Cyprus
University of
Technology

Faculty of Engineering
and Technology

Doctoral Dissertation

REDUCING STATISTICAL UNCERTAINTY IN GEOTECHNICAL ENGINEERING DESIGN RELYING ON TARGETED FIELD INVESTIGATION: A RANDOM FIELD APPROACH

Panagiotis Christodoulou

Limassol, May 2020

CYPRUS UNIVERSITY OF TECHNOLOGY

FACULTY OF ENGINEERING AND TECHNOLOGY

DEPARTMENT OF CIVIL ENGINEERING AND GEOMATICS

Doctoral Dissertation

REDUCING STATISTICAL UNCERTAINTY IN GEOTECHNICAL ENGINEERING DESIGN RELYING ON TARGETED FIELD INVESTIGATION: A RANDOM FIELD APPROACH

Panagiotis Christodoulou

Limassol, May 2020

Approval Form

Doctoral Dissertation

REDUCING STATISTICAL UNCERTAINTY IN GEOTECHNICAL ENGINEERING DESIGN RELYING ON TARGETED FIELD INVESTIGATION: A RANDOM FIELD APPROACH

Presented by

Panagiotis Christodoulou

Supervisor: Dr. Lysandros Pantelidis, Assistant Professor, Department of Civil Engineering and Geomatics, Cyprus University of Technology

Signature ____

Member of the committee: Dr. Christos Anagnostopoulos, Professor, School of Civil Engineering, Aristotle University of Thessaloniki

Signature ____

Member of the committee: Dr. Evangelos Akylas, Associate Professor, Department of Civil Engineering and Geomatics, Cyprus University of Technology

Signature ____

Cyprus University of Technology

Limassol, May 2020

Copyrights

Acknowledgements

This thesis would not have been possible without the assistance, guidance and encouragement of many people. First and foremost, I heartily wish to thank my supervisor, Dr. Lysandros Pantelidis for his supervision, advice and guidance. His passion, attitude and dedication to scientific research inspired me at every step of this research journey. I have been extremely lucky to have such a great supervisor who cared about me, my work, and who support me in all that I do.

The advice and comments given by Professor Vaughan Griffiths of Department of Civil and Environmental Engineering of Colorado School of Mines, and by Professor Phaedon C. Kyriakidis of the Department of Civil Engineering and Geomatics of Cyprus University of Technology, are acknowledged gratefully. I'm also grateful to all the staff of the Department of Civil Engineering and Geomatics of Cyprus University of Technology for their assistance.

A special thanks to Dr. Elias Gravanis for his encouragement and advice at many stages of this journey; time stands still when we work together. His positive outlook and confidence in my research inspired me and gave me confidence to continue.

I must express my thanks and gratitude to Dr. Kyriakos Kyrou and Dr. Christakis Onisiphorou, for their valuable support, comments and suggestions. I also wish to thank Skyramont Quarries Ltd, for giving me the opportunity to perform in-situ tests for site characterization.

Thanks also go to Antonia, my wife, for her continued understanding, patience, love, support and encouragement throughout these years.

Finally, I owe a great deal of thanks and respect to my parents, Anastasis and Margarita Christodoulou, to my sister Elena and to my brothers Nektarios, Fanos and George for their unconditional love, support and encouragement throughout my life and all these years of study.

Chapter Abstract

Chapter 1 – A comparative assessment of the methods-of-moments for estimating the correlation length of one-dimensional random fields

Due to geological processes, soil properties vary in vertical and horizontal directions which is defined as inherent uncertainty of soil. This type of uncertainty may seriously affect the reliability of all geotechnical structures. The inherent uncertainty of soil properties is modelled as a random field, which is described by the mean, standard deviation and correlation length (also known as scale of fluctuation, θ) of soil properties. In this Chapter, the effectiveness of eight methods-of-moments for estimating the correlation length θ is investigated. This is done by generating samples of one-dimensional random fields for pre-defined values of the correlation length, which is then estimated by the different methods. For each method, the influence of the sampling domain length D and sampling interval d_x on the estimation of θ were investigated, and the results are quantified in the form of errors over the parameter space, defined by the dimensionless ratios D/θ and θ/d_x. Through the present analysis, one is able to assess the reliability of θ estimations obtained in practice, by mapping the conditions of any given experiment i.e., sampling domain, interval and estimated correlation length, onto the parameter space. The expected error associated with each method used is also quantified. Through this analysis a comparative assessment of the methods is also obtained.

Chapter 2 – Spatial Correlation Length of Soils in Practice

This Chapter is a study on the spatial correlation length of soils in practice. On this basis, a series of Dynamic Probing Light (DPL) penetrations were carried out in cohesive and cohesionless soils in various sites in Cyprus. The reliability of θ estimations obtained were assessed by mapping the conditions i.e., sampling domain, the interval and estimated correlation length, into the findings of Chapter 1. Continuous sampling and laboratory tests on undisturbed samples (referring to clays) also took place. The findings verify that even the so-called “homogenous soils” can be far from uniform. Indeed, the spatial correlation length θ of the soils examined was found to be as low as a few tens of centimeters. These values are far from the $\theta = \infty$ value that might be used in a simplistic probabilistic analysis.

Chapter 3 – The effect of targeted field investigation on the reliability of axially loaded piles

This Chapter deals with the practical problem of the effect of targeted field investigation on the reliability of axially loaded piles, aiming at an optimal serviceability and ultimate limit state design. This is done in a Random Finite Element Method (RFEM) framework properly considering sampling in the analysis; the RFEM method combines finite element analysis with the random field theory. In this respect, the freely available program called RPILE1D has been modified as to consider sampling of soil and pile properties. In each RFEM realization, failure is considered to have occurred when the calculated shaft resistance of pile considering spatially uniform properties (average of sampled values from the soil and pile random fields), is greater than the respective one considering spatially random properties for both soil and pile. The necessary numerical demonstration of the proposed methodology is done by considering two sampling strategies: a) sampling from a single point and b) sampling from a domain, both along the pile, whilst the various parameters governing the statistical uncertainty of the problem are examined. As shown, by adopting the proper sampling strategy (defined by the number and location of sampling points along the pile), the statistical error can be minimized or even, eliminated. The error is quantified by the difference in the probability of failure comparing different sampling scenarios. Another main finding is that, the optimal horizontal sampling location is at location where the pile is going to be constructed. In addition, it was observed that, the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic soil property values.

Chapter 4 – The effect of targeted field investigation on the reliability of earth retaining structures

This Chapter introduces the concept of targeted field investigation on the reliability of earth-retaining structures in both active and passive state, which is implemented in RFEM framework. The open source RFEM software REARTH2D was used and modified suitably in order to accommodate the purposes of the present research. Soil properties are modelled as random fields, and measurements are modelled by sampling from different points of the field domain. Failure in the active state is considered to have occurred when the “actual” resultant earth pressure force on the retaining wall (calculated using the friction angle random field) is greater than the respective “predicted” force (calculated using an homogenous friction angle field characterized by the mean of the values sampled from the respective random field). For the passive state the failure is considered to have occurred when the “actual” resultant earth pressure force is less than the respective “predicted” one. Two sampling strategies are investigated, namely, sampling from a single point and sampling from a domain, through an extensive parametric analysis. As shown, the statistical uncertainty related to soil properties may be significant and can only be minimized by performing targeted field investigation. Among the main findings is that the optimal sampling location is immediately adjacent to the wall face and half wall height away from the wall face for the active and passive state respectively. In addition, it is advisable that the entire wall height be considered in sampling for both states. Finally, it was observed that the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic values in a Load and Resistance Factor Design framework.

Chapter 5 – Reducing statistical uncertainty in elastic settlement and bearing capacity analysis of shallow foundations relying on targeted field investigation

The present Chapter deals with the practical problem of reducing statistical uncertainty in elastic settlement and bearing capacity analysis of shallow foundations relying on targeted field investigation aiming at an optimal design. In a targeted field investigation, the optimal number and location of sampling points are a priory known. As samples are taken from the material field (i.e. the ground), which simultaneously is a stress field (stresses caused by the footing), the coexistence of these two fields allows for some points in the ground to better characterize the elastic settlement and the ultimate bearing capacity of foundation. These points are identified herein through an extensive parametric analysis of the factors controlling the magnitude of settlement and bearing capacity. This is done in an advance probabilistic framework using the RFEM properly considering sampling of soil properties. In this respect, the open source RSETL2D and RBEAR2D programs, has been modified as to include the function of sampling of soil property values from the generated random fields and return the failure probability of footing against excessive settlement and bearing capacity respectively. Two sampling strategies are examined: a)

sampling from a single point and b) sampling a domain (the latter refers to e.g. continuous Cone Penetration Test data). As shown, by adopting the proper sampling strategy (defined by the number and location of sampling points), the statistical error can be significantly reduced. The error is quantified by the difference in the probability of failure comparing different sampling scenarios. Finally, from the present analysis it is inferred that, the benefit from a targeted field investigation is much greater as compared to the benefit from the use of characteristic values in a Limit State design framework.

Table of contents

List of tables

List of figures

Notation List

General

A	pile cross-sectional area
B	footing width
b_B	smaller side of footing
b_F	smaller side of raft foundation
b_g	smaller side of the rectangle circumscribing the group of piles forming the foundation at the level of the pile base
COV	coefficient of variation
c	cohesion
c'	drained cohesion
$c_s + c_0$	sill of Semivariogram model
c_0	nugget effect
D	sampling domain length
D_f	pile base diameter
d_d	sampling domain length measured always from the uppermost point of the soil
d_p	depth of sampling point
d_x	sampling interval (constant value)
E	modulus of elasticity of soil
E_p	modulus of elasticity of the pile
$E()$	error function
e_r	relative error
F	resultant wall reaction force
FS	safety factor

H	wall height
h	excavation depth
K	lateral earth pressure coefficient
K_F	Fourier frequency
K_x	maximum number of lags up to which the $\hat{\rho}(\tau)$ and $\hat{\gamma}(h)$ should be calculated
K_0	coefficient of earth pressure at rest
L	pile length
M	resultant wall reaction moment
m	number of realizations
m_x	lag distance index at which the $\hat{\rho}(\tau)$ firstly become negative
N	dataset size
N_c	bearing capacity factor
n	number of samples
n_x	number of averaged measurements
n_c	number of times that two successive points that intersect the mean value
P	vertical applied force
p	perimeter of the pile
p_f	probability of failure
$p_{f_{max}}$	maximum probability of failure of pile
$p_{f_{min}}$	minimum probability of failure of pile
Q_s	Pile shaft resistance
q_u	ultimate bearing capacity
R	radius of pile
R^2	coefficient of determination

R_d	relative percentage difference
r	data considered in clusters of 2,3,4,…,N/2
r_B	Bartlett’s distance
S	skewness factor
SLS	serviceability limit state
S_d	sample standard deviation
S_p	pile stiffness
S_s	soil stiffness
s_u	undrained shear strength
T	averaging length for local averages
t	embedded length of the support
$t_{a;v_s}$	the Student t factor for a confidence level of α% in the case of v_s degrees of freedom
ULS	ultimate limit state
U_s	soil strength
X_d	design value of the geotechnical parameter X
X_k	characteristic value of the geotechnical parameter X
X_{K_F}	Fourier transform
X_m	sample mean
x	horizontal sampling location (measured from wall face or from the center of the footing)
x_k	value of property x at location k
z_a	investigation depth below the ground level (for the problem of axially loaded pile the investigation depth is below the tip)

Greek letters

α	empirical adhesion factor
α_s	range of influence
$ac_u(z)$	adhesion at depth z
γ	unit weight of soil
γ_M	partial material factor
γR	model factor
$\gamma(\tau)$	Semivariogram function at separation distance τ
$\hat{\gamma}(\tau)$	sample Semivariogram function at separation distance τ
$\gamma(T)$	Variance Function within the averaging window T
$\hat{\gamma}(T)$	sample Variance Function within the averaging window T
γ_ϕ	partial factor for the friction angle of soil
δ'	soil–pile friction angle
δ_{max}	maximum allowable settlement of pile
$\bar{\Delta}$	average distance between the intersections of the fluctuating property and its mean value
Δ_i	distance between two successive points that intersect the mean value
θ	spatial correlation length (also, it replaces the symbols θ_v and θ_h when $\theta_v = \theta_h$)
$\hat{\theta}_e$	estimated spatial correlation length
θ_h	horizontal spatial correlation length
θ_v	vertical spatial correlation length
θ_0	pre-defined spatial correlation length
μ	mean of a soil property
μ_γ	mean unit weight of soil

$\mu_{\hat{\theta}_e}$	mean of 5000 estimated correlation lengths
$\mu_{\hat{\rho}(\tau)}$	mean sample autocorrelation function based on 5000 random fields
μ_{ϕ}	mean friction angle
$\mu_{\phi'}$	mean of drained friction angle
μ_c	mean cohesion
μ_E	mean elastic modulus
μ_{S_p}	mean pile stiffness
μ_{S_s}	mean soil stiffness
μ_{U_s}	mean soil strength
μ_{s_u}	mean undrained shear strength
μ_x	mean of the dataset
ν	Poisson's ratio of soil
v_s	degrees of freedom
v_x	sequence number
ρ	elastic settlement of footing
$\rho(\tau)$	autocorrelation function between two points separated by lag distance τ
$\hat{\rho}(\tau)$	sample autocorrelation function between two points separated by lag distance τ
σ	standard deviation of soil properties
σ_{θ_e}	standard deviation of θ_e
$\sigma_{\hat{\rho}(\tau)}$	standard deviation of sample autocorrelation function based on 5000 random fields
σ_T	standard deviation of the local averages within the averaging window T
$\sigma_n(z)$	normal effective stress at depth z

σ_{s_u}	standard deviation of undrained shear strength
$\sigma_0^{'}$	effective stress
$\bar{\sigma}'_o$	average effective overburden pressure
σ_1	Major principal stress
τ	lag distance or separation distance between two measurements
ϕ	friction angle of soil
ϕ'	effective (drained) friction angle
$\psi(z)$	interface friction angle at depth z

1 Introduction

Unlike the variability of manufactured materials used in structures, geotechnical variability is a complex attribute that results from many disparate sources of uncertainties [1]. Uncertainties in soil properties arise from three main sources, namely, inherent variability, statistical uncertainty and systematic uncertainties [2,3]. The inherent variability results from the fact that, even in seemingly homogenous soil media, the soil properties exhibit variability by nature. Due to limited material soil testing, the statistics (i.e. the mean and standard deviation) of a soil property will be subjected to statistical uncertainty [4]. Discrepancies between the laboratory and in situ conditions, due to factors such as scale, anisotropy and degree of saturation of soil are related to systematic uncertainties [2,3,5]. Usually these three sources of uncertainties in soil properties are lumped together as the total variability of estimated geotechnical parameters [6–9]. Obviously, this leads to an overestimation of uncertainties and subsequently to unrealistic estimations of the failure probability of geotechnical structures. In this respect, the above-mentioned uncertainties, are addressed separately in geotechnical reliability analyses.

Several studies have investigated the effect of inherent variability of soil properties on the performance of geotechnical structures; e.g. of slopes [10-30], of bearing capacity [31-43] and settlement of shallow foundations [44–48], of retaining structures [49–51] and of deep foundations [52–55]. These studies have a common feature that they used random field theory to model the spatial variability of soils. In random field theory, the spatial variability is characterized by the parameter's mean value (μ), standard deviation (σ), and spatial correlation length (θ) [56–58]. θ is defined as the distance within which the soil property shows relatively strong correlation or persistence from point to point [59]. A large θ value implies that the soil property is highly correlated over a large spatial extent, resulting in a smooth variation within the soil mass. On the other hand, a small value indicates that the fluctuation of the soil property is large [60]. Therefore, for the proper description of spatial variability of soils, one is required to estimate the value of θ using at least one of the methods available in the literature. In addition to the need to properly describe the inherent variability, the value of θ, as will be shown later is a critical parameter in order to achieve a significant reduction of statistical uncertainty. In this respect, the present doctorate research, firstly, performs a comparative assessment of the most widely used methods for estimating the spatial correlation length (θ) (Chapter 3).

The effectiveness of the methods is investigated by comparing the estimated θ values of 5000 one-dimensional random fields with a specific predefined θ value used for generating these fields. The random fields were generated with the *tstlas1* program. The effect of sampling domain length, sampling interval and the method used for the estimation of sample autocorrelation function on the calculation of θ were systematically investigated. The vertical spatial variability of soils in practice based on field and laboratory studies, is subsequently examined. More specifically, a series of dynamic penetrations with a Dynamic Probing Light (DPL) apparatus were carried out in both cohesive and cohesionless soils, whilst unconfined compression tests on undisturbed clay samples also took place (Chapter 4).

Next in order, this doctorate research investigates numerically the effect of targeted field investigation on the reliability of deep foundation (Chapter 5), of earth-retaining structures (Chapter 6) and elastic settlement and bearing capacity analysis of shallow foundations (Chapter 7). The targeted field investigation involves sampling from a specific point or a set of points (i.e., adopting a sampling strategy) so that the statistical uncertainty in the design is minimized (the probability of failure is minimized). The specific sampling is called optimal. The studies are based on the random finite element method (RFEM) [61], properly considering soil sampling in the analyses. The RFEM method combines finite element method with the random field theory. The random fields are generated using the Local Average Subdivision Method [62] and mapped onto the finite element mesh, taking full account of element size in the local averaging process. Contrary to the common belief that statistical uncertainty decreases with increasing number of samples [63–66], the present doctorate research will show that the statistical error for the problems stated above can be minimized only by targeted field investigation.

Finally, a summary of salient conclusions is presented in Chapter 8. The dissertation is completed by four Appendices. The Appendices present the required number of realizations considered for the RFEM models used as well as the required element size.

2 Literature review

Inherent variability represents the natural randomness of soil property that arise out of depositional, geological and geomorphological processes. This type of uncertainty may seriously affect the reliability of geotechnical structures. Inherent uncertainty cannot be reduced or eliminated [4]; however, it can be statistically described by means of random fields. As mentioned previously, the random fields can be defined by their mean value (μ), standard deviation (σ) and spatial correlation length (θ). Although μ and σ are defined through simple and standard procedures in statistics, the same does not stand for θ which essentially requires estimation of the covariance function of the data. Over the last decades, several methods have been proposed for the estimation of the correlation length. Eight of the most commonly used methods in the literature are listed in Table 2.1. These methods may all be regarded as belong to methods-of-moments, as the θ estimations follows directly from the properties of either the autocorrelation function or the variance function. In contrast, methods such as the likelihood method, see e.g. [61,67], the estimation of θ is based on maximizing a certain probability function and can be regarded as belonging to a different category. Comparative studies regarding the effectiveness of the various methods estimating the correlation length are rather limited. Uzielli [68] and Tan et al. [69] analysed several Cone Penetration Test (CPT) profiles (sampling interval for the different profiles ranged from 0.002 to 0.10m) and compared five different methods for the calculation of θ, namely the ACF, $\mathrm{ACF_A}$, $\mathrm{ACF_B}$, VF and the MCA (for the definition of each method the readers are referred to Table 2.1). They concluded that none of the models rank consistently above or below the others; however, in some cases the scatter among estimates from different methods is significant. Kenarsari et al. [70] on their study on the evaluation of spatial variability of sandy soil compared the exact same methods as Uzielli [68]. They found that the coefficient of variation (COV) of the estimated values ranges from 12 to 27% and that the COV of θ decrease if the MCA and VF are excluded. On the basis of this assumption they concluded that the ACF, $\mathrm{ACF_A}$ and $\mathrm{ACF_B}$ give more consistent results. Onyejekwe et al. [71] used the ACF and SV methods to estimate the θ of seven CPTu profiles and concluded that the θ estimated using the SV method was mostly higher than that estimated using the ACF. However, the authors admitted that their study was by no means exhaustive and further studies involving more soundings are required to come to firm conclusions.

Table 2.1: Methods of moments available in the literature for estimation of *θ*

Method	Proposed by	Examples of application
Autocorrelation Function (ACF)	Vanmarcke [59,72]	[4,59,80–89,72,90,91,73–79]
Area under the Sample ACF (ACF_A)$^{(1)}$	Vanmarcke [72]	[72,92,93]
Bartlett's Limits (ACF_B)$^{(1)}$	Jaksa [83]	[76,83,94–96]
Variance Function (VF)	Vanmarcke [59]	[59,77,92,97–99]
Fluctuation Function (FF)$^{(2)}$	Campanella et al. [100]	[56,57,68,100–104]
Variance Reduction Function (VRF)$^{(2)}$	Uzielli [68]	[56,68,97]
Semivariogram (SV)	Clark [105]	[54,56,108–113,58,81,83,85,92,105–107]
Mean-Crossing Approximation (MCA)	Vanmarcke [59]	[56,70,87,90,91,113]

$^{(1)}$ modification of the ACF method; $^{(2)}$ modification of the VF method

In contrast to inherent variability, statistical uncertainty, can be reduced by proper field investigation. Generally, the effect of soil sampling on the performance of geotechnical structures has been studied only by a few researchers. In this respect, Griffiths et al. [49] studied the effect of sampling on the reliability of passive earth pressure by using the Random Finite Element Method (RFEM). Considering a limited number of sampling locations (four in number) they concluded that, a single sampling point located at horizontal distance equal to approximately one wall height from the wall, results to lower probability of failure independent of the spatial correlation length and that, the inclusion of additional sample points to characterize the soil properties reduces the probability of failure. Jaksa et al. [114] investigated the effect of soil variability and site investigation scope on footing's settlement of a three story building and observed that the likelihood of under-designing or over-designing a footing decreases as the scope of the investigation increases. Gong et al. [115] studied the level of site exploration effort on the predicted tunneling-induced ground settlement in soft clays. Through Monte Carlo Simulation, they generated randomly variable soil properties. Then by using the maximum likelihood method they evaluated the statistics of soil parameters, following by a probabilistic analysis of the tunneling-induced ground settlement. They found that, the level of site exploration effort has only a limited influence on the mean of the predicted ground settlement. However, it has a great influence on the variation of the predicted ground settlement. Moreover, they concluded that a greater level of site exploration yields a smaller variation in soil properties, which tends to reduce the variation of the predicted tunneling-induced ground settlement. Mašín, [116] developed a general approach that incorporate experimental and sampling uncertainties with random field methods which attributes the measured total soil variability to spatial variability. He has found that while a relatively large number of samples (greater than 40) is needed to properly characterise the total soil variability, a much smaller number of samples (greater than 5) is sufficient for the experimental uncertainty quantification. Lo and Leung [117] used extended Sobol's index (a probabilistic tool developed by Professor Ilya M. Sobol [118] to assess the influence of each input parameter in a physical model) to identify the optimal locations of samples that maximize reduction to the uncertainty in the factor of safety calculated from slope stability assessment. In this respect, using existing soil samples the mean, variance, and spatial correlation features of the parameters are firstly estimated, then a set of unconditional (or conditional) random field analyses are performed and the Sobol' indices are then evaluated for each location. The maximum Sobol' index value corresponds to the optimal sampling location. Yang et al. [119] used conditional random fields enabling the site investigation data to be incorporated directly in probabilistic analysis and they found that, the coefficient of variation of factor of safety can be reduced by incorporating more site investigation data. Ching and Phoon [120] addressed the statistical uncertainties associated with the estimation of a depth-dependent trend function and spatial variation about the trend function using limited site-specific geotechnical data. This study proposed a two-step approach to characterize the uncertainties in all parameters, including the functional form of the trend, within a consistent Bayesian framework. Yang et al. [121] studied the importance of sampling location on slope stability assessment based on statistical hypothesis testing concluding that the slope crest is the optimal location to conduct geotechnical site exploration. Fenton et al. [122] studied the effect of number of samples and type of trend removal on residual uncertainty. They found that, removing the sample mean outperforms removing the best linear unbiased estimate (BLUE) when the actual field correlation length is small, but the BLUE is better to use if the actual correlation length is large relative to the domain size. Also, they found that more samples reduce the uncertainty when the field correlation length is small but does not have much impact when the field correlation length is large. Li et al. [123] linked 3D conditional random fields with finite elements, within a Monte Carlo framework, to investigate optimum sampling locations and the cost-effective design of a slope. Their results clearly demonstrate the potential of 3D conditional simulation in directing exploration programmes and designing cost saving structures. More recently, Li et al. [124] examined the influence of soil strength mean, standard deviation and spatial correlation length on the risk of slope design failure for different levels of site investigation scope using conditional random fields and found that there is an optimal number of site investigation tests, beyond which the cost of additional boreholes does not justify the cost savings due to reduced slope failure risk.

Regarding the various design codes, these are limited to some general recommendations focusing mainly on the extend of the subsurface exploration and aiming at identifying possible unfavorable geological conditions. More specifically, for *deep foundations* in normal geological conditions, EN 1997-2:2007 [125] recommends two to six investigation points per foundation with minimum depth of investigation below the pile(s) tip $z_a \geq max\{b_g, 5\text{ m}, 3D_F\}$. AASHTO [126], in turn, refers to Geotechnical Engineering Circular #5 - Evaluation of Soil and Rock Properties [127]. The latter suggests one or two investigation points for substructure widths ≤30 m and 30 m respectively and depth of investigation below the pile(s) tip $z_a \geq max\{6\text{ m}, 2D_F\}$. For *retaining structures* EN 1997-2:2007 [125] recommends that, where the piezometric surface and the groundwater tables are below the excavation base, the investigation depth, z_a, be greater than or equal to $max\{0.4h, (t + 2.0)\text{m}\}$ and where the piezometric surface and the groundwater tables are below the excavation base, z_a, the recommendation is that it be greater than or equal to $max\{(H + 2.0)\text{ m}, (t + 2.0)\text{ m}\}$; if no stratum of low permeability is encountered down to these depths, then $z_a \geq t + 5\text{ m}$. For the same problem AASHTO [126] recommends a minimum of one exploration point per retaining wall; for retaining walls more than 30 m in length, investigation points spaced every 30 to 60 m with locations alternating from in front of the wall to behind the wall are recommended. For the minimum depth of exploration it suggests that the investigation to be extended at least to a depth below the bottom of the wall where the stress increase due to the estimated foundation load is less than 10% of the existing effective overburden stress at the depth and between 1 and 2 times the wall height. For the problems of *shallow foundations* EN 1997-2 [125] recommends that, for high-rise structures and civil engineering projects, the investigation depth below the ground level should be $z_a \geq max\{3.0b_F, 6\text{ m}\}$. For raft foundations and structures with several foundation elements whose effects in deeper strata are superimposed on each other the investigation depth below the ground level should be $z_a \geq 1.5b_B$. EN 1997-2, also, recommends that the investigation points for a building or structure should be placed at critical points relative to the shape, structural behavior and expected load distribution (e.g. at the corners of the foundation area). AASHTO [126], recommends a minimum of one exploration point for substructures (e.g. piers or abutments) less than or equal to 30m in width and for substructure widths greater than 30m, a minimum of two exploration points per substructure. Additional exploration points should be provided if erratic subsurface conditions are encountered. For the depth of exploration, AASHTO suggests that the depth should be great enough to fully penetrate unsuitable foundation soils, and that the investigation should be extended at least to a depth where the stress increase due to the estimated foundation load is less than 10% of the existing effective overburden stress at the depth. Also, in case where bedrock is encountered before the depth required by the above criterion is achieved, AASHTO recommends that the exploration depth should be great enough to penetrate a minimum of 3m into the bedrock, but rock exploration should be sufficient to characterize compressibility of infill material of near-horizontal to horizontal discontinuities.

Aiming to achieve an appropriate target of reliability in geotechnical designs, the current design codes applies partial factors to characteristic parameter values. In principle the characteristic values of geotechnical parameters are selected so as to take account of the inherent variability of the ground, the uncertainty in the determination of the soil parameters and the extent of the relevant failure mechanism. “Partial factors” are also applied to actions, material properties and/or resistances to provide safety and also to account for model uncertainties and dimensional variations [128]. While Eurocode 7 [129] defines the characteristic value of a geotechnical parameter as “a cautious estimate of … the mean of a range of values covering a large surface or volume of the ground”, in the various codes of North America the mean value of the measurements is used [126,130–132]. Eurocode 7 further notes that, "if statistical methods are used, the characteristic values should be derived such that the calculate probability of a worse value governing the occurrence of the limit state under consideration is not greater than 5%." In this respect, the following statistical equation is often used for the calculation of the characteristic value [128,133]:

$$X_k = X_m - \frac{t_{a;v_S} \cdot S_d}{\sqrt{n}} \tag{2.1}$$

where X_m is the sample mean, S_d is the sample standard deviation, n is the number of samples, $t_{a;v_s}$ is the Student t factor for a confidence level of α% in the case of v_s degrees of freedom and v_s is equal to $n-1$, assuming a normal distribution. $t_{a;v_s}$values for a confidence level of 95% and various degrees of freedom v_s can be found at Table 2.2.

Table 2.2: $t_{a;v_s}$values for a confidence level of 95% and various degrees of freedom v_s[134].

n	3	4	5	6	7	11	16	21	41
v_s=n-1	2	3	4	5	6	10	15	20	40
$t_{a(95\%),v_s}$	2.920	2.353	2.132	2.015	1.943	1.812	1.753	1.725	1.684

EN 1997-1 [129] dictates that the *design* values of the geotechnical parameters (X_d) be derived from the respective characteristic values using the following equation:

$$X_d = X_k/\gamma_M \tag{2.2}$$

where, X_k is the characteristic value of a material property X and the symbol γ_M denotes partial material factor. When partial factors are not applied to the material properties (i.e. γ_M =1), a model factor γ_R greater than 1 is applied to the resistances.

3 A comparative assessment of the methods-of-moments for estimating the correlation length of one-dimensional random fields

In this chapter, the effectiveness of the eight methods-of-moments presented in Table 2.1 for estimating the correlation length θ is investigated. This is done by generating samples of one-dimensional random fields for pre-defined values of the correlation length, which is then estimated by the different methods. For each method, the influence of the sampling domain length D and sampling interval d_x on the estimation of θ were investigated, and the results are quantified in the form of errors over the parameter space, defined by the dimensionless ratios D/θ and θ/d_x.

3.1 Methods of moments for estimating *θ*

3.1.1 The ACF method and its modifications

Probably the earliest application of the autocorrelation function (ACF) in geotechnical engineering was made by Atterberg [135]. However, the method became wider-known after Vanmarcke's work [59] and since then it remains one of the most popular methods for estimating θ. The procedure is briefly presented below.

The autocorrelation function, $\rho(\tau)$, is a measure of the correlation between two random measurements of the same property separated by lag distance τ [136]. However, because the available data are usually limited, only an estimation of the real ACF can be made [56]. In such a case, the "sample autocorrelation function" $\hat{\rho}(\tau)$ is used. The latter can be estimated by either the Method of Moments or the Fourier transform method.

Following the Method of Moments, $\hat{\rho}(\tau)$ is given by the following equation:

$$\hat{\rho}(\tau)=\frac{\sum_{k=1}^{N-v_x}(x_k-\hat{\mu}_x)(x_{k+v_x}-\hat{\mu}_x)}{\sum_{k=1}^{N}(x_k-\hat{\mu}_x)^2},\quad v_x=1,2,3,\ldots,K_x \tag{3.1}$$

In geotechnical engineering, it is common practice to assume data stationarity, which implies that the mean and variance are constant and the autocorrelation depends on the lag distance rather than the actual sample location [137–139]. The lag distance τ is an integral multiple of the sampling interval length: $\tau = v_x d_x$. K_x is the maximum number of lags up to which $\hat{\rho}(\tau)$ should be calculated. For lag τ=0 the autocorrelation function is equal to 1. The sample mean $\hat{\mu}_x$ is defined by

$$\hat{\mu}_x = \frac{1}{N}\sum_{k=1}^{N} x_k \tag{3.2}$$

Alternatively, the sample autocorrelation function can be obtained by computing the Fourier transform of the data. In the first step, the Fourier transform X_{K_F} of data x_k can be calculated as:

$$X_{K_F} = \sum_{k=0}^{N-1} x_k \, e^{-i2\pi \frac{k K_F}{N}} \tag{3.3}$$

where x_k is the value of property x at location k=0,1,2,…, N-1 and K_F =0, 1, 2…, N-1 is the Fourier frequency. In order to apply Fast Fourier Transformation algorithms, the number N is conveniently chosen a power of 2. Here it is taken to be $N = 2^{10} = 1024$. The autocorrelation function is given by the Fourier transform of the modulus-squared of X_{K_F}:

$$\hat{\rho}(\tau) = \sum_{K_F=0}^{N-1} \frac{\left|X_{K_F}\right|^2}{N^2} e^{-i2\pi \frac{v_x K_F}{N}} \tag{3.4}$$

The correlation length may then be estimated by i) fitting theoretical autocorrelation models to the $\hat{\rho}(\tau)$ ii) estimating the area under the positive part of $\hat{\rho}(\tau)$ or iii) using the Bartlett's limits method; ACF, ACF_A and ACF_B method respectively.

In the first case (ACF method), theoretical autocorrelation models are fitted to the sample autocorrelation function $\hat{\rho}(\tau)$ (Figure 3.1) using curve fitting methods, such as the Ordinary Least Square (OLS) method. The most commonly used theoretical autocorrelation models are given in Table 3.1. In this respect, Uzielli et al. [75] recommend that only the theoretical autocorrelation models giving $R^2 > 0.9$ be accepted with at least four initial autocorrelation coefficients greater than r_B:

$$r_B \approx \pm \frac{1.96}{\sqrt{N}} \tag{3.5}$$

In the second case (ACF$_A$ method), according to Vanmarcke [72], the correlation length θ is equal to the area under $\hat{\rho}(\tau)$, or twice the area under the positive part since $\hat{\rho}(\tau)$ is symmetric about τ=0, thus:

$$\theta = \int_{-\infty}^{+\infty} \hat{\rho}(\tau)d\tau = 2\int_{0}^{+\infty} \hat{\rho}(\tau)d\tau \tag{3.6}$$

The area is typically taken up to a certain distance $\tau = m_x d_x$ where the sample autocorrelation function $\hat{\rho}(\tau)$ first becomes negative (Figure 3.2a). Then, the correlation length θ is estimated by applying the trapezium rule to equation (3.6) [56]:

$$\theta = 2\sum_{k=2}^{m_x} \frac{\left(\hat{\rho}_{k-1} + \hat{\rho}_k\right)}{2}\Delta\tau = d_x \sum_{k=2}^{m_x} \left(\hat{\rho}_{k-1} + \hat{\rho}_k\right) \tag{3.7}$$

Regarding the ACF$_B$ method, Jaksa [83] working on stiff over-consolidated clays observed that θ can be estimated relatively easily by the Bartlett's distance r_B, which is the distance at which the sample autocorrelation function $\hat{\rho}(\tau)$ first intersects the limit (Figure 3.2b) given by Bartlett's formula (see equation (3.5)).

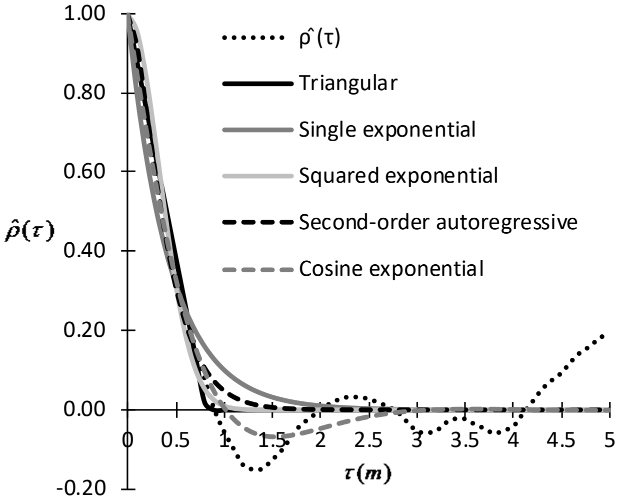


Figure 3.1: ACF method: Best fit of theoretical autocorrelation models $\rho(\tau)$ on $\hat{\rho}(\tau)$

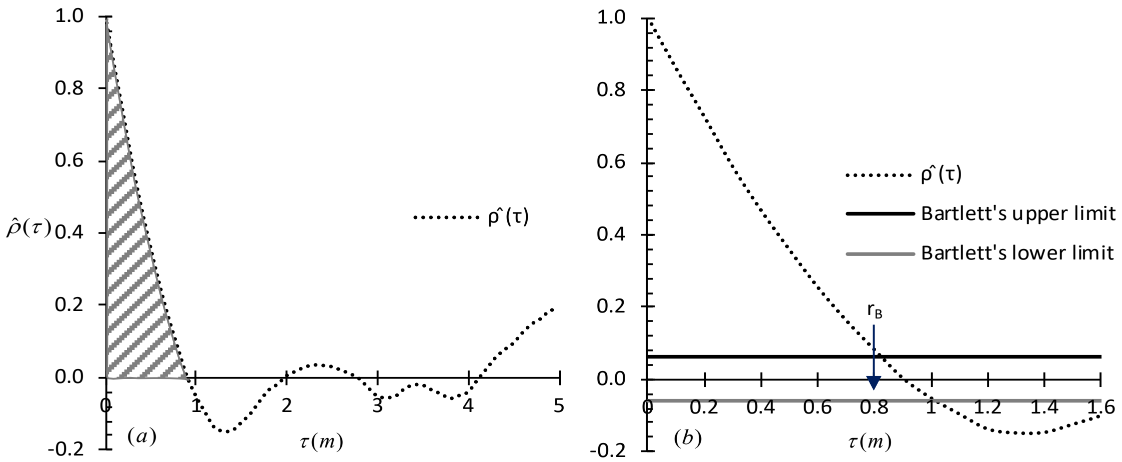


Figure 3.2: (a) ACF_A: Area under the positive part of $\hat{\rho}(\tau)$ and (b) ACF_B: Bartlett's Limits method (θ=r_B)

Table 3.1: Theoretical Autocorrelation models used for the estimation of *θ*

Model	Autocorrelation Function	Correlation length θ
Triangular	$\rho(\tau)=\begin{cases}1-\dfrac{\lvert\tau\rvert}{\alpha}, & \lvert\tau\rvert\le\alpha \\ 0 & \lvert\tau\rvert\ge\alpha\end{cases}$	a
Single Exponential (Markovian)	$\rho(\tau)=e^{-\frac{\lvert\tau\rvert}{b}}$	2b
Squared Exponential (Gaussian)	$\rho(\tau)=e^{-\left(\frac{\lvert\tau\rvert}{b}\right)^2}$	$\sqrt{\pi}c$
Second-order autoregressive	$\rho(\tau)=\left[1+\dfrac{\lvert\tau\rvert}{d}\right]e^{-\frac{\lvert\tau\rvert}{d}}$	4d
Cosine exponential	$\rho(\tau)=e^{-\frac{\lvert\tau\rvert}{d}}\cos\left(\dfrac{\lvert\tau\rvert}{\alpha}\right)$	a

3.1.2 The Variance Function method and its modifications

According to Vanmarcke [59,72] the correlation length can be estimated using the following equation:

$$\theta = T\hat{\gamma}(T) \quad \text{for} \quad T\to\infty \tag{3.8}$$

where, $\hat{\gamma}(T)$ is the sample Variance Function which measures the reduction of the point variance due to local averaging over regions of length $T=(n_x-1)d_x$, with n_x being the number of averaged measurements and d_x the constant sampling interval. The sample Variance Function $\hat{\gamma}(T)$ can be calculated as follows:

$$\hat{\gamma}(T)=\frac{\sigma_T^2}{\sigma^2} \tag{3.9}$$

where, σ_T is the standard deviation of the local averages within the averaging window T and σ is the standard deviation of data. In order to calculate $\hat{\gamma}(T)$ the data are first combined in pairs ($r=2$) and a moving average series is obtained. Then, the standard deviation σ of the resulted series is computed. The above procedure is repeated for increasing number of r. As r increases, σ becomes smaller than the standard deviation of the original function due to the spatial averaging, while μ remains identical to the original.

The correlation length θ is then estimated by best fitting theoretical Variance Function models, $\gamma(T)$, to the $\hat{\gamma}(T)$ as shown in Figure 3.3. The most commonly used theoretical $\gamma(T)$ models are listed in Table 3.2. Alternatively, θ can be estimated using the Fluctuation Function (FF) method and the Variance Reduction Function (VRF) method (see Table 2.1), which are variances of the Variance Function method. The Fluctuation Function can be defined as the product of the Variance Function $\hat{\gamma}(T)$ by the separation distance T. Theoretically, the Fluctuation Function curve $T\hat{\gamma}(T)$ will be constant when the separation distance T increases to a significant extent. However, in reality, as the separation distance increases, the value of $T\hat{\gamma}(T)$ decreases and irregular fluctuation often appears [103]. Campanella et al. [100] suggests that θ be taken as the peak point (i.e. point of inflection) of the $T\hat{\gamma}(T)-T$ curve (Figure 3.4a), whilst Uzielli [68] suggests that θ be taken as the value of $T\hat{\gamma}(T)$corresponding to $\hat{\gamma}(T)$ = 0.15 (Figure 3.4b).

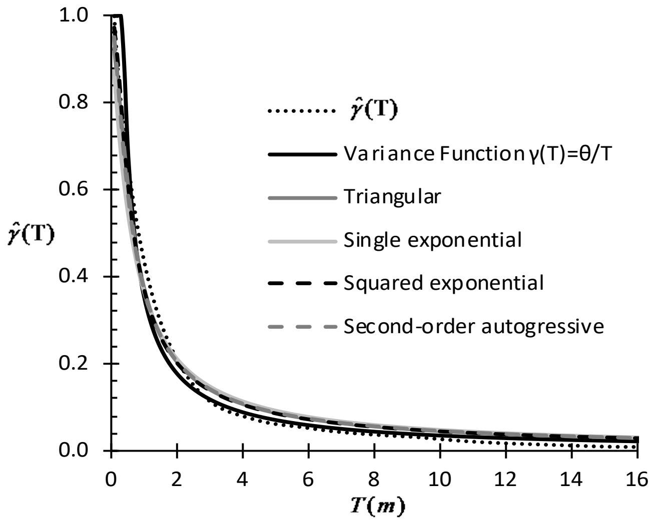


Figure 3.3: VF method: Best fit of theoretical variance function models $\gamma(T)$ on the $\hat{\gamma}(T)$

Table 3.2: Theoretical variance function models used for the estimation of *θ*.

Model	Variance Function	Correlation length θ
Triangular	$\gamma(\mathrm{T})=\begin{cases} 1-\dfrac{\mathrm{T}}{3a} & T\le a \\ \left(\dfrac{a}{T}\right)\left[1-\dfrac{a}{3T}\right] & T\ge a \end{cases}$	a
Single Exponential (Markovian)	$\gamma(\mathrm{T})=2\left(\dfrac{b}{T}\right)^2\left(\dfrac{T}{b}-1+e^{-\frac{T}{b}}\right)$	$2b$
Squared Exponential (Gaussian)	$\gamma(\mathrm{T})=\left(\dfrac{c}{T}\right)^2\left[\sqrt{\pi}\dfrac{\mathrm{T}}{c}E\left(\dfrac{T}{c}\right)+e^{-\left(\frac{T}{c}\right)^2}-1\right]$	$\sqrt{\pi}c$
Second-order autoregressive	$\gamma(\mathrm{T})=2\dfrac{d}{T}\left[2+e^{-\frac{T}{d}}-3\dfrac{d}{T}\left(1-e^{-\frac{T}{d}}\right)\right]$	$4d$

Note: $E(x)$is the usual error function which increases from 0 to 1 as its argument increase from 0 to ∞.

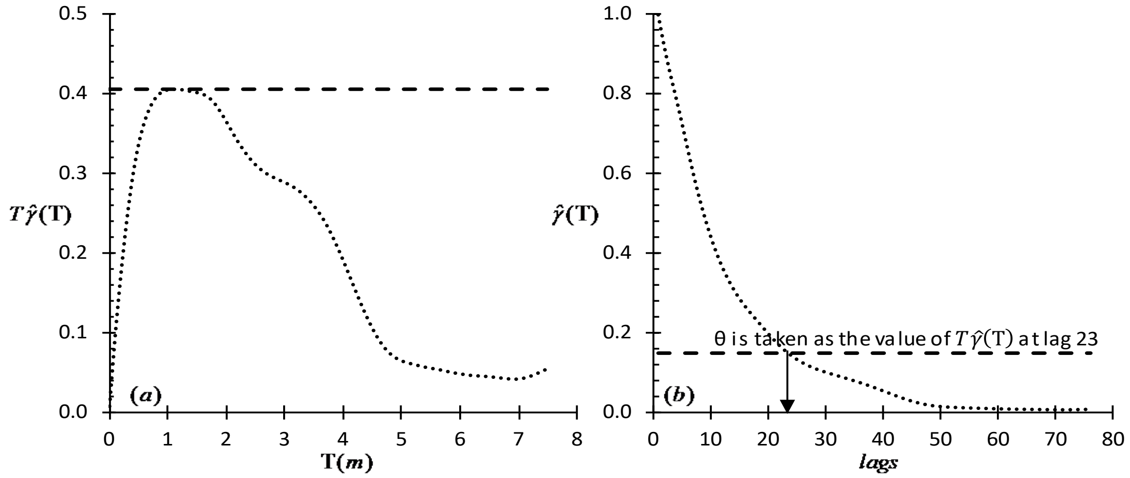


Figure 3.4: (*a*) FF method [100] and (*b*) VRF method [68]

3.1.3 The semivariogram method

To characterise the dependence between variables at different points, geostatistics employs a tool called semivariogram. According to Fenton [92] and Jaksa et al. [140], the semivariogram $\gamma(\tau)$ gives essentially the same information as the autocorrelation function and can be defined as half the expected value of the squared difference between two pairs of points [85]:

$$\gamma(\tau)=\frac{1}{2}E[(x_{k+\nu_x}-x_k)^2] \tag{3.10}$$

where $E[]$ is the expectation value operator. The sample semivariogram is defined by the following equation:

$$\hat{\gamma}(\tau)=\frac{1}{2\left(N-\nu_x\right)}\sum_{k=1}^{N-\nu_x}\left(x_{k+\nu_x}-x_k\right)^2 \tag{3.11}$$

To obtain reliable estimates, the $\hat{\gamma}(\tau)$ should be calculated for about half the total sampled extent [105,141–143]. In addition, the number of pairs at each lag, should be larger than 30 to 50 [142].

Moreover, in order to estimate θ, a continuous theoretical semivariogram model has to be fitted to the sample semivariogram, as shown in Figure 3.5; the most commonly used semivariogram models are listed in Table 3.3. In this respect, a number of authors (e.g. [54,83,105,142,144]) chose the theoretical semivariogram model that best fits to the sample semivariogram by eye based on a trial and error procedure.

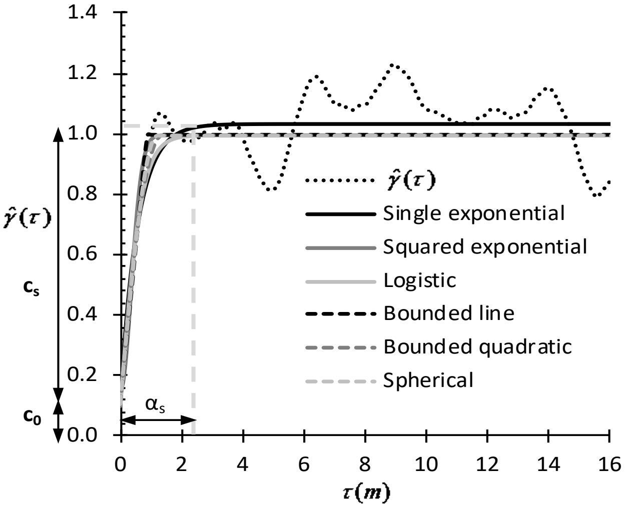


Figure 3.5: Best fit of theoretical semivariogram models $\gamma(\tau)$ on the $\hat{\gamma}(\tau)$

It is noted that, the theoretical semivariogram models are generally characterised by three parameters: (i) the nugget effect, c_0, (ii) the sill, $c_s + c_0$, and (iii) the range of influence, a_s(Figure 3.5). More specifically the nugget effect is the apparent discontinuity of the variogram at τ=0 m and it is mainly due to the microstructure, sampling and measurement error [109,144,145]. The sill is the distance in which the semivariogram ceases to increase sharply and its value approximates the data variance [105,146,147]. Sill defines the range of influence a_s with the latter being equivalent to θ.

3.1.4 The Mean-Crossing Approximation method

Based on the Mean-Crossing Approximation (MCA) method [59], θ can be approximated as:

$$\theta \approx \sqrt{\frac{2}{\pi}}\,\overline{\Delta} = \sqrt{\frac{2}{\pi}}\left(\frac{1}{n_c}\sum_{i=1}^{n_c}\Delta_i\right) \tag{3.12}$$

Schematically, this is illustrated in Figure 3.6.

Table 3.3: Theoretical semivariogram models used for the estimation of θ.

Model	Theoretical semivariogram function	Correlation length θ
Single Exponential (Markovian)	$\gamma(\tau) = c_0 + c_s\left(1 - \mathrm{e}^{-(\tau/a_s)}\right)$	$2a_s$
Squared Exponential (Gaussian)	$\gamma(\tau) = c_0 + c_s\left(1 - \mathrm{e}^{-(\tau/a_s)^2}\right)$	$\sqrt{\pi}a_s$
Logistic	$\gamma(\tau) = c_0 + c_s \tan\tau\left(\tau/a_s\right)$	$\ln(4)a_s$
Bounded line	$\gamma(\tau) = \begin{cases} c_0 + c_1\tau, & 0 < \tau \le a_s \\ c_0 + c_s, & \tau > a_s \end{cases}, \left(c_1 = \dfrac{c_s}{a_s}\right)$	a_s
Bounded quadratic	$\gamma(\tau) = \begin{cases} c_0 + c_s\left(2\left(\dfrac{\tau}{a_s}\right) - \left(\dfrac{\tau}{a_s}\right)^2\right), & 0 < \tau \le a_s \\ c_0 + c_s, & h > a_s \end{cases}$	$\frac{2}{3}a_s$
Spherical	$\gamma(\tau) = \begin{cases} c_0 + c_s\left(\dfrac{3}{2}\left(\dfrac{\tau}{a_s}\right) - \dfrac{1}{2}\left(\dfrac{\tau}{a_s}\right)^3\right), & 0 < \tau \le a_s \\ c_0 + c_s, & \tau > a_s \end{cases}$	$\frac{3}{4}a_s$

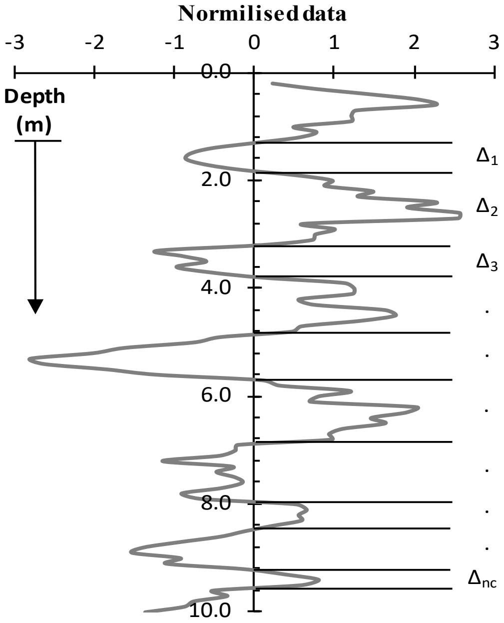


Figure 3.6: Fluctuation of normalised data along the mean axis.

3.1.5 Parameters influencing the estimation of θ

Vanmarcke [72] in his study on random fields shows that the optimal sampling intervals may be expected to be proportional to the correlation length. Also, through estimating the expectation value of the autocorrelation function estimated from finite domain samples, shows on general theoretical grounds that the latter, and hence the estimated correlation length, depends on the finiteness of the domain. DeGroot and Baecher [67] analysed vane shear strength data and assessed the spatial correlation using the maximum likelihood method. They found that θ was influenced by sample spacing and that the sample locations separated by distances within the range of θ are essential to providing good estimates of θ. Jaksa et al. [138] using the ACF nugget (which has the same meaning as c_0 of semivariogram method) studied the influence of sampling interval, which varied between 5mm up to 200mm, and showed that the estimated ACF nugget can vary between 3% and 62% for vertical spatial variability, and between 3% and 50% for horizontal spatial variability. In similar work, Jaksa et al.[148] showed that the horizontal spatial variability exhibits a nested structure, which means that the estimated θ increase with the size of the sampling domain. Fenton [149] analysed 143 regularly spaced CPT records and observed that θ depends on the sampling domain length, in that a larger sampling domain would result in a larger θ. Also, Fenton [92] stated that the sample correlation or covariance functions are acceptable measures of second moment (covariance) structure when the θ of the process is small relative to the sampling domain. Cafaro and Cherubini [57] studied the influence of large sampling interval on the assessment of vertical strength variability of a clayey soil. By adopting the FF method, they indicated that θ tends to increase as sampling interval increases. Nie et al.[150] studied the factors affecting the calculation of θ by generating one-dimensional Gaussian random fields using the Cholesky decomposition method. Using the ACF method, they estimated the predefine θ and concluded that the sampling interval should be at least 10 measurements within one θ, and the sampling extent should be wide enough such that it covers at least 100 times the θ. Sasanian [151] by adopting the ACF method, investigated the effect the sampling interval on the θ in the vertical direction using 70 CPT profiles and found that the reliability of the θ decreases with the increase in sampling interval. More specifically, for a specific dataset, which has a benchmark sampling interval of 0.05 m and θ =2.347 m , they extended the initial sampling interval to 0.1, 0.2, 0.4, 0.8, 0.16 and 0.32 m, and observed that by increasing the sampling interval to a value higher than 0.2 m, the reliability of the determined θ decreases.

3.2 Effectiveness of the methods estimating θ

As mentioned, the effectiveness of the methods given in Table 2.1 is investigated by comparing the estimated theta values (θ_e) of 5000 one-dimensional random fields with a specific θ_0 value used for generating these fields. As all the results of this analysis depend only on the dimensionless ratios D/θ_0 and θ_0/d_x where D is the domain length and d_x is the sampling interval, the input value θ_0 can be arbitrarily set; it has been set to 1.5m. The random fields were generated with the *tstlas1*[1] program. The *tstlas1* program uses the Local Averaging Subdivision (LAS) method developed by Fenton and Vanmarcke [62] to generate random fields with given mean, standard deviation, correlation length and correlation function. According to Fenton and Vanmarcke [62], the Local Average Subdivision method is a fast and generally accurate method of producing realizations of a discrete "local average" random process. Although, the generated samples of random fields by the *tstlas1* program were validated by comparing the average autocorrelation function of the samples to the theoretical estimate given by Vanmarcke [72] for arbitrary domain length D. The validation is presented in Appendix I. Moreover, the single exponential correlation function has been selected, whilst the mean and the standard deviation was set to zero and unity respectively. For the needs of the present research, minor modifications of the program were made as to consider 1024 cells instead of 512 ones and to export out each random field separately. For the assessment of each method, the first three moments of the 5000 estimated thetas were obtained and analysed in respect to sampling domain and interval.

Various sampling domain lengths ranging from D= 5m to 100m were considered whilst the sampling interval d_x was held constant and equal to 0.1m. The effect of sampling interval is further investigated by considering sampling intervals ranging from d_x = 0.1m to 2m, whilst the θ_0 and D were held constant and equal to 1.5m and 100m respectively. The Ordinary Least Square (OLS) method was adopted to best fit theoretical autocorrelation and variance function models to the sample ones. However, the Weighted

---

[1] http://www.engmath.dal.ca/rfem

Least Squares (WLS) method was used to best fit theoretical semivariogram models to the sample ones, since it gives more weight to early lags with great number of pairs and decrease bias resulting from lags with small number of pairs [152,153].

3.2.1 Effect of the method used for the estimation of $\hat{\boldsymbol{\rho}}(\boldsymbol{\tau})$

Before proceeding with the examination of the various methods, it is important to assess the reliability of the methods used for the estimation of the “sample autocorrelation function” $\hat{\rho}(\tau)$, i.e. the Method of Moments or the Fourier transform method.

In order to minimize the effect of sampling domain length (*D*) and the sampling interval (d_x) on the estimation of $\hat{\rho}(\tau)$, the following values were chosen: θ_0= 1m, *D* =50m, d_x =*D*/1024 =0.0488m, corresponding to *N*= 1024 sampling points. Hence, the sampling interval was 20 times smaller than the correlation length, while the latter was 50 times smaller than the sampling domain. The mean sample autocorrelation function, $\mu_{\hat{\rho}(\tau)}$, as well as the standard deviation, $\sigma_{\hat{\rho}(\tau)}$, of the 5000 random fields were computed and plotted against lag distance in Figure 3.7. From Figure 3.7a it is inferred that there is no difference in $\mu_{\hat{\rho}(\tau)}$ between the two methods. However, Figure 3.7b is more informative, showing that for τ >1m the standard deviation of the estimated autocorrelation function calculated by the method of moments behaves differently than the one of the Fourier transform method. One observes that the two curves in Figure 3.7b start deviating at the lag distance where $\hat{\rho}(\tau)$ firstly intersect the τ axis (see Figure 3.7a). This is an indication that reliable information is stored only in the positive part of $\hat{\rho}(\tau)$ up to the first zero. This is in agreement with the remarks of Lacasse and Nadim [4].

The maximum number of lags (denoted by K_x) up to which $\hat{\rho}(\tau)$ should be calculated (recall equation (3.1)) has been discussed by several authors suggesting that the maximum number of lags K_x be much smaller than the total number of the dataset *N* (e.g. [143,154,155] suggest K_x=*N*/4). This is due to the fact that in large separation distances the number of pairs contributing to $\hat{\rho}(\tau)$ diminish to such an extent that the values are unreliable [156]. Adopting the K_x=*N*/4 value, the $\hat{\rho}(\tau)$ for the current example was calculated for τ up to 12.5m. For this τ value, as shown in Figure 3.7b, the Fourier transform method gives appreciably more reliable results as compared to the Method of

Moments. In the present research the $\hat{\rho}(\tau)$ was estimated based on the Method of Moments using only the first positive part in the $\hat{\rho}(\tau) - \tau$ diagram.

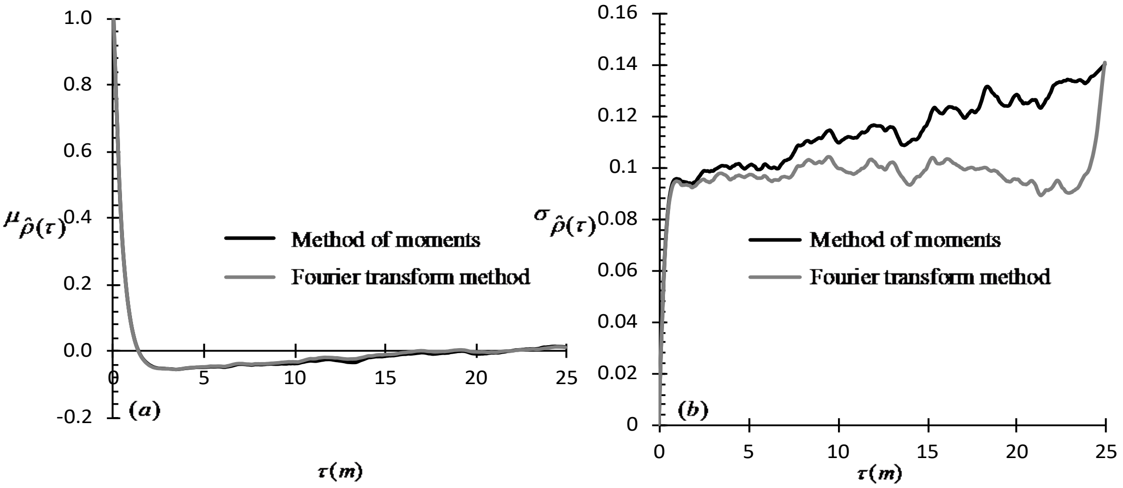


Figure 3.7: (a) mean (μ), and (b) standard deviation (σ), of $\hat{\rho}(\tau)$ based on 5000 random fields estimated using the Method of Moments and the Fourier transform method.

3.2.2 Effectiveness of the ACF method and its modifications

Figure 3.8a shows the variation of the μ_{θ_e}/θ_0 ratio (mean value of the 5000 estimated theta values) with the scaled domain length D/θ_0 for different theoretical $\rho(\tau)$ models. From this figure it is inferred that the single exponential (Markovian) correlation function is the best choice. However, this did not come as a surprise, as the random fields were generated with the same function. This highlights the need for investigating the model that best-fits to the sample autocorrelation function; Figure 3.8a shows that the difference in μ_{θ_e}/θ_0 for different model correlation functions can be as high as 25%. However, it can be said that all models examined here but the second-order autoregressive one provide in general a fairly good estimate of θ. Figure 3.8a also shows that the estimated θ depends strongly on D/θ_0. More specifically, as the D/θ_0 ratio increases, the μ_{θ_e} value tends to an asymptotic value which may be close or less close to the input value θ_0.

Figure 3.8b presents the results from ACFA method, which does not depend on theoretical autocorrelation function models. One may conclude that this method performs very well, behaving similarly to the ACF method for the single-exponential autocorrelation function, with the μ_{θ_e}/θ_0 curve showing tendency to converge to the desired value 1 for large domains. The results for the ACFB method are presented in Figure 3.8c. This method exhibits an over-estimation of the θ for large domains, crossing the desired value 1 at domains lengths of about $33\theta_0$.

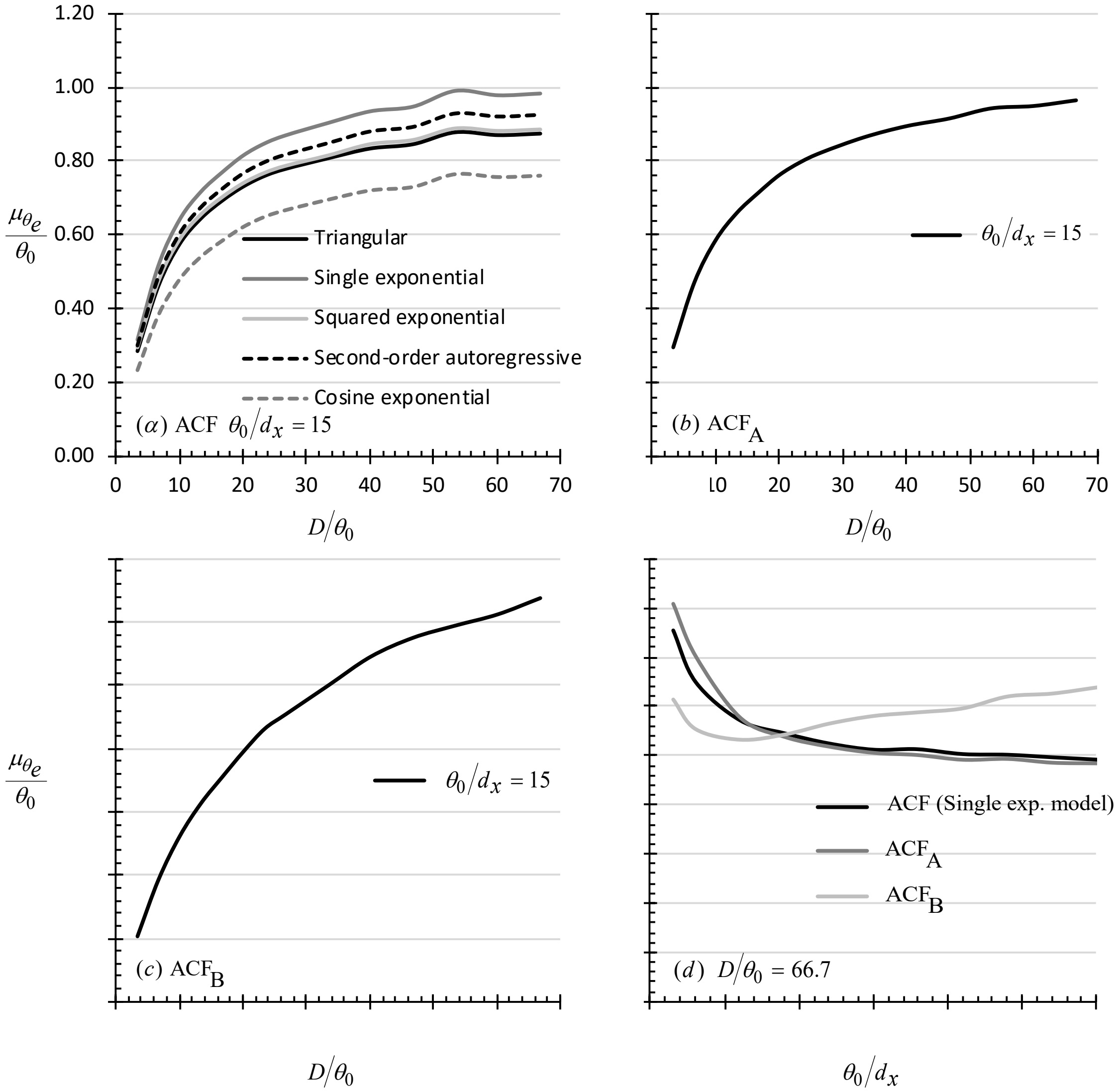


Figure 3.8: (a-c) μ_{θ_e}/θ_0 versus D/θ_0 charts for the ACF, ACF_A and ACF_B method respectively, for $\theta_0/d_x = 15$ (d) μ_{θ_e}/θ_0 versus θ_0/d_x chart for the ACF, ACF$_A$ and ACF$_B$ method for D/θ_0= 66.7

Figure 3.8d presents the dependence of μ_{θ_e}/θ_0 on θ_0/d_x for all three ACF methods, for D/θ_0= 66.7. The ACF method is given in this figure for the single-exponential case, which is best performing as discussed above. It is shown that the ACF and ACF$_A$ methods provide better prediction for theta as the sampling interval decreases i.e. the ratio θ_0/d_x increases, although a value of this ratio above 5 appears to be practically adequate. On the other hand, the ACF$_B$ method appears to result in over-estimations of the θ.

It is noted here that, Jaksa [83] investigated the relationship between ACF and ACF$_B$ and found a strong regression fit between the two. Nonetheless, it should be noted that the sampling domain length of the CPT profile examined by Jaksa was about 5m and the θ was found to be of the order of 0.2m, which means that D/θ =25. The sampling interval was set equal to 0.005m. By recalling the Figure 3.8 it can be seen that both methods result to about the same μ_{θ_e}/θ_0 value, which explains the strong regression found by Jaksa. Jaksa [83] observed that the r_B parameter (ACF_B) appears to be somewhat insensitive to sample interval when the interval is less than the θ and thus, this could be seen as a strength of this method [68]. On the other hand, according to our results, the estimations of theta by this method depend rather strongly on the sampling interval. In fact, the theta estimates increase with decreasing sampling interval.

3.2.3 Effectiveness of the Variance Function method and its modifications

The effectiveness of the original Variance Function (VF) method and its modifications (FF and VRF method; see Table 2.1) are examined in an analogous manner. Four comparison charts are given. Figure a shows the variation of the ratio μ_{θ_e}/θ_0 with D/θ_0 for different theoretical $\gamma(T)$ models. In this respect, the single exponential model appears to be again the best choice, for the same reason explained in the previous paragraph. The function $\gamma(\tau) = \theta/T$, on the other hand, appears to give the poorest prediction, whilst the triangular variance, the squared exponential and the second order autoregressive models appears to be less good than the single exponential model but much better than the function $\gamma(\tau) = \theta/T$. The FF and the VRF methods, which are independent of best fit models, appear to have similar behaviour with the original VF method if the latter is used along with the single exponential model. In every case, Figure a-c show that the VF methods require much longer domain length as compared to the ACF methods. Indeed, such domain lengths may be prohibitive in practice. Another important observation is that the VF methods do not appear to converge to an asymptotic value, at least for the range of sampling domain lengths considered here. That implies that the VF methods require much larger domains relatively to the correlation length than the ACF method.

Figure 3.9d shows the dependence of μ_{θ_e}/θ_0 on the sampling interval. An asymptotic μ_{θ_e}/θ_0 value is reached in every case, however this value is much smaller than the expected (μ_{θ_e}/θ_0 =1), at least for the largest sampling domain considered here. This behaviour may be attributed to the fact that, as mentioned above, the VRF methods require even larger sampling domains in order to perform well.

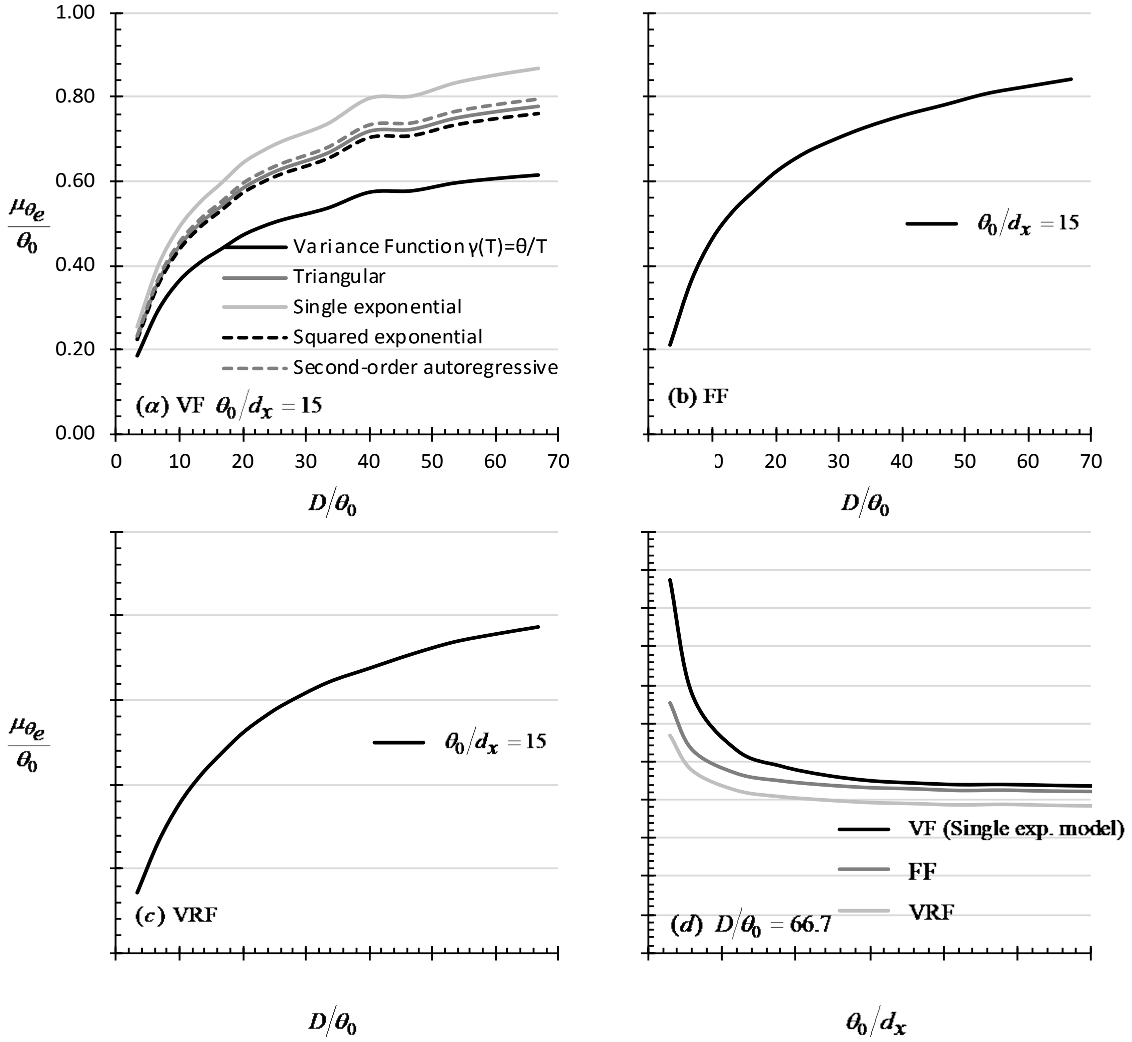


Figure 3.9: (a-c) μ_{θ_e}/θ_0 versus D/θ_0 chart for the VF, FF and VRF method respectively for $\theta_0/d_x = 15$ (d) μ_{θ_e}/θ_0 versus θ_0/d_x chart for the VF, FF and VRF method for D/θ_0= 66.7.

3.2.4 Effectiveness of the semivariogram method

Figure 3.10a shows the variation of the ratio μ_{θ_e}/θ_0 with D/θ_0 for different theoretical $\gamma(\tau)$ models. It turns out that the single exponential model, with which the random fields have been generated, gives μ_{θ_e}/θ_0 slightly higher than unity. The logistic model appears also to perform well, in the sense that the asymptotic value of μ_{θ_e}/θ_0 is close to unity. Figure 3.10b shows the effect of the sampling interval on the ratio μ_{θ_e}/θ_0 for D/θ_0 =66.7. It turns out that the semivariogram method is somewhat insensitive to sample interval.

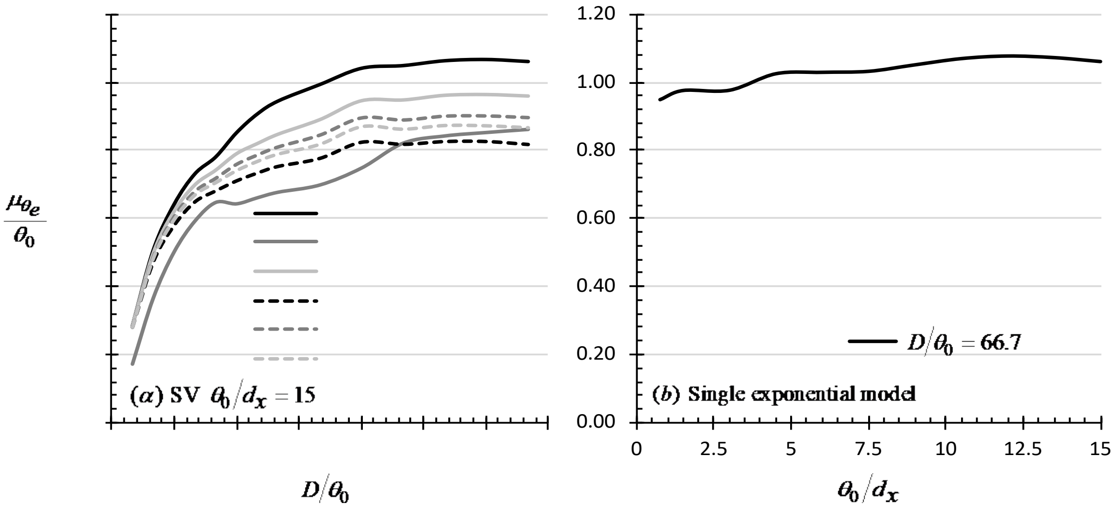


Figure 3.10: SV method: (a) μ_{θ_e}/θ_0versus D/θ_0 chart for $\theta_0/d_x = 15$ (b) μ_{θ_e}/θ_0 versus θ_0/d_x chart forD/θ_0= 66.7.

3.2.5 Effectiveness of the Mean-Crossing Approximation method

The effectiveness of the Mean-Crossing Approximation (MCA) method can be examined with respect to the sampling domain length and interval as shown in Figure 3.11. From Figure 3.11a, it appears that this method results to substantially smaller μ_{θ_e}/θ_0 values than the desired one for every D/θ_0 value examined. It should be noted that the theta estimate is virtually independent of the domain length for D/θ_0 greater than 10. Also, in contrast to the previous methods, the MCA is highly affected by the sampling interval (Figure 3.11b). As shown, the value μ_{θ_e}/θ_0 is significantly decreases as the θ_0/d_x ratio increases. Hence, one may conclude that the MCA method systematically under-estimates the correlation length.

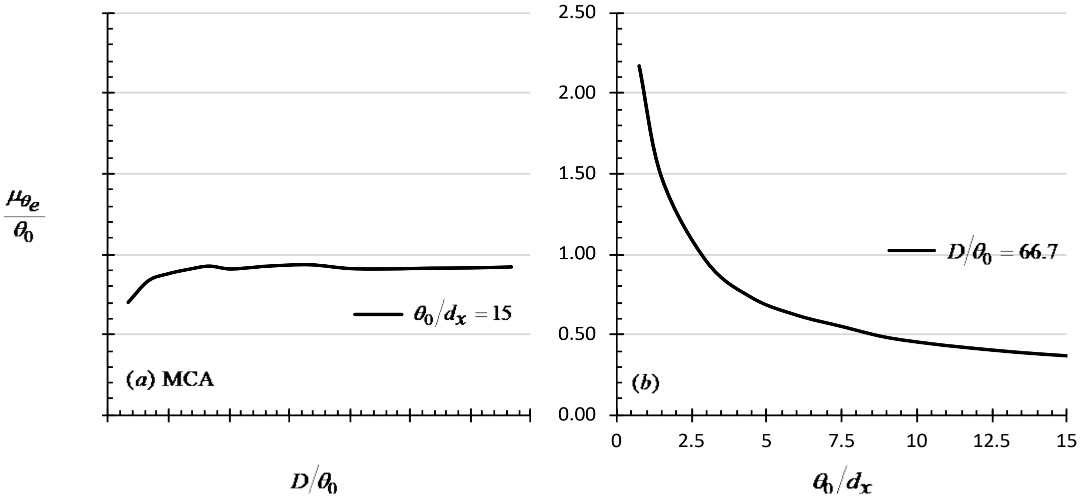


Figure 3.11: MCA method: (a) μ_{θ_e}/θ_0 versus D/θ_0 chart for various θ_0/d_x values, (b) μ_{θ_e}/θ_0 versus θ_0/d_x for D/θ_0 = 66.7.

3.3 Overall assessment of the methods

In what follows, the overall assessment of the various methods is based on the following four performance indices, defined over the sample of the 5000 estimated theta values for each method:

a) The sample mean μ_{θ_e} normalized with respect to the input value θ_0.

b) The relative error e_r of the mean value of the sample relatively to the input value.

c) The standard deviation $\sigma_{\theta_e}$of the sample.

d) The skewness factor S of the sample.

These indices are calculated for different domain lengths and sampling intervals, in order to quantify the influence of these parameters on the performance of each method. The cases of varying domain length correspond to sampling interval of 0.1m so that $\theta_0/d_x=15$ (Figure 3.12). The cases of varying sampling interval correspond to $D/\theta_0=66.7$ (Figure 3.13). Wherever applicable, the single exponential (Markovian) model has been used.

Figure 3.12a presents collectively the mean theta value curves shown already in the analysis of each method in Section 2 of the current Chapter. One observes that the autocorrelation function (ACF) and the area under the positive part of the autocorrelation function ($\mathrm{ACF_A}$) methods perform similarly and very well, approaching the desired value as the domain length increases. The desired is nearly attained for the largest domain considered in this analysis, $D/\theta_0\sim67$. The semivariogram method (SV) shows a consistent convergence towards a specific mean value, although overestimates the input value of theta by 6%, for the large domain lengths. The Fluctuation Function (FF) and variance reduction function (VRF) methods are variations of the variance function (VF) method. The FF method performs similarly to the VF one, while both exhibit a much slower approach to the input value for increasing domain length. The VRF method shows an even slower approach to the desired value. Bartlett’s limit method ($\mathrm{ACF_B}$) exhibits an increase without visible convergence, overestimating the input value of theta already at $D/\theta_0\sim33$. On the other hand, the mean crossing method (MCA) shows a strong convergence to a mean value ~35% of the input value almost independently of the domain length. Figure 3.12b mirrors the information contained in the curves of Figure 3.12a presenting the absolute value of the error of the mean value relatively to the input value. The curves in this graph quantify the required domain for a desired maximum error in the estimated mean value of theta.

Figure 3.12c presents the standard deviation of the sample of estimated thetas for each method. One observes that the ACF and VF methods, as well as their variations, exhibit an increase of the standard deviation in the small domain regime, but quickly converge to a nearly constant asymptotic value, between 0.3 and ~0.4 m, i.e. a constant coefficient of variation 20% to 25%. These limiting values are, most possibly, a result of inherent limitations of the estimation methods, limitations which are present in general in any estimation procedure of the correlation length [72]. The SV method presents a peculiar increase of the standard deviation of the estimated thetas up to D/θ_0~33 followed by a decrease for larger domains. Bartlett's limit method (ACF_B) shows a seemingly unbounded increase in the standard deviation for large domains. The MCA method exhibits a decrease of the standard deviation for increasing domain length, hence an ever-smaller variation around its (low) mean value.

Figure 3.12d shows the skewness factor of the estimated theta sample for each method. Skewness quantifies the asymmetry of the distribution of values of a certain random variable around its mean: a Gaussian distribution has zero skewness; positive skewness means that on the larger than mean values side exist less than 50% of the observation, while negative skewness means that on the larger than mean values side exist more than 50% of the observation. One observes that all methods exhibit negative skewness in the small domains regime i.e. more than half of the values are larger than the estimated mean. The ACF methods and its variations acquire positive skewness for D/θ_0~15-20 approaching a large positive value of skewness ~3 for large domains i.e. values of theta smaller than mean tend to arise more often than larger values. The VF and FF methods pass to the positive values of skewness at D/θ_0~35, while for the VRF method this occurs for D/θ_0~50, i.e. the VF method and its variations acquire positive skewness for large domains than the ACF type of methods. Hence the VF type of methods exhibit a sort of self-corrective behaviour so that their 'slower' convergence of their mean to desired value of theta, is compensated by a more frequent occurrence of thetas larger than the mean i.e. closer to the desired value. The same self-corrective behaviour is shown by the SV method, where a large positive skewness compensates the over-estimation of theta for large domains. Such a property is also shared by the MCA method, which exhibits negative skewness for all domains considered, while under-estimating the value of theta.

However, the under-estimation is so significant that the effect of skewness on the performance of the method is irrelevant.

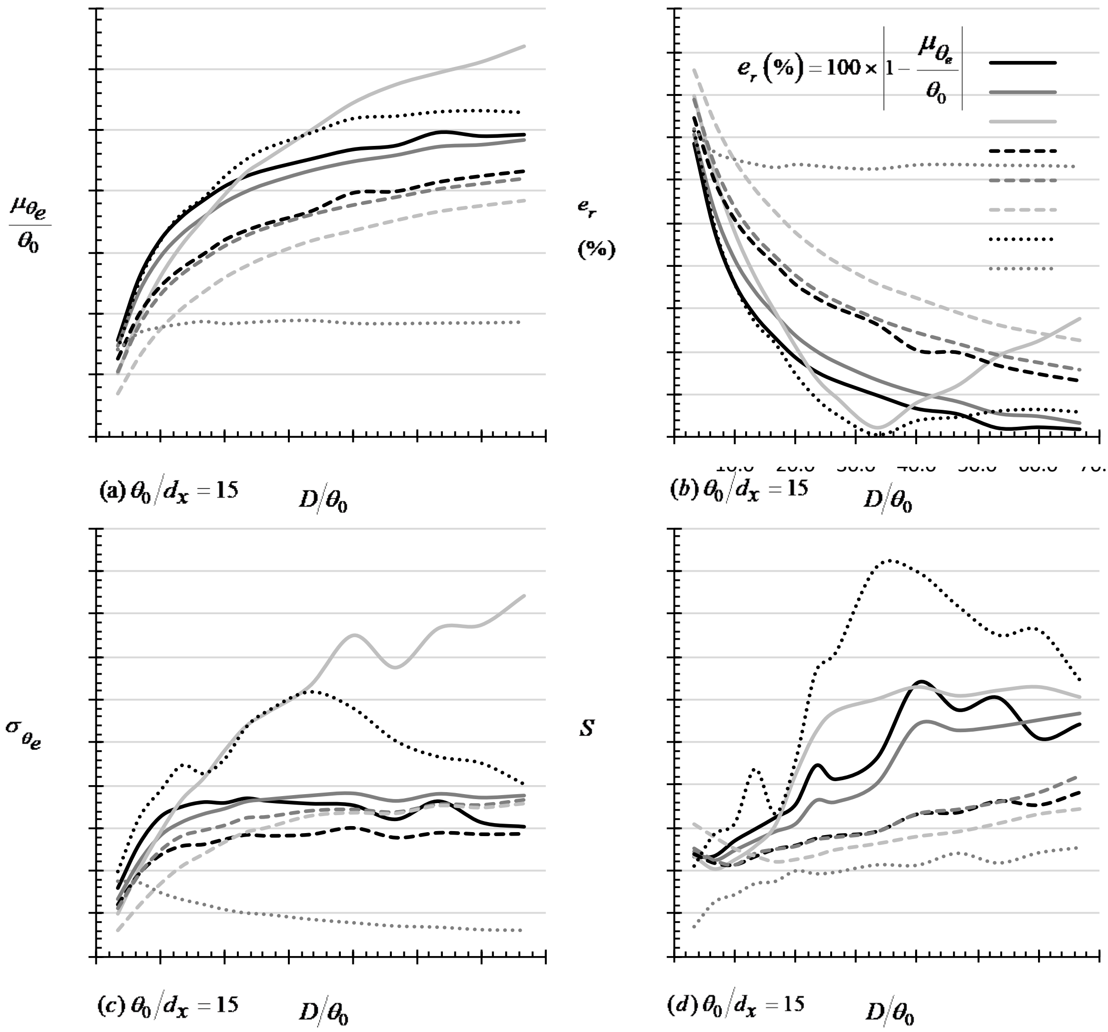


Figure 3.12: (a) normalized mean value μ_{θ_e}/θ_0 , (b) relative error e_r , (c) standard deviation σ_{θ_e} and (d) skewness S of θ_e corresponding to 5000 θ_e values for different D/θ_0 values.

Figure 3.13a presents the variation of the mean of the sample of the estimated thetas with sampling interval, expressed as a function of the ratio θ_0/d_x. First, apart from Bartlett's limit method, all other methods show a decrease of the mean with decreasing sampling interval. Secondly, except for the mean crossing method, all other methods exhibit convergence to a specific asymptotic value, shown already in Figure 3.12a; the semi-variogram method shows a small variation around the value 1. The mean crossing method shows a surprising decrease of the mean value for decreasing sampling interval. That is, same as with its behaviour for different domains lengths (Figure 3.12a), this method performs in some sense the opposite way than most other methods. Figure 3.13b merely presents the information contained in Figure 3.13a in terms of the relative error of the mean value of theta relatively to the input value.

Figure 3.13c shows the variation of the standard deviation of the sample of thetas for each method. The ACF and ACF_A methods, together with the VF method and its variations exhibit a smooth convergence of the standard deviation to a non-zero limiting value, shown also in Figure 3.12c. The SV method starts from high values of the standard deviation for large sampling intervals but quickly approaches the results obtained by the aforementioned group of the methods. The ACF_B shows a seemingly unbounded standard deviation for decreasing sampling interval. The MCA method exhibits decrease of the standard deviation for decreasing sampling interval. Recalling that the standard deviation decreases also with the domain length, one observes that the fluctuation around the mean value diminish while the sampling conditions (interval and domain) improve. That would be an ideal behaviour if the mean value was converging towards the desired value.

Figure 3.13d presents the skewness factor of the sample of the estimated thetas for each method. One observes that with the exception of the MCA method, all other methods have positive skewness, which means that they tend to produce theta estimates smaller than their mean. Consulting Figure 3.13a, that implies that the ACF and VF methods and their variations tend to produce theta estimates smaller than the desired value. Also, one observes that a common feature of the ACF and VF methods and their variations, and the MCA method, is that skewness is nearly constant for θ_0/d_x larger than (roughly) 5. Recalling Figures 13a-c, this value of the ratio θ_0/d_x can be considered as a minimum adequate value for these methods to be applied optimally. The SV method shows large positive skewness for all sampling intervals considered, which means that it tends to produce theta estimates smaller than the desired value (consulting also Figure 3.13a). Finally, the MCA method has a nearly constant negative skewness for θ_0/d_x larger than 5, that means that methods produces estimates larger than its mean value, which though is a too low to be useful. Overall, one may say that the optimum results are obtained when θ_0/d_x is larger than 5.

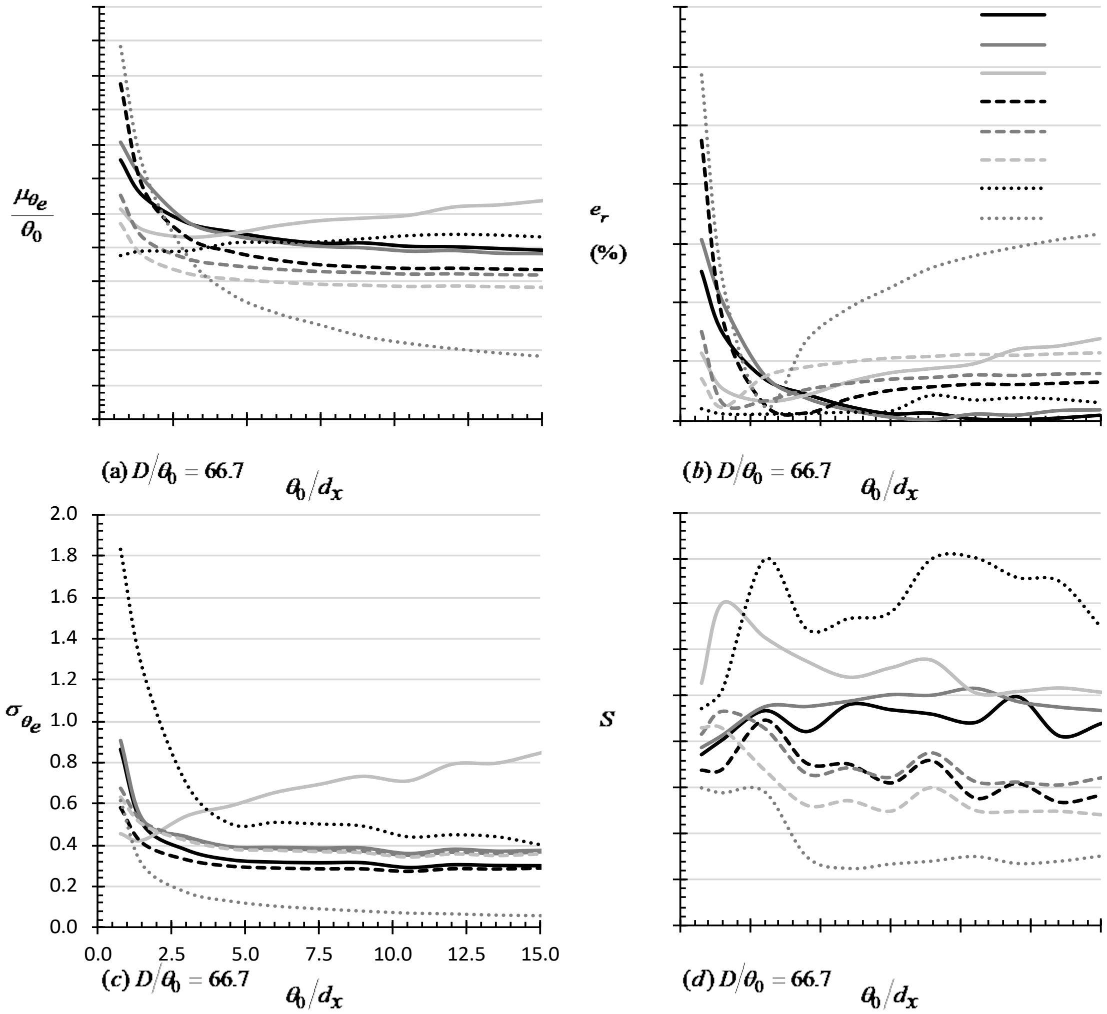


Figure 3.13: (a) normalized mean value μ_{θ_e}/θ_0 , (b) relative error e_r, (c) standard deviation σ_{θ_e} and (d) skewness S of θ_e corresponding to 5000 θ_e values for different θ_0/d_x values.

The estimated 5000 θ_e values for the case of D/θ_0= 66.7 and θ_0/d_x =15 were also plotted in histograms form in Figure 3.14 for the eight methods presented above. Each histogram contains 71 bins (following the “square root choice” rule) and refers to 5000 calculated theta values. From this figure it is generally inferred that the frequencies are log-normally distributed along the θ_e axis and that all methods except of that of MCA method, shows to have a right hand tail which means that they have a positive skewness (consulting also Figure 3.12d and Figure 3.13d).

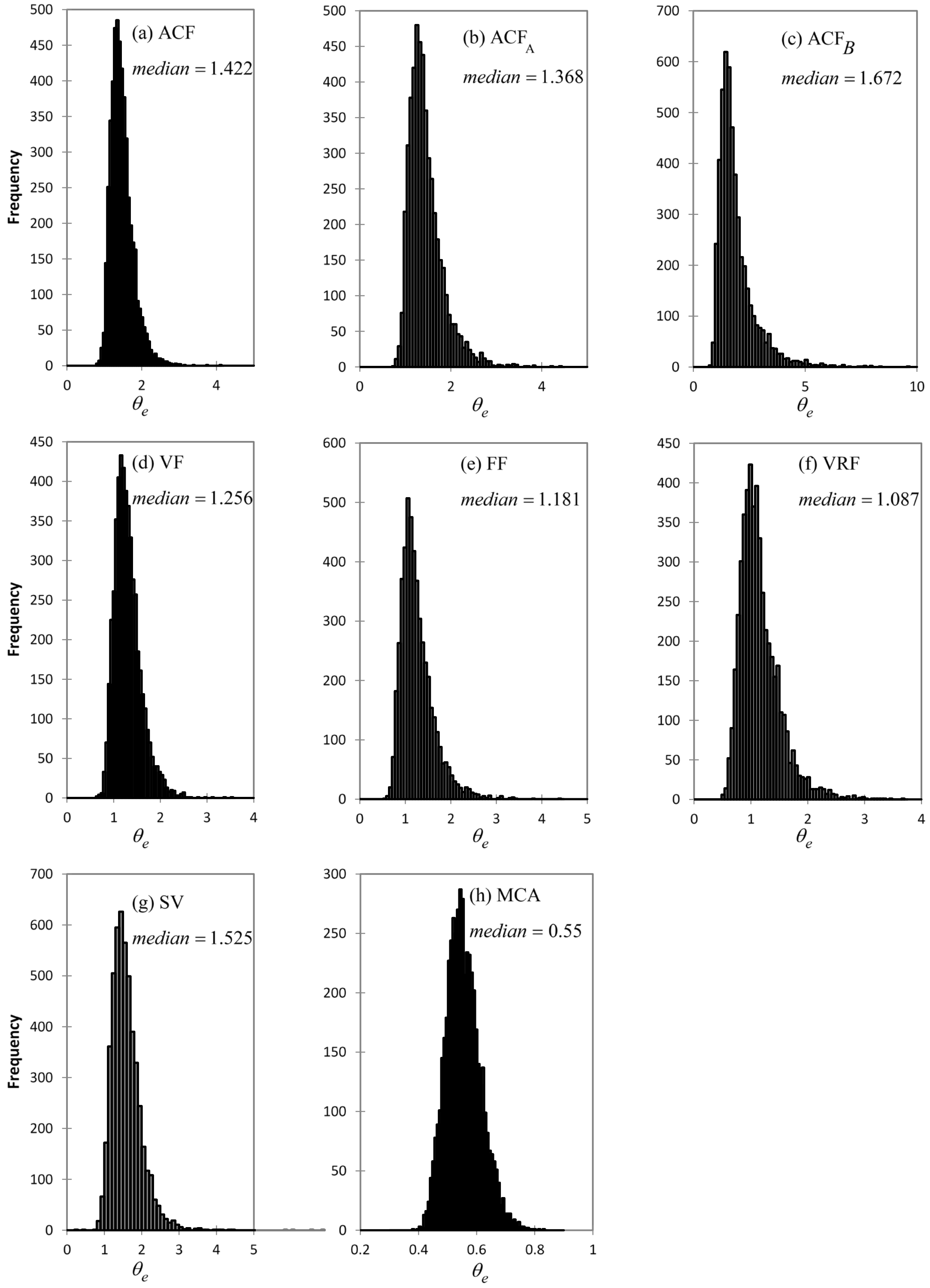


Figure 3.14: Histograms referring to the frequency of the estimated theta ($\boldsymbol{\theta_e}$) values for D/θ_0= 66.7 and θ_0/d_x=15

3.4 Summary and conclusions

In this Chapter, the effectiveness of the methods calculating the correlation length θ was investigated. This was done by generating 5000 one-dimensional random fields with the *tstlas1* program and comparing the estimated θ values with the respective one used as input value in the program. The effect of sampling domain length, sampling interval and the method used for the estimation of $\hat{\rho}(\tau)$ on the calculation of θ were systematically investigated.

The main conclusions are summarized as follows:

1. Referring to the original ACF method, $\hat{\rho}(\tau)$ should be calculated up to the point where $\hat{\rho}(\tau)$ firstly intersect the τ axis. Beyond this point the estimated $\hat{\rho}(\tau)$ behaves differently when calculated by the method of moments and the Fourier transform method, indicating that reliable information is stored only up to that point.
2. The domain length *D* strongly affects the performance of the methods. In general, larger domains improve the estimate of correlation length. In terms of the sample mean, the ACF and ACF_A methods converge to the input theta value for smaller domains than the other methods. The SV method converges to an over-estimate of the input value, while the VF type of methods exhibit a much slower convergence for increasing domain length. Regarding the fluctuations around the mean, the ACF, ACF_A and VF methods exhibit a nearly constant standard deviation for D/θ_0 larger than 15 for given sampling interval. The skewness factor of the samples produced by the ACF, ACF_A and VF methods, is positive in the large domain regime, that is, the estimated values of theta tend to be smaller than the sample mean. The opposite happens in the small domain regime. The SV method shows a peculiar behaviour regarding the standard deviation and skewness factor exhibiting peak values for D/θ_0 ~35.
3. The sampling interval d_x significantly affects the performance of the methods. In general, smaller intervals improve the estimate of correlation length. For θ_0/d_x greater than a specific value, say 7.5, the sample mean of the ACF, ACF_A and VF methods is nearly constant to the best estimate allowed by the given domain length. The SV method is nearly insensitive to the sampling interval exhibiting an over-estimation of the theta for small intervals. The standard deviation of the

predicted samples of theta for the ACF, ACF_A, VF and SV methods, becomes nearly constant for θ_0/d_x greater than 5 and the same can be observed for the skewness factor of the samples. The skewness factor corresponding to the SV method is significantly larger than the skewness of the other methods implying this method has the tendency to produce values smaller than the mean of the sample.

4 Spatial variability of soils in practice

Generally, the literature on estimating θ of soils based on real data is rather limited (e.g. [81,101,144,157]) and refers mainly to Cone Penetration Test (CPT) data as the test in question yields a large amount of closely spaced data. The Standard Penetration Test (SPT) have also been used in the past for the same reason (e.g. [54,158]), however, this test in not recommended due to the larger sampling interval [144,159]. Dynamic Probing (DP) is an alternative for indirect investigations of soils which provide continuous record with respect to depth. The objective of the dynamic probing tests is to determine the resistance of soil and soft rock in-situ to the dynamic penetration of a cone [125]. Five procedures are available according to EN ISO 22476-2 [160], DPL (Light), DPM (Medium), DPH (Heavy), DPSH-A and DPSH-B (Super Heavy), covering a wide range of specific work per blow. This chapter is a study on the vertical spatial variability of soils in practice. In this respect, a series of dynamic penetrations with a Dynamic Probing Light (DPL) apparatus were carried out in cohesive and cohesionless soils in various sites in Cyprus. The spatial correlation lengths of all readings were calculated using the $\mathrm{ACF_A}$ method and the θ values have been corrected by applying the findings presented in chapter 3. The spatial variability of clay soils has also been examined using the Unconfined Compression Strength (UCS) test on undisturbed samples (Figure 4.1b).

4.1 Estimation of the spatial correlation length of soils through dynamic probing tests

4.1.1 Correlation length of clay sites

Eight and five DPL penetrations were carried out in the wider area of Pentakomo (Limassol district; 34°43'16"N, 33°15'47"E) and Armou villages (Paphos district; 34°47'47"N, 32°28'24"E; see also Figure 4.1a) respectively. Both sites are consisted of stiff clay of high plasticity. The horizontal spacing between two successive DPL penetrations was kept constant and equal to 1.5m for both sites. The reached depth ranges from 4.9 to 5.6m for Pentakomo and 2.0 to 3.7m for Armou. EN ISO 22476-2 [160] for dynamic probing was adopted. According to this European Norm the number of blows shall be recorded every 100mm of penetration. Figure 4.4a shows the results of DPL test for the Pentakomo field.

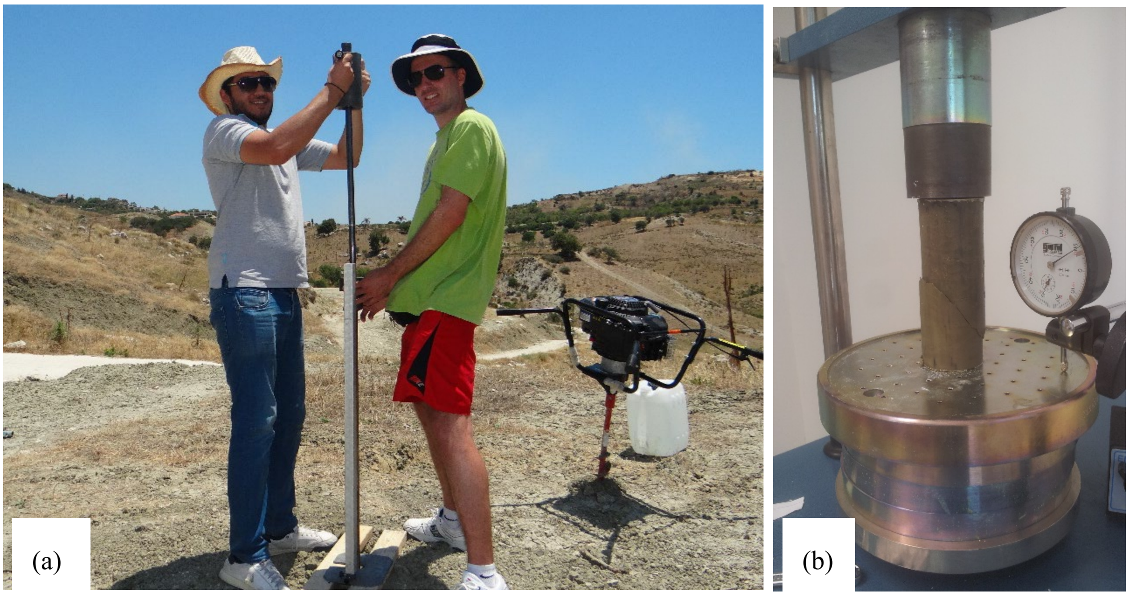


Figure 4.1: (a) DPL test performed at Armou village by Panagiotis Christodoulou (on the left) and Assistant Professor Dr. Lysandros Pantelidis (on the right), (b) Unconfined Compression Strength test of sample taken from Armou field.

An essential requirement for the application of Random Field theory for the estimation of spatial correlation length is the stationarity of the data. As Brockwell and Davis [137] stated, data are stationary if (i) the mean is constant with distance (no trend exists), (ii) the variance is constant with the distance, (iii) there are no seasonal variations and (iv) there are no irregular fluctuations. However, in practice the stationarity is defined in terms of weak stationarity, i.e. the mean and variance are constant and the autocorrelation is a function of separation (lag) distance only [56,83,137]. Besides, the data obtained from site investigation exhibit, more or less, trends and thus the data set should be transformed from non-stationarity into stationarity, simply by removing the trends using regression analysis [75,100].

In this section, the area under the positive part of Autocorrelation Function (ACF_A) method has been adopted for the estimation of the spatial correlation length θ; recall Equations (3.6) and (3.7). The values of spatial correlation length for both clay sites are summarized in Table 4.1. The reliability of θ estimations obtained were assessed by mapping the conditions i.e., sampling domain, the interval and estimated correlation length, into Figure 3.8b and d. The corrected θ values are also presented in Table 4.1. As can be seen, there is relatively small difference between corrected and uncorrected θ values. This is generally inferred from the fact that the scaled sampling domain length

(D/θ_0) of the soundings is relatively small, which as shown in Figure 3.8b results to an underestimation of θ values. In contrast, the small value of scaled sampling interval (θ_0/d_x) results to an overestimation (Figure 3.8d) of θ values. Therefore, these two conditions rule out each other and therefore no noticeable difference was observed.

It should be mentioned that these values are of the same scale with the values given by other researchers also for clay sites, e.g. Jaksa [83] for South Parklands (ADEL, AU), Akkaya and Vanmarcke [93] for Texas A&M University (TX, USA) and Cafaro and Cherubini [57] for Taranto, (IT).

Table 4.1: Estimated values of the correlation length θ of clay sites of Pentakomo and Armou.

Field	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Average θ (cm)**	**COV**
				No. of DPL penetration						
				θ (cm)						
Pentakomo	11.6	18.4	13.5	26.6	20.6	22.8	19.3	14.1	18.5	0.275
Armou	26.3	15.4	14.2	13.8	11.8	-	-	-	16.3	0.353
				Corrected θ						
Pentakomo	10.7	20.4	11.8	32.8	23.4	26.8	20.9	12.8	19.95	0.367
Armou	28.4	14.9	13.6	14.2	10.5	-	-	-	16.3	0.381

4.1.2 Correlation length of a quarry sand-heap

For the needs of the present research, the author examined the spatial variation of a quarry (Skyramont Quarries Ltd; 34°47'15.7"N, 33°09'16.3"E) sand-heap of height of 8.3m (Figure 4.2a) using the procedure described previously. The main purpose of this study was to investigate the spatial variation of a manmade soil, since its believing from the author that this type of soil may give the greatest spatial correlation length that can be found in practice. Five DPL penetrations were carried out with the horizontal spacing between two successive DPL penetrations to be at 2m (Figure 4.2b).

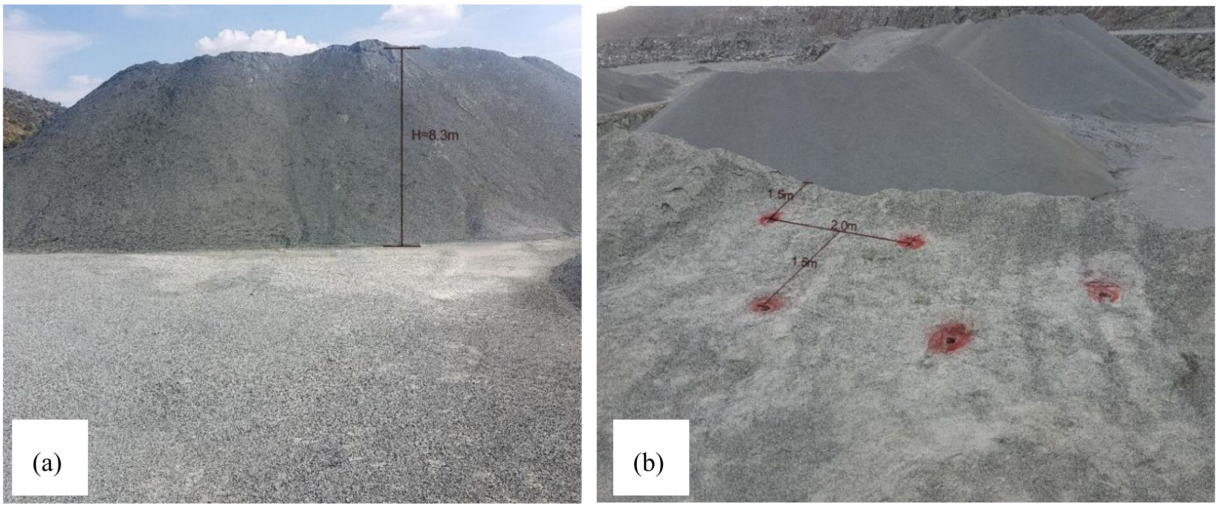


Figure 4.2: DPL test performed at sand-heap of height of 8.3m at a quarry plant.

The variation of number of blows per 100mm penetration (which is inverse analogous to the strength of soil) with the depth for the 5 penetrations performed is presented in Figure 4.3. The θ values refers to a depth up to 6m from the surface. As it is inferred from Figure 4.3, beyond this depth a different layer exists.

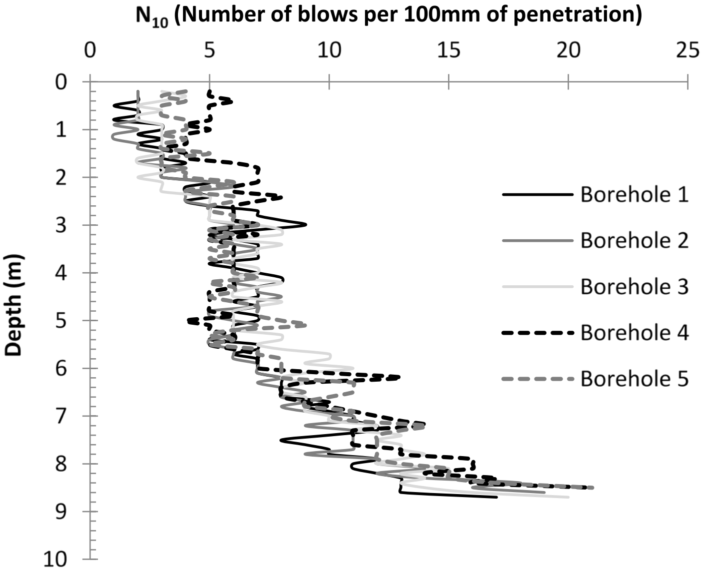


Figure 4.3: Variation in penetration resistance with DPL performed in 5 locations on a sand-heap.

The data presented above were analyzed using the ACF$_A$ method for the estimation of θ. As previously the reliability of θ estimations obtained were assessed by mapping the conditions i.e., D/θ_0 , θ_0/d_x and the estimated correlation length, into Figure 3.8b and d. In this case the θ_0/d_x values exhibit convergence to asymptotic value and hence it is concluded that the assumed d_x value leads to reliable results. On the other hand, the domain length D considered in this test is not sufficient enough and therefore it leads to an underestimation of θ values. Both corrected and uncorrected theta values are presented in Table 4.2.

Table 4.2: Estimated values of the correlation length θ of a sand-heap.

No. of DPL penetration					Average θ (cm)	COV
1	2	3	4	5		
		θ (cm)				
79.3	112.3	67.6	165.1	126.2	112.1	0.350
		Corrected θ				
111	170.1	91.9	264.2	191.8	165.9	0.370

4.1.3 Calculation of μ and σ of undrained shear strength (Su) through UCS tests

Laboratory tests were also performed on undisturbed clay samples (see Figure 4.1b) for further studying the spatial variability of cohesive soils. Continuous sampling took place adjacent to each mid DPL borehole with stainless steel tube samplers of 38.1mm internal diameter and of 200mm length. The procedure was performed in accordance with EN ISO 22475-1[161]. All soil samples along with their samplers were then submerged in water until saturation. The curing period lasted 30 days. Configuration similar to the one used in California Bearing Ratio Test [162] was used in order the saturation to take place in constant volume and not allowing soil particles to escape. This procedure allowed for the influence of moisture content on soil properties to be eliminated favouring objectivity. After the curing period, a series of Unconfined Compression Strength tests was carried out according to the procedure described in EN ISO/TS 17892-7 [163]. Figure 4.4b shows the stress-strain curves for the eight specimens of Pentakomo field, while Table 4.3 presents the undrained shear strength of specimens for both clay fields. From the results, it is inferred that the two clay materials have quite similar mean $s_u$value but the Armou clay appears to be more variable (see Table 4.3). It is reminded that the sampling was continuous, meaning that the sample No2 was immediately below No1 and so on (Figure 4.4, Table 4.3).

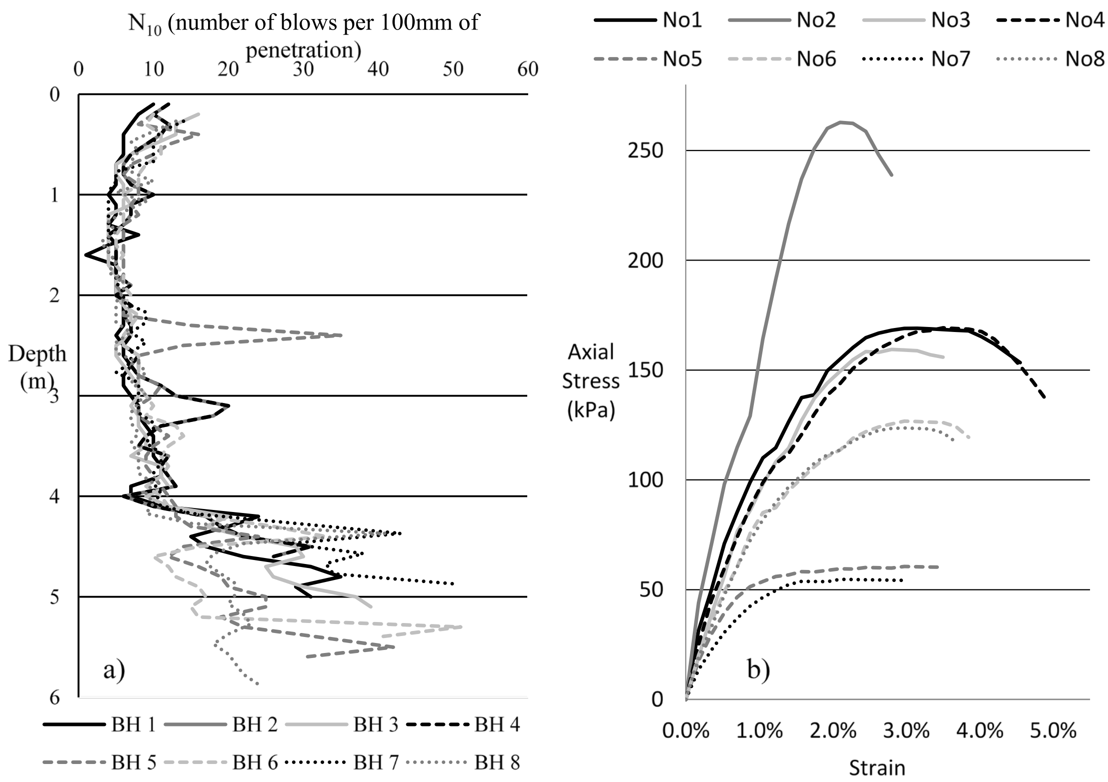


Figure 4.4: (a) Variation in penetration resistance with DPL; (b) Axial stress vs strain graph for the eight undisturbed samples from Pentakomo field.

Table 4.3: Statistical results for the undrained shear strength (s_u).

Field	**No. of sample**								μs_u **(kPa)**	σs_u **(kPa)**	**COV**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**			
	s_u **(kPa)**										
Pentakomo	84.5	131.4	79.7	84.6	30.3	63.4	27.4	61.8	70.4	33.4	0.47
Armou	39.2	43.3	148.7	52.1	36.8	-	-	-	64.0	47.7	0.74

4.2 Conclusions

The field and laboratory studies allowed the following conclusions to be drawn. Firstly, the one-dimensional random fields obtained by the DPL test allowed for the estimation of the vertical spatial correlation length, θ, of soils in various positions. For the two clay sites examined herein, θ was found to be in the order of a few tens of centimeters and more specifically it ranges from 10.7 to 32.8 cm for the Pentakomo field and from 10.5 to 28.4cm for the Armou field. In respect to the sandy field examined, the values of θ obtained are ranged from 91.9 to 264 cm. The variability of the clay sites is also indicated by the undrained shear strength values obtained in the laboratory. In every case, the spatial correlation length values calculated here are far from the $\theta = \infty$ value that might be used in a simplistic probabilistic analysis. Finally, it is interesting to be mentioned that the θ values obtained by the data from the two clay sites and from the sandy field, are in total agreement with the values given by other researchers.

5 The effect of targeted field investigation on the reliability of axially loaded piles

The present chapter deals with statistical error in the analysis of axially loaded piles. Only shaft resistance of piles is considered, as for the effective calculation of the tip resistance of a pile, soil property values referring to the very limited area affected by the tip must obviously be used.

5.1 Brief description of the procedure followed

The freely available RFEM program RPILE1D (http://www.engmath.dal.ca/rfem) has been extended as to consider sampling of soil and pile properties values. The original RPILE1D program is described in detail in Fenton and Griffiths [61,164], so only a brief description is given here. In this respect, the pile is divided into a series of elements with cross-sectional area, A, and modulus of elasticity, E_p. The axial stiffness assigned to the i-th element is, thus, the geometric average of the product $S_p = AE_p$ over the element domain. Also, the soil spring elements, which are attached to the nodes, are characterised by its initial stiffness, S_s, and its ultimate strength, U_s (bilinear relationship; see [61]). The initial stiffness is a function of soil's modulus of elasticity, E_s, while the ultimate strength is given by the following formula:

$$U(z) = p\left[ac_u(z) + \sigma_n(\mathrm{z})\tan\psi(\mathrm{z})\right] \tag{5.1}$$

The ultimate strength is simulated as a single random process due to the uncertainty on both empirical coefficients a and K[164]. In the same manner, the soil and pile stiffness (S_s and S_p respectively) are also simulated as random processes. Pile stiffness is considered as random field in the present analysis because, although a manmade material, according to experimental studies (e.g. [165–167]) the $COV$of concrete stiffness is as high as 0.1 (value equal to 0.3 has also been reported in the literature; e.g. [168]. The COV of timber stiffness is typically in the order of 0.2 (e.g. [169]); timber piles are mentioned here because driven timber piles for foundation purposes are considered in the AASHTO LRFD Bridge Design Specifications [126].

Therefore, the RFEM model consists of pile elements joined by nodes, a sequence of spring elements attached to the nodes (see Figure 5.1; [61]), and three independent 1-D random processes, described as follows:

- $S_s(z)$and $U_s(z)$ are the spring stiffness and strength contributions from the soil per unit length along the pile, and
- $S_p(z)$is the stiffness of the pile.

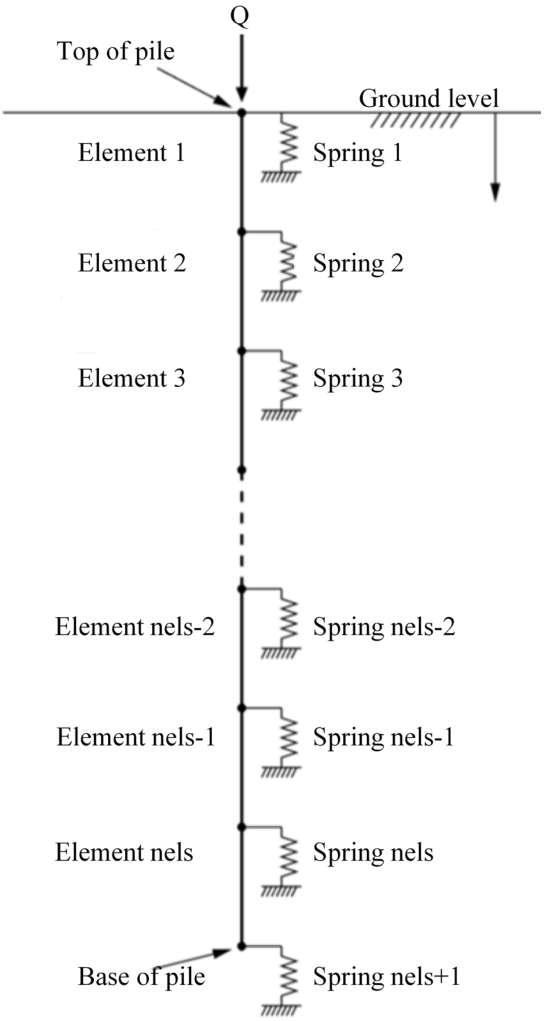


Figure 5.1: Finite element representation of pile-soil system [61].

The above one-dimensional random fields are assumed lognormally distributed having the same spatial correlation length (θ) and the same exponentially decaying (Markovian) correlation function (see [61]):

$$\rho(\tau) = \exp[-2\tau/\theta] \tag{5.2}$$

Eventually, the analysis results in the calculation of the load beyond which the given maximum settlement (δ_{max}) becomes intolerable or the ultimate load that the pile can carry just prior to failure. These two loads correspond to the serviceability and ultimate limit state respectively ($SLS$and ULS). The ULS state is defined by the sum of the ultimate strength (U_s)over all the springs, whilst, the calculation of SLS state is determined by imposing a displacement δ_{max} at the top of the pile. In general, the finite element analysis involves the computation of the spring force which yield the prescribed δ_{max}.

The RPILE1D program was modified as to:

1. have the function to virtually sample soil stiffness (S_s), soil strength (U_s) and pile stiffness (S_p) values from specific points or domains (it is noted that, the soil consists of a single material, i.e. there is no stratification; the same stands for the pile),
2. calculate the average of the sampled values (when the sampling points are more than one),
3. calculate the resistance of pile in both the SLS and ULS states, using the sampled value(s) and
4. calculate the failure probability of pile in either the SLS or the ULS state.

In addition to the above modifications, the base spring has been modified as to have the same properties with the lateral ones, meaning that the pile does not rest on firm stratum. The effect of firm stratum on the sampling strategy (see definition below) is subject matter of future investigation. The actions referring to the added features 1 to 3 are performed in each RFEM realization.

The validation of the modified program was done as follows. First a given pile was solved with the original RPILE1D program using deterministic property values for all materials. Then the same pile was solved with the modified program using values sampled from various places along the pile (since the problem is deterministic, all sampled values had the same value). The two programs gave exactly the same results for both the SLS and the ULS states, indicating that the function of sampling was embedded correctly into the original program. The rationality of the results of the present analysis are also indicative of the validity of the modified code.

The following definitions are given:

- the “optimal sampling strategy” refers to the number of sampling points and their location resulting to an optimal design. In an “optimal design” the error due to bad sampling strategy is the minimum possible (i.e. the statistical uncertainty is minimized). The error is quantified comparing the probability of failure (p_f) obtained by different sampling scenarios. The term “sampling” may refer to undisturbed specimens or to continuous probing test data (e.g. Cone Penetration Test, Standard Penetration Test).

- In each RFEM realization, “failure” is considered to have occurred when the calculated shaft resistance of pile considering spatially uniform properties sampled from the soil and pile random fields (average values), is greater than the respective one considering spatially random properties for both soil and pile.
- The “probability of failure” is defined by the fraction of the realizations resulted in failure over the total number of realizations.

5.2 Parametric study for determining the optimal sampling strategy

In the parametric study that follows, the input data used in the example presented in Fenton and Griffiths [164], are used herein as reference values (see Table 5.1). The reference pile length (L), here, is 15m. It is noted that, when no values are mentioned in the text below, the reference values are used.

Table 5.1: Input data abstracted by Fenton and Griffiths [164]**; called here as “reference data”.**

δ_{max} (mm)	Pile stiffness (S_p)		Soil stiffness (S_s)		Soil strength (U_s)	
25	μS_p (kN)	COV	μS_s (kN/m/m)	COV	μU_s (kN/m)	COV
	1000	0.1	100	0.2	10	0.2

The influence of the following parameters on the failure probability of pile has been investigated: the sampling depth (d_p) referring to a single sampling point or the sampling domain length (d_d) both measured from the surface (i.e. the uppermost point of the pile), the pile length (L), the spatial correlation length (θ), the maximum allowable settlement (δ_{max}) of pile, as well as, the pile stiffness (S_p), soil stiffness (S_s) and soil strength (U_s). The number of realizations for each RFEM model was set to 20,000 because this research deals with small differences in p_f values (this is further discussed in the Appendix II). The optimal sampling strategy, which could be a single point or a domain along the pile, is indicated by the minimum failure probability in the p_f vs depth charts.

5.2.1 Sampling soil and pile properties from a single point

5.2.1.1 Effect of spatial correlation length and pile length

The following spatial correlation and pile lengths were considered: θ =0.1, 0.2, 0.5, 1, 2, 5, 10 and 100m and L = 10, 15, and 20m respectively. It is noted that, the number of pile elements considered was always 100 for effectively finding the optimal sampling location; for comparison purposes, in their example, Fenton and Griffiths [164] considered 30 elements for a 10-meter long pile.

The variation of p_f with respect of the scaled sampling depth (d_p/L) for the various θ and L values is shown in Figure 5.2. In this paragraph, the pile and soil properties are sampled from the same depth, although the authors admit that sampling from a specific depth of a pile is unrealistic for cast in situ piles. As shown in Figure 5.2, the minimum p_f value was found near the top of the pile in the SLS state (Figure 5.2a, c and e) and at the centre in the ULS state (Figure 5.2b, d and f), the reference values given in Table 5.1 were used. It is interesting that, there is a worst-case theta value (theta giving the maximum statistical error), which is not the same for the two failure states. In addition, in the SLS state, the worst-case theta depends also on the location of the (single) sampling point. For the ULS state, it is observed that, for theta values in the order of 1 or smaller, p_f is independent of the d_p/L ratio (i.e. p_f is constant along the pile length). Moreover, in the SLS state, for the great θ values considered in the analysis, the p_f vs d_p/L curves present a characteristic minimum near the top. These behaviours are very interesting, and they are further discussed later.

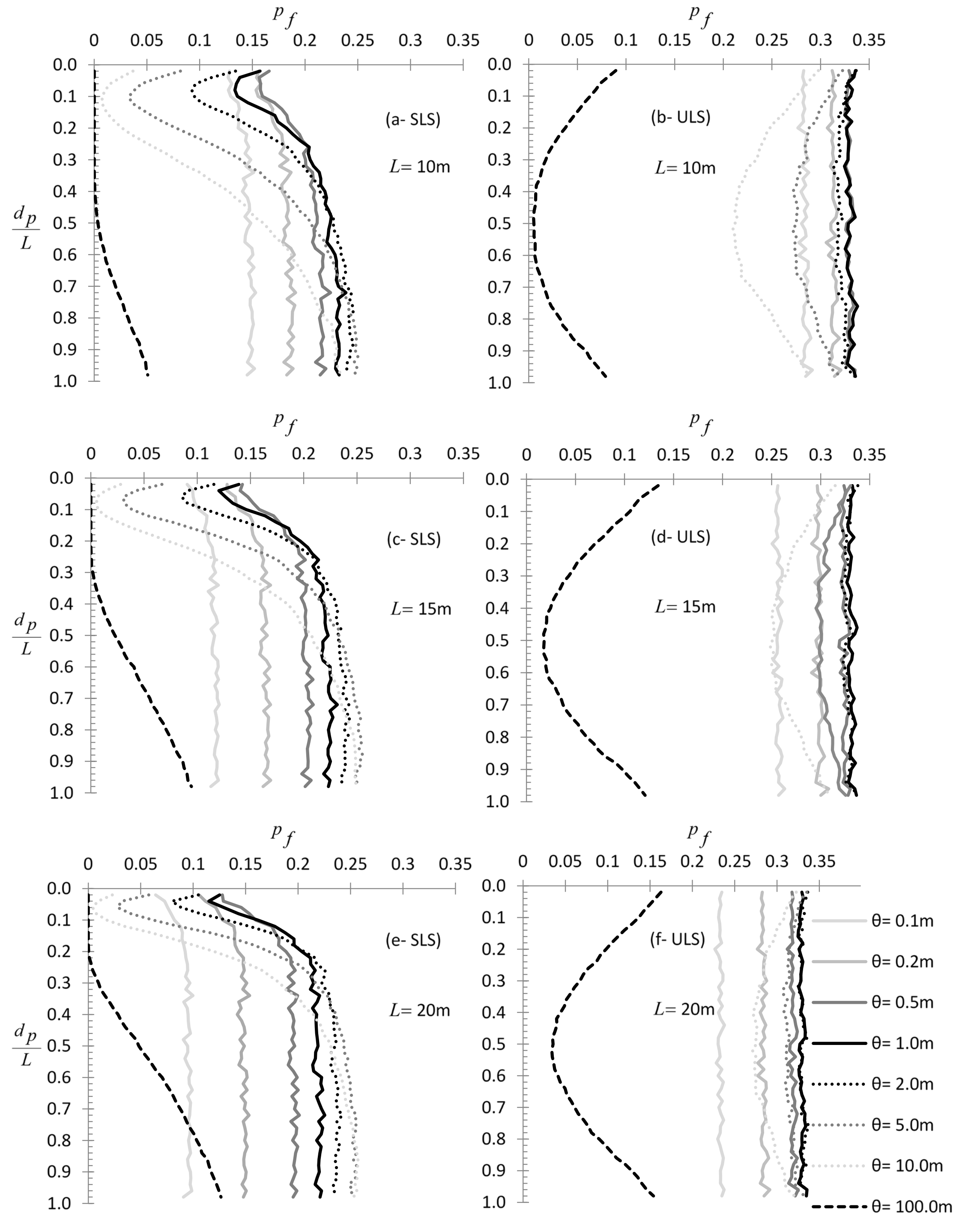


Figure 5.2: p_f vs d_p/L curves for various θ values for the case of sampling from a single point.

The effect of pile length on p_f is shown in Figure 5.3 both for the SLS and the ULS state. From Figure 5.3 it is inferred that as L increases, p_f also increases (recall the definition of p_f). Moreover, although for the ULS the minimum p_f is always found at the mid-height of pile (Figure 5.3b), for the SLS the characteristic minimum in p_f is shifted slightly to smaller d_p/L ratios as L increases (Figure 5.3a). Probably the increase of p_f with pile length is because the pile is affected by a longer random field. The fact that this increase appears to be smaller and smaller as the pile length increases is rather attributed to the fact that, a long pile (as compared to theta) meets repeating pattern of properties.

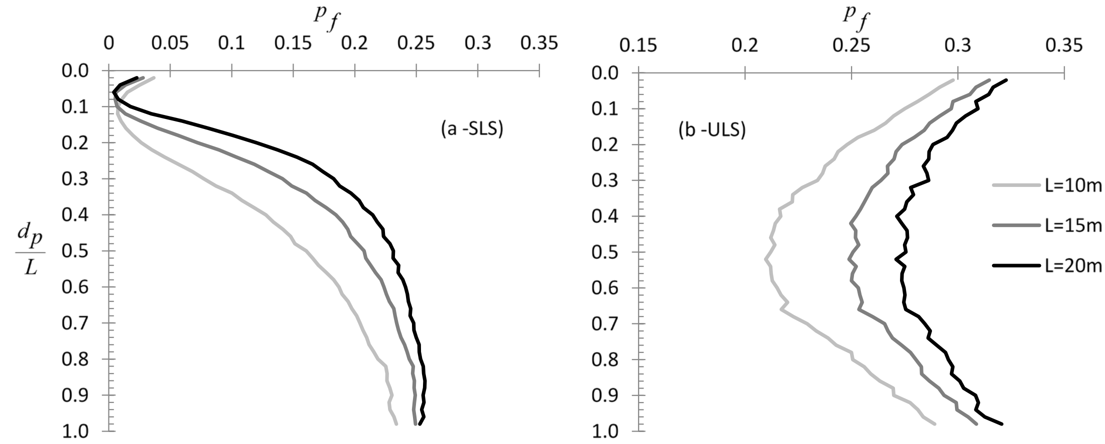


Figure 5.3: p_f vs d_p/L curves for (a) the *SLS* and (b) the *ULS*, for three different pile lengths (sampling from a single location).

From Figure 5.2 and Figure 5.3 it is inferred that, for the same pile and soil, the statistical error (expressed by the p_f), may vary significantly when the analysis is based on a single sample taken from different depths. The same is illustrated for both states in Figure 5.4, using as index the relative percentage difference (R_d) between the minimum and maximum p_f value; as shown,R_d depends strongly on θ.

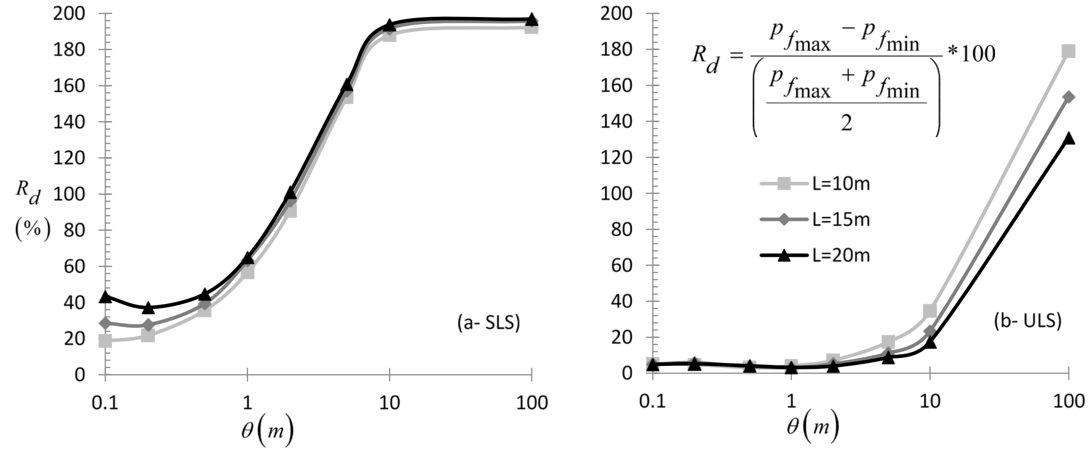


Figure 5.4: Chart indicating the relative percentage difference R_d between the minimum and maximum p_f value for θ values ranging from 0.1 to 100m and pile length L =10, 15 and 20m.

5.2.1.2 Effect of δ_{max}

The effect of the maximum allowable settlement (δ_{max}) on the optimal sampling location is examined in a similar way; apparently, this case is relevant only to the SLS state. Different δ_{max} values were considered, i.e. δ_{max} =0.025, 0.075, 0.100, 0.150 and 0.250m. The results are presented in Figure 5.5. The analysis was carried out for all input value combinations shown above; however only the scaled correlation length θ/L =2/3 case is illustrated here. In brief, it seems that, the δ_{max} value plays minor role in the location of the optimal sampling point. Moreover, the fact that the optimal sampling point is near the top is because the pile stiffness considered is relatively small. The authors add that, stronger piles (i.e. S_p>>S_s) call for sampling from the mid-height of pile.

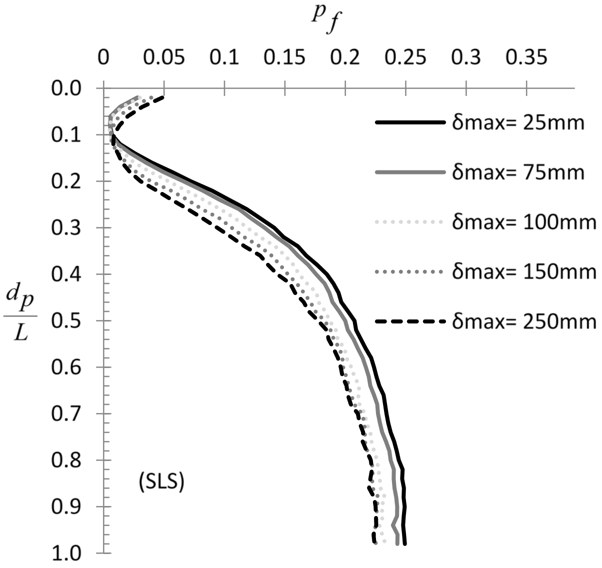


Figure 5.5: p_f vs d_p/L curves for the SLS state by considering different δ_{max} values.

5.2.1.3 Effect of pile stiffness (S_p)

The effect of pile stiffness on sampling strategy is illustrated in the p_f vs d_p chart of Figure 5.6 referring to the SLS state. Various S_p values ranging from 10^3 to 10^6 kN have been considered, whilst for all p_f - d_p curves, the θ/L ratio was equal to 2/3. From the figure in question it is inferred that the pile stiffness plays significant role in the selection of the optimal sampling location in the SLS state. More specifically, as S_p increases, the p_f vs d_p relationship tends to be symmetrical as for the mid-height of the pile; subsequently, the minimum p_f tends also to be in the mid-height of pile as S_p increases. On the other hand, as S_p decreases, the minimum p_f appears to be at shallower depths.

This is attributed to the fact that the axially loaded piles of low stiffness, (or better, of low S_p/S_s ratio) deform more near the top, while in very stiff piles the strain is distributed more evenly along their length. One should bear in mind that, excessive deformation near the top may cause either the pile or the soil to fail without influencing the underlain pile-soil system. Thus, the fact that the case of weak piles gives very low p_f values in the SLS state is only fictitious. It is noted that, the code has not yet been modified to include the pile strength; this is also subject matter of future research. In reality, piles have much greater stiffness than the surrounding soil, and thus, a failure of this kind is rather not possible. For the ULS state the optimal sampling location lies always at the mid-height of pile, exhibiting, in general, behaviour such as the respective one observed in Figure 5.8d.

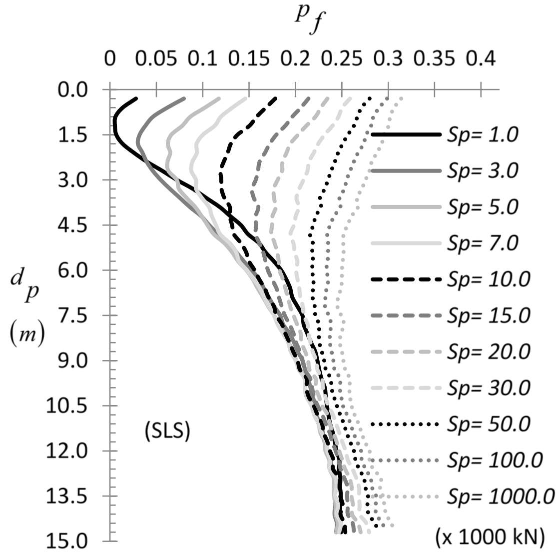


Figure 5.6: p_f vs d_p curves for different mean pile stiffness (S_p) values for the SLS state; figure referring to scaled correlation length θ/L =2/3.

5.2.1.4 Effect of soil stiffness (S_s)

The effect of soil stiffness (S_s) on the optimal sampling location has also been examined. In this respect, a number of p_f vs d_p curves for the SLS state was drawn for various S_s values ranging from 10 to 500 kN/m/m (Figure 5.7). Generally, it can be said that the soil stiffness affects the sampling location in a way similar to pile stiffness (see the previous paragraph), apparently because it is the relative difference between S_s and S_p that controls the behavior of the soil-pile system and not the absolute values of them.

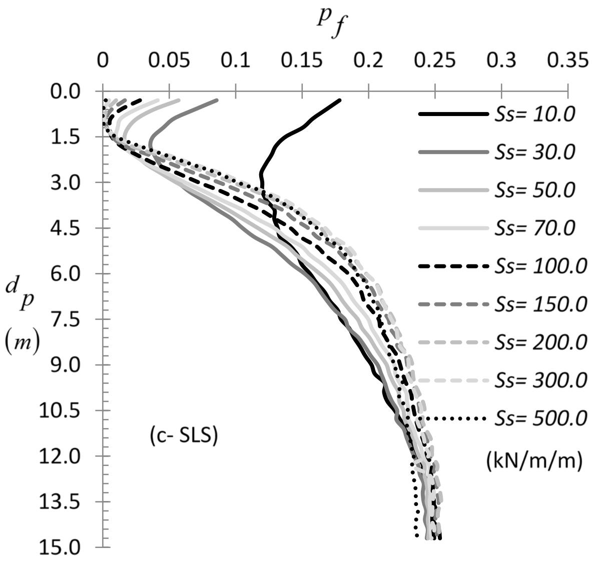


Figure 5.7: p_f vs d_p curves for different soil stiffness (S_s) values for the SLS state; figure referring to scaled correlation length θ/L =2/3.

5.2.1.5 Effect of soil strength (U_s)

The parametric study on the effect of soil strength on the optimal sampling point for both SLS and ULS states revealed that the U_s value does not affect the sampling strategy; $U_s$values ranging from 5 to 150 kN/m were examined.

Although the findings indicate that the soil strength from the parametric analysis point of view of considering different mean U_s values (while all the other parameters have been kept constant), does not affect the optimal sampling location, this is not absolutely true. The effect of soil strength on the sampling strategy should be better interpreted in relation to the effect of soil stiffness (see paragraph 5.2.1.4 of the current Chapter), because, as known, soils of high strength present also high stiffness and vice versa.

5.2.1.6 Effect of COV of pile stiffness (S_p)*, soil stiffness* (S_s) *and soil strength* (U_s)

In this paragraph the effect of COV of S_p, S_s and U_s on the optimal sampling location was examined for different COV values, i.e. COV=0.1, 0.2, 0.3, 0.4 and 0.5 with θ/L being equal to 2/3. The results are presented in Figure 5.8. From the figure in question it is inferred that, the statistical error is largely affected by the coefficient of variation of the various parameters. However, there is no influence on the location of the optimal sampling point.

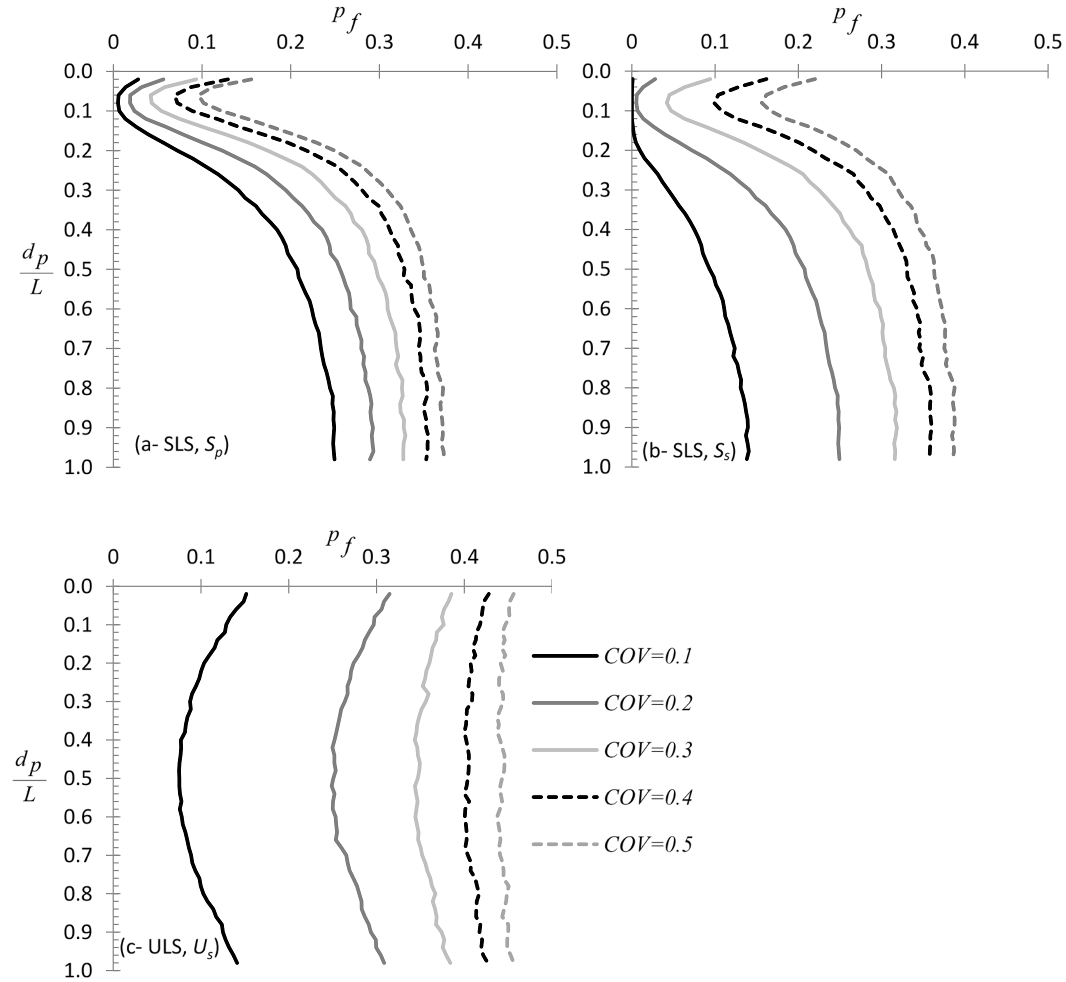


Figure 5.8: p_f vs d_p/L curves for different values of COV of (a) S_p (SLS state), (b) S_s and (SLS state) and (c) U_s (ULS state) for θ/L =2/3.

5.2.2 Sampling soil and pile properties from a domain

In this sampling strategy, sampling domains are chosen as fractions of the total pile length, L. All sampling domains are extended from the uppermost point of the pile to a certain depth; thus, the maximum sampling domain is equal to the total length of the pile. An incremental domain length of $0.1L$ was chosen in the analysis. The effect of pile stiffness (S_p), soil stiffness (S_s) and soil strength (U_s) on p_f of pile are examined separately below.

5.2.2.1 Effect of pile stiffness (S_p)

The effect of pile stiffness on the sampling strategy for the *SLS* state is shown in the p_f vs 'sampling domain length' (d_d) charts of Figure 5.9a. Generally, from this figure it is inferred that for the usual case where S_p>>S_s, the longer the sampling domain is, the smaller the statistical error is. Indeed, this error can be reduced to zero by exploiting the probing test data (e.g. CPT or SPT) along the entire pile length, if any. The same conclusion stands for any theta value (e.g. see Figure 5.9b).

Regarding the ULS state, it was found that the optimal sampling domain length is independent of the pile stiffness, since for any $S_p$value, the same p_f -d_d curves shown in Figure 5.11 are obtained. By considering, however, the entire pile length, if available, the statistical error is reduced to zero.

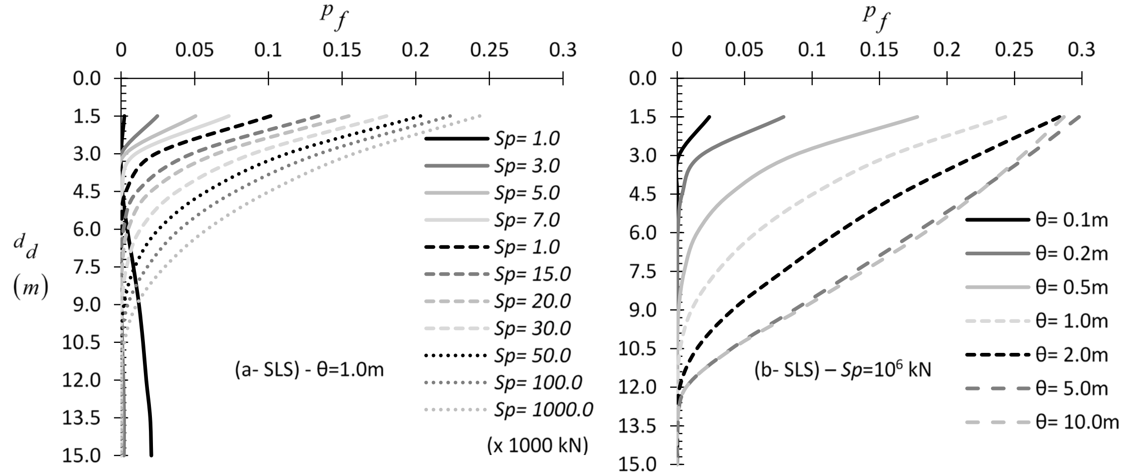


Figure 5.9: p_f vs d_d curves for different values of (a) pile stiffness (S_p) and θ =1.0m, (b) spatial correlation length θ and S_p =10^6 kN

5.2.2.2 *Effect of soil stiffness* (S_s)

Generally, it stands what it has been written for the case of S_p. The effect of $S_s$on the sampling strategy is shown in Figure 5.10a and b. From the figure in question it is inferred that, for the usual case of S_p>>S_s, the longer the sampling domain is, the smaller the statistical error is. For small S_p/S_s ratios, that is, for relatively weak piles, the optimal sampling domain length is very short, indicating a possible failure near the surface. Regarding the ULS state, it was found that the optimal sampling domain length is independent of S_s; indeed, for any $S_s$value, the same p_f -d_d curves shown in Figure 5.11 are obtained.

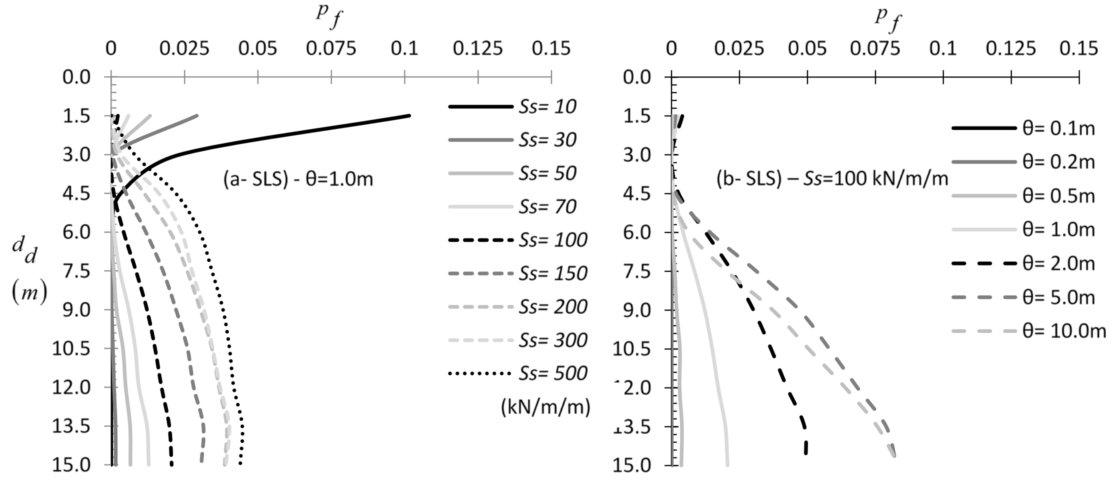


Figure 5.10: p_f vs d_d curves for different values of (a) soil stiffness (S_S) and θ =1.0m, , (b) spatial correlation length $(\theta)$and S_S =100 kN/m/m

5.2.2.3 Effect of soil strength (U_s)

The parametric study on the effect of soil strength on the optimal sampling domain length for both SLS and ULS states revealed that the U_s value does not affect the sampling strategy (Figure 5.11; see also paragraph 5.2.1.5).

5.2.2.4 Effect of δ_{max}

The effect of the maximum allowable settlement (δ_{max}) on the optimal sampling domain is shown in Figure 5.12; for the obvious reason, this case is relevant only to the SLS state. Different δ_{max} values were considered, i.e. δ_{max} =0.025, 0.075, 0.100, 0.150 and 0.250m, while the curves in Figure 5.12 refer to θ/L =2/3. In brief, it seems that, the δ_{max} value plays rather minor role in the location of the optimal sampling point. Moreover, the fact that the optimal domain length ranges between $L/5$ and $L/3$ is because the pile stiffness considered is relatively small. Stronger piles with S_p>>S_s call for d_d values equal to the entire pile length.

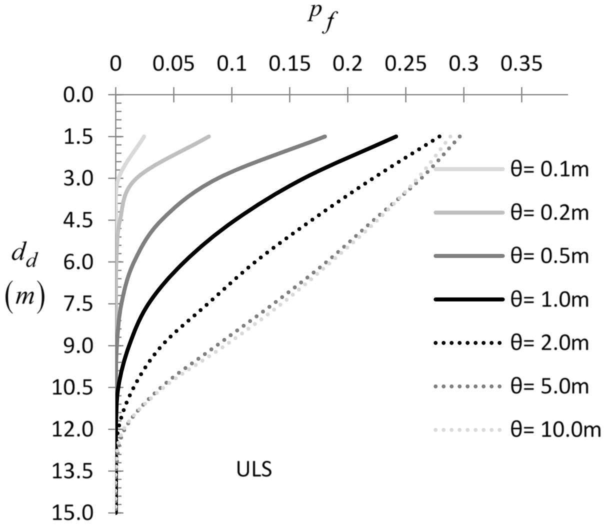


Figure 5.11: p_f vs d_d curves for different values of spatial correlation length (θ); chart standing for any S_p and S_s value.

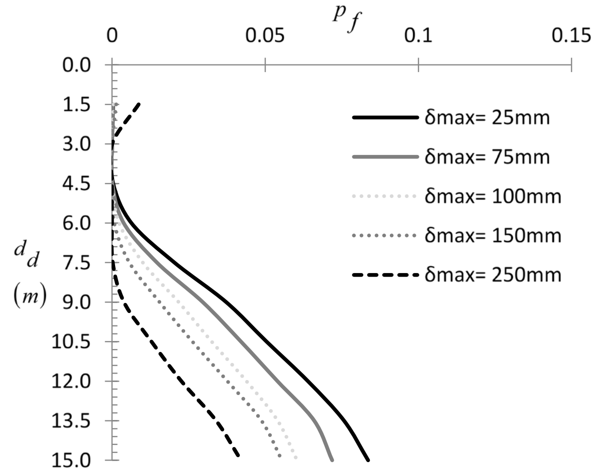


Figure 5.12: p_f vs d_d curves for various δ_{max} values; chart referring to the SLS state

5.2.2.5 Effect of COV of pile stiffness (S_p), soil stiffness (S_s) and soil strength (U_s)

In this paragraph the effect of COV of S_p, S_s and U_s on the optimal sampling domain length was examined for different COV values, i.e. COV=0.1, 0.2, 0.3, 0.4 and 0.5 with θ/L being equal to 2/3. The results are presented in Figure 5.13. From the figure in question it is inferred that, the statistical error is largely affected by the coefficient of variation of the various parameters. However, the influence on the optimal sampling domain length is rather minor.

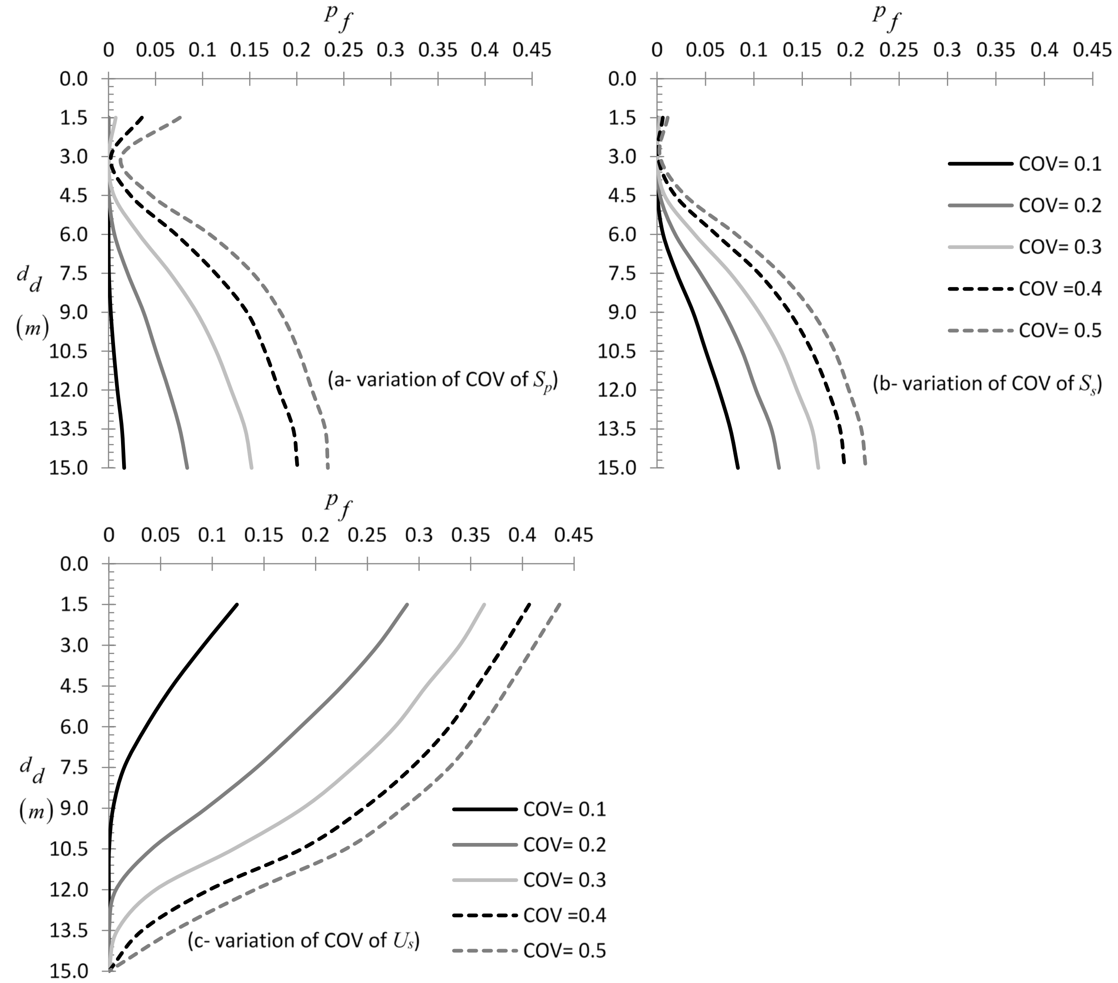
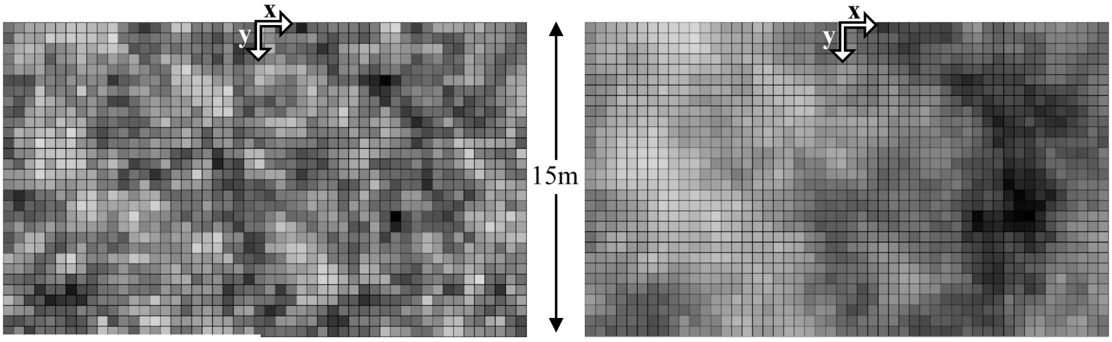


Figure 5.13: p_f vs d_d curves for different values of COV values of (a) S_p , (b) S_s and (c) U_s. Figure (a) and (b) refer to the SLS state, whilst figure (c) to the ULS state; also θ/L = 2/3 in all figures.

5.3 The importance of targeted field investigation in practice

The importance of targeted field investigation, where samples are taken from a priory known optimal locations, is highlighted here. A random material field referring to a specific RFEM realization, such the ones presented in Figure 5.14 (light areas correspond to lower friction angles and vice versa), it can be said that it convincingly represents a real field. For both examples presented in this paragraph, the various values of the parameters used are summarised in Table 5.2 (both soil materials are assumed cohesionless, whilst the unit weight of soil is considered constant throughout the soil mass).



a) $\theta/L = 1/15$ b) $\theta/L = 1/1.5$

Figure 5.14: Graphical representation of the two example random fields of friction angle of soil (recall Table 5.2). a) $\theta/L = 1/15$ and b) $\theta/L = 1/1.5$. Light areas correspond to lower friction angles and vice versa. The pile is located at $x/L = 0$.

Table 5.2: Summary of the values of parameters defining the two examples.

Example	Random field	Distribution	$\mu_{\phi'}$	$COV_{\phi'}$	L	θ/L	Figure
#1	ϕ'	Log-normal	30°	0.2	15	1/15	Figure 5.14a
#2	ϕ'	Log-normal	30°	0.2	15	1/1.5	Figure 5.14b

As known the shaft resistance of piles is given by the following equation:

$$Q_S = (2\pi R) \cdot L \cdot (K\bar{\sigma}'_O \tan\delta') \tag{5.3}$$

where, R and L are the radius and length of pile respectively, $\bar{\sigma}'_O$ is the average effective overburden pressure, K is the coefficient of earth pressure and $\delta'$is the soil–pile friction angle. For bored, cast in-situ piles, $K=K_o= 1 - \sin\phi'$ [170] and δ'/ϕ'=0.98 [171]. Rearranging the terms in Equation (5.3), after making the relevant substitutions:

$$\frac{Q_S}{(2\pi R) \cdot L \cdot \bar{\sigma}'_O} = (1 - \sin\bar{\phi}')\tan(0.98\bar{\phi}') = \Lambda \tag{5.4}$$

Thus, for a given bored pile (expressed by the dimensions L and R) and deterministic unit weight of soil, the shaft resistance of pile is directly analogous to the term shown on the right-hand side of Equation (5.4). $\bar{\phi}'$ is the mean of the values sampled (values derived from empirical correlation with Cone Penetration Test tip resistance data).

Using as reference $\bar{\phi}'$value the value corresponding to the entire length of the pile at the location where the pile is going to be constructed (x/L=0), the relative difference R_d (=Λ/Λ_{ref}-1) versus x/L charts of Figure 5.15 was drawn. It is reminded that these figures refer to the specific realizations given in Figure 5.14, which have been randomly chosen.

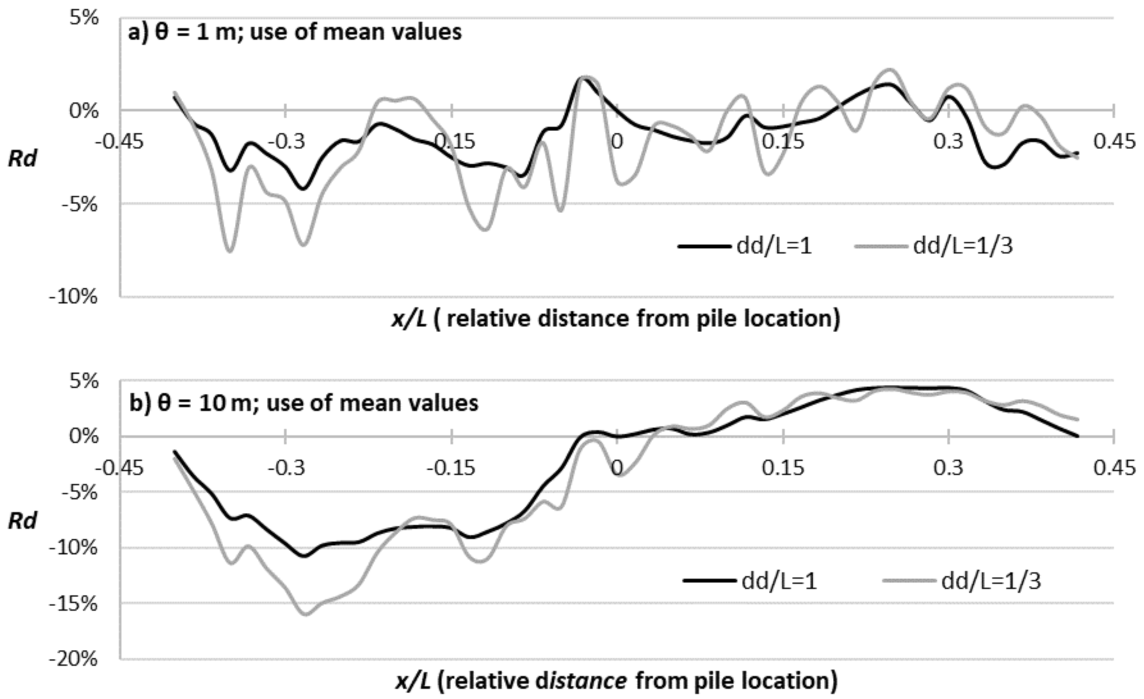


Figure 5.15: Relative difference Rd (=Λ/Λref-1; recall Equation (5.3)) versus x/L charts for a) $\theta/L = 1/15$ and b) $\theta/L = 1/1.5$. Charts drawn using characteristic friction angle values.

From Figure 5.15 it is inferred that, not only the sampling domain length strongly affects the statistical error but also the location (in plan-view) of sampling. Also, it is mentioned that, the decrease and increase of the R_d values in Figure 5.15b around *x/L*= -0.25 and 0.25 respectively are due to the existence of a weak (elements) and a strong area (light and dark elements respectively) at these locations (see Figure 5.14b).

5.4 Designing with LRFD (Load and Resistance Factor Design) codes

The discussion on the design of piles based on characteristic soil property values (please see equations (2.1) and (2.2) in Chapter 2) instead of the respective mean values is facilitated by the two example charts of Figure 5.16 (same as Figure 5.15 but with characteristic values). From the latter it is clear that, the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic values. Moreover, despite the conservatism which is inserted in the analysis using the characteristic value concept, the characteristic values alone, as shown, cannot guaranty a conservative enough study.

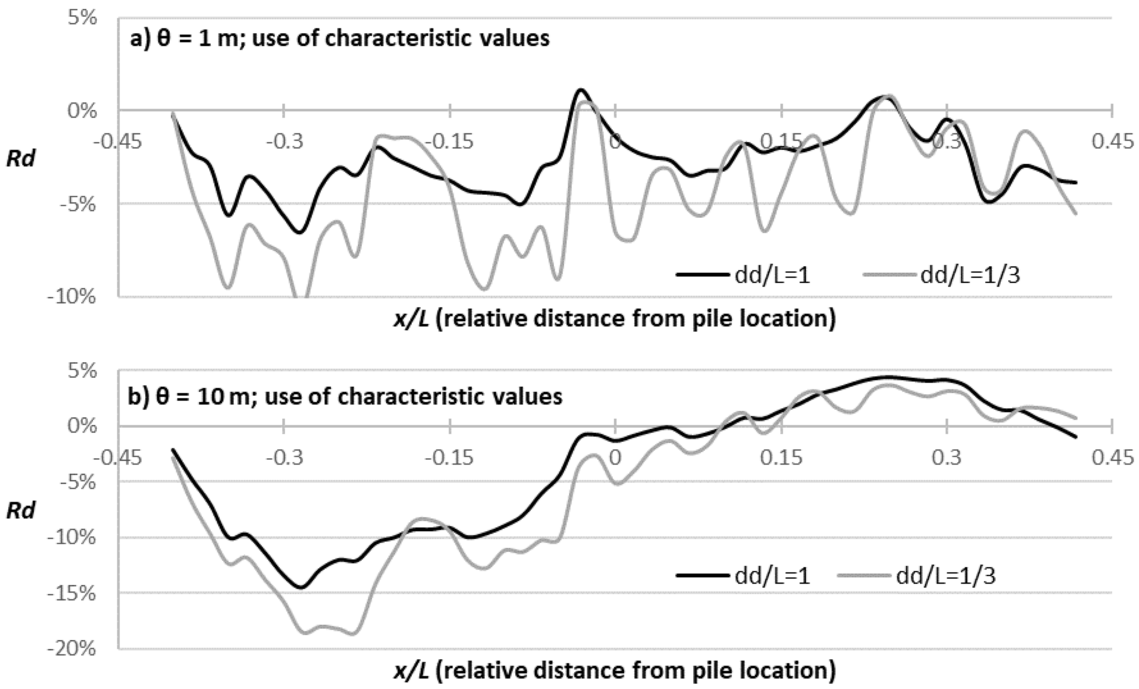


Figure 5.16: Relative difference R_d (=Λ/Λ_{ref}-1; recall Equation (5.2)) versus x/L charts for a) $\theta/L = 1/15$ and b) $\theta/L = 1/1.5$. Charts drawn using characteristic friction angle values.

5.5 Summary and conclusions

This Chapter deals with the practical problem of the effect of targeted field investigation on the reliability of axially loaded piles, aiming at an optimal design. In this respect, the freely available Random Finite Element Method (RFEM) program called RPILE1D has been modified by the authors as to consider sampling of both soil and pile properties. Two sampling strategies were considered, namely, sampling from a single point and sampling a domain, both along the pile, whilst the various parameters governing the statistical uncertainty of the problem were examined.

The analysis showed that statistical uncertainty in designing axially loaded piles can be very high and that, statistical error is not necessarily reduced by increasing the number of sampling points considered. Indeed, the opposite may happen. As shown, the statistical error can be minimized or even, eliminated by adopting the proper sampling strategy (defined by the number and location of sampling points along the pile).

More specifically, when continuous probing test data are used and the pile stiffness is much greater than the stiffness of the surrounding soil, the entire pile length is advised to be taken into account both in the SLS and the ULS state, as shorter sampling domain lengths may insert great statistical error. Also, weak piles, such as, timber piles, require much shorter domain lengths (measured from the top), as longer sampling domains may increase the error.

On the other hand, when the design is based on sampling points and not on continuous probing test data, the best practice for minimizing the statistical error is sampling from the mid-height of the pile in both SLS and ULS states. However, for the SLS state, when the pile is relatively weak, the optimal sampling point lies near the top of the pile. For the ULS state, the optimal sampling point lies at the mid-height of pile, for any pile stiffness.

Another main finding is that, the optimal horizontal sampling location is at the exact location where the pile is going to be constructed. In addition, it was observed that, the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic soil property values. Moreover, despite the conservatism which is inserted in the analysis using the characteristic value concept, the characteristic values alone, as shown, cannot guaranty a conservative enough engineering study. The safety level can be increased by applying a statistical uncertainty partial factor (similar to the model factor γ_R used by Eurocode 7) or a unified and more conservative model factor to the resistances, which will absorb the statistical uncertainties related to the soil.

6 The effect of targeted field investigation on the reliability of earth retaining structures

This chapter investigates numerically the effect of targeted field investigation on the reliability of earth retaining structures in active and passive state. This is done based on the RFEM [61], properly considering soil sampling in the analysis. The present analysis will show that the statistical error in an active or passive earth pressure analysis can be minimized only by targeted field investigation. Apparently, this study refers to structures, such as, sheet pile and bored pile walls, retaining undisturbed soil and not to backfilled retaining structures. As stated in the introduction, the current design codes are limited to some general recommendations, focusing mainly on the extend of the subsurface exploration and aiming at identifying possible unfavorable geological conditions. In this respect, AASHTO [126] additionally recommends that samples be taken from locations alternating from in front of the wall to behind the wall. Apparently, this recommendation refers to structures such as sheet walls and pile walls, where undisturbed soil is retained and both active and passive states exist simultaneously. No recommendation is given for the sampling depth or distance from the wall.

6.1 Brief description of the RFEM program used

The open source RFEM program REARTH2D (see Fenton and Griffiths [61] and Griffiths et al. [49]) is used and modified suitably in order to accommodate the purposes of the present research. The program involves generation and mapping of soil properties (cohesion, friction angle and/or unit weigh; at least one of these parameters is required to be random) onto a finite element mesh. For a specific set of material random fields, the program returns the wall reaction force and overturning moment caused by the self-weight of a spatially random soil. In addition, from the same set of random fields, the REARTH2D program is able to sample soil property values for calculating the respective wall reaction force and overturning moment based on Rankine's [172] earth pressure theory considering that the soil medium is homogenous and characterized by the sampled values (mean of the values sampled for each soil property). The procedure is repeated m times; m is the number of realizations, where, each realization refers to a new set of random fields for c', ϕ' and γ. Then, the failure probability (p_f) of the wall against sliding or overturning is calculated. The “failure probability” is defined by the fraction of the number of realizations resulted to a specific type of failure (sliding or overturning) over the total number of realizations. In the active state and for each RFEM realization, “failure” is considered to have occurred when the “actual” wall reaction force (or overturning moment) referring to the spatially random soil (value calculated using the RFEM method), is greater than the respective (factored or unfactored) predicted value referring to soil having spatially uniform properties sampled from the RFEM random fields (value calculated based on Rankine’s earth pressure theory as mentioned above). That is, it stands that

$$p_f = P\left[X_{a,RFEM}/FS > X_{a,Rankine}\right] \text{ or } P\left[X_{a,\text{"actual"}}/FS > X_{a,predicted}\right] \tag{6.1}$$

where, the symbol X denotes either wall reaction force or overturning moment (F and M respectively), the subscript a denotes active state of failure and FS is the user-defined safety factor. In the passive state, “failure” is considered to have occurred when the “actual” wall reaction force (or overturning moment), is less than the respective (factored or unfactored) predicted value and therefore, it stands that

$$p_f = P\left[X_{p,RFEM} < X_{p,Rankine}/FS\right] \text{ or } P\left[X_{p,\text{"actual"}} < X_{p,predicted}/FS\right] \tag{6.2}$$

where, the subscript p denotes passive state of failure.

In the REARTH2D program the active and passive states are reached incrementally. However, in practice, retaining structures do not always work under large wall movements; thus, the active or passive state may not be fully reached. Ni et al. [173] highlighted the importance of the intermediate active or passive state in design. Reducing the increments in the finite element analysis and thus not allowing the active or passive state to be fully reached, the authors observed, as expected, greater wall reaction forces for the active state and smaller for the passive state, but no change in the optimal sampling location. Thus, avoiding any confusion, the analysis that follows refers to fully reached active and passive states.

Favoring objectivity in the comparison between the “actual” and the respective “predicted” values, the original REARTH2D program has been modified so as to calculate the wall reaction force and overturning moment based on the finite element method instead of Rankine’s theory. That is, the failure probability for the active state is defined as follows:

$$p_f = P\left[X_{a,RFEM}/FS > X_{a,FEM}\right] \text{ or } P\left[X_{a,\text{"actual"}}/FS > X_{a,predicted}\right] \tag{6.3}$$

and for the passive state as:

$$p_f = P\left[X_{p,RFEM} < X_{p,FEM}/FS\right] \text{ or } P\left[X_{p,\text{"actual"}} < X_{p,predicted}/FS\right] \tag{6.4}$$

The finite-element earth pressure analysis in REARTH2D uses an elastic, perfectly plastic Mohr–Coulomb constitutive model with stress redistribution achieved iteratively using an elasto-viscoplastic algorithm essentially similar to that described in the text by Smith and Griffiths [174]. The boundary conditions on the right side of the mesh (across the wall) are such that they allow vertical but not horizontal movement, while the base of the mesh is fully restrained. The top and left sides of the mesh are unrestrained, except for the nodes adjacent to the wall, which are as described immediately below.

The active state against sliding is modelled by translating the nodes of the mesh next to the wall horizontally and uniformly, away from the soil. While, the passive state against sliding is modelled by translating the nodes towards the soil mass. These nodes have fixed horizontal components of displacement.

As active (or passive) conditions are mobilized, the vertical components of these displaced nodes are either free to move down (or up for the passive state) or restrained depending on whether a perfectly smooth or perfectly rough wall is modelled. Following each increment of displacement, the mobilized active (or passive) reaction force on the wall is computed by integrating the stress in the elements attached to the displaced nodes. Considering a rough, rigid wall, the active (or passive) state against rotation is modelled by imposing the same angular displacement to the nodes next to the wall, having as pivot point the lower point of wall; in this respect, the (cross-sectional) width of the wall is considered to be infinitesimally small. For smooth rotating walls, these nodes are allowed to slip downwards (or upwards for passive state) along the wall surface. The translation or rotation of the wall is performed incrementally. For the active state, the finite element analysis is terminated when the incremental displacements have resulted in the active reaction force reaching its minimum limiting value, while for the passive state, is terminated when the passive reaction force flattens out and reaches a maximum value.

The sub-chapters that follows examine the case of a wall retaining a fully drained cohesionless soil against active and passive failure in plane strain conditions. Different sampling strategies are examined, i.e. sampling from a single point and sampling from a domain, through a parametric analysis for defining the strategy that minimizes the probability of failure and thus the statistical error (i.e. the optimal sampling strategy).

In both studies (i.e. active and passive state), both ϕ' and γ are treated as log-normal random fields. The soil is assumed cohesionless with $\mu_{\phi'}$ =30° and μ_{γ} =20 kN/m^3, while various standard deviation and θ values are examined. Moreover, since K_o depends on ϕ' (the initial horizontal stresses are defined by Jaky's [175] $K_o = 1 - sin\,\phi\,'$ equation), this is also treated as random field. In addition, it is mentioned that, although the elastic parameters of soil (ν and E) affect the required wall movement in order the active or passive state to be fully reached, preliminary parametric analysis carried out by the author showed that they have no influence on the optimal sampling strategy. Thus, these values have been kept constant and equal to 0.3 and 10^5 kN/m^2 respectively throughout the entire analysis presented herein. Furthermore, the random fields are assumed to have the same spatial correlation length (θ) and an exponentially decaying (Markovian) correlation structure (see Equation (5.2)). Finally, a safety factor *FS* equal to 1.3 is generally assumed in the active state analysis, while a *FS* equal to 1.25 for the passive state; the effect of *FS* on the sampling strategy is also investigated later.

6.2 Active state of stress

6.2.1 Parametric study for determining the optimal sampling strategy

The soil mass is discretised into a 48x34 mesh (number of elements at the horizontal and vertical direction respectively) consisting of eight-noded square elements with side length equal to 0.1 (Figure 6.1). Various wall heights are considered ranging from *H*=1.4 m to 2.9 m (meaning that the wall extends to a depth ranging from 14 to 29 elements), while the mesh geometry has been kept the same for all cases. A 24-element wall will be generally considered in the analysis (hereafter called a "reference wall"); the other wall heights will be used for the investigation of the effect of wall height on the optimal sampling strategy. The 48-element mesh in the horizontal direction was chosen so that the failure mechanism in the RFEM analysis will not to be affected by the proximity of the right boundary. In this respect, for the highest wall considered (29-element-high wall), as the ϕ' value approaches zero (extreme case), according to Rankine's theory, the failure mechanism in the active state will occupy a horizontal distance from the wall face equal to one wall height (that is, 29 elements; value 40% smaller as compared to the 48 horizontal elements of the geometry).

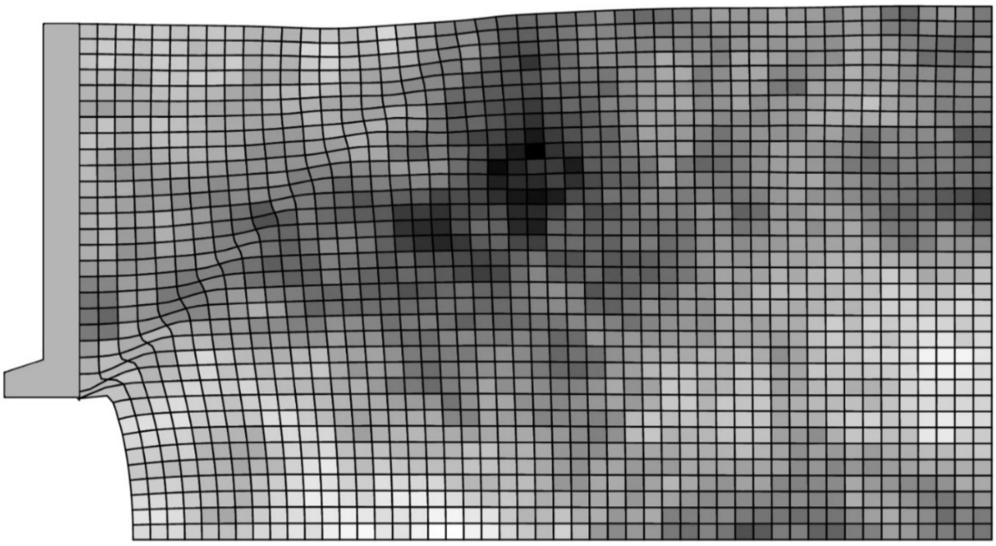


Figure 6.1: Active earth failure of the "reference" wall. Graphical representation of a random field of ϕ' (this is a typical RFEM realisation); light areas correspond to lower friction angles and vice versa. For the soil shown, $\theta/H = 8.3$ and COV_{φ} = 0.3.

As samples are taken from a material field (i.e. the ground), which simultaneously is a stress field (stresses caused by the self-weight of the soil and also, any external load), the location of the optimal sampling points is affected by the coexistence of these two fields. Aiming at finding the optimal sampling strategy, the following parameters will be examined: the sampling depth (d_p) and horizontal distance (x) for the case of sampling from a single point (measured from the soil surface and the wall face respectively), the sampling domain length (d_d) and horizontal distance (x) of the continuous probing test location for the case of sampling from a domain (measured as in the previous case), the spatial correlation length of soil (θ), the wall roughness (perfectly smooth or perfectly rough wall), the wall height (H), the coefficient of variation (COV) of ϕ' and γ, the mean value of ϕ', the safety factor value (FS) considered and the soil mass anisotropy $(\theta_h \neq \theta_v)$. Hereafter, it is noted that the symbol θ (that is, without subscript) denotes isotropic conditions $(\theta_h = \theta_v)$. Four sampling scenarios are indicatively shown in Figure

6.2 (scenarios A and B refer to a single sampling point, whilst scenarios C and D to continuous probing tests).

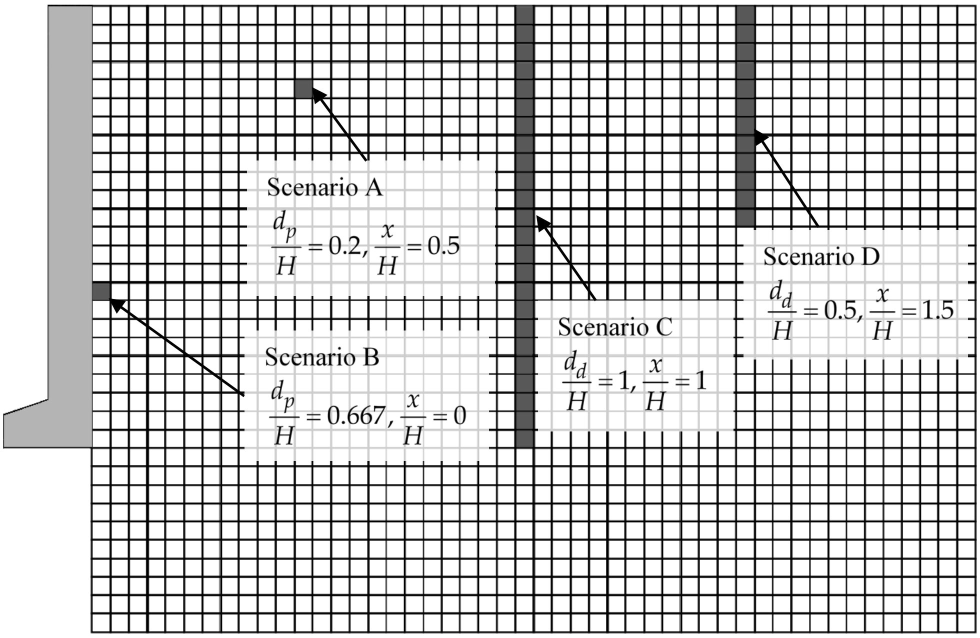


Figure 6.2: Graphical representation of different sampling scenarios: Scenarios A and B refer to a single sampling point (each located at depth d_p), whilst Scenarios C and D to sampling domains (each having length d_d).

The optimal sampling point or domain will be identified by comparing the failure probability (p_f)values derived by various sampling strategies. Apparently, when dealing with small differences in p_f values, the stability of the results is very important. In this respect, the number of realizations was set equal to 3000; this number, as discussed in

Appendix III, can be considered adequate for the needs of the present research. The effect of the element size is also examined in the same appendix.

6.2.2 Sampling from a single point

6.2.2.1 Effect of spatial correlation length (θ)

Example charts showing the variation of p_f with respect to d_p/H for various θ/H values are given in Figure 6.3. From this figure it is inferred that, the optimal sampling location for the active case is clearly for zero horizontal distance from the wall, both for the case of translating and rotating wall and for any θ value. Isolating the curves for $x/H = 0$ (see Figure 6.4) it seems that, there is a "worst case spatial correlation length"—see also [176], where the failure probability becomes maximum; for example, for the various θ/H ratios shown in Figure 6.4 (ranging from θ/H =0.04 to 8.3), the θ/H =0.21 case gives the higher p_f values. From Figure 6.4 it is also inferred that, as θ tends to zero, the p_f value tends to a single value for any depth (that is, p_f is independent of the sampling depth). However, as θ increases, the p_f value becomes more dependent on the sampling depth.

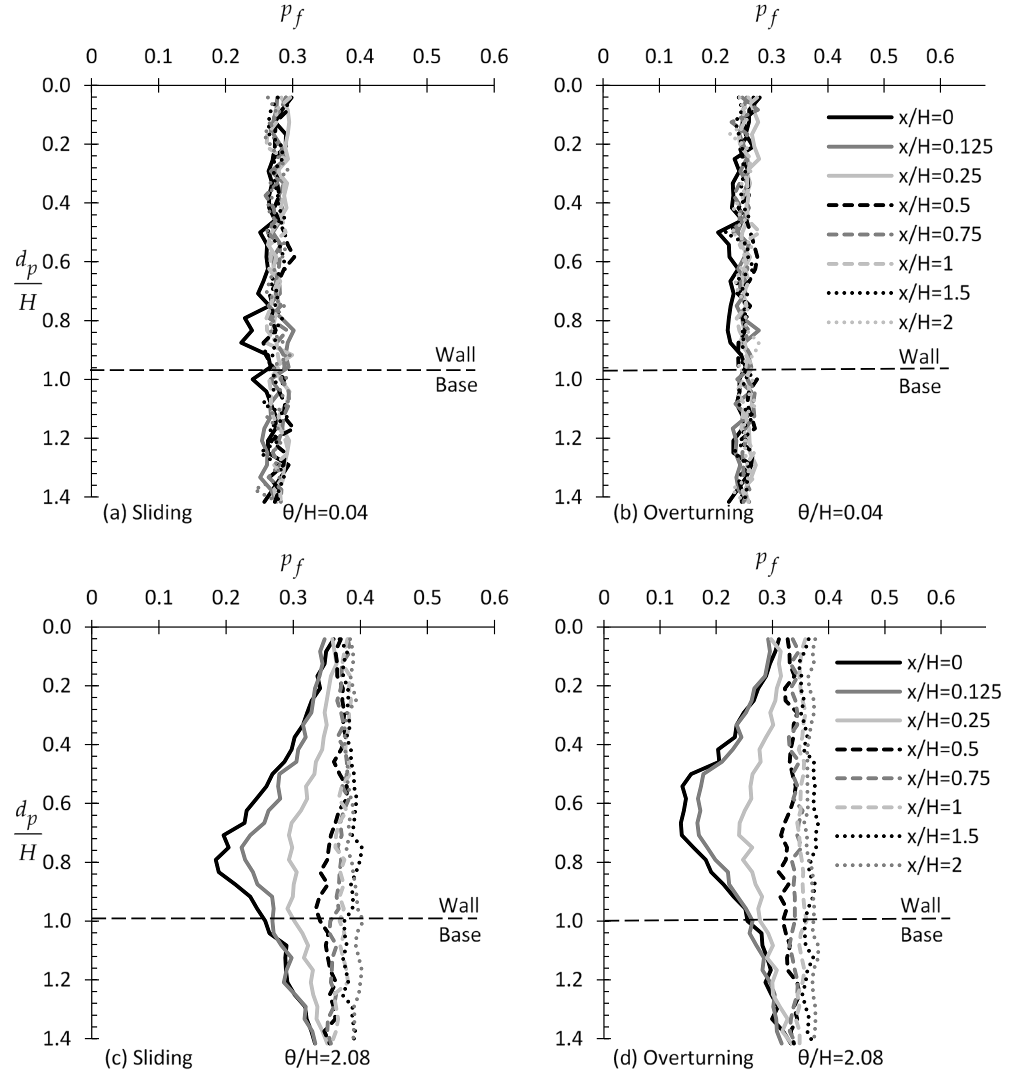


Figure 6.3: p_f versus d_p/H example curves for various θ/H and x/H values for the case of sliding (figures (a), (c) and (e)) and overturning wall (figures (b), (d) and (f)).

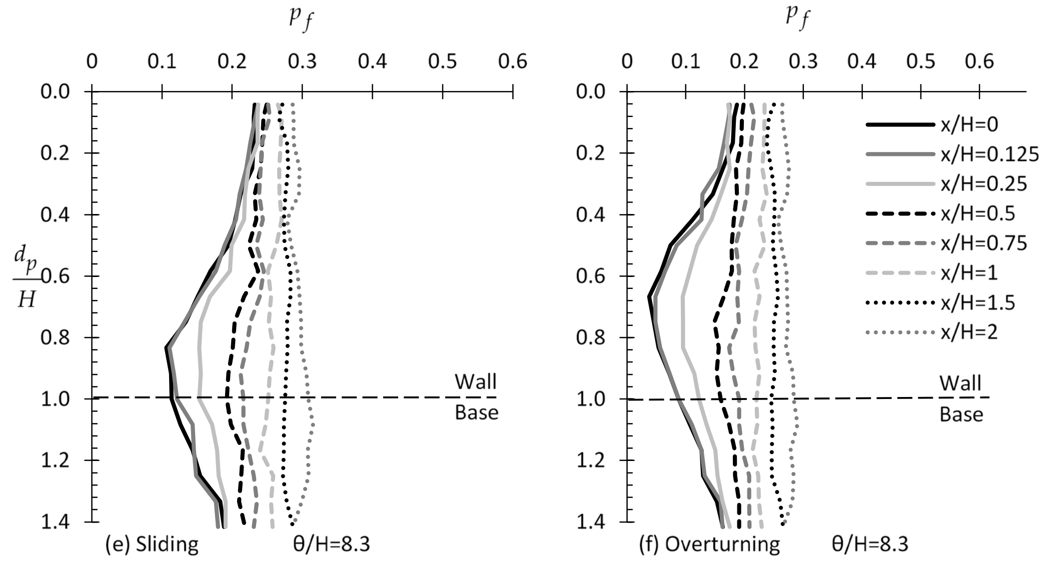


Figure 6.3: Continued.

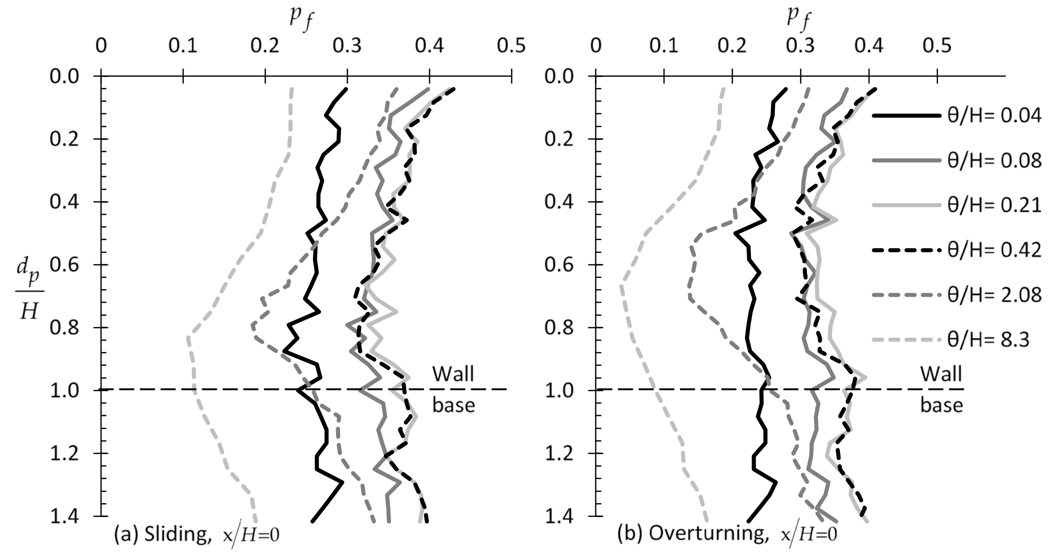


Figure 6.4: p_f versus d_p/H example curves for various θ/H and $x/H = 0$ (smooth wall) for the case of (a) sliding and (b) overturning wall.

6.2.2.2 Effect of wall roughness

The effect of wall roughness on the location of the optimal sampling point is examined herein. The optimal sampling distance was also found to be at x/H =0, therefore, results are presented only for this case (see Figure 6.5). Comparing Figure 6.4 with Figure 6.5 (for perfectly smooth and perfectly rough wall respectively) it is inferred that, the statistical error is, generally, less sensitive to the sampling depth in the case of rough wall. However, comparing similar soil-wall systems but with different wall roughness it can be said that, by choosing the proper sampling depth, smaller p_f can be obtained when the wall is smooth.

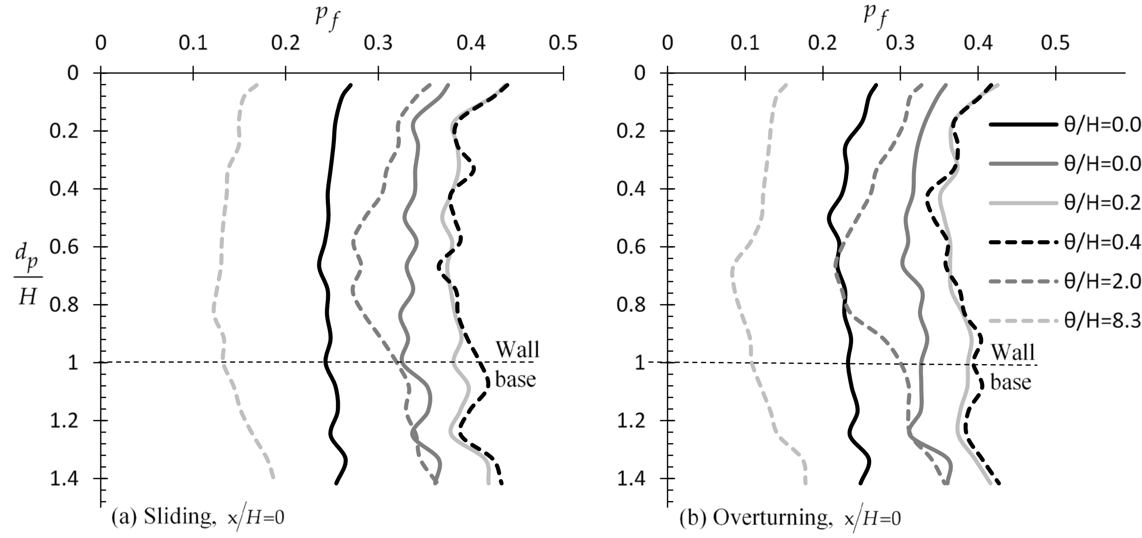

(a) Sliding, $x/H=0$ (b) Overturning, $x/H=0$

Figure 6.5: p_f versus d_p/H example curves for various θ/H values and $x/H = 0$ (perfectly rough wall) for the case of (a) sliding and (b) overturning wall; please compare with Figure 6.4 (perfectly smooth wall).

6.2.2.3 *Effect of wall height*

The variation of p_f with d_p/H for various wall height values , i.e. H =1.4, 1.9, 2.4, 2.9 m, is shown in Figure 6.6. From this figure it is clear that, the wall height has minor influence on the location of the optimal sampling point.

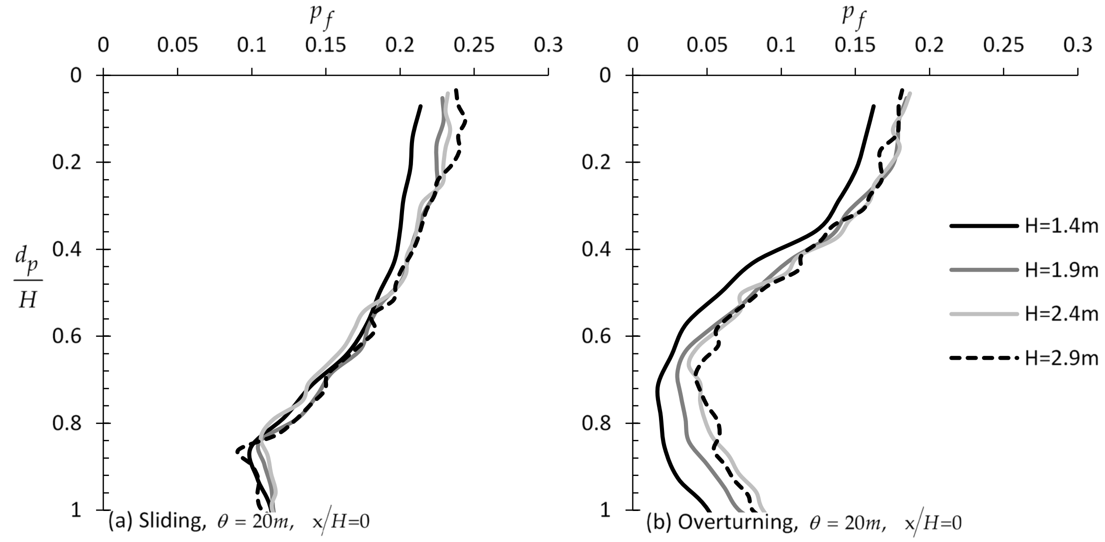

(a) Sliding, $\theta = 20m$, $x/H=0$ (b) Overturning, $\theta = 20m$, $x/H=0$

Figure 6.6: p_f versus d_p/H example curves for different wall heights, H, and θ =20 m for the case of (a) sliding and (b) overturning wall.

6.2.2.4 *Effect of COV of ϕ' and γ*

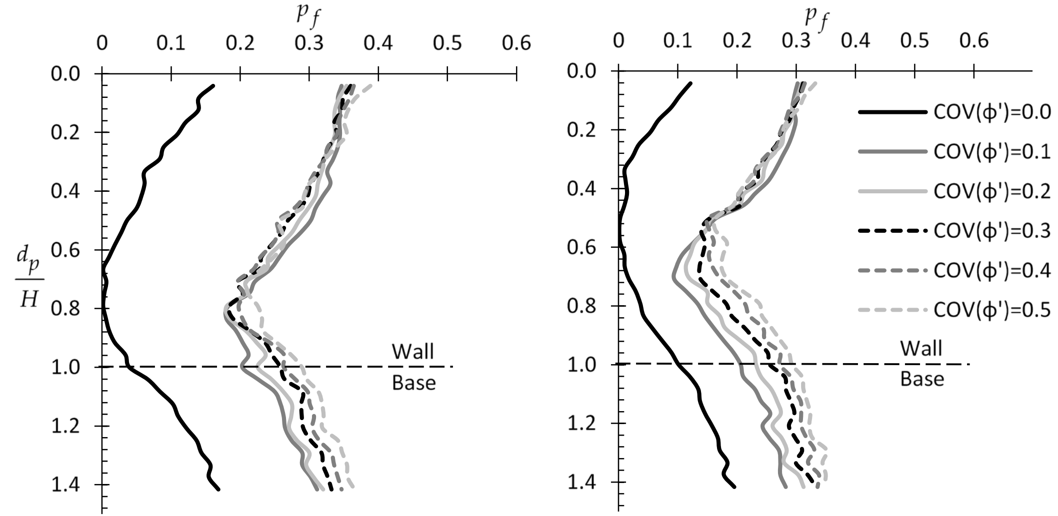
In this paragraph, six COVvalues for ϕ' and γ were considered, i.e. COV =0, 0.1, 0.2, 0.3, 0.4 and 0.5. The optimal sampling distance from the wall was found not to be affected by the COV of ϕ' and γ, where again the x/H =0 case leads to the smaller statistical error. Thus, only the x/H =0 case will be presented here. From Figure 6.7 it is, generally, inferred that the $COV$of ϕ' and γ has no or minor effect on the optimal sampling depth for the sliding and rotating case respectively.

6.2.2.5 *Effect of* $\mu_{\phi'}$ *and* μ_{γ}

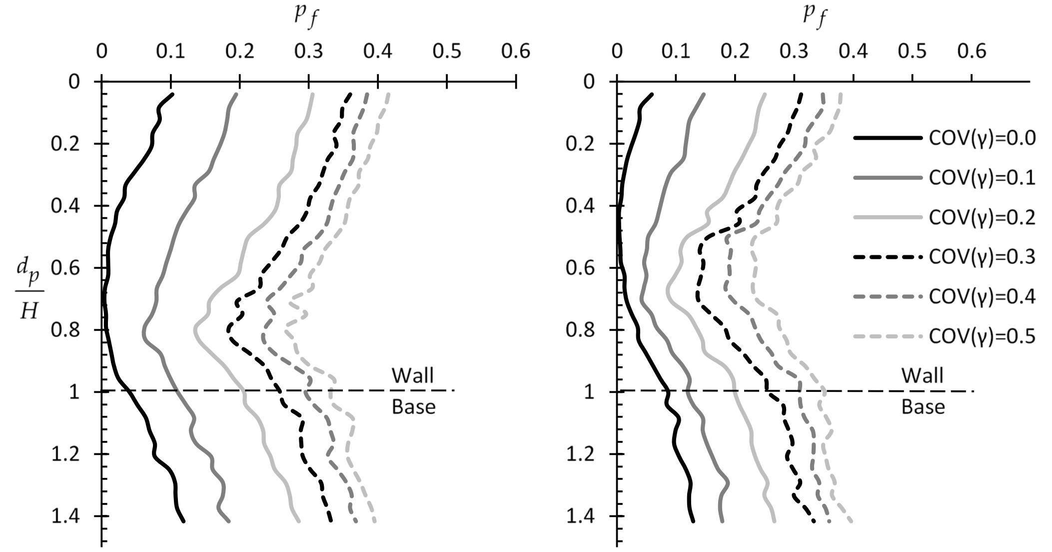
So far, the mean value of ϕ' was equal to 30° in all cases considered in the analysis. The influence of $\mu_{\phi'}$ on the optimal sampling location is examined herein. In this respect, three $\mu_{\phi'}$ values were considered, i.e. $\mu_{\phi'}$ = 20°, 30°, and 40°. The $COV$of ϕ' was set equal to 0.3, while the $COV$of γ was set to zero. The authors found that, the optimal sampling distance was again at x/H =0 for any $\mu_{\phi'}$ value (not shown here for space economy). In addition, from Figure 6.8 it is inferred that the $\mu_{\phi'}$ of soil has no effect on the optimal sampling depth both in the case of translating and rotating wall. The same stands for μ_{γ}.

6.2.2.6 *Effect of the Factor of Safety* (FS)

The variation of p_f with respect to d_p/H for different FS values is shown in Figure 6.9; the optimal sampling distance from the wall face was also found to be at x/H =0 for any FS value, thus, only the x/H =0 case is presented here. From Figure 6.9 it is obvious that the failure probability decreases as FS increases, but what it is not trivial is that, the positive effect of targeted field investigation on the reduction of the statistical error is greater for greater FS values. For example, considering the case of rotating wall, as shown in Figure 6.9, the p_f for FS=1.1 is approximately 0.45 and independent of the d_p/H ratio. For the same soil-wall system, if FS=1.5, p_f ranges from 0.18 for d_p/H =0 (and 1.4) to almost zero for d_p/H =0.6.



(a) Sliding $\theta/H = 2.08$ $x/H = 0$ $\mathrm{COV}_\gamma = 0.3$ (b) Overturning $\theta/H = 2.08$ $x/H = 0$ $\mathrm{COV}_\gamma = 0.3$



(c) Sliding $\theta/H = 2.08$ $x/H = 0$ $\mathrm{COV}_{\phi'} = 0.3$ (d) Overturning $\theta/H = 2.08$ $x/H = 0$ $\mathrm{COV}_{\phi'} = 0.3$

Figure 6.7: p_f versus d_p/H example relationships by considering different values of COV of ϕ' and γ; figures (a) and (b) refer to COV of ϕ', whilst figures (c) and (d) refer to COV of γ for the sliding and overturning moments, respectively.

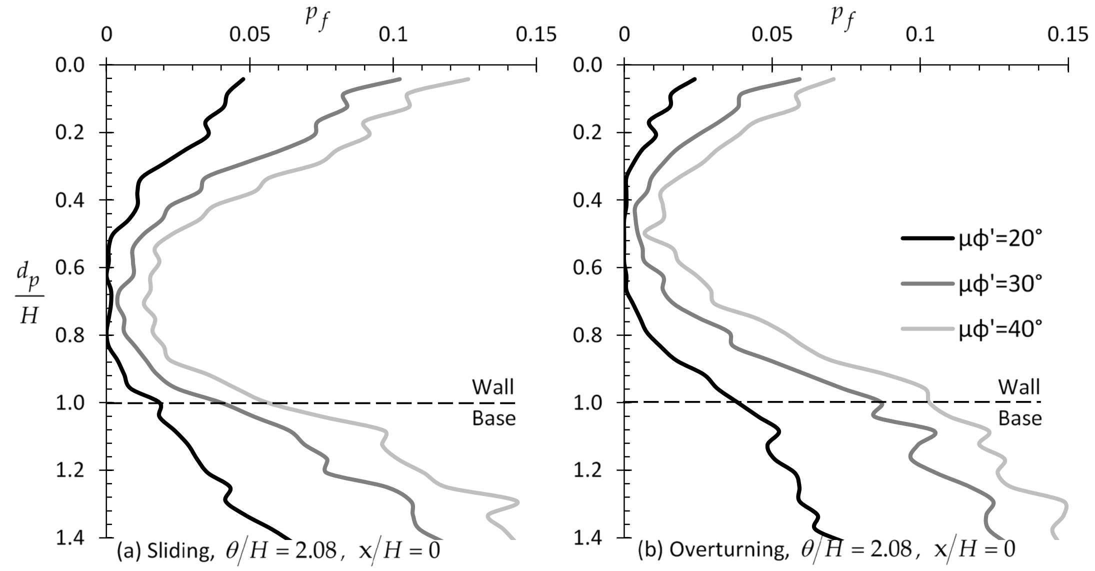


Figure 6.8: p_f versus d_p/H example relationships for (a) sliding and (b) overturning moment considering different $\mu_{\phi'}$ values.

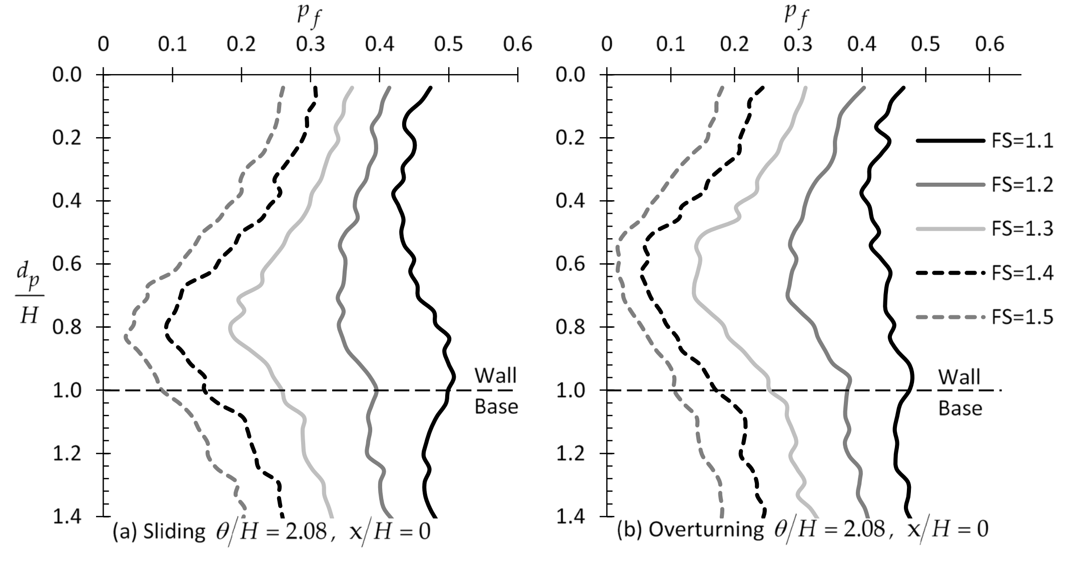


Figure 6.9: p_f versus d_p/H example curves for different FS values for the case of (a) sliding and (b) overturning wall.

6.2.2.7 Effect of soil anisotropy

According to the literature, the spatial variability of soil in the horizontal direction is much greater than the respective one in the vertical direction due to natural deposition and soil formation processes. In this respect, these studies mention that the horizontal spatial correlation length (θ_h) is generally about 10 times the vertical one (θ_v). For example, $\theta_h \approx 9\theta_v$ for Vanmarcke [177], $\theta_h \approx 10\theta_v$ for Soulie et al. [178] and Cherubini [179], $\theta_h \approx 12\theta_v$ for Popescu et al. [180], $\theta_h \approx 13\theta_v$ for Phoon and Kulhawy [181], and $\theta_h \approx$

2 to 7 times the θ_v value for Firouzianbandpey et al. [84]. Driven from these findings, the effect of soil anisotropy on the optimal sampling location will be investigated here by comparing the $\theta_h = \theta_v$ case with the $\theta_h = 10\theta_v$ case. The reference wall-soil system with $\theta_v/H = \theta_h/H$ =2.08 will be compared with a respective one having θ_v/H =2.08 and θ_h/H =20.8. The variation of p_f with d_p/H for various x/H values is shown in Figure 6.10. From this figure it is inferred that, the statistical error practically remains the same for horizontal sampling distances less than one wall height $(x \leq H)$ for both the translating and rotating wall cases. Although for $x \leq H$ the difference in the p_f values is very small, again the optimal sampling location is at $x/H = 0$. Comparing Figure 6.10a with Figure 6.3c and Figure 6.10b with Figure 6.3d, it can be said that the effect of the horizontal sampling location is significantly higher in the isotropic case. Regarding the optimal sampling location, the soil anisotropy has no effect on the optimal sampling depth both in the case of translating and rotating wall.

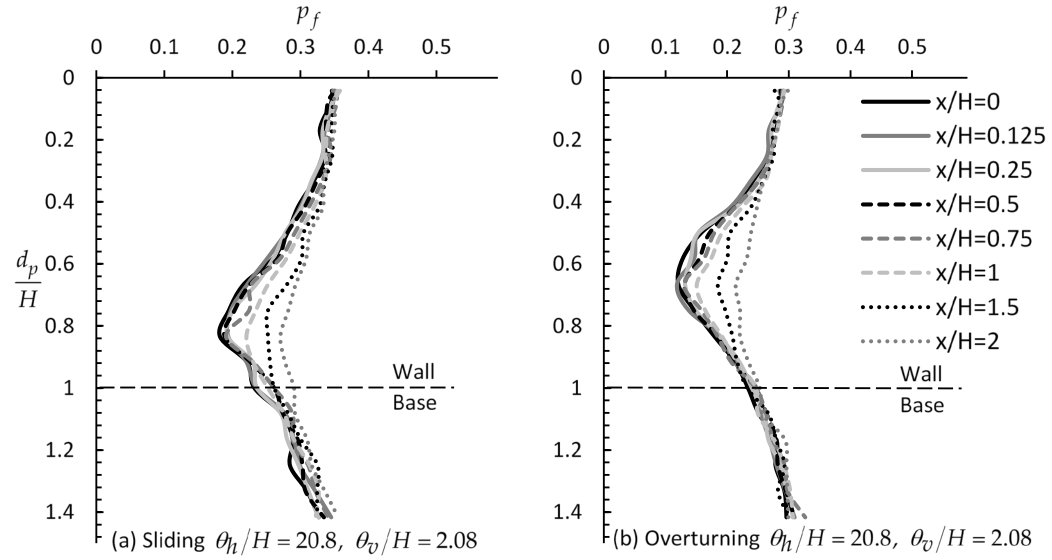


Figure 6.10: p_f versus d_p/H example curves for various x/H values for (a) sliding and (b) overturning wall considering anisotropic soil (to be compared with Figure 6.3c and d respectively).

6.2.3 Sampling from a domain

This sampling strategy refers to data referring to continuous probing tests (e.g. the Cone Penetration Test or Standard Penetration Test). The length of the sampling domain is always measured from the soil surface, whilst arithmetic mean values are used for the various soil properties in the Finite Element Method (FEM) analysis. Since each (finite) element (see Figure 6.1) has edge 0.1 length units (in this respect, meters), sampling is considered to take place every 0.1m (along the vertical direction). The minimum and maximum sampling domain length considered were 0.1m (rather referring to a single point) and 3.4m respectively. It is noted that for all cases examined in this section the optimal sampling distance was found again to be at x/H =0. Thus, for space economy, the analysis below generally refers to the x/H =0 case.

6.2.3.1 Effect of spatial correlation length ($\boldsymbol{\theta}$)

Example charts showing the variation of p_f with respect to d_d/H for various θ/H and x/H values are given in Figure 6.11, both for the case of sliding and rotation of wall; it is reminded that FS was set equal to 1.3 (recall Equation (6.3)). From this figure it is inferred that, the optimal horizontal sampling distance from the wall is again for x/H =0, although for very small theta values the horizontal sampling distance makes no noticeably difference. However, as the theta increases the role of horizontal distance becomes more significant. Given now that, soil samples will be taken from x/H =0, it is advisable, as it is inferred from Figure 6.11, that the entire domain length along the wall to be considered, especially for the rotational mode of failure. This practice may significantly reduce the statistical error. It is also interesting that, extending the sampling domain beyond the maximum depth of wall (i.e. d_d/H>1), the statistical error remains constant. Finally, from Figure 6.11 it is inferred that, a “worst case theta” exists. This is more obvious in Figure 6.12 showing the variation of p_f with d_d/H for various θ/H values and x/H =0.

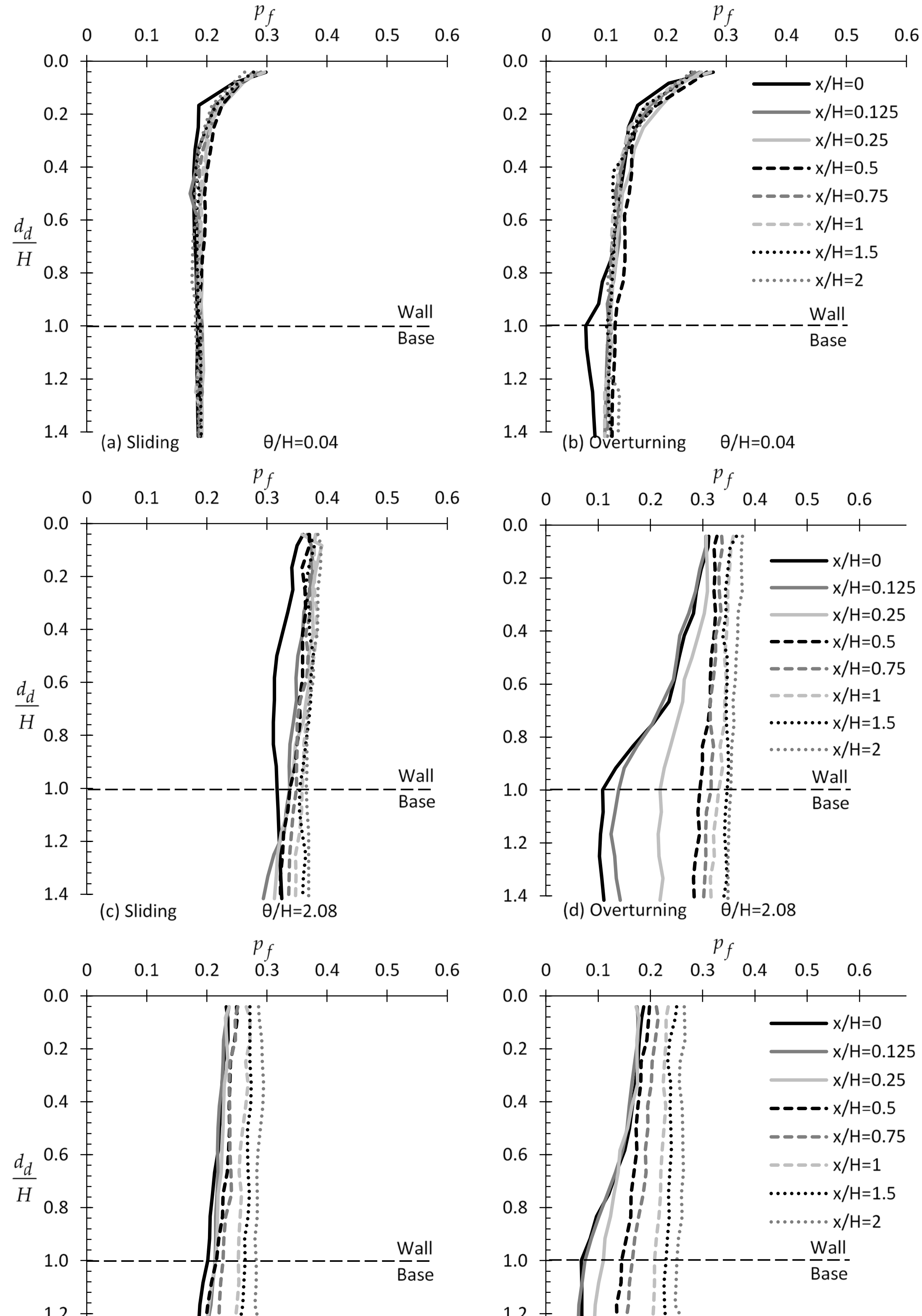


Figure 6.11: p_f versus d_d/H example relationships for different values of scaled correlation length θ/H and lateral distance from the wall face (x/H). Figures (a), (c), and (e) refer to the case of sliding wall whilst figures (b), (d) and (f) to the case of overturning wall.

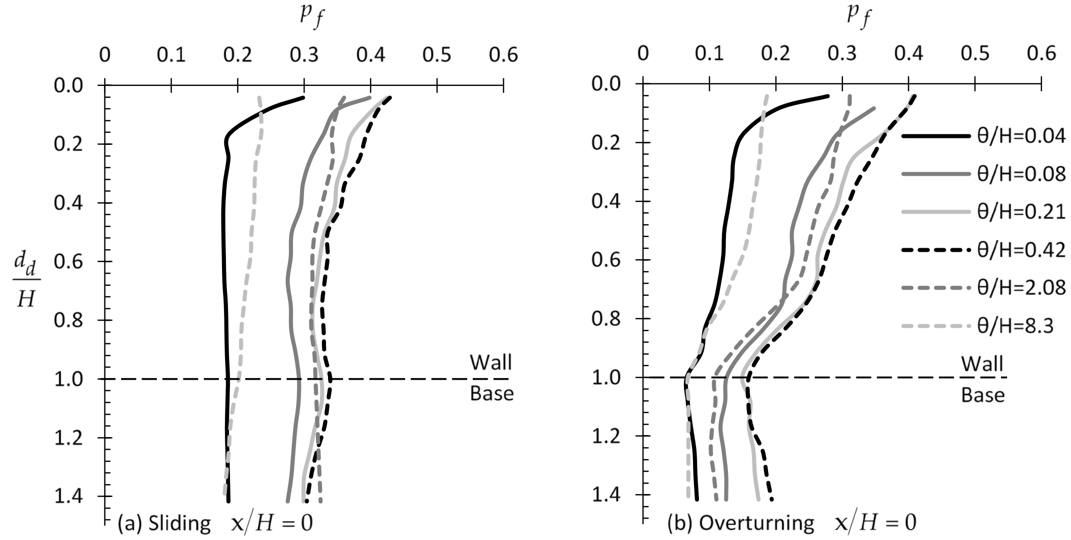


Figure 6.12: p_f versus d_d/H example relationships for the case of (a) sliding and (b) overturning wall by considering different scaled θ/H values.

6.2.3.2 *Effect of wall roughness*

Generally, the wall roughness has a minor effect on the optimal sampling domain length, although, as expected (see Figure 6.13), it noticeably affects the failure probability. As shown in Figure 6.13, a great reduction in the statistical error can be achieved only in the case of smooth rotating wall, with the optimal sampling domain length being the entire wall height. Characteristically it is mentioned that, the minimum failure probability is obtained for d_d/H=1 and that, this probability remains constant for greater d_d/H values.

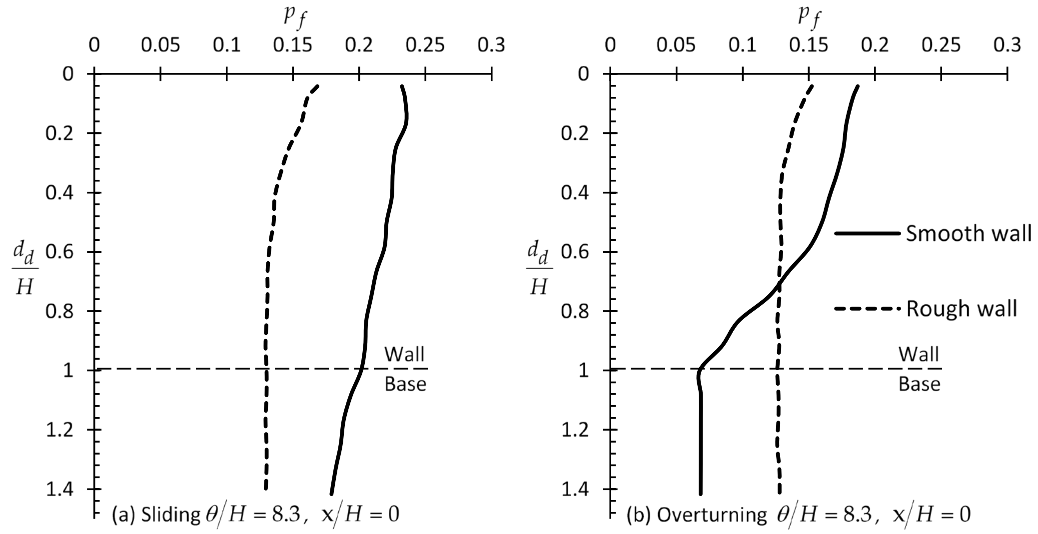
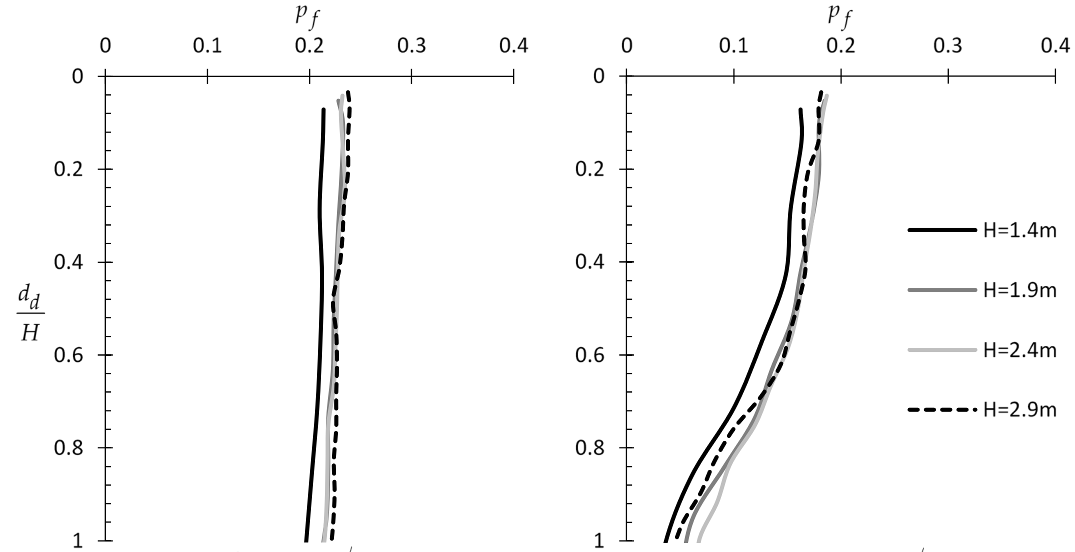


Figure 6.13: p_f versus d_d/H example curves for $\theta/H = 8.3$ and $x/H = 0$ (rough and smooth wall) for the case of a) sliding and b) overturning wall.

6.2.3.3 *Effect of wall height*

In this paragraph four wall heights were considered, i.e. H =1.4, 1.9, 2.4, 2.9 m. Figure 6.14 presents the variation of p_f with d_d/H for these four cases. From this figure it is clear that the wall height has only minor influence on the sampling domain length (see also Section 3.1.3).

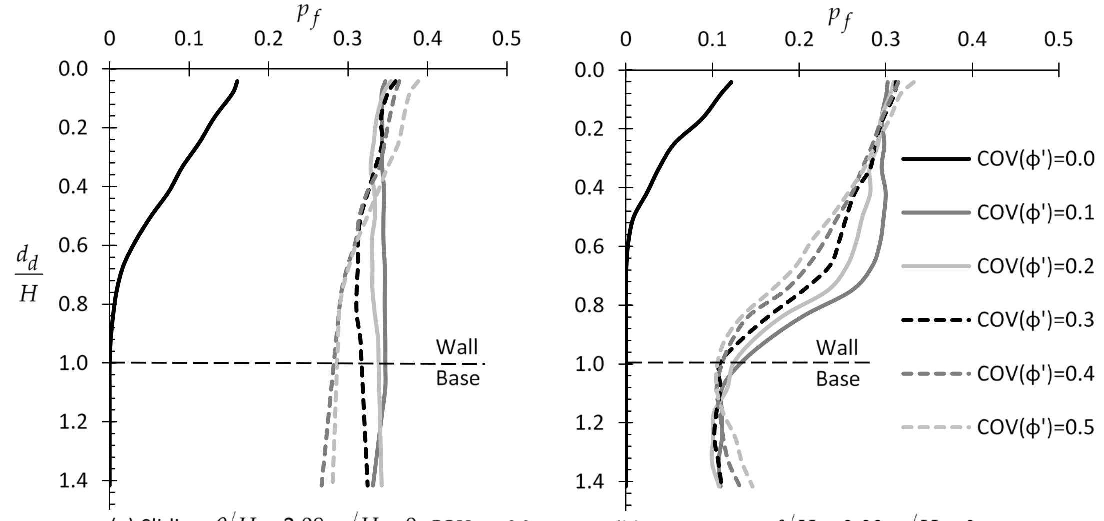


(a) Sliding $\theta = 20m,\ x/H = 0$ (b) Overturning $\theta = 20m,\ x/H = 0$

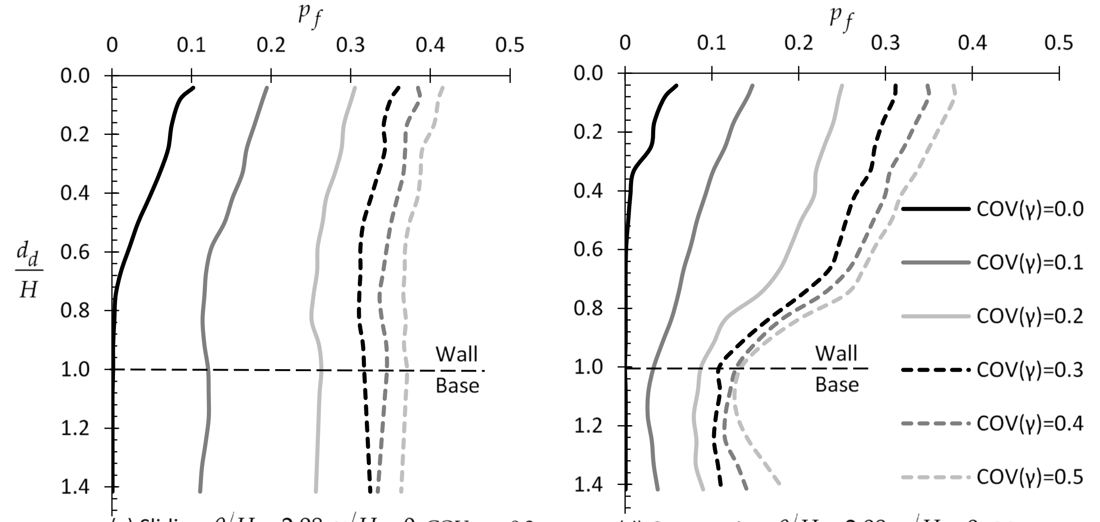
Figure 6.14: p_f versus d_d/H example curves for different wall heights H and $\theta = 20m$ for the case of a) sliding and b) overturning wall.

6.2.3.4 *Effect of COV of ϕ' and γ*

In this paragraph, six COV values for ϕ' and γ were considered, i.e. COV =0.0, 0.1, 0.2, 0.3, 0.4 and 0.5. The optimal horizontal sampling distance from the wall was found not to be affected by the COV of ϕ' or γ, where again the x/H =0 case leads to the smaller probabilities of failure. Thus, only the x/H =0 case will be presented here. From Figure 6.15 it is, generally, inferred that the COV of ϕ' and γ has no or minor effect on the optimal sampling length for the sliding and rotating case respectively.



(a) Sliding $\theta/H = 2.08$ $x/H = 0$ $\text{COV}_\gamma = 0.3$ (b) Overturning $\theta/H = 2.08$ $x/H = 0$ $\text{COV}_\gamma = 0.3$

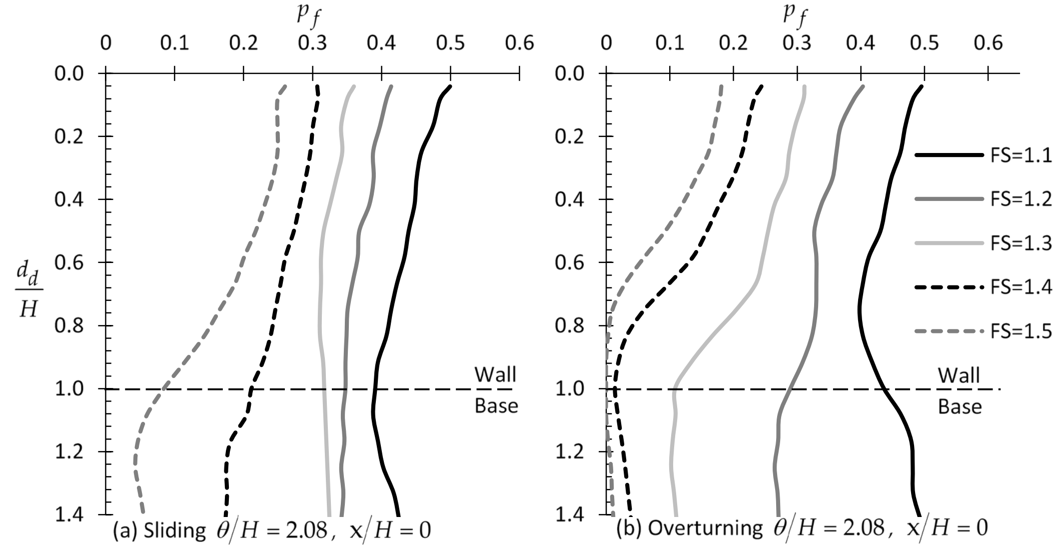


(c) Sliding $\theta/H = 2.08$ $x/H = 0$ $\text{COV}_{\phi'} = 0.3$ (d) Overturning $\theta/H = 2.08$ $x/H = 0$ $\text{COV}_{\phi'} = 0.3$

Figure 6.15: p_f versus d_d/H example relationships by considering different values of COV of ϕ'(figures (a) and (b)) and γ (figures (c) and (d)); figures (a) and (c) refer to the case of sliding wall, whilst figures (b) and (d) refer to the case of overturning wall.

6.2.3.5 Effect of the Factor of Safety (FS)

The variation of p_f with respect to d_d/H for different FS values is shown in Figure 6.16; the optimal sampling distance was also found to be at x/H =0 for any FS value, thus, only this case is presented here. From Figure 6.16 it is obvious that the failure probability decreases as FS increases, but what it is not trivial is that, the positive effect of targeted field investigation on p_f (that is, decrease in p_f) is greater for greater FS values.

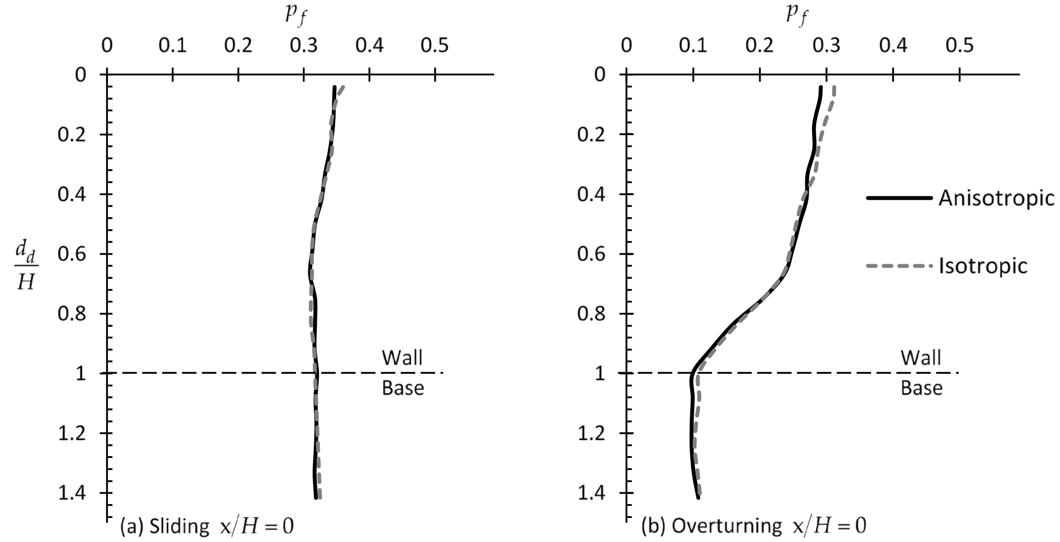


(a) Sliding $\theta/H = 2.08$, $x/H = 0$ (b) Overturning $\theta/H = 2.08$, $x/H = 0$

Figure 6.16: p_f versus d_d/H example curves for different FS values for the case of a) sliding and b) overturning wall.

6.2.3.6 Effect of soil anisotropy

In this paragraph the retaining soil is considered to be highly anisotropic having θ_h/H =20.8 and θ_v/H =2.08; for the isotropic case is stands that $\theta_h/H = \theta_v/H = \theta/H$= 2.08. The optimal horizontal sampling distance from the wall was found not to be affected by the anisotropy of soil, where again the x/H =0 case leads to the smallest probabilities of failure; thus, only the x/H =0 case is presented here. From Figure 6.17, it is generally inferred that the soil anisotropy has also no effect on the sampling domain length for sliding or rotating cases.



(a) Sliding $x/H = 0$ (b) Overturning $x/H = 0$

Figure 6.17: p_f versus d_d/H example curves for the case of (a) sliding and (b) overturning wall considering anisotropic soil (θ_h/H =20.8 and θ_v/H =2.08) and isotropic soil ($\theta_h/H = \theta_v/H = \theta/H$= 2.08).

6.2.4 Discussion

6.2.4.1 Optimal sampling locations

One of the main outcomes derived from the present analysis is that the optimal horizontal sampling location in the active state of stress is at x/H =0, that is, immediately adjacent to the wall face. This came as a surprise as someone would expect the optimal location to lie on or in the close vicinity of Rankine's 45°+ϕ' /2 failure plane passing through the lower point of the wall. Actually, based on author' findings (see Section 3 of the current Chapter), this is the case for the passive state. Regarding Rankine's earth pressure theory it is reminded that, there is not a single failure plane, but an infinite number of such planes, parallel to the one mentioned above encompassing all the other [171]. A common characteristic of these planes is that their lower point is in contact with the wall. In this respect, it seems that the optimal sampling location in the active case shows preference to this array of points. Regarding now the depth of the optimal sampling *point*, it was found that this lies at depth greater than the 2/3 or 1/2 of the wall height for the sliding and rotational mode of failure respectively; the exact depth depends on the spatial correlation length of the soil. For the optimal sampling domain length, it is advisable that the entire wall height be considered.

6.2.4.2 The importance of targeted field investigation in practice

The importance of targeted field investigation, where samples are taken from a priory known optimal locations, is highlighted here. *A random material field referring to a specific RFEM realization* (such as the one presented in Figure 6.1), it can be said that it convincingly represents a real field. For the three examples presented in this paragraph, the reference wall (and the finite element mesh) of Figure 6.1 will be used, whilst the material properties are given in Table 6.1. These materials differ from each other, in essence, in the spatial correlation length and only for the first material, in addition to the friction angle of soil, the unit weight is a random field. Besides, as it is inferred from the present research, the mean and COV values of ϕ' and γ have no effect on the optimal sampling location. The random field of ϕ' used in each example is shown in Figure 6.1, Figure 6.18 and Figure 6.19 respectively. It is reminded that the light areas correspond to lower friction angles and vice versa. The FS value is assumed unity (recall Equation (6.3)); this factor is discussed in the next paragraph.

Table 6.1: Summary of the characteristics of the soils used in the three examples (wall height H= 2.4 m).

Example	Random field(s)	Distribution	$\mu_{\phi'}$	μ_{γ}	COV	θ/H	Figure [(1)]
#1	ϕ', γ	Log-normal	30°	20 kN/m^3	0.3	8.3	Figure 6.1
#2	ϕ'	Log-normal	30°	20 kN/m^3	0.3	4.2	Figure 6.18
#3	ϕ'	Log-normal	30°	20 kN/m^3	0.3	0.42	Figure 6.19

[(1)] Figures shown the random fields of ϕ'.

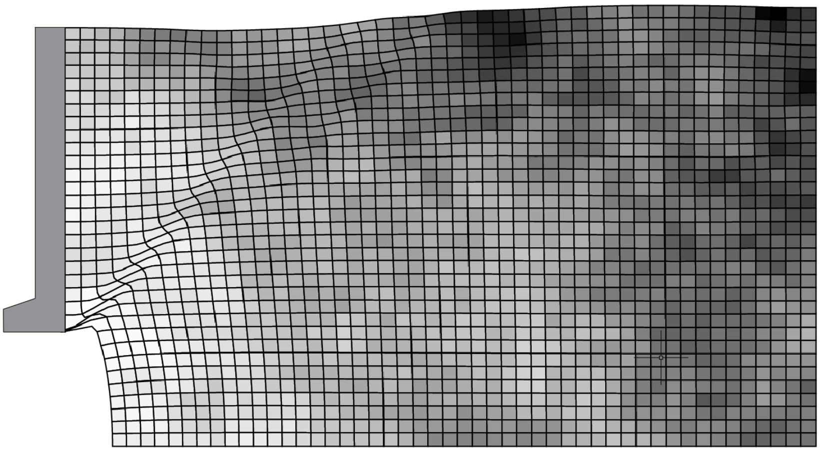


Figure 6.18: Graphical representation of the random field of ϕ' of Example #2 (θ/H=4.2; see Table 6.1). Light areas correspond to lower friction angles and vice versa.

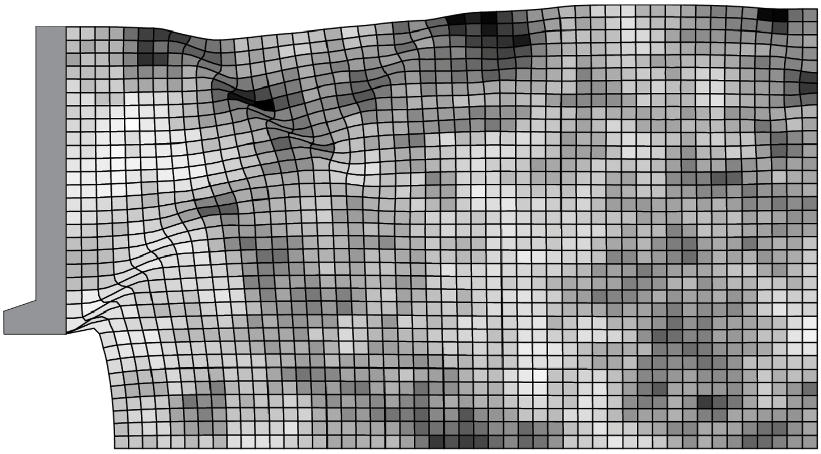


Figure 6.19: Graphical representation of the random field of ϕ' of Example #3 (θ/H=0.42; see Table 6.1). Light areas correspond to lower friction angles and vice versa.

The *predicted* resultant driving force (F) or moment (M) acting on the wall is compared against the respective "*actual*" ones. For each one of the examples presented herein, the latter derives from the respective random field of ϕ' using the RFEM method (in Example #1, in addition to ϕ', γ is also a random field). The *predicted* F and M values derive from a homogenous soil field characterized by the mean of the values sampled from the original (random) field. The results are presented in Figures 20-22 in $F_{predicted}/F_{"actual"}$ or $M_{predicted}/M_{"actual"}$ versus x/H form for various d_d/H values. The relative difference R_d is also given in each chart (secondary vertical axis; a positive value indicates design on the safe side and vice versa; R_d is equal to $F_{predicted}/F_{"actual"}$-1 or $M_{predicted}/M_{"actual"}$-1 for the case of forces and moments respectively).

If the suggestions related to the horizontal distance from the wall and the domain length (x/H=0 and d_d/H=1 respectively) are valid, the $F_{predicted}/F_{"actual"}$ and $M_{predicted}/M_{"actual"}$ ratios for this specific sampling scenario should, logically, be equal to unity or very close to this value. The readers should bear in their mind that, a $F_{predicted}/F_{"actual"}$ or $M_{predicted}/M_{"actual"}$ value close to unity or equal to unity for a x/H value other than zero does not indicate that this x/H location is an optimal sampling location. As the soil retained by the wall is a spatially random field, a set of samples taken from points away from the wall face, may also give (coincidentally) mean value equal (or approximately equal) to the respective one obtained from a set of samples taken from the x/H=0 location.

As shown in Figure 6.20 - Figure 6.22, the $F_{predicted}/F_{"actual"}$ and $M_{predicted}/M_{"actual"}$ ratio values for x/H=0 are very close to unity or equal to unity, indicating the validity of author' suggestions. Indicatively, it is mentioned that the abrupt drop of the $F_{predicted}/F_{"actual"}$ (or $M_{predicted}/M_{"actual"}$) versus x/H curves in Figure 6.20 between x/H=0.5 and 1.5 is attributed to the "dark" (strong) area appearing at this particular location, as shown in Figure 6.1. From Figure 6.20 - Figure 6.22 it is also confirmed that a vertical sampling domain of length equal to the wall height gives better prediction for the destabilizing forces acting on the wall.

A comparison between the figures given in Section 2.1 with the respective ones given in Section 2.2 shows clearly that statistical uncertainty does not necessarily decrease with the increasing number of samples. Indeed, the opposite may easily happen. For example, comparing the $p_f \approx 0.03$ value for $x/H = 0$ shown in Figure 6.3f (single point case) with the $p_f \approx 0.26$ value for $x/H = 2$ (case of 24 sampling points) shown in Figure 11f, it is obvious that statistical uncertainty can only be minimized by targeted field investigation. Such examples can also be found in the present section; please compare the case of $\{x/H = 2, d_d/H = 1\}$ with the $\{x/H = 0, d_d/H = 0.25\}$ in Figure 21b giving $R_{d,M} \approx -0.42$ and -0.12, respectively.

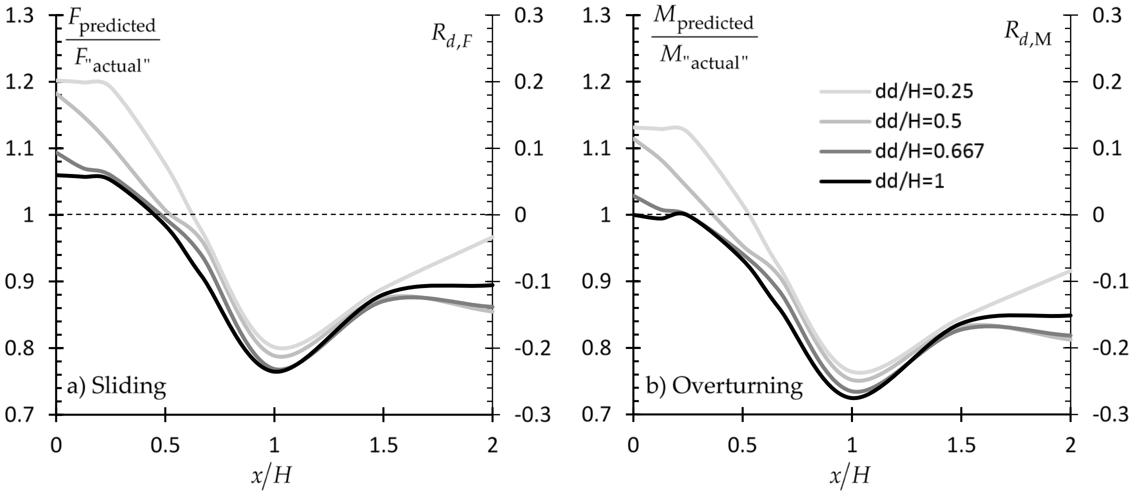


Figure 6.20: Example #1: $F_{predicted}/F_{"actual"}$ and $M_{predicted}/M_{"actual"}$ vs x/H curves for various d_d/H values and for both the sliding and overturning failure case (see also Table 6.1 and Figure 6.1).

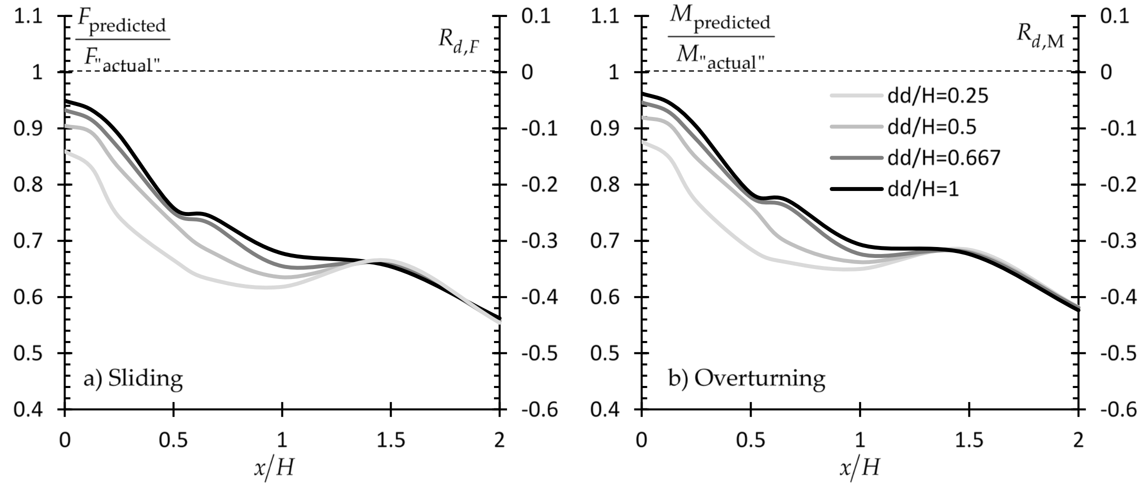


Figure 6.21: Example #2: $F_{predicted}/F_{"actual"}$ and $M_{predicted}/M_{"actual"}$ vs x/H curves for various d_d/H values and for both the sliding and overturning failure case (see also Table 6.1 and Figure 6.18).

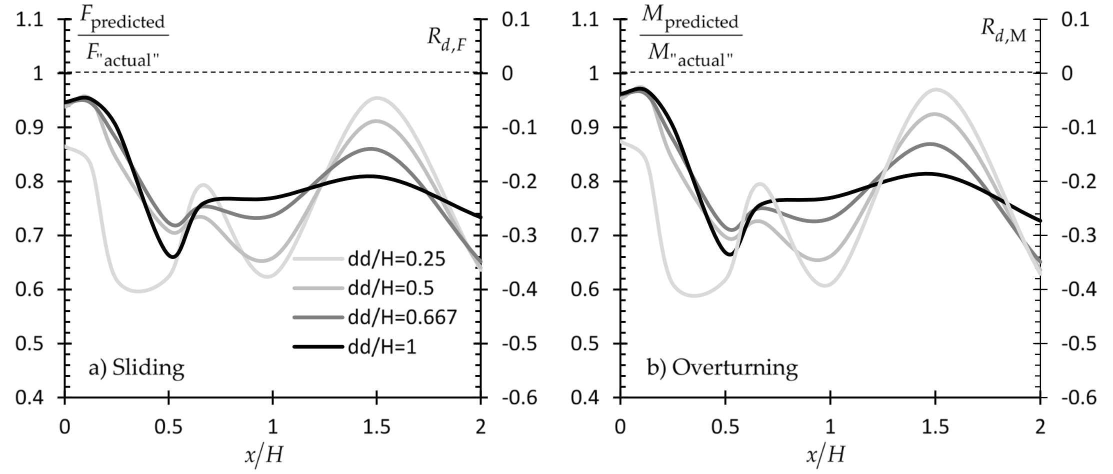


Figure 6.22: Example #3: $F_{predicted}/F_{"actual"}$ and $M_{predicted}/M_{"actual"}$ vs x/H curves for various d_d/H

values and for both the sliding and overturning failure case (see also Table 6.1 and Figure 6.19).

6.2.4.3 Designing with Load and Resistance Factor Design (LRFD) codes

The discussion on the design of earth retaining structures based on characteristic soil property values (please see equations (2.1) and (2.2) in the literature review of this Thesis) instead of the respective mean values is facilitated by the two example charts of Figure 6.23. These charts refer to the case #3 presented in the previous paragraph (see also Table 6.1). This specific case was chosen because of the relatively low θ value (i.e. θ/H=0.42), which indicates a rather highly spatially variable soil; thus, the use of the characteristic value makes more sense. Two cases are presented, the d_d/H=1 and the d_d/H=0.25. The figure in question refers to the sliding mode of failure, however, the respective curves for the overturning mode of failure do not differ appreciably. It is also mentioned that in the example presented here the partial material factor for the friction angle $\gamma_M = \gamma_\phi$ was set equal to unity.

From Figure 6.23 it is clear that the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic values. Moreover, despite the conservatism which is inserted in the analysis using the characteristic value concept, the characteristic values alone, as shown, cannot guaranty a conservative enough engineering study. The safety level can be increased by applying a statistical uncertainty partial factor (similar to the model factor γ_R used by Eurocode 7) or a unified and more conservative model factor to the resistances, which will absorb the statistical uncertainties related to the soil. In this respect, a partial factor equal to 1.3 has also been applied (FS=1.3; recall Equation (6.3)) in the present example. As shown in Figure 6.23, the use of such a factor simply displaces upwards (that is to the safe side) the $F_{predicted}/F_{"actual"}$ (or $M_{predicted}/M_{"actual"}$) versus x/H curves. The inclusion of the "characteristic value" in the REARTH2D code has been done by the authors.

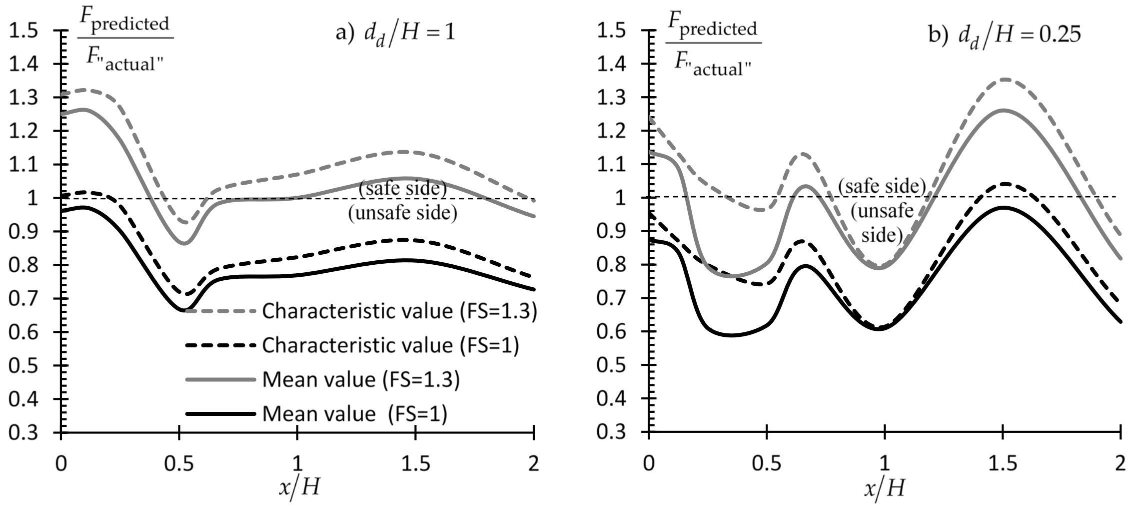


Figure 6.23: $F_{predicted}/F_{"actual"}$ vs. x/H curves using both mean and characteristic values (dashed and solid lines respectively) for $FS = 1$ and 1.3. Figure referring to the case of a sliding wall and to two sampling domain cases ($d_d/H = 1$ (figure a) and to $d_d/H = 0.25$ (figure b)). The reference wall was used. Soil characteristics as shown in Table 6.1 (Example #3).

6.3 Passive state of stress

6.3.1 Parametric study for determining the optimal sampling strategy

For the determination of optimal sampling strategy in passive state of failure, the soil mass is discretized into a 60x34 mesh (number of elements at the horizontal and vertical direction respectively) consisting of eight-noded square elements with side length equal to 0.1m (Figure 6.24). As in the previous section, various wall heights are considered ranging from H=1.4 m to 2.9 m, while the mesh geometry has been kept the same for all cases. The reference wall (i.e. 24-element wall) will be generally considered in the analysis. The 60-element mesh in the horizontal direction was chosen so that the failure mechanism in the RFEM analysis not to be affected by the proximity of the right boundary.

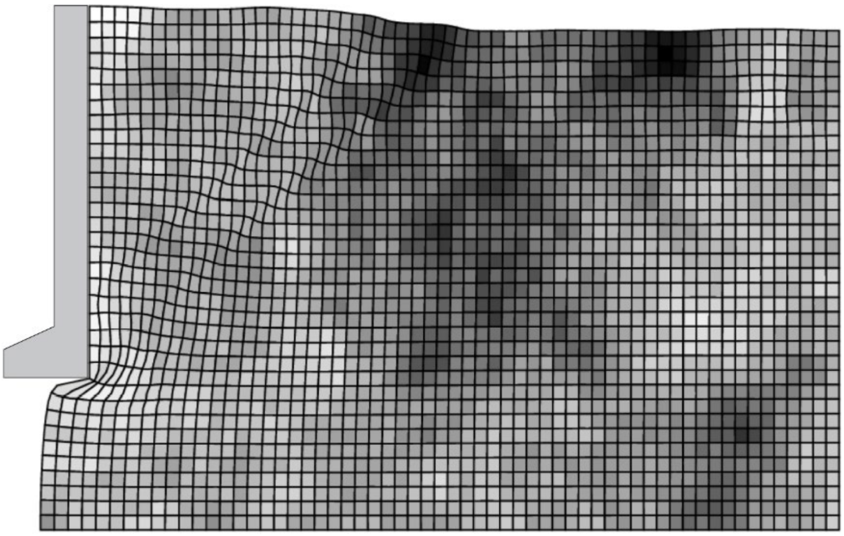


Figure 6.24: Passive earth failure of the "reference" wall. Graphical representation of a random field of ϕ' (this is a typical random finite element method (RFEM) realisation); light areas correspond to lower friction angles and vice versa. For the soil shown, $\theta/H =$ 8.3 and $COV_{\phi'}$ = 0.3.

Similarly to active state, the following parameters will be examined: the sampling depth (d_p) and horizontal distance (x) for the case of sampling from a single point, the sampling domain length (d_d) and horizontal distance (x) of the continuous probing test location for the case of sampling from a domain, the spatial correlation length of soil (θ), the wall roughness (perfectly smooth or perfectly rough wall), the wall height (H), the coefficient of variation (COV) of ϕ', the mean value of ϕ', the safety factor value (FS) considered and the soil mass anisotropy $(\theta_h \neq \theta_v)$. The number of realizations was set equal to 3000; this number, as discussed in Appendix III, can be considered adequate for the needs of the present research. The element size was chosen to be half of the smallest spatial correlation length considered in this study (see e.g. [182]).

6.3.2 Sampling from a Single Point

6.3.2.1 Effect of Spatial Correlation Length $(\boldsymbol{\theta})$

Example charts showing the variation of p_f with respect to d_p/H for various θ/H values are given in Figure 6.25. From this figure it is inferred that, the optimal sampling location for the passive case is 0.5H away from the wall both for the case of translating and rotating wall and for any θ value. Isolating the curves for x/H =0.5 (see Figure 6.26) it seems that, there is a "worst case spatial correlation length" (see also [176]), where, the failure probability becomes maximum; for example, for the various θ/H ratios shown in Figure

6.26 (ranging from $\theta/H = 0.08$ to 8.3), the $\theta/H = 0.42$ case gives the higher p_f values. From Figure 6.26 it is also inferred that, as θ tends to zero, the p_f value tends to a single value for any depth (that is, p_f is independent of the sampling depth). However, as θ increases, the p_f value becomes more dependent on the sampling depth.

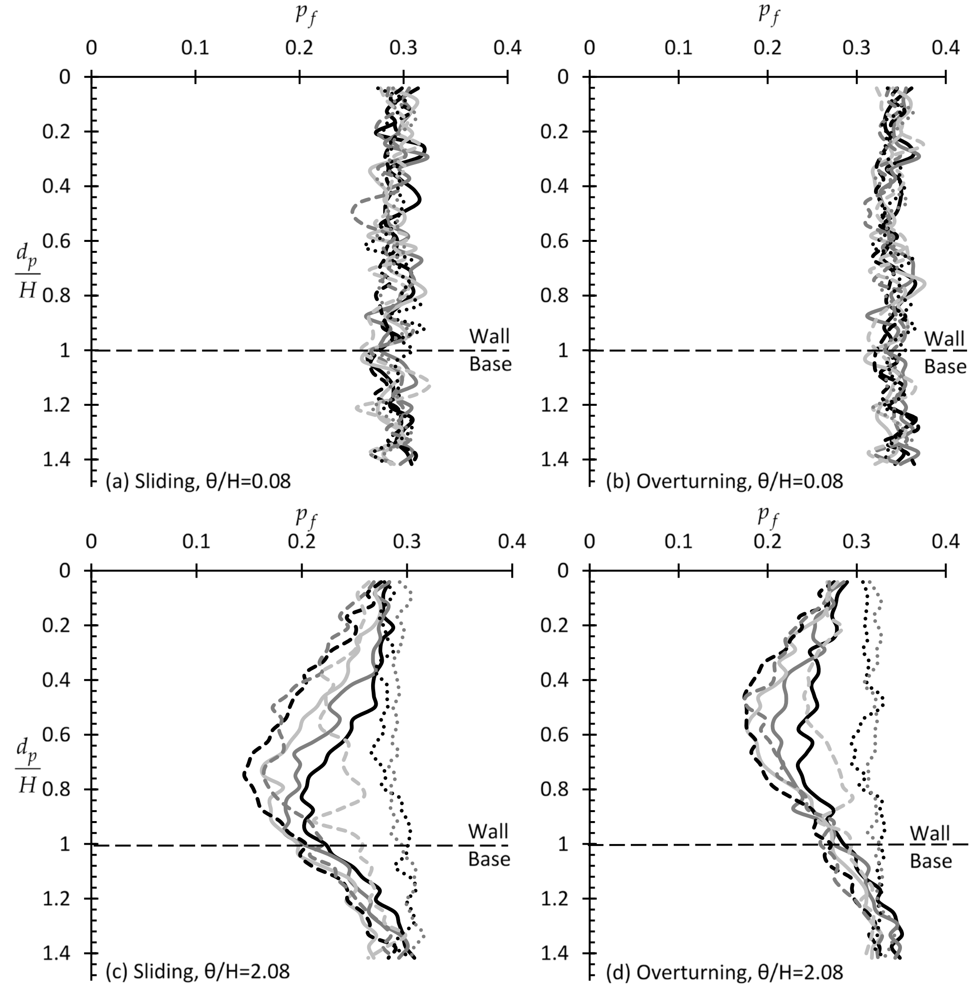


Figure 6.25: p_f versus d_p/H example curves for various θ/H and x/H values for the case of sliding (figure (a, c, e)) and overturning wall (figures (b, d, f)).

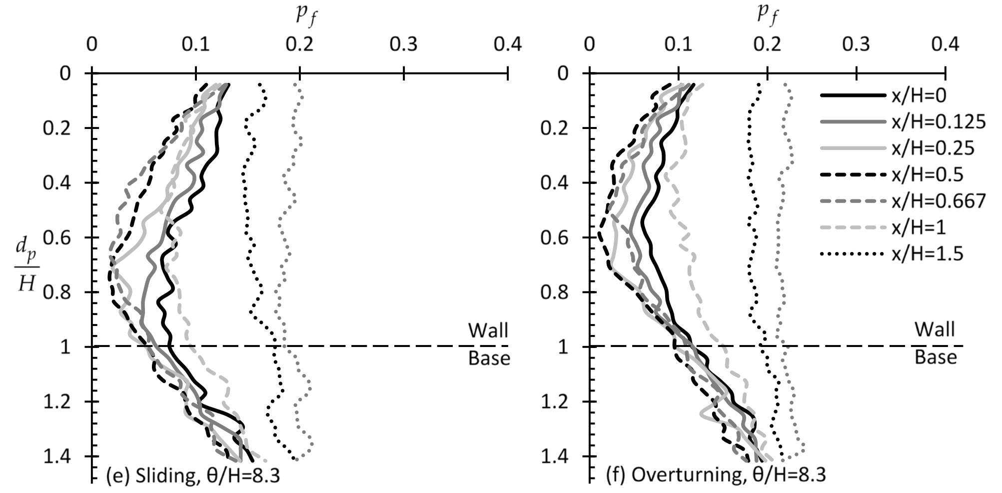


Figure 6.25: Continued.

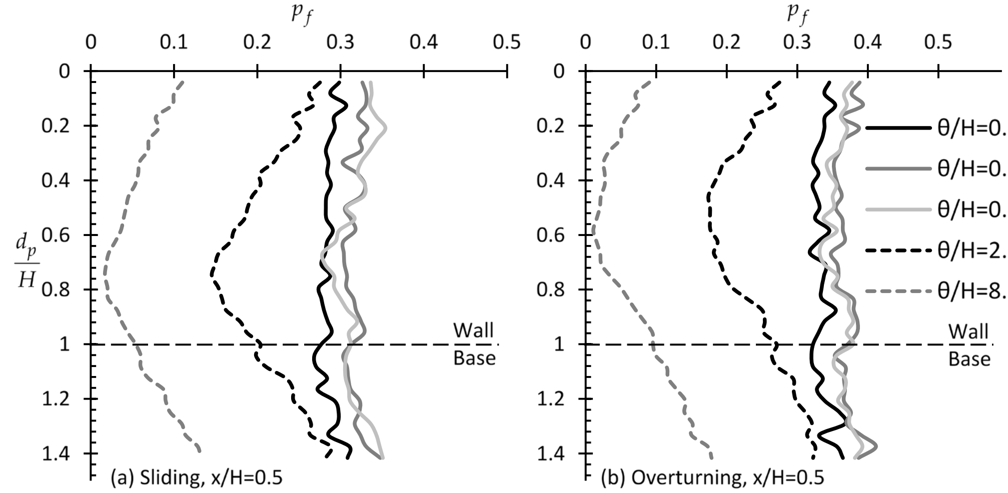


Figure 6.26: p_f versus d_p/H example curves for various θ/H and $x/H = 0$ (smooth wall) for the case of (a) sliding and (b) overturning wall.

6.3.2.2 Effect of Wall Roughness

The effect of wall roughness on the location of the optimal sampling point is examined here. The optimal sampling distance was also found to be at x/H =0.5, therefore, results are presented only for this case. From Figure 6.27 it is inferred that the wall roughness has no effect on the optimal sampling depth or on the optimal sampling location for the sliding and rotating case respectively. However, comparing Figure 6.26 with Figure 6.27 (for perfectly smooth and perfectly rough wall respectively) it is inferred that, by choosing the proper sampling depth, smaller p_f can be obtained when the wall is smooth.

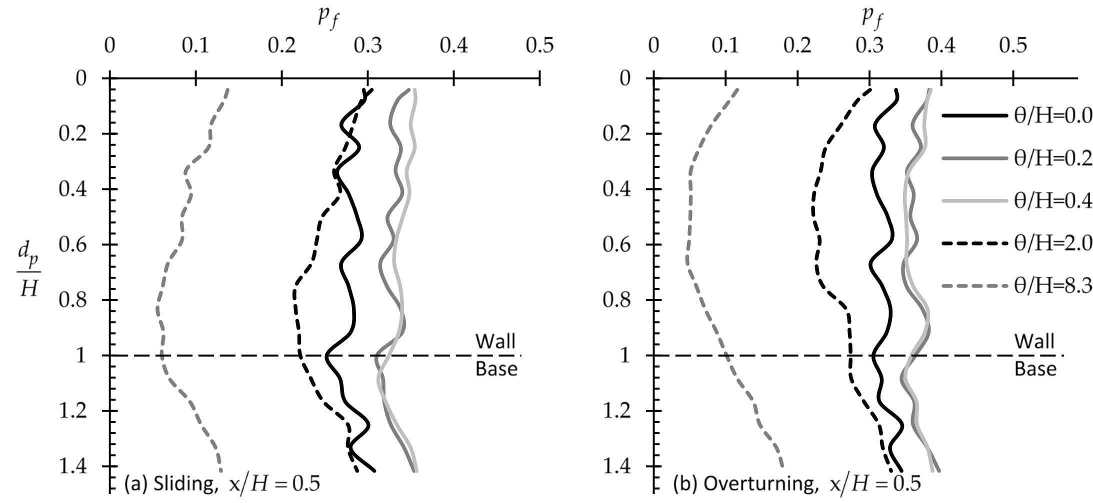


Figure 6.27: p_f versus d_p/H example curves for various θ/H and $x/H = 0$ (perfectly rough wall) for the case of (a) sliding and (b) overturning wall; please compare with Figure 6.26 (perfectly smooth wall).

6.3.2.3 Effect of Wall Height

The variation of p_f with d_p/H for various wall height values , i.e. H =1.4, 1.9, 2.4, 2.9m, is shown in Figure 6.28. From this figure it is clear that, the wall height has a minor influence on the location of the optimal sampling point.

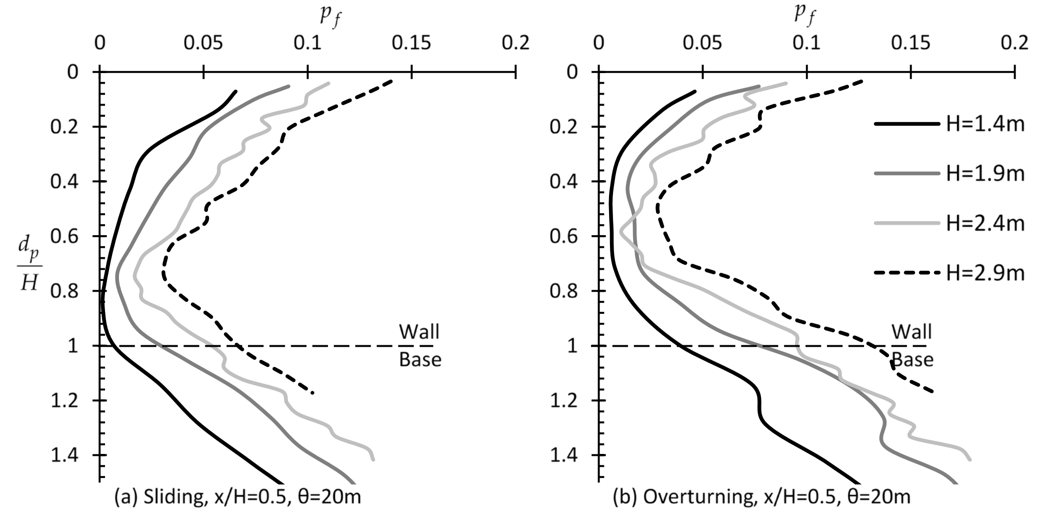


Figure 6.28: p_f versus d_p/H example curves for different wall heights, H, and θ =20 m for the case of (a) sliding and (b) overturning wall.

6.3.2.4 Effect of COV of ϕ'

In this paragraph, five COVvalues for ϕ' were considered, i.e. $COV = 0.1$, 0.2, 0.3, 0.4 and 0.5. The optimal sampling distance from the wall was found not to be affected by the $COV$of ϕ', where again the $x/H = 0.5$ case leads to the smaller statistical error. Thus, only the $x/H = 0.5$ case will be presented here. From Figure 6.29 it is, generally, inferred that the COV of ϕ' has no effect on the optimal sampling depth both for the sliding and rotating cases.

6.3.2.5 Effect of $\mu_{\phi'}$ value

The influence of $\mu_{\phi'}$ on the optimal sampling location is examined here. In this respect, three $\mu_{\phi'}$ values were considered, i.e. $\mu_{\phi'} = 20$, 30 and 40°. The $COV$of ϕ' was set equal to 0.3, while the $COV$of γ was set to zero. The authors found that, the optimal sampling distance was again at $x/H = 0.5$ for any $\mu_{\phi'}$ value. In addition, from Figure 6.30 it is inferred that the $\mu_{\phi'}$ of soil has no effect on the optimal sampling depth both in the case of translating and rotating wall.

6.3.2.6 Effect of the Factor of Safety (FS)

The variation of p_f with respect to d_p/H for different FS values is shown in Figure 6.31; the optimal sampling distance from the wall face was also found to be at $x/H = 0.5$ for any FS value. Thus, only the $x/H = 0.5$ case is presented here. From Figure 6.31, it is obvious that the failure probability decreases as FS increases, but what is not trivial is that the positive effect of targeted field investigation on the reduction of the statistical error is greater for greater FS values. For example, considering the case of the translating wall, as shown in Figure 6.31a, the p_f value when $FS = 1.05$ is approximately 0.47 and independent of the d_p/H ratio. For the same soil–wall system, if $FS = 1.45$, p_f ranges from 0.14 for $d_p/H = 0$ (and 1.4) to 0.03 for $d_p/H \approx 0.8$.

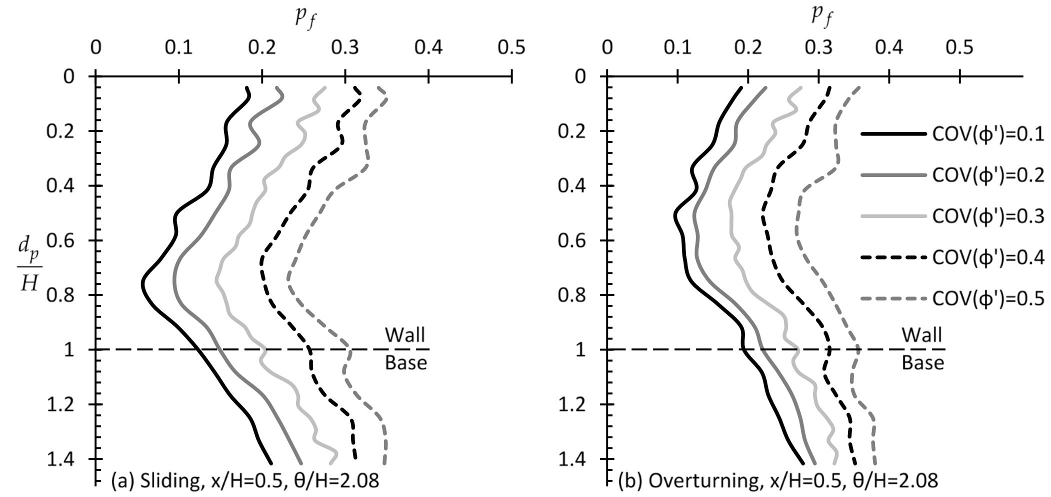


Figure 6.29: p_f versus d_p/H example relationships considering different values of COV of ϕ' for the case of (a) sliding and (b) overturning wall.

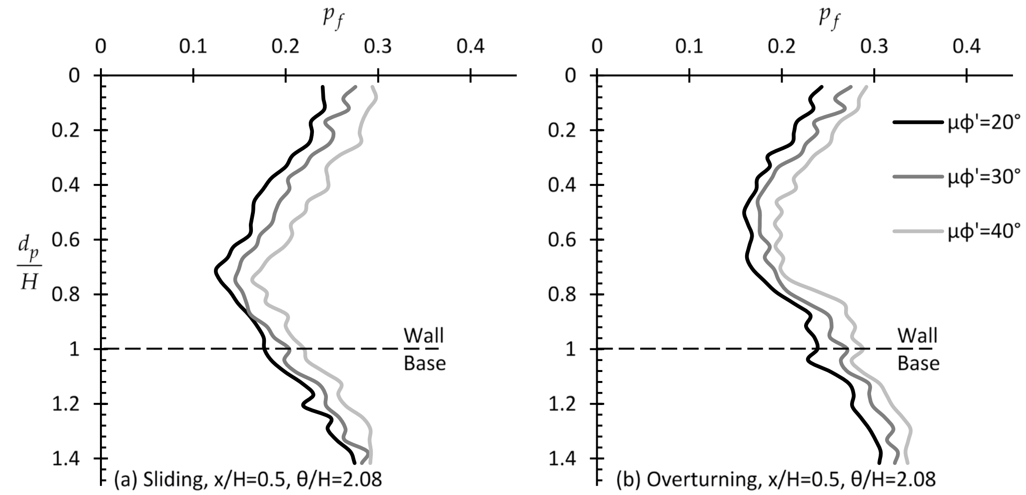


Figure 6.30: p_f versus d_p/H example relationships for (a) sliding and (b) overturning wall considering different $\mu_{\phi'}$ values.

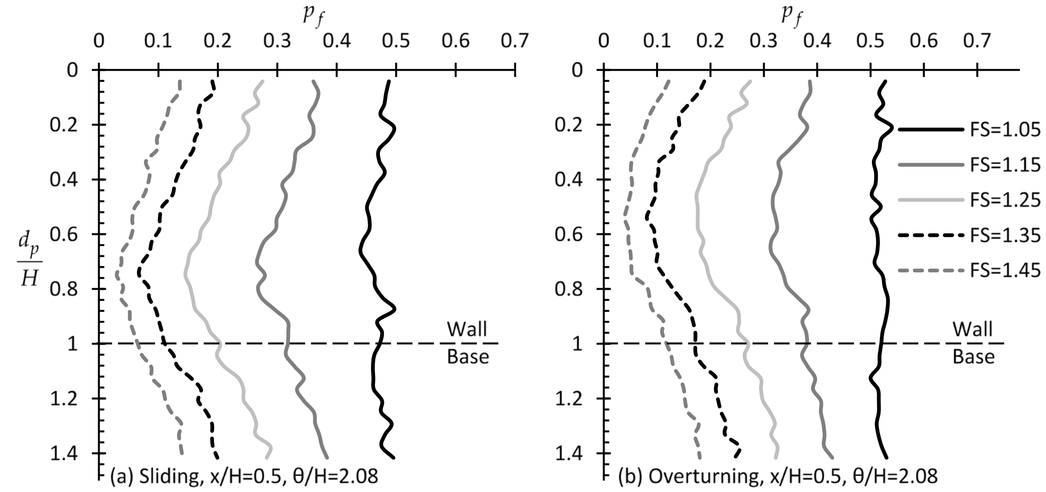


Figure 6.31: p_f versus d_p/H example curves for different FS values for the case of (a) sliding and (b) overturning wall.

6.3.2.7 Effect of Soil Anisotropy

The effect of soil anisotropy on the optimal sampling location will be investigated here by comparing the $\theta_h = \theta_v$ case with the $\theta_h = 10\theta_v$ case (see also 2.1.1.7). The reference wall-soil system with $\theta_v/H = \theta_h/H$ =2.08 will be compared with a respective one having $\theta_v/H = 2.08$ and $\theta_h/H = 20.8$. The variation of p_f with d_p/H for various x/H values is shown in Figure 6.32. From this figure it is inferred that, the horizontal distance plays, generally, no role in the statistical uncertainty, especially when x is smaller than H. Regarding the optimal sampling location, the soil anisotropy seems to have no effect on the optimal sampling depth both in the case of translating and rotating wall. However, for the anisotropic case smaller p_f are obtained.

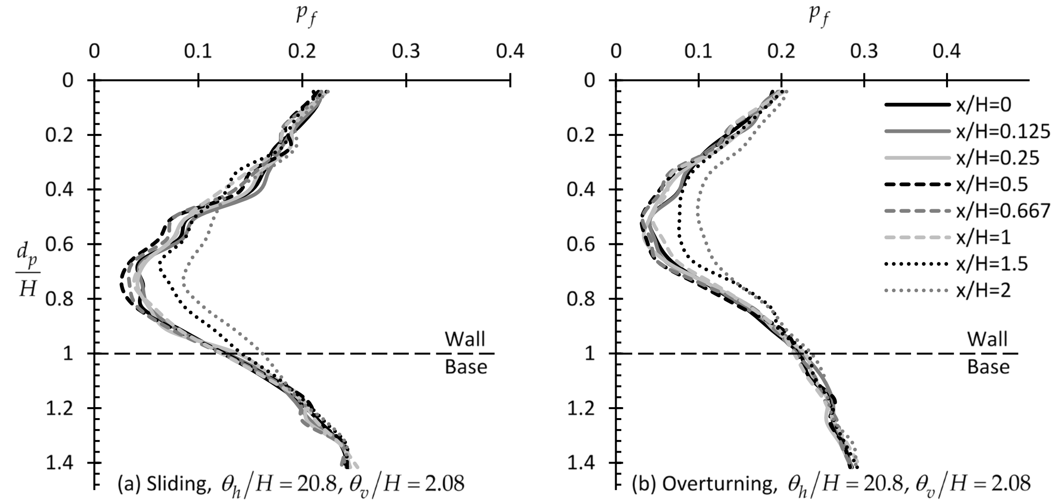


Figure 6.32: p_f versus d_p/H example curves for various x/H values for (a) sliding and (b) overturning wall considering anisotropic soil (to be compared with Figure c and Figure d respectively).

6.3.3 Sampling from a Domain

The minimum and maximum sampling domain length considered were 0.1m (rather referring to a single point) and 3.4m respectively. It is noted that, for all cases examined in this section the optimal sampling distance was found again to be at $x/H = 0.5$. Thus, the analysis below, generally, refers to the $x/H = 0.5$ case.

6.3.3.1 Effect of Spatial Correlation Length (θ)

From the p_f versus d_d/H charts of Figure 6.33 it is clear that the optimal horizontal sampling distance from the wall is again for x/H=0.5. Indeed, when the spatial correlation length is very small, a different sampling distance seems not to affect greatly the statistical error. On the contrary, as theta increases the role of horizontal distance becomes more and more significant. In the figure in question, x/H values ranging from 0 to 2 were considered, whilst FS was set equal to 1.25 (recall Equation (6.2)). In addition, it is advisable that the entire domain length for x/H=0.5 to be considered for minimizing the statistical error. Indeed, this practice may significantly reduce the statistical error. It is also interesting that, extending the sampling domain beyond the maximum depth of wall (i.e. d_d/H>1), the statistical error remains constant. Finally, from Figure 6.33 it is inferred that, a “worst case theta” exists. This is more obvious in Figure 6.34 showing the variation of p_f with d_d/H for various θ/H values and x/H=0.5.

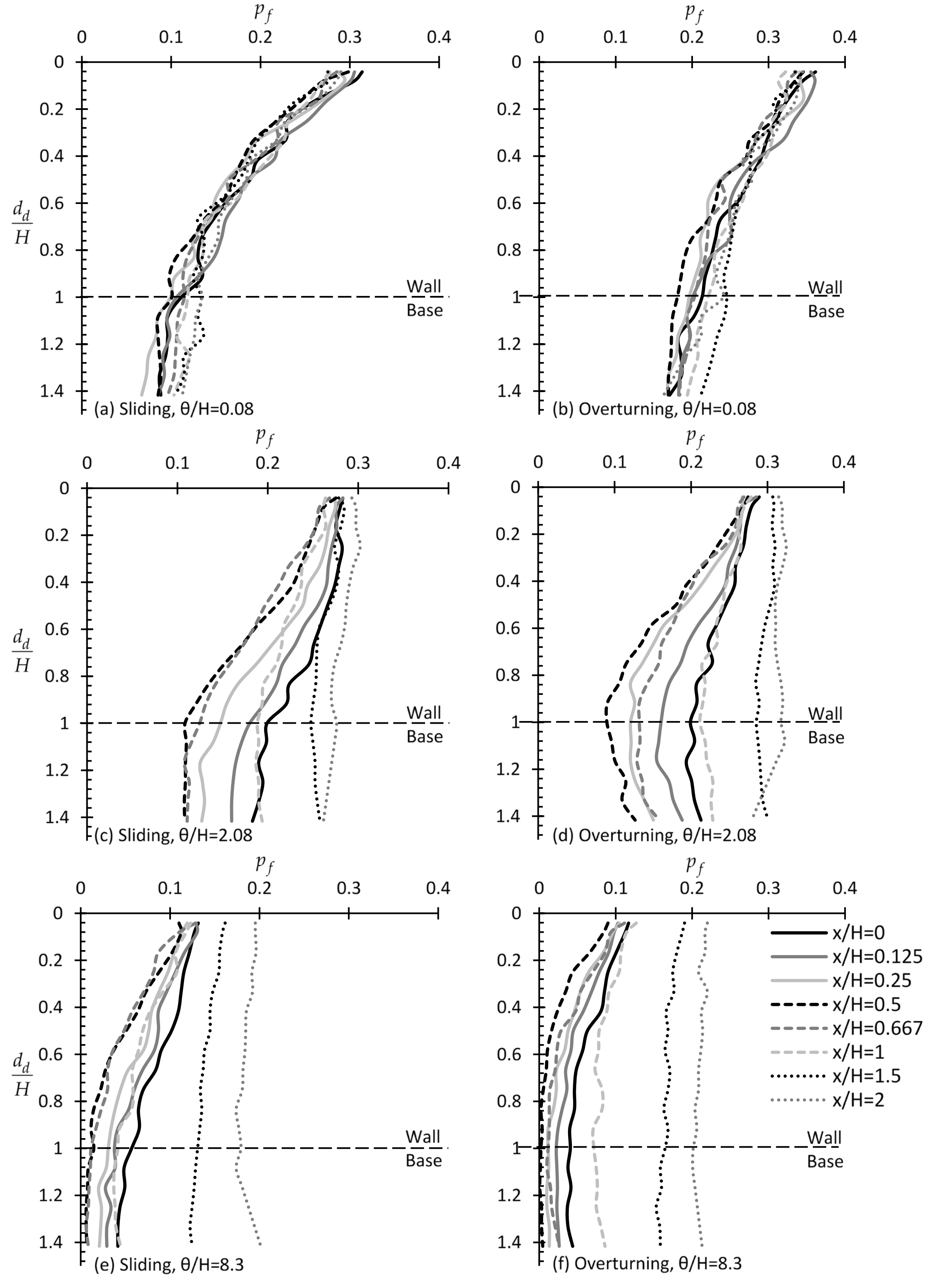


Figure 6.33: p_f versus d_d/H example relationships for different values of scaled correlation length (θ/H) and lateral distance from the wall face (x/H). Figure (a, c, e) shows the p_f variation for the case of sliding and (b, d, f) to the case of overturning wall.

.

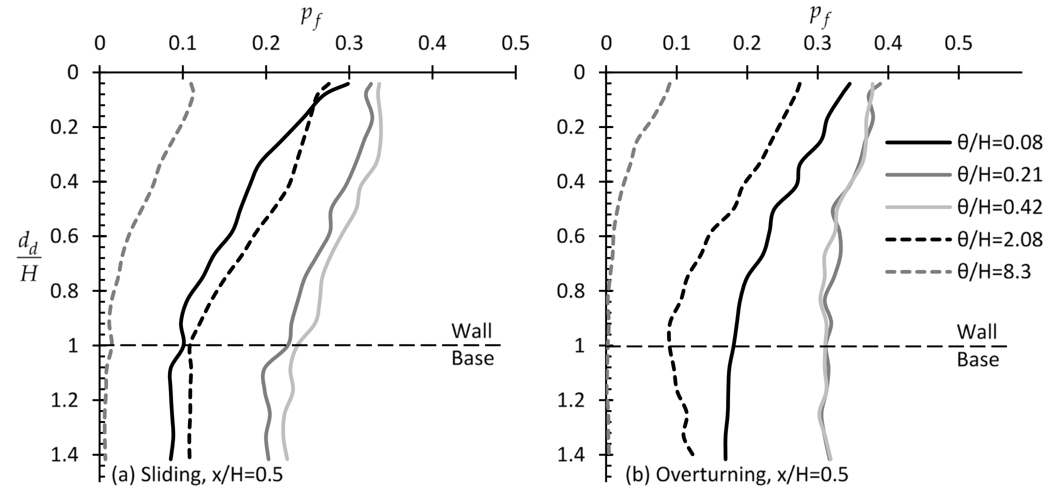


Figure 6.34: p_f versus d_d/H example relationships for the case of (a) sliding and (b) overturning wall by considering different scaled θ/H values.

6.3.3.2 Effect of Wall Roughness

Generally, the wall roughness has minor effect on the optimal sampling domain length, although, as expected (see Figure 6.35), it noticeably affects the failure probability. As shown in Figure 6.35, smaller p_f can be obtained when the wall is smooth, with the optimal sampling domain length being the entire wall height. Characteristically it is mentioned that, the minimum failure probability is obtained for d_d/H=1 and that, this probability remains constant for greater d_d/H values.

6.3.3.3 Effect of Wall Height

In this paragraph four wall heights were considered, i.e. H =1.4, 1.9, 2.4, 2.9m. Figure 6.36 presents the variation of p_f with d_d/H for these four cases. From this figure it is clear that the wall height largely affects the statistical error with the higher walls requiring greater sampling domain length, especially in the case of sliding wall. For sliding walls, it is advisable that the whole domain length be taken into account, whilst for rotating walls at least the half wall height.

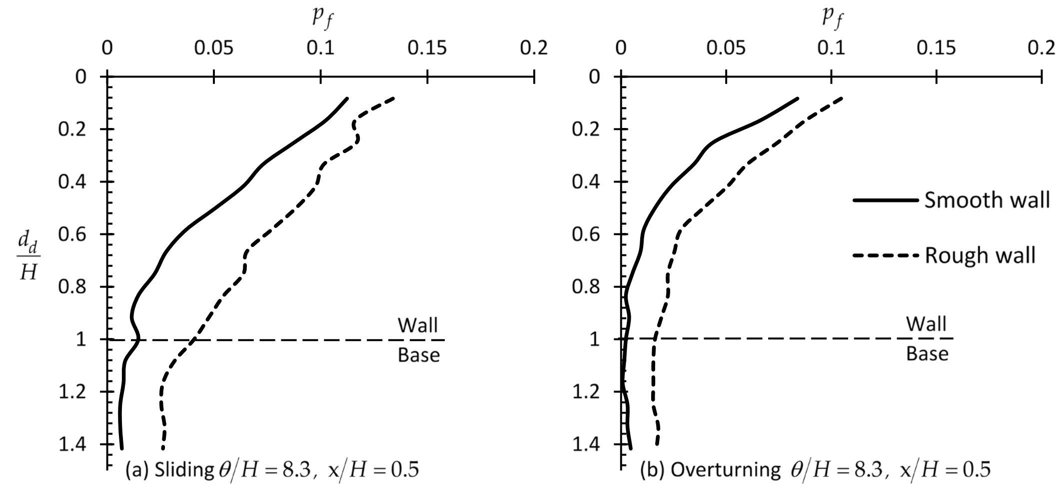


Figure 6.35: p_f versus d_d/H example curves for $\theta/H = 8.3$ and $x/H = 0$ (rough and smooth wall) for the case of (a) sliding and (b) overturning wall.

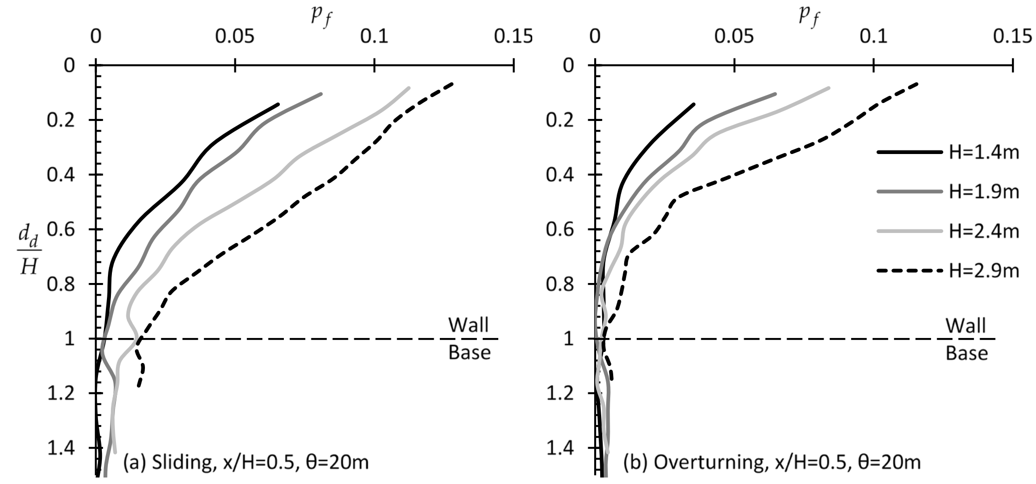


Figure 6.36: p_f versus d_d/H example curves for different wall heights H and $\theta = 20m$ for the case of (a) sliding and (b) overturning wall.

6.3.3.4 Effect of COV of ϕ'

In this paragraph, five COV values for ϕ' were considered, i.e. $COV = 0.1$, 0.2, 0.3, 0.4 and 0.5. The optimal horizontal sampling distance from the wall was found not to be affected by the COV of ϕ', where again the $x/H = 0.5$ case leads to the smaller probabilities of failure. Thus, only the $x/H = 0.5$ case will be presented here. From Figure 6.37 it is, generally, inferred that the COV of ϕ' has minor effect on the sampling domain length for both the case of sliding or rotating wall.

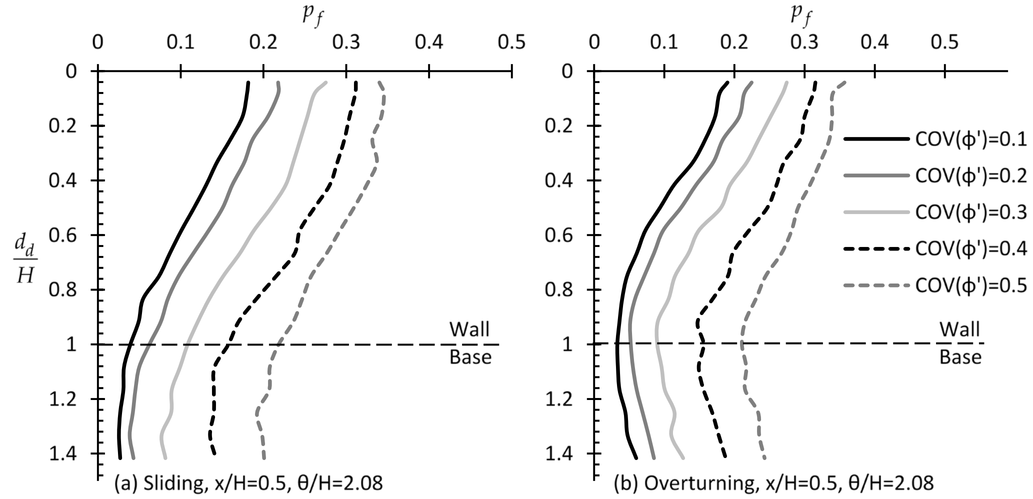


Figure 6.37: p_f versus d_d/H example relationships considering different values of COV of ϕ'. Figure (a) refers to the case of sliding wall, whilst figure (b) to the case of overturning wall.

6.3.3.5 Effect of the Factor of Safety (FS)

The variation of p_f with respect to d_d/H for different FS values is shown in Figure 6.38; the optimal sampling distance was also found to be at $x/H = 0.5$ for any FS value, thus, only this case is presented here. From Figure 6.38 it is obvious that the failure probability decreases as FS increases. However, the FS has no effect on the optimal sampling domain.

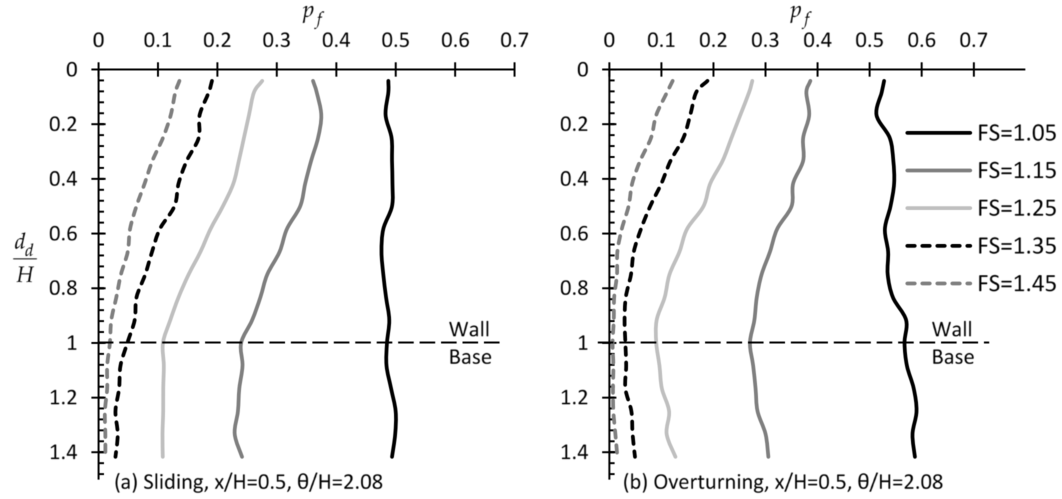


Figure 6.38: p_f versus d_d/H example curves for different FS values for the case of (a) sliding and (b) overturning wall.

6.3.3.6 Effect of Soil Anisotropy

In this paragraph the retaining soil is considered to be anisotropic having θ_h/H =20.8 and θ_v/H =2.08; for the isotropic case is stands that $\theta_h/H = \theta_v/H = \theta/H$= 2.08. As it is inferred from Figure 6.39, the soil anisotropy has minor effect both on the failure probability and the optimal sampling domain length. For both cases (translating and rotating wall) the statistical error remains relatively constant for horizontal sampling distances less than the wall height. By taking samples in distance greater than that, the statistical error increases slightly.

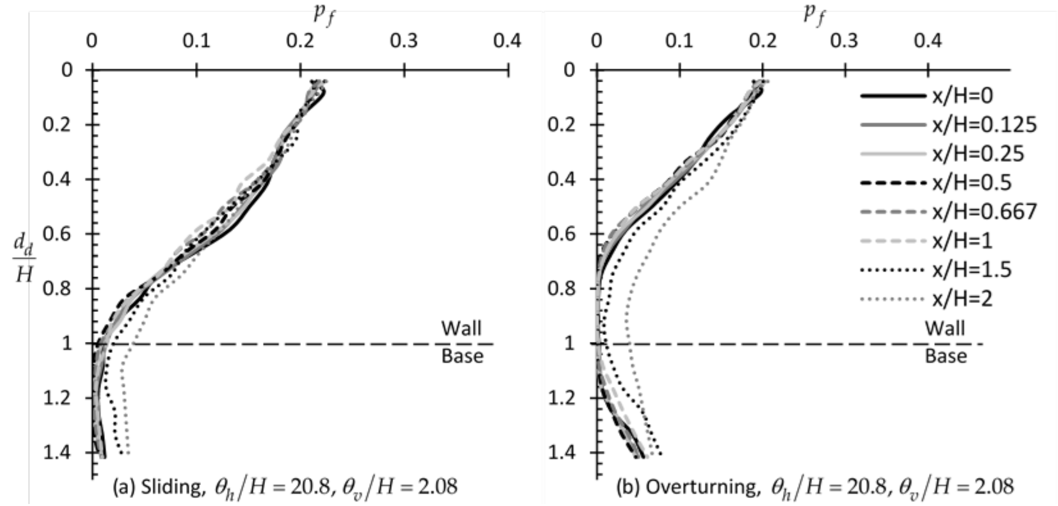


Figure 6.39: p_f versus d_d/H example curves for the case of (a) sliding and (b) overturning wall considering anisotropic soil (θ_h/H =20.8 and θ_v/H =2.08) and isotropic soil ($\theta_h/H = \theta_v/H = \theta/H$= 2.08).

6.3.4 Discussion

6.3.4.1 Optimal sampling locations

One of the main outcomes derived from the present analysis is that, the optimal horizontal sampling distance in the passive state of stress was found to be at x/H =0.5. Regarding the depth of the optimal sampling point, it was found that this lies at depth greater than the 2/3 or 1/2 of the wall height for the sliding and rotational mode of failure respectively; the exact depth depends on the spatial correlation length of the soil. This location lies on or in the close vicinity of Rankine's *45°-ϕ' /2* failure plane passing through the lower point of the wall. For the optimal sampling domain length, it is advisable that the entire wall height be considered.

6.3.4.2 The Importance of Targeted Field Investigation in Practice

The importance of targeted field investigation, where samples are taken from a priory known optimal locations, is highlighted here. For the three examples presented in this paragraph, the reference wall (and the finite element mesh) of Figure 6.24 will be used, whilst the material properties are given in Table 6.2. These materials differ from each other, in essence, in the spatial correlation length and only for the first material, in addition to the friction angle of soil, the unit weight is random field. Besides, as stated earlier, the mean and COV values of ϕ' have no effect on the optimal sampling location. The random field of ϕ' used in each example is shown in Figure 6.24, Figure 6.40 and Figure 6.41 respectively. The FS value is assumed unity (recall Equation (6.4)); this factor is discussed in the next paragraph.

Table 6.2: Summary of the characteristics of the soils used in the three examples (wall height *H*= 2.4 m).

Example	Random field(s)	Distribution	$\mu_{\phi'}$	μ_γ	COV	θ/H	Figure (1)
#1	ϕ', γ	Log-normal	30°	20 kN/m³	0.3	8.3	Figure 6.24
#2	ϕ'	Log-normal	30°	20 kN/m³	0.3	4.2	Figure 6.40
#3	ϕ'	Log-normal	30°	20 kN/m³	0.3	0.42	Figure 6.41

(1) Figures shown the random fields of ϕ'.

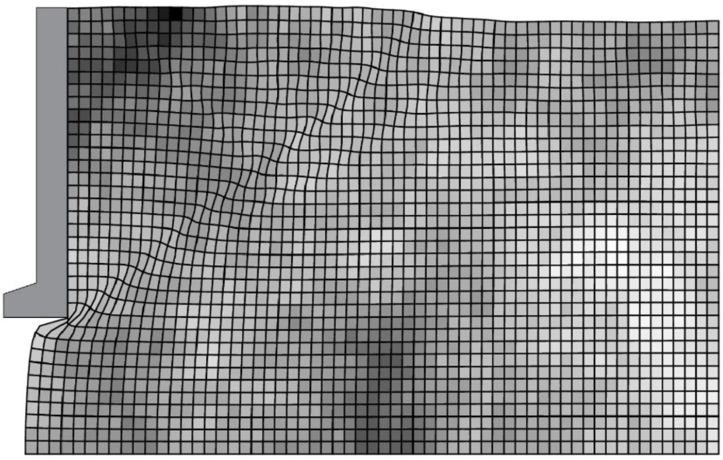


Figure 6.40: Graphical representation of the random field of ϕ' of Example #2 (θ/H= 4.2; see Table 6.2). Light areas correspond to lower friction angles and vice versa.

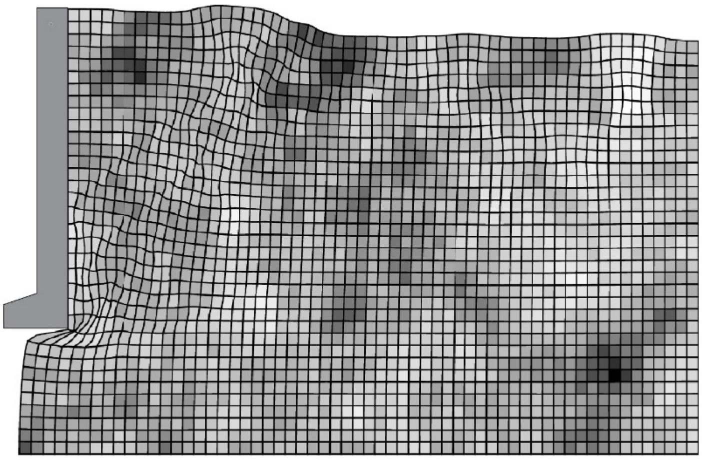


Figure 6.41: Graphical representation of the random field of ϕ' of Example #3 (θ/H= 0.42; see Table 6.2). Light areas correspond to lower friction angles and vice versa.

As on the active state analysis, the *predicted* resultant driving force (*F*) or moment (*M*) acting on the wall is compared against the respective "*actual*" ones. The *predicted F* and *M* values derive from a homogenous soil field characterized by the mean of the values sampled from the original (random) field. The results are presented in Figures 20-22 in $F_{predicted}/F_{"actual"}$ or $M_{predicted}/M_{"actual"}$ versus *x*/*H* form for various d_d/H values; the relative difference R_d defined as $F_{predicted}/F_{"actual"}$-1 and $M_{predicted}/M_{"actual"}$-1 for the forces and moments respectively is also given in each chart (secondary vertical axis; a negative value indicate design on the safe side and vice versa).

If the suggestions related to the horizontal distance from the wall and the domain length (*x*/*H*=0.5 and d_d/H=1 respectively) are valid, the $F_{predicted}/F_{"actual"}$ and $M_{predicted}/M_{"actual"}$ ratios for this specific sampling scenario should, logically, be equal to unity or very close to this value.

As shown in Figures 20-22, the $F_{predicted}/F_{"actual"}$ and $M_{predicted}/M_{"actual"}$ ratio values for *x*/*H*=0.5 are very close to unity or equal to unity, indicating the validity of author' suggestions. Indicatively it is mentioned that the abrupt increase of the $F_{predicted}/F_{"actual"}$ (or $M_{predicted}/M_{"actual"}$) versus *x*/*H* curves in Figure 6.42 between *x*/*H*=1 and 2 is attributed to the "dark" (strong) area appearing at this particular location, as shown in Figure 6.24. From Figure 6.42- Figure 6.44 it is also confirmed that a vertical sampling domain of length equal to the wall height gives better prediction for the destabilizing forces acting on the wall.

Comparison between the figures given in Section 3.1 with the respective ones given in Section 3.2, shows clearly that statistical uncertainty does not necessarily decreases with increasing number of samples. Indeed, the opposite may easily happen. For example, comparing the p_f ≈0.015 value for *x/H*=0.5 shown in Figure 6.25e (single point case;$d_p/H = 0.7$) with the p_f ≈0.17 value for *x/H*=2 (case of 24 sampling points; d_d/H =1) shown in Figure 6.33e, it is obvious that statistical uncertainty can only be minimized by targeted field investigation. Such examples can also be found in the present section; please compare the case of {*x/H* =2.5, d_d/H=1} with the {*x/H* =0.5, d_d/H =0.25} in Figure 6.44b giving $R_{d,M}$≈ -0.52 and -0.04 respectively.

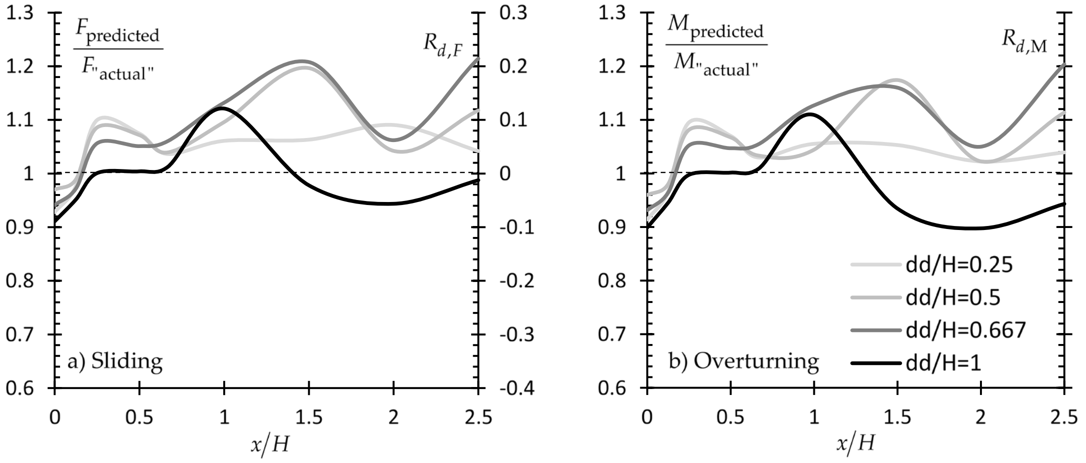


Figure 6.42: Example #1: $F_{predicted}/F_{"actual"}$ and $M_{predicted}/M_{"actual"}$ vs *x/H* curves for various d_d/H values and for both the sliding and overturning failure case (see also Table 6.2 and Figure 6.24).

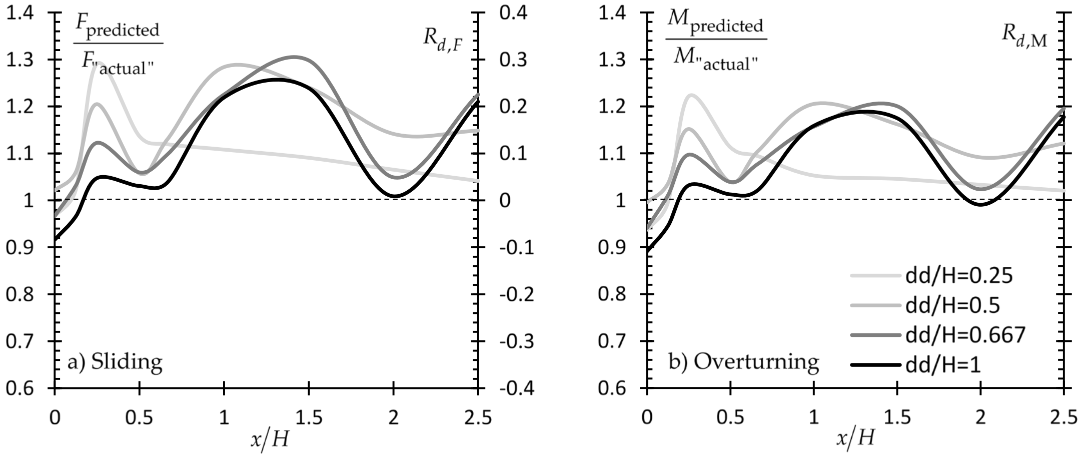


Figure 6.43: Example #2: $F_{predicted}/F_{"actual"}$ and $M_{predicted}/M_{"actual"}$ vs *x/H* curves for various d_d/H values and for both the sliding and overturning failure case (see also Table 6.2 and Figure 6.40).

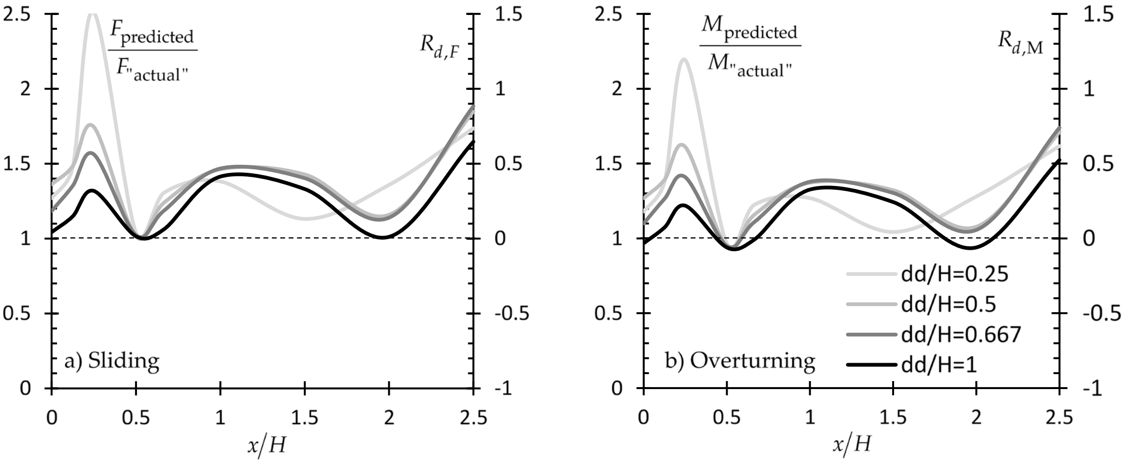


Figure 6.44: Example #3: $F_{predicted}/F_{\text{"}actual\text{"}}$ and $M_{predicted}/M_{\text{"}actual\text{"}}$ vs x/H curves for various d_d/H values and for both the sliding and overturning failure case (see also Table 6.2 and Figure 6.41).

6.3.4.3 Designing with Load and Resistance Factor Design (LRFD) Codes

The discussion on the design of earth retaining structures based on characteristic soil property values instead of the respective mean values is facilitated by the two example charts of Figure 6.45 (see also discussion of Chapter 3 and the relevant equations (2.1) and (2.2)). These charts refer to the case #3 presented in the previous paragraph (see also Table 6.2). Two cases are presented, the d_d/H=1 and the d_d/H=0.25. The figure in question refers to the sliding mode of failure, however, the respective curves for the overturning mode of failure do not differ appreciably. It is also mentioned that, in the example presented here, the partial material factor for the friction angle $\gamma_M=\gamma_\phi$ was set equal to unity.

From Figure 6.45 it is clear that, the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic values. The same conclusion as on the active state of failure can be drawn, i.e. the safety level can be increased by applying a statistical uncertainty partial factor (similar to the model factor γ_R used by Eurocode 7) or a unified and more conservative model factor to the resistances, which will absorb the statistical uncertainties related to the soil. In this respect, a partial factor equal to 1.25 has also been applied (FS=1.25; recall Equation (6.4)) in the present example. As shown in Figure 6.45, the use of such a factor simply displaces downwards (that is, to the safe side) the $F_{predicted}/F_{\text{"}actual\text{"}}$ (or $M_{predicted}/M_{\text{"}actual\text{"}}$) versus x/H curves.

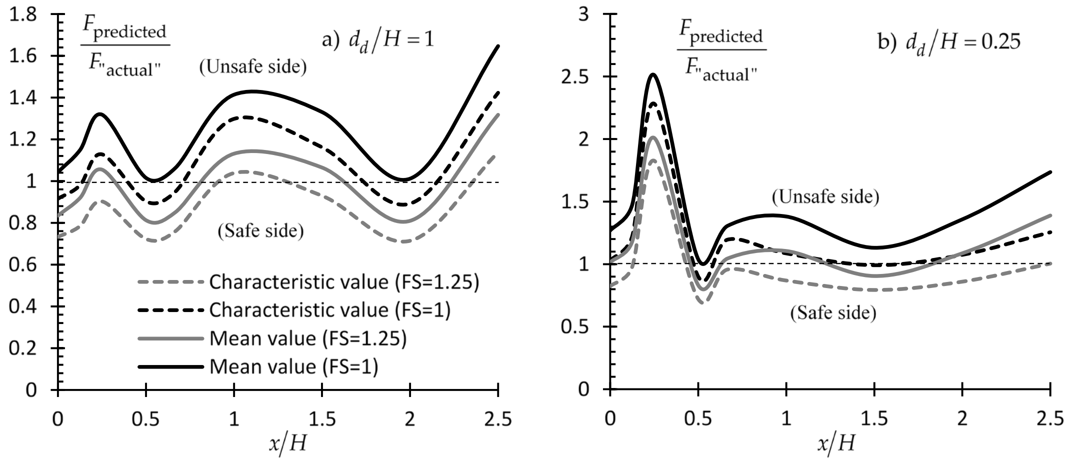


Figure 6.45: $F_{predicted}/F_{"actual"}$ vs. x/H curves using both mean and characteristic values (dashed and solid lines respectively) for *FS*=1 and 1.25. Figure referring to the case of a sliding wall and to two sampling domain cases (d_d/H = 1 (figure a) and to d_d/H = 0.25 (figure b)). The reference wall was used. Soil characteristics as shown in Table 6.2 (Example #3).

6.4 Conclusions

The present research clearly shows that statistical uncertainty inserted in to analysis of retaining structures, related to soil properties, may be significant and that it can only be minimized by performing targeted field investigation; the latter is defined by the number and location of sampling points. Two main sampling strategies were investigated, namely, sampling from a single point and sampling from a domain, through an extensive parametric analysis.

One of the main findings of the present Chapter is that the optimal horizontal sampling location in the active state is at $x/H = 0$ -that is, immediately adjacent to the wall face, while the optimal horizontal sampling location in the passive state is half wall height away from the wall face. Regarding the depth of the *optimal sampling point*, for both states it was found that this lies at depth greater than the 2/3 or 1/2 of the wall height for the sliding and rotational mode of failure respectively; the exact depth depends on the spatial correlation length of the soil. For the *optimal sampling domain length*, it is advisable that the entire wall height be considered.

In addition, it was observed that, the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic soil property values. Moreover, despite the conservatism which is inserted in the analysis using the characteristic value concept, the characteristic values alone, as shown, cannot guaranty a conservative enough engineering study. The safety level can be increased by applying a statistical uncertainty partial factor (similar to the model factor γ_R used by Eurocode 7) or a unified and more conservative model factor to the resistances, which will absorb the statistical uncertainties related to the soil.

7 Reducing statistical uncertainty in elastic settlement and bearing capacity analysis of shallow foundations relying on targeted field investigation

The present chapter deals with the practical problem of reducing statistical uncertainty in elastic settlement and bearing capacity analysis of shallow foundations relying on targeted field investigation aiming at an optimal design. This is done in an advance probabilistic framework using the Random Finite Element Method (RFEM) [61] properly considering sampling of soil properties. In this respect, the open source RSETL2D and RBEAR2D programs, has been modified as to include the function of sampling of soil property values from the generated random fields and return the failure probability of footing against excessive settlement and bearing capacity respectively. Two sampling strategies are examined: a) sampling from a single point and b) sampling a domain. The present analysis will show that the statistical error in an elastic settlement and bearing capacity analysis can be effectively minimized only when targeted field investigation is carried out.

7.1 Elastic settlement analysis

7.1.1 Two-dimensional probabilistic elastic settlement analysis based on RFEM

As mentioned, the present analysis was based on the open source RSETL2D program. The program involves generation and mapping of the elastic modulus of soil (E; which is treated as random field) onto a finite element mesh using the Local Average Subdivision method. The RSETL2D program calculates the settlement induced by a single strip footing (or a pair of strip footings) founded on a soil having spatially random E; the RSETL2D program can return the settlement induced at any finite element node, however, in the parametric analysis that follows, the settlement is calculated at the centre of the footing. The procedure is repeated m times; m is the number of realizations, where, each RFEM realization refers to a new random field of E.

The footing(s) is (are) assumed to be founded on a soil layer underlain by bedrock. The physical problem is represented using a two-dimensional (plane-strain) model as shown in Figure 7.1. The soil mass is discretized into 4-noded quadrilateral elements. The nodes along the left and right boundary of the finite-element model are constrained against horizontal displacement but are free to slide vertically, while the nodes on the horizontal boundary are fixed. The footing(s) is (are) assumed to be rough and rigid, undergoing no rotation. A unit force *P* (per unit length in the out-of-plane direction) is applied to each footing — since elastic settlement is directly proportional to *P*.

For the needs of the present research, the original RSETL2D program has been extended by the author as to:

- virtually sample elastic modulus values from the random field generated in each RFEM realization,
- calculate the footing settlement (again in each RFEM realization) considering that the soil is homogenous, having *E* equal to the mean of the values sampled (this settlement is calculated in addition to the settlement of footing lying on spatially random soil) and
- estimate the failure probability of the footing.

The latter is defined by the fraction of the realizations resulted in failure over the total number of realizations. In each RFEM realization, “failure” is considered to have occurred when the “actual” settlement value, referring to the spatially random soil, is greater than the respective predicted value, referring to the spatially uniform soil. That is, it stands that

$$p_f = P\left[\rho_{\text{"actual"}} > \rho_{predicted}\right] \tag{7.1}$$

where, the symbol ρ denotes footing settlement.

The modified program was validated as follows. First a given footing was solved using in the original RSETL2D program a deterministic soil modulus value. Then the same footing was solved with the modified program using values sampled from various points (because the same deterministic soil modulus value was spread out in the finite element mess, all *E* values sampled were the same). The two programs gave exactly the same results, indicating that the sampling function was correctly embedded into the original program.

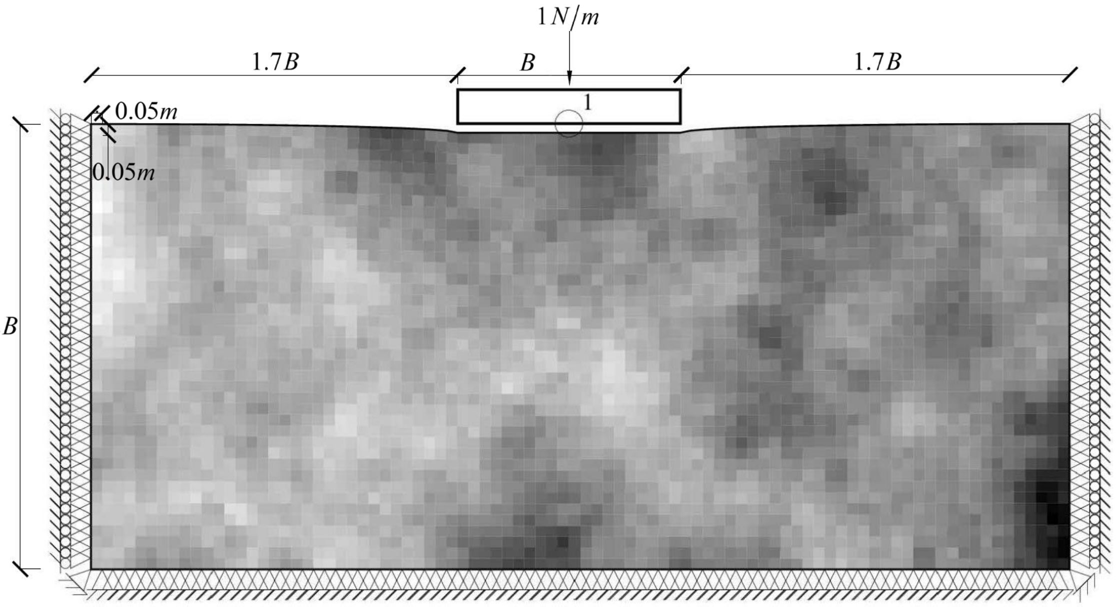


Figure 7.1: Typical deformed mesh considering $\theta/B = 10$. Figure showing the 20-element footing on the surface of the 88 x 40 (H x V) mesh. The darker elements indicate stiffer soil (greater *E* value).

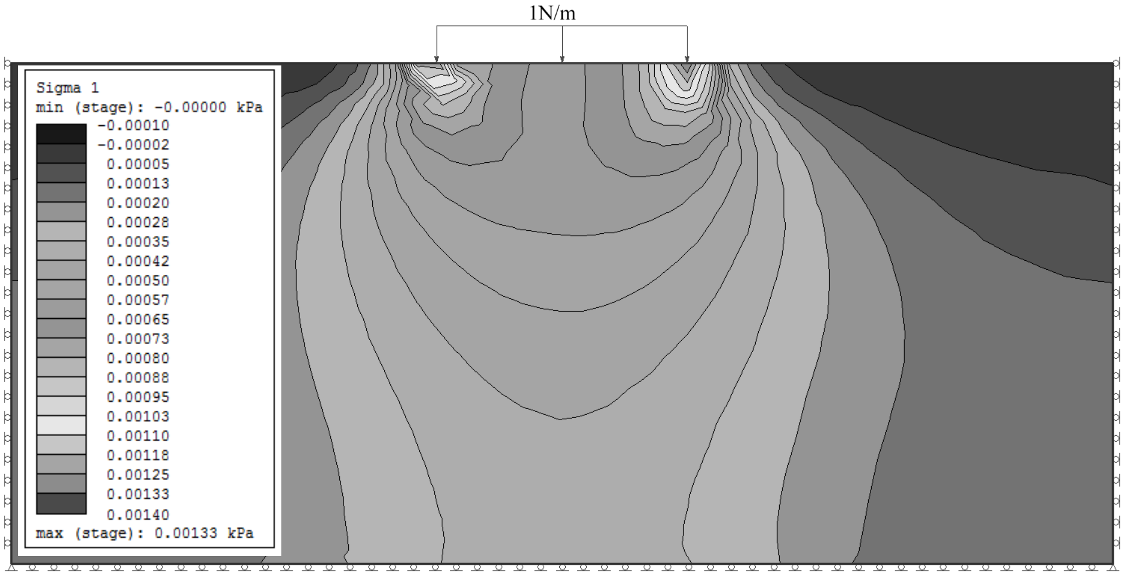


Figure 7.2: Stress regime in the vicinity under the footing. Contours of major principal stress (σ_1) obtained from the elastic analysis performed using the Rocscience software RS2 [183]. Figure referring to homogenous material field.

7.1.2 Parametric study for determining the optimal sampling strategy

This Chapter deals with the case of a single strip footing. Both the sampling from a single point and the sampling from an entire domain strategies are investigated through an extensive parametric analysis of the factors controlling the magnitude of settlement for defining the optimal sampling strategy that minimizes the probability of failure and thus the statistical error. The error is quantified comparing the probability of failure value obtained based on different sampling scenarios.

In the finite analysis that follows, the soil mass is discretized into a 88 (horizontal direction) by 40 (vertical direction) mesh, consisting of four-noded square elements having edge 0.05m. The strip footing occupies width on the surface of the finite element mesh equal to 20 elements (i.e. B=20 x 0.05m=1 m, called hereafter "reference footing"; other footing widths will also be considered in a later sections). A typical deformed mesh is shown in Figure 7.1. The effect of the distance between the edge of the footing and the respective lateral boundary was investigated prior to the analysis. As shown in Appendix IV, the error inserted considering a 20-element footing centered on the surface of a 88-element mesh is negligible; that makes a free distance between each edge of footing and the respective lateral boundary equal to $1.7B$. The same distance of $1.7B$ was kept the same for the other footing widths considered (i.e. B=1, 2 and 3m); however, the element size for the B= 2 and 3m footings was 0.1 m and 0.15 m respectively.

In the present Section, only E is treated as random field. According to Fenton and Griffiths [61], the Poisson's ratio, v, have a smaller relative spatial variability and only a second-order importance to settlement statistics. Generally, when not mentioned herein, v=0.25, whilst $\mu_E = 1$Pa (E is assumed that follows a log-normal distribution). Moreover, it is mentioned that the footing is subjected to a centrally applied vertical force of $P = 1$ N/m (unit force per unit length in the out-of-plane direction). A Markovian spatial correlation function has been adopted (see Equation (5.2)).

Aiming at finding the optimal sampling strategy, the following parameters will be examined: the sampling depth (d_p) and horizontal distance for the case of sampling from a single point (measured from the soil surface and the axis of strip footing respectively), the sampling length (d_d) and horizontal distance of the continuous probing test location for the case of sampling an entire domain (measured as in the previous case), the spatial correlation length of soil (θ), the footing width (B), the coefficient of variation (COV) of E of soil, the effect of the Poisson's ratio (v) value of soil and the anisotropy of soil mass $(\theta_h \neq \theta_v)$. Four sampling scenarios are indicatively shown in Figure 7.3 (scenarios A and B refer to a single sampling point, whilst scenarios C and D to continuous probing tests).

The optimal sampling point or domain will be identified by comparing the failure probability values (p_f) derived from various sampling strategies. The number of realizations was set equal to 10,000; this number, as shown in Appendix IV, can be considered adequate for the needs of the present analysis.

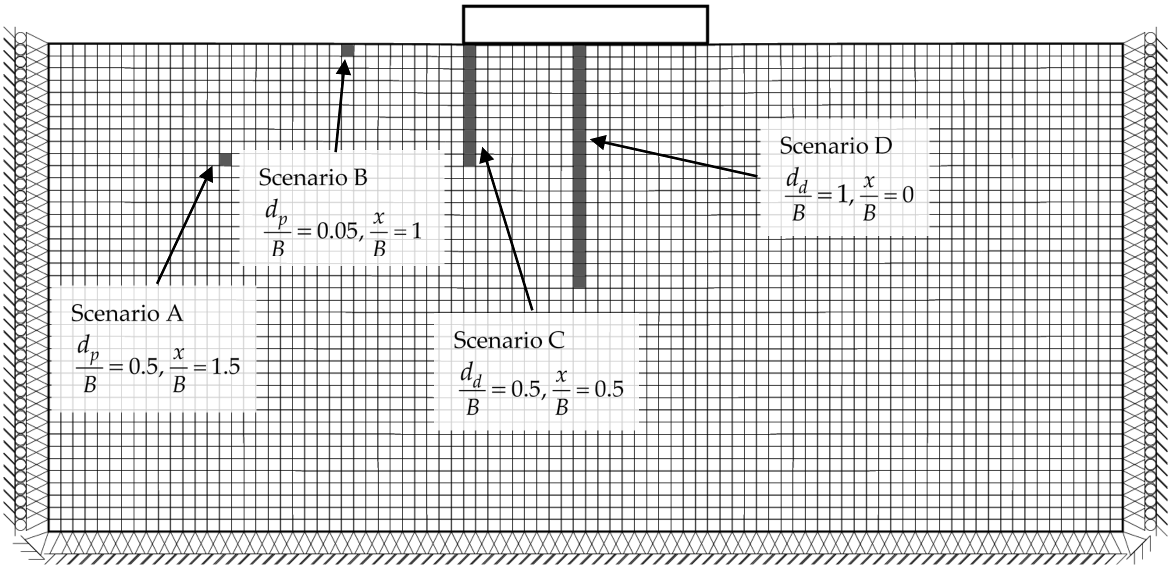


Figure 7.3: Graphical representation of single sampling point (d_p; Scenario A and B) and sampling domain length (d_d; Scenario C and D) sampling strategies.

7.1.3 Sampling from a single point

7.1.3.1 Effect of spatial correlation length ($\boldsymbol{\theta}$)

Example charts showing the variation of p_f with respect to d_p/B for various θ/B values are given in Figure 7.4. The soil mass is considered isotropic. From the figure in question it is inferred that the optimal sampling location lies at the center of strip foundation ($x/B = 0$) and at depth approximately 0.7B. It is also observed that, as θ/B tends to zero, the p_f value tends to a single value, that is, p_f becomes independent of the sampling depth. The effect of soil anisotropy (for more details about soil anisotropy please see Chapter 4 , Section 2.1.1.7) on the optimal sampling location is investigated herein by comparing the $\theta_h = \theta_v$ case with the $\theta_h = 10\theta_v$ case, i.e. the reference footing-soil system with $\theta_v/B = \theta_h/B$ =5 will be compared with a respective one having $\theta_v/B = 5$ and $\theta_h/B = 50$. In this respect, the optimal sampling location was found not to be affected by soil anisotropy (see heavy bold line in Figure 7.4c).

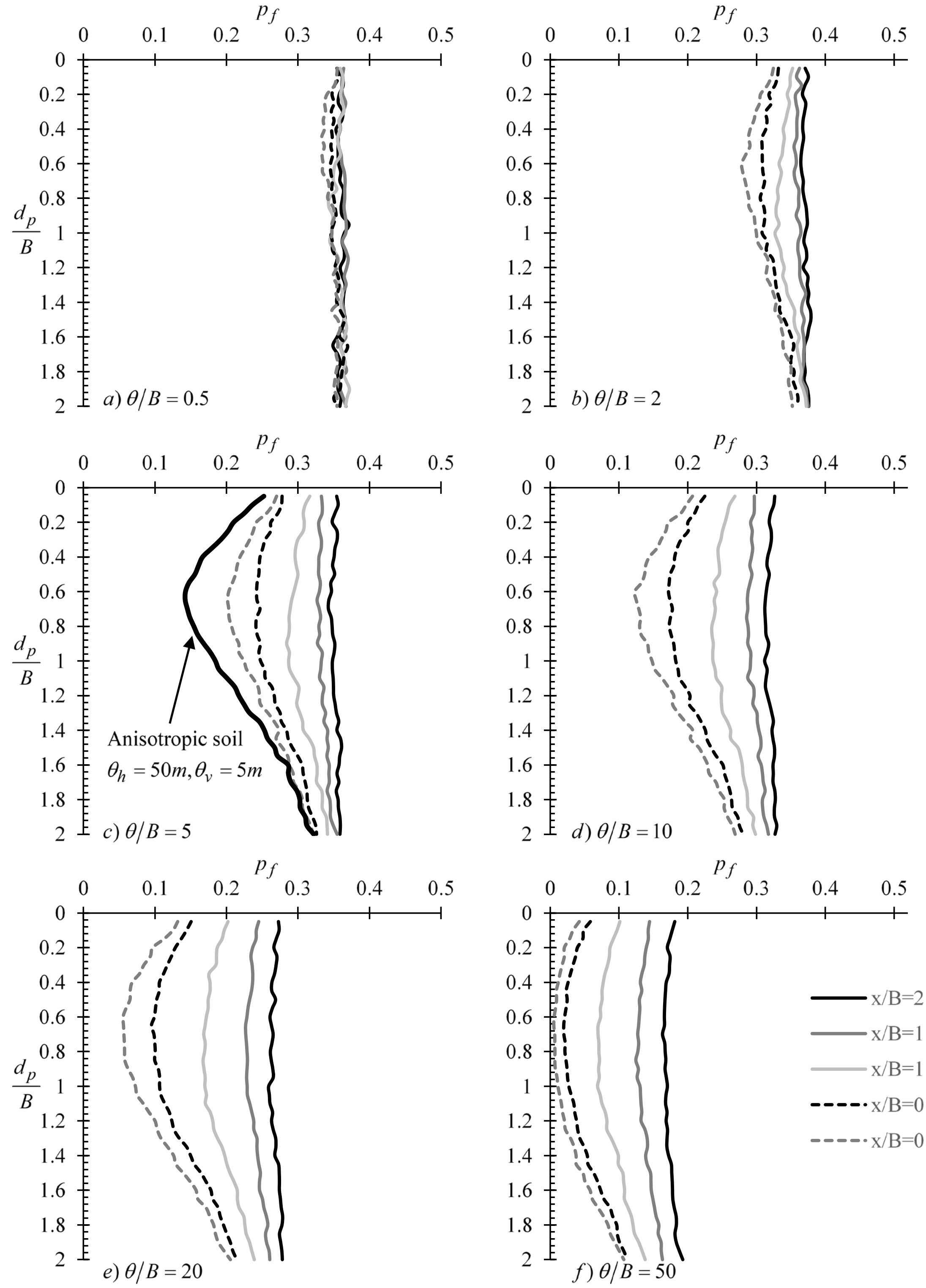


Figure 7.4: p_f versus d_p/B example curves for various θ/B and x/B values. The thick continuous line in Figure c is for the anisotropic case and refers to the x/B =0 case.

7.1.3.2 Effect of footing width (B)

The variation of p_f with d_p/B is shown in Figure 7.5 for three footing widths, i.e. B =1, 2 and 3 m. The analysis showed that, the optimal sampling depth is not affected by the footing width, whilst again the x/B =0 case leads to the smaller statistical error. The three curves in Figure 7.5 refer to x/B =0.

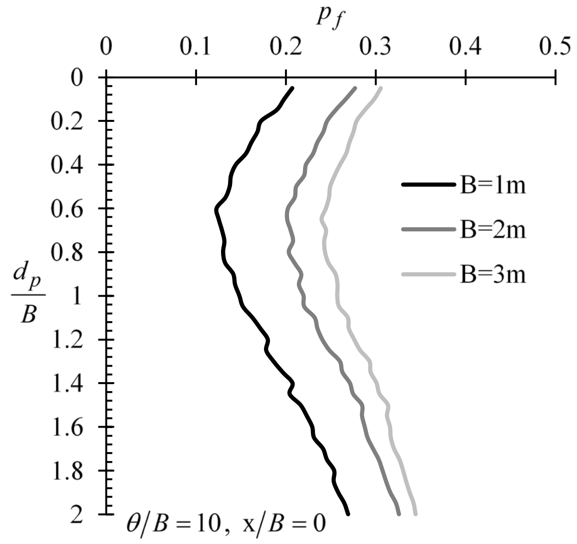


Figure 7.5: p_f versus d_p/B example curves for different footing widths B and $\theta/B = 10$.

7.1.3.3 Effect of COV of the elastic constants of soil

In this paragraph five COV values of soil modulus, E, were considered, namely COV = 0.1, 0.2, 0.3, 0.4 and 0.5. The analysis showed that the optimal sampling depth is not affected by the COV of E, where again the smaller statistical error is found for x/B =0. The five curves in Figure 7.6 refer to x/B =0.

7.1.3.4 Effect of the elastic constant values of soil

For all cases considered above, the Poisson's ratio of soil was equal to 0.25. Parametric study on the effect of the v on the optimal sampling location, however, showed that the latter is not affected by the parameter in question. The following v values were considered, i.e. v= 0, 0.1, 0.25, 0.4 and 0.495. The five curves shown in Figure 7.7 refer to x/B =0. Similarly to the Poisson's ratio value, the optimal sampling location is not affected by the mean value of the elastic modulus of soil.

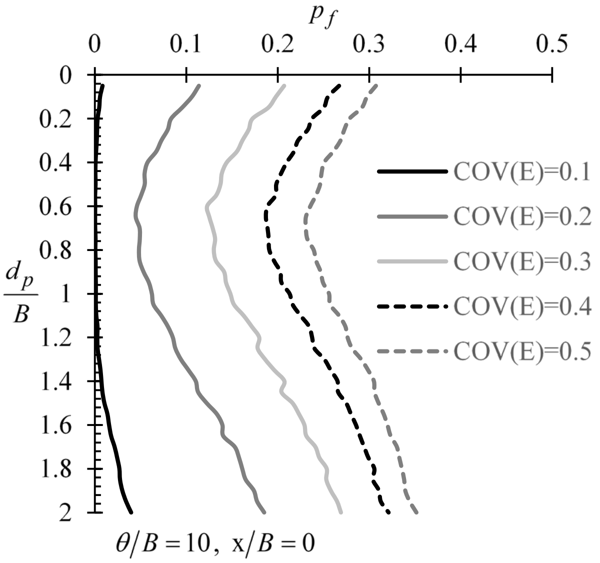


Figure 7.6: p_f versus d_p/B example relationships by considering different values of COV of E.

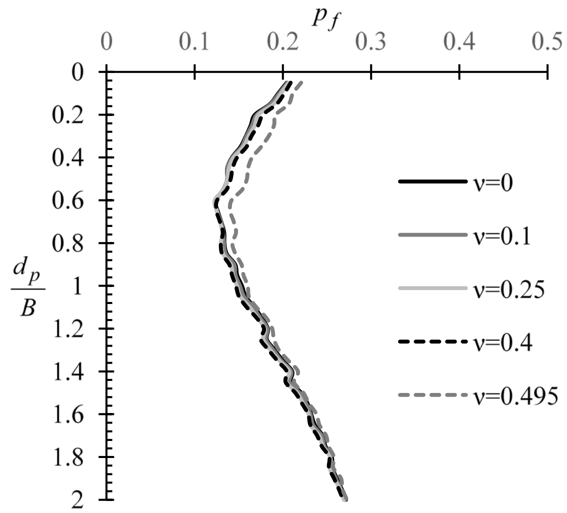


Figure 7.7: p_f versus d_p/B example curves for different v values and $\theta/B = 10$.

7.1.4 Sampling from an entire domain

In the parametric analysis carried out, three footing widths were considered, i.e. B=1, 2 and 3 m, whilst the distance between two successive sampling points (in the vertical direction) was B/20 for all cases. The maximum sampling domain length considered was always two times the footing width B. The arithmetic mean of the soil elastic modulus values sampled was used in the analysis. It is noted that, for all cases examined, the optimal sampling distance was found again to be at x/B =0. Thus, the analysis below generally refers to x/B =0.

7.1.4.1 Effect of spatial correlation length ($\boldsymbol{\theta}$)

Example charts showing the variation of p_f with respect to d_d/B for various θ/B and x/B values are given in Figure 7.8. From this figure it is inferred that the optimal horizontal sampling distance from the footing center for every θ/B value is again for x/B =0. It is also observed that the required domain length is smaller for greater θ values of soil. Given now that, soil samples will be taken from x/B =0, it is advisable a domain length at least $2B$ to be considered. This practice may significantly reduce the statistical error. The variation of p_f with d_d/B for the case of anisotropic soil mass ($\theta_v/B = 5$ and $\theta_h/B = 50$) is shown in Figure 7.8c (heavy bold line). Generally, it can be said that, the required sampling domain length was found to be somewhat smaller as compared to the case of the isotropic soil mass.

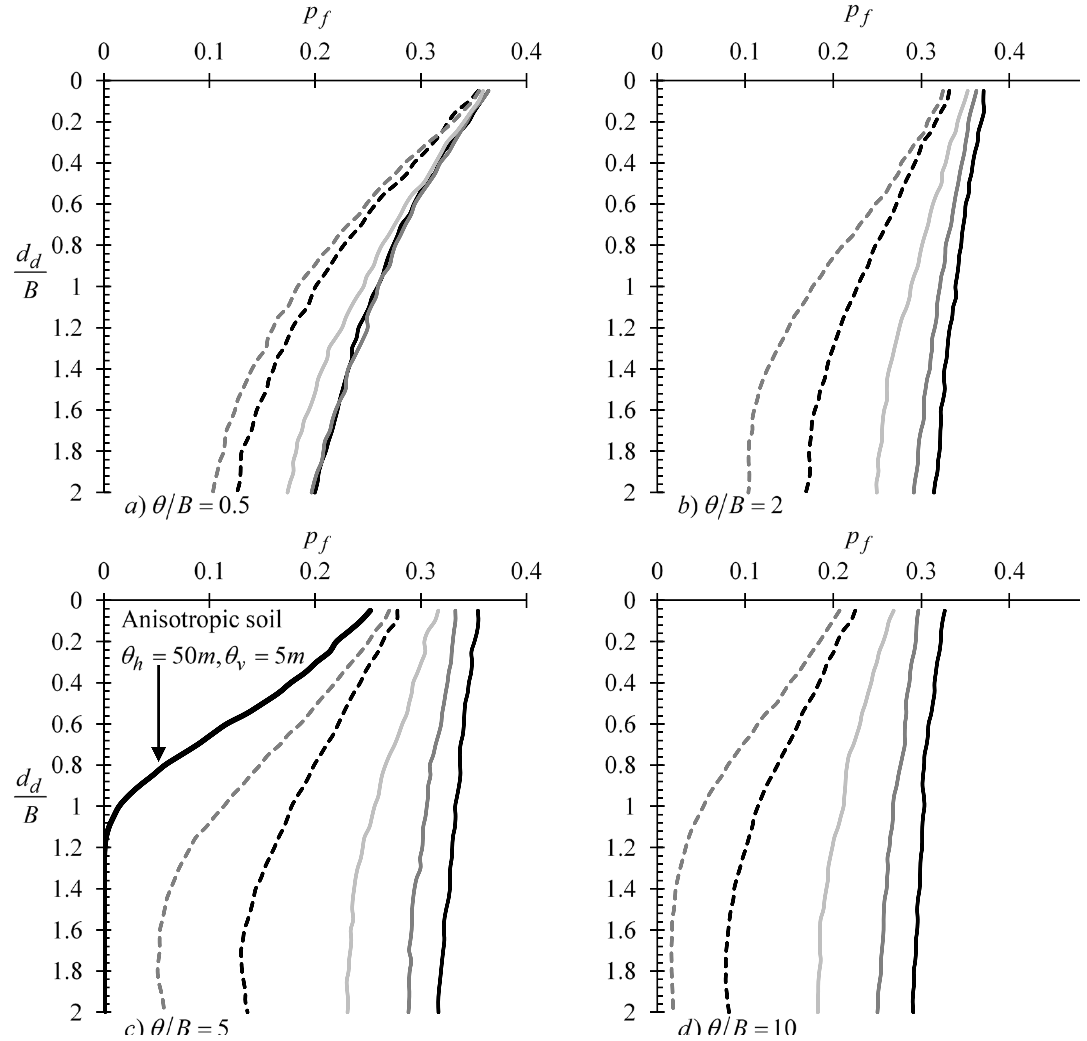


Figure 7.8: p_f versus d_p/B example curves for various θ/B and x/B values. The thick continuous line in Figure 7.8c is for the anisotropic case and refers to the x/B =0 case.

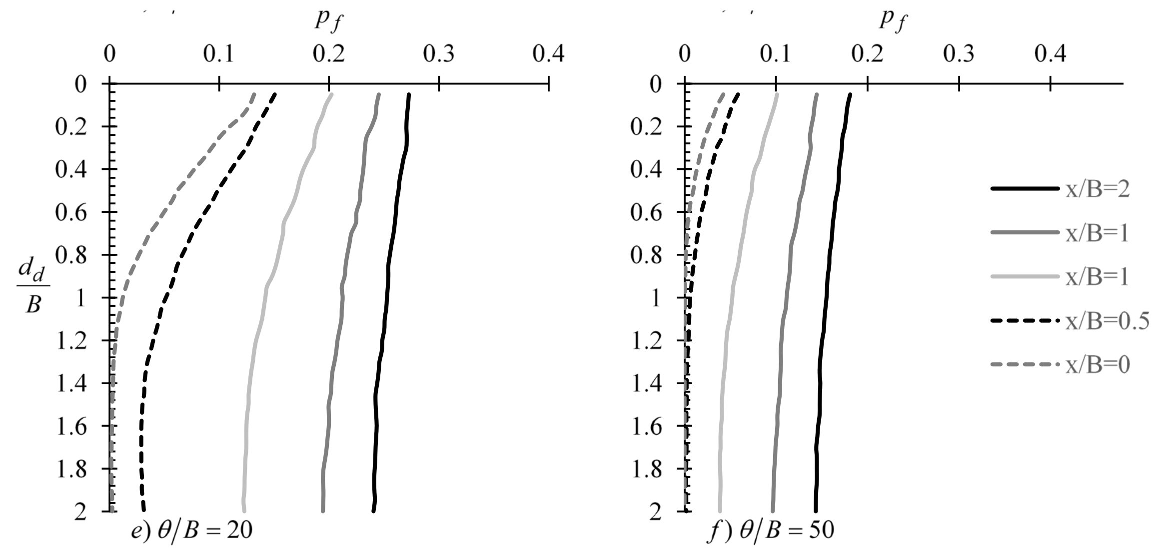


Figure 7.8: Continued.

7.1.4.2 Effect of footing width (B)

In this paragraph three footing widths were considered, i.e. B =1, 2, 3m. Figure 7.9 presents the variation of p_f with d_d/B for these three cases. From this figure, it is clear that the footing width has only minor influence on the required sampling domain length.

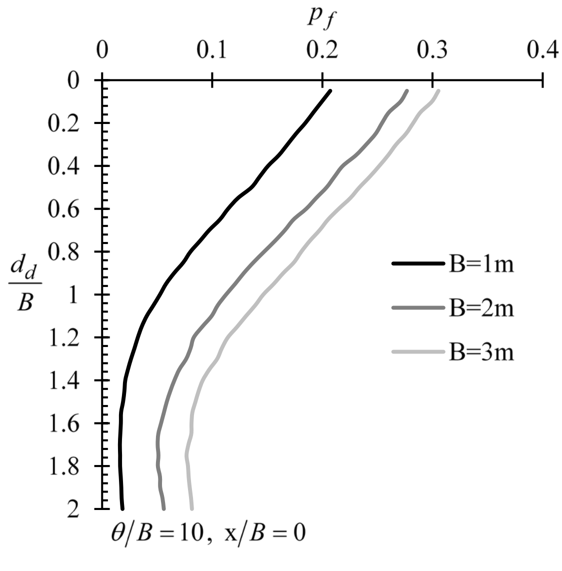


Figure 7.9: p_f versus d_p/B example curves for different footing widths B and $\theta/B = 10$.

7.1.4.3 Effect of COV of the elastic constants of soil

In this paragraph, five COV values of E were considered, i.e. COV = 0.1, 0.2, 0.3, 0.4 and 0.5. The optimal horizontal sampling distance from the footing center was found not to be affected by the COV of E, whilst again the x/B =0 case leads to the smaller probabilities of failure. Thus, only the x/B =0 case will be presented here. From Figure 7.10 it is generally inferred that the $COV$of E largely affects the optimal sampling domain length. More specifically, as the $COV$of E increases the sampling domain length required to minimize the statistical error also increases.

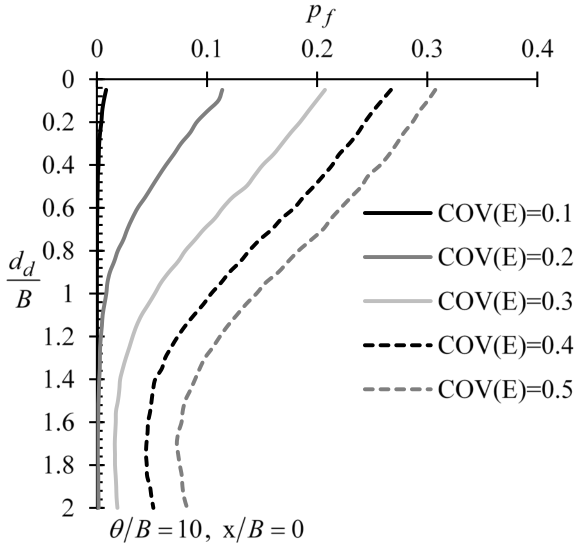


Figure 7.10: p_f versus d_p/B example relationships by considering different values of COV of *E*.

7.1.4.4 *Effect of the elastic constant values of soil*

The variation of p_f with respect to d_d/B for different Poisson's ratio values is shown is Figure 7.11; the optimal sampling distance was also found to be at x/B =0 for any *v* value, thus, only this case is presented in this paragraph. From Figure 7.11 it is obvious that the Poisson's ratio value has no effect on the optimal sampling domain length. The same stands for the mean value of elastic modulus of soil.

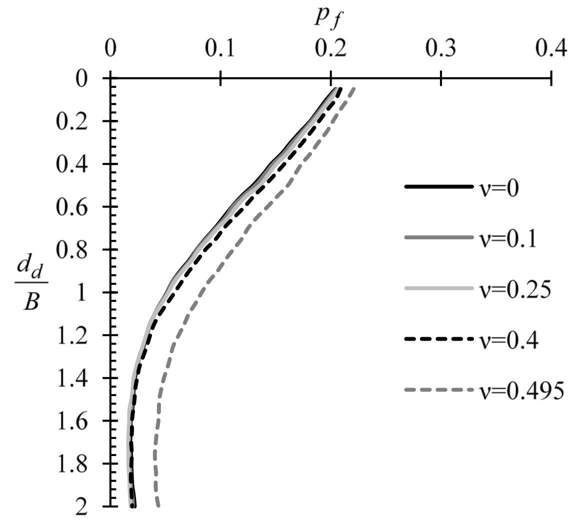


Figure 7.11: p_f versus d_p/B example curves for different v values and $\theta/B = 10$.

7.1.5 Discussion

7.1.5.1 The importance of targeted field investigation in practice

In the four examples presented below, the footing and the mesh / boundary conditions are the same with those presented in Section 7.1.2 of the current Chapter (i.e. the reference footing). The four examples, in essence, differ in the spatial correlation length, since as shown in the present research, the mean and COV values of E have no effect on the optimal sampling location. The random field of E used in each example is shown in Figures 7.1, 7.12, 7.13 and 7.14 respectively. It is reminded that, the darker elements indicate stiffer soil and vice versa.

Table 7.1: Summary of the characteristics of the soils used in the four examples (footing width *B*= 1 m).

Example	Random field(s)	Distribution	μ_E (Pa)	v	COV	θ/B	Figure ⁽¹⁾
#1	*E*	Log-normal	1	0.25	0.3	10	Figure 7.1
#2	*E*	Log-normal	1	0.25	0.3	0.5	Figure 7.12
#3	*E*	Log-normal	1	0.25	0.3	5	Figure 7.13
#4	*E*	Log-normal	1	0.25	0.3	50	Figure 7.14

(1) Figures shown the random fields of *E*.

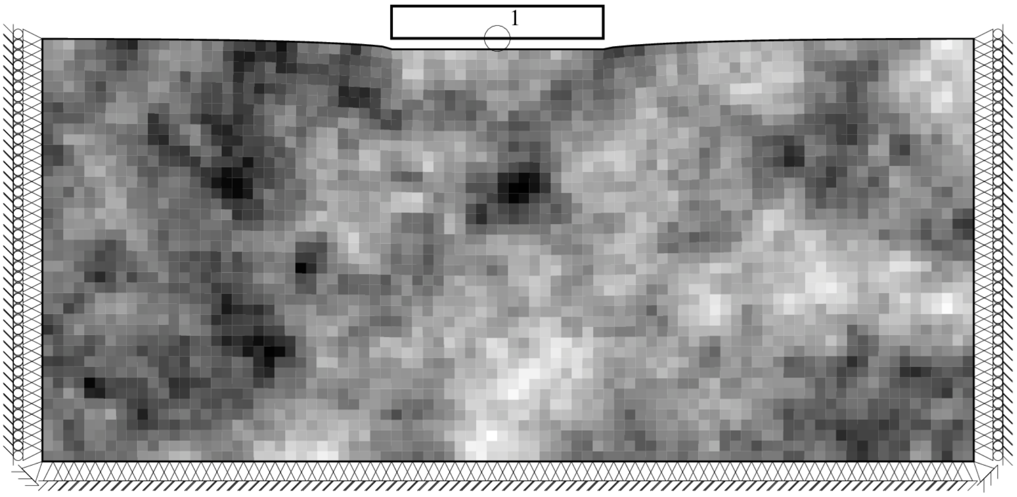


Figure 7.12: Graphical representation of the random field of *E* of Example #2 (θ/B=0.5; see Table 7.1).

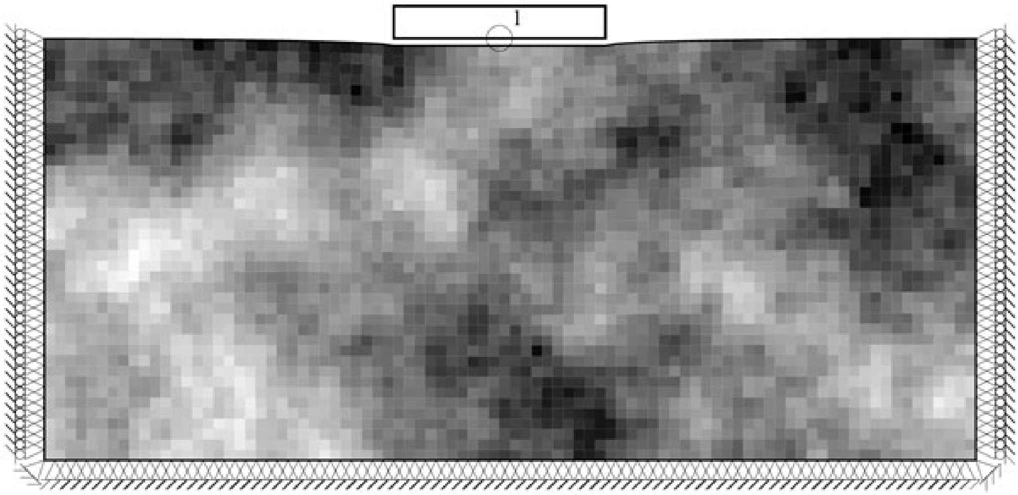


Figure 7.13: Graphical representation of the random field of *E* of Example #3 (θ/B=0.5; see Table 7.1).

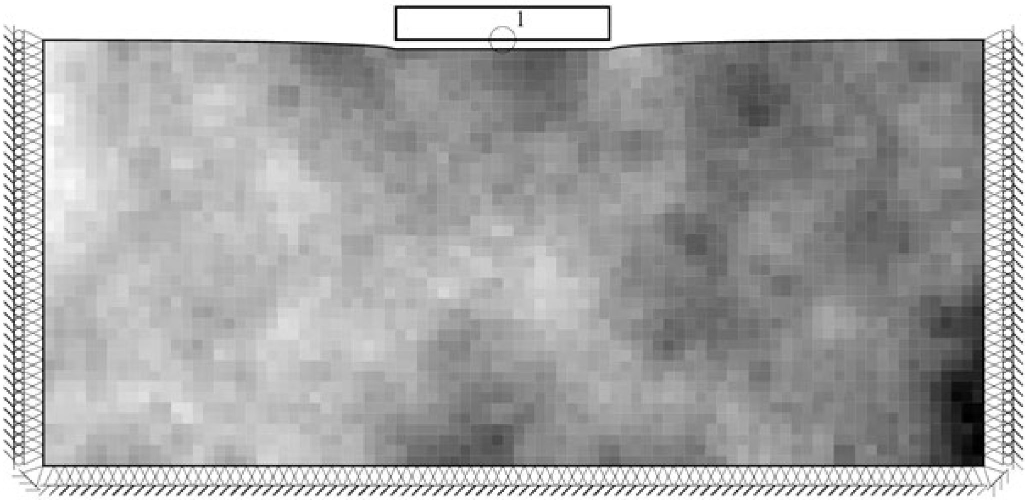


Figure 7.14: Graphical representation of the random field of *E* of Example #4 (θ/B=0.5; see Table 7.1).

The *predicted* settlement (ρ) is compared against the respective "*actual*" one. For each one of the examples presented herein, the latter derives from the respective random field of E using the RFEM method. The *predicted* ρ value derive from a homogenous soil field characterized by the mean of the values sampled from the original (random) field. The results are presented in Figure 7.15 in $\rho_{predicted}/\rho_{"actual"}$ versus x/B form for various d_d/B values; the relative difference R_d defined as $\rho_{predicted}/\rho_{"actual"}$ -1 is also given in each chart (see secondary vertical axis; a positive value indicate design on the safe side and vice versa).

If the suggestions related to the horizontal distance from the footing centre and the domain length (x/B=0 and d_d/H=2 respectively) are valid, the $\rho_{predicted}/\rho_{"actual"}$ ratio for this specific sampling scenario should, logically, be equal to unity or very close to this value. As shown in Figure 7.15 the $\rho_{predicted}/\rho_{"actual"}$ ratio value for x/B=0 are very close to unity or equal to unity, indicating the validity of author' suggestions. Indicatively it is mentioned that, some locations (x/B) present a high $\rho_{predicted}/\rho_{"actual"}$ ratio values and these are attributed to the "light" (weak) areas appearing at these particular locations (see Figures 7.12 and 7.14). From Figure 7.15 it is confirmed that a vertical sampling domain of length equal to $2B$ leads to significantly lower statistical uncertainty. Also, sampling away from the center of footing may lead to significant statistic error, especially if the optimal sampling domain length is not used.

Finally, comparison between the figures given in Section 7.1.3 and 7.1.4 with the respective ones given in Section 7.1.5, shows clearly that statistical uncertainty does not necessarily decreases with increasing number of samples. Indeed, the opposite may easily happen. For example, comparing the $p_f \approx 0.2$ value for x/B=0 shown in Figure 7.4c (single point case; isotropic soil) with the $p_f \approx 0.32$ value for x/B=2 (case of 40 sampling points; also isotropic soil) shown in Figure 7.8c, it is obvious that statistical uncertainty can only be minimized by targeted field investigation. Such examples can also be found in the present section; please compare the case of {x/B =1, d_d/B=2} with the {x/B =0, d_d/B =0.5} in Figure 7.15b giving $R_d \approx$ +0.27 and -0.045 respectively.

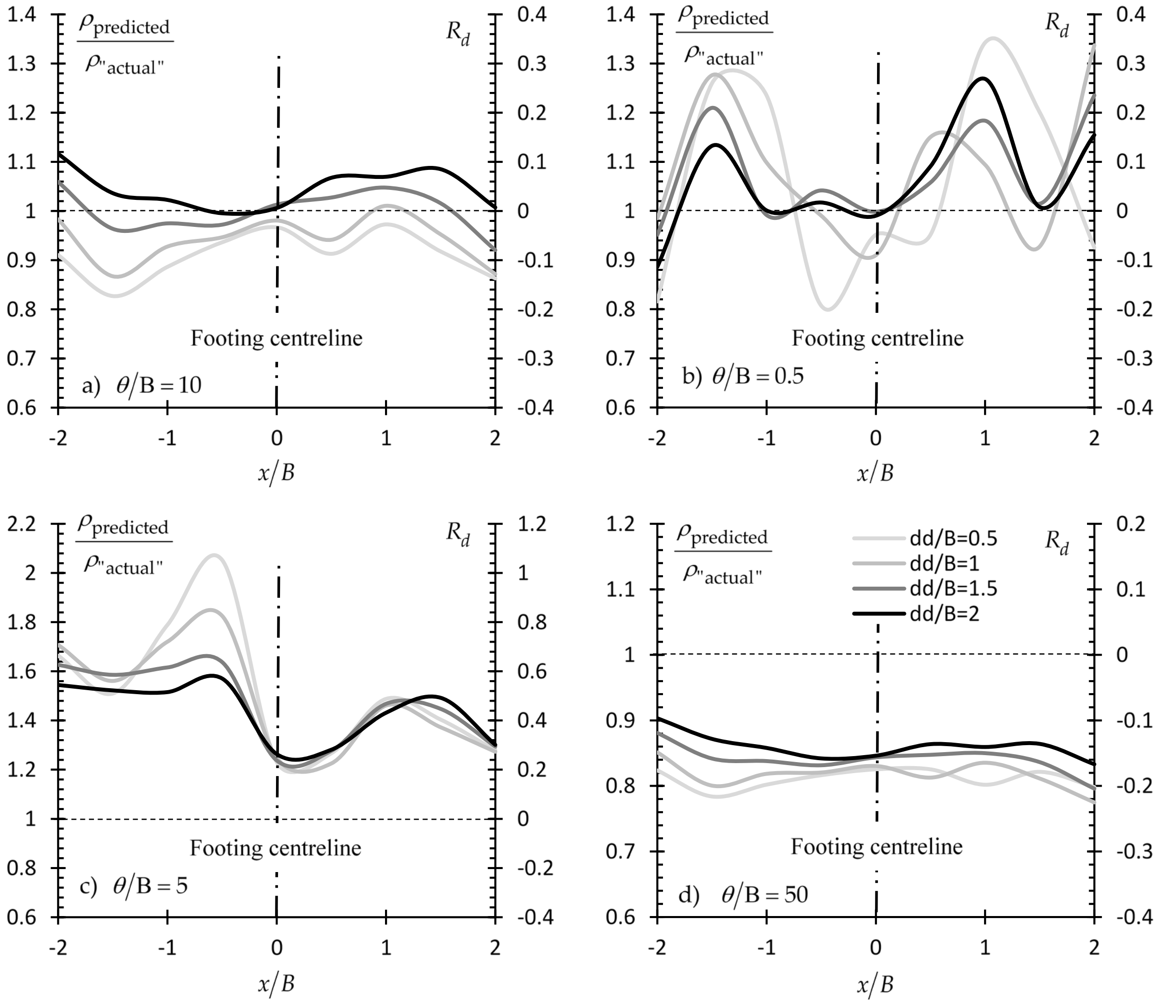


Figure 7.15: $\rho_{predicted}/\rho_{"actual"}$ vs *x/B* curves for various d_d/B values for Example a) #1, b) #2, c) #3 and d) #4, (see also Table 7.1 and Figure 7.1, 7.12, 7.13 and 7.14).

7.1.5.2 Designing with Load and Resistance Factor Design (LRFD) codes.

The discussion on the design of foundation against settlement based on characteristic soil property values instead of the respective mean values is facilitated by the two example charts of Figure 7.16 (see also discussion of Chapter 3 and the relevant equations (2.1) and (2.2)). These charts refer to the case #2 presented in the previous paragraph (see also Table 7.1). This specific case was chosen because of the relatively low θ value (i.e. θ/B=0.5), which indicates a rather highly spatially variable soil; thus, the use of the characteristic value makes more sense. Two cases are presented, the d_d/B=2 and the d_d/B=0.5.

From Figure 7.16 it is clear that, the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic values. Moreover, despite the conservatism which is inserted in the analysis using the characteristic value concept, the characteristic values alone, as shown, cannot guaranty a conservative enough engineering study. The inclusion of the "characteristic value" in the RSETL2D code has also been done by the author.

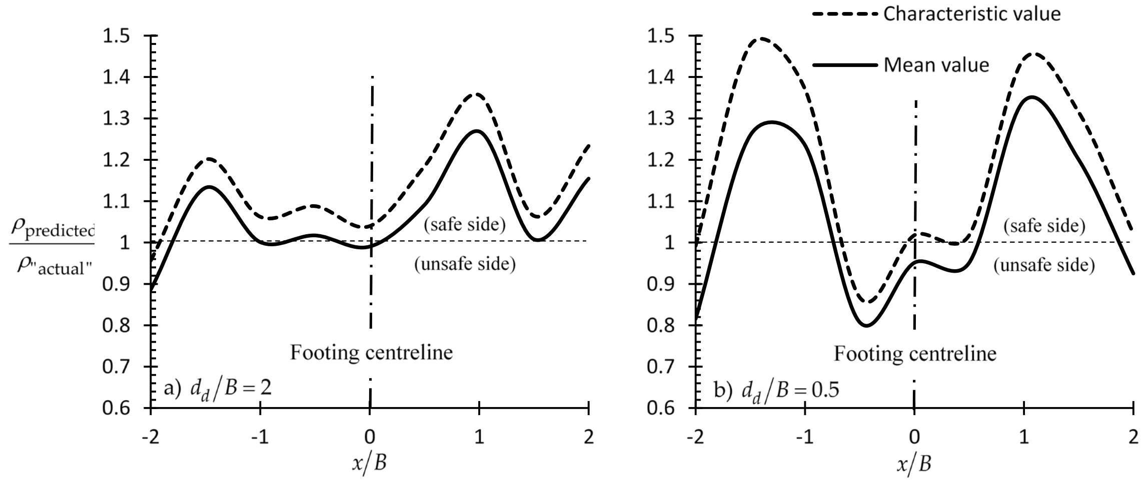


Figure 7.16: $\rho_{predicted}/\rho_{"actual"}$ vs x/B curves using both mean and characteristic values (solid and dashed lines respectively).

7.2 Bearing capacity analysis

The present Chapter deals with the practical problem of reducing statistical uncertainty in bearing capacity analysis of shallow foundations relying on targeted field investigation, aiming at an optimal design. This is done in an advance probabilistic framework using the RFEM [61] method properly considering sampling of soil properties. In this respect, the open source RBEAR2D program, which combines elasto-plastic finite element analysis with the random field theory, has been modified as to include the function of sampling of soil property values from the generated random fields and return the failure probability of footing against bearing capacity failure. The present analysis will show that the statistical error in an bearing capacity analysis can be effectively minimized only when targeted field investigation is carried out.

7.2.1 Brief description of the RFEM program used

As mentioned, the present analysis was based on the open source RBEAR2D program. The RBEAR2D program calculates the ultimate bearing capacity of a smooth strip footing (or a pair of strip footings) founded on a soil having spatially random properties. The procedure is repeated m times; m is the number of realizations, where, each RFEM realization refers to a new random field of soil properties.

The footing(s) is (are) assumed to be founded on a soil layer underlain by bedrock. The physical problem is represented using a two-dimensional (plane-strain) model as shown in Figure 7.17. The finite-element bearing capacity analysis in RBEAR2D uses an elastic-perfectly plastic stress-strain law with a Mohr-Coulomb failure criterion. Plastic stress redistribution is accomplished using a viscoplastic algorithm [174]. The program uses 8-node quadrilateral elements and reduced integration in both the stiffness and stress redistribution parts of the algorithm. In particular, the footing is incrementally displaced vertically into the soil and the sum of the nodal reactions back-figured from the converged stress stated. When the sum of the nodal reactions levels out to within quite strict tolerances, “failure” is said to have occurred and the sum of the nodal reactions is considered to be the “bearing capacity” of the particular realization [184].

For the needs of the present research, the original RBEAR2D program has been extended by the author as to:

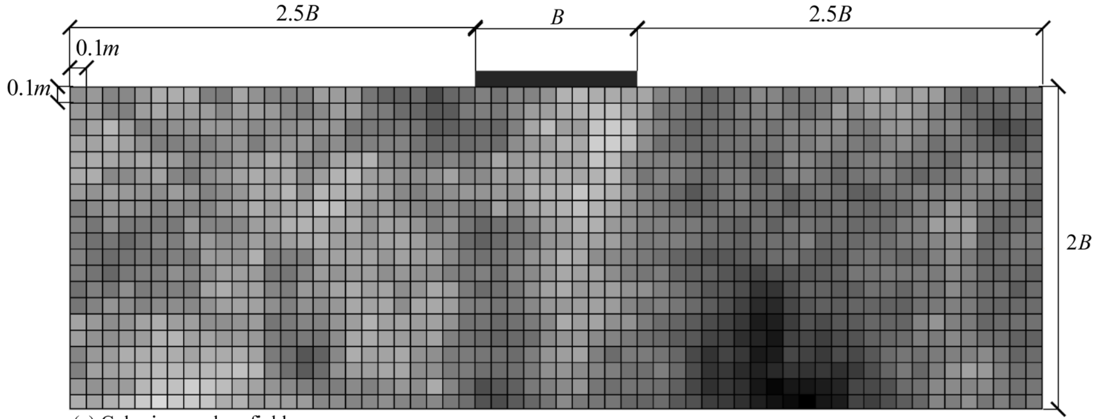
- virtually sample soil property values from the random fields generated in each RFEM realization,
- calculate the footing bearing capacity (again in each RFEM realization) considerring that the soil is homogenous, having soil property values equal to the mean of the values sampled (this bearing capacity is calculated in addition to the bearing capacity of footing lying on spatially random soil) and
- estimate the failure probability of the footing.

The latter is defined by the fraction of the realizations resulted in failure over the total number of realizations. In each RFEM realization, “failure” is considered to have occurred when the “actual” bearing capacity referring to the spatially random soil, is smaller than the respective (factored or unfactored) predicted value, referring to the spatially uniform soil. That is, it stands that

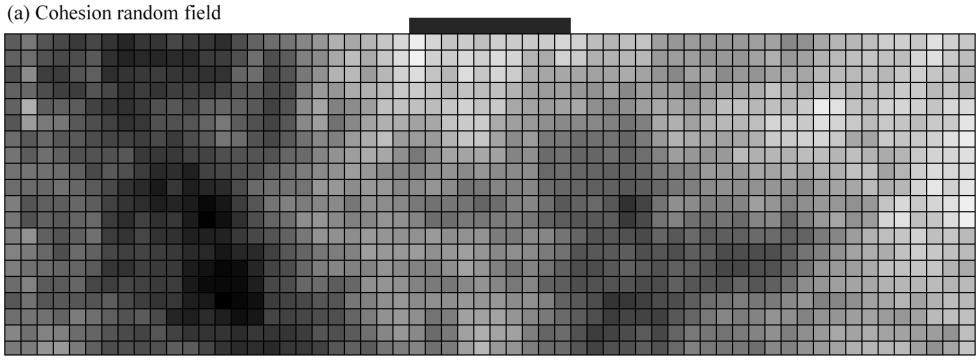
$$p_f = P\left[q_{u\,\text{"actual"}} < q_{u\,predicted} / FS\right] \tag{7.2}$$

where, q_u is the ultimate bearing capacity and FS is the user-defined safety factor.

The modified program was validated as follows. First a given footing was solved with the original RBEAR2D program using deterministic property values. Then the same footing was solved with the modified program using values sampled from various places (because the same deterministic soil properties value was spread out in the finite element mess, all sampled values had the same value). The two programs gave exactly the same results, indicating that the function of sampling was embedded correctly into the original program.



(a) Cohesion random field



(b) Friction angle random field

Figure 7.17: Typical random finite element meshes of (a) cohesion, c, and (b) friction angle, ϕ, considering $\theta/B = 5$. Figure showing the 10-element footing on the surface of the 60 x 20 (H x V) mesh. The soil's spatially random shear strength parameters (c and ϕ) are shown qualitatively using a grayscale representation, where lighter regions indicate weaker soil.

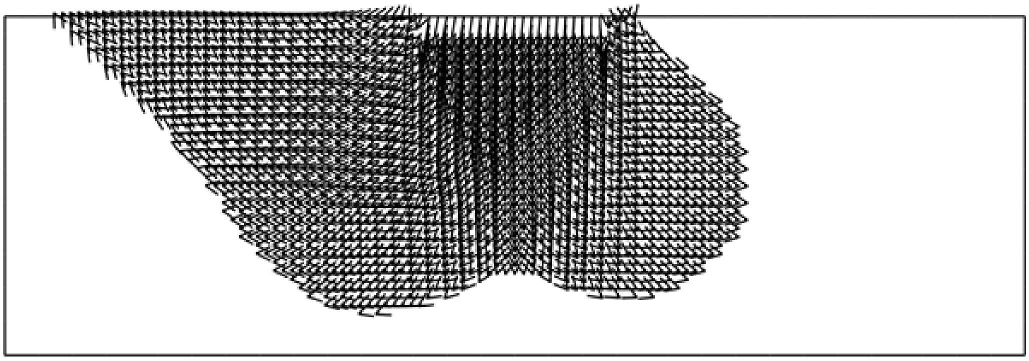


Figure 7.18: Displacement vector plot of bearing failure on spatially variable soil.

7.2.2 Parametric study for determining the optimal sampling strategy

This Chapter deals with the case of a single strip footing. Both the sampling from a single point and the sampling from an entire domain strategies are investigated through an extensive parametric analysis of the factors controlling the bearing capacity for defining the strategy that minimizes the probability of failure and thus, the statistical error (this strategy is called hereafter "optimal sampling strategy"). The error is quantified comparing the probability of failure value obtained based on different sampling scenarios.

In the finite analysis that follows, the soil mass is discretized into a 60 (horizontal direction) by 20 (vertical direction) mesh consisting of eight-noded square elements with side length equal to 0.1m. The strip footing occupies width on the surface of the finite element mesh equal to 10 elements, giving it a width of B=1m (caller hereafter "reference footing"; other footing widths will also be considered in a later sections). A typical random finite element mesh of cohesion and friction angle is presented in Figure 7.17. The 60-element mesh in the horizontal direction was chosen so that the failure mechanism in the RFEM analysis not to be affected by the proximity of the side boundaries. The effect of the distance between the edge of the footing and the respective lateral boundary was investigated prior to the analysis. As shown in Appendix IV, the error inserted considering a 10-element footing centered on the surface of a 60-element mesh is negligible; that makes a free distance between each edge of footing and the respective lateral boundary equal to 2.5B. The same distance of 2.5B was kept the same for the other footing widths considered (i.e. B=1, 2 and 3m); however, the element size for the B= 2 and 3 m footings was 0.2 m and 0.3 m respectively.

In the present analysis, c and ϕ are treated as random fields assuming that they follow a log-normal distribution; the soil is assumed weightless with $\mu_c = 100$ kPa and μ_ϕ =25°, whilst various standard deviation and θ values are examined. In addition, it is mentioned that, based on a preliminary parametric analysis carried out by the author, the modulus of elasticity E have no influence on the optimal sampling strategy. Thus, E have been kept constant and equal to 10^5 kN/m^2 throughout the entire analysis presented herein. Furthermore, it is assumed that the random fields have the same spatial correlation length and the same type of correlation function; in this respect, Markovian spatial correlation function has been adopted (see Equation (5.2)).

Aiming at finding the optimal sampling strategy, the following parameters will be examined: the sampling depth (d_p) and horizontal distance for the case of sampling from a single point (measured from the soil surface and the axis of strip footing respectively), the sampling length (d_d) and horizontal distance of the geotechnical borehole for the case of sampling an entire domain (measured as in the previous case), the spatial correlation length of soil (θ), the footing width (B), the coefficient of variation (COV) of c and ϕ, the effect of Poisson's ratio (v) value of soil and the soil mass anisotropy $(\theta_h \neq \theta_v)$. Four sampling scenarios similar to those presented in Figure 7.3 were consider. The number of realizations was set equal to 3000; this number, as shown in Appendix IV, can be considered adequate for the needs of the present research. Finally, a safety factor *FS* equal to 1.2 is generally assumed in the analysis.

7.2.3 Sampling from a single point

7.2.3.1 Effect of spatial correlation length $(\boldsymbol{\theta})$

Example charts showing the variation of p_f with respect to d_p/B for various θ/B values are given in Figure . The soil mass is considered isotropic. From the figure in question it is inferred that the optimal sampling location lies at the center of the strip footing $(x/B = 0)$ and at depth approximately 0.8B. Isolating the curves for $x/B = 0$ it seems that, there is a "worst case spatial correlation length", where, the failure probability becomes maximum; for example, for the various θ/B ratios shown in Figure 7.20 (ranging from θ/B =0.2 to 50), the θ/B =1 case gives the higher p_f values. This outcome is in agreement with the conclusions made by Fenton and Griffiths [31,61]. Moreover, from Figure 7.20 it is inferred that, as θ tends to zero the p_f value tends to a single value for any depth (that is, p_f is independent of the sampling depth). However, as θ increases, the p_f value becomes more dependent on the sampling depth.

Furthermore, the effect of soil anisotropy on the optimal sampling location is investigated herein by comparing the $\theta_h = \theta_v$ case with the $\theta_h = 10\theta_v$ case, i.e. the reference footing-soil system with $\theta_v/B = \theta_h/B$ =5 will be compared with a respective one having $\theta_h/B = 50$and $\theta_v/B = 5$. In this respect, the soil anisotropy has only a minor influence on the location of the optimal sampling point (see Figure 7.19d, heavy bold line).

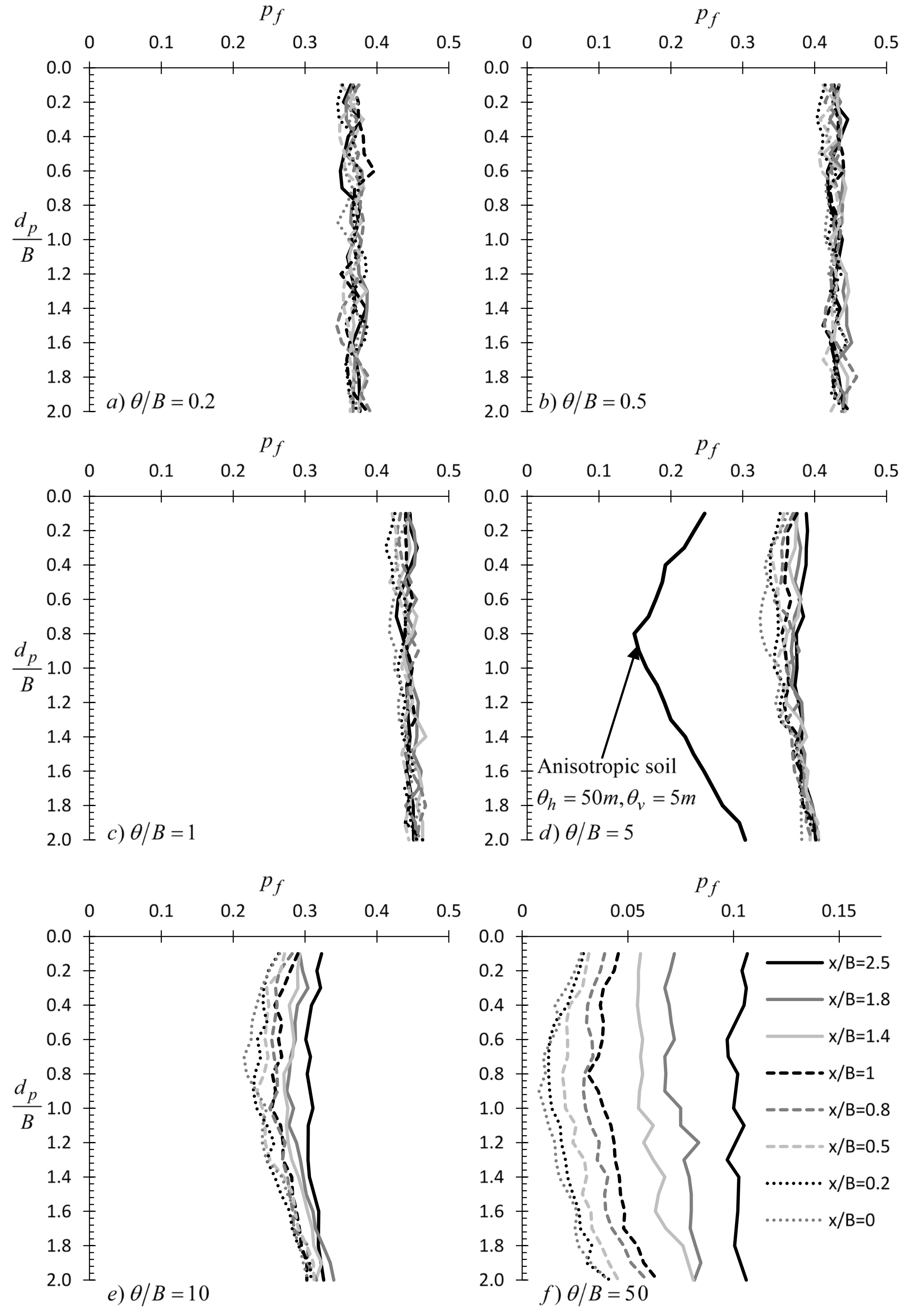


Figure 7.19: p_f versus d_p/B example curves for various θ/B and x/B values. The thick continuous line in Figure d is for the anisotropic case and refers to the x/B =0 case.

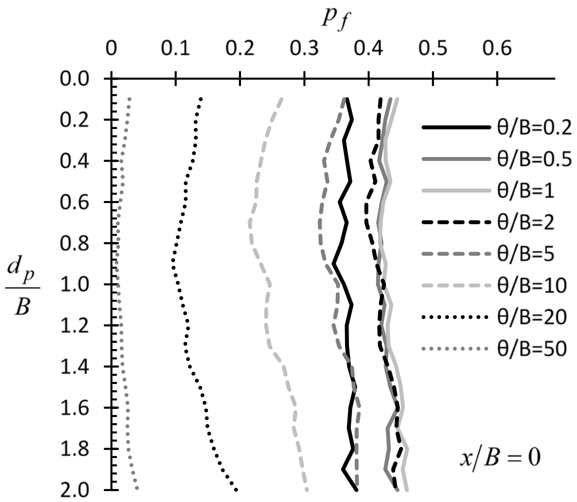


Figure 7.20: p_f versus d_p/B example curves for various θ/B values and $x/B=0$.

7.2.3.2 Effect of footing width *(B)*

The variation of p_f with d_p/B is shown in Figure 7.21 for three footing widths, i.e. B =1, 2 and 3m. From this figure, it is inferred that the footing width has only a minor influence on the location of the optimal sampling point. The three curves in Figure 7.21 refer to $x/B=0$.

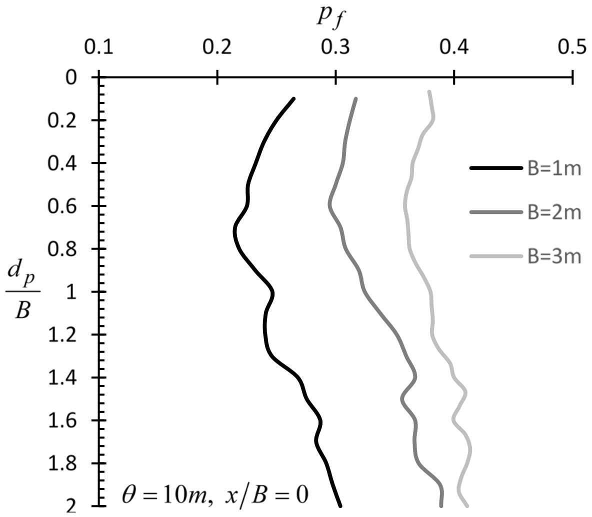


Figure 7.21: p_f versus d_p/B example curves for different footing widths B and $\theta/B=10$.

7.2.3.3 Effect of COV of c and ϕ

In this paragraph, five COV values for c and ϕ were considered, namely, COV = 0.1, 0.2, 0.3, 0.4 and 0.5. From Figure 7.22, it is obvious that the failure probability increases as COV of c and ϕ increases, but what is not trivial is that the positive effect of targeted field investigation on the reduction of the statistical error is greater for smaller COV values. For example, considering the different values of COV of c, as shown in Figure 7.22a, the p_f value when COV(c) = 0.5 is approximately 0.48 and independent of the d_p/B ratio. In contrast, if COV(c) = 0.1, p_f ranges from 0.074 for $d_p/B = 0$ to 0.04 for $d_p/B = 0.8$. The curves in Figure 7.22 refer to $x/B = 0$.

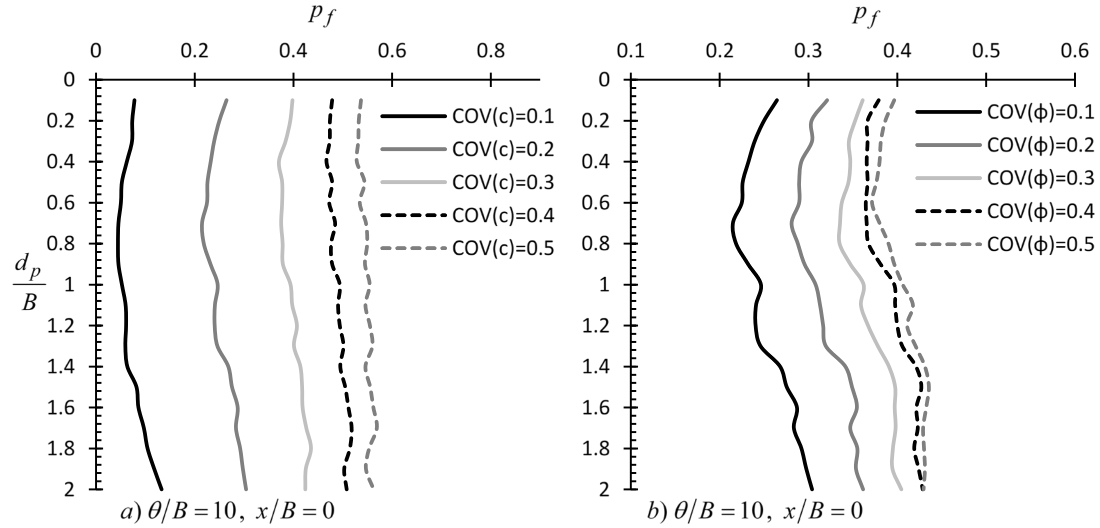


Figure 7.22: p_f versus d_p/B example relationships by considering different values of COV of (a) cohesion (c) and (b) friction angle (ϕ).

7.2.3.4 Effect of the elastic constant values of soil

For all cases considered above, the Poisson's ratio of soil was equal to 0.3. Parametric study on the effect of the elastic modulus *E* on the optimal sampling location, however, showed that the latter is not affected by the parameter in question. The influence of v on the optimal sampling location is examined herein. In this respect, six v values were considered, i.e. v= 0, 0.1, 0.2, 0.3, 0.4 and 0.495. The six curves shown in Figure 7.23 refer to $x/B = 0$. From Figure 7.23 it is inferred that the v has no effect on the optimal sampling depth. However, it is interesting that there is a unfavourable v value where the failure probability becomes maximum; for example, for the various v values shown in Figure 7.23, the v=0.2 case gives the maximum p_f value.

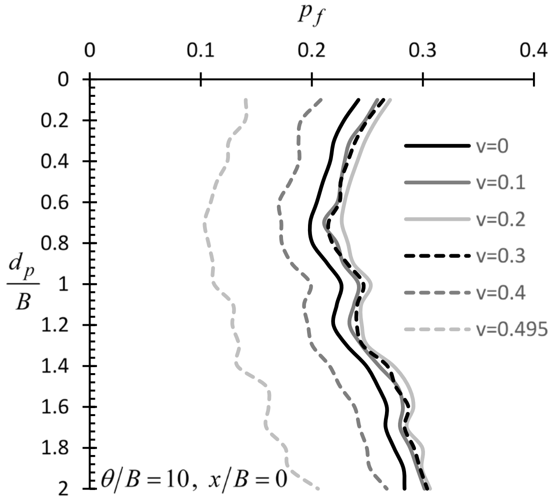


Figure 7.23: p_f versus d_p/B example curves for different v values and $\theta/B = 10$.

7.2.4 Sampling from an entire domain

In the parametric analysis carried out, three footing widths were considered, i.e. B=1, 2 and 3m, whilst the distance between two successive sampling points (in the vertical direction) was B/10 for all cases. The maximum sampling domain length considered, was always two times the footing width B. It is noted that, for all cases examined, the optimal sampling distance was found again to be at x/B =0. Thus, the analysis below, generally, refers to x/B =0.

7.2.4.1 Effect of spatial correlation length ($\boldsymbol{\theta}$)

Example charts showing the variation of p_f with respect to d_d/B for various θ/B and x/B values are given in Figure 7.24; it is reminded that FS was set equal to 1.2 (recall Equation (7.2)). From this figure it is inferred that, the optimal horizontal sampling distance from the footing center is again for x/B =0, although for very small theta values the horizontal sampling distance makes no noticeably difference. However, as theta increases the role of horizontal distance becomes more significant. Given now that, soil samples will be taken from x/B =0, it is advisable, a domain length at least $2B$ to be considered. Indeed, this practice may significantly reduce the statistical error. Finally, from Figure it is inferred that, a “worst case theta” exists. This is more obvious in Figure 7.25 showing the variation of p_f with d_d/B for various θ/B values and x/B= 0; see also the relevant discussion in single point sampling strategy (Section 7.2.3.1).

The variation of p_f with d_d/B for the case of anisotropic soil mass ($\theta_h/B = 50$ and $\theta_v/B = 5$) is shown in Figure 7.24d (heavy bold line). Generally, it can be said that, the anisotropic soil mass affects the optimal sampling domain length in a manner that it calls for greater domain length compared to the isotropic soil mass.

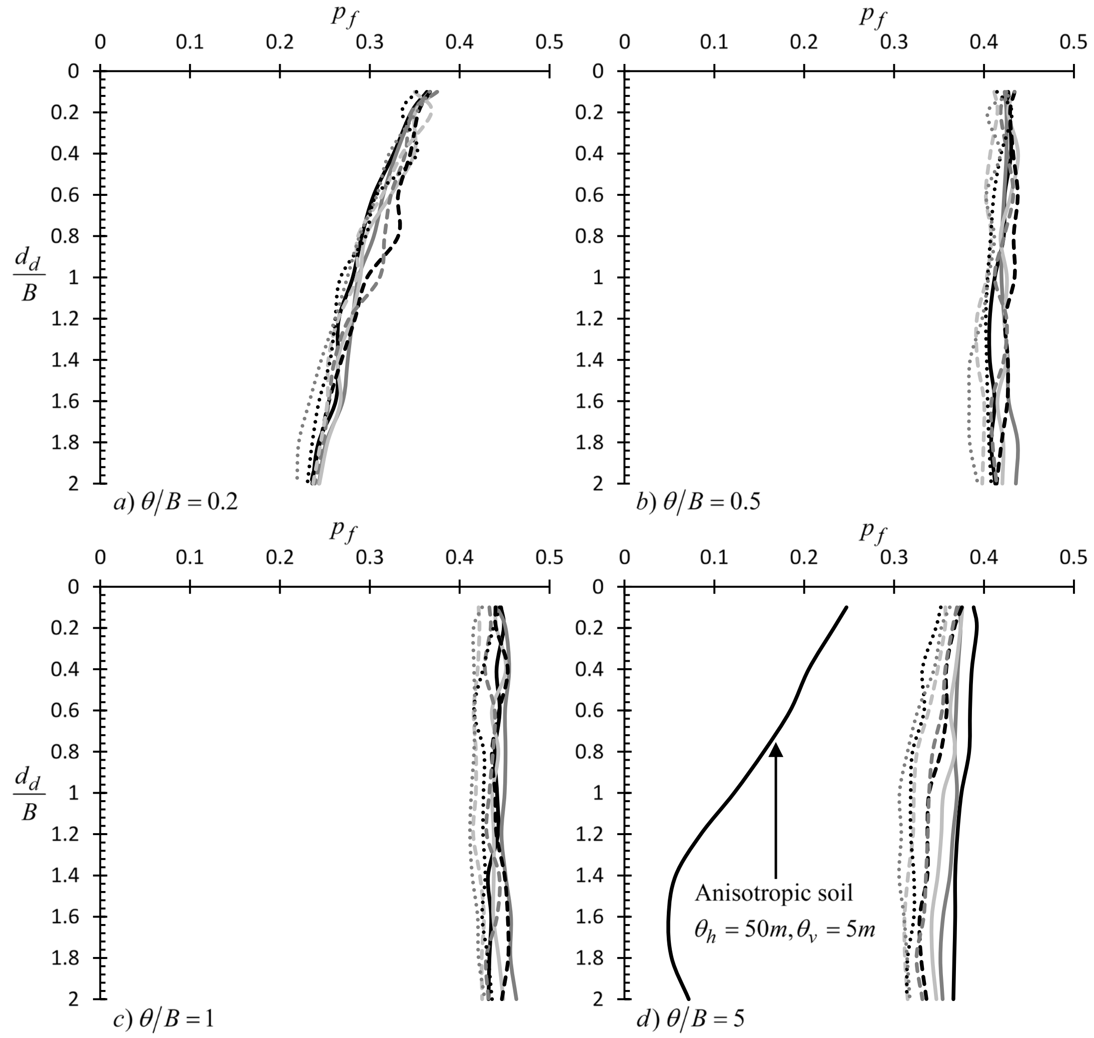


Figure 7.24: p_f versus d_d/B example curves for various θ/B and x/B values. The thick continuous line in Figure d is for the anisotropic case and refers to the x/B =0 case.

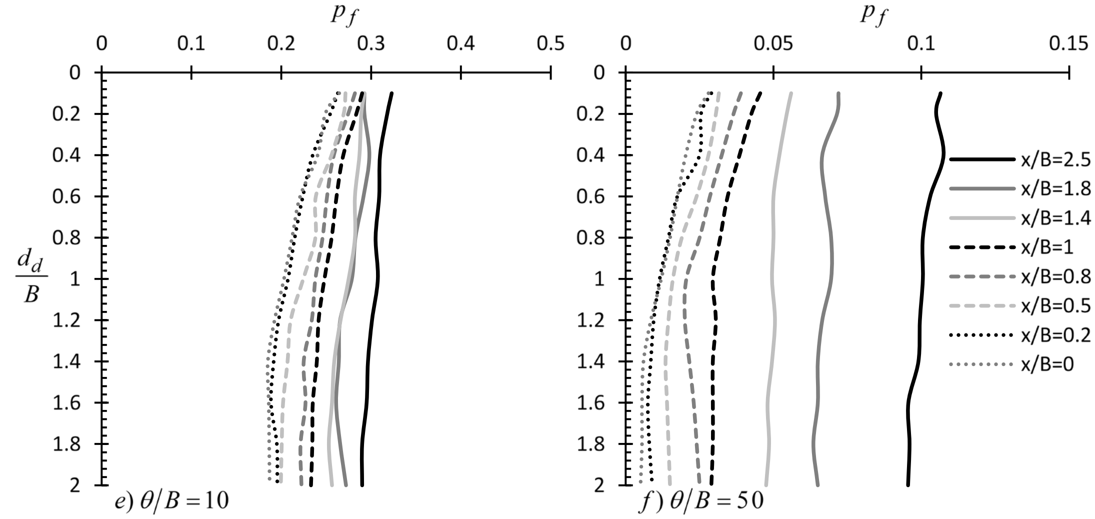


Figure 7.24: Continued.

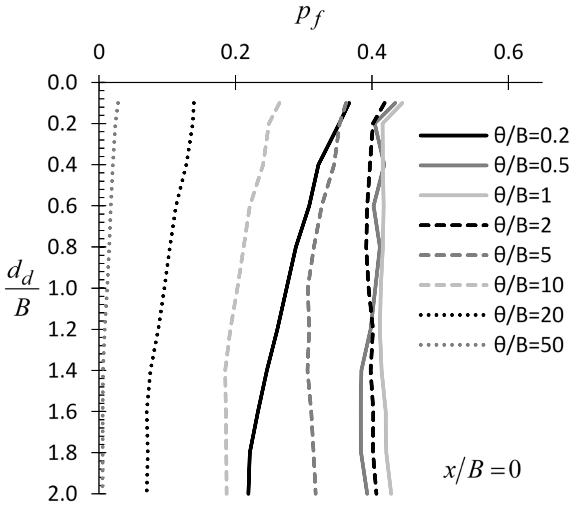


Figure 7.25: p_f versus d_d/B example curves for various θ/B and x/B=0.

7.2.4.2 Effect of footing width (B)

In this paragraph three footing widths were considered, i.e. B =1, 2, 3m. Figure 7.26 presents the variation of p_f with d_d/B for these three cases. From this figure, it is inferred that the footing width has only a minor influence on the location of the optimal sampling point.

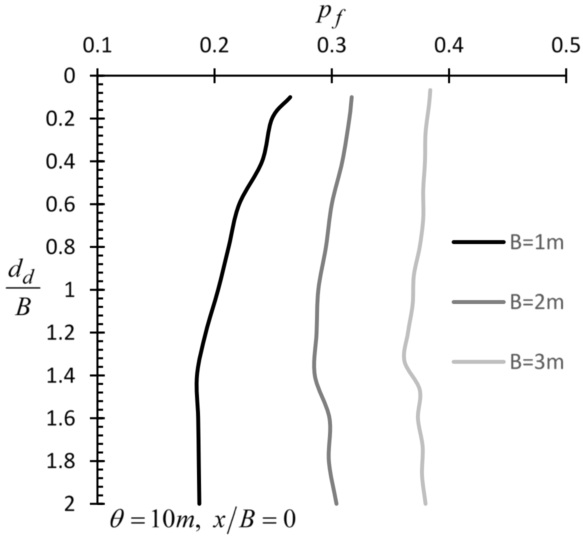
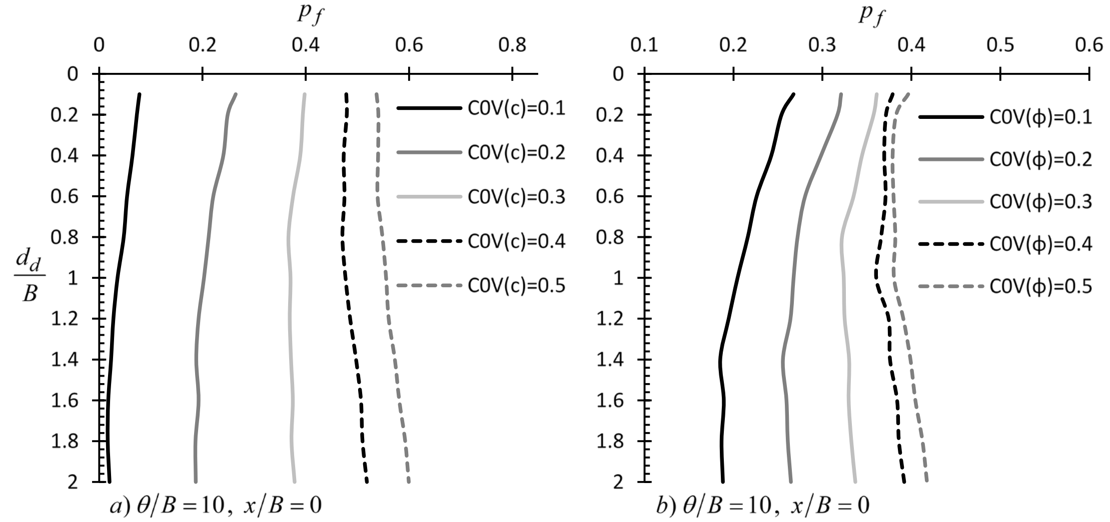


Figure 7.26: p_f versus d_d/B example curves for different footing widths B and $\theta/B = 10$.

7.2.4.3 Effect of COV of c and ϕ

In this paragraph, five COVvalues of c and ϕ were considered, i.e. $COV = 0.1$, 0.2, 0.3, 0.4 and 0.5. The optimal horizontal sampling distance from the footing center was found not to be affected by the COV of c and ϕ, where again the x/B =0 case leads to the smaller probabilities of failure. Thus, only the x/B =0 case will be presented here. From Figure 7.27 it is, generally, inferred that the required sampling domain length decreases as the $COV$of c or ϕ increases. However, as discussed earlier, a domain length of at least $2B$ may significantly reduce the statistical error.



a) $\theta/B = 10,\ x/B = 0$

b) $\theta/B = 10,\ x/B = 0$

Figure 7.27: p_f versus d_d/B example relationships by considering different values of COV of c and ϕ.

7.2.4.4 Effect of the elastic constant values of soil

The variation of p_f with respect to d_d/B for different Poisson's ratio values is shown in Figure 7.28; the optimal sampling distance was also found to be at x/B =0 for any v value, thus, only this case is presented here. From Figure 7.28 it is obvious that the Poisson's ratio value has no effect on the optimal sampling domain length. The same stands for the mean value of elastic modulus of soil.

Although, from the curves of the figure in question it seems that, there is a "unfavorable" Poisson's ratio value, where, the failure probability becomes maximum; see also the relevant discussion in single point sampling strategy (Section 7.2.3.4).

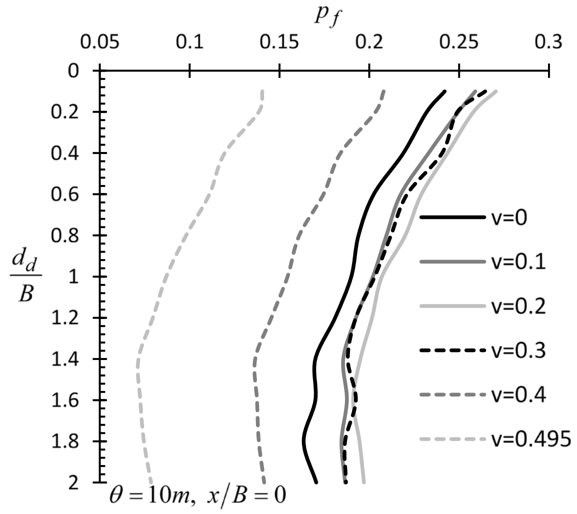


Figure 7.28: p_f versus d_d/B example curves for different v values and $\theta/B = 10$.

7.2.5 Discussion

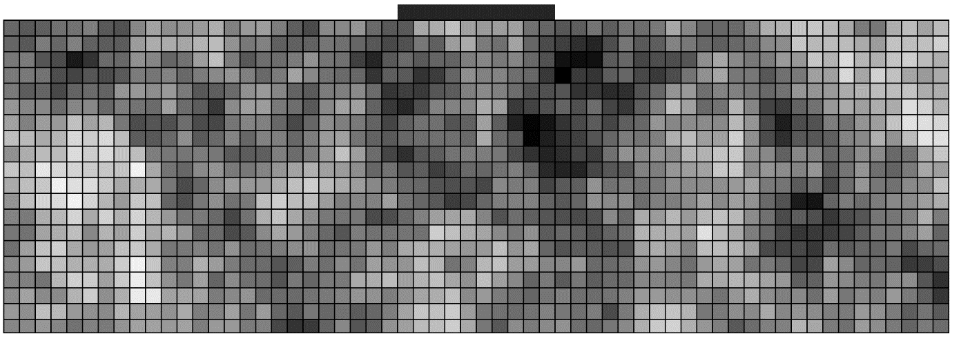
7.2.5.1 The importance of targeted field investigation in practice

In the four examples presented below, the footing and the mesh / boundary conditions are the same with those presented in Section 7.2.2 (i.e. the reference footing). The material properties are given in Table 7.2. The four examples, in essence, differ only in the spatial correlation length. The random field of c and ϕ used in each example is shown in Figures 7.17, 7.29, 7.30 and 7.31. It is reminded that, the light areas correspond to lower shear strength values and vice versa. The *FS* value is assumed unity (recall Equation (7.2)); this factor is discussed in the next paragraph.

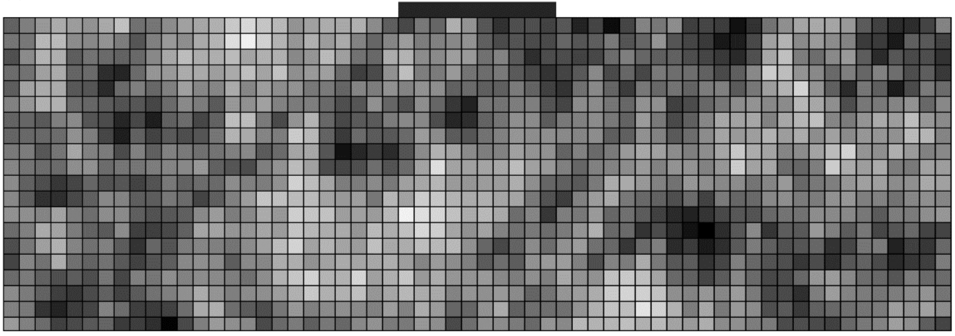
Table 7.2: Summary of the characteristics of the soils used in the four examples (footing width B= 1 m).

Example	Random field(s)	Distribution	μ_ϕ	μ_c	COV	θ/B	Figure (1)
#1	ϕ, c	Log-normal	25°	100 kPa	0.3	5	Figure 7.17
#2	ϕ, c	Log-normal	25°	100 kPa	0.3	0.5	Figure 7.29
#3	ϕ, c	Log-normal	25°	100 kPa	0.3	10	Figure 7.30
#4	ϕ, c	Log-normal	25°	100 kPa	0.3	50	Figure 7.31

(1) Figures shown the random fields of c and ϕ.

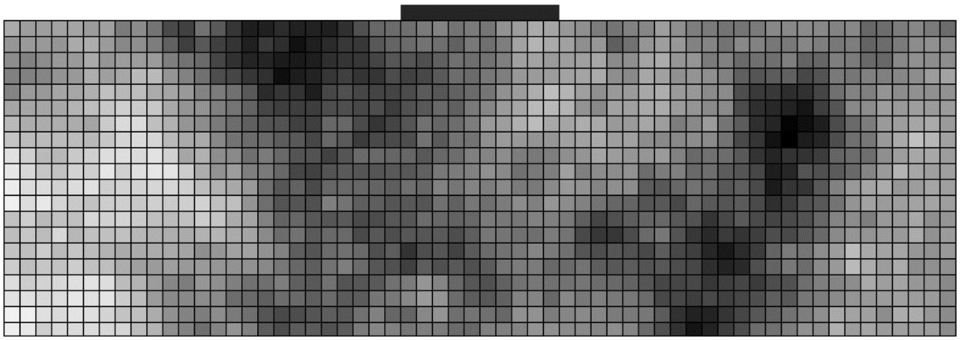


(a) Cohesion random field

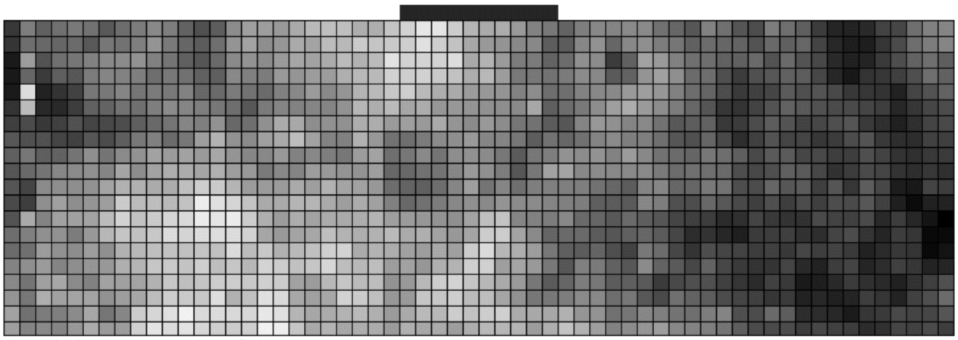


(b) Friction angle random field

Figure 7.29: Graphical representation of the random fields of (a) cohesion, c, and (b) friction angle, ϕ, of Example #2 (θ/B=0.5; see Table 7.2).

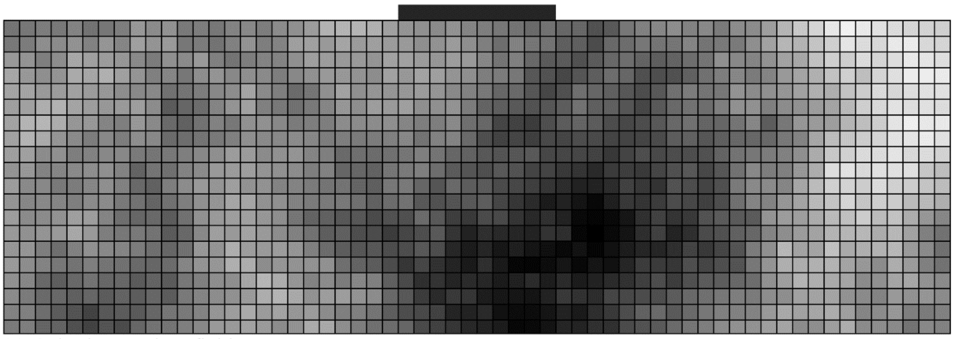


(a) Cohesion random field

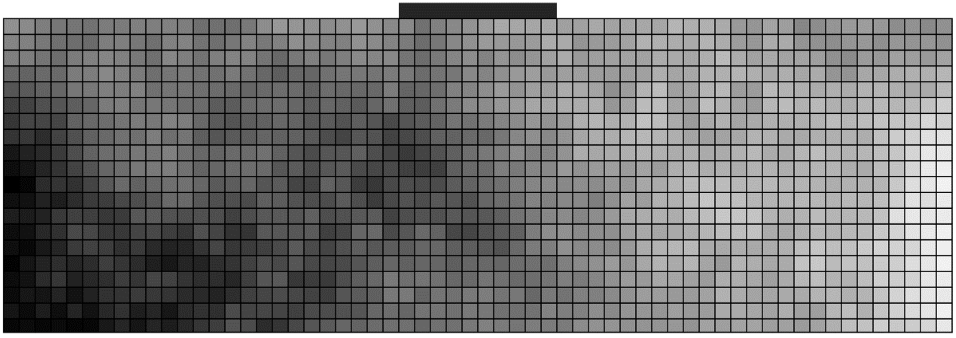


(b) Friction angle random field

Figure 7.30: Graphical representation of the random fields of (a) cohesion, c, and (b) friction angle, ϕ , of Example #3 (θ/B=10; see Table 7.2).



(a) Cohesion random field



(b) Friction angle random field

Figure 7.31: Graphical representation of the random fields of (a)) cohesion, c, and (b) friction angle, ϕ , of Example #4 (θ/B=50; see Table 7.2).

The *predicted* bearing capacity (q_u) is compared against the respective "*actual*" one. For each one of the examples presented herein, the latter derives from the respective random fields of c and ϕ using the RFEM method. The *predicted* q_u value derive from a homogenous soil field characterized by the mean of the values sampled from the original (random) fields. The results are presented in Figure 7.32 in $q_{u\ predicted}/q_{u\ \text{"}actual\text{"}}$versus *x*/*B* form for various d_d/Bvalues; the relative difference *R_d* defined as$q_{u\ predicted}/q_{u\ \text{"}actual\text{"}}$" -1 is also given in each chart (see secondary vertical axis; a negative value indicate design on the safe side and vice versa).

If the suggestions related to the horizontal distance from the footing centre and the domain length (*x*/*B*=0 and d_d/B=2 respectively) are valid, the $q_{u\ predicted}/q_{u\ \text{"}actual\text{"}}$ ratio for this specific sampling scenario should, logically, be equal to unity or very close to this value. As shown in Figure 7.32 the $q_{u\ predicted}/q_{u\ \text{"}actual\text{"}}$ ratio value for *x*/*B*=0 are very close to unity or equal to unity, indicating the validity of authors' suggestions. Indicatively it is mentioned that, some locations (*x*/*B)* present a high $q_{u\ predicted}/q_{u\ \text{"}actual\text{"}}$ ratio values and these are attributed to the "dark" (strong) areas appearing at these particular locations (see e.g. Figure 7.30 and Figure 7.31). From Figure 7.32 it is confirmed that a vertical sampling domain of length equal to 2*B* leads to significantly lower statistical uncertainty. Also, sampling away from the center of footing may lead to significant statistic error, especially if the optimal sampling domain length is not used.

Finally, comparison between the figures given in Section 7.2.3 and 7.2.4 with the respective ones given in Section 7.2.5, shows clearly that statistical uncertainty does not necessarily decreases with increasing number of samples. Indeed, the opposite may easily happen. For example, comparing the p_f ≈0.21 value for *x/B*=0 shown in Figure 7.19e (single point case; isotropic soil) with the p_f ≈0.30 value for *x/B*=2.5 (case of 20 sampling points; also isotropic soil) shown in Figure 7.24e, it is obvious that statistical uncertainty can only be minimized by targeted field investigation. Such examples can also be found in the present section; please compare the case of {*x/B* =1, d_d/B=2} with the {*x/B* =0, d_d/B =0.5} in Figure 17b giving R_d≈ +0.49 and 0.11 respectively.

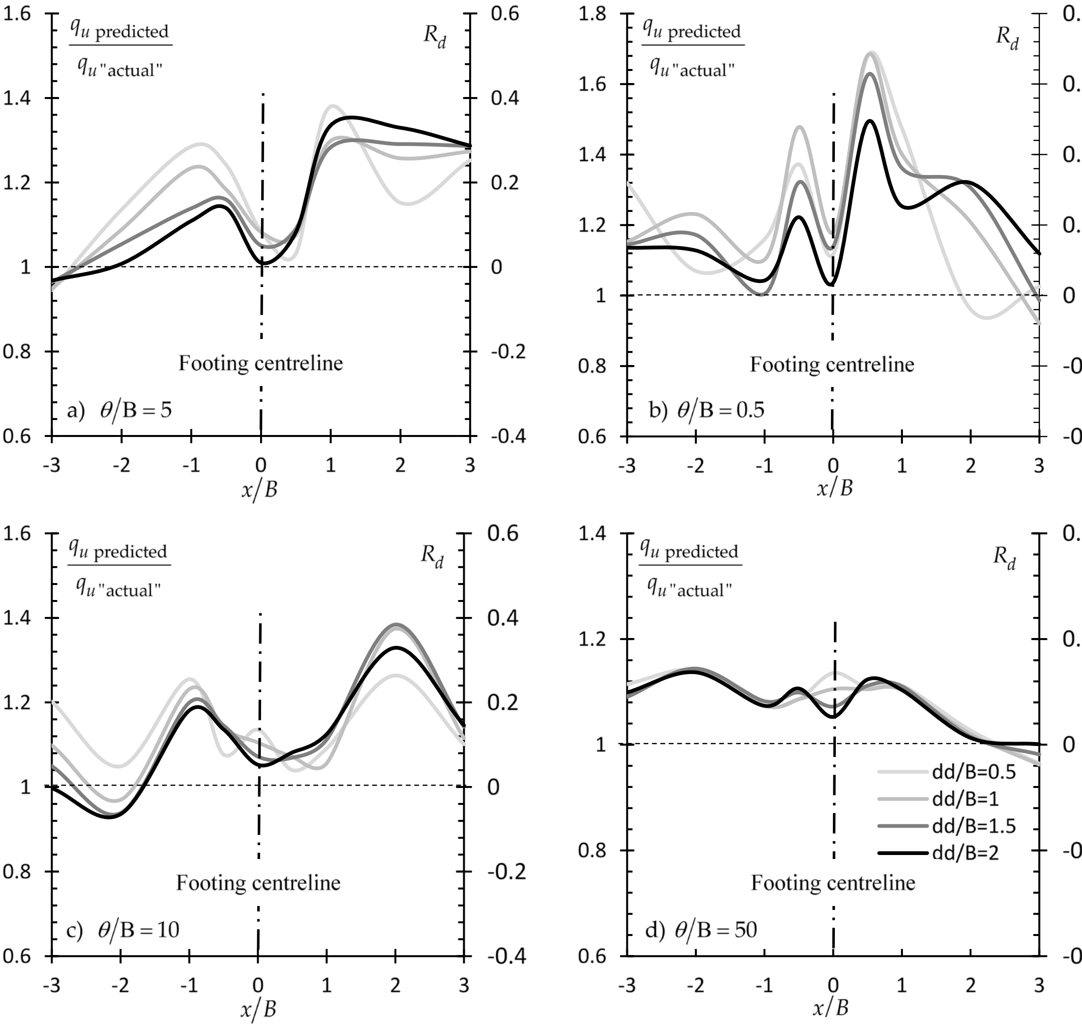


Figure 7.32: $q_{u\ predicted}/q_{u\ \text{"}actual\text{"}}$ vs *x*/*B* curves for various *d*d_d/B values for Example a) #1, b) #2, c) #3 and d) #4, (see also Table 7.2 and Figures 7.17, 7.29, 7.30 and 7.31).

7.2.5.2 Designing with Load and Resistance Factor Design (LRFD) Codes

The discussion on the bearing capacity design of shallow footings based on characteristic soil property values (please see equations (2.1) and (2.2) in the literature review of this Thesis) instead of the respective mean values is facilitated by the two example charts of Figure 7.33. These charts refer to the case #2 presented in the previous paragraph (see also Table 7.2). This specific case was chosen because of the relatively small θ value (i.e. $\theta/B=0.5$), which indicates a rather highly spatially variable soil; thus, the use of the characteristic value makes more sense. Two cases are presented, the $d_d/B=0.5$ and the $d_d/B=2$. It is also mentioned that, in the example presented here, the partial material factor γ_M for the shear strength parameters was set equal to unity.

From Figure 7.33 it is clear that, the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic values. Moreover, despite the conservatism which is inserted in the analysis using the characteristic value concept, the characteristic values alone, as shown, cannot guaranty a conservative enough engineering study. The safety level can be increased by applying a statistical uncertainty partial factor (similar to the model factor γR used by Eurocode 7) or a unified and more conservative model factor to the resistances, which will absorb the statistical uncertainties related to the soil. In this respect, a partial factor equal to 1.2 has also been applied (FS=1.2; recall Equation (7.2)) in the present example. As shown in Figure 7.33, the use of such a factor simply displaces downwards (that is, to the safe side) the $q_{u\ predicted}/q_{u\ \text{"}actual\text{"}}$versus x/B curves. The inclusion of the "characteristic value" in the RBEAR2D code has also been done by the author.

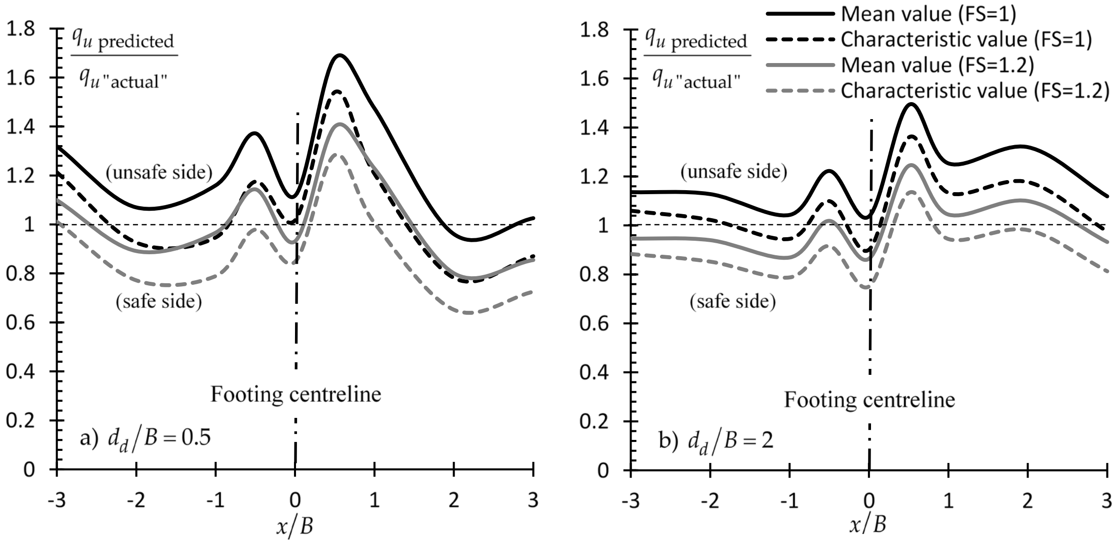


Figure 7.33: $q_{u\ predicted}/q_{u\ \text{"}actual\text{"}}$vs. x/B curves using both mean and characteristic values (solid and dashed lines respectively) for *FS*=1 and 1.2. Figure referring to two sampling domain cases (d_d/H = 1 (figure a) and to d_d/H = 0.25 (figure b)). The reference footing was used. Soil characteristics as shown in Table 7.2 (Example #2).

7.3 Summary and conclusions

The present Chapter deals with the practical problem of reducing statistical uncertainty in elastic settlement and bearing capacity analysis of shallow foundations relying on targeted field investigation aiming at an optimal design. The results of the present research clearly show that statistical uncertainty may significantly affect the reliability of shallow foundations and that it can only be minimized by adopting the proper sampling strategy; the latter is defined by the number and location of sampling points. As samples are taken from a material field (i.e. the ground), which simultaneously is a stress field (stresses caused by the footing), the location of the optimal sampling points is affected by the coexistence of these two fields. Two main sampling strategies were investigated, namely, sampling from a single point and sampling from a domain, through an extensive parametric analysis.

One of the main findings of the present analysis is that for the case of a single footing, the geometric centre on its plan-view is the optimal sampling location. Regarding the depth of the *optimal sampling point*, it was found that it lies at depth equal to approximately $0.8B$; the exact depth depends on the spatial correlation length of the soil. Furthermore, it was observed that, the sampling domain length strategy leads usually to significantly lower statistical uncertainty than the sampling from a single point strategy, given that an adequate sampling length will be considered. In this respect, for the *optimal sampling domain length*, it is advisable that a domain length of at least $2B$ to be taken into account in the analysis.

In addition, it was observed that the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic soil property values. Moreover, despite the conservatism which is inserted in the analysis using the characteristic value concept, the characteristic values alone, as shown, cannot guaranty a conservative enough engineering study. The safety level can be increased by applying a statistical uncertainty partial factor (similar to the model factor γ_R used by Eurocode 7) or a unified and more conservative model factor to the resistances, which will absorb the statistical uncertainties related to the soil.

8 Conclusions

The present dissertation deals with the practical problem of reducing statistical uncertainty in geotechnical structures relying on targeted field investigation aiming at an optimal design. This is done in an advance probabilistic framework using the RFEM method properly considering sampling of soil properties. The RFEM method combines finite element analysis with the random field theory. Generally, the random field theory describes the inherent variability of soil properties. The inherent variability is modelled as a random field, which is described by the mean (μ), standard deviation (σ), spatial correlation length (θ) of soil properties. θ is defined as the distance within which the soil property shows relatively strong correlation or persistence from point to point. As shown in this study, the value of θ, is a critical parameter in order to achieve a significant reduction of statistical uncertainty.

Hence, this dissertation firstly examined the effectiveness of eight of the most commonly used methods for the estimation of the value of θ. This was done by generating 5000 one-dimensional random fields and comparing the estimated θ values with the respective one used as input value on the generation of random fields. The method used for the estimation of sample autocorrelation function $\hat{\rho}(\tau)$ is examined. Also, the effect of sampling domain length and sampling interval on the estimation of θ were systematically investigated. The main conclusions that can be drawn are summarized as follows: (a) the $\hat{\rho}(\tau)$ should be calculated up to the point where $\hat{\rho}(\tau)$ firstly intersect the τ axis, (b) the domain length D strongly affects the performance of the methods. In general, larger domains improve the estimate of correlation length. In terms of the sample mean, the ACF and ACF$_A$ methods converge to the input theta value for smaller domains than the other methods, and (c) the sampling interval d_x significantly affects the performance of the methods. Smaller intervals improve the estimate of correlation length. For θ_0/d_x greater than a specific value, say 7.5, the sample mean of the ACF, ACF$_A$ and VF methods is nearly constant to the best estimate allowed by the given domain length.

Thereafter, the inherent spatial variability of the clay sites of Pentakomo and Armou (Cyprus) were examined through a number of Unconfined Compression Strength (UCS) tests on undisturbed samples and a series of dynamic penetrations with a Dynamic Probing Light (DPL) apparatus. The latter test was also used for the investigation of spatial variability of a quarry sand-heap. For the two clay fields examined here it was found that θ is of the order of a few tens of centimetres and more specifically it ranges from 10.7 to 32.8 cm for the Pentakomo field and from 10.5 to 28.4 cm for the Armou field. For the sand field examined it was found that θ ranges from 91.9 to 264 cm. The high variability of the two clay sites is also indicated by the undrained shear strength values obtained in the laboratory.

Regarding now the statistical uncertainty, the parametric analyses performed showed that statistical uncertainty in designing geotechnical structures can be very high and that, statistical error is not necessarily reduced by increasing the number of sampling points considered. Indeed, the opposite may happen. As shown the statistical error can be minimized or even, eliminated by adopting the proper sampling strategy. The problem of reducing the statistical uncertainty in axially loaded piles, earth retaining structures in active and passive state and finally in elastic settlement and bearing capacity analysis of shallow foundations, relying on targeted field investigation aiming at an optimal design was implemented in a random finite element method (RFEM) framework. The open source RFEM software RPILE1D, RSETL2D and RBEAR2D, have been modified as to include the function of sampling of soil property values from the generated random fields. It is noted that the REARTH2D software has been modified so as to calculate the wall reaction force and overturning moment based on the finite element method instead of Rankine's theory. Two main sampling strategies were investigated, namely, sampling from a single point and sampling from a domain (the latter refers to e.g. continuous Cone Penetration Test, Standard Penetration Test data). Based on the results the following conclusions can be drawn:

For axially loaded piles:

1. when continuous probing test data are used and the pile stiffness is much greater than the stiffness of the surrounding soil, the entire pile length is advised to be taken into account both in the SLS and the ULS state, as shorter sampling domain lengths may insert great statistical error.
2. weak piles, such as, timber piles, require much shorter domain lengths (measured from the top), as longer sampling domains may increase the error.

3. when the design is based on sampling points and not on continuous probing test data, the best practice for minimizing the statistical error is sampling from the mid-height of the pile in both SLS and ULS states.
4. when the pile is relatively weak, the optimal sampling point for the SLS state lies near the top of the pile. For the ULS state, the optimal sampling point lies at the mid-height of pile, for any pile stiffness.
5. the optimal horizontal sampling location is at location where the pile is going to be constructed.

For earth retaining structures:

1. the optimal horizontal sampling location in the active state is at x/H =0 that is, immediately adjacent to the wall face, while the optimal horizontal sampling location in the passive state is half wall height (x/H =0.5) away from the wall face.
2. regarding the depth of the optimal sampling point, for both states it was found that this lies at depth greater than the 2/3 or 1/2 of the wall height for the sliding and rotational mode of failure respectively; the exact depth depends on the spatial correlation length of the soil.
3. for the optimal sampling domain length, it is advisable that the entire wall height to be considered. This practice may significantly reduce the statistical error.

For elastic settlement and bearing capacity analysis of shallow foundations:

1. for the case of a single footing (no interference with adjacent footings), the geometric centre on its plan-view is the optimal sampling location.
2. regarding the depth of the optimal sampling point, it was found that it lies at depth equal to approximately $0.8B$; the exact depth depends on the spatial correlation length of the soil.
3. the optimal sampling domain length, it is advisable that a domain length of at least $2B$ to be taken into account.
4. sampling away from the centre of footing may lead to significant statistic error.

For all problems examined above it was observed that the benefit from a targeted field investigation is much greater as compared to the benefit gained using characteristic soil property values. Moreover, despite the conservatism which is inserted in the analysis using the characteristic value concept, the characteristic values alone, as shown, cannot guaranty a conservative enough engineering study. The safety level can be increased by applying a statistical uncertainty partial factor (similar to the model factor γ_R used by Eurocode 7) or a unified and more conservative model factor to the resistances, which will absorb the statistical uncertainties related to the soil.

Appendix

Appendix I

The generated samples of random fields by the *tstlas1* program may be validated by comparing the average autocorrelation function of the samples to the theoretical estimate given by Vanmarcke [72] for arbitrary domain length D:

$$\rho_{\text{theoretical}} = \left(1 - \frac{\tau}{D}\right) \frac{e^{-2\tau/\theta} - \gamma(D/\theta)}{1 - \gamma(D/\theta)} \tag{A.1}$$

where, the autocovariance γ is given by

$$\gamma(D/\theta) = \frac{e^{-2D/\theta} - 1 + 2D/\theta}{2(D/\theta)^2} \tag{A.2}$$

In Figure A. 1 the average autocorrelation function of samples of 5000 random fields are given for four cases which are combinations of different domain lengths and sampling intervals. Although, the result (A.1) is derived for infinitesimal sampling interval, the agreement between the numerical and analytical results is excellent. The fluctuations of the numerical autocorrelation function are due to the finite number of realizations.

Figure A. 2 presents the dependence of the mean and standard deviation of the samples of thetas, estimated by each one of the eight methods discussed here, as a function of the number of realizations. One may infer that 5000 realizations used are an adequate number for convergent results. Also, one may observe that the convergence of the standard deviation requires a larger number of realizations than the mean value.

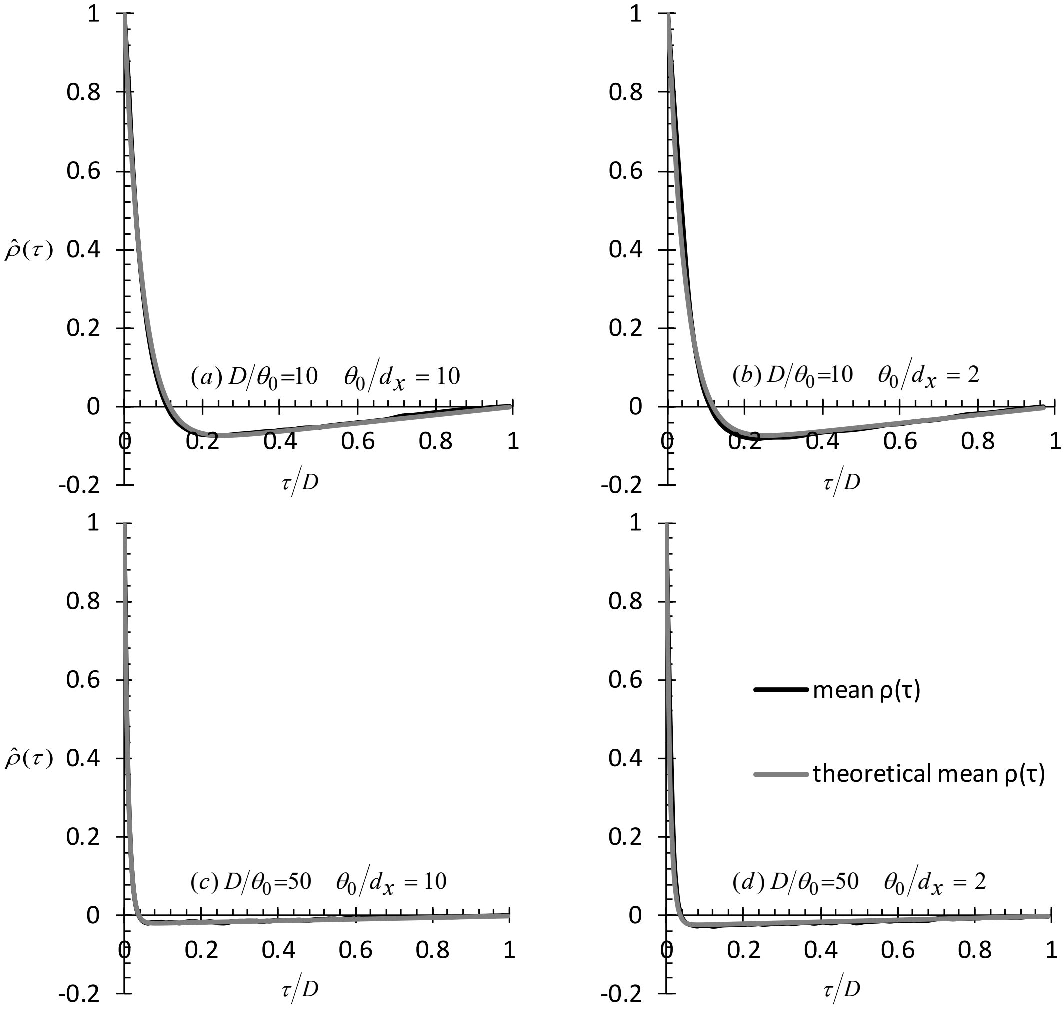


Figure A. 1: Comparison of the sample mean ACF with the theoretical mean ACF.

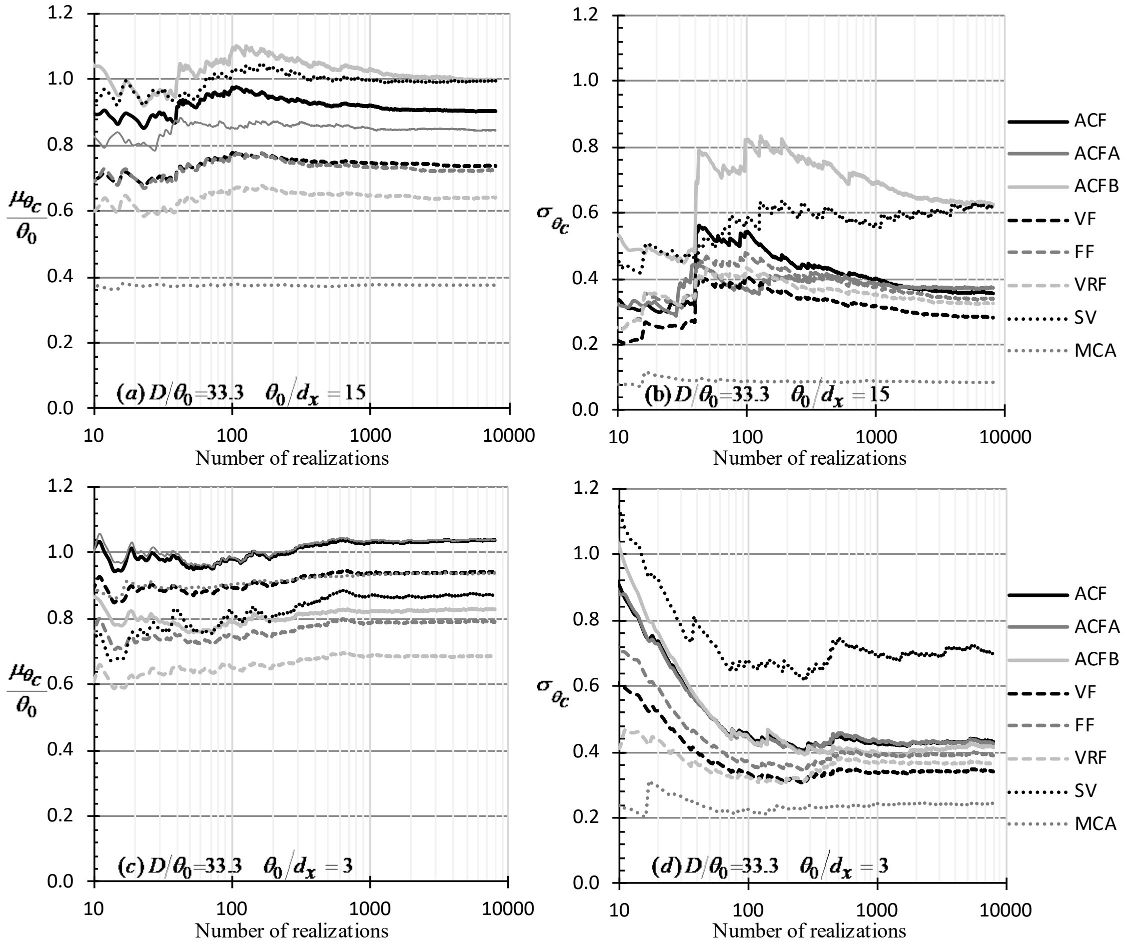


Figure A. 2: Convergence of sample mean value and standard deviation with number of realizations.

Appendix II

Stability of numerical results (number of realizations considered for the RFEM models)- Axially loaded piles.

The parametric analysis carried out is, in essence, a comparison study between different scenarios for determining the optimal sampling location or domain length for effectively designing axially loaded piles. When dealing with small differences in p_f, the need for statistically stable results is even greater. In this respect, 20,000 realizations were considered. As shown in the example curves of Figure B.1, this number of realizations is adequate.

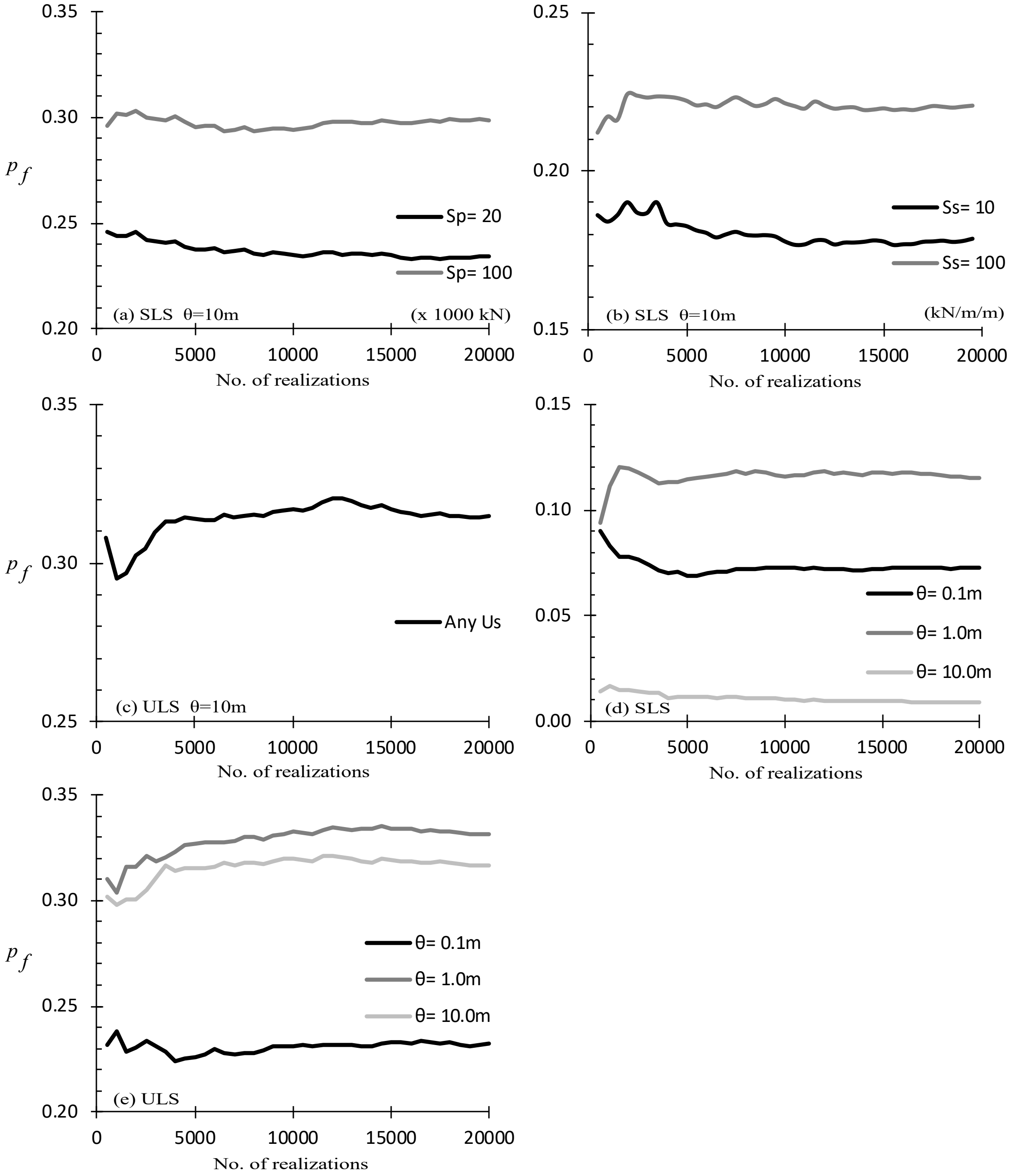


Figure B. 1: p_f vs number of realizations for different values of pile stiffness (S_p), soil stiffness (S_s),soil strength (U_s) and spatial correlation length (θ).

Appendix III

Stability of numerical results (number of realizations considered in the RFEM models) and effect of element size of active earth pressure analysis.

As shown in Figure C. 1, 3000 realizations can be deemed sufficient for the problem of active earth pressure analysis. Although, the p_f versus number of realizations curves of the figure in question refer to d_p/H=0.667 and (when not mentioned) to $x/H = 0$, the same results are taken for other sampling locations.

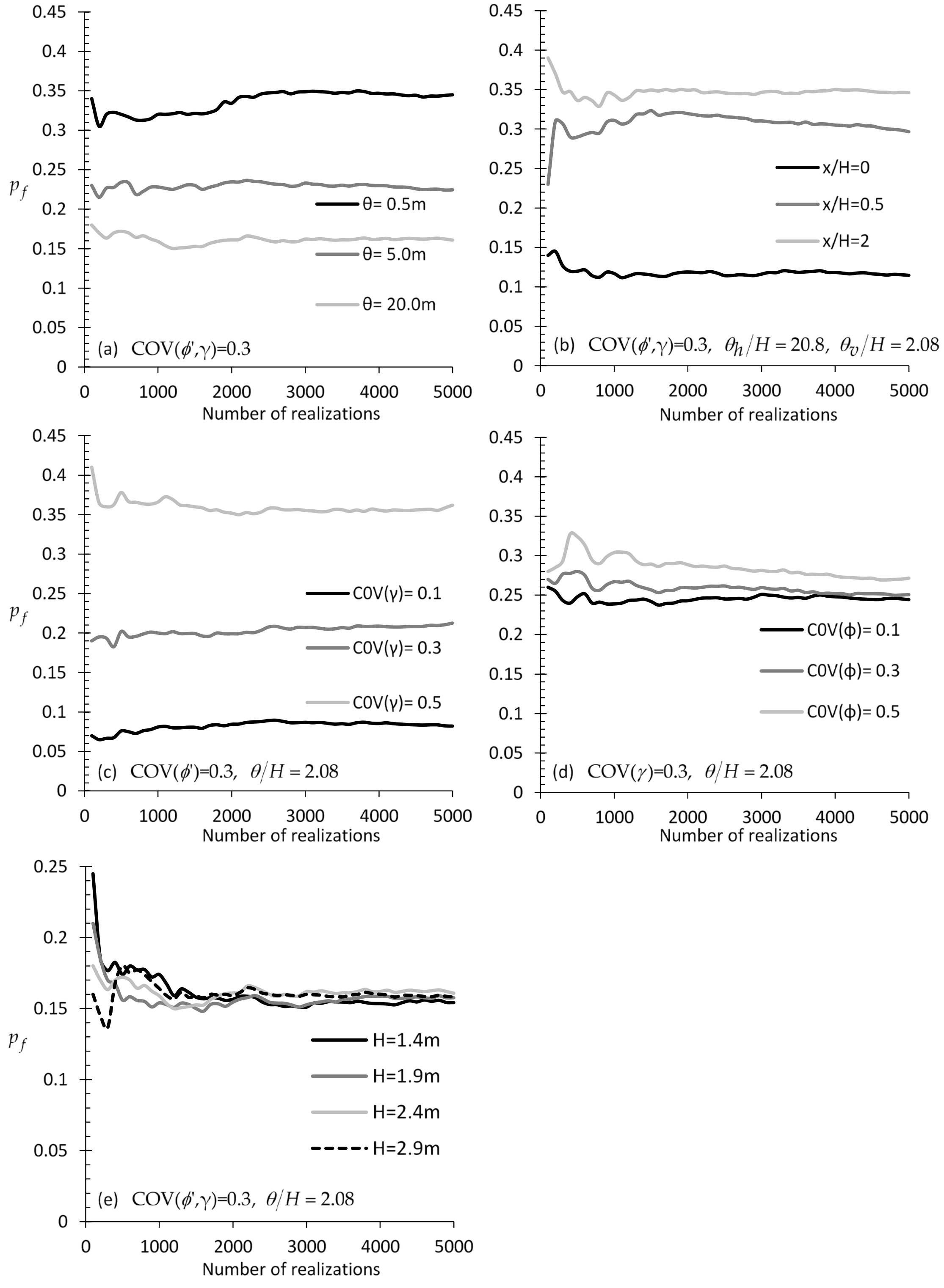


Figure C. 1: p_f versus number of realizations charts for different (a) correlation length values (case of isotropic field), (b) x/H values (case of anisotropic field), c) COV value of γ

(isotropic field), (d) COV value of ϕ (isotropic field), and (e) wall heights; figures referring to sampling from $x/H = 0$ and $d_p/H = 0.667$.

Regarding the element size, it is recommended that this be less than the half of the spatial correlation length, although an element size equal to the θ does not introduce great error in the analysis (e.g. see the comparison chart of Figure C. 2 where the reference wall has been used).

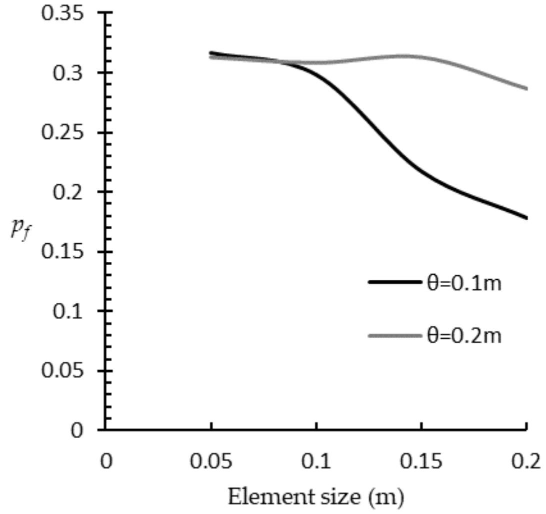


Figure C. 2: Effect of element size on the failure probability.

Stability of Numerical Results (Number of Realizations Considered in the RFEM Models) of passive earth pressure analysis.

As shown in Figure C. 3, 3000 realizations can be deemed sufficient for the problem of passive earth pressure analysis. Although the p_f versus number of realizations curves of the figure in question refer to d_p/H = 0.667 and (when not mentioned) to $x/H = 0.5$, the same results are taken for other sampling locations.

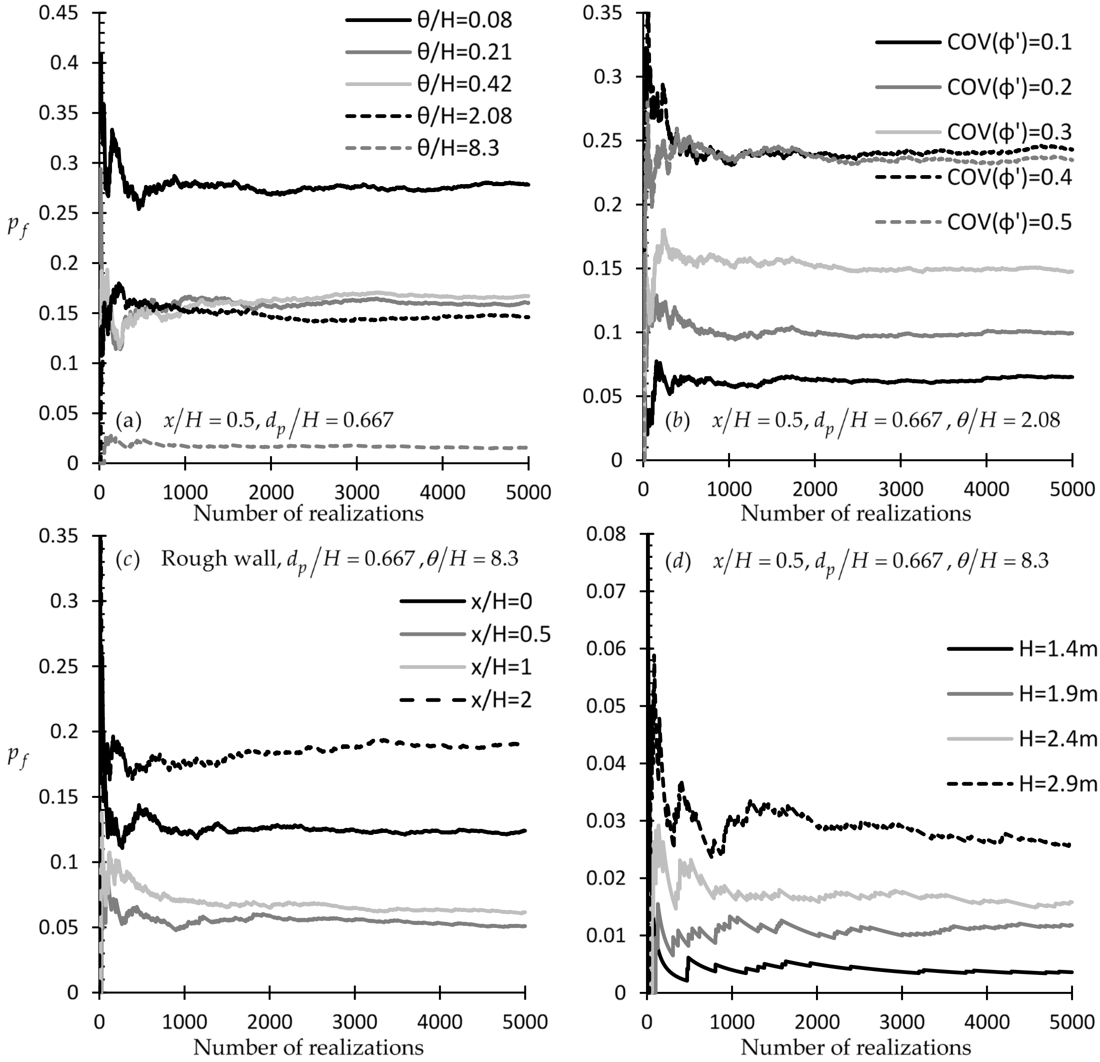


Figure C. 3: p_f versus number of realizations charts for different (a) correlation length values, (b) COV value of ϕ (isotropic field), (c) x/Hvalues and (d) wall heights; if not mentioned otherwise, figures referring to sampling from x/H=0.25 and $d_p/H = 0.667$.

Appendix IV

Stability of numerical results (distance of lateral boundaries and number of realizations considered in the RFEM models) for elastic settlement analysis.

Figure D. 1 shows the variation of p_f with the number of realizations for various cases examined in this study. As it is inferred, the 10,000 realizations considered in the present analysis can be deemed sufficient. The p_f versus number of realizations curves of the figure in question refer to d_p/B=0.7 and x/B =0, however, the same outcome is obtained for any sampling location.

The effect of the distance between the footing and the lateral boundaries in the finite element analysis carried out in the RFEM framework is shown in Figure D. 2. From this figure it is inferred that, a distance of $1.1B$ can be considered adequate in order to obtain reliable results. In this respect, the authors considered a distance equal to $1.7B$.

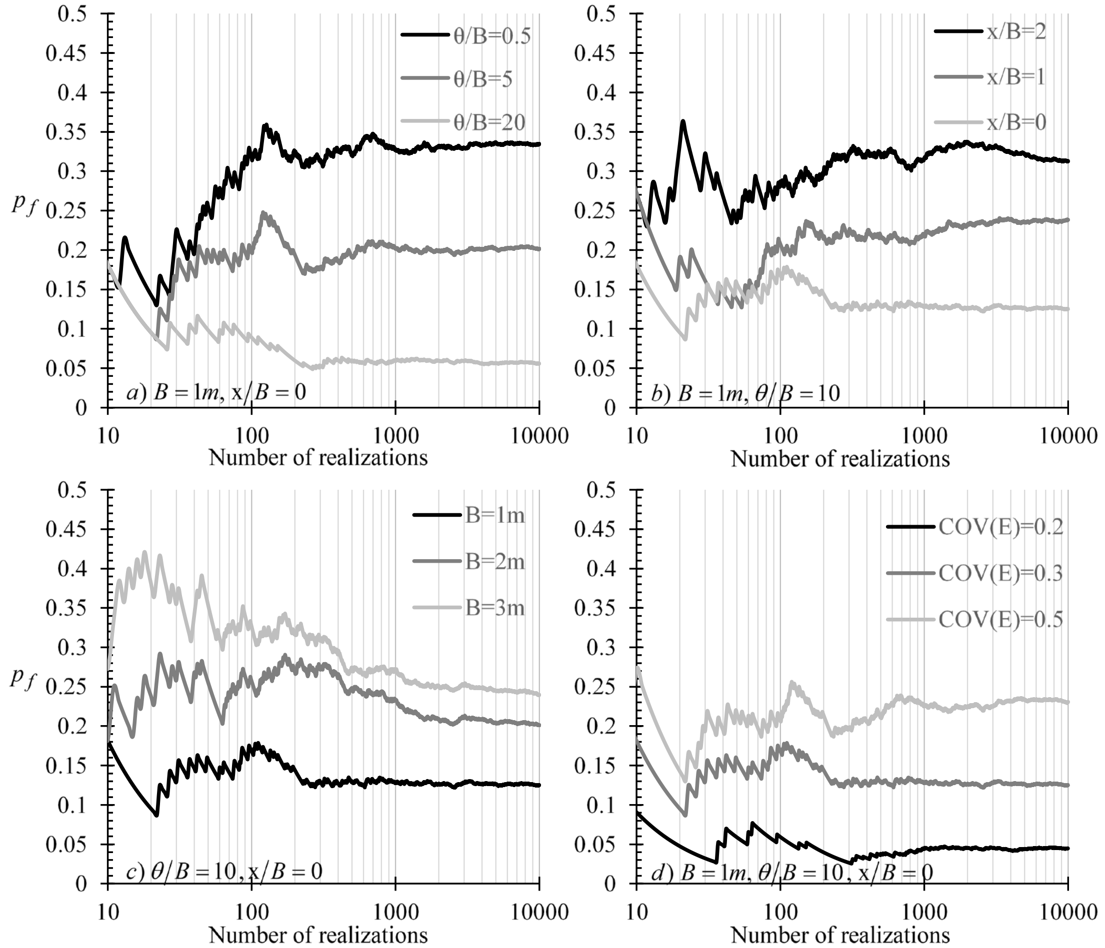


Figure D. 1: p_f versus number of realizations by considering different (a) scaled correlation length values, (b) scaled horizontal sampling location values, c) footing widths and (d) COV of E; charts referring to $d_p/B = 0.7B$ and x/B =0.

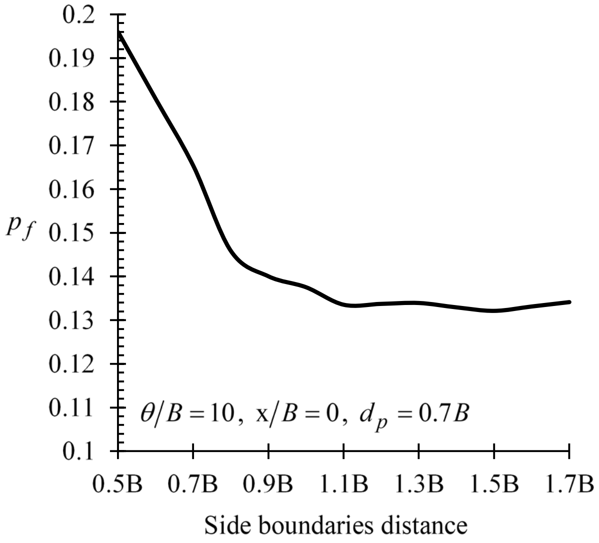


Figure D. 2: p_f versus side boundaries distance from the edge of footing of width B=1m.

Stability of numerical results (side boundaries distance and number of realizations considered in the RFEM models) for bearing capacity analysis.

The effect of the distance between the footing and the lateral boundaries in the finite element analysis carried out in the RFEM framework is shown in Figure D. 3 . From this figure it is inferred that, a distance of 2B can be considered adequate in order to obtain reliable results. In this respect, the authors considered a distance equal to 2.5B.

In addition, Figure D. 4 shows the variation of p_f with the number of realizations for various cases examined in this study. As it is inferred, the 3,000 realizations considered in the present analysis can be deemed sufficient. The p_f versus number of realizations curves of the figure in question refer to d_p/B=1, however, the same outcome is obtained for any sampling depth.

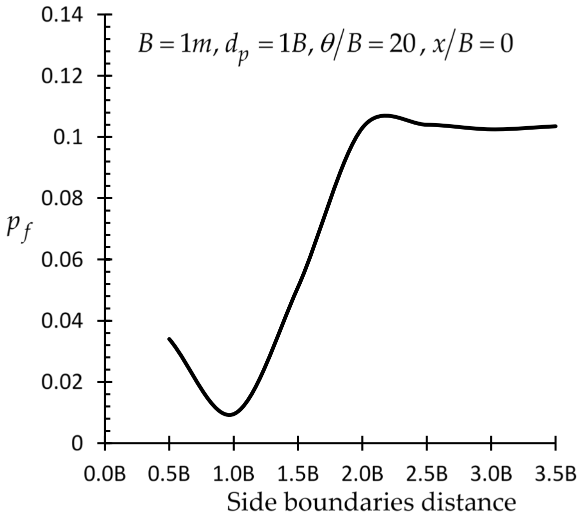


Figure D. 3: $p_f$versus side boundaries distance from the edge of footing with width B=1m.

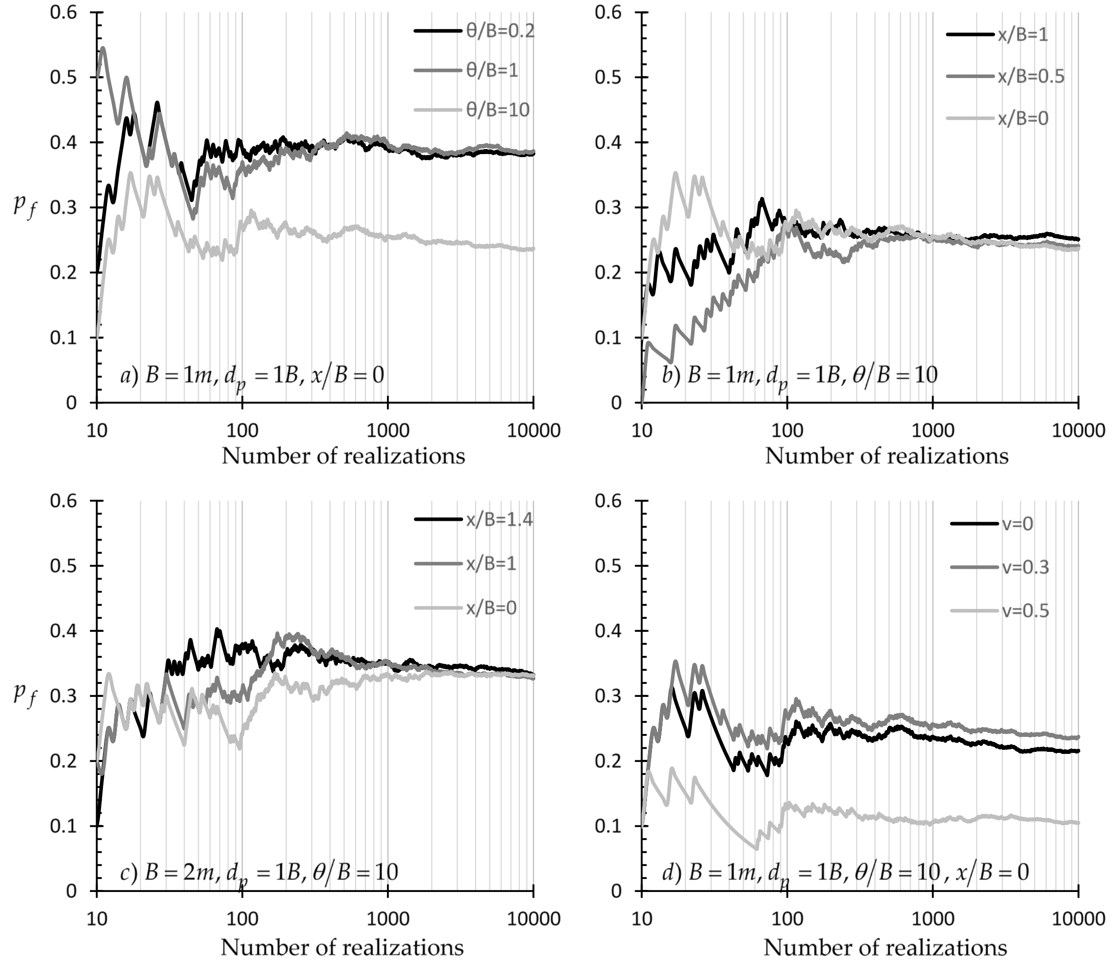


Figure D. 4: p_f versus number of realizations by considering different (a) scaled correlation length values, (b) (normalized) distance of soil sampling from the footing axis, c) footing width and (d) Poisson's ratio values.

References

1. Phoon, K.-K.; Kulhawy, F.H.; Grigoriu, M.D. Reliability-based design for transmission line structure foundations. *Comput. Geotech.* **2000**, *26*, 169–185.
2. Tang, W.H.; Zhang, L. Quality assurance and risk reduction in foundation engineering. In *Frontier Technologies for Infrastructures Engineering: Structures and Infrastructures Book Series*; Chen, S.-S., Ang H-S., A., Eds.; CRC Press: Boca Raton, FL, USA, 2009; Vol. 4, pp. 123–142.
3. Duzgun, H.S.B.; Yucemen, M.S.; Karpuz, C. A probabilistic model for the assessment of uncertainties in the shear strength of rock discontinuities. *Int. J. rock Mech. Min. Sci.* **2002**, *39*, 743–754.
4. Lacasse, S.; Nadim, F. Uncertainties in characterising soil properties. *Publ. Geotek. Inst.* **1996**, *201*, 49–75.
5. Christian, J.T.; Ladd, C.C.; Baecher, G.B. Reliability applied to slope stability analysis. *J. Geotech. Eng.* **1994**, *120*, 2180–2207.
6. Shen, M.-Y.; Cao, Z.-J.; Li, D.-Q.; Wang, Y. Probabilistic Characterization of Site-Specific Inherent Variability of Undrained Shear Strength Using Both Indirect and Direct Measurements. *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.* **2017**, *4*, 4017038.
7. Phoon, K.-K.; Kulhawy, F.H. Characterization of geotechnical variability. *Can. Geotech. J.* **1999**.
8. Phoon, K.-K.; Kulhawy, F.H. Evaluation of geotechnical property variability. *Can. Geotech. J.* **1999**, *36*, 625–639.
9. Wang, Y.; Cao, Z.; Li, D. Bayesian perspective on geotechnical variability and site characterization. *Eng. Geol.* **2016**, *203*, 117–125.
10. Cho, S.E. Probabilistic assessment of slope stability that considers the spatial variability of soil properties. *J. Geotech. geoenvironmental Eng.* **2010**, *136*, 975–984.
11. El-Ramly, H.; Morgenstern, N.R.; Cruden, D.M. Probabilistic slope stability analysis for practice. *Can. Geotech. J.* **2002**, *39*, 665–683.

12. Hicks, M.A.; Spencer, W.A. Influence of heterogeneity on the reliability and failure of a long 3D slope. *Comput. Geotech.* **2010**, *37*, 948–955.

13. Jamshidi Chenari, R.; Alaie, R. Effects of anisotropy in correlation structure on the stability of an undrained clay slope. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2015**, *9*, 109–123.

14. Ji, J.; Liao, H.J.; Low, B.K. Modeling 2-D spatial variation in slope reliability analysis using interpolated autocorrelations. *Comput. Geotech.* **2012**, *40*, 135–146.

15. Jiang, S.-H.; Li, D.-Q.; Zhang, L.-M.; Zhou, C.-B. Slope reliability analysis considering spatially variable shear strength parameters using a non-intrusive stochastic finite element method. *Eng. Geol.* **2014**, *168*, 120–128.

16. Le, T.M.H. Reliability of heterogeneous slopes with cross-correlated shear strength parameters. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2014**, *8*, 250–257.

17. Low, B.K.; Lacasse, S.; Nadim, F. Slope reliability analysis accounting for spatial variation. *Georisk* **2007**, *1*, 177–189.

18. Srivastava, A.; Babu, G.L.S.; Haldar, S. Influence of spatial variability of permeability property on steady state seepage flow and slope stability analysis. *Eng. Geol.* **2010**, *110*, 93–101.

19. Gravanis, E.; Pantelidis, L.; Griffiths, D. V An analytical solution in probabilistic rock slope stability assessment based on random fields. *Int. J. Rock Mech. Min. Sci.* **2014**, *71*, 19–24.

20. Pantelidis, L.; Gravanis, E.; Griffiths, D. V Influence of spatial variability on rock slope reliability using 1-D random fields. In *Engineering Geology for Society and Territory-Volume 2*; Springer, 2015; pp. 1235–1238.

21. Chok, Y.H.; Jaksa, M.B.; Griffiths, D. V; Fenton, G.A.; Kaggwa, W.S. A parametric study on reliability of spatially random cohesive slopes. *Aust. Geomech.* **2007**, *42*, 79–85.

22. Griffiths, D. V; Huang, J.; Fenton, G.A. Comparison of slope reliability methods of analysis. In *GeoFlorida 2010: Advances in Analysis, Modeling & Design*; 2010; pp. 1952–1961.

23. Griffiths, D. V; Fenton, G.A. Influence of soil strength spatial variability on the stability of an undrained clay slope by finite elements. In *Slope stability 2000*; 2000; pp. 184–193.

24. Griffiths, D. V; Huang, J.; Fenton, G.A. Influence of Spatial Variability on Slope Reliability Using 2-D Random Fields. **2009**, 1367–1378.

25. Griffiths, D. V; Huang, J.; Fenton, G.A. On the reliability of earth slopes in three dimensions. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **2009**, *465*, 3145–3164.

26. Griffiths, D. V; Huang, J.; Fenton, G.A. Probabilistic infinite slope analysis. *Comput. Geotech.* **2011**, *38*, 577–584.

27. Griffiths, D. V.; Fenton, G.A. Probabilistic Slope Stability Analysis by Finite Elements. *J. Geotech. Geoenvironmental Eng.* **2004**, 507–518.

28. Griffiths, D. V; Huang, J.; Fenton, G.A. Modeling of stability and risk of geotechnical systems in highly variable soils. In Proceedings of the International conference on advances in geotechnical engineering; 2011.

29. Huang, J.; Griffiths, D. V; Fenton, G.A. System reliability of slopes by RFEM. *Soils Found.* **2010**, *50*, 343–353.

30. Zhu, H.; Griffiths, D. V; Fenton, G.A.; Zhang, L.M. Undrained failure mechanisms of slopes in random soil. *Eng. Geol.* **2015**, *191*, 31–35.

31. Fenton, G.; Griffiths, D. V. Bearing-capacity prediction of spatially random c - ϕ soils. *Can. Geotech. J.* **2003**, *40*, 54–65.

32. Griffiths, D. V.; Fenton, G.A. Bearing capacity of spatially random soil: the undrained clay Prandtl problem revisited. *Géotechnique* **2001**, *51*, 351–359.

33. Griffiths, D. V.; Fenton, G. a.; Manoharan, N. Bearing Capacity of Rough Rigid Strip Footing on Cohesive Soil: Probabilistic Study. *J. Geotech. Geoenvironmental Eng.* **2002**, *128*, 743–755.

34. Griffiths, D. V.; Fenton, G. a.; Manoharan, N. Undrained Bearing Capacity of Two-Strip Footings on Spatially Random Soil. *Int. J. Geomech.* **2006**, *6*, 421-427Griffiths, D. V., Fenton, G. a. and Manohar.

35. Pieczyńska, J.; Puła, W.; Griffiths, D. V; Fenton, G. a Probabilistic characteristics

of strip footing bearing capacity evaluated by random finite element method. *Civ. Eng.* **2011**, 1673–1682.

36. Pieczyńska, J.; Puła, W.; Griffiths, D. V; Fenton, G.A. Influence of embedment, self-weight and anisotropy on bearing capacity reliability using the random finite element method. **2015**, *67*, 229–238.

37. Li, J.H.; Cassidy, M.J.; Tian, Y.; Huang, J.; Lyamin, A. V; Uzielli, M. Comparative study of bearing capacity of buried footings using random limit analysis and random finite element method. In Proceedings of the Computer Methods and Recent Advances in Geomechanics: Proceedings of the 14th International Conference of International Association for Computer Methods and Recent Advances in Geomechanics, 2014 (IACMAG 2014); Taylor & Francis Books Ltd, 2015; pp. 1301–1305.

38. Kasama, K.; Whittle, A.J. Bearing Capacity of Spatially Random Cohesive Soil Using Numerical Limit Analyses. *J. Geotech. Geoenvironmental Eng.* **2011**, *137*, 989–996.

39. Popescu, R.; Deodatis, G.; Nobahar, A. Effects of random heterogeneity of soil properties on bearing capacity. *Probabilistic Eng. Mech.* **2005**, *20*, 324–341.

40. Sivakumar Babu, G.L.; Srivastava, A.; Murthy, D.S.N. Reliability analysis of the bearing capacity of a shallow foundation resting on cohesive soil. *Can. Geotech. J.* **2006**, *43*, 217–223.

41. Nainegali, L.S.; Basudhar, P.K.; Ghosh, P.; ASCE, A.M.; Ghosh, P. Interference of two asymmetric closely spaced strip footings resting on nonhomogeneous and linearly elastic soil bed. *Int. J. Geomech.* **2013**, *13*, 840–851.

42. Puła, W.; Zaskórski, Ł. Estimation of the probability distribution of the random bearing capacity of cohesionless soil using the random finite element method. *Struct. Infrastruct. Eng.* **2015**, *11*, 707–720.

43. Vessia, G.; Cherubini, C.; Pieczyńska, J.; Puła, W. Application of random finite element method to bearing capacity design of strip footing. *J. Geoengin.* **2009**, *4*, 103–112.

44. Fenton, G.A.; Griffiths, D. V. Probabilistic Foundation Settlement on Spatially

Random Soil. *J. Geotech. Geoenvironmental Eng.* **2002**, *128*, 381–390.

45. Fenton, G. a.; Griffiths, D. V. Three-Dimensional Probabilistic Foundation Settlement. *J. Geotech. Geoenvironmental Eng.* **2005**, *131*, 232–239.

46. Griffiths, D. V; Fenton, G.A. Probabilistic settlement analysis of rectangular footings.; AA BALKEMA PUBLISHERS, 2005; Vol. 16, p. 1041.

47. Griffiths, D. V; Fenton, G.A. Probabilistic settlement analysis by stochastic and random finite-element methods. *J. Geotech. geoenvironmental Eng.* **2009**, *135*, 1629–1637.

48. Ahmed, A.; Soubra, A.-H. Probabilistic analysis of strip footings resting on a spatially random soil using subset simulation approach. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2012**, *6*, 188–201.

49. Griffiths, D. V; Fenton, G.A.; Ziemann, H.R. Reliability of passive earth pressure. *Georisk* **2008**, *2*, 113–121.

50. Chauhan, V.B.; Dasaka, S.M.; Dasgupta, U.S. Effect of Spatial Variability on the Earth Pressure of a Rigid Retaining Wall. In *Geo-Risk 2017*; pp. 350–359.

51. Dodagoudar, G.; Sayed, S.; Rajagopal, K. Random field modeling of reinforced retaining walls. *Int. J. Geotech. Eng.* **2015**, *9*, 229–238.

52. Fenton, G.A.; Griffiths, D. V Reliability-based deep foundation design. *Geo-Denver 2007 New Peaks Geotech.* **2007**.

53. Fenton, G.; Griffiths, D. Reliability-based geotechnical engineering. *GeoFlorida Conf. Keynote Lect. ...* **2010**.

54. Dasaka, S.M.; Zhang, L.M. Spatial variability of in situ weathered soil. *Géotechnique* **2012**, *62*, 375–384.

55. Breysse, D.; La Borderie, C.; Elachachi, S.M.; Niandou, H. Spatial variations in soil properties and their influence on structural reliability. *Civ. Eng. Environ. Syst.* **2007**, *24*, 73–83.

56. Uzielli, M.; Lacasse, S.; Nadim, F.; Phoon, K.K. Soil Variability Analysis for Geotechnical Practice. **2006**.

57. Cafaro, F.; Cherubini, C. Large sample spacing in evaluation of vertical strength

variability of clayey soil. *J. Geotech. Geoenvironmental Eng.* **2002**, *128*, 558–568.

58. Onyejekwe, S.; Ge, L. Scale of fluctuation of geotechnical parameters estimated from CPTu and laboratory test data. *Found. Eng. Face* **2013**.

59. Vanmarcke, E.H. Probabilistic modeling of soil profiles. *J. Geotech. Eng. Div.* **1977**, *103*, 1227–1246.

60. Cho, S.E. Probabilistic analysis of seepage that considers the spatial variability of permeability for an embankment on soil foundation. *Eng. Geol.* **2012**, *133*, 30–39.

61. Fenton, G.; Griffiths, D. V. *Risk assessment in geotechnical engineering*; Wiley: Hoboken, NJ, USA, 2008; ISBN 9780470178201.

62. Fenton, G.A.; Vanmarcke, E.H. Simulation of random fields via local average subdivision. *J. Eng. Mech.* **1990**, *116*, 1733–1749.

63. Orr, T.L.L. Managing risk and achieving reliable geotechnical designs using Eurocode 7. In *Risk and Reliability in Geotechnical Engineering*; Phoon, K.K., Ching, J., Eds.; CRC Press: Boca Raton, FL, USA, 2014; pp. 395–433.

64. Ching, J.; Phoon, K.-K. Constructing multivariate distributions for soil parameters. *Risk Reliab. Geotech. Eng.* **2015**, 3–76.

65. Huang, X.C.; Zhou, X.P.; Ma, W.; Niu, Y.W.; Wang, Y.T. Two-dimensional stability assessment of rock slopes based on random field. *Int. J. Geomech.* **2016**, *17*, 4016155.

66. Imanzadeh, S.; Denis, A.; Marache, A. Settlement uncertainty analysis for continuous spread footing on elastic soil. *Geotech. Geol. Eng.* **2015**, *33*, 105–122.

67. DeGroot, D.J.; Baecher, G.B. Estimating Autocovariance of In-Situ Soil Properties. *J. Geotech. Eng.* **1993**, *119*, 147–166.

68. Uzielli, M. Variability of stress-normalized CPT Parameters and application to seismic liquefaction initiation analysis, Ph.D. dissertation, University of Florence, Italy, 2004.

69. Tan, T.S.; Phoon, K.-K.; Hight, D.W.; Leroueil, S.; (Eds.) *Characterisation and Engineering Properties of Natural Soils*; CRC Press, 2006;

70. Kenarsari, A.E.; Chenari, R.J.; Eslami, A. Characterization of the correlation

structure of residual CPT profiles in sand deposits. *Int. J. Civ. Eng.* **2013**, *11*, 29–37.

71. Onyejekwe, S.; Kang, X.; Ge, L. Evaluation of the scale of fluctuation of geotechnical parameters by autocorrelation function and semivariogram function. *Eng. Geol.* **2016**, *214*, 43–49.

72. Vanmarcke, E.H. Random fields: analysis and synthesis. *MIT Press. Cambridge, MA* **1983**.

73. Zhu, H.; Zhang, L.M. Characterizing geotechnical anisotropic spatial variations using random field theory. *Can. Geotech. J.* **2013**, *50*, 723–734.

74. Zhang, L.L.; Zhang, L.M.; Tang, W.H. Similarity of soil variability in centrifuge models. *Can. Geotech. J.* **2008**, *45*, 1118–1129.

75. Uzielli, M.; Vannucchi, G.; Phoon, K.K. Random field characterisation of stress-normalised cone penetration testing parameters. *Géotechnique* **2005**, *55*, 3–20.

76. Stuedlein, A.W.; Kramer, S.L.; Arduino, P.; Holtz, R.D. Geotechnical characterization and random field modeling of desiccated clay. *J. Geotech. Geoenvironmental Eng.* **2012**, *138*, 1301–1313.

77. Popescu, R.; Prevost, J.H.; Vanmarcke, E.H. Numerical simulations of soil liquefaction using stochastic input parameters. **1995**.

78. Phoon, K.-K. Bootstrap estimation of sample autocorrelation functions. In *GeoCongress 2006: Geotechnical Engineering in the Information Technology Age*; 2006; pp. 1–6.

79. Müller, R. Probabilistic stability analysis of embankments founded on clay 2013.

80. Liu, C.-N.; Chen, C.-H. Mapping liquefaction potential considering spatial correlations of CPT measurements. *J. Geotech. geoenvironmental Eng.* **2006**, *132*, 1178–1187.

81. Liu, C.N.; Chen, C.H. Spatial correlation structures of CPT data in a liquefaction site. *Eng. Geol.* **2010**, *111*, 43–50.

82. Keaveny, J.M.; Nadim, F.; Lacasse, S. Autocorrelation functions for offshore geotechnical data. In Proceedings of the Structural Safety and Reliability; ASCE,

1989; pp. 263–270.

83. Jaksa, M.B. The influence of spatial variability on the geotechnical design properties of a stiff, overconsolidated clay. **1995**.

84. Firouzianbandpey, S.; Griffiths, D. V; Ibsen, L.B.; Andersen, L. V Spatial correlation length of normalized cone data in sand : case study in the north of Denmark. *Can. Geotech. J.* **2014**, *51(8)*, 844–857.

85. Baecher, G.B.; Christian, J.T. *Reliability and Statistics in Geotechnical Engineering*; 2003; ISBN 0471498335.

86. Ching, J.; Wu, T.J.; Stuedlein, A.W.; Bong, T. Estimating horizontal scale of fluctuation with limited CPT soundings. *Geosci. Front.* **2017**, 1–12.

87. Pieczyńska-Kozłowska, J.M.; Puła, W.; Vessia, G. A Collection of Fluctuation Scale Values and Autocorrelation Functions of Fine Deposits in Emilia Romagna Plain, Italy. In *Geo-Risk 2017*; 2017; pp. 290–299.

88. Oguz, E.A.; Huvaj, N.; Griffiths, D.V. Vertical spatial correlation length based on standard penetration tests. *Mar. Georesources Geotechnol.* **2018**, *0*, 1–12.

89. Yue, Q.; Yao, J.; Ang, A.H.S.; Spanos, P.D. Efficient random field modeling of soil deposits properties. *Soil Dyn. Earthq. Eng.* **2018**, *108*, 1–12.

90. Nuttall, J.D. Estimating spatial correlations from CPT data using neural networks and random fields. **2018**.

91. Vessia, G.; Russo, S. Random field theory to interpret the spatial variability of lacustrine soils. *Biosyst. Eng.* **2018**, *168*, 4–13.

92. Fenton, G.A. Estimation for stochastic soil models. *J. Geotech. Geoenvironmental Eng.* **1999**, *125*, 470–485.

93. Akkaya, A.D.; Vanmarcke, E.H. Estimation of spatial correlation of soil parameters based on data from the Texas A&M University NGES. *Probabilistic site Charact. Natl. Geotech. Exp. Sites (eds GA Fent. EH Vanmarcke)* **2003**, 29–40.

94. Onyejekwe, S. Characterization of soil variability for reliability-based design, Missouri University of Science and Technology, 2012.

95. Jamshidi Chenari, R.; Oloomi Dodaran, R. New method for estimation of the scale of fluctuation of geotechnical properties in natural deposits. *Comput. Methods Civ. Eng.* **2010**, *1*, 55–64.

96. Lingwanda, M.I.; Prästings, A.; Larsson, S.; Nyaoro, D.L. Comparison of geotechnical uncertainties linked to different soil characterization methods. *Geomech. Geoengin.* **2017**, *12*, 137–151.

97. Onisiphorou, C. Stochastic analysis of saturated soils using finite elements. **2000**.

98. Li, X.Y. Statistic Analyses of Correlation Distances of Soil Parameters. In Proceedings of the Key Engineering Materials; Trans Tech Publ, 2011; Vol. 480, pp. 1412–1417.

99. Czmoch, I. Discretization of Random Fields Representing Material Properties and Distributed Loads in FORM Analysis. *Dependability Eng.* **2018**.

100. Campanella, R.G.; Wickremesinghe, D.S.; Robertson, P.K. Statistical Treatment of Cone Penetrometer test data 1987, *2*, 1010–1019.

101. Wickremesinghe, D.; Campanella, R. Scale of fluctuation as a descriptor of soil variability. *Probabilistic Methods Geotech.* **1993**.

102. Cherubini, C.; Vessia, G.; Pula, W. Statistical soil characterization of Italian sites for reliability analysis. In Proceedings of the Proc. of 2nd Intern. workshop on charact. & eng. prop. of natural soils (Invited Paper), Singapore; 2007; Vol. 4, pp. 2681–2706.

103. Yan, S.; Guo, L. Calculation of scale of fluctuation and variance reduction function. *Trans. Tianjin Univ.* **2015**, *21*, 41–49.

104. Pantelidis, L.; Christodoulou, P. Spatial Correlation Length of Clay Soils in Practice and Its Influence in Probabilistic Bearing Capacity Analysis. **2017**, 487–496.

105. Clark, I. *Practical geostatistics*; Applied Science Publishers, 1979; ISBN 085334843X.

106. Chiasson, P.; Lafleur, J.; Soulié, M.; Law, K.T. Characterizing spatial variability of a clay by geostatistics. *Can. Geotech. J.* **1995**, *32*, 1–10.

107. DeGroot, D.J. Analyzing spatial variability of in situ soil properties. In Proceedings of the Uncertainty in the geologic environment: From theory to practice; ASCE, 1996; pp. 210–238.

108. Meek, D.W. A semiparametric method for estimating the scale of fluctuation. *Comput. Geosci.* **2001**, *27*, 1243–1249.

109. Jones, A.L.; Kramer, S.S.L.; Arduino, P. Estimation of uncertainty in geotechnical properties for performance-based earthquake engineering. **2002**.

110. Elkateb, T.; Chalaturnyk, R.; Robertson, P.K. Simplified geostatistical analysis of earthquake-induced ground response at the Wildlife Site, California, USA. *Can. Geotech. J.* **2003**, *40*, 16–35.

111. Jaksa, M. *Modeling the natural variability of over-consolidated clay in Adelaide, South Australia*; 2007; ISBN 041542691X.

112. Chen, Q.; Wang, C.; Hsein Juang, C. CPT-based evaluation of liquefaction potential accounting for soil spatial variability at multiple scales. *J. Geotech. Geoenvironmental Eng.* **2015**, *142*, 4015077.

113. Faisal, A.H.M.H. Evaluation of Site Variability and its Application in Design. **2018**.

114. Jaksa, M.B.; Goldsworthy, J.S.; Fenton, G.A.; Kaggwa, W.S.; Griffiths, D. V; Kuo, Y.L.; Poulos, H.G. Towards reliable and effective site investigations. *Géotechnique* **2005**, *55*, 109–121.

115. Gong, W.; Luo, Z.; Juang, C.H.; Huang, H.; Zhang, J.; Wang, L. Optimization of site exploration program for improved prediction of tunneling-induced ground settlement in clays. *Comput. Geotech.* **2014**, *56*, 69–79.

116. Mašín, D. The influence of experimental and sampling uncertainties on the probability of unsatisfactory performance in geotechnical applications. *Géotechnique* **2015**, *65*, 897–910.

117. Lo, M.K.; Leung, Y.F. Reliability Assessment of Slopes Considering Sampling Influence and Spatial Variability by Sobol'Sensitivity Index. *J. Geotech. Geoenvironmental Eng.* **2018**, *144*, 4018010.

118. Sobol, I.M. Global sensitivity indices for nonlinear mathematical models and their

Monte Carlo estimates. *Math. Comput. Simul.* **2001**, *55*, 271–280.

119. Yang, R.; Huang, J.; Griffiths, D. V; Sheng, D. Probabilistic Stability Analysis of Slopes by Conditional Random Fields. *Geo-Risk 2017* **2017**, 450–459.

120. Ching, J.; Phoon, K.-K. Characterizing uncertain site-specific trend function by sparse Bayesian learning. *J. Eng. Mech.* **2017**, *143*, 4017028.

121. Yang, R.; Huang, J.; Griffiths, D.V.; Li, J.; Sheng, D. Importance of soil property sampling location in slope stability assessment. *Can. Geotech. J.* **2019**, *56*, 335–346.

122. Fenton, G.A.; Naghibi, F.; Hicks, M.A. Effect of sampling plan and trend removal on residual uncertainty. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2018**, *12(4)*, 253–264.

123. Li, Y.J.J.; Hicks, M.A.A.; Vardon, P.J.J. Uncertainty reduction and sampling efficiency in slope designs using 3D conditional random fields. *Comput. Geotech.* **2016**, *79*, 159–172.

124. Yang, R.; Huang, J.; Griffiths, D. V; Meng, J.; Fenton, G.A. Optimal geotechnical site investigations for slope design. *Comput. Geotech.* **2019**, *114*, 103–111.

125. EN 1997-2 *Eurocode 7-Geotechnical design-Part 2: Ground investigation and testing.*; CEN (European Committee for Standardization): Brussels, Belgium, 2007;

126. AASHTO *AASHTO LRFD Bridge Design Specifications*; American Association of State Highway and Transportation Officials: Washington, DC, USA, 2010; ISBN 9781560514510.

127. Sabatini, P.J.; Bachus, R.C.; Mayne, P.W.; Schneider, J.A.; Zettler, T.E. *Geotechnical Engineering Circular 5 (GEC5)-Evaluation of Soil and Rock Properties*; Report No FHWA-IF-02-034. Federal Highway Administration, US Department of …, 2002;

128. Orr, T.L.L. Defining and selecting characteristic values of geotechnical parameters for designs to Eurocode 7. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2016**, *11*, 103–115.

129. EN 1997-1 *Eurocode 7 Geotechnical design—Part 1: General rules*; CEN

(European Committee for Standardization): Brussels, Belgium, 2004;

130. Canadian Standards Association *Canadian Highway Bridge Design Code. CAN/CSA-S6-14*; Canadian Standards Association: Mississauga, Ont., 2014;

131. Kok-Kwang, P.; Kulhawy, F.H.; Grigoriu, M.D. Development of a Reliability-Based Design Framework for Transmission Line Structure Foundations. *J. Geotech. Geoenvironmental Eng.* **2003**, *129*, 798–806.

132. Fenton, G.A.; Naghibi, F.; Dundas, D.; Bathurst, R.J.; Griffiths, D. V Reliability-based geotechnical design in 2014 Canadian Highway Bridge Design Code. *Can. Geotech. J.* **2015**, *53*, 236–251.

133. JGS *Principles for foundation design grounded on a performance-based design concept (GeoGuide)*; Japanese Geotechnical Society: Tokyo, 2004;

134. White, J.; Yeats, A.; Skipworth, G. *Tables for statisticians*; 3rd ed.; Nelson Thornes: Cheltenham, UK, 1979; ISBN 085950462X.

135. Agterberg, F.P. Autocorrelation functions in geology. In *Geostatistics*; Springer, 1970; pp. 113–141.

136. Bowerman, B.L.; O’Connell, R.T. *Time Series Forecasting*; Duxbury Press: New York, USA, 1979;

137. Brockwell, P.J.; Davis, R.A. *Time Series: Theory and Methods*; SpringerVerlag: New York, USA, 1987;

138. Jaksa, M.B.; Brooker, P.I.; Kaggwa, W.S. Inaccuracies associated with estimating random measurement errors. *J. Geotech. Geoenvironmental Eng.* **1997**.

139. Lumb, P. *Application of statistics in soil mechanics*; Lee, I.K., Ed.; New York, 1974;

140. Jaksa, M.B.; Yeong, K.S.; Wong, K.T.; Lee, S.L. Horizontal Spatial Variability of Elastic Modulus in Sand from the Dilatometer. *Proc. 9th Aust. New Zeal. Conf. Geomech.* **2004**, 289–294.

141. Pyrcz, M.; Deutsch, C. V. *Geostatistical reservoir modeling*; 2nd ed.; Oxford university press, 2014; ISBN 9780199731442.

142. Journel, A.G.; Huijbregts, C.J. *Mining geostatistics*; Academic press, 1978; ISBN

0123910501.

143. Kelkar, M.; Perez, G. *Applied geostatistics for reservoir characterization*; Society of Petroleum Engineers, 2002; ISBN 978-1555630959.

144. Jaksa, M.B. Assessing Soil Correlation Distances and Fractal Behavior. **2013**.

145. Marechal, A.; Serra, J. Random kriging. In *Geostatistics A Colloquium*; Springer, 1970; pp. 91–112.

146. McBratney, A.B.; Webster, R. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. *Eur. J. Soil Sci.* **1986**, *37*, 617–639.

147. Jaksa, M.B.; Brooker, P.I.; Kaggwa, W.S.; Holst Pellekaan, P.D.A.; Cathro, J.L. Modelling the Lateral Spatial Variation of the Undrained Shear Strength of a Stiff, Overconsolidated Clay Using an Horizontal Cone Penetration Test. **1994**, 34.

148. Jaksa, M.B.; Kaggwa, W.S.; Brooker, P.I. Experimental evaluation of the scale of fluctuation of a stiff clay. In Proceedings of the Proc. 8th Int. Conf. on the Application of Statistics and Probability; Sydney, AA Balkema, Rotterdam, 1999; Vol. 1, pp. 415–422.

149. Fenton, G.A. Random field modeling of CPT data. *J. Geotech. Geoenvironmental Eng.* **1999**, *125*, 486–498.

150. Nie, X.; Zhang, T.; Huang, H.; Liu, Z.; Lacasse, S. Scale of Fluctuation for Geotechnical Probabilistic Analysis. **2015**, 816–821.

151. Sasanian, S.; Soroush, A.; Chenari, R.J. Effect of Sampling Interval on the Scale of Fluctuation of CPT Profiles Representing Random Fields. *Int. J. Civ. Eng.* **2018**, 1–10.

152. Cressie, N. Fitting variogram models by weighted least squares. *J. Int. Assoc. Math. Geol.* **1985**, *17*, 563–586.

153. Cressie, N. *Statistics for Spatial Data*; Revised Ed.; John Wiley & Sons, Inc., 1993; ISBN 0-471-00255-0.

154. Box, G.E.P.; Jenkins, G.M. Time Series Analysis, Prediction and Control 1970.

155. Anderson, O.D. Time Series Analysis and Forecasting: The Box. *Jenkins*

Approach, London **1976**.

156. Jaksa, M.B.; Fenton, G.A. Discussion of "Random Field Modeling of CPT Data." *J. Geotech. Geoenvironmental Eng.* **2000**.

157. Lloret-Cabot, M.; Fenton, G.A.; Hicks, M.A. On the estimation of scale of fluctuation in geostatistics. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2014**, *8*, 129–140.

158. DeGroot, D.J. Analyzing spatial variability of insitu soil properties. *Uncertain. Geol. Environ. from theory to Pract. Proc. Uncertain. '96, Geotech. Spec. Publ.* **1996**, *58*, 210–238.

159. Orchant, C.J.; Kulhawy, F.H.; Trautmann, C.H. Critical Evaluation of In-Situ Test Methods and their Variability. Report EL-5507, Vol. 2. *EPRI , Palo Alto* **1988**.

160. EN ISO 22476-2 *Geotechnical Investigation and Testing–Field Testing–Part 2: Dynamic probing*; CEN: Brussels, Belgium, 2005;

161. EN ISO 22475-1 *Geotechnical investigation and testing-Sampling methods and groundwater measurements. Technical principles for execution*; CEN: Brussels, Belgium, 2006;

162. ASTM D1883 Standard Test Method for California Bearing Ratio (CBR) of Laboratory-Compacted Soils. *ASTM Int.* **2016**.

163. EN ISO/TS 17892-7 Geotechnical investigation and testing - Laboratory testing of soil - Part 7: Unconfined compression test on fine-grained soils. **2004**.

164. Fenton, G.A.; Griffiths, D. V Reliability-based deep foundation design. In *Probabilistic applications in geotechnical engineering*; 2007; pp. 1–12.

165. Pacheco, J.; de Brito, J.; Chastre, C.; Evangelista, L. Experimental investigation on the variability of the main mechanical properties of concrete produced with coarse recycled concrete aggregates. *Constr. Build. Mater.* **2019**, *201*, 110–120.

166. Aït-Mokhtar, A.; Belarbi, R.; Benboudjema, F.; Burlion, N.; Capra, B.; Carcasses, M.; Colliat, J.-B.; Cussigh, F.; Deby, F.; Jacquemot, F. Experimental investigation of the variability of concrete durability properties. *Cem. Concr. Res.* **2013**, *45*, 21–36.

167. Christou, G.; Tantele, E.A.; Votsis, R.A. Effect of environmental deterioration on buildings: a condition assessment case study. In Proceedings of the RSCy2014; International Society for Optics and Photonics, 2014; Vol. 9229, p. 92290Y.

168. Masi, A.; Chiauzzi, L. An experimental study on the within-member variability of in situ concrete strength in RC building structures. *Constr. Build. Mater.* **2013**, *47*, 951–961.

169. Wood, L.W. Variation of strength properties in woods used for structural purposes. **1960**.

170. Das, B.M. *Principle of Geotechnical Engineering*; 7th ed.; Cengage Learning: Stamford, CT, USA, 2010; ISBN 9780495668121.

171. Venkatramaiah, C. *Geotechnical engineering*; Revised 3r.; New Age International: New Delhi, India, 2006; ISBN 812240829X.

172. Rankine, W.J.M. II. On the stability of loose earth. *Philos. Trans. R. Soc. London* **1857**, 9–27.

173. Ni, P.; Mangalathu, S.; Song, L.; Mei, G.; Zhao, Y. Displacement-dependent lateral earth pressure models. *J. Eng. Mech.* **2018**, *144*, 4018032.

174. Smith, I.M.; Griffiths, D. V. *Programming the Finite Element Method*; 4th ed.; Wiley: Hoboken, NJ, USA, 2004; ISBN 0470011246.

175. Jaky, J. The coefficient of earth pressure at rest. *J. Soc. Hungarian Archit. Eng.* **1944**, 355–388.

176. Griffiths, D. V; Zhu, D.; Huang, J.; Fenton, G.A. Observations on probabilistic slope stability analysis. In Proceedings of the APSSRA 2016; Huang, J.L., Zhang, J., Chen, J., Eds.; Tongji University: Shanghai, China, 2016.

177. Vanmarcke, E.H. Reliability of earth slopes. *J. Geotech. Eng. Div.* **1977**, *103*, 1247–1265.

178. Soulie, M.; Montes, P.; Silvestri, V. Modelling spatial variability of soil parameters. *Can. Geotech. J.* **1990**, *27*, 617–630.

179. Cherubini, C. Data and considerations on the variability of geotechnical properties of soils. In Proceedings of the Proceedings of the international conference on safety

and reliability, ESREL; 1997; Vol. 97, pp. 1583–1591.

180. Popescu, R.; Prévost, J.H.; Deodatis, G. Effects of spatial variability on soil liquefaction: some design recommendations. *Geotechnique* **1997**, *47*, 1019–1036.

181. Phoon, K.; Kulhawy, F. Characterization of geotechnical variability. *Can. Geotech. J.* **1999**, *36*, 612–624.

182. Huang, J.; Griffiths, D. V Determining an appropriate finite element size for modelling the strength of undrained random soils. *Comput. Geotech.* **2015**, *69*, 506–513.

183. Rocscience Inc. RS2 Version 9.0 - Finite Element Analysis for Excavations and Slopes 2019.

184. Griffiths, D.V.; Fenton, G.A. Bearing capacity of heterogeneous soils by finite elements. In Proceedings of the 5th International Congress on Numerical Methods in Engineering and Scientific Applications; Troyani, N., Cerrolaza, M., Eds.; Sociedad Venezolana de Metodos Numericos en Ingenieria, 2000; pp. 27–37.