverse problem response, having a less significant impact, too. As a result, they are fixed at: γ=20 kN/m^3, ψ=0° and $\nu \approx 0.50$ for a soft clay in undrained conditions.

4.2. Induction of soil heterogeneity

In the scope of this project, inverse analysis has to deal with soil heterogeneity. Spatial variability of soil parameters is chosen to be simulated with the Random Finite Element Method (RFEM). According to this stochastic approach, random fields of parameter values picked from a prior distribution are compiled, and then these values are matched to points of the FEM mesh. In this way, randomness is included in the model, contrary to the deterministic concept(Fenton & Griffiths, 2007).

The basis of creating a random field for a parameter is the transformation of a standard normal random field. In order to produce a more variable field, requiring though the same computational power, each random cell is matched to an integration point of the mesh (N in number).

The synthesis of the random fields is achieved through the superposition of components taken from the analysis of heterogeneity patterns (Lloret-Cabot, Fenton, & Hicks, 2014). The scale of fluctuation (ϑ) is the distance of reappearance of similar optima for a parameter in a given direction (Vanmarcke, 2010). Based on this norm, the covariance matrix [C_{NxN}] of the random field, expressing the relationship between the standard normal values of every cell, can now act as the chart of field interdependency. Considering a continuous, underlying random field, each integration point is assigned the average of the cell. For that, the spatial correlation function of the field is integrated over the cell (Eq. 4.1 - 4.3), therefore reducing the variability of the field. Subsequently, the covariance matrix of the field is composed of entries, each one connecting the values of two cells.

$$\tau = \sqrt{(\frac{\Delta x}{\vartheta_h})^2 + (\frac{\Delta y}{\vartheta_v})^2} \tag{4.1}$$

$$\rho = e^{-2\tau} \tag{4.2}$$

$$C_{A,B} = \frac{1}{V_A V_B} \int_{\Omega_A} \int_{\Omega_B} \rho(\mathrm{x}_B - x_A)\,\mathrm{dV}\mathrm{dV} \tag{4.3}$$

$$C = \Phi \Lambda \Phi^T \tag{4.4}$$

$$V = \Phi \Lambda^{1/2} \tag{4.5}$$

$$Field = \mu_{\mathrm{Prior}} + \sigma_{\mathrm{Prior}} V \xi \tag{4.6}$$

(Where ξ is a random standard draw)

The method applied for compiling the random field is centered around the eigen-decomposition of the covariance matrix (Eq. 4.4), allowing for the calculation of its eigenvectors [Φ_{NxN}] and eigenvalues [Λ_{NxN}] (van den Eijnden & Hicks, 2017). Every produced eigenvector expresses a form of parameter spatial variation, attributing a value to every integration point, and by being scaled to its eigenvalue (Eq. 4.5), the proportion of its participation to the final field is assessed. Then, a vector of standard normal random values is picked, with each value coinciding to an eigenvector of the covariance matrix. Eventually, the random field is created by the multiplication of the ensemble of scaled eigenvectors and the standard random field, which is interpreted as the superposition of N random fields, each one formed by the corresponding scaled eigenvector and a random value(Bracewell & Bracewell, 1986). Also, the field mean and standard deviation have to be taken into account (Eq. 4.6). These random values now act as the variables of the inverse analysis.

In a structured mesh, the eigenvector plots (Fig. 18a and Fig. 19a) exhibit a smooth transition, as field cells are adjacent in both the mesh and the plot axes. On the other hand, the respective plots in unstructured meshes tend to seem quite random (Fig. 18b Fig. 19b). Besides, plotting an eigenvector in the mesh can provide insight on the dynamics of cell interdependency for a variable (Fig. 20). The first eigenvector has concentrated zones of similar values, while later eigenvectors have high and low values seemingly randomly scattered.

Although this approach provides an efficient attempt at modeling soil heterogeneity, the computational strain is relatively high, considering that the inverse analysis would have to deal with a number of variables equal to the entries of the random vector (N). In order to reduce the cost of the calculation, the eigenvectors equivalent to 95% of the total energy are employed. While this device effectively drops the size of the variable state (random vector), only a minor impact on the accuracy of the simulation is inflicted. Besides, having assumed uncorrelated soil properties, every random field reproduction demands a unique variable state, meaning that the variables increase linearly to the number of simulated parameters.

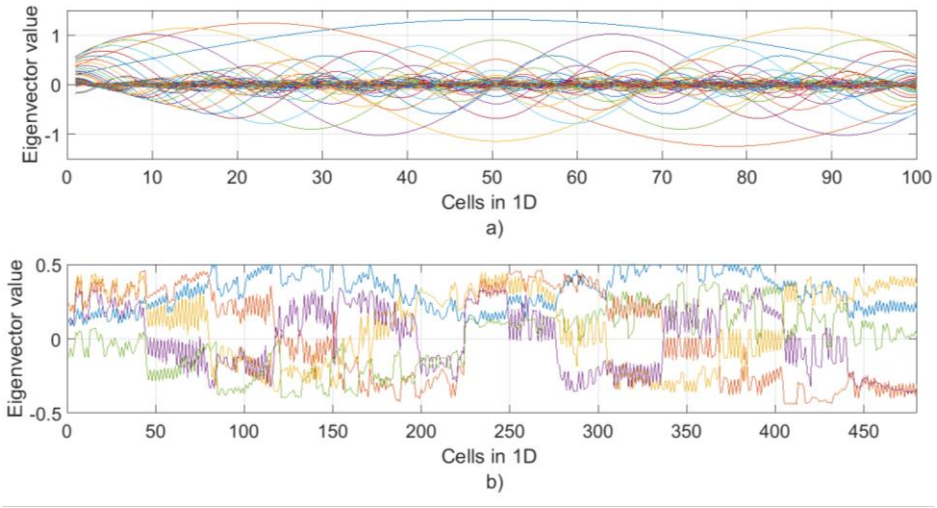

Fig. 18: Comparison between the scaled eigenvectors [*V*] plots over the field cells for: a) an 1D structured mesh, b) a 2D unstructured mesh

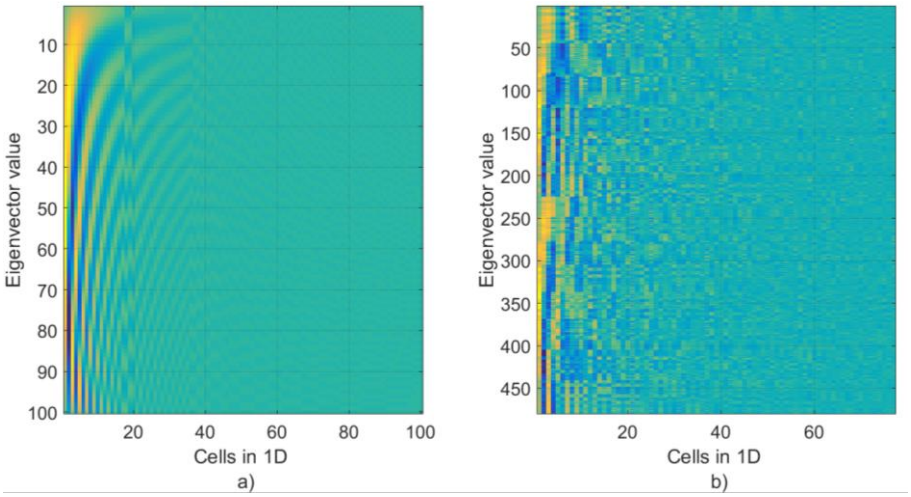

Fig. 19: Comparison between the eigenvectors image plots for: a) an 1D structured mesh, b) a 2D unstructured mesh

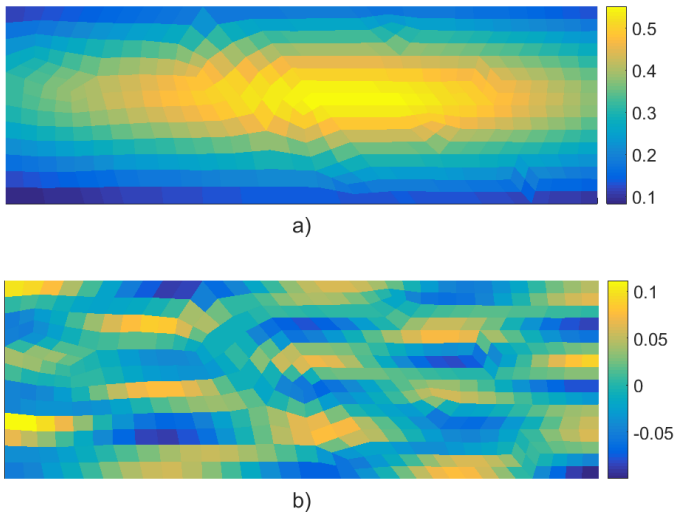

Fig. 20: Example comparison of eigenvector plot in mesh for: a) eigenvector no.1, b) eigenvector no.52

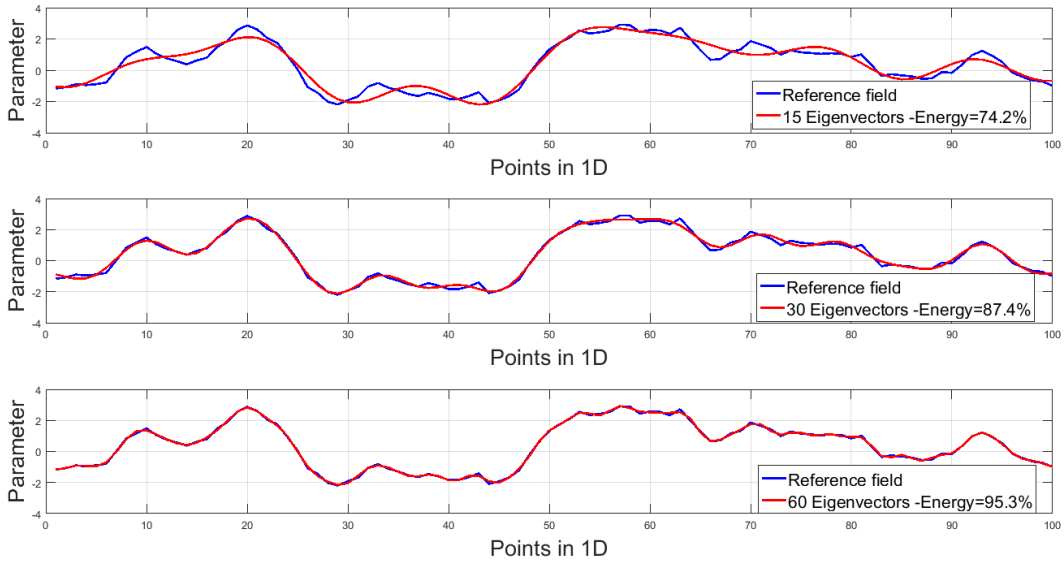

Fig. 21: Comparison of fit between reference and simulated field in an 1D mesh for several eigenvector energy levels

Moreover, a critical point of the analyses lies within the compilation of the reference field. In each analysis, a field has to be used to produce the displacements acting as measurements. In order to set a goal reflecting reality for the back-calculation, the reference field is to be assembled using a detailed random field, or in other words, a field generated by the full eigenvector ensemble. Therefore, the inverse analysis is attempting to describe a decent approximation of the exact field, by utilizing a less accurate, yet computationally efficient, collection of eigenvectors.

Following the aforementioned concept for including heterogeneity, the specifics of this mesh can be exhibited. In the initial analysis, fixed values of the scales of fluctuation are employed, as the main goal here is to assess the ability of the methods to approach the posterior and evaluate their convergence. Thus, a pre-specified lateral scale of fluctuation of ϑ_v*=1 m* and a horizontal one of ϑ_h*=10 m*, realistic values for a clay deposit, are utilized for calculating the covariance matrix of the random field. At this point, it should be noted that in the main analyses, the scales of fluctuation are treated as constants.

4.3. Error formulation

In a single execution, the FEM program provides as an output a displacement vector, which includes the displacements of all monitoring points at every construction stage. Notably, the vector is normalized to the displacements of the initial stage, which plainly consists of gravity loading. As a result, a generated random field is given as input in the FEM, simulated in all stages and eventually, the calculated displacement vector is compared to the observation vector as a standard of a successful run. Then, the displacement error is evaluated as the first norm of the difference between the displacement and measurement vectors, normalized by the greatest measurement and the length of the vectors (Eq. 4.7).

Adopting the pseudo time concept, two possible criterion schemes for field validation ensue. The first one involves evaluating the error of all stages in total and requiring it to be within the user-defined tolerance. This approach is the simplest, allowing for a relatively fast convergence of the inverse analysis. However, it suffers a major setback; the error expresses an average between all stages, implying that some them could be above the tolerable margin.

$$Error = \frac{\sqrt{\sum_{i=1}^{n}(d_i - Obs_i)^2}}{\max(|Obs|) * n} \tag{4.7}$$

(Where n is the number of measurement points)

Finally, the reasoning for selecting the displacement error tolerance lies within the comparison of a random heterogeneous and a homogenous fields. Firstly, a homogenous field with a high strength value, so that no yielding occurs, is set as reference, and then a field with a heterogeneous strength and stiffness is generated, delivering a displacement error is in the order of 0,3%. Therefore, the tolerance of the analyses should be set an order of magnitude less (circa 0,03%). This investigation primarily aimed into pinpointing the errors that the analysis could typically generate, and thus confirm that the defined tolerance is effective.

4.4. Target distribution schemes

Fundamentally, the goal of the inverse analysis is for the posterior of variable states sampled from the prior distribution to approach the target one, used to create the reference field. At this point, two analyses schemes can be established:

- The first one is applying the inverse analysis in a case study mimicking a real-life project. Assuming that soil testing and site investigation is implemented on the embankment, then its results can provide the prior distributions of the soil parameters. Moreover, the distribution of the reference field, the target, will have to comply to the results, as it has to reflect the in-situ conditions. In this predicament, the prior matches the target (and posterior), hence the analysis is faster, but also degenerated to merely identifying soil heterogeneity.
- The second approach follows a different philosophy, being based on the notion that the posterior is different than the prior. In this case, no data is available about the soil, on the contrary a guess is made on the distributions of its parameters. Thus, different prior and target (and posterior) distributions can be supposed, allowing for the algorithm to not only allocate the heterogeneity of the mesh but also approximate the distribution used to generate the reference field.

While the first scenario seems more relevant for a realistic application, the second scheme will be utilized in order to evaluate the competency of the method. Starting from a prior distribution, the main goal is the approximation of the effective, in-situ distribution of the parameters.

4.5. Mesh selection

The Finite Element Method is employed to solve a mechanical boundary value problem. The model is a 2D simulation, assuming plane strain conditions. This simplification is realistic in the case of a long slope, whose nodes move in a similar manner[1]. Model behavior is dictated by the two dimensional force equilibrium governing equation (Eq. 4.8). After selecting an appropriate mesh, the relevant boundary conditions can be defined so that the problem acquires a unique solution.

$$\begin{aligned} \frac{d\sigma_x}{dx} + \frac{d\tau_{xy}}{dy} + F_x = 0 \\ \frac{d\sigma_y}{dy} + \frac{d\tau_{yx}}{dx} + F_y = 0 \end{aligned} \tag{4.8}$$

[1] In this way the monitored displacements are relevant to the entire slope, so the plane strain assumption is acceptable.

The mesh of the FEM analyses will represent the staged excavation of a 5m high slope at an inclination of 51°. Due to symmetricity, the simulation is allowed to focus on only half of the excavation site. After introducing the requirements of mesh generation, the mesh utilized in the following analyses is compiled.

Firstly, the mesh is consisted of 8-noded elements with 4 integration points. Employing 8 nodes, alongside the utilization of second order (quadratic) shape functions, greatly enhances the accuracy of the mesh, with the benefit outbalancing the computational strain (Smith, Griffiths, & Margetts, 2013). Each node has two degrees of freedom: lateral and vertical displacement, with their load counterparts Also, the boundary conditions at the left and right are horizontal fixations, while the bottom nodes are fixed in both directions (Fig. 23). When soil is excavated, nodes that are left exposed have a zero loading and are able to move freely. Thus, the boundary value problem is now adequately defined.

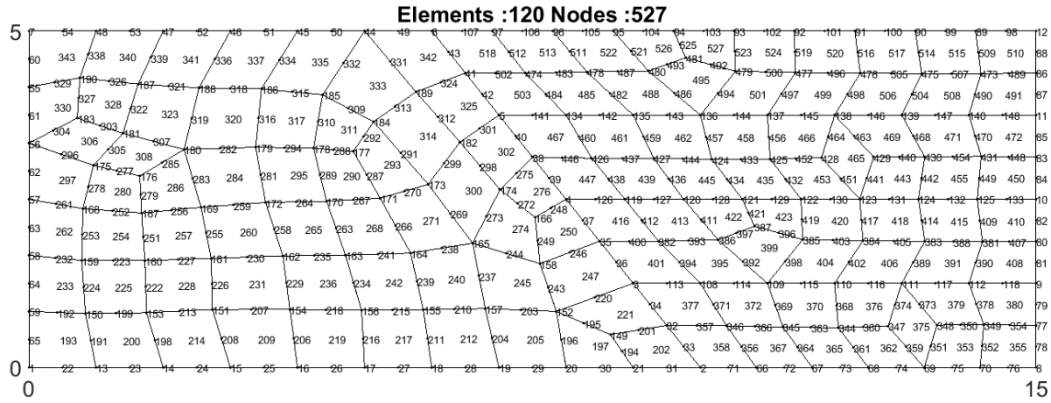

Fig. 22: Mesh of main RFEM model – Nodes

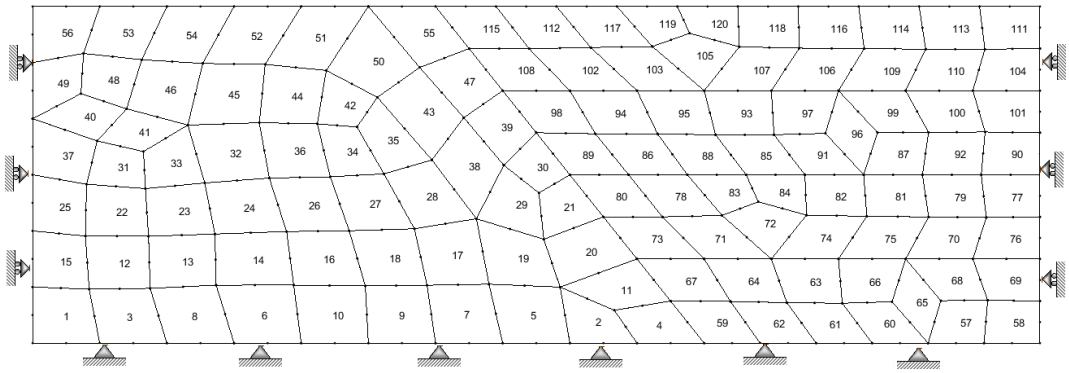

Fig. 23: Mesh of main RFEM model – Elements and boundary conditions

The analysis takes place in 4 excavation stages of equal depth progress, plus a starting one for initializing the stresses with gravity loading. The depth reached in each stage can be seen in the following table (Table 3):

Table 3: Excavation plan (depth reached after each excavation stage)

EXCAVATION STAGE	DEPTH (m)
1	-1,25
2	-2,50
3	-3,75
4	-5,00

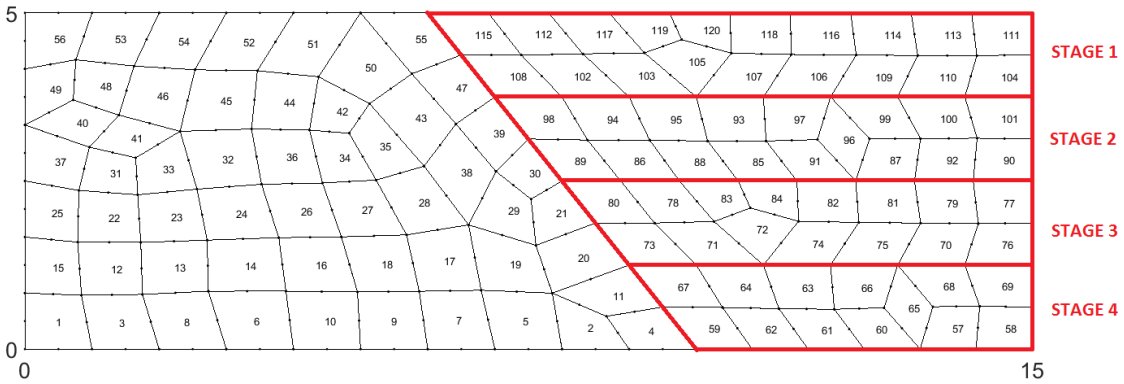

Fig. 24: Excavation plan for the main RFEM model

In order to set the characteristics of the FEM simulation, a sensitivity analysis is run on them. The FEM displacement tolerance, maximum number of iterations and maximum displacement allowed by the model are treated as variables. Then, each variable is assigned an initial value and displacements of the reference simulation at selected measurement points are calculated. In each iteration of the sensitivity analysis, the error of displacements at the measurement points guarantees the independency of the inverse analysis, as well as provides insight about the minimum intrinsic error of the method. Moreover, the soil of the simulation is homogenous and the analysis is executed for two strength values (20 kPa and 13 kPa), in order to ensure model behavior in domains with both light and severe yielding. Therefore, the FEM model is ran with a 10^{-6} tolerance, a maximum of 3000 plastic iterations and a maximum displacement of 1m, values that seem potent to allow plasticity to fully develop (Fig. 27 and Fig. 28). Besides, with these values the minimum intrinsic observation error is achieved, while setting a computationally affordable analysis.

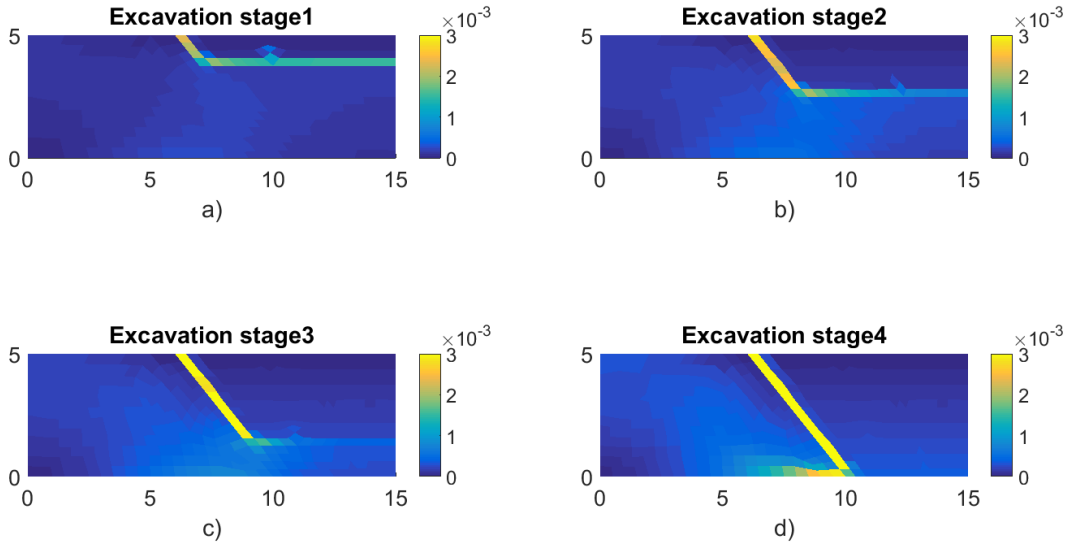

Fig. 25: Evolution of deviatoric strains - Homogenous strength field of 20 kPa

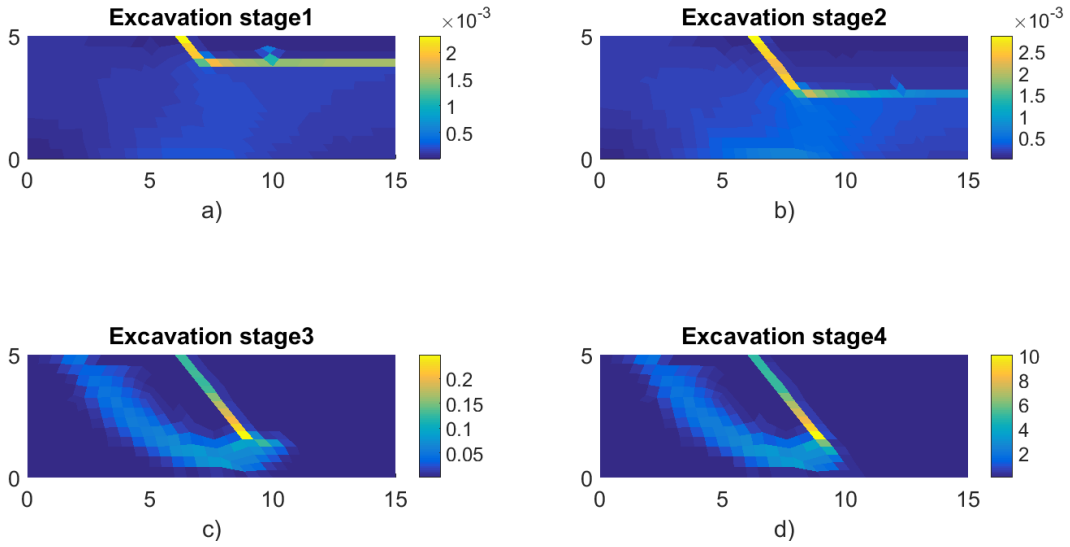

Fig. 26: Evolution of deviatoric strains - Homogeneous strength field of 13 kPa

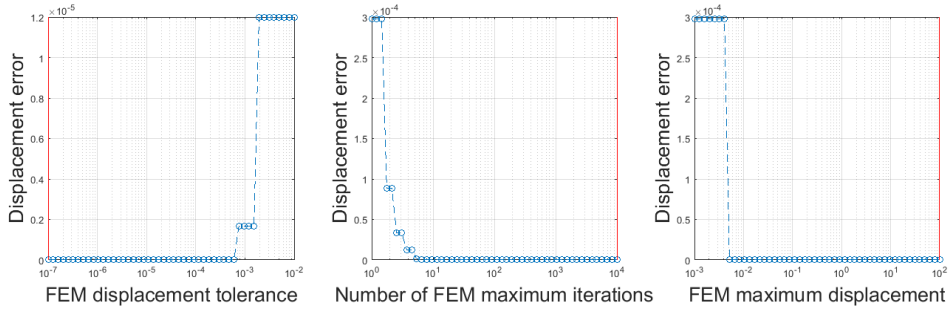

Fig. 27: Sensitivity analysis of maximum tolerance, maximum number of plastic iterations and maximum displacement implement in the RFEM model - Homogenous strength field of 20 kPa

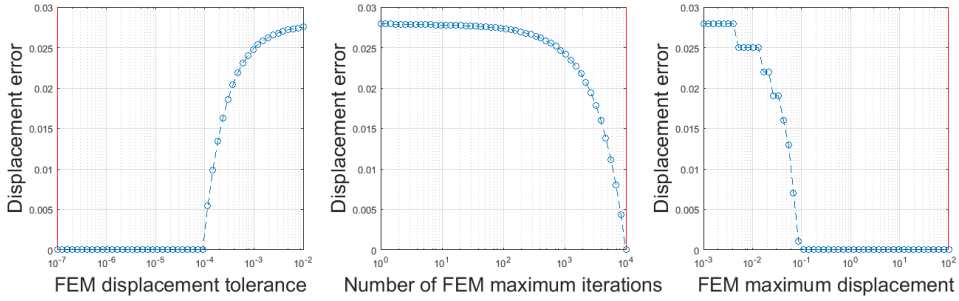

Fig. 28: Sensitivity analysis of maximum tolerance, maximum number of plastic iterations and maximum displacement implement in the RFEM model - Homogenous strength field of 13 kPa

Ultimately, a significant criterion is adopted for evaluating the results of each analysis. The efficiency of each application is judged by its ability to generate a posterior distribution matching the target one. Thus, for each realization the mean and coefficient of variation over the field is assessed. Taking into consideration that in the FEM calculation some integration points have a reduced or non-existent effect on the outcome, then the inverse analysis may refrain from optimizing them. Then, if the mean of the whole field is considered for the calculation of the posterior, a divergence from the target distribution is expected because of this condition. As a result, when presenting the data taken from each analysis, the posterior of the entire field, as well as that of selected integration points will be evaluated. In order to achieve this, two possible areas of high importance are identified:

- The first area is close to the slope where the largest portion of measurement points is concentrated (Fig. 29). It is suitable for a case of elastic-like soil response.
- The second area is identified in the plot of deviatoric strains, which allows for the recognition of the developing failure mechanism (Fig. 30). The points of interest are usually at the toe of the formed shear band.

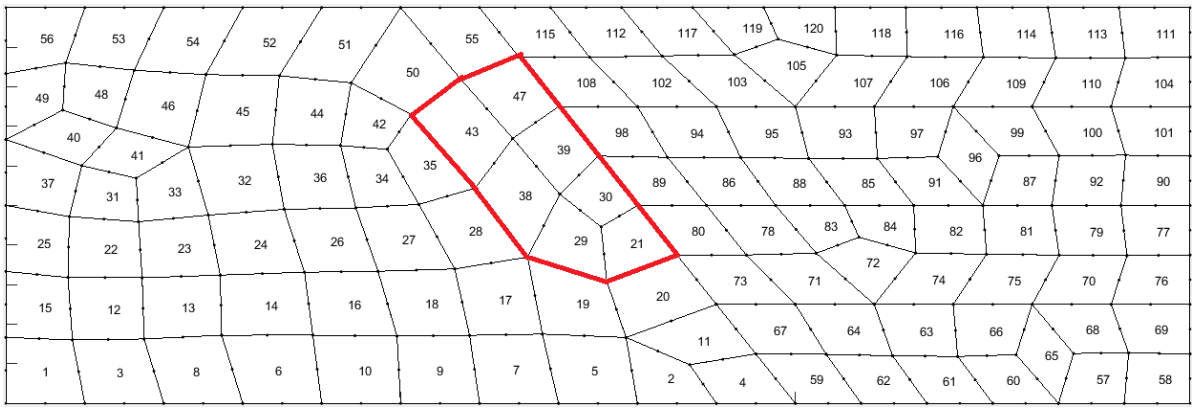

Fig. 29: Selected elements of interest for strong soil response

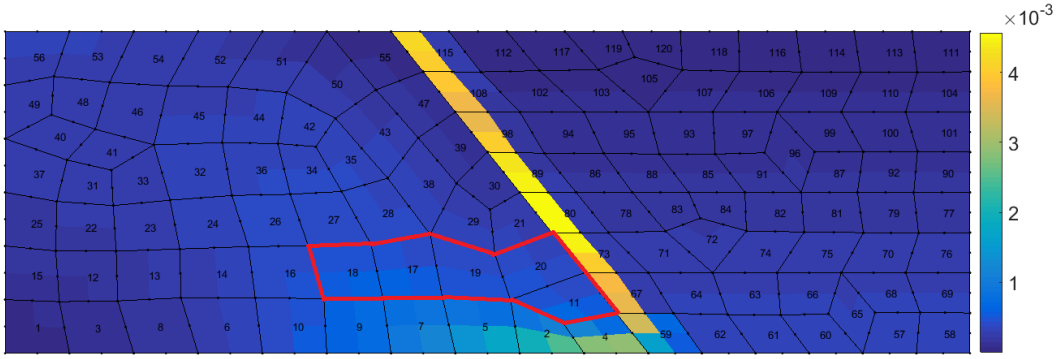

Fig. 30: Selected elements of interest for weak soil response

Lastly, the soil heterogeneity scheme presented in paragraph 4.2 is applied on the constructed mesh. There are 120 elements, so 480 Gauss points, equal to the number of field cells simulated. After the compilation and eigen-decomposition of the heterogeneity covariance matrix, a vector [*V*] of scaled eigenvectors with 480x480 entries is calculated. Attempting to reduce the number of inverse analysis variables, it is estimated that only the first 77 eigenvectors have to be applied for achieving 95% energy of the total matrix. Hence, the variables were dropped to only 16% of their initial value.

4.6. Monitoring scheme

Following, the measurement points of the mesh should be defined. Conceptually, these are the points whose displacements will act as the goal of the inverse analysis, and compared to a real situation, the points where the monitoring equipment operates.

Firstly, it is chosen that the horizontal displacements will be applied in the inverse analysis, as this is the type of measurement produced by the usual monitoring tool, the inclinometer. Moreover, horizontal displacements are greater when a failure mechanism starts to develop, attributing to it a more significant

impact on the analysis. Secondly, the measurement points are be positioned in columns, a formation mimicking an inclinometer.

Executing the FEM code for the strong soil field described in the previous paragraph, the plot of displacements is interpreted (Fig. 31). The greatest displacements are gathered near the slope, and so this should be an area of observation. Measurement points should not be located at nodes of excavated soil volume, so that they produce displacements at all simulation stages. Besides, displacements at the area between the left boundary and the center of mesh are still considerable and so they should also have an effect on the inverse analysis. Moreover, including measurements covering different mesh regions may allow the analysis to chart soil heterogeneity in a more efficient manner. Lastly, it should be noted that the number of measurements near the slope should always be a significant portion of the total observation vector, so that the area of greatest interest is sufficiently represented in the analysis. Following the above, the monitoring plan of 20 points is compiled (Fig. 32).

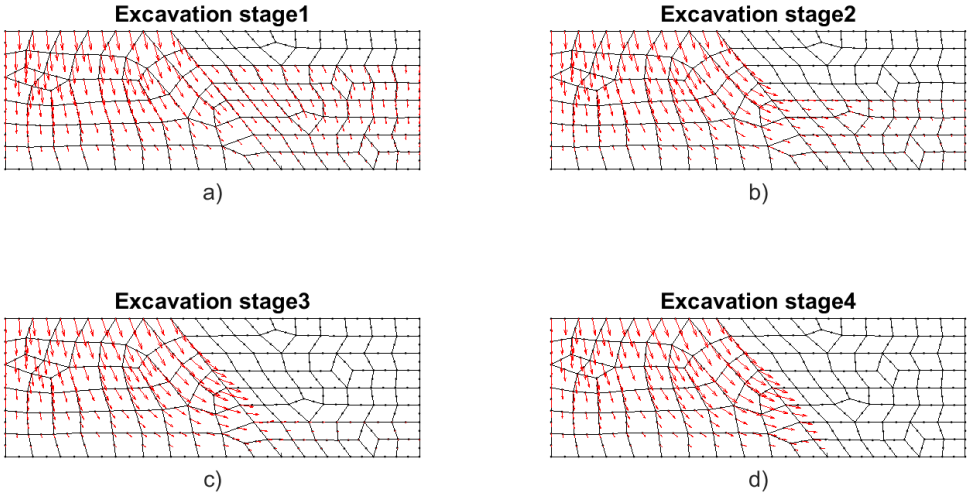

Fig. 31: Evolution of displacements in the main RFEM simulation

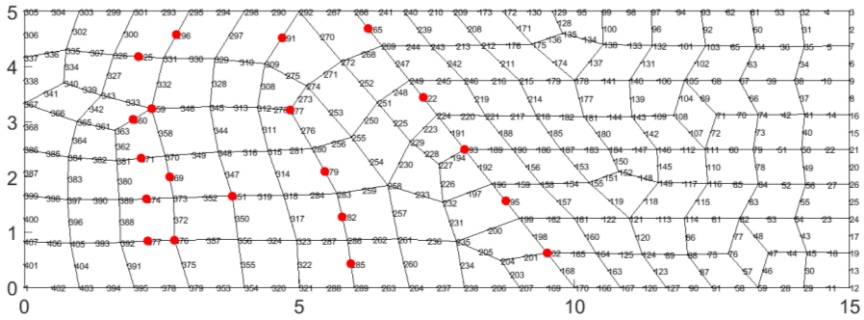

Fig. 32: Final selection of monitoring points in mesh for the main model

4.7. Result visualization

The output of every inverse analysis has to be objectively evaluated on some critical points. First and foremost, the approximation of the target distribution is the main objective of the investigation. Also, a major goal of the analysis is mapping parameter variability in mesh. Several techniques are employed in order to illustrate and interpret method behavior and efficiency.

4.7.1. Posterior distribution plots

In order to visualize method competence, the posterior distributions of soil parameters and inverse analysis variables are plotted. The first part includes drawing the mean and coefficient of variation of the field for every iteration and parameter. Then, the path of the parameter can be compared to the reference and method convergence and success can be determined. The second part illustrates the probability density function of an inverse analysis variable, using its values in every iteration as a sample. For a successful approximation, the distribution should be thin and concentrated around the reference value, ideally taking the form of a "spike". Also, it should be considered that with a fixed scale of fluctuation, matching variables would lead to a perfect field fit.

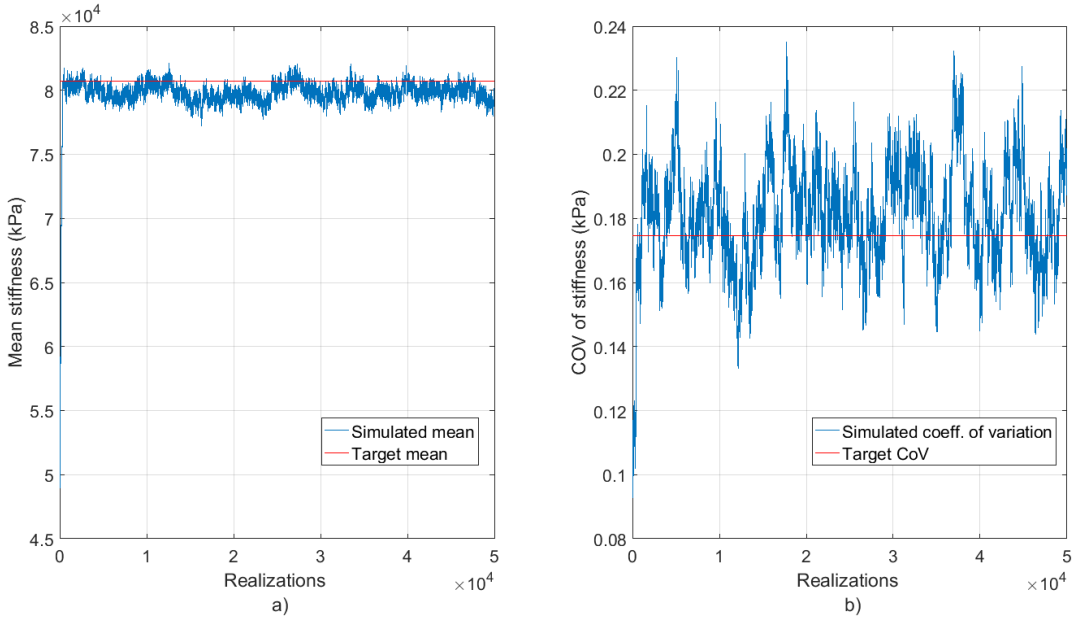

Fig. 33: Example of plot for mean and CoV of field along realizations

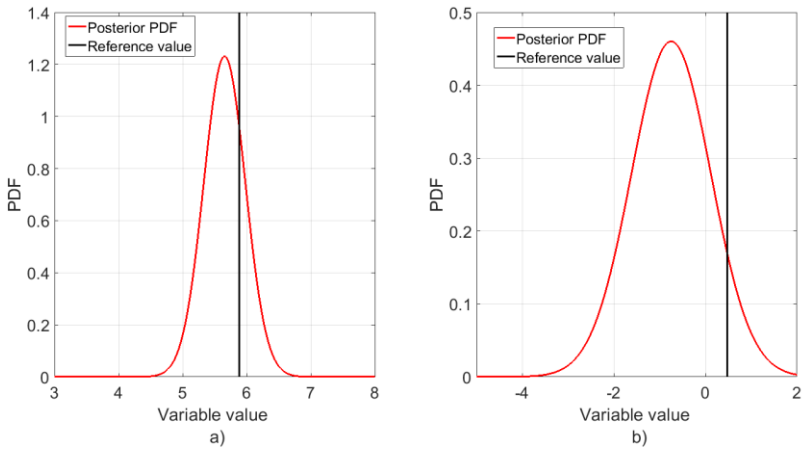

Fig. 34: Example of variable posterior PDF plot with reference (target) values: a) variable no.77, b) variable no.82

4.7.2. Principal Component Analysis

The Principal Component Analysis (PCA) is a method of re-expressing a data set based on patterns between them. After collecting a sample for n variables, its covariance matrix is analyzed to its eigenvectors, which are

scaled to the corresponding eigenvalues. Then, the eigenvectors define axes and a new coordinate system for the data. Besides, the scaled eigenvectors form an ellipse that includes the data for a given level of confidence. The first eigenvector (corresponding to the highest eigenvalue) is the principal component and the first axis of the ellipse (Smith, 2002).

The PCA is utilized on this matter for representing local convergence of soil strength and stiffness. Supposing two variables: the reference field and the mean of Gauss points for a parameter, then a sample can be compiled. Ideally, the scatter of the sample would be exactly on the first bisector of the original axes system, meaning that the mean at every point matches the reference value. Applying the PCA generates the ellipse that includes 68% of the data and allows for sample examination. A perfect field fit would require the center of the ellipse to be on the bisector, the first axis to collide with it and the second axis to be minimized, meaning that the ellipse would degenerate to the line. Besides, the ellipses for several confidence levels can be illustrated, allowing a superior visualization of sample statistics (Levasseur, Malecot, Boulon, & Flavigny, 2010).

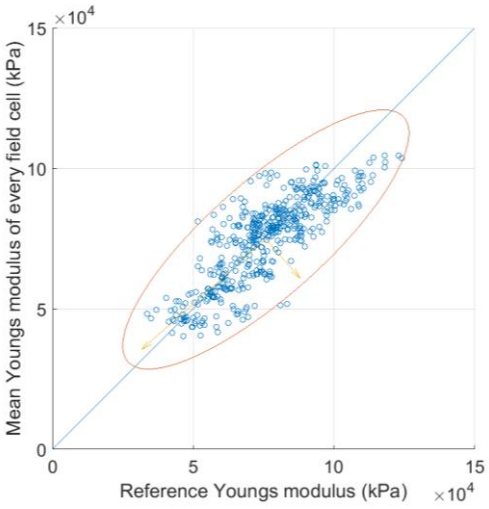

Fig. 35: Example of PCA ellipsis for 68% confidence level.

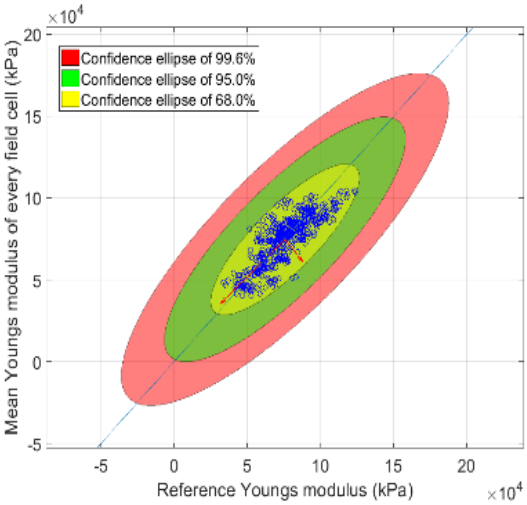

Fig. 36: Example of PCA ellipsis for 68%,95% and 99.6%confidence levels.

4.7.3. Field plots

The simplest way to visualize field variability is plotting the values of each property at integration points. This can be achieved by either drawing the contour of properties in the mesh, or by presenting the same data in chart form. For these plots, the reference field of each property is compared to the mean of each Gauss point for all successful iterations. Also, in the chart, the lines of ± the standard deviation are added, so that field approximation can be judged based on sample spread. Typically, the Z value of 2 is applied, so that the band formed represents 95% of the posterior distribution values. The field error is simply defined as the norm of the difference between mean and reference field normalized by the standard deviation of each point and the number of points included (Eq. 4.9). The meaning of the confidence interval is that when sampling 100 fields from the produced distributions, 95 of them are going to have a field error equal to or lower than the one of the mean and the limits of the band. Establishing this, when the reference is within the band, the local field error is not higher than the marginal value defined. The visualization techniques can be applied for all or selected integration points of interest.

$$Field\ error = \frac{\sqrt{\sum_{i=1}^{n}(\frac{\mu_i - Ref_i}{\sigma_i})^2}}{n}, where\ n\ is\ the\ number\ of\ examined\ points \quad (4.9)$$

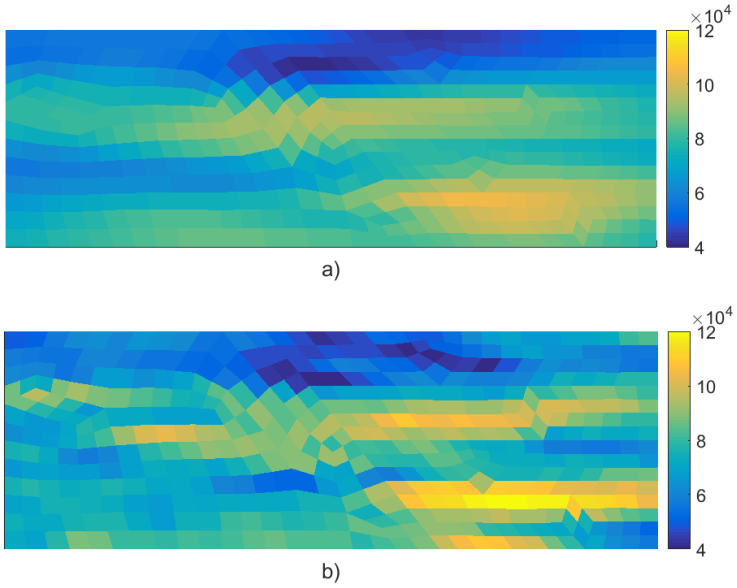

Fig. 37: Example of simulated field (a) and reference field (b) comparison plotted in mesh

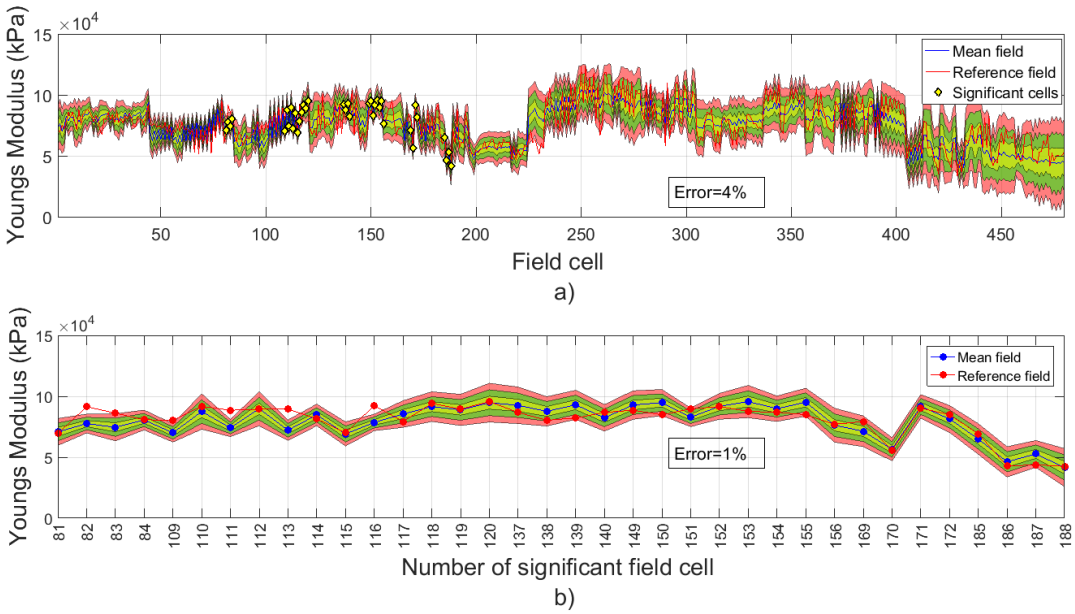

Fig. 38: Example of simulated field with 68%, 95% and 99.6% confidence intervals to reference field comparison plotted over : a) all cells, b)cells of interest

4.8. Result post-processing

A major goal of the inverse analysis, except from recognizing the existing soil properties, lies in providing the media in order to reproduce several successful random fields. As a result of the analysis a sample of variable states is compiled *[S]*. All entries of *[S]* are successful, in the sense that after the required transformation, they produce an error equal to or less than the tolerance applied in the inverse analysis. Regarding the variable posteriors of the inverse analysis, the variable state is described by a multivariate posterior distribution, a notion that takes into account state member interdependency, and not by multiple univariate ones corresponding to each variable. Considering these features, then a random sample picked from a random multivariate distribution, based on *[S]*, is expected to be within tolerance (usually with a probability of at least 50%). Besides, steps can be taken into bypassing the complex multivariate sampling and allowing for a faster and simpler field generation.

As previously explained, in order to generally create a field, a random draw [ξ] must be multiplied with the ensemble *[V]* of eigenvectors resulting from the eigen-decomposition of the field heterogeneity matrix. Then, the prior distribution parameters are applied, producing the field. However, in order to sample random fields taken from the inverse analysis results, the random univariate sample should be equivalent to one taken from a multivariate distribution, so that the variable interdependency of the successful sample is represented. So, the covariance matrix *[C]* of *[S]* is calculated (Eq. 4.10), defining variance and relationships between parameters of the sample. The method proceeds under a significant assumption; variable relations are linear, and so *[C]* is able to capture them. Then, the eigen-decomposition of the sample covariance (Eq. 4.11) provides the components needed for the transformation of the random univariate standard draw in a sample equivalent to a multivariate pick (Eq. 4.12).

$$C = \Phi_C \Lambda_C \Phi_C^T \tag{4.10}$$

$$A = \Phi_C \Lambda_C^{1/2} \tag{4.11}$$

$$f = V\mu + VA\xi \tag{4.12}$$

(Where μ is the mean of the sample for each variable and ξ is A univariate standard sample)

$$Field = \mu_{\text{Prior}} + \sigma_{\text{Prior}} f \tag{4.13}$$

Thus, a matrix [A] is compiled with a size of *nxm*, where n is the number of field cells and m the size of the variable state. It is the key to combining variable interdependency to the existing transformations for soil variability, finally managing to produce a field with a displacement error close to the tolerance. Lastly, this matrix is now unique to the model, based on both its heterogeneity and measured displacements.

Apart from recognizing and charting soil heterogeneity, the benefits of applying an inverse analysis expand due to this feature. Generating multiple fields with displacements relatively close to in-situ measurements enables quicker and easier result interpretation and enhances further applications of the outcome. Firstly, the confidence curve of the measurement approximation can be calculated through a Monte Carlo analysis, allowing for the selection of a measurement spread satisfying a required reliability level. Secondly, a Strength Reduction Factor (SRF) analysis connected to monitoring observations can be performed. A Monte Carlo analysis including the compilation of multiple fields, as presented, and then the application of the SRF method, is potent to create an SRF reliability curve strongly associated to in-situ observations and also reduce the vast number of realizations required. Thus, proper conduct of the inverse analysis significantly enhances geotechnical applications of the examined slope.

5. Inverse Analysis Applications

5.1. Markov Chain Monte Carlo

The first inverse analysis method utilized is the Markov Chain Monte Carlo (MCMC). The FEM model applied is exactly as described in chapter 4. The embankment excavation takes place in 4 stages and soil behavior is constituted in undrained conditions based on the Tresca failure surface. The vertical scale of fluctuation is set at 1m and the horizontal at 10m. Also, the points of significant influence on the results have been defined. The inverse analysis operates with a 0,2‰ and 0,3‰ observation error tolerance, which was the lowest possible to achieve.

5.1.1. Method variations

Regarding the application of the MCMC, three feasible variations exist for the Metropolis-Hastings algorithm:

- The Original, where a variable might adopt the proposal state or preserve the former one. The proposal is generated always with the same standard deviation – proposal reach (*SD*) (Eq. 5.1).
- The Univariate Adaptive, in which the proposal is produced with the standard deviation taken from previous successful samples of the procedure. Variables are independent as the proposal is taken from multiple univariate proposal distributions.
- The Multivariate Adaptive, which resembles the previous case, with the exception that now a proposal state is taken from a single multivariate distribution. For that, the covariance matrix of previous samples is exploited [*C*]. Until the sample is large enough, an initial covariance matrix [C_o] is utilized (Eq. 5.2 and 5.3).

The proposal reach of the Original Metropolis-Hastings as well as the initial value used in the Adaptive version is taken as (*SD*). When the sample is potent to produce a dependable covariance matrix (roughly having entries 3 times more than the size of the variable state), then the adaptive procedure takes effect with a weighting between the calculated covariance and the initial one, based on the ratio of successful and total states sampled.

$$SD = \frac{2.4^2}{v} \tag{5.1}$$

(Where v is the number of variables)

$$C_o = SD * I_{vxv} \tag{5.2}$$

(The initial covariance matrix, no interaction between variables is available)

$$C = \beta * SD * \mathrm{cov}(Sample) + (1-\beta) * C_o \tag{5.3}$$

(Where β is a factor weighting the effect of the sample with its successful realizations)

$$\beta = \frac{realizations_{(\leq tolerance)}}{realizations} \tag{5.4}$$

5.1.2. Markov Chain Monte Carlo performance

After calibrating the three Metropolis-Hastings approaches, an analysis set is ran in order to assess the behavior and potency of each one. Firstly, the analyses are executed for 10.000, 25.000 and 50.000 realizations, in order to identify the number required for convergence.

In the sake of simplicity for this introductory investigation, all distributions are normal. The soil is chosen to be a lightly consolidated clay, which is a typical embankment material. Considering the assumption on the prior distribution, its undrained strength has a mean of 40 kPa and a standard deviation of 4 kPa, while its stiffness has values of 50.000 kPa and 12.500 kPa respectively. For the target distribution, the reference field is generated with strength distribution identical to the prior, whereas the stiffness has a mean of 75.000 kPa and a standard deviation of 7500 kPa (Table 4). Of course, allowing the same distribution representing a stronger soil in both positions, means that elastic behavior dominates the development of displacements and soil stiffness is the actual target of the inverse analysis. This simplified inquiry will provide insight on the methods and act as a basis for later analyses.

Table 4: Prior and target distribution properties of the main MCMC analysis

PRIOR DISTRIBUTION

DISTRIBUTION / PARAMETERS	Mean (kPa)	Standard deviation (kPa)	Coefficient of Variation (-)
UNDRAINED SHEAR STRENGTH	40	4	0,10
YOUNG'S MODULUS	50.000	12.500	0,25

TARGET DISTRIBUTION

DISTRIBUTION / PARAMETERS	Mean (kPa)	Standard deviation (kPa)	Coefficient of Variation (-)
UNDRAINED SHEAR STRENGTH	40	4	0,10
YOUNG'S MODULUS	75.000	7.500	0,10

Also, 20 measurement points are established in the domain, as explained in paragraph “Monitoring scheme”. The observations are taken from areas of interest, which are expected to wield a heavy impact on the analysis, and at an adequate distance from the boundaries, mimicking the measurements that would be given by inclinometers. Besides, the measurements act as boundaries to the inverse analysis and this population is both realistic and capable of providing a decent estimation of the posterior, even though they are far less than the used variables.

Following, the results of the first analysis set for every Metropolis-Hastings variation are illustrated. Since strength was not in the scope of these analyses, only the results of stiffness act as criteria for method performance.

The outcome is firstly judged based on the posterior distribution parameters of the relevant field points for each realization. Every version of the Metropolis-Hastings algorithm is able to approximate the target mean of 75.000 kPa starting from a prior mean of 50.000 kPa (Fig. 39). The univariate approach is by little the most

efficient, while the multivariate has the largest deviation from the target (almost 1%). It is evident that the multivariate variation has the largest spread, which is also the reason why it cannot operate with a 2% tolerance as the other approaches. Furthermore, the original algorithm exhibits a fluctuation of the mean around a fixed value along realizations, as does the univariate one, with the exception that in this case the fluctuation width is considerably decreased in the end (almost 50%).

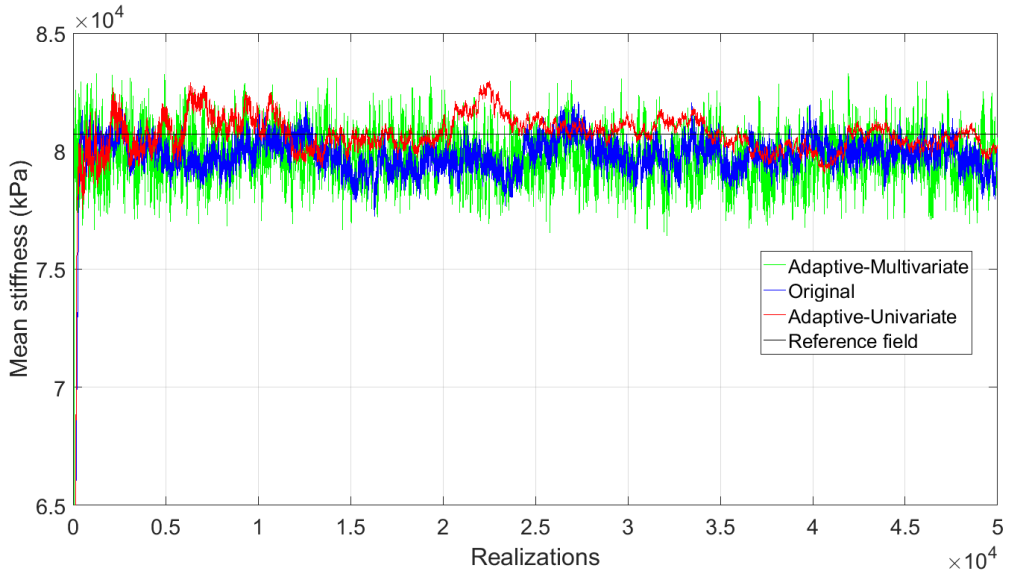

Fig. 39: Mean of stiffness field for every realization - Comparison over the MCMC variations

As with the field mean, the coefficient of variation (CoV) has a different prior and target value (Fig. 40). All methods start from a 0.10 and have to reach the target of 0.17 for the relevant Gauss points. However, only the Original version manages to fluctuate around the proper value. The adaptive univariate approach takes into account the history of the produced Markov chain through the covariance matrix, thus tending to "lock" on specific standard deviation values. On the other hand, the original version only regards the last step of the Markov path, and thus is more flexible and seemingly able to better approximate the proper CoV value. The multivariate approach seems to have the largest spread of values, even though it employs the adaptive component, and it fluctuates around a wrong value, possibly due to the larger tolerance used.

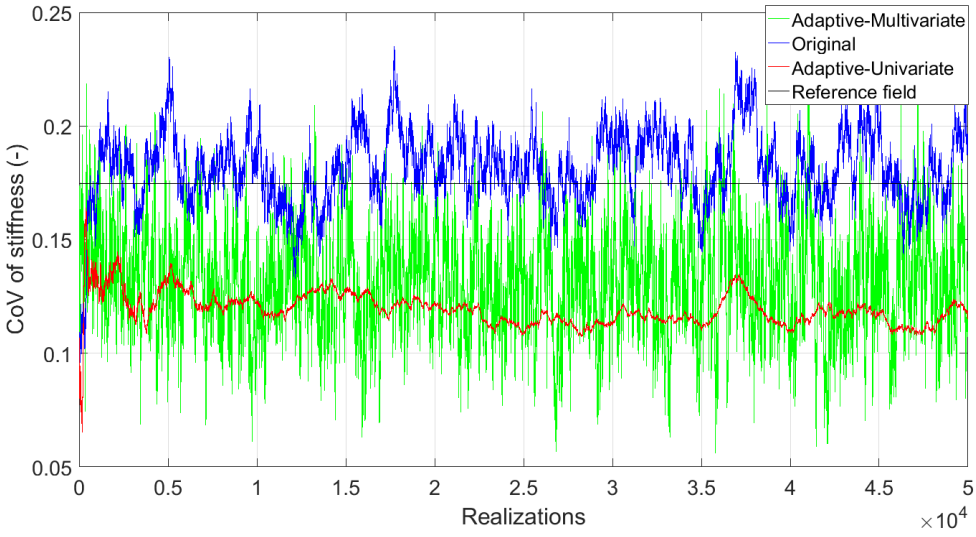

Fig. 40: Mean CoV of stiffness field for every realization - Comparison over the MCMC variations

At this point, it should be noted that the histograms generated by the sample of variables or field cells largely resemble normal distributions (Fig. 41). Thus, it is safe to approximate it with this distribution type.

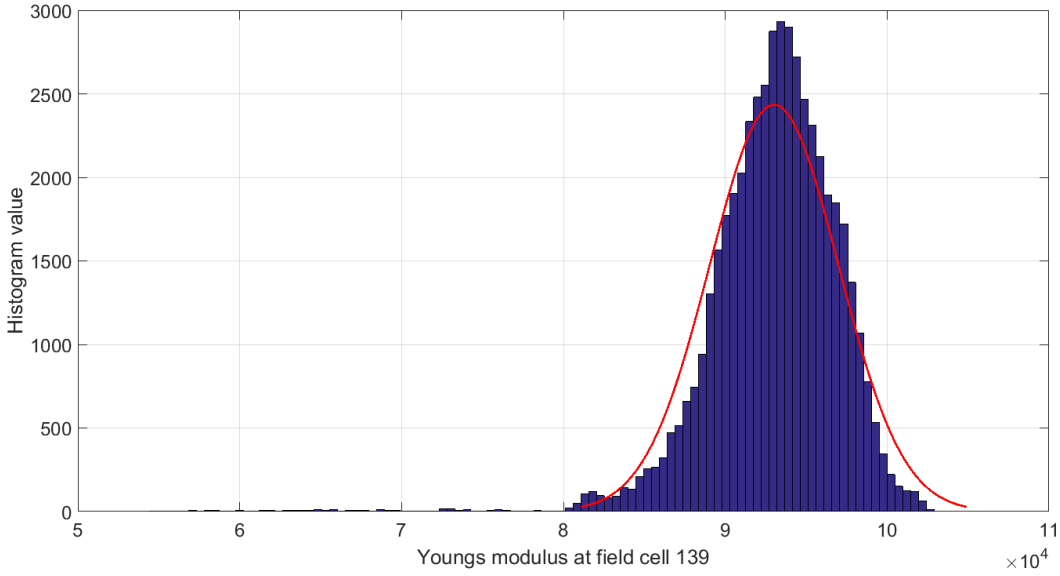

Fig. 41: Example: histogram and PDF fit for variable no.139

Concluding on this part, the probability density function plot of the three posteriors and the target distribution provides the final resolution. While all versions of the MCMC are close to the target, the Original Metropolis-Hastings provides the closest fit (Fig. 42), as also verified by the distribution fit error (Fig. 43).

$$\text{error} = \frac{\sqrt{\sum_{i=1}^{n}(PDF_{POST}^{i} - PDF_{TARGET}^{i})^2}}{n} \tag{5.5}$$

(Where n is the number of examined points)

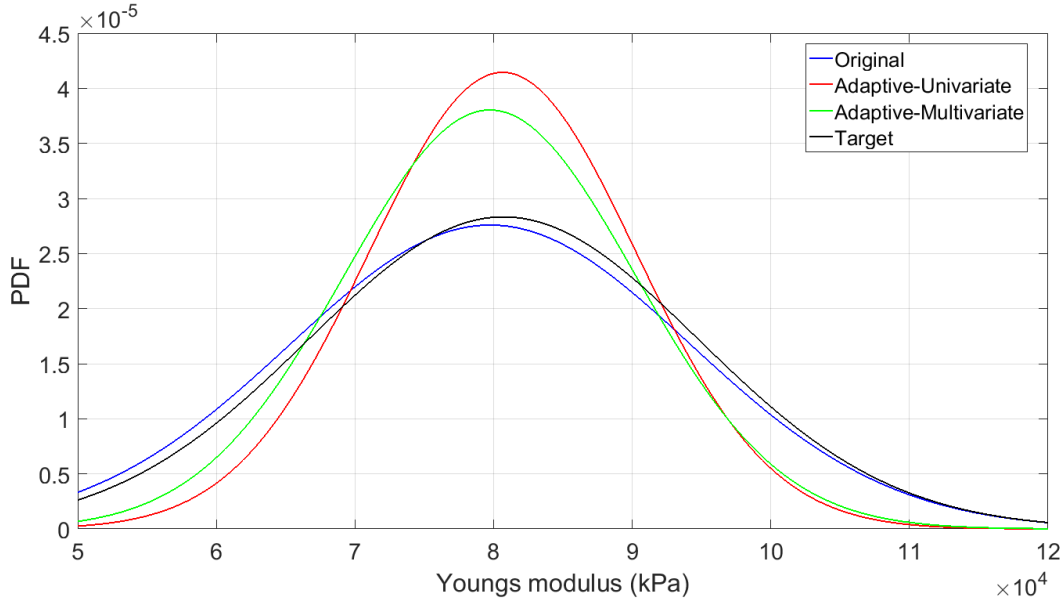

Fig. 42: Comparison of PDFs of means over stiffness field for the MCMC variations

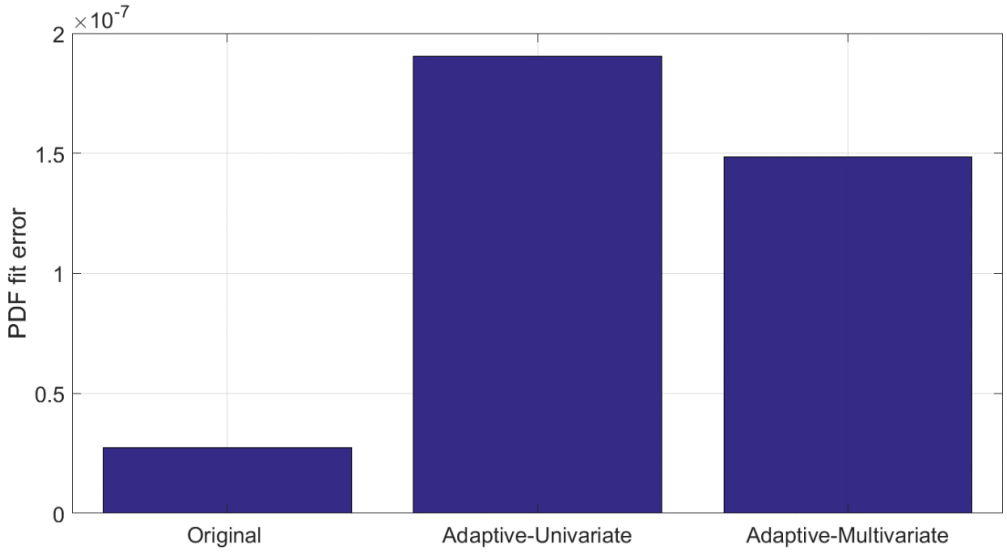

Fig. 43: Error of posterior fit to target for MCMC variations

Aside from the posteriors of field properties, the posteriors of the inverse analysis variables hold some noteworthy value, being compared to the variable values utilized to generate the random field. For example, in figure (Fig. 44), the posterior of variables 78 and 79, the first for stiffness, are illustrated. Since they are connected to eigenvectors with higher eigenvalues, they are expected to bear a heavier effect on the process. The Original Metropolis is yet the best fitting method, however, this is possible to differ for some variables. Besides, it is noteworthy that all methods are able to track the variable reference value, even when it significantly deviates from a normal Gaussian.

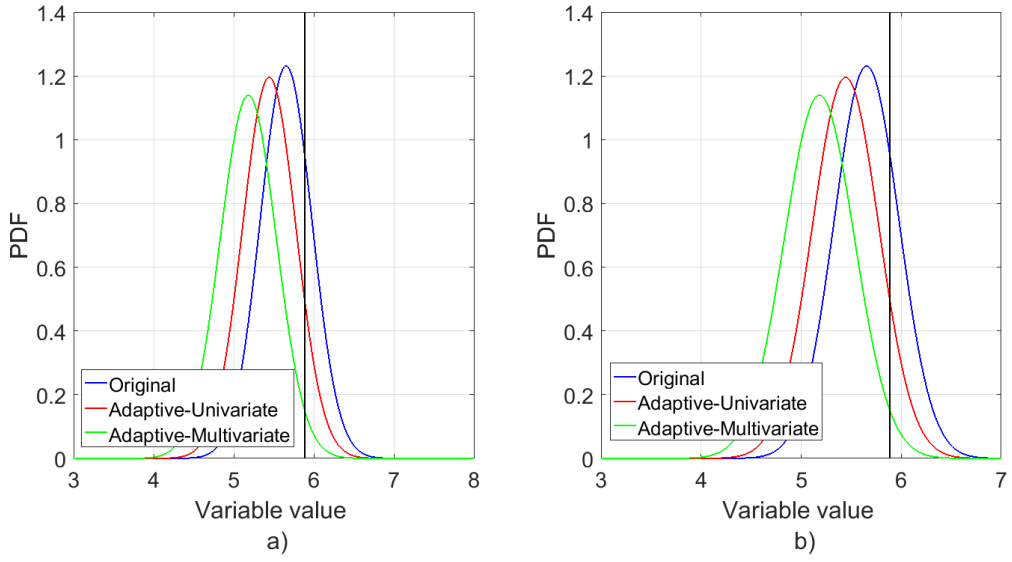

Fig. 44: Examples of posterior variable distributions and their target value for: a) variable no.78, b) variable no.79

Following, the minimum number of realizations needed to ensure method convergence has to be determined. Based on (Fig. 45 and Fig. 46), all three methods have acquired their final values of field mean and coefficient

of variation by 15.000 realizations. Hence, the following analyses can be safely assumed converged at this number of iterations.

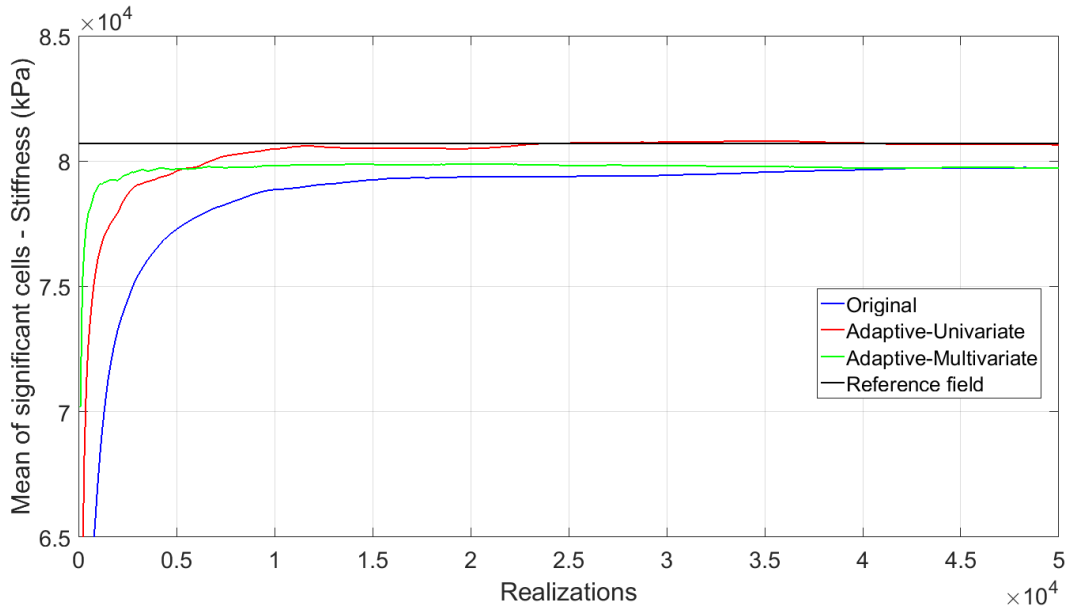

Fig. 45: Mean of field calculated retrospectively for every realization – Deduction of when is convergence achieved

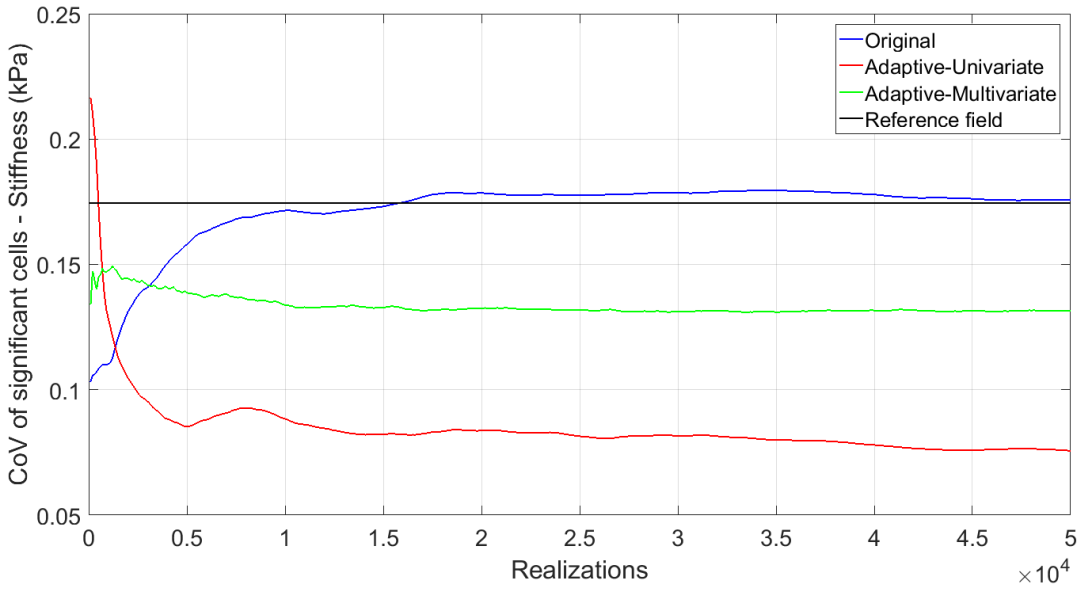

Fig. 46: Mean CoV of field calculated retrospectively for every realization - Deduction of when is convergence achieved

Furthermore, the error plot over the realizations for the Original Metropolis-Hastings can lead to important deductions (Fig. 47). At realization no. 569 the error drops below tolerance and analysis has reached a solution domain. After that, the Markov path randomly wanders in the domain, collecting the MCMC sample. Error values vary as samples can be drawn closer to the solution. Notably, the point of entering the solution domain is close but not identical to the point where the method approaches the target mean (Fig. 39). This means the difference between the posterior and target means is compensated by soil heterogeneity.

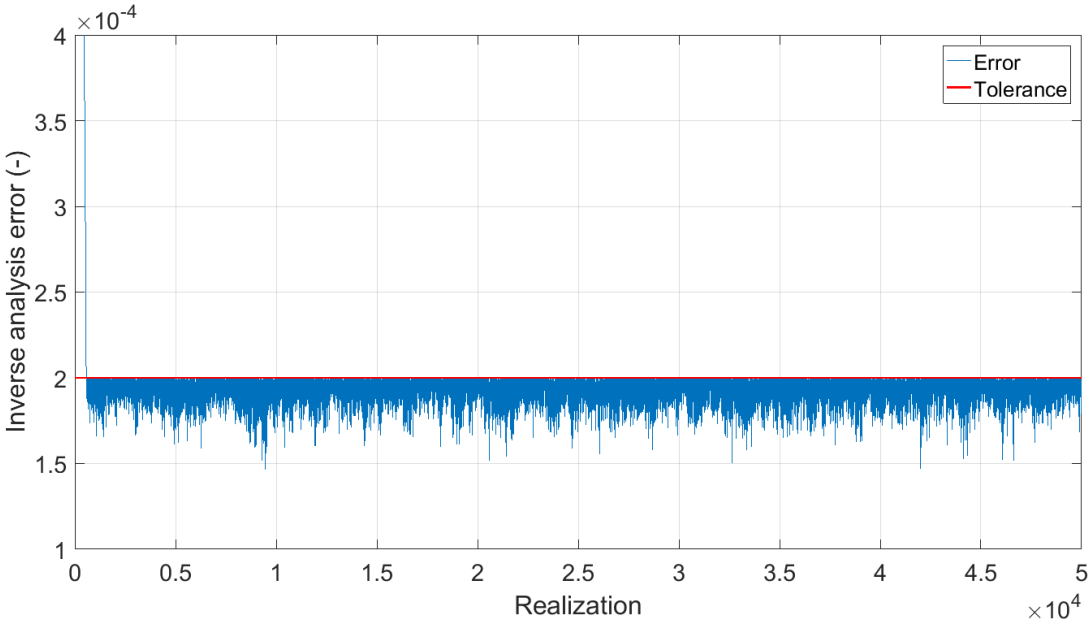

Fig. 47: Error convergence in the original MCMC

Finally, the computational efficiency of each approach is addressed. The acceptance ratio is simply the number of variable states accepted by the Markov criterion to the total number of FEM iterations. A low value is interpreted to a Markov chain that sampled an abundance of rejected variable states. As expected, the adaptive algorithms have a higher acceptance ratio, as the method evolves and chooses the proposal variable states based on previous successes (Fig. 48). The univariate approach, which stays focused in a domain of successful states due to its lack of variable interdependency, is established it as the most efficient method.

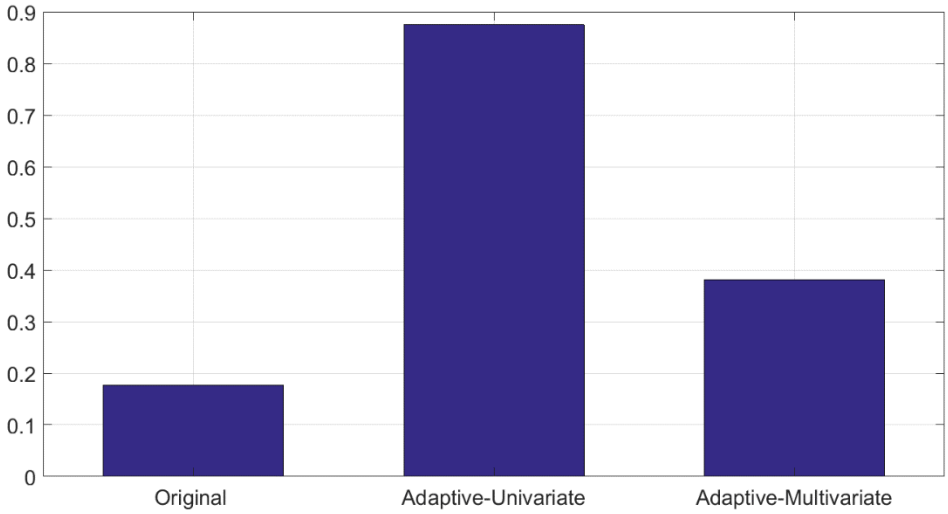

Fig. 48: Acceptance ratio of each MCMC variation

After determining the Original Metropolis as the most competent version of the Markov Chain Monte Carlo, its competency in allocating soil heterogeneity is put under examination. As seen in Fig. 49 where the mean stiffness value of each Gauss point is compared to its reference counterpart, the field approximation is quite successful. The location of stiff and non-stiff zones is almost exactly positioned, with the simulated field being smoother than the reference one, which covers fully the detail of soil heterogeneity. This is expected,

considering that the latter was generated with the full eigenvector set taken from the decomposition of the heterogeneity covariance matrix. The degree of field replication can be quantified when plotted along cells (Fig. 51) and the simulated and reference fields are compared at each Gauss point in the entire mesh and the areas of interest. In this analysis, the overall field error is almost 4% and the error of the relative points is estimated at 1%, results that signify a competent field approach. In order to grasp the measure of field fit error, a random field generated from the posterior distribution exhibits errors of 12% and 4% respectively (Fig. 50), meaning that the inverse analysis generated a posterior whose means approximate the original field roughly three times as well. Also, the reference field is at most points within the 95% confidence band. This fact implies significant confidence for the approximation of the reference field when sampling from the inverse analysis sample. Lastly, this accurate result is also validated by the PCA plot of (Fig. 52).The ellipse exhibits a center on the bisector, a first axis matching it and a relatively short second axis, meaning that the values of Gauss points are close to the reference values. Moreover, the data is centered on the mean, as also seen when plotting the ellipses for several confidence intervals (Fig. 53).

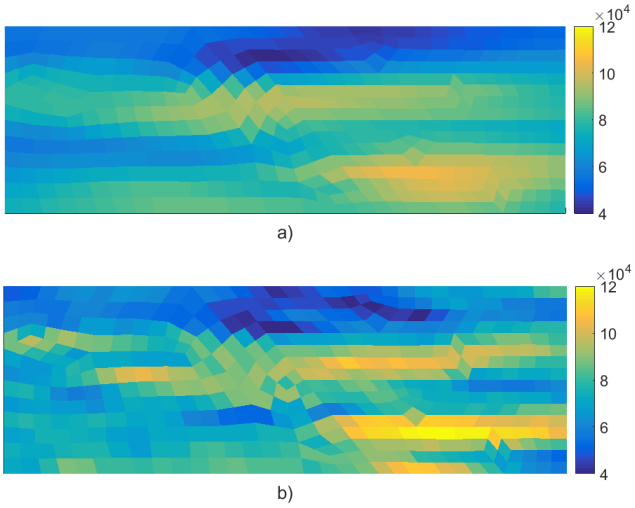

Fig. 49: Comparison of mean field (a) based on MCMC posterior sample to reference field (b) in mesh for Young's Modulus

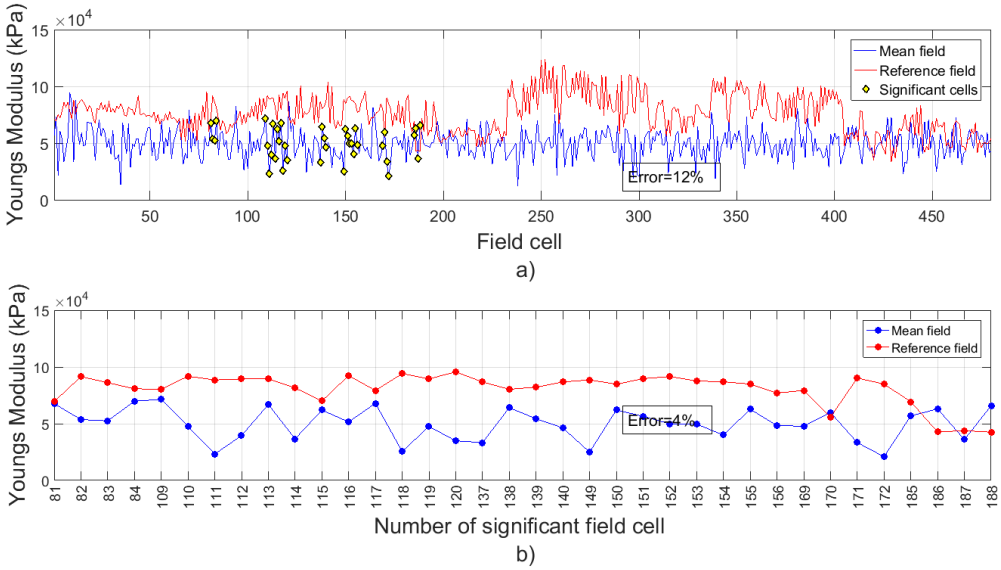

Fig. 50: A random field based on the prior compared to the reference field for the: a) entire mesh, b) cells of interest – Young's Modulus

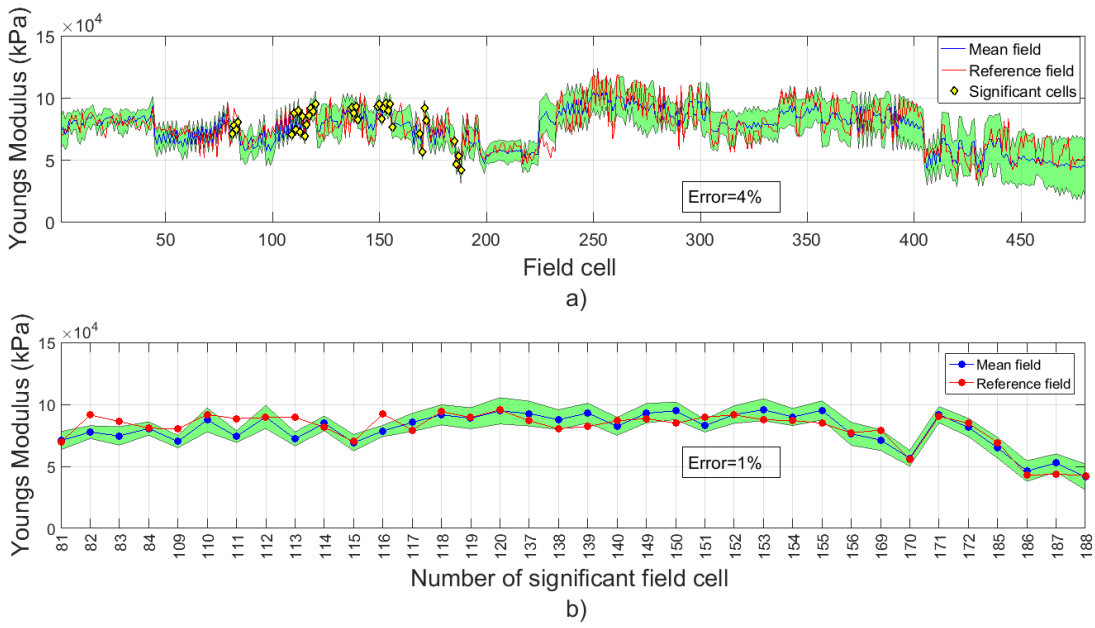

Fig. 51: Mean field of the MCMC posterior with the 95% confidence band compared to the reference field for the: a) entire mesh, b) cells of interest – Young's Modulus

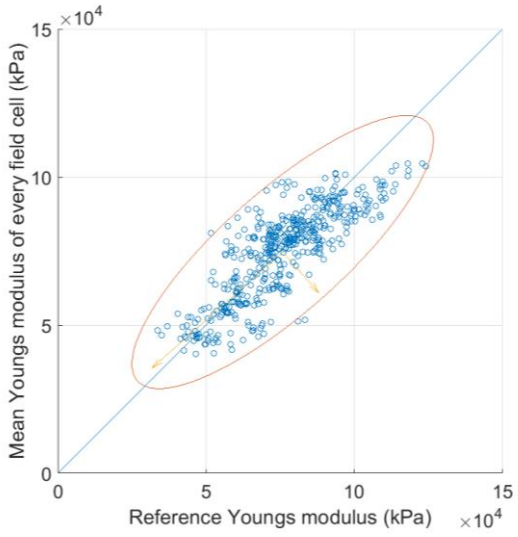

Fig. 52: PCA ellipsis for the MCMC posterior for the 68% confidence interval - Young's Modulus

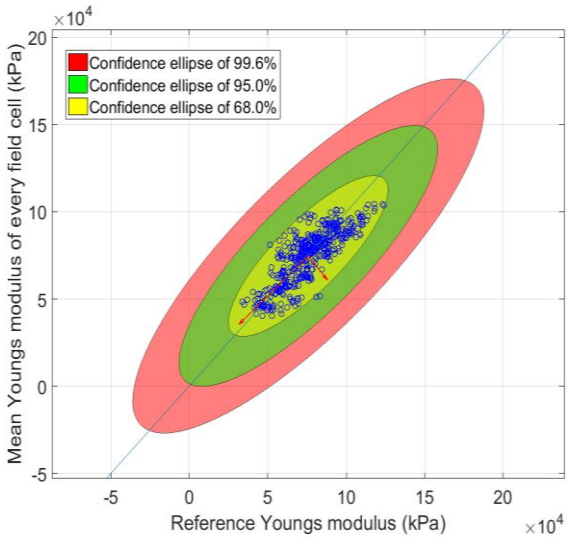

Fig. 53: PCA ellipsis for the MCMC posterior for the 68%, 95% and 99.6% confidence intervals - Young's Modulus

Moreover, upon inspecting Fig. 54, it is evident that the standard deviation of every field cell attempts to change so that when divided by the corresponding mean, the CoV revolves around the target value. Also, the Gauss points that are near the boundary and are excavated early in the FEM simulation, have little effect on the inverse analysis, and this is why their standard deviation does not follow the said pattern.

All in all, the application of the Original Metropolis–Hastings algorithm variation in the Markov Chain Monte Carlo method delivers a robust and competent inverse analysis scheme. This method succeeds in approximating both the parameters of the stiffness distribution, as well as the heterogeneity of the parameter. Thus, it is able to pose as a reference for the following analyses.

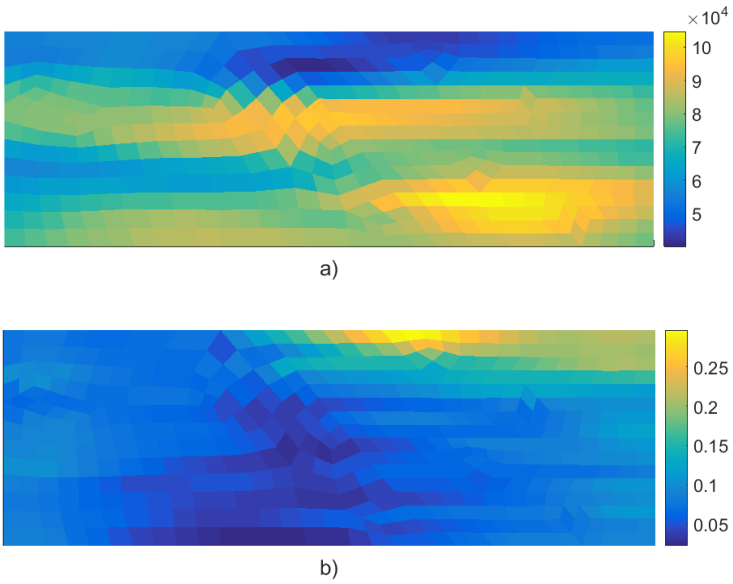

Fig. 54: Mean (a) and CoV (b) of MCMC posterior plotted in the mesh – Young's Modulus

5.1.3. Model for strength approximation

Following, the analyses have to include strength as an active variable of the inverse analysis. For that, a model is compiled where the target distribution of strength is different than the prior. Since the Tresca models dictates soil behavior, the effect of strength on the development of the displacements is only evident through failure. So, the target mean chosen should be low enough in order to allow for plasticity to develop, and thus enable the undrained strength to have a measureable effect on the mesh. According to this, the prior distributions, as well as the target of stiffness are taken as previously, while the target of strength is now set at 20 kPa mean and 2 kPa standard deviation (Table 5).

Table 5: Prior and target distribution properties of the MCMC strength scheme

PRIOR DISTRIBUTION

DISTRIBUTION / PARAMETERS	Mean (kPa)	Standard deviation (kPa)	Coefficient of Variation (-)
UNDRAINED SHEAR STRENGTH	40	4	0,10
YOUNG'S MODULUS	50.000	12.500	0,25

TARGET DISTRIBUTION

DISTRIBUTION / PARAMETERS	Mean (kPa)	Standard deviation (kPa)	Coefficient of Variation (-)
UNDRAINED SHEAR STRENGTH	20	2	0,10
YOUNG'S MODULUS	75.000	7.500	0,10

As seen from the following figures (Fig. 55 and Fig. 56), the results are unsatisfactory. The inverse analysis scheme is totally unable to track strength, and as a result of this, the approximation of stiffness is also spoiled,

while still maintaining a mean and CoV close to the reference ones[2]. The outcome is similar when compared for the entire field, or even both selections of the most significant Gauss points. Since the target distribution of strength has a lower mean than the previous analyses, plasticity now plays a bigger role than before in the creation of displacements, alongside stiffness. Hence, the impotency to approximate strength upsets the method's efficiency in simulating the latter. So, a question is now raised on why the MCMC is unable to approach the target distribution of strength.

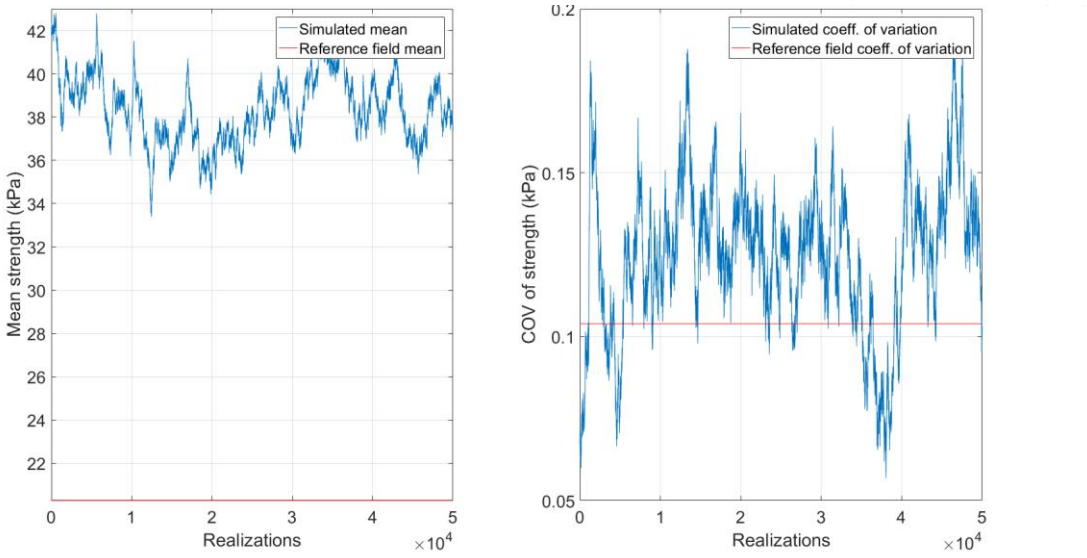

Fig. 55: Mean (a) and CoV (b) of strength over field for each realization – Original model and 20 kPa target mean

The most valid explanation lies in the constitutive behavior of the soil. The Tresca model is Linearly Elastic – Perfectly Plastic, meaning that plasticity is only accredited at soil element failure, and until then stiffness dictates the strain development. Thus, in MC the two variables of a Gauss point never have a simultaneous effect on the generation of its strains and no connection exists between them. Besides, strength has no part in the evolution of displacements until failure and only signifies its yielding, with no effect on its post-failure behavior.

But then, what are the conditions that have to be met in order to achieve a successful approach to the strength distribution? Intuitively, the proper way to manage this issue would be to describe soil behavior through an advanced, elastoplastic model. Then, at every stress – strain level, soil stiffness would also be a function of mobilized strength, taking into account both the undrained shear strength and Young's modulus. Then, this device would be able to properly account for the effect of strength. On the other hand, an elastoplastic constitutive model would expand the area of solutions, and so care should be delivered in order to track the reference distributions. Unfortunately, developing such a model in FORTRAN is out of the scope of this project.

Furthermore, additional evidence on this argument can be provided. Soil strength would be able to be tracked once its effect would be significant and regulated. In order to accomplish this, a new mesh is composed. The main idea is that a taller and steeper slope would develop yielding and fail in a lower undrained shear strength value, but would do so in a more controlled manner. Hence, a slope of m height 10m and 75° inclination is

[2] When soil yields, it seems that the initial stages have a greater impact on the identification of stiffness, whereas the latter mostly affect strength. More stages with elastic response could provide a better stiffness approximation (just as happened in the main analysis).

generated. Judging by Fig. 57 showing the development of deviatoric strains along the excavation stages, yielding develops in a smoother manner.

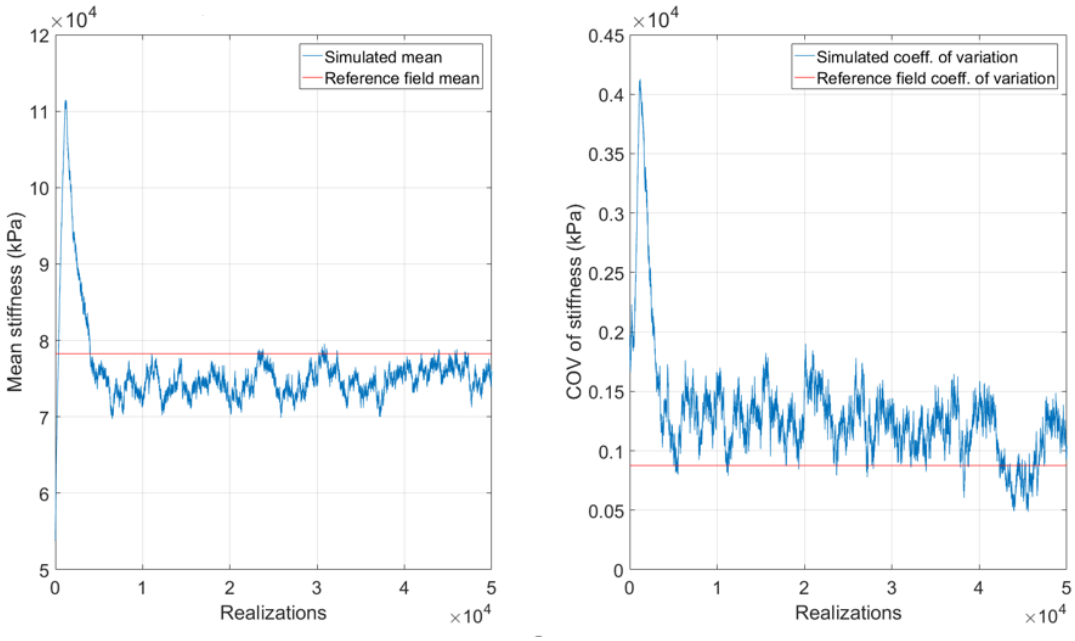

Fig. 56: Mean (a) and CoV (b) of stiffness over field for each realization – Original model and 40 kPa target mean

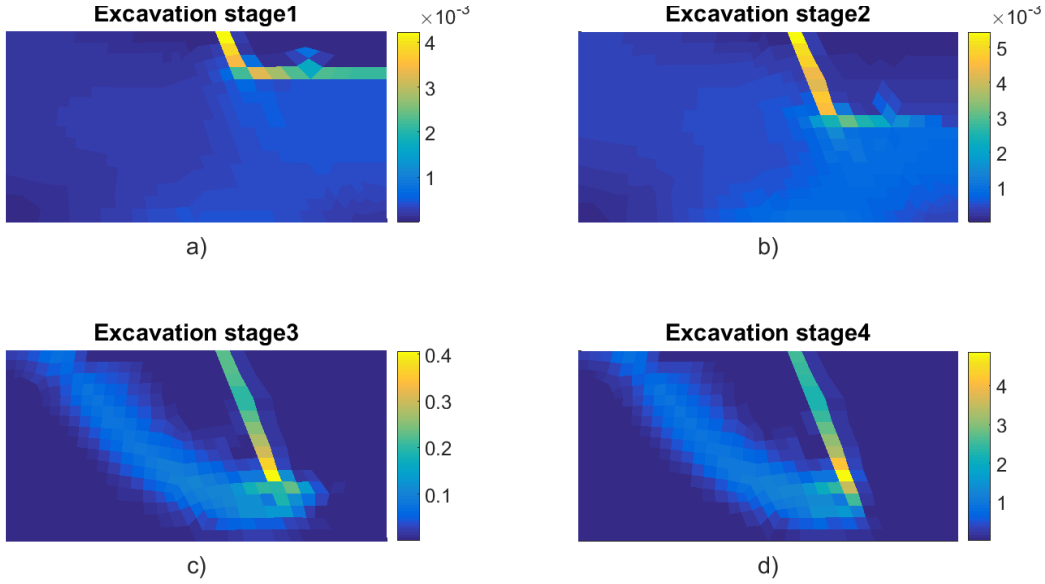

Fig. 57: Evolution of deviatoric strains in the new mesh - Homogenous strength field of 30 kPa

An inverse analysis is set up applying the Original Metropolis-Hastings algorithm in MCMC for the new mesh. Due to its complexity and time requirement, only 5000 realizations are performed, while the tolerance is fixed at 10^{-3}. The target distribution of strength has a 30 kPa mean and 3 kPa standard deviation, values potent to provide excessive plasticity. All other analysis parameters are set as before. Notably, the last excavation stage is left out of the simulation, as large displacements take place, denoting failure. The scope of this analysis is to inquire whether the MCMC has any potency to discover the target distribution and allocate strength heterogeneity.

The analysis is deemed successful. The posterior of stiffness is properly approaching its target (Fig. 59). More importantly, the mean of the strength distribution is now following a path, and finally reaches the target mean, while on the other hand the CoV is random. Thus, it is confirmed that the MCMC is actually able to approach the strength target distribution, even when soil response is described by the Tresca model. Although the analysis achieves its goal, the accuracy of the results is rather poor. Their quality is now a matter of required tolerance and number of iterations executed.

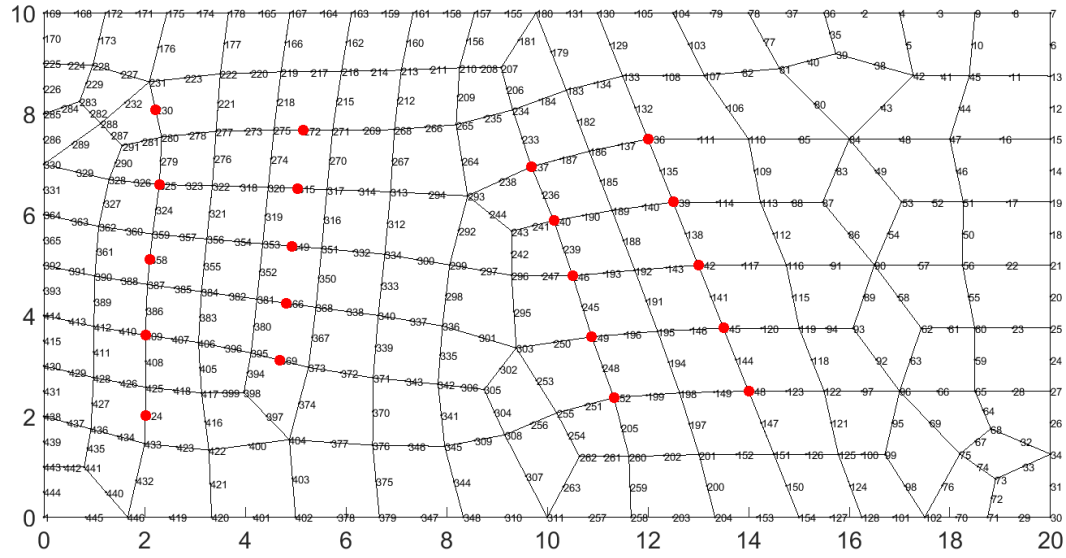

Fig. 58: Monitoring plan of the new mesh

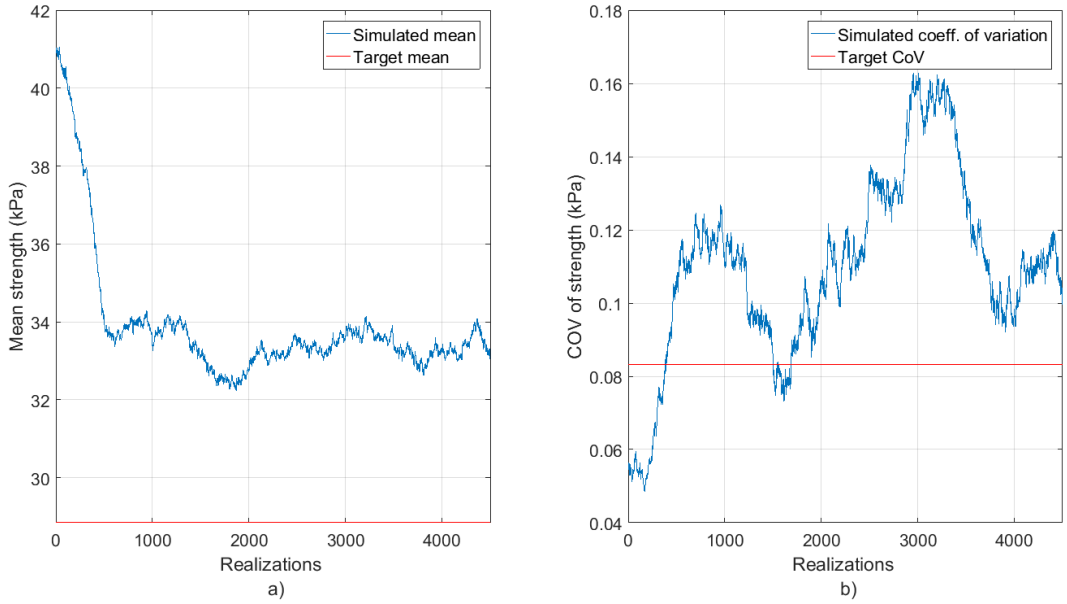

Fig. 59: Mean (a) and CoV (b) of strength over field for each realization –New mesh and 30 kPa target mean

The product of this analysis is not properly qualified as an inverse analysis result due to its inaccuracy. However, it is proof that the possibility of approximating strength distributions exists even when applying the Tresca constitutive model. This achievement is feasible under regulated circumstances, specifically chosen in order to set up a paradigm for the competency of MCMC. After all the adequate approach to dealing with strength detection in the inverse analysis, as mentioned previously, should be the application of advanced soil models.

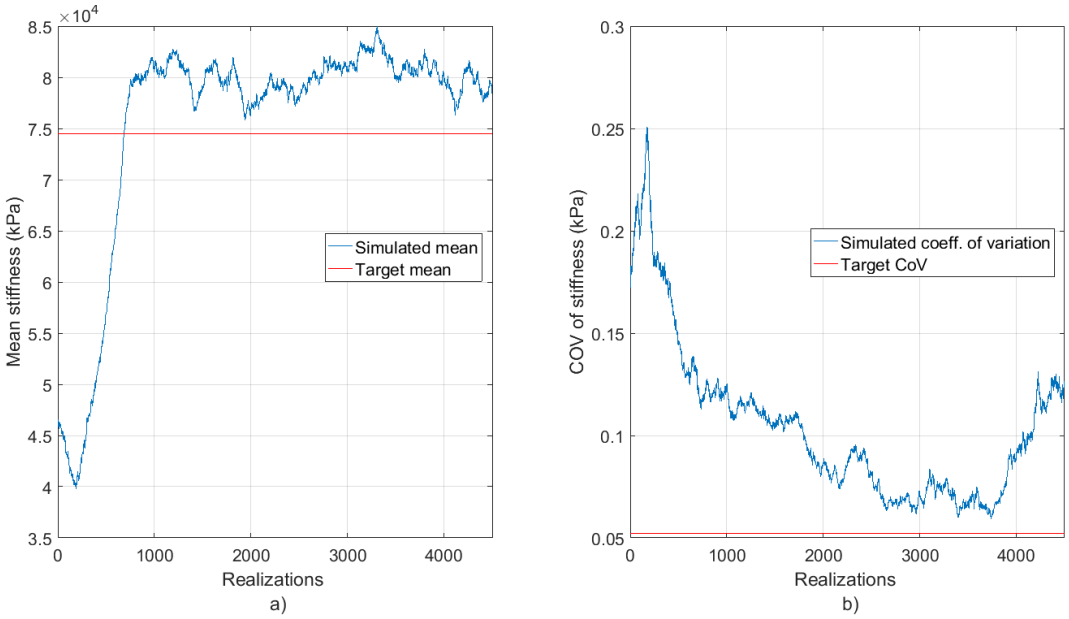

Fig. 60: Mean (a) and CoV (b) of stiffness over field for each realization –New mesh and 30 kPa target mean

5.2. Genetic Algorithm

The second inverse analysis scheme utilized is the Genetic Algorithm (GA). The boundary value problem is exactly the same as for the MCMC and the one presented in paragraph 4.5. The excavation of the slope takes place in 4 stages and soil behavior is described by the Tresca failure criterion. The tolerance allowed for this analysis is 0,2‰, which was the minimum manageable value. 20 measurement points are selected according to the presented monitoring plan. Moreover, the prior and target distributions are selected as previously (Table 6). The sampled reference field of the MCMC analysis is adopted, so that method comparison is possible, and thus the exact same target mean and CoV are reused. Still, strength is out of the scope of this analysis, and so the same target and prior are employed.

Table 6: Prior and target distribution properties for the main GA analysis

PRIOR DISTRIBUTION

DISTRIBUTION / PARAMETERS	Mean (kPa)	Standard deviation (kPa)	Coefficient of Variation (-)
UNDRAINED SHEAR STRENGTH	40	4	0,10
YOUNG'S MODULUS	50.000	12.500	0,25

TARGET DISTRIBUTION

DISTRIBUTION / PARAMETERS	Mean (kPa)	Standard deviation (kPa)	Coefficient of Variation (-)
UNDRAINED SHEAR STRENGTH	40	4	0,10
YOUNG'S MODULUS	75.000	7.500	0,10

The method operates as described in paragraph 2.4. The main parameters setting up the GA analysis are the crossover and mutation standard deviation, the population size and the number of required solution states. When it comes to calibrating some of them, the experience gained from applying MCMC is exploited.

Specifically, the standard deviation used for the crossover and mutation subroutines is set equal to that of the proposal distribution employed in the Metropolis-Hastings algorithm (Eq. 5.1). Moreover, after experimentation, the population is selected to be consisted of 100 individuals and the collection of at least 1000 successful variable states signify method completion. In this manner, the algorithm is able to adequately approach the target distribution and reference field while maintaining an acceptable computational toll.

Table 7: Genetic Algorithm parameters

PARAMETER	VALUE
CROSSOVER ST. DEV.	0,037
MUTATION ST. DEV.	0,037
POPULATION	100
NUMBER OF SUCCESSFUL STATES	1.000

5.2.1. Genetic Algorithm performance

The analysis proceeds until a sample of 1.000 successful individuals is collected and then its approximation of target distribution is judged (Fig. 61). For each one of them, the mean Young's modulus over the selected significant cells (80.235 kPa) is quite close to the target one (78.928 kPa), whereas the first generation started with a prior mean equal to 40.000 kPa. On the other hand, the method is unable to follow the coefficient of variation, exhibiting lower and randomized CoV values than the target.

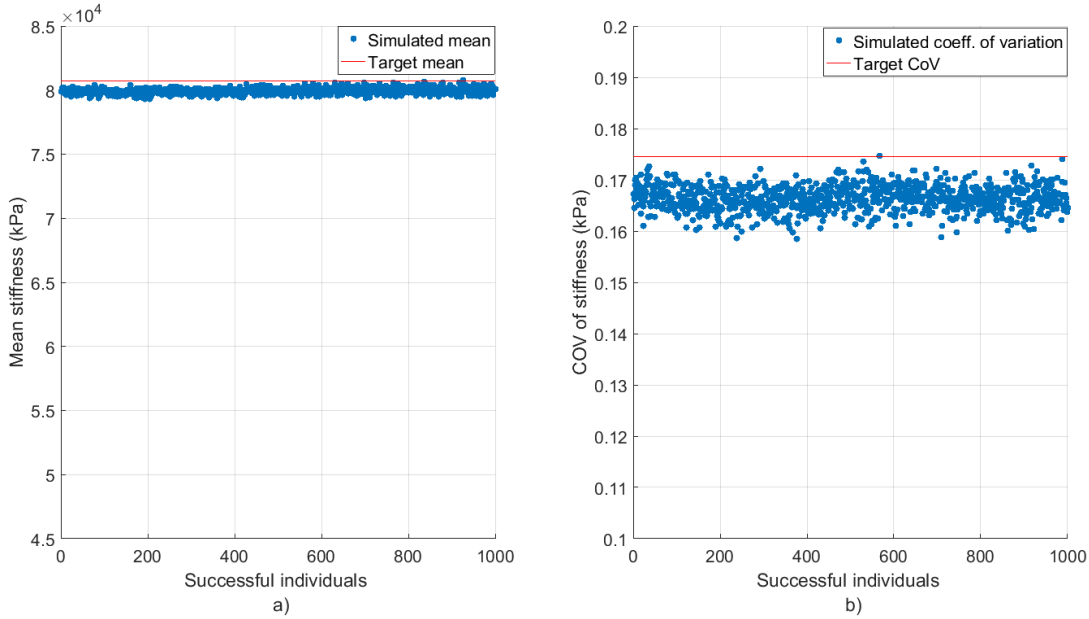

Fig. 61: Mean (a) and CoV (b) of stiffness over field for each successful individual

In Fig. 62 the development of the mean error of the population is compared to error of the fittest individual in each generation. While the two line almost match at the beginning, they deviate in latter generations. It is until almost generation 800 out of 960 that the fittest individual achieves an error lower than the tolerance. After that point, (mostly) its offsprings are able to create a sample of 1.000 solution in 160 generations, thus reducing the diversity of the GA sample. On the other hand, the mean error of the population never drops below tolerance, meaning that the GA population is diverse, but possesses states that are not acceptable as an inverse analysis result.

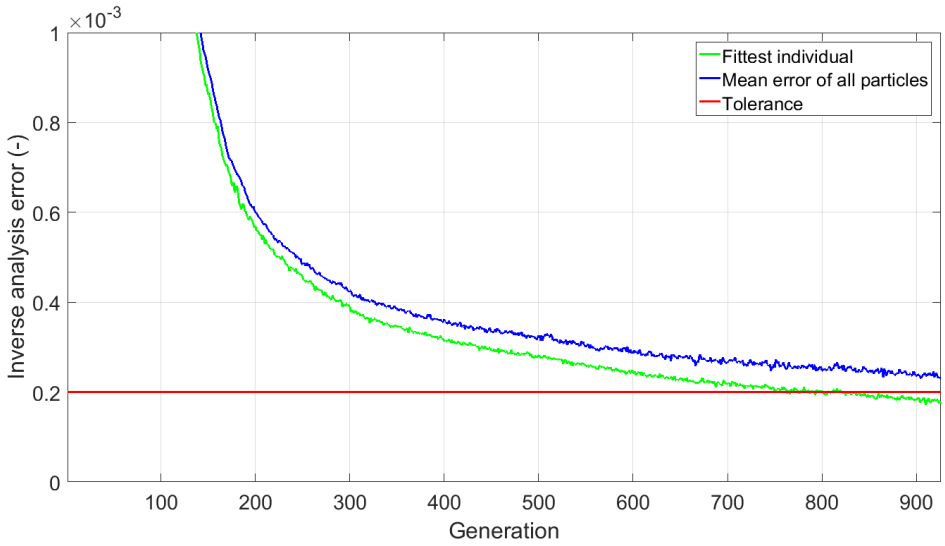

Fig. 62: Mean error of population and error of fittest individual evolution over generations

When it comes to the variable PDFs, method performance varies (Fig. 63). Although the posterior mean is always close to the target mean, the reference value is sometimes beyond the spread of ±3σ, meaning that sampling in its proximity has a reduced probability.

All in all, method efficiency can be judged by the comparison between the mean stiffness at each Gauss point and the reference field (Fig. 64). The GA manages a great reproduction of the reference field, with high and low stiffness areas having a similar distribution of the mesh. Mesh regions close to the boundaries or masses of soil excavated early on have a limited effect on the analysis, and so their emulation is poorer. Moreover, the simulated field has a smoother complexion, as a reduced eigenvector matrix was employed for its compilation.

After a PCA on the sample composed by the mean and reference of each field cell, the ellipse of Fig. 65 is drawn. The contour for 68% confidence interval includes almost the entire sample, which does not exhibit any entries especially far from the mean. Furthermore, the ellipse displays a tendency to match the first bisector; the major axis is on the line and the minor one has a low length ratio to the first. Ultimately, this interpretation of the GA outcome is positive on the method's efficiency.

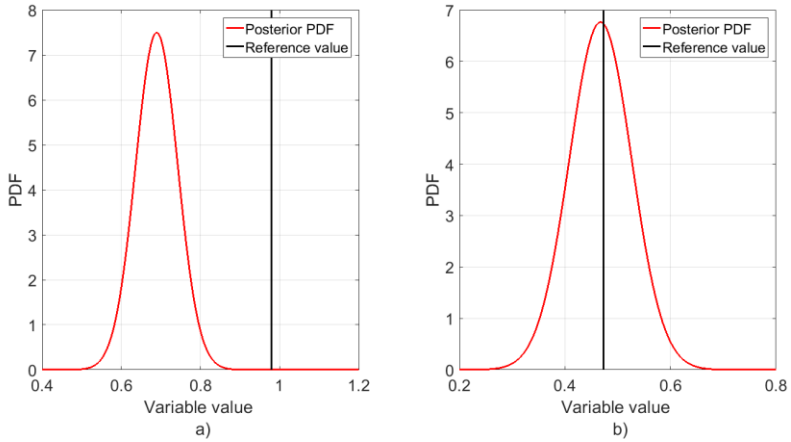

Fig. 63: Examples of GA posterior variable distributions and their target value for: a) variable no.82, b) variable no.154

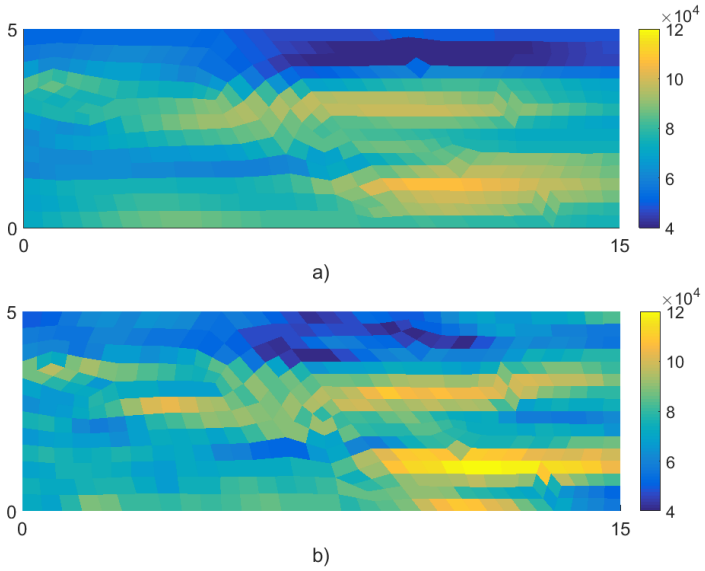

Fig. 64: Comparison of mean field (a) based on the GA posterior to reference field (b) in mesh for Young's Modulus

While the previous plots verify the favorable performance of the GA, the comparison of the mean and reference fields presents an arguable point (Fig. 66). The field fit is similar to the MCMC at an error of 4% for the entire field and 1% for the significant cells, 3 times lower than the error achieved by a random field, and thus confirms the method's ability to properly track the reference values. Nevertheless, it is evident that the 95% confidence band is unable to encompass the reference field, as the standard deviation of each cell is deemed too low. A short variance implies that the proximity of the mean area is reliably calculated, but when the target is far off, approaching it seems rather impossible. The answer to this issue lies in the user defined configurations made on the GA. The method runs with a population of 100 individuals and stops when it collects a sample of 1.000 successes. As a result, 774 generations where simulated until the analysis was finished. While the GA is able to efficiently explore the variable space, when a solution is found, it is gradually attracted to it. Thus, after this large number of generations all individuals seem to have collapsed to the exact same solution, drastically reducing variable variance.

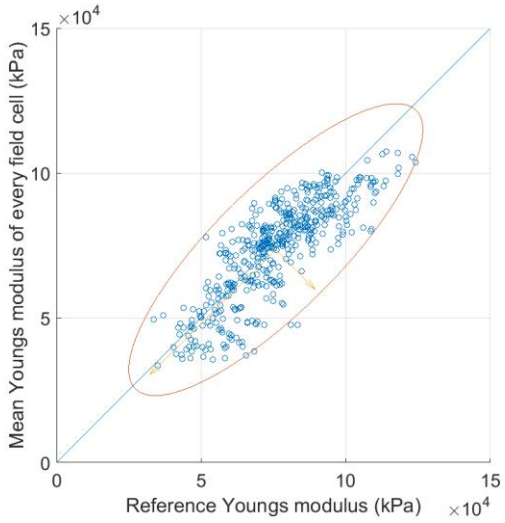

Fig. 65: PCA ellipsis for the GA posterior for the 68% confidence interval - Young's Modulus

Even though the reference values usually have a considerable distance of much more than three standard deviations, the produced fields still comply with the set inverse analysis tolerance. Also, the mean field has a

low fit error and largely imitates the reference one. More deductions about these results can be drawn when elaborating with them in a stochastic concept (chapter 6).

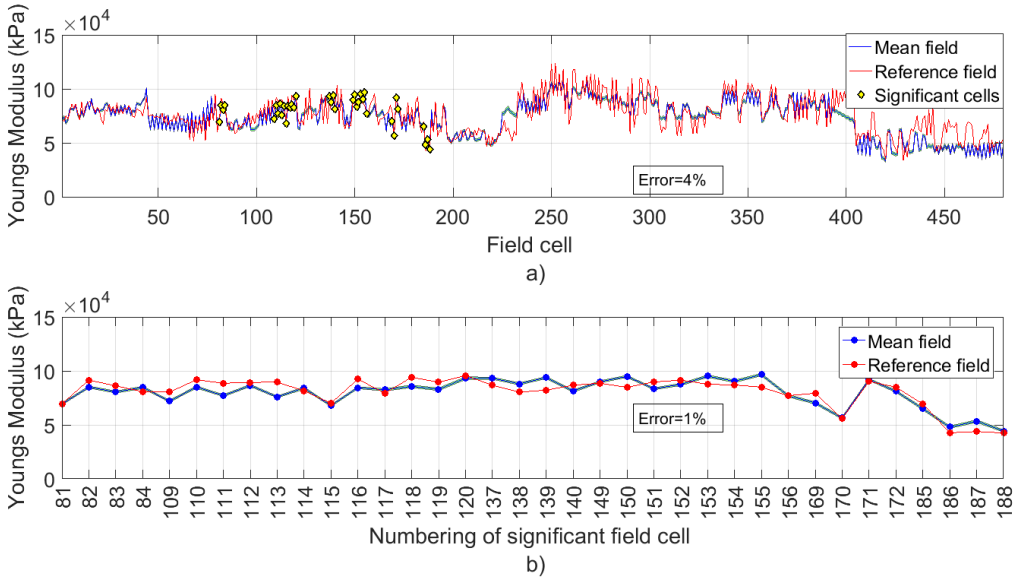

Fig. 66: Mean field of the GA posterior with the 95% confidence band compared to the reference field for the: a) entire mesh, b) cells of interest – Young's Modulus

5.2.2. Genetic Algorithm variation for strength approximation

After applying the Genetic Algorithm for the recognition of the stiffness field, an extension of method is set up in order to deal with soil strength, since the normal analysis is not potent to allocate it. The main notion adopted is that a diverse population is generated, with every individual having a different strength prior distribution. Whereas directly approaching the strength target distribution seems impossible when employing the Tresca model, now individuals with strength closer to the target are expected to thrive in their generation, and thus dominate the population. As a first comment, it is obvious that this concept is applicable only when the soil is weak enough to yield, hence rendering the approximation of stronger soils futile.

The exact same model as before is applied but the strength reference field is reduced to an undrained shear strength of 20 kPa. As proven in "Mesh selection", this value is enough to achieve extensive yielding in the mesh. Then, the population of the first generation is created, with every individuals having a different prior distribution. The mean is sampled from a uniform distribution between 15 kPa and 40 kPa, while the CoV is always fixed at 0,10. Even though the strength field cannot be simulated, the stiffness posterior is relatively close to the target, even in cases where significant plasticity take place. As a result, a series of analyses is composed for this scheme; firstly an ordinary run is executed in order to estimate the posterior of Young's modulus and the result is then employed as a stiffness prior for allocating the undrained shear strength.

Table 8: Target distribution properties for the GA strength scheme

TARGET DISTRIBUTION			
DISTRIBUTION / PARAMETERS	Mean (kPa)	Standard deviation (kPa)	Coefficient of Variation (-)
UNDRAINED SHEAR STRENGTH	40	4	0,10
YOUNG'S MODULUS	75.000	7.500	0,10

Observing the means and CoVs of all 235 successful individuals (Fig. 67), the method seems to efficiently discover the target distribution. The successful individuals are in the proximity of the target values after simulating 29 generations. The evident trends in both plots allow for a prediction of the posterior, assuming that the inverse analysis continued to collect acceptable variable states. Moreover, it should be noted that the described results are for the selected significant cells of the mesh, as presented in paragraph 4.44.4. Since yielding primarily develops at the toe of the slope, this area is considered as more important, having a powerful impact on the produced displacements. The logic behind this choice is confirmed by the illustration of the mean and reference fields in the mesh (Fig. 68). Apparently, areas that offer a better replication of the strength field are the ones near the about-to-be-formed shear band (Fig. 30Fig. 25). As previously stated, the analysis has an upper limit for the strength distributions it is able to approach, due to the nature of the Tresca constitutive model. Thus, all areas that do not yield have a limited effect to the analysis, meaning that the respective field cells end up with a seemingly random result for strength.

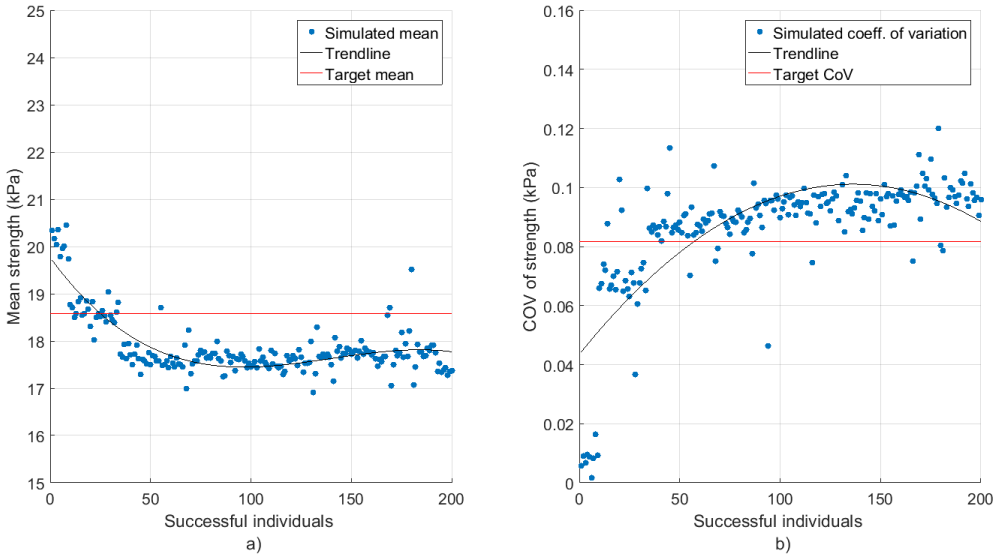

Fig. 67: Mean (a) and CoV (b) of strength over field for each successful individual with their respective trendlines

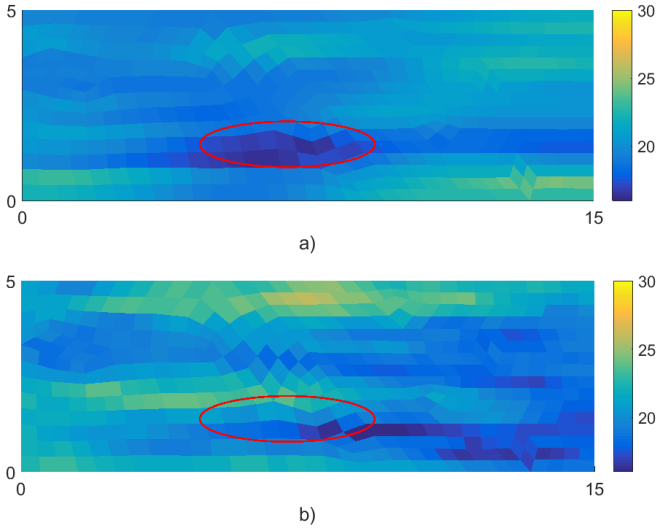

Fig. 68: Comparison of mean field (a) based on the GA posterior to reference field (b) in mesh for Young's Modulus – The area of interest is highlighted

Finally, the plot of the mean and reference fields along all Gauss points is able to quantify the outcome of the method (Fig. 69). When it comes to the cells of interest, the field approximation has an error of 1% as in the analyses for stiffness. Nevertheless, the entire mean field now has a 6% deviation from the reference. Besides, Gauss points that do not yield have a randomized posterior distribution, meaning that they fail to approximate the actual strength field, rendering most of the plot irrelevant.

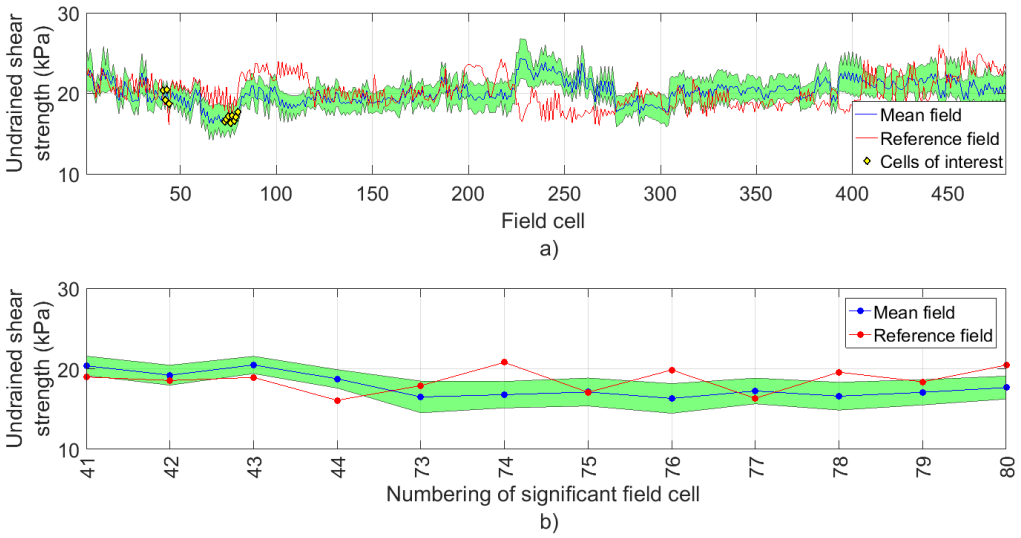

Fig. 69: Mean field of the GA posterior with the 95% confidence band compared to the reference field for the: a) entire mesh, b) cells of interest – Undrained shear strength

Overall, this application enables the recognition of the strength field at points of yielding, even when the Tresca model is used. It provides a decent approximation at the area of interest and is capable of allocating its target distribution. However, the method is not properly equipped to cope with higher strength values, which poses as the main reason of model inaccuracy. In order to overcome this issue, an advanced soil model would have to be employed.

5.3. Regularized Particle Filter

The final method to be tested in the inverse analysis problem is the Regularized Particle Filter (RPF). When employed in the main inverse problem, this approach failed to deliver acceptable results. The RPF is mainly used in cases where an adequately long timeline of observations is available, such as in meteorology and oceanography. Thus, the particles have enough time to be conditioned in order to produce errors within tolerance. On the other hand, the excavation investigated here only simulated four stages, a sequence that did not allow the error of the ensemble to drop satisfactorily. Hence, a new model is compiled as an argument for the incompatibility of the RPF in the main case.

The new model includes a higher slope of 10m with an inclination of 45°. The construction takes place in 10 excavation stages of equal depth progress. Also, the FEM model parameters are all set as in the previous analyses (paragraph 4.5). Still, the construction stages do not seem to be enough in order to track the target distribution when it is different than the prior. Therefore, the first scheme described in paragraph 4.4 is adopted; the prior is equal to the target. Essentially, this implies that there is some prior knowledge of the soil properties, e.g. site investigation has been performed, and the inverse analysis attempts to properly replicate soil heterogeneity. The prior and target distributions have a mean of 75.000 kPa and CoV of 7.500 kPa. Since strength is not easily approximated, stiffness is the only unknown of the analysis.

The analysis operates retrospectively, in the sense that particle error is calculated for all stages up to the current iteration. In this manner, particle performance at the final iteration is representative of their competency along the entire excavation. Finally, successful variable states are defined those that after the final iteration are able to produce an error below the tolerance of 0,5‰.

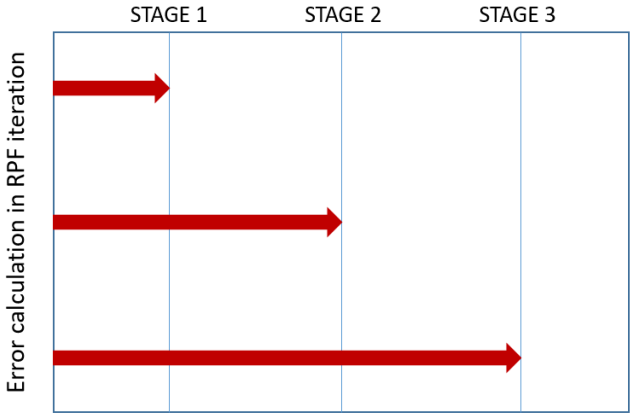

Fig. 70: Illustration of retrospective error calculation in each RPF iteration

The RPF analysis utilizes an ensemble of 500 particles, which are simulated over 10 excavation stages. Fig. 71 presents the histogram of errors for the last iteration of the method. Specifically, they appear to follow a log-normal distribution and most of them are concentrated at 0,6‰. In addition, Fig. 72 is an example for ensemble evolution of a variable in each iteration. Particles start rather scattered by the Monte Carlo initialization of the sample, but they steadily converge to a final value, while also exploring the domain by acquiring some distant values. In particular, since the first eigenvector is mainly responsible for the mean of the field and the prior is equal to the posterior, the first variable is seen converging to a value close to zero.

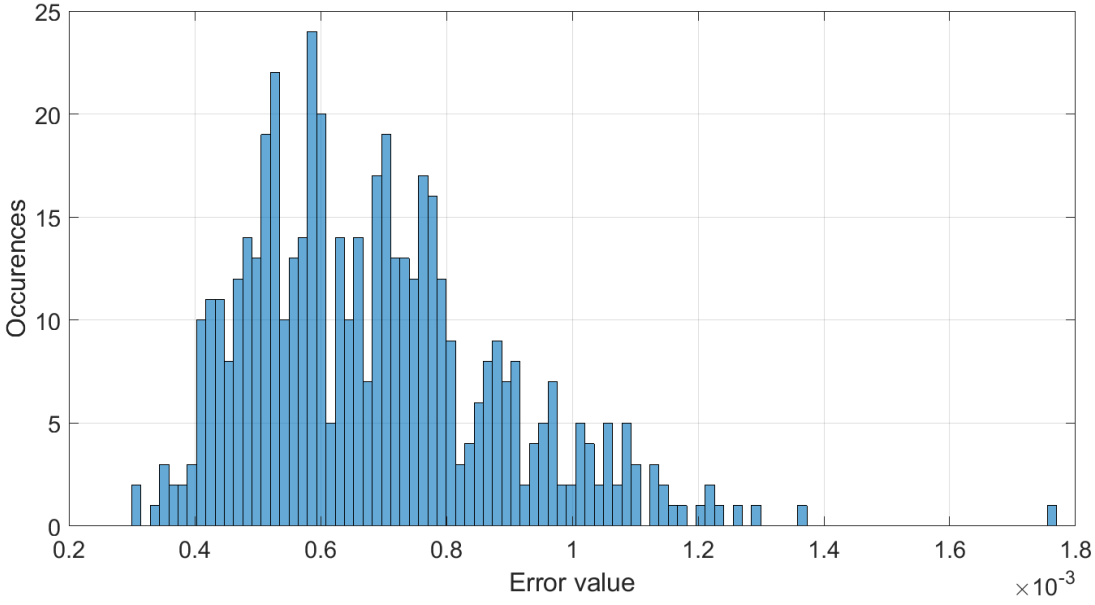

Fig. 71: Histogram of errors for the final ensemble

As previously, a quantification of method efficiency can be expressed by plotting the mean of the field for each particle of the final ensemble (Fig. 73a). While in this instance the scatter plot seems random, the nature of Filter methods is highlighted. Means of particles vary greatly; unsuccessful particles may have values close to

the target and vice-versa, as spatial distribution of parameters compensates for the inconsistencies. However, the field created by the mean of all particles, which according to Particle Filter theory is considered the solution, exhibits impressive results. It is not only a perfect much for the target distribution, but its displacements error is calculated at 0,2‰, which by far lower than the error of any particle. Moreover, the RPF is able to satisfactorily approximate the CoV of the target (Fig. 73b).

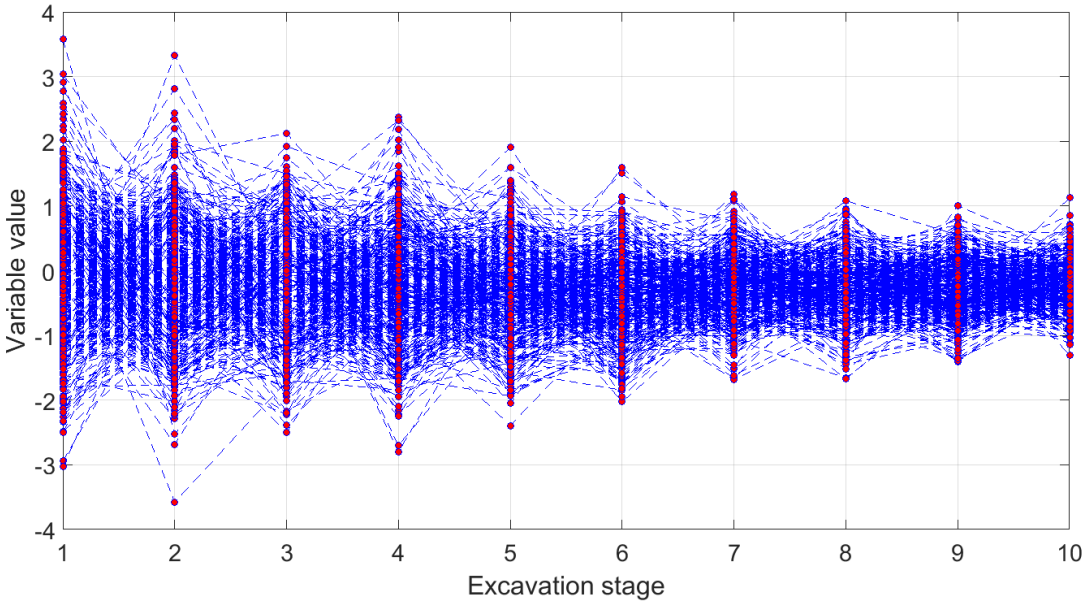

Fig. 72: Example of variable evolution illustrating particle values in each stage – Variable no.1

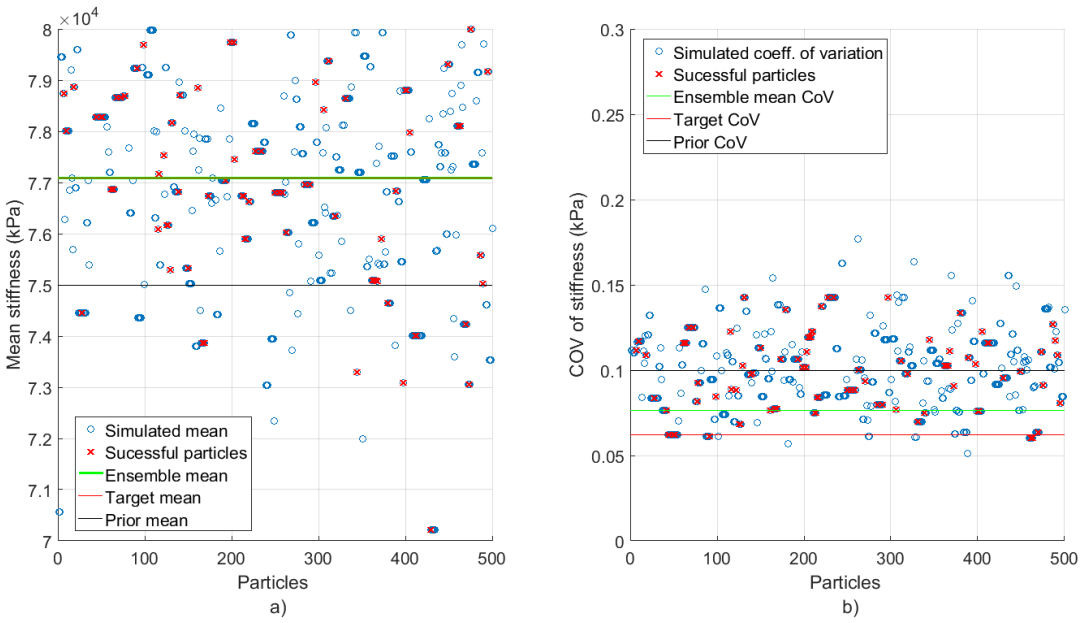

Fig. 73: Mean (a) and CoV (b) of stiffness over field for each particle along with the value of the ensemble mean

In Fig. 74 the evolution of the mean error of the ensemble and the error of the mean particle is displayed. Evidently, the results are as previously described. The mean error of the ensemble is relatively high and never drops below the selected tolerance of 0,5‰. On the other hand, the mean particle always provides a significantly better behaviour; it reaches the tolerance at the 5th excavation stage and achieves an error of 0,2‰ by the 10th. Notably, both methods exhibit their largest error values at the 2nd assimilation step, where

the particles seem to expand, attempting to explore the variable domain (Fig. 72). All in all, error convergence cannot be discussed in the RPF as it was for the other methods, since it now it is applied in a dynamic problem with no steady state loops. As a comment, it is possible that more assimilation steps could allow the mean error of the ensemble to reduce below tolerance.

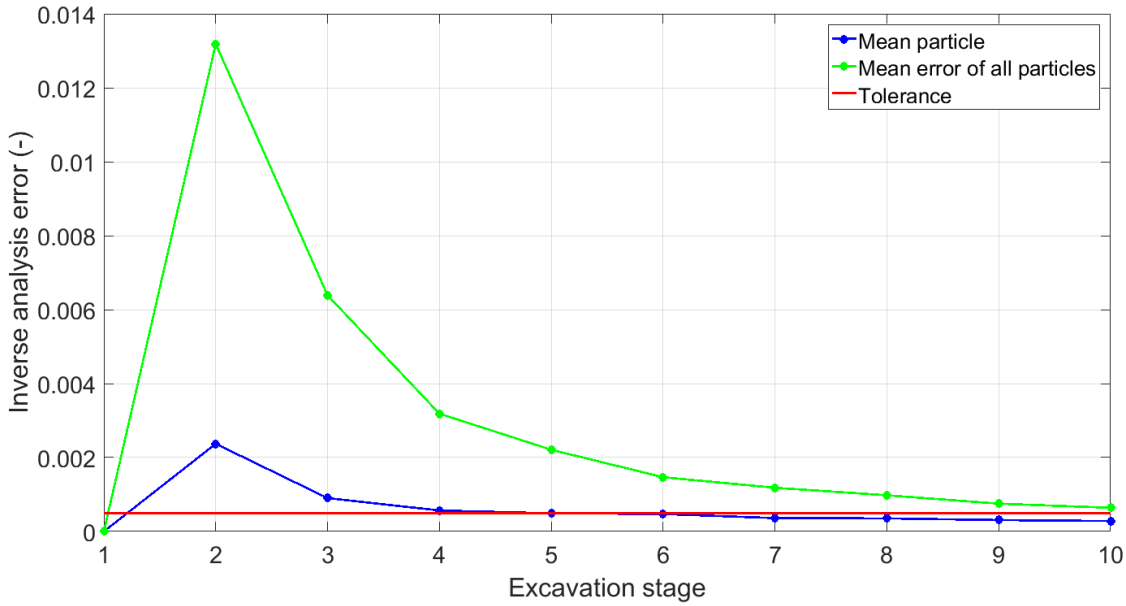

Fig. 74: Mean error and error of mean particle evolution over RPF assimilation steps

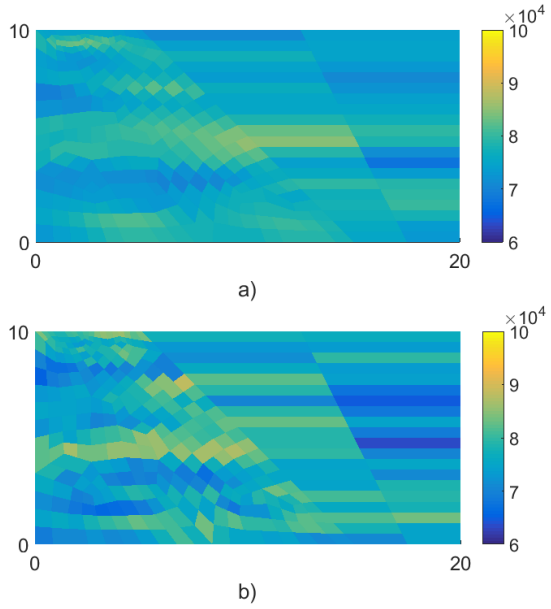

Fig. 75: Comparison of mean field (a) based on the RPF posterior to reference field (b) in mesh for Young's Modulus

Following, the comparison of the reference field and the field of the mean of the ensemble provide insight on the efficiency of the RPF (Fig. 75). Evidently, the approximation is quite successful, taking into account the inherent error of the simulation due to the reduced number of eigenvectors used. The sequence of more and less stiff layers is akin in both plots, especially in areas that bear a heavier impact on the outcome.

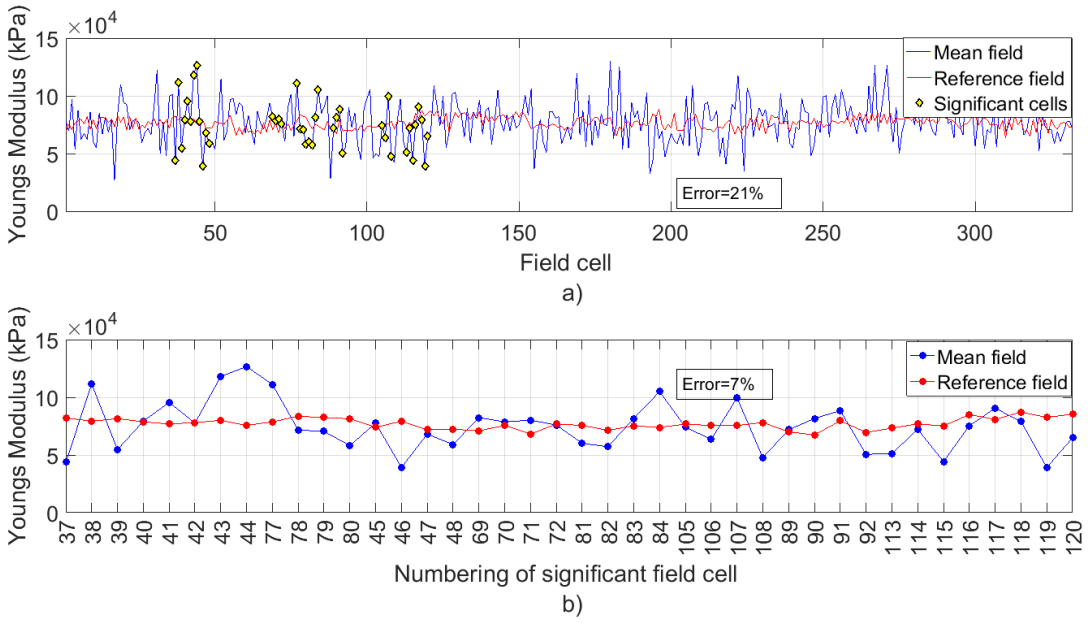

Fig. 76: A random field based on the prior compared to the reference field for the: a) entire mesh, b) cells of interest – Young's Modulus

Moreover, the benefits of the RPF inverse analysis can be witnessed in the field comparison plot. Despite using a prior distribution identical to the target, a random field cannot capture the reference (Fig. 76). However, the RPF is able to efficiently reproduce it and mimic its spatial variation (Fig. 77). The error of the entire field is now reduced from 21% to 1%, and the respective change for the cells of interest is from 6% to 0,3%. These dramatic alterations of the field fit error are proof of the RPF's competency as an inverse analysis method. Lastly, another feature of the RPF is that the 95% confidence band includes the reference field at all times, which is partly accredited to the efficient approximation achieved by the mean field and partly to the diversity of the sample. While this means that the reference field moves closely to the mean, approximating the mean field upon sampling is now harder, which might yield some significant drawbacks for method reliability.

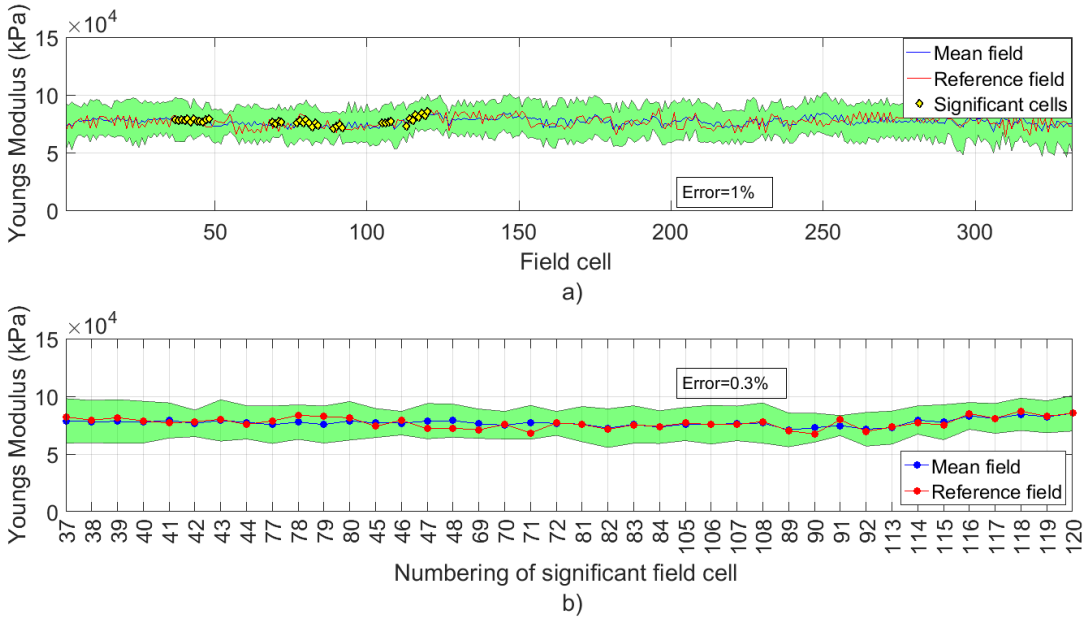

Fig. 77: Mean field of the RPF posterior with the 95% confidence band compared to the reference field for the: a) entire mesh, b) cells of interest – Young's Modulus

5.4. Conclusion

To sum up, this chapter focused on testing the performance of the three inverse analysis schemes on the FEM simulation of the staged slope excavation. In order to be put into comparison, five objective criteria are defined, quantifying the computational power and time consumption of each algorithm, as well as their ability to approximate the target distribution and the reference field of soil stiffness (Table 9 and Fig. 78). As a reminder, the RPF, while seemingly competent, operated on a quite different FEM model, therefore direct comparison to the other methods is improper. The performance of the GA is equally competent to that of the MCMC in identifying the target mean and CoV, along with the reference field. Besides, the GA is significantly faster in achieving convergence and with only one fourth of the FEM iterations. Furthermore, the stochastic interpretation of the results bears major importance and will be discussed in the following chapter.

Table 9: Inverse method performance based on five criteria

METHOD / CRITERIA	MCMC	GA	RPF
TIME(MIN)	600	240	90
FEM RUNS	283.406	77.400	5.000
MEAN APPROX. (%)	99	98	100
COV APPROX (%)	91	98	77
FIELD APPROX (%)	96	96	99

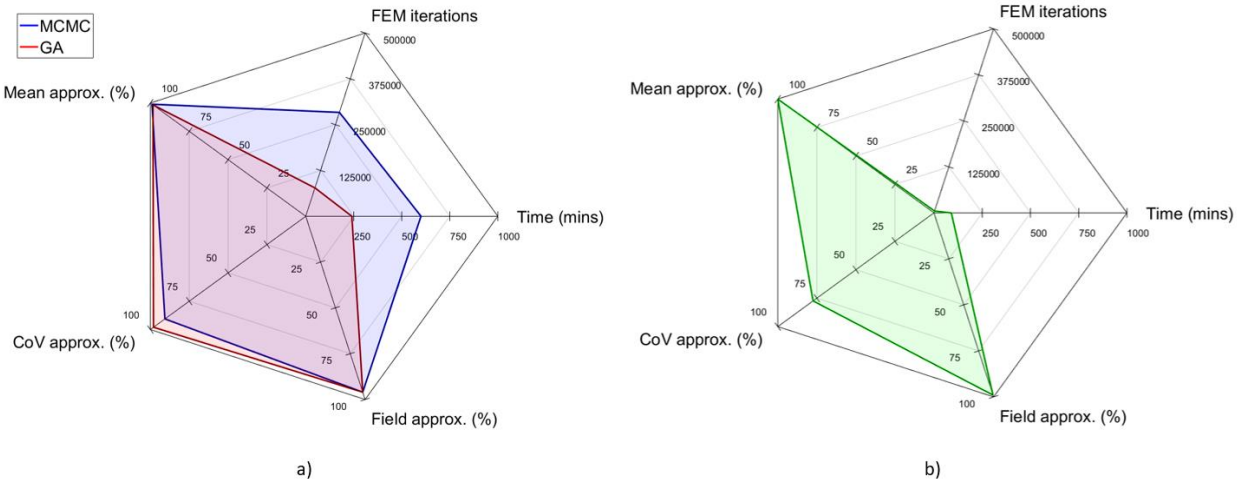

Fig. 78: Radar plots for method performance based on five criteria, a) MCMC and GA comparison, b) RPF

6. Reliability Analysis

The results taken from the inverse analysis are not singly defined, they are rather in the form of a distribution. Their stochastic nature requires the assessment of each method's reliability and therefore, a Monte Carlo analysis is applied.

6.1. Introduction

The main objective of the inverse analysis is collecting a sample of successful variable states, able to generate fields whose displacements are close to the monitored ones. In addition to evaluating method efficiency and the replication of the reference field, the benefits of compiling this sample can be extended in particular operations.

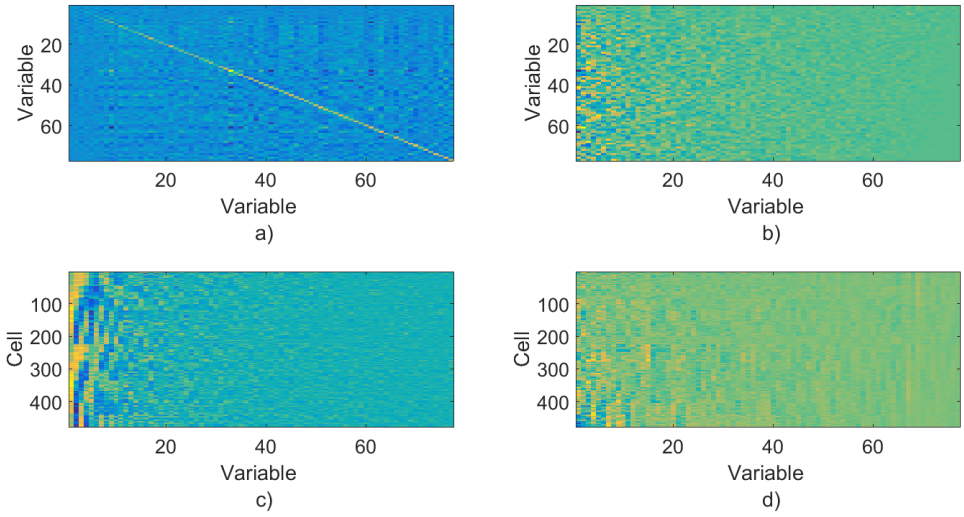

Fig. 79: Illustration of the: a) sample covariance matrix, b) decomposed matrix, c) decomposed heterogeneity matrix, d) and final field generation matrix in image plot

As presented in paragraph 4.8, the mean and covariance matrix of the variable sample can be utilized in order to compile fields which perform similarly to the reference one. The covariance matrix describes the relations between variables in a linear manner, and when applied along with the mean in a multivariate draw, it is potent to provide such a field. In order to replace the sampling with a univariate draw, a transformation is performed as shown in (paragraph 4.8). Then, *[f]* is converted to a field by applying the prior distribution parameters used in the inverse analysis. Notably, since strength was not properly simulated in the main analyses, the covariance matrix for every soil property is different, and thus, no interdependency between them is assumed.

The ability of generating fields based on the outcome of the inverse analysis allows for the evaluation of result reliability. Specifically, a Monte Carlo simulation is utilized, whose every realization includes sampling from a standard normal univariate distribution and transforming to a field as described above. The resulting field is applied in an FEM run and its error to the observations of the inverse analysis is calculated. After collecting the errors of all realizations, their respective Cumulative Distribution Function (CDF) can be plotted. In this graph, the error axis can be interpreted as the deviation from the measurements in a ratio form, while the other axis is the confidence of achieving a value equal to or lower than this spread. Furthermore, the error

axis can also be presented in measure form, as a displacement spread around the monitored values. Notably, there are 3 points of interest on this S-shaped curve:

- The confidence level at the inverse analysis tolerance: Originally this value should at minimum be 50%, and practically is always demonstrated to be more, since the error of every successful iteration in the inverse analysis is somewhat below the tolerance (however, exception of this rule are possible, i.e. due to solution domain shape, but seem to lie outside of scope).
- The spread for a confidence of 50%: This value represent the mean error, or simply the value splitting the population of sampled fields in half.
- The spread for a confidence of 95%: This is the largest error expected to be met when sampling from the inverse analysis posterior, since engineering practice generally assumes that a 0,95 probability for encountering a lower value generally renders it as an almost sure event (Hicks & Samy, 2002).

In addition, if the strength had also been approximated, then correlations between the effective strength and stiffness of the soil could have been made. Also, a Monte Carlo SRF analysis would be performed, including the generation of successful fields and testing them against failure. Likewise, the estimation of the safety factor for the slope construction would take advantage of the available monitoring data, tightly connecting it to real, in-situ measured displacements.

The following analyses aim into quantifying the dependability of the inverse analysis schemes. Moreover, they attempt to examine and confirm the applicability of the results in practice, in accordance to the engineering thinking administered in project design.

6.2. Markov Chain Monte Carlo reliability analysis

The first reliability analysis assesses the results of the MCMC Original Metropolis-Hastings method. Firstly, a Monte Carlo analysis consisted of 10.000 runs is performed, collecting the error sample. The random fields generated are based on the posterior distribution of the said inverse analysis, employing the covariance matrix and mean of its successful variable state sample. Then the CDF of the error is plotted, resulting in line of Fig. 81. Evidently, the curve is an almost symmetric S-Shaped CDF, meaning that the respective PDF is close to a Gaussian. Also, the three points of interest are marked, allowing for a quick examination of the result. The tolerance of the inverse analysis (*0,2‰ or +/-0,02mm spread*) has a confidence level of 64%, meaning that this is the chance that a generated field will produce an error within tolerance. The 50% confidence interval is met at *0,19‰ or +/-0,019mm spread*, implying that this is the mean error of the inverse analysis. Also, the desired confidence level of 95% is achieved at a *0,23‰ or +/-0,023mm spread*, meaning that the error of any simulation can be safely expected to be lower than this value. Extending this thought to the operation of the inverse analysis, once the Markov Chain had achieved an error below the tolerance, the deviation of most rejected realizations did not surpass *0,23‰,* exhibiting that sampling was not wandering far from the posterior. Notably, the greatest achievable value appears to be *0,3‰ or +/-0,03mm spread*, where 100% confidence is reached.

In order to conceive the competency of the MCMC method, the confidence curves for the posterior and the prior of the inverse analysis are plotted (Fig. 82). It was already known that the prior produces errors that are greater than the tolerance by two orders of magnitude (paragraph 5.1) and now the Monte Carlo analysis is able to illustrate it in a confidence curve. Comparing the two curves, the efficiency of the inverse analysis is highlighted. The method was able to guide the sampling, vastly reducing the error to the monitored displacements. Thus, the confidence curve of the posterior seems as a sharp increase at much lower values compared to that of the prior.

Considering the 95% confidence level, which acts as a reference value due to its usual adoption in engineering practice, the inverse analysis provides major benefits. The prior achieves this reliability level at a tolerance 200 times higher than the one of the posterior (*5% or +/-5mm spread*). Therefore, the displacements calculated when utilizing the prior, which was supposed to be selected based on the designer's view, would be expected to differ from actual displacements by up to 5mm.This value drops dramatically to 0,023mm (outside of the geo-engineering scope) when employing an inverse analysis with monitored displacements, emphasizing the method's competency in allocating the reference parameter fields.

Similar conclusions can be drawn for the comparison between the prior and the posterior of the Genetic Algorithm analysis.

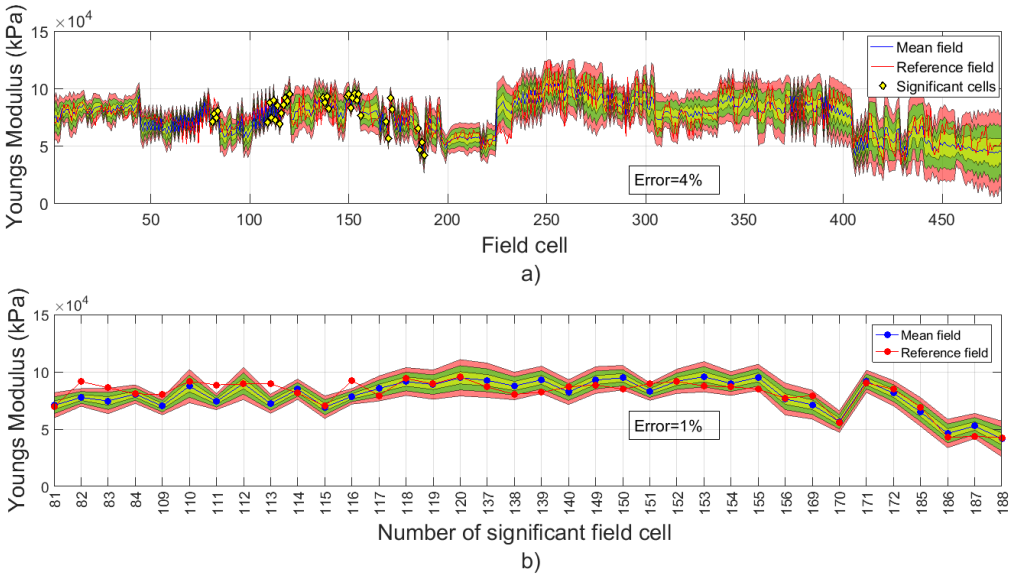

Fig. 80: Mean field of the MCMC posterior with the 68%, 95% and 99.6% confidence bands compared to the reference field for: a) all cells, b) cells of interest – Young's Modulus

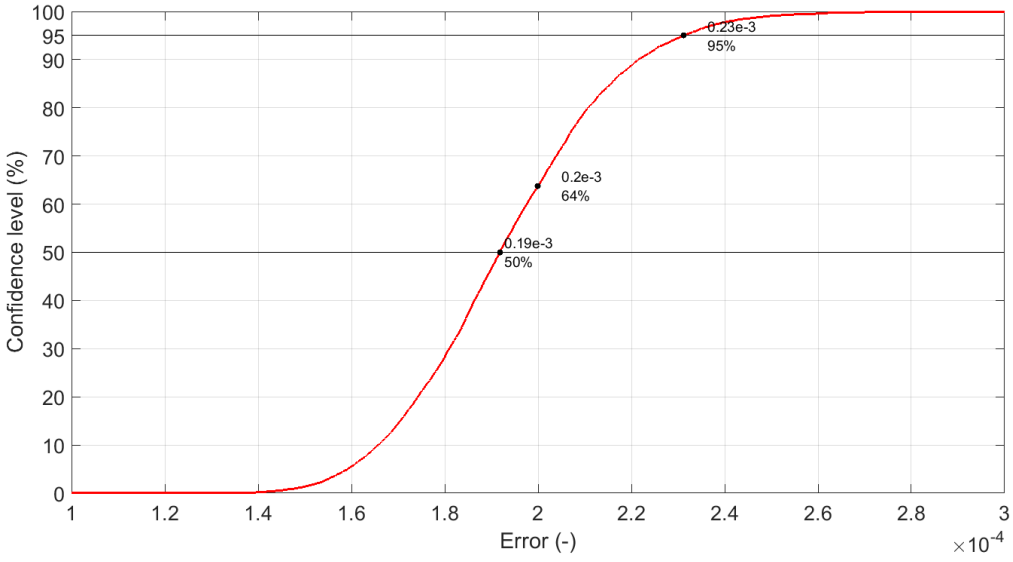

Fig. 81: Plot of confidence curve for the MCMC posterior to error with points of interest

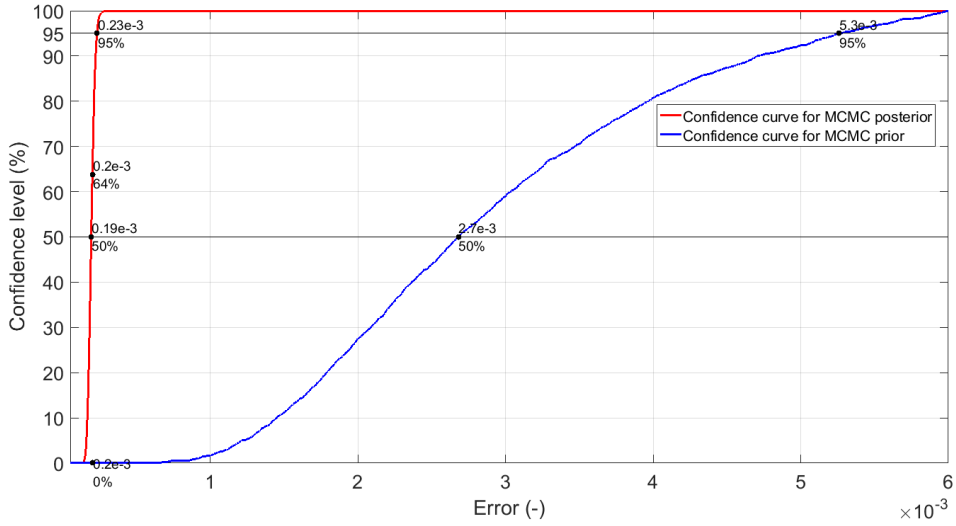

Fig. 82: Comparison between the confidence curves for the MCMC posterior and prior to error with points of interest

6.3. Markov Chain Monte Carlo – Genetic Algorithm reliability curve comparison

Following, the same Monte Carlo scheme is applied on the results of the GA inverse analysis. While the reliability curve seems proper when evaluated independently, a significant difference is highlighted upon comparison to the curve of the MCMC (Fig. 83) The GA confidence curve is steeper than the one for MCMC, a feature accredited to the lower variance of the GA sample. Specifically, a reduced variance means that sampled fields tend to be similar and so the errors calculated in the Monte Carlo analysis of the GA do not deviate as much as in the MCMC.

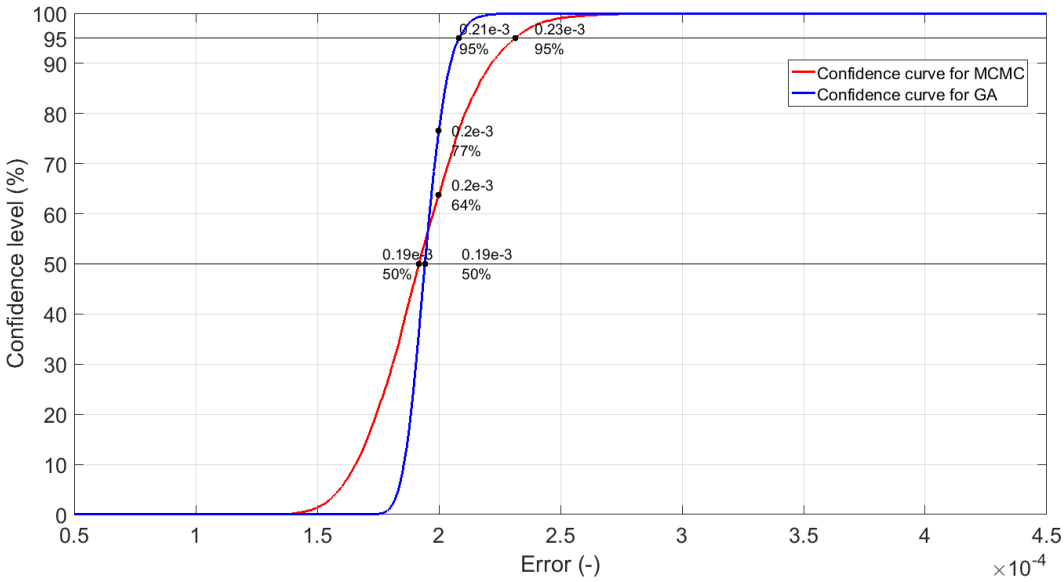

Fig. 83: Comparison of the MCMC and GA confidence curves

Evidently, the GA solution provides a much thinner 95% confidence band in the field plot than the MCMC (Fig. 84).The thick band of the MCMC means that there is higher chance of approximating the reference field, which is almost fully enveloped, than in the GA, where the reference is mostly outside of the band. On the other

hand, the low standard deviation of each field cell means that the GA produces a dependable simulated field, which is both close to the reference and strictly-defined. Ultimately, this behavior draws a crucial deduction; the result of the GA inverse analysis tends to degenerate into a deterministic one. Consequently, applying a reliability evaluation is no longer relevant.

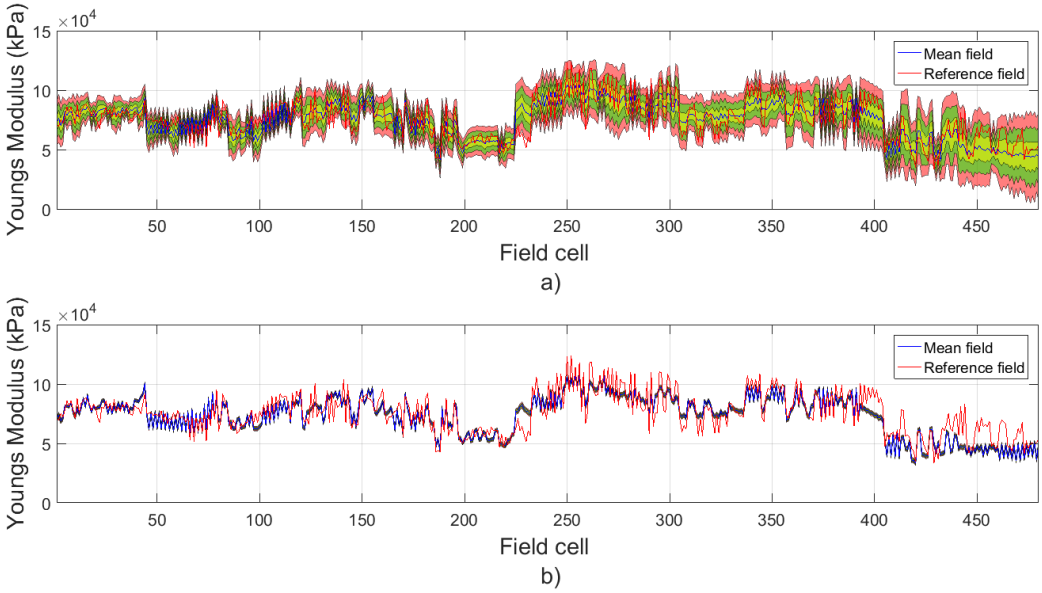

Fig. 84: Comparison of the MCMC and GA simulated fields (sample field means of field cells)

The reason for these contrasting behaviors lies within the philosophy of the methods. In order to allow for a thorough clarification, a simple optimization problem is set up. The MCMC and GA are utilized in an attempt to identify the minimum of a parabola in a bi-variable domain. Both methods are initialized with the same prior distribution, unequal to the target, and both use an error tolerance of 1, defined as the difference between the function value of the variable state and the minimum. Also, the MCMC is ran for 1.000 realizations, while the GA employs a population of 100 individuals, requires a solution sample of 1.000 states and thus is simulated for 35 generations (method parameters taken by experimenting for a representative illustration). The analysis provides overview of method operation in an illustratable domain of two variables.

Firstly, result visualization takes place in 3D (Fig. 85), exhibiting the development of every scheme in regard to the examined surface. Secondly, result data is illustrated in a 2D contour plot (Fig. 86),which represents the variable plane projection of the 3D counterpart and enables for a clearer view on method movement. Now the difference in sample collection between the schemes is evident.

The MCMC is attracted to the minimum and thus follows a sampling path from the initial point to it. When within tolerance, the Markov criterion which enforces the said attraction is disabled, and the chain starts randomly wandering. As a result, a diverse sample is collected from within the tolerable zone. Moreover, it should be noted that the mean of the successful sample may provide an error lower than any realization.

Contradictorily, the GA attempts to "herd" the individuals toward the solution. In the case where the prior and the target do not match, then the population translated on the 2D plane, while also shrinking in area. Eventually, it reaches the tolerance and tends to collapse to a single variable state, thus producing a sample of much lower variance. A point of importance is that the GA continues to be attracted by the optimum even when having achieved an error below tolerance (evident in a more complex inverse problem). Also, in the examined case the mean of the sample is expected to give the mean error.

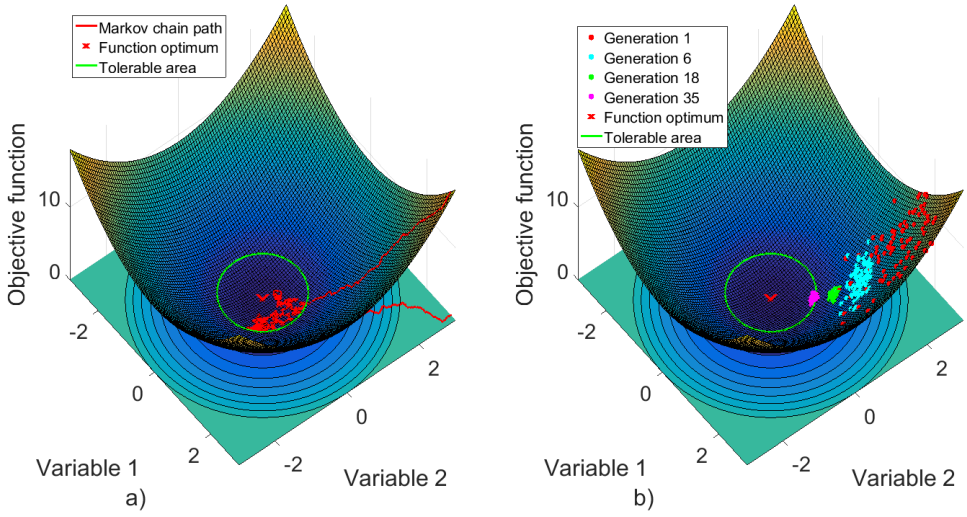

Fig. 85: 3D illustration of method operation in example inverse problem: a) MCMC path and variable plane projection, b) GA population convergence

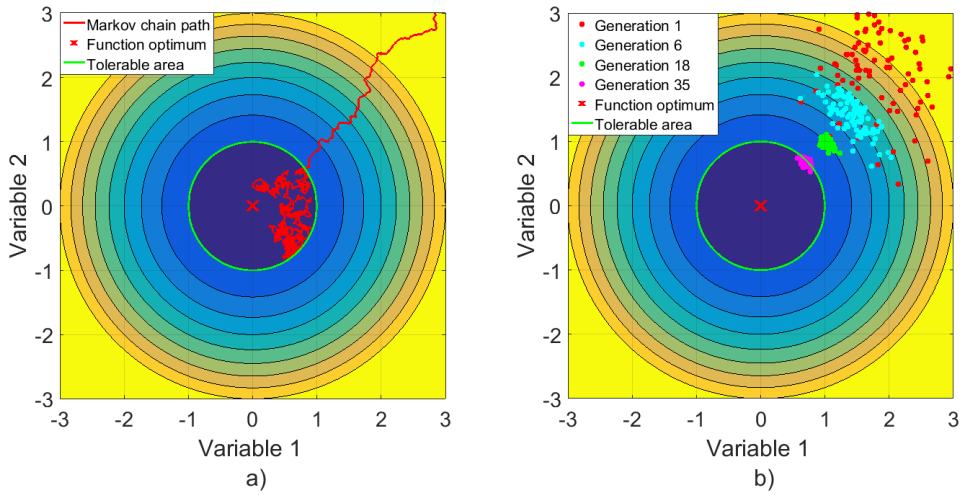

Fig. 86: 2D illustration of method operation in example inverse problem: a) MCMC path, b) GA population convergence, both in variable plane

Although the research was brought to the conclusion that the GA results are not dependable for a Monte Carlo analysis, measures can be taken in order to condition them. Hence, this inverse scheme can again attain a strong reliability-oriented aspect. The first step is increasing the variance of the sample. In order to do so, an inflation must be applied, and an inflation factor needs to be defined, to account for the fact that the spread in the GA individuals has degenerated, since the inverse problem is not able to provide any more conditions for the sample. After a experimentation, the constant is picked equal to the ratio of the mean of the simulated field (mean of every cell) and the prior standard deviation (Eq. 6.2). Notably, only the diagonal of the covariance matrix is scaled, so that the interdependency between variables is left intact. Then, the new curve of the GA is calculated (Fig. 87). Admittedly, it approaches the MCMC curve, meaning that the sample can be treated again as a stochastic result. However, the reliability curve is translated to higher error values.

$$C = \mathrm{cov}(sample) \tag{6.1}$$

$$C_{i,i} = C_{i,i} * \frac{\sigma_{Simulated\_field}}{\sigma_{Prior}} \tag{6.2}$$

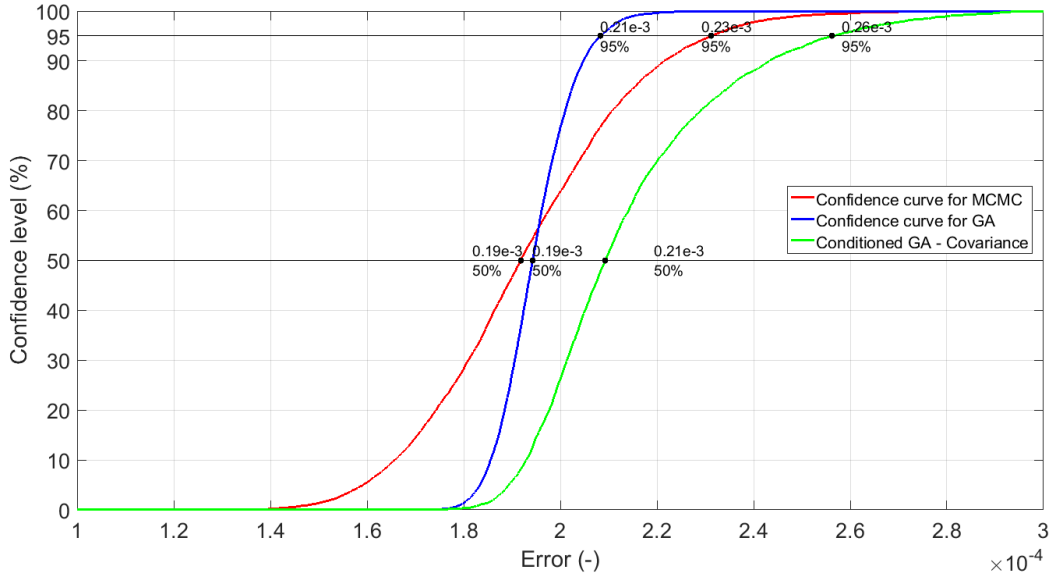

Fig. 87: Comparison of the MCMC, original GA and GA with a conditioned covariance confidence curves

A critical point in understanding the conditioning procedure prevails in the definition of the sample mean. It should be clarified that the sample mean does not always coincide with the variable state producing the mean error, a predicament actually met in this case. The mean state provides an excellent approach of the reference and so, any other field sampled is more likely to have a higher error. As a result, when increasing the variance of the sample, the curve does not rotate around the error of 50% confidence level. Specifically, as variable states further from the mean are sampled, the mean error increases and the curve inherently translates. This is due to the definition of error; as it is composed of a geometrical average, when displacements are both below and above the mean value, they still produce a higher error.

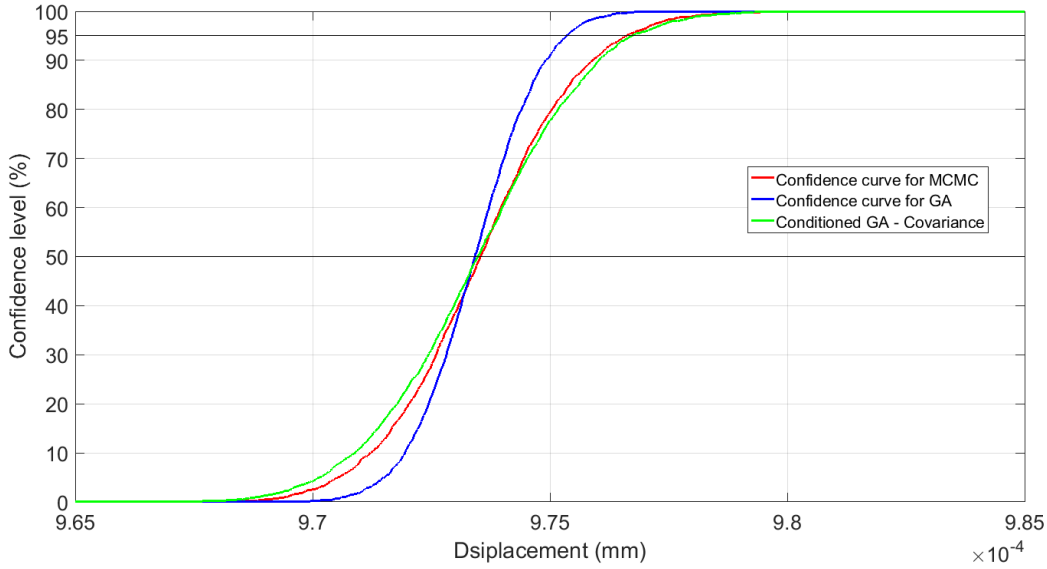

Fig. 88: Comparison of the MCMC, original GA and GA with a conditioned covariance confidence curves for displacement values (displacement at monitoring point 2 during the 3rd excavation stage)

Insight can be gained when performing a similar analysis but for a displacement value instead of the total error. As a general rule, now the conditioned GA curve rotates around the mean displacement, because the analysis only investigates on the effect of stiffness. Therefore a linear relationship is established between variables and displacement, meaning that the mean variable state coincides with the state of mean error. In the case where the mean errors of the MCMC and GA are close, then the new GA curve efficiently approaches that of the MCMC (Fig. 88). In contrast, the conditioned GA curve only follows the MCMC curvature when the GA and MCMC errors significantly differ (Fig. 89).

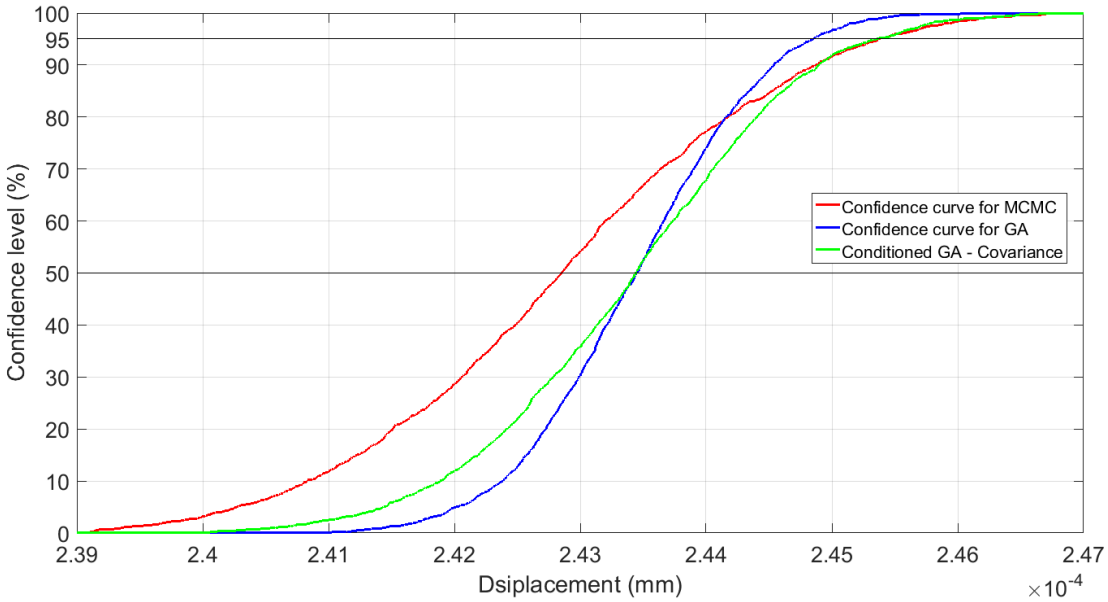

Fig. 89Comparison of the MCMC, original GA and GA with a conditioned covariance confidence curves for displacement values (displacement at monitoring point 10 during the 3rd excavation stage)

However, it is expected that in the curve for total error the rotation point will stay almost fixed at 50% (accounting for the said inherent translation) when sampling around the variable state which provides the mean error of the curve. So, the sample mean is also put under conditioning. For this cause, the mean error of the original GA curve is identified and then all variable states adequately close to it are selected. This new sample (for ease named "T") undergoes the same handling as the original (Eq. 6.3 and 6.4), employing its mean and the eigen-decomposition of its covariance. So, the random univariate pick once used is now interpreted to an equivalent multivariate of "T" (Eq. 6.5) before put in the procedure of (Eq. 4.10). As a result of experimenting, the best fit is identified for a normalizing constant 0.9 times the original value proposed (Eq. 6.2).

$$C_T = \Phi_T \Lambda_T \Phi_T^T \tag{6.3}$$

$$A_T = \Phi_T \Lambda_T^{1/2} \tag{6.4}$$

$$f = V\mu_{GA\_sample} + VA(\mu_T + A_T\xi) \tag{6.5}$$

(Where ξ is a univariate standard draw.)

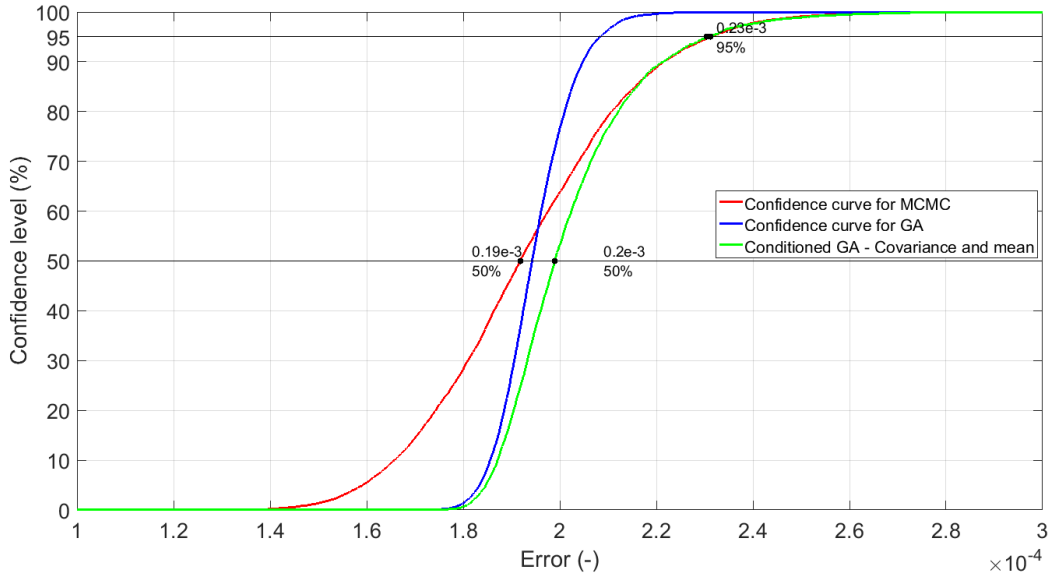

Fig. 90: Comparison of the MCMC, original GA and GA with a conditioned covariance and mean confidence curves

Subsequently, the problem is fixed (Fig. 90). While initially the curves do not match, the 50% confidence points are now sufficiently close. Following, the curve of the conditioned GA converges to the MCMC one, totally matching it after the confidence level of 80%. Although this analysis may lack the exactly proper mathematical background, it is still able to improve the reliability of the GA results, hence composing an interesting and insightful experiment. Ultimately, the results of the GA can dismiss their deterministic nature when subjected to proper treatment, therefore exhibiting a meaningful reliability performance.

6.4. Regularized Particle Filter reliability analysis

Finally, the confidence analysis of the RPF results is put under inspection. Following the same Monte Carlo routine, the S-shaped curve of (Fig. 92) is formed and some strong deductions are drawn. Notably, the sample utilized is the entire final ensemble, as its mean matches the target, and not just the successful particles. The first observation made is that the curve is relatively wide (the confidence level of 95% is achieved at an error almost 1,4 times greater than the mean), meaning that there is a strong variation among generated fields, accredited to the high standard deviation of the results (Fig. 91). Secondly, predictions about the error of the 50% confidence level were ambiguous. While this value could have been achieved at the error of *0,17‰*, which is the error of the mean field, instead the mean error of the ensemble appears. Extending this result, ensemble convergence is speculated. While the mean of the ensemble provides an excellent field fit with a low error, it is virtually never approximated. This behavior is attributed to extreme ensemble variance. Illustratively, particles are randomly scattered in the variable domain with significant distances in between, as every iteration of the RPF attempts to bring them closer to the solution. However, only the mean of the ensemble is actually within acceptable tolerance to it.

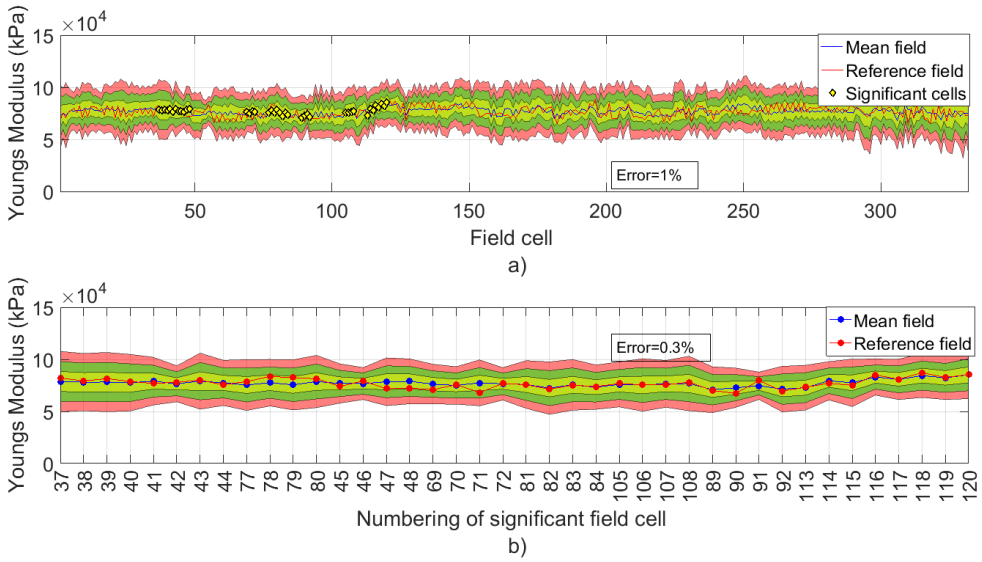

Fig. 91: Mean field of the RPF posterior with the 68%, 95% and 99.6% confidence bands compared to the reference field for: a) all cells, b) cells of interest – Young's Modulus

At this point, the opposite behavior of the GA and RPF is highlighted. While the latter produces a more variable sample able to contain the reference field (Fig. 91), it actually rarely achieves to approximate it, leading to a wide reliability curve. On the other hand, the GA sample does not encompass the reference at all (Fig. 84, b), a fact that should imply that the reference field is never approached. Nevertheless, the mean field is much more dependably defined, seemingly becoming a deterministic result of the analysis. Since the mean field gives off an acceptable error, the GA is able to produce less uncertain and sufficient results.

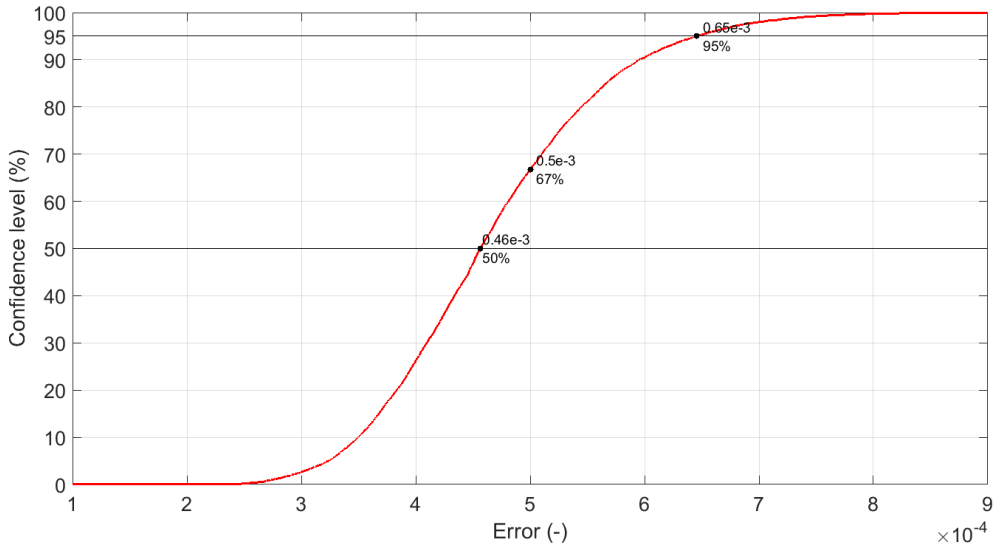

Fig. 92: Plot of confidence curve for the RPF posterior to error with points of interest

While the RPF exhibits a unique behavior in the reliability analysis, its benefits cannot be overlooked. Specifically, the method is indeed able to vastly reduce the uncertainty of the design utilizing the prior distribution by narrowing the spread between computed and actual displacements (Fig. 93). So, this comparison poses as evidence of the RPF's competency in allocating the reference parameter values.

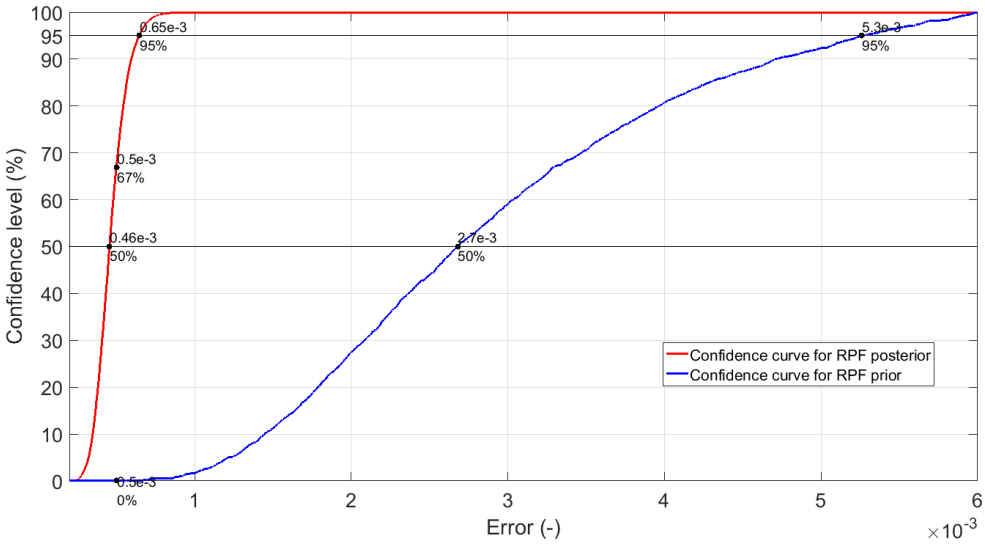

Fig. 93: Comparison between the confidence curves for the RPF posterior and prior to error with points of interest

6.5. Prediction of next stage response

This analysis aims into expanding the results of an inverse analysis and exhibiting their potential utilization in a slope construction design. The main notion is that monitoring the displacements during the excavation and applying an inverse analysis on the data, can provide an updated prediction on the performance of the design in the following stage.

Firstly, an MCMC inverse analysis is performed with every factor exactly the same as before. The only difference is that now only three out of the four construction stages are simulated. The collected successful variable state sample is employed in a Monte Carlo analysis in order to assess the confidence curve for the three-stage approach. Furthermore, an extra reliability analysis is executed, which takes into account only the error of the last stage for every generated field.

Since the inverse analysis applies the same tolerance for both the main MCMC and the three-staged variations, as well as because their resulting standard deviations appear to be similar, their confidence curves are almost identical (Fig. 81 and Fig. 94). On the other hand, the second analysis has not taken into account the last stage, meaning that results are not regulated to it. Simulating for less stages means that the process of identifying the reference field has less boundary conditions and so its accuracy decreases. This fact can justify why the confidence curve of the final stage prediction is transferred to much higher tolerance values, as well as why it becomes significantly wider, meaning that uncertainty is induced.

This outcome can have a significant impact on the continuation of the slope construction. The practically applied limit of 95% confidence is met at almost *2,5‰* error or *0,25mm*.This value expresses the largest anticipated deviation between the displacements expected to be met in construction and those calculated upon designing with the mean field of the three-staged inverse analysis posterior. In case this error reflects unacceptable displacements, the reevaluation of the last stage design is mandatory.

Finally, the prediction confidence curve of the MCMC is compared to the one established by the prior (Fig. 96). Evidently, applying the prior in the design induces a greater uncertainty. Firstly, the curve is shifted towards a higher error scale, implying that the mean field deviates from the performance of the reference, as expected. Secondly, the shape is wider, meaning that the fields generated for the Monte Carlo analysis are more diverse.

Accordingly, the 95% confidence level is available for a vastly larger error, enhancing the doubt on the prediction of the next stage.

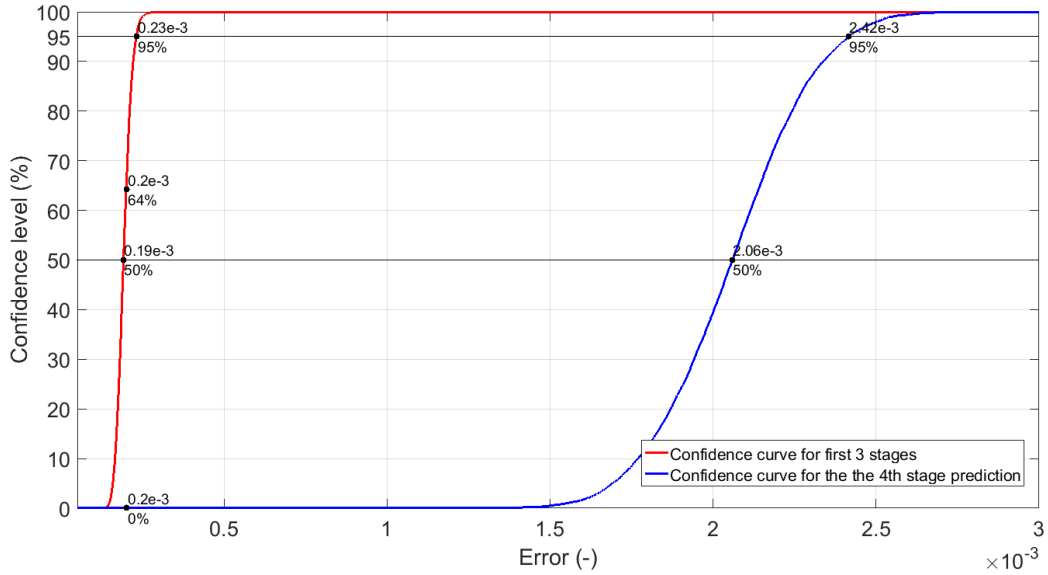

Fig. 94: Plot of confidence curves for the MCMC (3 stage inverse analysis) posterior and prediction of 4th stage to error with points of interest

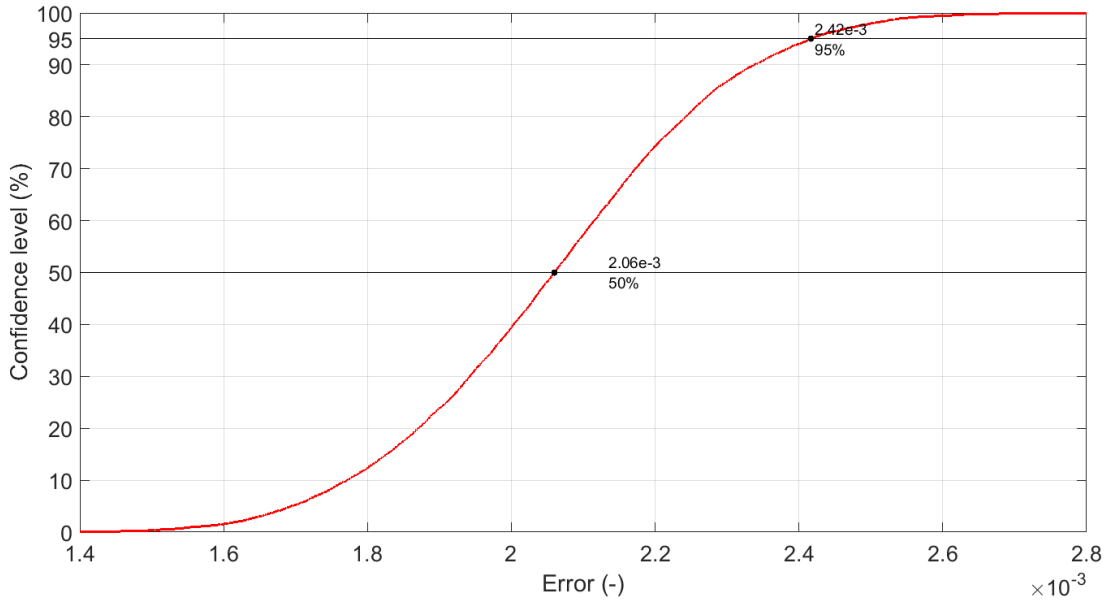

Fig. 95: Confidence curve of the 4th stage prediction for the MCMC with points of interest

The transition from the prior confidence curve to the ones of the inverse analyses is an immense step towards reducing the uncertainty of the design. The back-calculation is able to diminish the error at which the reliability of 95% is attained. from *10mm* to merely *0,25mm*. Consequently, the merits of a dynamic project plan, which includes feedback loops between design and in-situ measurements, are highlighted. Monitoring data is utilized in order to allocate the effective parameter values of the soil, hence complementing the design with measurements that reflect actual soil behavior. Therefore, the maximum error spread is significantly narrowed, enhancing project reliability and allowing for design optimization.

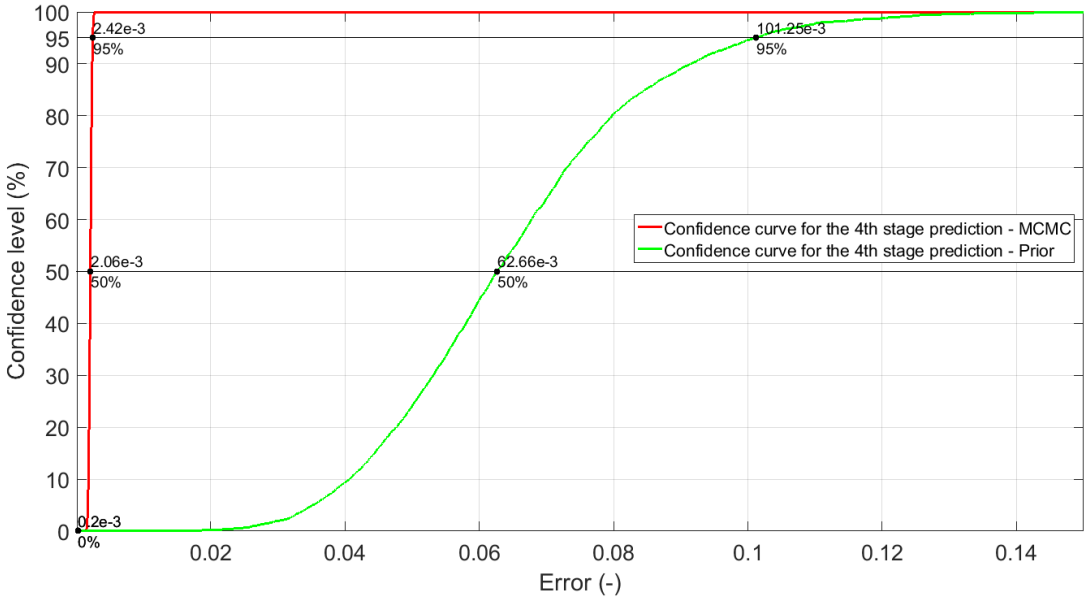

Fig. 96: Comparison between the confidence curves of the 4th stage prediction for the MCMC and prior with points of interest

6.6. Conclusion

In this chapter, a stochastic investigation was applied on the results of the inverse analysis. A Monte Carlo framework was developed in order to estimate the reliability curve of each inverse method posterior, thus connecting measurement deviations to confidence levels. After comparing the curve of each approach to the respective one for the prior distribution, the efficiency of the methods along with the benefits of an inverse analysis are highlighted. In addition, a comparison between the line of the MCMC and the GA is performed, followed by a conditioning of the latter, allowing for the identification of the merits, shortcomings and possibilities inherent in their outcomes. Finally, the reliability analysis is employed on a practice-oriented topic; what is the uncertainty in foreseeing the development of the construction? After executing an inverse analysis only on the first excavation phases, the error associated to the 95% confidence level in the prediction of the next stage is evaluated.

7. Conclusions

Essentially, the mission of this thesis was to provide a strong inverse analysis framework enhancing its potency in real-life applications, while also accounting for the concept of reliability driven design, adopted in modern geotechnical engineering. Over the course of the report, the set research goals have been administered in a complete and critical manner. Chapter “Inverse Analysis” dealt with the fruitful implementation of the inverse methods, as well as judging their efficiency and analyzing their benefits and shortcomings. Section “Reliability Analysis” was occupied with introducing the results into a stochastic investigation, thus evaluating their reliability and addressing the applicability of the inverse analysis in practice.

The application of the three inverse analysis methods on the staged slope excavation RFEM model was thoroughly described. After collecting and comparing their outcomes, the GA appears as the most attractive approach, a deduction standing on two pillars. Firstly, this method performs equally effectively with the MCMC when identifying the target distribution and approximating the reference field, whereas it is more efficient, since its time and calculation tolls are significantly less. Besides, the comparison between the GA and the RPF is not representative, since the latter operates on a fundamentally different case. Secondly, the GA is potent to provide reliable results. Arguably, the outcome of the inverse analysis is almost deterministic, and so a employing it in a reliability analysis seems rather unorthodox and pointless. However, it was experimentally shown that when converted properly, the GA sample offers a reliability curve similar to that of the MCMC, signaling its applicability in a stochastic analysis. Ultimately, the GA is potent to achieve equivalent reliability results in a more computationally-efficient manner.

Moreover, research on this topic has encountered serious drawbacks. Firstly, strength cannot be properly approximated, since the Tresca (or generally Mohr-Coulomb) soil constitutive model is unsuitable for this operation. Still, a new FEM model was set up in order to capture soil yielding in a more appropriate manner, allowing for the accomplishment of the MCMC. Additionally, a GA scheme that exploited the advantages of the evolutionary concept was successful in grasping the strength target distribution. Secondly, the employment of the RPF on a shorter observation timeline was not feasible, leading to its application on a different case. Specifically, a new mesh able to simulate more construction stages was introduced, orchestrating a simulation well adapted to the RPF process. While such obstacles hinder the goals of the thesis, their causes have been explained and measures have been taken for their mitigation. In particular, the conditions leading to method attainment have been provided, along with thorough examples on their implementation.

Considering the optimization of the presented methods, a hybrid model is proposed. The GA exhibited the greatest performance when attempting to simulate with a target distribution different than the prior. On the other hand, the RPF was operable only in a regulated mesh and demanded matching prior and target. Nevertheless, it was able to produce the most qualified field approximation among all methods. As a result, a serial scheme combining these methods could prove to be the most viable approach to the inverse problem so far. Essentially, the GA will take up the approximation of the target distribution. Following, its posterior is going to be utilized as the prior distribution of the RPF, which will aim into improving the reproduction of the reference field. However, the behavior of the two methods varies widely in the reliability analysis, and so deeper examination of the combination notion is strongly advised.

Summing up, this report successfully addresses a variety of points regarding the fulfillment of inverse analysis. After providing a full walkthrough of method implementation and comparing their outcomes, the most competent inverse analysis scheme is chosen. Definitely, the selection criteria are oriented towards algorithm efficiency along with minimization of uncertainty. Ultimately, modern geotechnical engineering demands a reliability-driven design; integrating the inverse analysis poses as a competent way to reduce uncertainty and promotes an adaptable construction, while exploiting actual soil response.

8. Recommendations for Future Research

Even though this thesis dealt with some major issues in the field of inverse analysis, there is still an abundance of unanswered questions. In an effort to inspire future research, a list of suggestions is composed.

- The adoption of an advanced constitutive model would enable the identification of the strength field. Generally, any model that connects stiffness to mobilized strength would account for the impact of the strength parameter in the inverse analysis even before the failure criterion has been met. In this case, the possibility that multiple parameter combinations being successful exists, hence expanding the solution domain. Accordingly, allocating strength would mean that a correlation between effective strength and stiffness would be possible. In addition, knowledge of the strength field would allow for a stochastic slope stability calculation connected to in-situ observations. For this FEM simulation, the Hardening Soil model would be ideal as a first attempt.

- The inverse analysis should be adapted to deal with measurement noise. In practice, inclinometer readings always include random fluctuations and observations are then considered in a stochastic manner. The analysis must be able converge to the target distributions while neglecting these perturbations in order to emphasize its applicability in actual design.

- The scale of fluctuation should be treated as a variable of the inverse analysis. While this thesis deals with a case where the heterogeneity pattern of the soil is considered known (e.g. through experience), a full analysis should be able to identify it. Essentially, a database of heterogeneity eigenvector ensembles can be collected beforehand for a range of scales of fluctuation, so that calculation costs are preserved in every inverse analysis iteration. Additionally, in this occasion an attempt for method combination can prove to be an excellent course of action.

- The accomplishment of the inverse analysis should also be examined for various types of prior and target distributions. The investigation of method operation under these circumstances would provide advantageous insight on the proper application of the analysis in real-life cases.

- A serial combination of the GA and the RPF may optimize the final result. While the GA is efficient in identifying the target distribution, the RPF is the greatest solution when locally modelling soil heterogeneity. Therefore, utilizing the advantages of each approach can allow for a complete and largely more competent inverse analysis.

- A deeper examination of the monitoring plan should be performed. The impact of different monitoring points should be studied and their combination providing the most efficient analysis in terms of computational toll and accuracy should be identified.

Bibliography

Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing, 50*(2), 174-188.

Bishop, G., & Welch, G. (2001). An introduction to the Kalman filter. *Proc of SIGGRAPH, Course 8*.

Bracewell, R. N., & Bracewell, R. N. (1986). *The Fourier transform and its applications* (Vol. 31999): McGraw-Hill New York.

Brinkgreve, R., & Vermeer, P. (1998). Plaxis manual. *Version 7*, 5.1-5.18.

Carr, J. (2014). An introduction to genetic algorithms. *Senior Project*, 1-40.

Chib, S., & Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician, 49*(4), 327-335.

Coley, D. A. (1999). *An introduction to genetic algorithms for scientists and engineers*, Singapore, World Scientific Publishing Co.

Doucet, A., & Johansen, A. M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of nonlinear filtering, 12* (656).

Evensen, G. (2003). The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics, 53*(4), 343-367.

Evensen, G. (2009). *Data assimilation : the ensemble Kalman filter, Springer Science and Business Media*

Fenton, G. A., & Griffiths, D. (2007). Review of probability theory, random variables, and random fields, *Probabilistic methods in geotechnical engineering* (pp. 1-69): Springer.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1996). *Markov chain Monte Carlo in practice*. London :: Chapman & Hall.

Glickman, M. E., & Van Dyk, D. A. (2007). Basic bayesian methods. *Topics in Biostatistics*, 319-338.

Griffiths, D., & Fenton, G. A. (2004). Probabilistic slope stability analysis by finite elements. *Journal of Geotechnical and Geoenvironmental Engineering, 130*(5), 507-518.

Groetsch, C. W. (1999). *Inverse problems: activities for undergraduates*: Cambridge University Press.

Haupt, R. L., & Haupt, S. E. (2004). The continuous genetic algorithm. *Practical Genetic Algorithms, Second Edition*, 51-66, John Wiley & Sons, Inc.

Hicks, M. A., & Samy, K. (2002). Reliability-based characteristic values: a stochastic approach to Eurocode 7. *Ground Engineering, 35* (12), 30-34.

Jeffreys, H. (1931). Scientific inference. *Cambridge Univer*.

Ledesma, A. (2014). Geotechnical back-analysis using a maxi-mum likelihood approach, *Stochastic Analysis and Inverse Modelling*, Hicks M. A., Jommi C., 209-238.

Levasseur, S., Malecot, Y., Boulon, M., & Flavigny, E. (2010). Statistical inverse analysis based on genetic algorithm and principal component analysis: applications to excavation problems and pressuremeter tests. *International journal for numerical and analytical methods in geomechanics, 34* (5), 471-491.

Lloret-Cabot, M., Fenton, G. A., & Hicks, M. A. (2014). On the estimation of scale of fluctuation in geostatistics. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 8*(2), 129-140.

Mitchell, M. (1996). *An introduction to genetic algorithms*. Cambridge, Mass, MIT Press.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology, 47*(1), 90-100.

Noordam, A. F. (2016). *Numerical Investigation of Inverse Analyses Using Hydraulic Measurements of Embankments,* Master thesis, TU Delft, Delft.

Orhan, E. (2012). Particle filtering. *Center for Neural Science, University of Rochester, Rochester, NY, 8*(11).

Papaioannou, I., Betz, W., Zwirglmaier, K., & Straub, D. (2015). MCMC algorithms for subset simulation. *Probabilistic Engineering Mechanics, 41*, 89-103.

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

Schofield, A., & Wroth, P. (1968). *Critical state soil mechanics*, 310: McGraw-Hill London.

Smith. (2002). A tutorial on principal components analysis. *Cornell University, USA*.

Smith, I. M., Griffiths, D. V., & Margetts, L. (2013). *Programming the Finite Element Method*

Soubra, A.-H., & Bastidas-Arteaga, E. (2014). Advanced reliability analysis methods, *ALERT Doctoral School 2014: Stochastic Analysis and Inverse Modelling*, Michael A. Hicks; Cristina Jommi, 79-94.

Terejanu, G. A. (2008). Extended kalman filter tutorial. *Department of Computer Science and Engineering, University at Buffalo*.

van den Eijnden, A., & Hicks, M. (2017). Efficient subset simulation for evaluating the modes of improbable slope failure. *Computers and Geotechnics, 88*, 267-280.

Vanmarcke, E. (2010). *Random fields: analysis and synthesis*: World Scientific.

Vardon, P., Liu, K., & Hicks, M. (2016). Reduction of slope stability uncertainty based on hydraulic measurement via inverse analysis. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 10*(3), 223-240.

Wikle, C. K., & Berliner, L. M. (2007). A Bayesian tutorial for data assimilation. *Physica D: Nonlinear Phenomena, 230*(1), 1-16.

Wood, D. M. (1990). *Soil behaviour and critical state soil mechanics*: Cambridge university press.