

JOINT TC205/TC304 WORKING
GROUP ON “DISCUSSION OF
STATISTICAL/RELIABILITY
METHODS FOR EUROCODES”

TC205 (Safety and serviceability in geotechnical design)
TC304 (Engineering practice of risk assessment and management)

International Society for Soil Mechanics and Geotechnical
Engineering (ISSMGE)

FINAL REPORT
September 2017

Table of Contents

Preface	Kok-Kwang Phoon & Brian Simpson
Geotechnical Information	
1. Transformation models & multivariate soil data	Jianye Ching <i>jyching@gmail.com</i>
2. Evaluation and consideration of model uncertainties in reliability based design	Kerstin Lesny <i>kerstin.lesny@hcu-hamburg.de</i>
Practical design methods	
3. Practical geotechnical design using reliability based design methods	Peter Day <i>day@jaws.co.za</i>
4. EXCEL-based direct reliability analysis and its potential role to complement Eurocodes	Bak-Kong Low <i>CBKLOW@ntu.edu.sg</i>
Bayesian methods	
5. Selection of characteristic values for rock and soil properties using Bayesian statistics and prior knowledge	Yu Wang <i>yuwang@cityu.edu.hk</i>
6. Bayesian method: a natural tool for processing geotechnical information	Jie Zhang <i>cezhangjie@gmail.com</i>
Advanced methods	
7. Incorporating spatial variability into geotechnical reliability based design	Dianqing Li <i>dianqing@whu.edu.cn</i>
8. Imprecise probabilistic and interval approaches applied to partial factor design	Sónia H. Marques <i>sonia.h@fe.up.pt</i>
Role of reliability in geotechnical practice	
9. Robustness in geotechnical design	Brian Simpson <i>brian.simpson@arup.com</i>
Future work	
10. Future directions and challenges	

Preface

A joint meeting between four ISSMGE technical committees was held on 14 Oct 2015 at the conference venue of the 5th International Symposium on Geotechnical Safety and Risk (ISGSR2015), 13-16 Oct 2015, Rotterdam, Netherlands. Participants came from TC304 (Engineering practice of risk assessment and management), TC205 (Safety and serviceability in geotechnical design), TC212 (Deep foundations), and TC 302 (Forensic geotechnical engineering) (Figure 1). The meeting was graced by the attendance of three TC chairs: KK Phoon (TC304 Chair) (Figure 2), Brian Simpson (TC205 Chair) (Figure 3), and G.L. Sivakumar Babu (TC302 Chair) (Figure 4). A report of the meeting was published in ISSMGE Bulletin Volume 10, Issue 1, February 2016, pp. 21-22.

A Joint TC205/TC304 Working Group on “Discussion of statistical/reliability methods for Eurocodes” was convened after the meeting to continue the conversation, particularly with regards to the following constructive goals in mind:

1. Identify the benefits and scope of applicability for reliability methods in geotechnical design
2. Identify gaps (e.g. unknown unknowns) and review existing literature for possible solutions.
3. Compile information on soil property and model uncertainties – this information is useful for design in both deterministic and reliability context.
4. Based on the above, can we refine and build on Annex D of ISO2394 to produce a useful annex for future codes such as the next revision of Eurocode 7? (e.g., further clarification of characteristic value? reliability calibrated partial factors? Bayesian updating to combine limited site data with global data from comparable sites?).

This Final Report was prepared in the spirit of open consultation and inclusive collaboration. An open call for participation in the Joint TC205/TC304 Working Group was made at the 21st TC304 meeting and the 25th TC205 meeting (both February 2016). Lead discussers who were interested to moderate a discussion topic were requested to submit a 1-page brief containing: (a) title of topic, (b) name of lead and other contributors, and (c) an abstract describing what the group hoped to achieve within a time frame of 3 months. The lead discussers invited group members through personal invitation and open invitation by posting their 1-page briefs on the TC304 website (March 2016). A preliminary draft containing around 10 pages was prepared by each group and made available to all groups and other interested participants (August 2016). Comments on these 10-page preliminary drafts were invited to fill in gaps in the coverage, to reduce overlap in the coverage of the topics between groups, and to voice other concerns between September and November 2016. The first draft of the full report was published in April 2017. A discussion session was organized in Georisk 2017/ISGSR 2017, June 4-6 2017, Denver, Colorado for participants to share ideas for finalizing the report. Public consultation took place between April and July 2017. This Final Report is published in September 2017 and made available to the ISSMGE community during the 19th International Conference on Soil Mechanics and Geotechnical Engineering (19ICSMGE), Sep 17-22 2017, Seoul, Korea. A Joint TC205/TC304 Workshop is planned for 20 September 2017 in conjunction with the 19ICSMGE to present the key findings in each chapter and to discuss future directions and challenges.

Given the scope and complex nature of this discussion spanning reliability theory, code drafting, design practice, risk management, and others, diverse perspectives are to be expected. Discussion leaders have taken onboard comments by revising the main body of the chapter or inviting the commentator to submit a “Discussion”. Separate discussion threads possibly reflecting alternate perspectives are appended at the end of each chapter in the form of “Discussion” and “Reply to discussion” (standard journal practice). In addition, discussion leaders were encouraged to read other chapters and add citations to other chapters/other works in the literature pointing to complementary findings or conflicting views to ensure both mainstream and alternate views were represented in a balanced and inclusive manner.

The Chairs of TC205 and TC304 are exceedingly grateful to the participants (list of members provided below) who have enthusiastically contributed to this joint working group and to the lead discussers who have invested significant time and efforts to moderate the discussions and to take the lead in drafting their chapters. Professor Malcolm Bolton graciously contributed a preliminary draft in the form of PPT on “Limits of reliability analysis”. However, the full chapter is not available. This Final Report is by no means a “finished” product – the discussion threads appended to the end of some chapters amply demonstrate a healthy exchange of views to improve our state of practice. As

discussed during the 2015 Rotterdam meeting, one advantage for adopting reliability analysis despite known limitations is to sharpen our appreciation of the gap between the calculated and observed probability of failure. The current “baseline” reliability approach that primarily focuses on dealing with known unknowns (uncertainties quantifiable statistically from data) as rationally as possible may point the way towards closing the gap in the future by updating reliability with more information such as monitoring (which is philosophically aligned to the observational approach) or stimulating research to get a better handle on unknown unknowns (e.g. human errors). Ultimately, no design approach can completely do away with engineering judgment given the challenging conditions that a geotechnical engineer routinely encounters in a project (spatially variable site with unique conditions that may change with time, scarce information, geologic surprises, etc.). Nonetheless, a more rational design approach, be it reliability or otherwise, arguably can improve our state of practice in a more directed way by combining with other risk management strategies, taking advantage of data analytics, and perhaps most importantly, focusing our engineering judgment on what it does best - setting up the right lines of scientific investigation, selecting the appropriate models and parameters for calculations, and verifying the reasonableness of the result. Directions for future work are presented in the last chapter, including the outcomes from the ISSMGE Global Survey conducted between 10 March and 30 April 2017 that are relevant to TC205 and TC304.

The exclusive rights to use and distribute each chapter belong to the authors. This Final Report can be downloaded from <http://140.112.12.21/issmge/tc304.htm> for use in education and research with permission from the authors.

Kok-Kwang Phoon (Chair, TC304) and Brian Simpson (Chair, TC205)
September 2017



Figure 1. Attendees of joint meeting between ISSMGE TC304, TC205, TC212 and TC 302 at the 5th International Symposium on Geotechnical Safety and Risk, 13-16 Oct 2015, Rotterdam, Netherlands



Figure 2. Prof Kok-Kwang Phoon (Chair of TC304) welcoming participants to the Rotterdam joint meeting.



Figure 3. Dr Brian Simpson (Chair of TC205) presenting his views on role of reliability analysis in design calculations.



Figure 4. Prof G.L. Sivakumar Babu (TC302 Chair) presenting his views on role of probabilistic design and analysis methods in forensic geotechnical engineering.

Acknowledgments

The contributions of the following TC205 and TC304 members that participated in the discussion and preparation of the chapters are gratefully acknowledged below. We would like to thank Dr Chong Tang for his editorial assistance in compiling this report.

Joint TC205/TC304 Working Group Members

Sami O. Akbas
Marcos Arroyo
Gregory Baecher
Richard Bathurst
Witold Bogusz
Malcolm Bolton
Philip Boothroyd
Sébastien Burlon
Michele Calvello
Zi-Jun Cao
Jianye Ching (Lead Discussor, Chapter 1)
Jieru Chen
Zuyu Chen
Satyanarayana Murty Dasaka
Peter Day (Lead Discussor, Chapter 3)
Mike Duncan
Malcolm Eddleston
Herbert Einstein
Robert Gilbert
D.V. Griffiths
Antonio Canavate Grimal
Hongwei Huang
Jinsong Huang
Papaioannou Iason
Mark Jaksa
Celeste Jorge
Charnghsein Juang
Karim Karam
Nico de Koker
Suzanne Lacasse
Tim Länsivaara
Kerstin Lesny (Lead Discussor, Chapter 2)
Dian-Qing Li (Lead Discussor, Chapter 7)
Zhe Luo
Bak Kong Low (Lead Discussor, Chapter 4)
Sónia H. Marques (Lead Discussor, Chapter 8)
Paul Mayne
Shadi Najjar
Shin-ichi Nishimura
Lars Olsson
Trevor Orr
Kok-Kwang Phoon
Widjojo Prakoso
Adrian Rodriguez-Marek
Hansruedi Schneider
Bernd Schuppener
Timo Schweckendiek
Brian Simpson (Lead Discussor, Chapter 9)
Daniel Straub

Armin Stuedlein
Chong Tang
Marco Uzielli
Giovanna Vessia
Paul Vardanega
Norbert Vogt
Yu Wang (Lead Discusser, Chapter 5)
Tien Wu
Jie Zhang (Lead Discusser, Chapter 6)
Limin Zhang

Chapter 1 Transformation Models and Multivariate Soil Databases

Lead discussor:

Jianye Ching (contributor of CLAY/10/7490 & SAND/7/2794)

jyching@gmail.com

Discussors (alphabetical order):

Marcos Arroyo, Jieru Chen (contributor of SAND/7/2794), Celeste Jorge, Tim Länsivaara (contributor of F-CLAY/7/216), Dianqing Li, Paul Mayne, Kok-Kwang Phoon, Widjojo Prakoso, Marco Uzielli

1.1 INTRODUCTION

Transformation models (Phoon and Kulhawy 1999) are valuable because they serve as “prior” information for correlation behaviors among various soil parameters. Useful compilations of these models are available in the literature (e.g., Djoenaidi 1985; Kulhawy and Mayne 1990; Mayne et al. 2001). For instance, it is common to estimate the friction angle (ϕ') of sand based on its SPT N value through a transformation model derived from data points obtained in the literature, as is showed in Figure 1-1. Here, the SPT N value is the site-specific information, and the friction angle ϕ' is the design soil parameter, also assumed as site-specific. However, the SPT N- ϕ' transformation model is not site-specific and is typically developed using a SPT N- ϕ' database collected from the literature. It is customary to adopt such a transformation model, which is not site-specific, in the process of estimating site-specific ϕ' . This process is illustrated in Figure 1-1. Suppose that the $(N_1)_{60}$ value (corrected N value) of a sand at certain depth at the design site is known to be 25. A vertical line is drawn at this $(N_1)_{60}$ value in Figure 1-1, and there are quite a few SPT N- ϕ' data points that are with similar $(N_1)_{60}$ values (circles). Although these data points are not site-specific, their ϕ' values may be meaningful. If the design site characteristics are within the coverage of the SPT N- ϕ' database, it is reasonable to think that the site-specific ϕ' value can be captured by the ensemble of these non-specific ϕ' values. By doing so, a single measurement of site-specific SPT N is converted into several “equivalent” ϕ' values that are viewed as posterior information for the actual site-specific ϕ' value.

Although this ensemble of “equivalent” ϕ' values (posterior information) may be a meaningful and realistic representation of the actual site-specific ϕ' value, there are occasion concerns expressed in the literature that the transformation models are applied too liberally in practice without careful consideration of their limitations. The purpose of this report is therefore to address the following practical questions:

1. What does the design soil parameter estimated from a transformation model really mean? The estimate can be a point estimate (e.g., the average of the equivalent ϕ' values) or an interval estimate (e.g., the range of the equivalent ϕ' values).
2. In what conditions will a transformation model produce meaningful estimates that are closely related to the actual site-specific design soil parameter?
3. In what conditions will a transformation model produce meaningless results that have very little to do with the actual site-specific design soil parameter?

This report will address the above questions through the “leave-one-out” design exercise based on a real soil database. The soil database is divided into two subsets: the first subset contains data points from a single site (design site), whereas the second subset contains the remaining sites in the database (training sites). The purpose of the leave-one-out exercise is to construct the transformation model based on the training sites, and then estimate the design soil parameter for the design site. The effectiveness of the transformation model can then be verified by comparing the estimation result and the actual value of the design soil parameter. To understand the effect of adopting a “general” soil database versus a “regional” soil database, two clay databases were investigated: one general database

for generic clays, and one regional database for Finland clays.

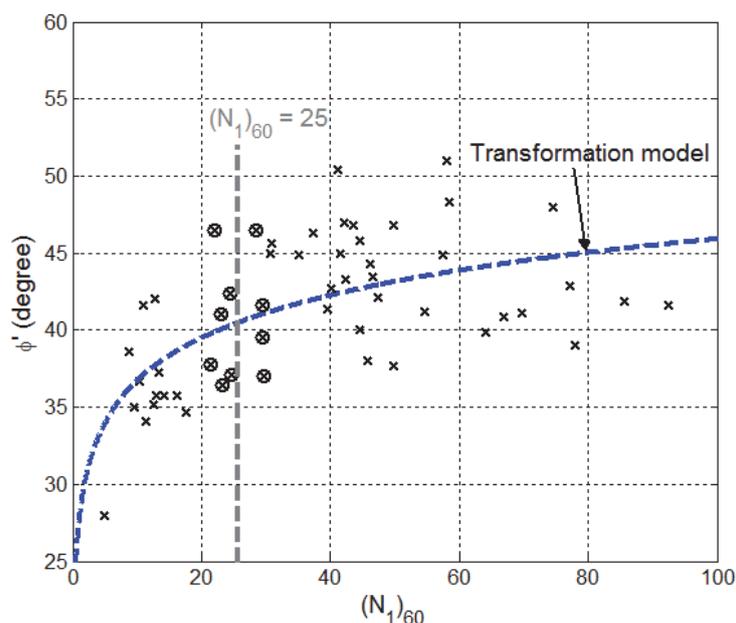


Figure 1-1 Transformation model between $(N_1)_{60}$ and ϕ' derived from data points in the literature.

1.2 MULTIVARIATE SOIL DATABASES

Two clay databases and one sand database, shown with light grey background in Table 1-1, were investigated. For completeness, three other clay databases available in the literature are also shown in the table. The databases are labelled as (soil type)/(number of parameters of interest)/(number of data points). The two clay databases, CLAY/10/7490 and F-CLAY/7/216, will be adopted to conduct the leave-one-out design exercise. CLAY/10/7490 is a general clay database, whereas F-CLAY/7/216 is a regional (Finland) clay database. SAND/7/2794 is a general sand database.

Table 1-1 Several multivariate soil databases.

Database	Reference	Parameters of interest	# data points	# sites/ studies	Range of properties		
					OCR	PI	S_t
CLAY/5/345	Ching and Phoon (2012a)	$LI, s_u, s_u^{re}, \sigma'_p, \sigma'_v$	345	37 sites	1~4		Sensitive to quick clays
CLAY/6/535	Ching et al. (2014)	$s_u/\sigma'_v, OCR, (q_t - \sigma_v)/\sigma'_v, (q_t - u_2)/\sigma'_v, (u_2 - u_0)/\sigma'_v, B_q$	535	40 sites	1~6	Low to very high plasticity	Insensitive to quick clays
CLAY/7/6310	Ching and Phoon (2013, 2015)	s_u from 7 different test procedures	6310	164 studies	1~10	Low to very high plasticity	Insensitive to quick clays
CLAY/10/7490	Ching and Phoon (2014)	$LL, PI, LI, \sigma'_v/P_a, S_t, B_q, \sigma'_p/P_a, s_u/\sigma'_v, (q_t - \sigma_v)/\sigma'_v, (q_t - u_2)/\sigma'_v$	7490	251 studies	1~10	Low to very high plasticity	Insensitive to quick clays
F-CLAY/7/216	D'Ignazio et al. (2016)	$s_u^{FV}, \sigma'_v, \sigma'_p, w_n, LL, PL, S_t$	216	24 sites	1~7.5	Low to very high plasticity	Insensitive to quick clays
SAND/7/2794	Ching et al. (2017)	$D_{50}, C_u, D_r, \sigma'_v/P_a, \phi', q_{t1}, (N_1)_{60}$	2794	176 studies	1~15	$D_{50} = 0.1 \sim 40$ mm $C_u = 1 \sim 1000+$ $D_r = -0.1 \sim 117\%$	

Note: LL = liquid limit; PL = plastic limit; PI = plasticity index; LI = liquidity index; w_n = natural water content; D_{50} = median grain size; C_u = coefficient of uniformity; D_r = relative density; σ'_v = vertical effective stress; σ'_p = preconsolidation stress; s_u = undrained shear strength; s_u^{FV} = undrained shear strength from field vane; s_u^{re} = remoulded s_u ; ϕ' = effective friction angle; S_t = sensitivity; OCR = overconsolidation ratio, $(q_t - \sigma'_v)/\sigma'_v$ = normalized cone tip resistance; $(q_t - u_2)/\sigma'_v$ = effective cone tip resistance; u_0 = hydrostatic pore pressure; $(u_2 - u_0)/\sigma'_v$ = normalized excess pore pressure; B_q = pore pressure ratio = $(u_2 - u_0)/(q_t - \sigma'_v)$; P_a = atmospheric pressure = 101.3 kPa; $q_{t1} = (q_t/P_a) \times C_N$ (C_N is the correction factor for overburden stress); $(N_1)_{60} = N_{60} \times C_N$ (N_{60} is the N value corrected for the energy ratio).

1.2.1 CLAY/10/7490

The CLAY/10/7490 database (Ching and Phoon 2014) is a general clay database consisting of data points from 251 studies. The geographical regions cover Australia, Austria, Brazil, Canada, China, England, Finland, France, Germany, Hong Kong, India, Iraq, Italy, Japan, Korea, Malaysia, Mexico, New Zealand, Norway, Northern Ireland, Poland, Singapore, South Africa, Spain, Sweden, Thailand, Taiwan, United Kingdom, United States, and Venezuela. The clay properties cover a wide range of overconsolidation ratio (OCR) (but mostly 1~10), a wide range of sensitivity (S_t) (sites with $S_t = 1$ ~tens or hundreds are fairly typical), and a wide range of plasticity index (PI) (but mostly 8 ~ 100). Ten dimensionless parameters of clays are of primary interest: liquid limit (LL), plasticity index (PI), liquidity index (LI), normalized vertical effective stress (σ'_v/P_a) (P_a is one atmosphere pressure = 101.3 kN/m²), normalized preconsolidation stress (σ'_p/P_a), normalized undrained shear strength (s_u/σ'_v) (s_u converted to the “mobilized” s_u defined by Mesri and Huvaj 2007), S_t , normalized piezocone tip resistance $(q_t - \sigma'_v)/\sigma'_v$, and normalized effective piezocone tip resistance $(q_t - u_2)/\sigma'_v$, and piezocone pore pressure ratio B_q . Some other dimensionless parameters of interest, such as s_u/σ'_p , OCR, and s_u^{re}/P_a , can be derived from the above 10 parameters. The basic statistics of all these parameters (10 basic parameters together with s_u/σ'_p , OCR, and s_u^{re}/P_a) are listed in Table 1-2.

Table 1-2 Statistics for the CLAY/10/7490 database (Table 3 in Ching and Phoon 2014)

Variable	n*	Mean	COV*	Min	Max
LL	3822	67.7	0.80	18.1	515
PI	4265	39.7	1.08	1.9	363
LI	3661	1.01	0.78	-0.75	6.45
σ'_v/P_a	3370	1.80	1.47	4.13E-3	38.74
σ'_p/P_a	2028	4.37	2.31	0.094	193.30
s_u/σ'_v	3538	0.51	1.25	3.68E-3	7.78
S_t	1589	35.0	2.88	1	1467
B_q	1016	0.58	0.35	0.01	1.17
$(q_t - \sigma'_v)/\sigma'_v$	862	8.90	1.17	0.48	95.98
$(q_t - u_2)/\sigma'_v$	668	5.34	1.37	0.61	108.20
s_u/σ'_p	1467	0.23	0.55	3.68E-3	1.34
OCR	3531	3.85	1.56	1.0	60.23
s_u^{re}/P_a	1143	0.075	2.86	9.67E-5	2.47

*n is the number of data points; COV stands for the coefficient of variation.

1.2.2 F-CLAY/7/216

The F-CLAY/7/216 database (D’Ignazio et al. 2016) is a regional clay database consisting of 216 field vane (FV) data points from 24 different test sites from Finland. Each data point contains genuine multivariate information on 7 clay parameters measured at comparable depths and sampling locations:

FV undrained strength (s_u^{FV}), vertical effective stress (σ'_v), preconsolidation stress (σ'_p), water content (w), liquid limit (LL), plastic limit (PL), and sensitivity (S_t). The clay properties cover wide ranges of sensitivity ($S_t = 2\sim 64$), plasticity ($PI = 2\sim 95$), overconsolidation ratio ($OCR = 1\sim 7.5$), and water content ($w = 25\sim 150$). To be consistent with Table 1-2, these parameters are converted to dimensionless parameters PI , LI , σ'_v/P_a , σ'_p/P_a , s_u/σ'_v , etc., in which $s_u = (\text{design value of } s_u) = s_u^{FV} \times (\text{correction factor proposed by Bjerrum 1972})$. The basic statistics for these dimensionless parameters are listed in Table 1-3.

Table 1-3 Statistics for the F-CLAY/7/216 database

Variable	n	Mean	COV	Min	Max
LL	216	66.3	0.30	22.0	125.0
PI	216	38.5	0.48	2.0	95.0
LI	216	1.44	0.46	0.42	4.80
σ'_v/P_a	216	0.46	0.48	0.074	1.61
σ'_p/P_a	216	0.79	0.50	0.20	2.27
s_u/σ'_v	216	0.40	0.74	0.11	2.71
S_t	216	17.4	0.79	2	64
s_u/σ'_p	216	0.22	0.31	0.058	0.52
OCR	216	1.84	0.51	1.0	7.5
s_u^{re}/P_a	216	0.016	0.99	0.0011	0.14

1.2.3 Comparison between CLAY/10/7490 and F-CLAY/7/216

The main difference between the two databases is that CLAY/10/7490 is a general database, whereas F-CLAY/7/216 is a regional database. A preliminary comparison between Tables 1-2 and 1-3 indicates the following distinct features between the general and regional databases:

1. The number of data points for the general database is significantly larger than that for the regional database.
2. The range spanned between the minimum and maximum values for the general database is significantly wider than that for the regional database. As a result, the COVs for the general database are significantly larger than those for the regional database.

Figure 1-2 shows the $LI-S_t$, $OCR-s_u/\sigma'_v$, $LI-\sigma'_p/P_a$, and $PI-s_u/\sigma'_p$ relationships for the two databases. It is clear that the coverage of the general database (CLAY/10/7490) is wider than the coverage of the regional database (F-CLAY/7/216). There are six data points from four Finland sites (annotated in Figure 1-2) with $LI > 3$, but with exceptionally low S_t . These six data points in F-CLAY/7/216 are not within the coverage of the general database CLAY/10/7490.

1.3 USE OF DATABASE IN ESTIMATING SITE-SPECIFIC DESIGN SOIL PARAMETER

In the design process, a soil database can be adopted to develop a transformation model that can be further used to estimate the design soil parameter (e.g., s_u) for the design site based on site-specific information. For instance, based on the site-specific OCR information of a clay at a design site, its s_u/σ'_v value can be estimated from an $OCR-s_u/\sigma'_v$ transformation model developed from a clay database. Note that the clay database and the resulting transformation model are not site-specific. A question may arise: is the resulting s_u/σ'_v estimate site-specific or not? This question can be answered by comparing the s_u/σ'_v estimate with the actual site-specific s_u/σ'_v value. If the s_u/σ'_v estimate can capture the actual site-specific s_u/σ'_v value, the s_u/σ'_v estimate is site-specific. Otherwise, it is not. This comparison can be realized by the leave-one-out design exercise. The details for this leave-one-out exercise will be presented later. Consider two scenarios:

1. Scenario 1: A regional database, such as F-CLAY/7/216, is adopted to develop the $OCR-s_u/\sigma'_v$ transformation model.
2. Scenario 2: A regional database is not available. A general database, such as CLAY/10/7490, is used to develop the $OCR-s_u/\sigma'_v$ transformation model.

For both scenarios, the question whether the resulting s_u/σ'_v estimate is site-specific or not will be addressed. The effect for adopting a regional database against a general database will be also illustrated.

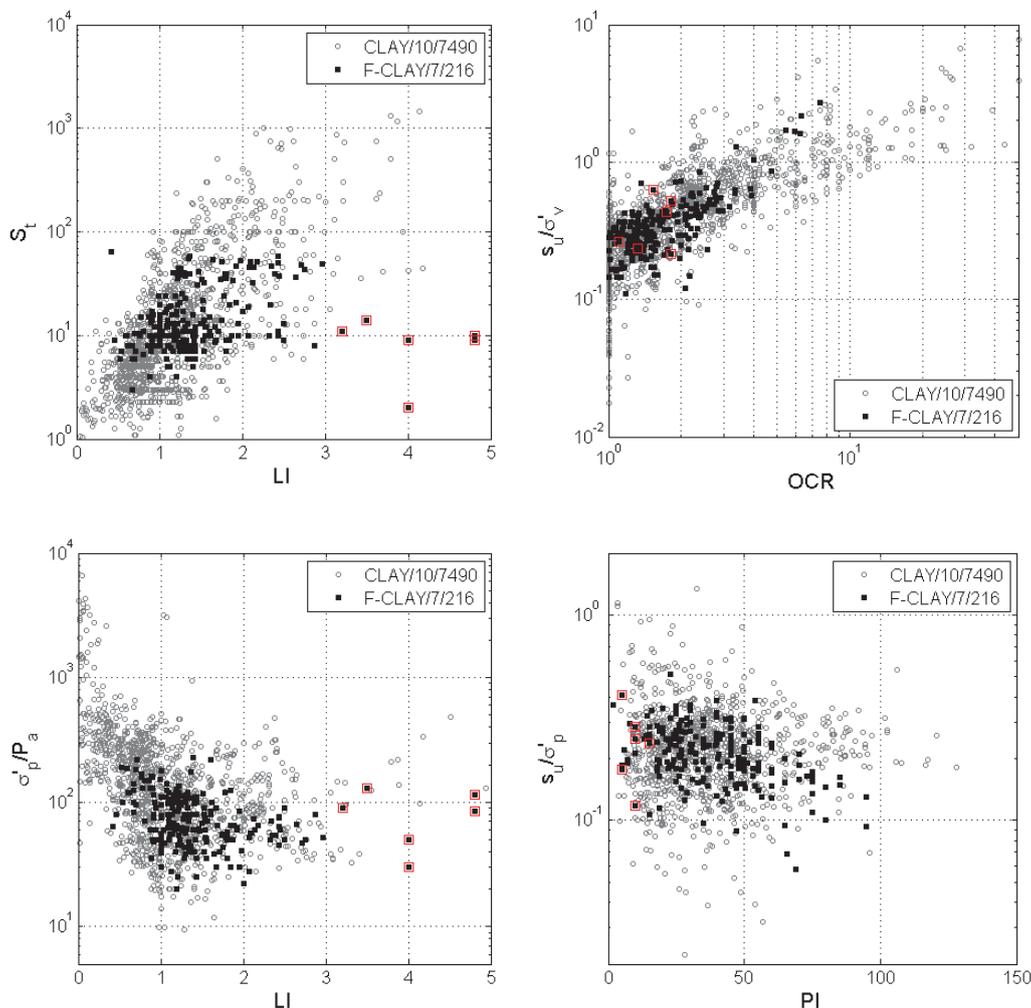


Figure 1-2 LI- s_t , OCR- s_u/σ'_v , LI- σ'_p/P_a , and PI- s_u/σ'_p relationships for the two databases.

1.3.1 Scenario 1

Let us consider a new Finland site, and suppose the design engineer has the regional database F-CLAY/7/216. Consider a clay at that site with a known site-specific OCR, denoted by OCR_{new} . The goal is to estimate its site-specific s_u/σ'_v , denoted by $(s_u/\sigma'_v)_{new}$. The estimate can be either a point estimate or an interval estimate. The design engineer can adopt the $OCR-s_u/\sigma'_v$ data points in the database to develop the following transformation model:

$$\ln(s_u/\sigma'_v) = a + b \times \ln(OCR) + \varepsilon \quad (1-1.)$$

where (a, b) are unknown coefficients to be estimated; ε is the transformation error, modeled as a zero-mean normal random variable with standard deviation σ_ε . (a, b) can be estimated by least squares:

$$\begin{aligned}
 b^* &= S_{xy}/S_{xx} & a^* &= y_m - b^* \times x_m \\
 x_m &= \frac{1}{n} \sum_{i=1}^n \log(\text{OCR}_i) & y_m &= \frac{1}{n} \sum_{i=1}^n \log[(s_u/\sigma'_v)_i] \\
 S_{xx} &= \sum_{i=1}^n [\log(\text{OCR}_i) - x_m]^2 & S_{xy} &= \sum_{i=1}^n [\log(\text{OCR}_i) - x_m] \times (\log[(s_u/\sigma'_v)_i] - y_m)
 \end{aligned} \tag{1-2}$$

where (a^*, b^*) denote the least square estimates for (a, b) ; OCR_i and $(s_u/\sigma'_v)_i$ denote the OCR and (s_u/σ'_v) values of the i -th data point in the clay database; n is the number of data points in the database. σ_ε can be estimated as well:

$$\sigma_\varepsilon^* = \sqrt{\frac{1}{n-2} \sum_{i=1}^n \left\{ \log[(s_u/\sigma'_v)_i] - a^* - b^* \times \log(\text{OCR}_i) \right\}^2} \tag{1-3}$$

The design engineer can then obtain two useful estimates for $\ln[(s_u/\sigma'_v)_{\text{new}}]$: (a) the point estimate $a^* + b^* \times \ln(\text{OCR}_{\text{new}})$; and (b) the 95% confidence interval (CI) estimate defined by:

$$a^* + b^* \times \ln(\text{OCR}_{\text{new}}) \pm t_{0.975} \times \sigma_\varepsilon^* \times \sqrt{1 + \frac{1}{n} + \frac{[\ln(\text{OCR}_{\text{new}}) - x_m]^2}{S_{xx}}} \tag{1-4}$$

where $t_{0.975}$ is the 97.5% percentile for the Student-t distribution with $(n-2)$ degrees of freedom. In the following illustration, we will focus on the 95% confidence interval (CI) estimate. The 95% CI in Eq. (1-4) is a “nominal” 95% CI. It is unclear whether it is a genuine 95% CI. Namely, it is unclear whether the actual chance for $\ln[(s_u/\sigma'_v)_{\text{new}}]$ to be within the interval is indeed close to 95%.

To illustrate that the nominal 95% CI is genuine, consider the following “leave-one-out” design exercise. There are 24 sites in the F-CLAY/7/216 database. Each time, one site is treated as the new design site, whereas the remaining 23 sites are treated as the training sites. Note that the new design site and the training sites belong to the same “population”: they are all Finland sites. However, the design site is independent of the 23 training sites. There may be several clay data points in the design site. For each clay at the design site, its OCR_{new} is considered known, e.g., an oedometer test is conducted to determine its OCR. However, we “pretend” its $(s_u/\sigma'_v)_{\text{new}}$ to be unknown. First, $(a^*, b^*, \sigma_\varepsilon^*)$ are estimated based on the 23 training sites using Eqs. (1-2) and (1-3). The point estimate for $\ln[(s_u/\sigma'_v)_{\text{new}}]$ is $a^* + b^* \times \ln(\text{OCR}_{\text{new}})$ and the nominal 95% CI for $\ln[(s_u/\sigma'_v)_{\text{new}}]$ is obtained using Eq. (1-4). Because $\ln[(s_u/\sigma'_v)_{\text{new}}]$ for the clay is actually known, we can compute the prediction error $e = \ln[(s_u/\sigma'_v)_{\text{new}}] - a^* - b^* \times \ln(\text{OCR}_{\text{new}})$ and also determine whether $\ln[(s_u/\sigma'_v)_{\text{new}}]$ is within the nominal 95% CI. This leave-one-out exercise is repeated for all 216 data points in F-CLAY/7/216. Figure 1-3 shows the histogram of the 216 prediction errors. The prediction errors have a mean value that is roughly zero. Among the 216 leave-one-out trials, $\ln[(s_u/\sigma'_v)_{\text{new}}]$ is within the nominal 95% CI for 202 times. This means that the CI is in effect a $202/216 = 93.5\%$ CI: it is reasonably close to a genuine 95% CI. The difference between 93.5% and 95% may be partially due to the statistical uncertainty.

The above leave-one-out design exercise shows that the nominal 95% CI developed by the 23 training sites is close to genuine. This is probably because the design site and the 23 training sites belong to the same population, e.g., Finland sites. When this happens (same population), the nominal 95% CI is theoretically the genuine 95% CI. This conclusion should not change if all 24 sites in F-CLAY/7/216 are adopted to develop the transformation model and goal is to estimate the $(s_u/\sigma'_v)_{\text{new}}$ for a 25th site that is not in F-CLAY/7/216. This justifies the use of a transformation model in the case where the design site belongs to the same population.

There is a caveat here: the training soil database needs to have a sufficient coverage to represent the population. In the above leave-one-out design exercise, there are $24 - 1 = 23$ training sites. If the number of sites is small, the training sites can no longer represent the Finland population, and the nominal 95% CI can cease to be genuine, so the actual chance for $\ln[(s_u/\sigma'_v)_{\text{new}}]$ to be within the

nominal 95% CI will not be close to 95%. Figure 1-4 shows how this actual chance varies with respect to the number of training sites (see the line with legend “Scenario 1”). Consider a subset database with $(n_t + 1)$ sites randomly sampled from the 24 sites in F-CLAY/7/216. The leave-one-out design exercise is conducted on the subset database with n_t training sites and one design site, and the chance for $\ln[(s_u/\sigma_v)_{new}]$ to be within the 95% CI can be evaluated. This actual chance is itself random because it depends on the random sampling effect of the $(n_t + 1)$ sites. Therefore, the subset database with $(n_t + 1)$ sites is randomly sampled for 100 times to obtain 100 samples for the actual chance. The horizontal axis in Figure 1-4 is the number of training sites (n_t) and the vertical axis is the average of the 100 samples for the actual chance. The (averaged) actual chance seems to converge to 95% with increasing number of sites. The actual chance is significantly less than 95% if the database contains less than 4 Finland sites, whereas the actual chance is close to 95% if the database contains more than 10 Finland sites. This means that for the Finland case, a regional database with more than 10 sites should have a sufficient coverage.

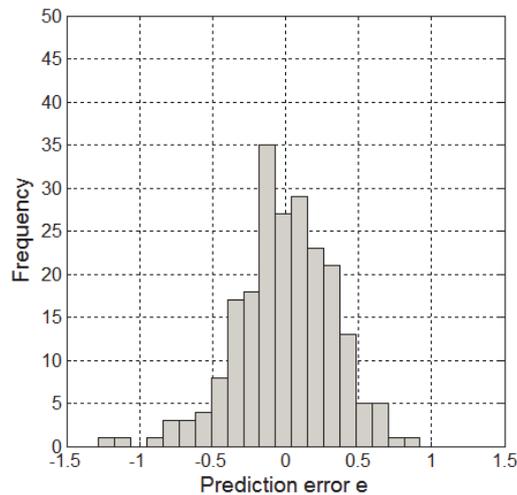


Figure 1-3 Histogram of the prediction error e.

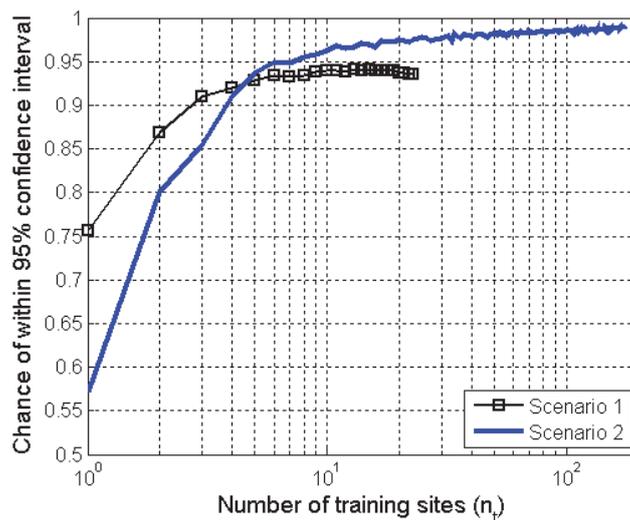


Figure 1-4 Chance for $\ln[(s_u/\sigma_v)_{new}]$ to be within the nominal 95% CI.

1.3.2 Scenario 2

Suppose that the new design site is a Finland site, but a Finland database is not available. Yet, suppose the design engineer has the general database CLAY/10/7490. Note here that now the design site and

the training sites do not belong to the same population: the training sites obviously have a wider coverage because they are global sites. The OCR- s_u/σ'_v data points in CLAY/10/7490 are from 179 global sites from Americas, Europe, Asia, etc. The design engineer can still adopt the OCR- s_u/σ'_v transformation model developed from CLAY/10/7490 to obtain the nominal 95% CI, but is it still a genuine 95% CI with respect to the Finland design site?

To understand the significance of the nominal 95% CI obtained from a general database, the following design exercise is taken. It is not necessary to do leave-one-out, because the design site is not within the general database. First, (a^* , b^* , σ_ε^*) are estimated using Eqs. (1-2.) and (1-3.) based on the general database. Each clay data point in F-CLAY/7/216 is a Finland design case. Its OCR value is treated as known (denoted by OCR_{new}), whereas we pretend its s_u/σ'_v value to be unknown [denoted by $(s_u/\sigma'_v)_{new}$]. The nominal 95% CI for this $\ln[(s_u/\sigma'_v)_{new}]$ can be obtained using Eq. (1-4.) based on (a^* , b^* , σ_ε^*) and OCR_{new}. Nonetheless, $\ln[(s_u/\sigma'_v)_{new}]$ is actually known, and we can determine whether $\ln[(s_u/\sigma'_v)_{new}]$ is within the nominal 95% CI. This exercise is repeated for all 216 data points in F-CLAY/7/216. It turns out that the actual chance for $\ln[(s_u/\sigma'_v)_{new}]$ of a Finland clay to be within the nominal 95% CI is 99.1%, significantly larger than 95%. This is probably because the design site and the 179 training sites belong to different populations. When this happens (different populations), there is no guarantee that the nominal 95% CI is genuine.

There are 179 sites in the general database CLAY/10/7490. It is interesting to know that the actual chance will change if there is a different number of training sites. Figure 1-4 shows how the actual chance varies with respect to the number of training sites (n_t) (see the line with legend “Scenario 2”). Again, the actual chance is random due to the random sampling effect of the n_t sites. Therefore, the subset database with n_t sites is randomly sampled from the 179 sites for 100 times to obtain 100 samples for the actual chance. The actual chance seems to converge to 100% with increasing number of training sites, rather than converge to 95%. The 95% CI developed from a general database with many sites is “conservative” with respect to a Finland site, in the sense that the actual chance for $\ln[(s_u/\sigma'_v)_{new}]$ of a Finland site to be within the 95% CI is more than 95%. The nominal 95% CI is wider than the genuine 95% CI, because CLAY/10/7490 has a wider coverage than the Finland database F-CLAY/7/216. This wider coverage can be clearly seen in Figure 1-2. Nonetheless, the actual chance can be significantly less than 95% if the general database contains less than 4 sites.

1.3.3 Scenario “A”

Let us consider a rather academic scenario: the new design site belongs to the “general population” containing all global sites. We call this scenario “Scenario A”. The purpose of this scenario is to further verify the significance of the nominal 95% CI developed from the general database. Let a clay at the design site have a known site-specific OCR, denoted by OCR_{new}. The goal is to estimate its unknown s_u/σ'_v , denoted by $(s_u/\sigma'_v)_{new}$. The nominal 95% CI for $\ln[(s_u/\sigma'_v)_{new}]$ is constructed by the general database CLAY/10/7490. Note that the sites in the general database CLAY/10/7490 also belong to the general population. Therefore, the design site and the training sites belong to the same general population. Is this nominal 95% CI the genuine 95% CI with respect to the new design site?

The following leave-one-out design exercise is adopted to illustrate the significance of the nominal 95% CI. The OCR- s_u/σ'_v data points in the CLAY/10/7490 database are from 179 sites. Each time, one site is treated as the design site, whereas the remaining 178 sites are treated as the training sites. First, (a^* , b^* , σ_ε^*) are estimated based on the 178 training sites using Eqs. (1-2.) and (1-3.), and the nominal 95% CI for $\ln[(s_u/\sigma'_v)_{new}]$ is obtained using Eq. (1-4.). Nonetheless, the $\ln[(s_u/\sigma'_v)_{new}]$ for the design site is actually known so that we can determine the chance for $\ln[(s_u/\sigma'_v)_{new}]$ to be within the nominal 95% CI. The actual chance for $\ln[(s_u/\sigma'_v)_{new}]$ to be within the nominal 95% CI is about 94.4%, close to 95%. Again, the caveat is that there are a sufficient number of sites in the database. Figure 1-5 shows how the actual chance varies with respect to the number of training sites (n_t). The (averaged) actual chance seems to converge to 95% with increasing number of sites. The actual chance is close to 95% if the number of sites is more than 50.

1.3.4 Discussions

The key questions that this study aims to address are:

1. What does the design soil parameter estimate from a transformation model really mean?

2. In what conditions will a transformation model produce meaningful estimates that are closely related to the actual site-specific design soil parameter?

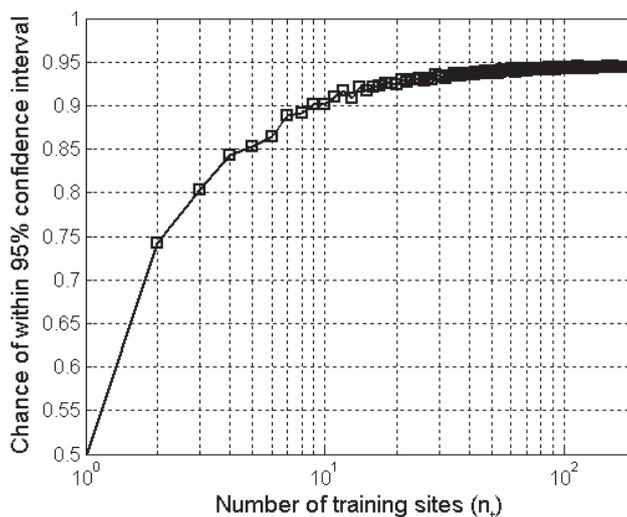


Figure 1-5 Chance for $\ln[(s_u/\sigma'_v)_{new}]$ to be within the nominal 95% CI (Scenario A).

Based on the above results, it can be concluded that the nominal 95% CI produced by the transformation model is meaningful because it has a large chance to include the actual site-specific $\ln[(s_u/\sigma'_v)_{new}]$. Moreover, the nominal 95% CI is close to genuine if the design site and the training sites belong to the same population. The following two scenarios exemplify the concept of “same population”:

1. The design site is a Finland site, whereas the soil database is a Finland (regional) soil database that has a sufficient coverage. This is Scenario 1. For the Finland case, 10 sites in the regional database seem sufficient.
2. The design site belongs to the general population, whereas the soil database is a general database with a sufficient coverage. This is Scenario A. In the above illustration, 50-100 sites in the general database seem sufficient.

Although the nominal 95% CI provides a satisfactory estimate for the site-specific $\ln[(s_u/\sigma'_v)_{new}]$, it is an interval estimate, not a point estimate. It is possible to obtain the point estimate, i.e., the point estimate = $a^* + b^* \times \ln(\text{OCR}_{new})$, but certain inaccuracy is to be expected (see the prediction error in Figure 1-3).

3. In what conditions will a transformation model produce meaningless results that have little to do with the actual site-specific design soil parameter?

If the design site and the training sites do not belong to the same population, there is no guarantee that the nominal 95% CI derived from the training sites is genuine. If the design site belongs to the Finland population but the training sites are general with a sufficient number of sites, the nominal 95% confidence interval derived from the general database will be wider than the genuine 95% CI. When this happens, the nominal 95% CI is still meaningful, but it is conservative (it has a very large chance to include the actual $\ln[(s_u/\sigma'_v)_{new}]$) and less effective.

The nominal 95% CI may become completely meaningless if the design site and training sites belong to two populations occupying completely different regions in the $\text{OCR}-(s_u/\sigma'_v)$ space. For instance, the design site contains fissured clays, whereas the training sites only contain non-fissured clays.

Appendix 1A (Transformation Models Calibrated by Soil Databases) shows some transformation models calibrated by the F-CLAY/7/216 regional database and by the CLAY/10/7490 general database. These transformation models were originally developed in the literature, but their biases and variabilities are calibrated by the soil databases. Given the site-specific investigation information of a new design site, the point estimate and nominal 95% CI can be obtained from these transformation models (details given in Appendix 1A). The 95% CI estimate is meaningful in the sense that the actual design soil parameter will have a large chance to be within the confidence interval. Appendix 1A also

shows some transformation models for sands as well as their biases and variabilities calibrated by the SAND/7/2794 general database.

1.3.5 Other transformation models

For other transformation models, the qualitative conclusions obtained above remain unchanged. Consider the LI- S_t transformation model. Figure 1-6 shows how the actual chance for $\ln[(S_t)_{new}]$ to be within the nominal 95% CI varies with the number of sites in the database. The left plot is for Scenarios 1 and 2, whereas the right plot is for Scenario A. Those plots are qualitatively similar to Figures 1-4 and 1-5.

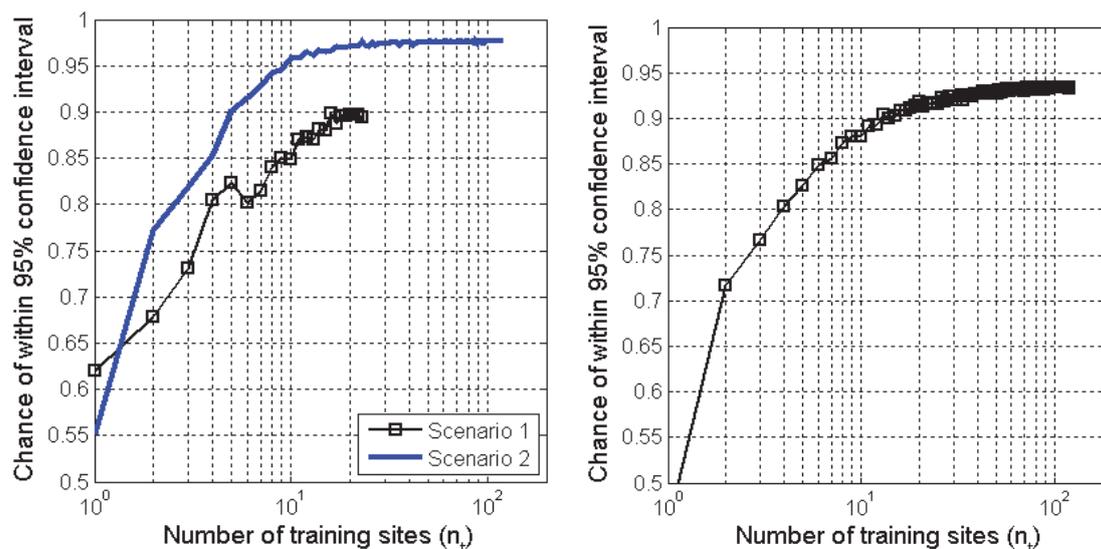


Figure 1-6 Chance for $\ln[(S_t)_{new}]$ to be within the nominal 95% CI: (left) Scenarios 1 & 2; (right) Scenario A.

1.3.6 Multivariate correlations

We have illustrated how $\ln[(s_u/\sigma'_v)_{new}]$ can be estimated based on the site-specific OCR_{new} . It was shown that the nominal 95% CI developed from a soil database can be useful and meaningful. However, it can happen that the resulting 95% CI is very wide so that $\ln[(s_u/\sigma'_v)_{new}]$ is still very uncertain. Multivariate information is usually available in a typical site investigation. For instance, when undisturbed samples are extracted for oedometer tests to determine OCR, piezocone test (CPTU) may be conducted in close proximity. These multiple data sources are typically correlated to the design soil parameter, e.g., the undrained shear strength (s_u). Figure 1-7 shows the data points for the two transformations in the CLAY/10/7490 database. It is clear that both OCR and $(q_t - \sigma'_v)/\sigma'_v$ are positively correlated to s_u/σ'_v . These multiple correlations can be exploited to reduce the uncertainty in the design soil parameter. In the previous sections, we have illustrated a framework where the site-specific OCR information can be used to obtain the 95% CI for (s_u/σ'_v) . This univariate framework is extended to account for multivariate framework, e.g., both OCR and $(q_t - \sigma'_v)/\sigma'_v$ are known, in the following.

Suppose that a design engineer has a multivariate OCR- $[(q_t - \sigma'_v)/\sigma'_v]$ - (s_u/σ'_v) database. For each data point, OCR, $(q_t - \sigma'_v)/\sigma'_v$, and s_u/σ'_v are simultaneously known. The engineer can adopt the data points in the database to develop the following multivariate transformation model:

$$\ln(s_u/\sigma'_v) = a + b \times \ln(OCR) + c \times \ln[(q_t - \sigma'_v)/\sigma'_v] + \varepsilon \quad (1-5.)$$

where (a, b, c) are unknown coefficients to be estimated; ε is the transformation error, modeled as a zero-mean normal random variable with standard deviation σ_ε . (a, b, c) can be estimated by least

squares:

$$\begin{bmatrix} a^* \\ b^* \\ c^* \end{bmatrix} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}$$

$$\mathbf{A} = \begin{bmatrix} 1 & \log(\text{OCR}_1) & \log\left(\frac{q_t - \sigma_v}{\sigma'_v}\right)_1 \\ 1 & \log(\text{OCR}_2) & \log\left(\frac{q_t - \sigma_v}{\sigma'_v}\right)_2 \\ \vdots & \vdots & \vdots \\ 1 & \log(\text{OCR}_n) & \log\left(\frac{q_t - \sigma_v}{\sigma'_v}\right)_n \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} \log\left(\frac{s_u}{\sigma'_v}\right)_1 \\ \log\left(\frac{s_u}{\sigma'_v}\right)_2 \\ \vdots \\ \log\left(\frac{s_u}{\sigma'_v}\right)_n \end{bmatrix} \quad (1-6.)$$

σ_ε can be estimated as well:

$$\sigma_\varepsilon^* = \sqrt{\frac{1}{n-3} \sum_{i=1}^n \left\{ \log\left(\frac{s_u}{\sigma'_v}\right)_i - a^* - b^* \times \log(\text{OCR}_i) - c^* \times \log\left(\frac{q_t - \sigma_v}{\sigma'_v}\right)_i \right\}^2} \quad (1-7.)$$

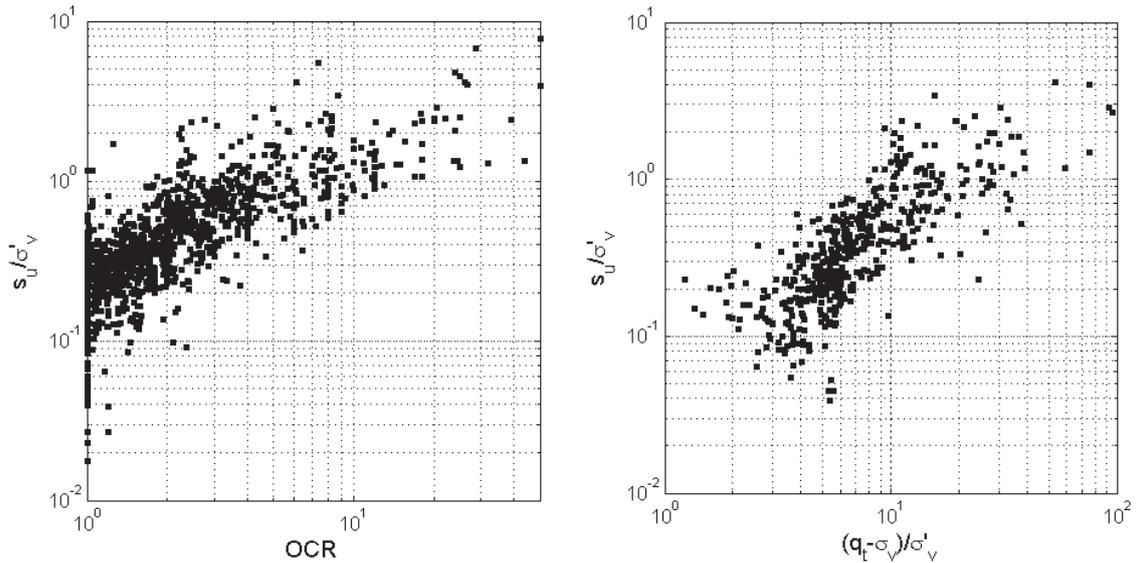


Figure 1-7 OCR- (s_u/σ'_v) and $[(q_t - \sigma_v)/\sigma'_v]$ - (s_u/σ'_v) data points in the CLAY/10/7490 database.

Based on the OCR- $[(q_t - \sigma_v)/\sigma'_v]$ - (s_u/σ'_v) data points in CLAY/10/7490, the estimated σ_ε^* is equal to 0.46. With the OCR- (s_u/σ'_v) information from the same data points, the estimated σ_ε^* for the univariate OCR- (s_u/σ'_v) transformation model in Eq. (1-1) is equal to 0.51. This shows that the transformation uncertainty in the multivariate model (Eq. 1-5) is less than that in the univariate model (Eq. 1). The resulting 95% CI for $\ln[(s_u/\sigma'_v)_{\text{new}}]$ from the multivariate model (to be presented below) will be also narrower than that from the univariate model.

Now consider a new design site with known site-specific OCR and $(q_t - \sigma_v)/\sigma'_v$, denoted by OCR_{new} and $[(q_t - \sigma_v)/\sigma'_v]_{\text{new}}$. The goal is to estimate its site-specific s_u/σ'_v , denoted by $(s_u/\sigma'_v)_{\text{new}}$. Based on $(a^*, b^*, c^*, \sigma_\varepsilon^*)$, the design engineer can obtain two useful estimates for $\ln[(s_u/\sigma'_v)_{\text{new}}]$: (a) the point estimate $a^* + b^* \times \ln(\text{OCR}_{\text{new}}) + c^* \times \ln([(q_t - \sigma_v)/\sigma'_v]_{\text{new}})$ and (b) the nominal 95% CI estimate defined by:

$$a^* + b^* \times \ln(\text{OCR}_{\text{new}}) + c^* \times \log\left(\left[\frac{(q_t - \sigma_v)}{\sigma'_v}\right]_{\text{new}}\right) \pm t_{0.975} \times \sigma_e^* \times \sqrt{1 + \mathbf{A}_{\text{new}}^T (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}_{\text{new}}} \quad (1-8.)$$

where $t_{0.975}$ is the 97.5% percentile for the Student-t distribution with $(n-3)$ degrees of freedom; $\mathbf{A}_{\text{new}} = [1 \ln(\text{OCR}_{\text{new}}) \ln\left(\left[\frac{(q_t - \sigma_v)}{\sigma'_v}\right]_{\text{new}}\right)]^T$.

The following leave-one-out exercise based on CLAY/10/7490 is adopted to verify whether the nominal 95% CI is genuine. There are 50 sites in the CLAY/10/7490 database containing 417 multivariate OCR- $[(q_t - \sigma'_v)/\sigma'_v] - (s_u/\sigma'_v)$ data points. Each time, one site is treated as the design site, whereas the remaining 49 sites are treated as the training sites that are further used to obtain the nominal 95% CI. This is similar to Scenario A above. For the leave-one-out exercise, $\ln[(s_u/\sigma'_v)_{\text{new}}]$ is actually known. Therefore, we can determine whether $\ln[(s_u/\sigma'_v)_{\text{new}}]$ is within the nominal 95% CI. Among the 417 leave-one-out trials, $\ln[(s_u/\sigma'_v)_{\text{new}}]$ is within the nominal 95% CI for 381 times. This means that the CI is in effect a $381/417 = 91.4\%$ CI. The difference between 91.4% and 95% may be partially due to the statistical uncertainty. It is also possible that 50 sites are not yet sufficient for the convergence. Figure 1-8 shows how the actual chance for $\ln[(s_u/\sigma'_v)_{\text{new}}]$ to be within the nominal 95% CI varies with respect to the number of sites in the database. The convergence behavior in this figure is similar to those in Figures 1-4 to 1-6. It is possible that the qualitative conclusions obtained for the univariate framework above still apply to the multivariate framework.

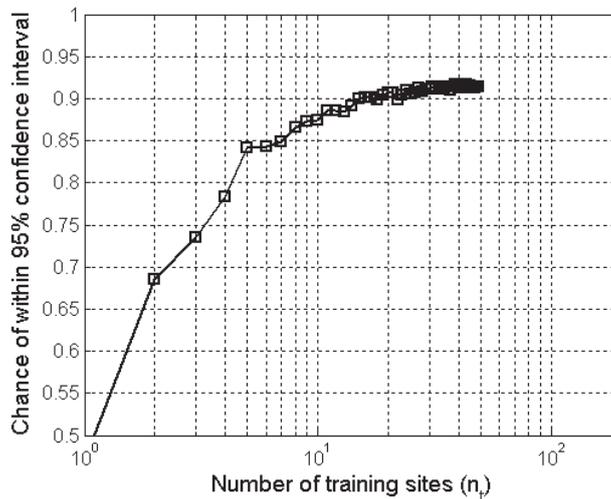


Figure 1-8 Chance for $\ln[(s_u/\sigma'_v)_{\text{new}}]$ to be within the nominal 95% CI (multivariate scenario).

1.4 CONCLUSIONS

A transformation model is frequently used to estimate the design soil parameter. However, it is not clear what the estimated design soil parameter really means. A possible concern for such a soil parameter estimate is that a transformation model is constructed by non-site-specific data points. Can these non-site-specific data points be used to derive any meaningful site-specific estimate? The purpose of this report is to address this question and to verify the significance of this design soil parameter estimate based on the so-called “leave-one-out” design exercise.

The leave-one-out exercise emulates the process of estimating the design soil parameter: the design soil parameter for a “design site” is estimated based on the transformation model constructed by a set of “training sites”. The design site is not within the training sites. Basically, a large soil database with N sites is divided into two subsets: one subset only contains the design site, and the other subset contains $N-1$ training sites. A transformation model is first calibrated by the training sites, then it is adopted to estimate the design soil parameter for the design site. This process is repeated for all data points in the soil database. Because the actual value of the design soil parameter for the design site is in fact known, the performance and significance of the design soil parameter estimate obtained from the transformation model can be verified. In this report, we focus on the 95% confidence interval (CI) estimate obtained from the transformation model. This 95% CI may or may not be the genuine

95% CI, so it is called, in this report, the “nominal” 95% CI.

The results show that the nominal 95% CI estimate obtained from the transformation model is meaningful, albeit the transformation model is derived from non-site-specific data points. The concept of “population” in statistics is central to our conclusions. It is concluded that as long as the design site and training sites belong to the same population, the nominal 95% CI estimate obtained from the transformation model is close to a genuine 95% CI, meaning that the chance for the actual design parameter to be within the nominal 95% CI is close to 95%. A radical view is that only the data points at the design site (site-specific data points) can be used to derive the design soil parameter and that all non-site-specific data points are irrelevant. Nonetheless, the findings in this report suggest that this view may be incorrect. In fact, non-site-specific data points can be still useful if they are in the same “population” for the design site. This means that if the design site is a Finland site, the transformation model developed by Finland training sites (i.e., a Finland database) can be useful and meaningful in the sense that the resulting nominal 95% CI is close to a genuine 95% CI.

A more controversial scenario is that the design site and training sites do not belong to the same population, e.g., the design site is a Finland site, yet the training sites are general (global) sites. In the case that the design site population is a subset of the training site population (e.g., the Finland population is a subset of the general population), the results in this report suggest that the resulting nominal 95% CI is no longer a genuine 95% CI. Moreover, the nominal 95% CI is wider than the genuine 95% CI. In one previous illustration in this report (Scenario 2), the chance for the actual design parameter to be within the nominal 95% CI is close to 95% is 99.1%. Yet, this does not suggest that the nominal 95% CI is completely useless and meaningless. Instead, this only suggests that the nominal 95% CI is less effective and more conservative.

An even worse scenario is that the design site and training sites not only belong to different populations but also the design site population is not a subset of the training site population. For instance, the design site is with fissured clays, yet the training sites do not contain fissured clays. When this occurs, the resulting nominal 95% CI can become useless and meaningless.

Appendix 1A shows some transformation models calibrated by some soil databases. The guideline for deriving the point estimate and 95% CI estimate is also provided.

1.5 REFERENCES

- Bjerrum, L. (1954). Geotechnical properties of Norwegian marine clays. *Géotechnique*, 4(2), 49-69.
- Bjerrum, L. (1972). Embankments on soft ground. Proc. Specialty Conference on Performance of Earth and Earth-Supported Structures, ASCE, Purdue University, Lafayette, USA, 2,1-54.
- Bolton, M.D. (1986). The strength and dilatancy of sands. *Géotechnique*, 36(1), 65-78.
- Chen, B.S.Y. and Mayne, P.W. (1996). Statistical relationships between piezocone measurements and stress history of clays. *Canadian Geotechnical Journal*, 33 (3), 488-498.
- Chen, J.R. (2004). Axial Behavior of Drilled Shafts in Gravelly Soils. Ph.D. Dissertation, Cornell University, Ithaca, NY.
- Ching, J. and Phoon, K.K. (2012a), Modeling parameters of structured clays as a multivariate normal distribution. *Canadian Geotechnical Journal*, 49(5), 522-545.
- Ching, J. and Phoon, K.K. (2012b), Establishment of generic transformations for geotechnical design parameters. *Structural Safety*, 35, 52-62.
- Ching, J. and Phoon, K.K. (2013). Multivariate distribution for undrained shear strengths under various test procedures. *Canadian Geotechnical Journal*, 50(9), 907-923.
- Ching, J. and Phoon, K.K. (2014). Transformations and correlations among some clay parameters – the global database, *Canadian Geotechnical Journal*, 51(6), 663-685.
- Ching, J., Phoon, K.K., and Chen, C.H. (2014). Modeling CPTU parameters of clays as a multivariate normal distribution. *Canadian Geotechnical Journal*, 51(1), 77-91.
- Ching, J. and Phoon, K.K. (2015). Reducing the transformation uncertainty for the mobilized undrained shear strength of clays. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 141(2), 04014103.
- Ching, J., Lin, G.H., Chen, J.R., and Phoon, K.K. (2017). Transformation models for effective friction angle and relative density calibrated based on a multivariate database of cohesionless soils. *Canadian Geotechnical Journal*, 54(4), 481-501.
- D'Ignazio, M., Phoon, K.K., Tan, S.A., and Lansivaara, T. (2016). Correlations for undrained shear strength of Finnish soft clays. *Canadian Geotechnical Journal*, 53(10), 1628-1645.

- Djoenaidi, W.J. (1985). A Compendium of Soil Properties and Correlations. MEngSc thesis, University of Sidney, Sidney, Australia.
- Hatanaka, M, Uchida, A., Kakurai, M., and Aoki, M. (1998). A consideration on the relationship between SPT N-value and internal friction angle of sandy soils. *Journal of Structural Construction Engineering*, Architectural Institute of Japan, 506, 125-129. (in Japanese)
- Hatanaka, M. and Uchida, A. (1996), Empirical correlation between penetration resistance and internal friction angle of sandy soils. *Soils and Foundations*, 36(4), 1-9.
- Jamiolkowski, M., Ladd, C.C., Germain, J.T., and Lancellotta, R. (1985). New developments in field and laboratory testing of soils. *Proceeding of the 11th International Conference on Soil Mechanics and Foundation Engineering*, San Francisco, 1, 57-153.
- Kootahi, K. and Mayne, P.W. (2016). Index test method for estimating the effective preconsolidation stress in clay deposits. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 10.1061/(ASCE)GT.1943-5606.0001519, 04016049.
- Kulhawy, F.H. and Mayne, P.W. (1990). *Manual on Estimating Soil Properties for Foundation Design*, Report EL-6800, Electric Power Research Institute, Cornell University, Palo Alto.
- Locat, J. and Demers, D. (1988). Viscosity, yield stress, remoulded strength, and liquidity index relationships for sensitive clays. *Canadian Geotechnical Journal*, 25, 799-806.
- Marcuson, W.F. III and Bieganousky, W.A. (1977). SPT and relative density in coarse sands. *ASCE Journal of the Geotechnical Engineering Division*, 103(11), 1295-1309.
- Mayne, P.W., Christopher, B.R., and DeJong, J. (2001). *Manual on Subsurface Investigations*. National Highway Institute Publication No. FHWA NHI-01-031, Federal Highway Administration, Washington, D.C.
- Mesri, G. (1975). Discussion on "New design procedure for stability of soft clays". *Journal of the Geotechnical Engineering Division, ASCE*, 101(4), 409-412.
- Mesri, G. and Huvaj, N. (2007). Shear strength mobilized in undrained failure of soft clay and silt deposits. *Advances in Measurement and Modeling of Soil Behaviour (GSP 173)*, Ed. D.J. DeGroot et al., ASCE, 1-22.
- Phoon, K.K. and Kulhawy, F.H. (1999). Characterization of geotechnical variability. *Canadian Geotechnical Journal*, 36(4), 612-624.
- Robertson, P.K. and Campanella, R.G. (1983). Interpretation of cone penetration tests: part I - sands. *Canadian Geotechnical Journal*, 20(4), 718-733.
- Salgado, R., Bandini, P., and Karim, A. (2000). Shear strength and stiffness of silty sand. *Journal of Geotechnical and Geoenvironmental Engineering*, 126(5), 251-462.
- Stas, C. V. and Kulhawy, F. H. (1984). Critical evaluation of design methods for foundations under axial uplift and compressive loading. Electric Power Research Institute, Palo Alto, Calif. Report EL-3771.
- Terzaghi, K. and Peck, R.B. (1967). *Soil Mechanics in Engineering Practice*, 2nd Edition. John Wiley and Sons, New York.

Appendix 1A: Transformation models calibrated by soil databases

This appendix presents the calibration results for some transformation models in the literature. The calibrated models can be used to develop the point estimate and 95% confidence interval (CI) for the design soil parameter. The bias and variability for the clay transformation models are calibrated by the F-CLAY/7/216 and CLAY/10/7490 databases (see Table 1A-1), whereas the sand transformation models are calibrated by the SAND/7/2794 database (see Table 1A-2).

To explain the significance of the bias and variability for a transformation model, consider the first model in Table 1A-1, the LI-(s_u^{re}/P_a) model proposed by Locat and Demers (1988). The actual target value is s_u^{re}/P_a , and the predicted target value is $0.0144 \times LI^{2.44}$. For each data point in the database with simultaneous knowledge of (LI, s_u^{re}), (actual target value)/(predicted target value) = $(s_u^{re}/P_a)/(0.0144 \times LI^{2.44})$ can be computed. The sample mean of this ratio is called the bias factor (b) for the transformation model. The sample coefficient of variation (COV) of this ratio is called the COV (δ) of the transformation model. To be specific,

$$\text{Actual target value} = \text{predicted target value} \times b \times \varepsilon \quad (1A-1)$$

where b is the bias factor (b = 1 means unbiased), and ε is the variability term with mean = 1 and COV = δ . If $\delta = 0$, there is no data scatter about the transformation model, i.e. the prediction is single-valued or deterministic, rather than a distribution. The calibrated bias factors and COVs for all clays and sand transformation models are shown in the last two columns of Tables 1A-1 and 1A-2, respectively. The number of data points used for each calibration is listed in the table (‘n’ in the third column).

The calibrated bias and COV of a transformation model can be adopted to develop the point estimate and 95% CI, described as follows. Consider again the LI-(s_u^{re}/P_a) model, let the site-specific LI value for the new design site be denoted by LI_{new} , the point estimate for $(s_u^{re}/P_a)_{new}$ is simply $b \times (\text{predicted target value}) = b \times (0.0144 \times LI_{new}^{2.44})$. By assuming ε to be lognormal, the 95% CI for $(s_u^{re}/P_a)_{new}$ can be expressed as

$$\begin{aligned} & \frac{b \times (\text{predicted target value})}{\sqrt{1 + \delta^2}} \times \exp\left(\pm 1.96 \times \sqrt{\ln(1 + \delta^2)}\right) \\ & = \frac{b \times (0.0144 \times LI_{new}^{2.44})}{\sqrt{1 + \delta^2}} \times \exp\left(\pm 1.96 \times \sqrt{\ln(1 + \delta^2)}\right) \end{aligned} \quad (1A-2)$$

If the design site is a Finland site, the chance for the actual target value to be within the above nominal 95% CI (with b and δ calibrated by F-CLAY/7/216) should be close to 95%. If the design site is a general site, the chance for the actual target value to be within the above nominal 95% CI (with b and δ calibrated by CLAY/10/7490) should be close to 95%. The numbers of calibration data points (n) for some sand transformation models are quite limited (see Table 1A-2). For those transformation models, their nominal 95% CI may not be genuine.

Table 1A-1 Transformation models in the literature for some clay parameters.

Target parameter	Measured parameter(s)	Literature	Transformation model	Calibration database		Calibration results	
				n	Bias (b)	COV (δ)	
s_u^{re}	LI	Locat and Demers (1988)	$s_u^{re}/P_a \approx 0.0144 \times LI^{-2.44}$	CLAY/10/7490	899	1.92	1.25
S_t	LI	Bjerrum (1954)	$S_t \approx 10^{0.8 \times LI}$	F-CLAY/7/216	216	2.23	1.08
S_t	LI	Ching and Phoon (2012a)	$S_t \approx 20.726 \times LI^{1.910}$	CLAY/10/7490	1279	2.06	1.09
σ'_p	LI, S_t	Stas and Kulhawy (1984)	$\sigma'_p/P_a \approx 10^{1.1-1.62 \times LI}$	F-CLAY/7/216	216	1.56	1.40
σ'_p	LI, S_t	Ching and Phoon (2012a)	$\sigma'_p/P_a \approx 0.235 \times LI^{-1.319} \times S_t^{0.556}$	CLAY/10/7490	1279	0.88	1.28
			If $5.512 \log_{10}(\sigma'_v/P_a) - 0.061 \times LL - 0.093 \times PL + 6.219 \times e_n > 1.123$	F-CLAY/7/216	216	0.57	1.94
			$\Rightarrow \sigma'_p/P_a \approx 1.62 \times (\sigma'_v/P_a)^{0.89} (LL)^{0.12} (w_n)^{-0.14}$	CLAY/10/7490	489	1.32	0.78
			Otherwise	F-CLAY/7/216	216	1.35	0.94
σ'_p	w_n, PL, LL	Kootahi and Mayne (2016)	$\Rightarrow \sigma'_p/P_a \approx 7.94 \times (\sigma'_v/P_a)^{0.71} (LL)^{0.53} (w_n)^{-0.71}$	CLAY/10/7490	1242	1.10	0.67
σ'_p	qt	Kulhawy and Mayne (1990)	$\sigma'_p/P_a \approx 0.33 \times (q_t - \sigma'_v)/P_a$	F-CLAY/7/216	216	1.02	0.38
			$\sigma'_p/P_a \approx 0.54 \times (u_2 - u_0)/P_a$	CLAY/10/7490	690	0.97	0.39
			$\sigma'_p/P_a \approx 0.227 \times [(q_t - \sigma'_v)/P_a]^{1.200}$	CLAY/10/7490	690	1.18	0.75
			$\sigma'_p/P_a \approx 0.490 \times [(q_t - u_2)/P_a]^{1.053}$	CLAY/10/7490	690	0.99	0.42
σ'_p	qt	Chen and Mayne (1996)	$\sigma'_p/P_a \approx 1.274 + 0.761 \times (u_2 - u_0)/P_a$	CLAY/10/7490	542	1.08	0.61
OCR	qt	Kulhawy and Mayne (1990)	OCR $\approx 0.32 \times (q_t - \sigma'_v)/\sigma'_v$	CLAY/10/7490	690	0.49	0.59
			OCR $\approx 0.259 \times [(q_t - \sigma'_v)/\sigma'_v]^{1.107}$	CLAY/10/7490	690	1.00	0.39
OCR	qt	Chen and Mayne (1996)	OCR $\approx 0.545 \times [(q_t - u_2)/\sigma'_v]^{0.969}$	CLAY/10/7490	690	1.01	0.42
			OCR $\approx 1.026 \times B_q^{-1.077}$	CLAY/10/7490	542	1.06	0.57
s_u	PI	Mesri (1975)	$s_u/\sigma'_p \approx 0.22$	CLAY/10/7490	779	1.28	0.86
s_u	OCR	Jamiolkowski et al. (1985)	$s_u/\sigma'_v \approx 0.23 \times OCR^{0.8}$	CLAY/10/7490	1155	1.04	0.55
			$s_u/\sigma'_v \approx 0.229 \times OCR^{0.823} \times S_t^{0.121}$	F-CLAY/7/216	216	1.08	0.28
				CLAY/10/7490	1402	1.11	0.53
				F-CLAY/7/216	216	1.15	0.29
				CLAY/10/7490	395	0.84	0.34
				F-CLAY/7/216	216	0.84	0.32

s_u	q_t	Ching and Phoon (2012b)	$\frac{[(q_t - \sigma_v)/\sigma'_v]/(s_u/\sigma'_v)}{[(q_t - u_2)/\sigma'_v]/(s_u/\sigma'_v)} \approx 29.1 \times \exp(-0.513B_q)$	CLAY/10/7490	423	0.95	0.49
			$\frac{[(u_2 - u_0)/\sigma'_v]/(s_u/\sigma'_v)}{[(u_2 - u_0)/\sigma'_v]/(s_u/\sigma'_v)} \approx 21.5 \times B_q$	CLAY/10/7490	428	1.11	0.57
				CLAY/10/7490	423	0.94	0.49

* All s_u are the “mobilized” s_u defined by Mesri and Huvaj (2007); e_n : natural void ratio.

Table 1A-2 Transformation models in the literature for some sand parameters.

Target parameter	Measured parameter(s)	Literature	Transformation model	Calibration database		Calibration results	
				n	Bias (b)	COV (δ)	
D_r	$(N_1)_{60}$	Terzaghi and Peck (1967)	$D_r(\%) \approx 100 \times \sqrt{(N_1)_{60}} / 60$	SAND/7/2794	198	1.05	0.231
D_r	N_{60} , OCR, C_u	Marcuson and Bieganousky (1977)	$D_r(\%) \approx 100 \times \left\{ 12.2 + 0.75 \sqrt{\frac{222 \times N_{60} + 2311 - 711 \times \text{OCR}}{-779(\sigma'_{v0}/P_s) - 50 \times C_u^2}} \right\}$	SAND/7/2794	132	1.00	0.211
D_r	$(N_1)_{60}$, OCR, D_{50}	Kulhawy and Mayne (1990)	$D_r(\%) \approx 100 \times \sqrt{\frac{(N_1)_{60}}{[60 + 25 \log_{10}(D_{50})] \times \text{OCR}^{0.18}}}$	SAND/7/2794	199	1.01	0.205
D_r	q_{ti}	Jamiolkowski et al. (1985)	$D_r(\%) \approx 68 \times [\log_{10}(q_{ti}) - 1]$	SAND/7/2794	681	0.84	0.327
D_r	q_{ti} , OCR	Kulhawy and Mayne (1990)	$D_r(\%) \approx 100 \times \sqrt{\frac{q_{ti}}{305 \times Q_c \times \text{OCR}^{0.18}}}$	SAND/7/2794	840	0.93	0.339
ϕ'	D_r , ϕ'_{cv}	Bolton (1986)	$\phi' \approx \phi'_{cv} + 3 \times (D_r [10 - \ln(p'_r)] - 1)$	SAND/7/2794	391	1.03	0.052
ϕ'	D_r , ϕ'_{cv}	Salgado et al. (2000)	$\phi' \approx \phi'_{cv} + 3 \times (D_r [8.3 - \ln(p'_r)] - 0.69)$	SAND/7/2794	127	1.08	0.054
ϕ'	$(N_1)_{60}$	Hatanaka and Uchida (1996)	$\phi' \approx \sqrt{15.4 \cdot (N_1)_{60}} + 20$	SAND/7/2794	28	1.04	0.095
ϕ'	$(N_1)_{60}$	Hatanaka et al. (1998)	$\phi' \approx \begin{cases} \sqrt{15.4 \cdot (N_1)_{60}} + 20 & (N_1)_{60} \leq 26 \\ 40 & (N_1)_{60} > 26 \end{cases}$	SAND/7/2794	58	1.07	0.090
ϕ'	$(N_1)_{60}$	Chen (2004)	$\phi' \approx 27.5 + 9.2 \times \log_{10} [(N_1)_{60}]$	SAND/7/2794	59	1.00	0.095
ϕ'	q_t	Robertson and Campanella (1983)	$\phi' \approx \tan^{-1} [0.1 + 0.38 \times \log_{10}(q_t / \sigma'_{v0})]$	SAND/7/2794	99	0.93	0.056
ϕ'	q_{ti}	Kulhawy and Mayne (1990)	$\phi' \approx 17.6 + 11 \times \log_{10}(q_{ti})$	SAND/7/2794	376	0.97	0.081

* ϕ'_{cv} : critical-state friction angle (in degrees); p'_r is the mean effective stress at failure = $(\sigma'_{1r} + \sigma'_{2r} + \sigma'_{3r})/3$; $Q_c = 1.09, 1.0, 0.91$ for low, medium, high compressibility soils, respectively.

Discussion 1A – Sedimentary formations with heterogeneous 3D architecture

Celeste Jorge (Laboratório Nacional de Engenharia Civil - LNEC, Lisboa - Portugal)

Geological formations formed in a basin or a lacustrine environment, with endorheic regimen, are worldwide distributed, and are composed by continental detritic materials. These materials were deposited on a band, more or less wide, near the basin bank, giving rise to multiple alluvial fans. These geological formations are predominated by coarse materials (such as sandstones and conglomerates) and also by intermitted layers of calcareous materials (such as limestones and marls). The coarse materials are deposited during/immediately after torrential climatic periods, whereas the calcareous ones are deposited during calm periods. Furthermore, the basins were invaded successively by the ocean (advancements and retreatments of sea level over long geological periods). The episodes of advancements and retreatments of the sea level caused the deposition process highly variable and produce geological formations with significant three-dimensional (3D) heterogeneity.

The irregular structure and texture of these geological formations make it almost impossible to predict their spatial distribution. This feature is reflected in the great heterogeneity of the geotechnical characteristics. Thus, the use of probabilistic models for extrapolating parameters at unexplored locations may introduce a large error. Therefore, it is not advisable to use this type of extrapolation approach in such geological formations.

Reply to Discussion 1A

Jianye Ching

Celeste correctly pointed out that caution should be taken when parameters are extrapolated at unexplored locations. The highly spatially variable geological formations mentioned by Celeste serve as a very good example why such caution should be taken. I agree very much with Celeste. However, the development and implementation of a transformation model do not involve in spatial extrapolation.

To explain why the development of a transformation model does not involve in spatial extrapolation, let us consider the $(N_1)_{60}-\phi'$ model in Figure 1-1 as an example. Any data point in this figure is based on the test results of two sands in the literature. One sand is tested by the standard penetration test in situ to derive its $(N_1)_{60}$, whereas another sand sample is extracted using the ground freezing method and tested in laboratory to obtain its ϕ' . More importantly, the spatial locations of these two sands are typically very close, e.g., at the same depth and with few meters apart horizontally, to minimize the effect of spatial variation. As a result, these two sands are practically considered as the “same” sand, and the effect of spatial variation does not really enter into the data points in Figure 1-1.

The implementation of a transformation model does not involve in spatial extrapolation, either. For instance, suppose that an engineer would like to implement the transformation model (the blue dashed line in Figure 1-1) to estimate ϕ' for a sand at depth of 10 m with the knowledge of $(N_1)_{60}$ at the same depth. This process does not involve spatial extrapolation, either, because 10 m is the only depth of concern.

However, extrapolation can still happen when implementing a transformation model, but the extrapolation is not in space but in soil/rock type. For instance, the data points in Figure 1-1 are mostly from siliceous sands and there is no calcareous sand. If an engineer would like to implement the transformation model in Figure 1-1 to estimate ϕ' for a calcareous sand, extrapolation may happen.

The geological formations example mentioned by Celeste is indeed challenging in the aspect of spatial variability. However, transformation models are not for modeling and predicting spatial variability. The Discussion Group led by Dianqing Li (Incorporating spatial variability into geotechnical reliability based design) is more relevant to the subject of spatial variability.

Chapter 2 Evaluation and Consideration of Model Uncertainties in Reliability Based Design

Lead discussor:

Kerstin Lesny

kerstin.lesny@hcu-hamburg.de

Discussors (alphabetical order):

Sami Oguzhan Akbas, Witold Bogusz, Sébastien Burlon, Giovanna Vessia, Kok-Kwang Phoon, Chong Tang, Limin Zhang

2.1 INTRODUCTION

The quality and reliability of geotechnical design depends among other aspects on the quality of the chosen design method. This can be a simple analytical or empirical equation or a sophisticated numerical model with different degrees of complexity and accuracy in modeling the behavior of a geotechnical structure. However, exact modeling of the real behavior is impossible and the deviation of the predicted from the real behavior is expressed by the model uncertainty. According to e.g. Nadim (2015) model uncertainty is of epistemic nature as it is based on the simplifications, assumptions and approximations made in the respective design model. It therefore can be reduced by improving the model. Nadim (2015) also used the term transformation uncertainty for the uncertainties associated with the model as the model transfers input parameters to output parameters. A recent summary of geotechnical model uncertainty was presented by Dithinde et al. (2016).

It should be explicitly noted, that model uncertainty in the context of this contribution and according to the definition given above only refers to the uncertainties of the applied calculation procedure or method. It shall not consider uncertainties from the overall geotechnical model which also includes the definition of the ground model and the related soil mechanical parameters used in the design. However, the model uncertainty often is intrinsically tied to these uncertainties as will be shown in the following.

For the geotechnical engineering design practice, as considered by Eurocode 7, calculation models in many countries were calibrated on the basis of an overall factor of safety resulting from previous standards and past experience leading to design approaches that are often regarded as being conservative. In such cases failures of geotechnical structures are usually attributed to insufficient ground investigations or external factors (human error, overloading, water levels exceeding design assumptions, etc.), but not to an insufficient calculation model. However, derivation of partial factors with the aim of achieving a consistent reliability level within geotechnical design requires the separation of different sources of uncertainties and therefore establishes a need to analyze the uncertainties associated with the respective design methods.

This chapter presents the current status of model uncertainty assessment for different types of geotechnical structures and calculation models based on a discussion among members of the ISSMGE technical committees TC 205 and TC 304.

2.2 PROCEDURES FOR MODEL UNCERTAINTY ASSESSMENT

The need for separate assessment of model uncertainties was lately emphasized by Eurocode 7 code drafters involved in the current revision process of the Eurocodes. Following this postulation, for geotechnical design according to Eurocode 7 or other standards utilizing load and resistance factor design (LRFD), a pre-defined target reliability shall be achieved by the combination of partial factors and model factors. The latter shall be used to calibrate specific calculation models to provide the same level of confidence in the validity of the prediction for possible design situations prior to the use of

partial safety factors. It further allows the engineers to better appreciate the reliability of the applied calculation models, and thus improve their designs.

In this regard, Phoon and Ching (2015) argued that simplified reliability based design (RBD) methods such as LRFD with constant factors can only cover the design situations inside the domain the factors have been calibrated for. For other situations the target reliability may not necessarily be reached. Consequently, the evaluation of model uncertainties must also reflect all possible design situations to be covered by the code.

The way to introduce model uncertainties into design very much depends on the character of the applied design method or approach. Consequently, various procedures for model uncertainty evaluation and the associated difficulties are outlined in the following.

2.2.1 Model uncertainty expressed by model factors

For simple analytical, empirical or semi-empirical design methods, which result in a unique design quantity, model uncertainties can be considered relatively straightforward by a model factor according to Eq. 2-1, where the model factor or bias M is defined as the ratio of a measured (e.g. in a load test) to a calculated quantity X .

$$M = X_{\text{meas}} / X_{\text{cal}} \quad (2-1)$$

The quantity X can be a load, a resistance, a displacement, etc. The model factor itself is not constant, but a random value. Hence, it can be introduced in reliability analyses, where it is usually assumed to be log-normally distributed with a mean and a COV to be defined (e.g. Phoon and Kulhawy 2005; Juang et al. 2012). In this context, a mean close to 1.0 would represent the ideal solution, whereas the COV represents the scatter and therefore the uncertainty in the calculation model. Hence, by introducing a model factor into the design a certain reliability of the prediction using a specific calculation model can be provided, thus ensuring that there is only a $p\%$ probability that the real value is lower than calculated one (Figure 2-1).

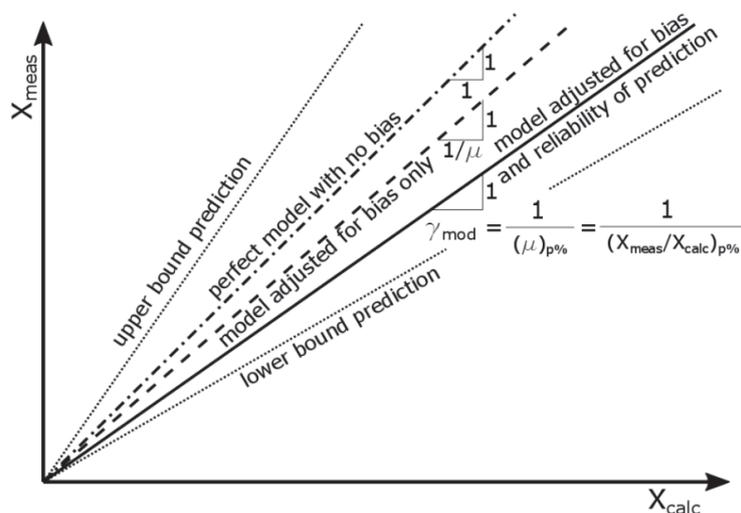


Figure 2-1 Model bias implemented into geotechnical design according to Eurocode 7 (courtesy of Witold Bogusz).

Literature review revealed that the simple definition in Eq. 2-1 has been often used for model uncertainty evaluation for

- (1) Axially loaded piles (settlement prediction or bearing capacity prediction)
- (2) Spread foundations (settlement prediction or bearing capacity prediction)

Here, the design problem can be reduced to one single quantity which is the result of the applied calculation model, e.g. the pile bearing capacity or the foundation bearing capacity. In case of pile foundations pile loading and pile capacity are usually treated as independent quantities, hence the bias can be clearly defined (though depending on the type of the prediction model). In the case of shallow

foundations the bearing resistance is not an independent quantity as it depends on the type of load. Therefore, the bias cannot be unique.

2.2.2 Model uncertainty prediction for more complex design situations

For more complex design situations or design models not resulting in a unique design quantity or where results have to be interpreted from the calculations (e.g. FEM calculations) model uncertainty prediction may be based on representative quantities which properly characterize the performance as well as the design of the structure.

For example, Zhang et al. (2015) chose the top deflection as the representative quantity for the design of a cantilever wall and evaluated the model uncertainty of the mobilized strength design (MSD) proposed by Osman and Bolton (2004). However, for more complex braced excavations, it may be necessary to consider the maximum deflection and the maximum bending moment – both quantities do not appear at a fixed depth location and change with the excavation stage. It is worthwhile to note that the maximum deflection and the maximum bending moment are likely to be correlated and there is a need to consider how the model uncertainties related to both quantities can be updated based on monitoring which is routinely carried out using inclinometers.

Laterally loaded pile design is driven by the anticipated flexibility of the pile which determines the type of calculation model. For rigid piles evaluation of the model uncertainty related to the bearing capacity prediction is relatively straightforward and follows the procedure outlined in section 2.2.1. Phoon and Kulhawy (2005), for example, analyzed the model uncertainties of different earth pressure models by calculating the ultimate lateral and moment capacity which were the design quantities.

For flexible piles it is questionable, if the pile head displacement (or rotation) as a single representative quantity is sufficient or if the whole deflection curve is a better choice. On the other hand, it seems possible to compare measured and calculated load-displacement curves, but only very few load tests on laterally loaded piles are available. Such a work has been presented for axially loaded piles by Abchir et al. (2016) to assess the uncertainty of t-z curves obtained from pressuremeter test results.

A similar problem arises for settlement prediction of footings. In case of single rigid footings a representative settlement and rotation can be defined. In case of flexible footings this is at least difficult.

2.2.3 Model uncertainty assessment by reliability based sensitivity analyses

Another procedure to assess the model uncertainty of a given design method is to perform a RBD based sensitivity analysis to evaluate the relative influence random parameters have on the reliability index or the probability of failure, respectively, when adopting a specific design method. Such a work was proposed e.g. by Teixeira et al. (2012) who analyzed the bearing capacity of piles calculated by an empirical method directly correlating the blow counts N_{SPT} from standard penetration tests (SPT) to the pile capacity. The sensitivity analysis was based on calculations using the first order reliability method (FORM) and Monte Carlo Simulations (MCS). The soil variability expressed by the blow count N_{SPT} , model factors for base and shaft resistance of the pile and factors for load uncertainties were considered as random variables with statistical parameters adopted from literature. The resulting reliability indices were compared for different scenarios in which the different sources of uncertainties were considered or not. Such a procedure allows an assessment of the importance of different sources of uncertainties within design. For the investigated design method it was shown that model uncertainties were the predominant sources of uncertainties emphasizing the importance for appropriately considering them in design.

2.2.4 Interpretation of load-displacement curves for model uncertainty assessment in the serviceability limit state

At the ultimate limit state, a consistent load test interpretation procedure should be used to produce a single “measured capacity” from each measured load-displacement curve. The ratio of the measured capacity to the calculated capacity is called a model factor as defined in Eq. (2-1). The same approach applies to the serviceability limit state (SLS). The capacity is replaced by an allowable capacity that

depends on the allowable displacement. The distribution of the SLS model factor is established from a load test database in the same way. Notice that the SLS model factor has to be re-evaluated when a different allowable displacement is prescribed. If the allowable settlement is treated as a random variable in the serviceability limit state, a more general approach involving fitting measured load-displacement data to a normalized hyperbolic curve is recommended as detailed below:

$$\frac{Q}{Q_m} = \frac{y}{a + by} \quad (2-2)$$

in which Q = applied load, Q_m = failure load or capacity interpreted from a measured load-displacement curve, “a” and “b” = curve-fitting parameters, and y = pile butt displacement. Note that the curve-fitting parameters are physically meaningful, with the reciprocals of “a” and “b” equal to the initial slope and asymptotic value of the hyperbolic curve, respectively.

The curve-fitting equation is empirical and other functional forms can be considered (Phoon and Kulhawy 2008). However, the important criterion is to apply a curve-fitting equation that produces the least scatter in the measured normalized load-displacement curves. Each measured load-displacement curve is thus reduced to two curve-fitting parameters. Based on “a” and “b” statistics estimated from the load test database (Table 2-1), one can construct an appropriate bivariate probability distribution for (a, b) that can reproduce the scatter in the normalized load over the full range of displacements. Details are given in Phoon and Kulhawy (2008). It is evident that this approach can be used in conjunction with a random allowable settlement. This approach has been applied to various foundation types (Phoon et al. 2006, Phoon et al. 2007, Akbas & Kulhawy 2009, Dithinde et al. 2011, Stuedlein and Reddy 2013, Huffman and Stuedlein 2014, Huffman et al. 2015).

Table 2-1 Summary of statistics for hyperbolic parameters

Reference	Foundation type	Load	Soil type	N	a (mm)		b		ρ
					Mean	COV	Mean	COV	
Phoon et al.(2006)	ACIP (D/B>20)	C	Sand	40	5.15	0.6	0.62	0.26	-0.67
Phoon e al. (2007)	Spread footing	U	Clay+sand	85	7.13	0.65	0.75	0.18	-0.24
	Drilled shaft	U	Clay+sand	48	1.34	0.54	0.89	0.07	-0.59
	Pressure-injected footing	U	Sand	25	1.38	0.68	0.77	0.27	-0.73
Dithinde et al. (2011)	Driven pile	C	Sand	28	5.55	0.54	0.71	0.14	-0.78
	Bored pile	C	Sand	30	4.1	0.78	0.77	0.21	-0.88
	Driven pile	C	Clay	59	3.58	0.57	0.78	0.11	-0.89
	Bored pile	C	Clay	53	2.79	0.57	0.82	0.11	-0.8
Stuedlein and Reddy (2013)	ACIP	C	Sand	87	0.16	0.49	3.4	0.23	-0.73
Huffman and Stuedlein (2014)	Spread footing	C	Reinforced Clay	30	2	0.79	1.15	0.25	—
Huffman et al. (2015)	Spread footing	C	Clay	30	1.3	0.53	0.7	0.16	-0.95
Tang et al. (2017b)	Square foundation	C	Sand	64	1.47	0.4	0.72	0.17	-0.76
Tang and Phoon (2017b)	Small-diameter helical pile (square shaft)	C	Clay	53	5.54	0.36	0.78	0.14	-0.46
		C	Sand	49	5.84	0.27	0.76	0.14	-0.36
	Large-diameter helical pipe pile (single-helix)	C	Clay	11	6.2	0.5	0.83	0.13	-0.85
Tang and Phoon (2017b)	Large-diameter helical pipe pile (multi-helix)	C	Clay	18	5.62	0.59	0.83	0.12	-0.88
Tang and Phoon (2017c)	Steel H-pile	C	Clay	47	2.77	0.61	0.81	0.11	-0.72
		C	Sand	53	5.35	0.71	0.68	0.21	-0.73
		C	Layered	51	3.9	0.74	0.74	0.14	-0.76

Note:

- (1) The bivariate load-settlement model used by Huffman and Stuedlein (2014), Huffman et al. (2015), and Tang et al. (2017b) is $Q/Q_m = (s/B) / [a + b(s/B)]$.
- (2) The model factor in Stuedlein and Reddy (2013) is transformed model factor, namely, $a_t = a \times (B/D)$ and $b_t = b \times (D/B)^{0.5}$.
- (3) ACIP=augered cast-in-place pile, D=embedment depth, B=foundation width or diameter, C=compression, U=uplift, N=number of load tests, ρ =correlation between the hyperbolic parameters (a, and b).

2.2.5 Difficulties in model uncertainty evaluation

Prediction of model uncertainties based on measurements of the real behavior as provided by the simple definition given in section 2.2.1 reveals some fundamental problems:

- (1) The bias cannot be separated from the inherent variability of parameter values used within the model.

The inherent variability of parameters refers on one the hand to the way they are determined (from theoretical formulas, indirect correlations or direct measurements). On the other hand, the spatial variability of soil characteristics plays an important role.

Fenton and Griffith (2005) showed in an analysis of the reliability of traditional retaining wall design, assuming one particular failure mode, how the spatial variability of the soil affects the failure mechanism which in turn affects the uncertainty of the chosen design method. On the other hand, Teixeira et al. (2012) in their study on axially loaded piles showed that soil variability was not as important as expected. One may conclude that the different results are related to the failure modes assumed in both cases. In the pile design the shaft friction and the end bearing of the pile are not or not as much affected by soil variability as in case of the retaining wall design where large failure zones in the soil have to be considered. The same applies e.g. to the design of shallow foundations. Hence, especially in the case where large failure zones are involved, the model uncertainty is intrinsically tied to the uncertainties related to the soil characteristics and cannot be separated.

- (2) The bias cannot be separated from measurement errors.

Measurement errors are related to the experimental device used to measure the variables of the model and other related errors. On the other side, especially in the case of complex numerical calculations the prediction is also affected by other sources of uncertainties related to the modeling itself which indirectly contribute to the model uncertainty (e.g. user experience, choice of constitutive models, mesh generation procedures).

It may be concluded that there is no unique model uncertainty. So as already stated above, the evaluation of model uncertainties and the derivation of model factors shall cover all possible design situations.

- (3) The bias cannot be separated from definition and determination of X_{meas} from the test (e.g. consistency of failure criteria and its application to the test results).

Phoon and Kulhawy (2005) demonstrated using a large load test database that the mean of the model factor for the lateral capacity of a rigid drilled shaft is a function of the capacity interpretation method (lateral or moment limit, hyperbolic limit). To characterize model uncertainties, a consistent load test interpretation procedure should be used to produce a single “measured” quantity X_{meas} from the test.

- (4) The model factor given by Eq. 2-1 could be a function of input parameters.

Ideally, a theoretical model should capture the key features of physical processes, and the remaining difference between the model and reality should be random in nature as it is caused by numerous minor factors that were left out of the model. In practice, however, the theoretical modelling of physical processes generally entails making unrealistic assumptions and simplifications physically and geometrically just to create a useable and oftentimes an analytically tractable model. In this context, the ratio M between the measured value X_{meas} and the calculated value X_{cal} may not be random in the sense that it is systematically affected by input parameters such as the problem geometry and soil mechanical properties. Examples are given by Zhang et al. (2015) (see Figure 2-2), Phoon and Tang (2015, 2017) (see Figure 2-3), Tang and Phoon (2016, 2017a), and Tang et al. (2017a).

- (5) Finally, the model factor approach reaches its limits where obvious deficiencies in the design

model exist.

An example of this may be the pile design for combined loading which is a typical design situation for offshore pile foundation structures. Available design methods account only for one load component (i.e. axial or lateral). The interaction of the load components is usually neglected arguing that lateral loading affects predominantly the upper part of the pile whereas axial loading mobilizes resistances predominantly in greater depths. The uncertainties related to such a design procedure can hardly be addressed by simple model factors.

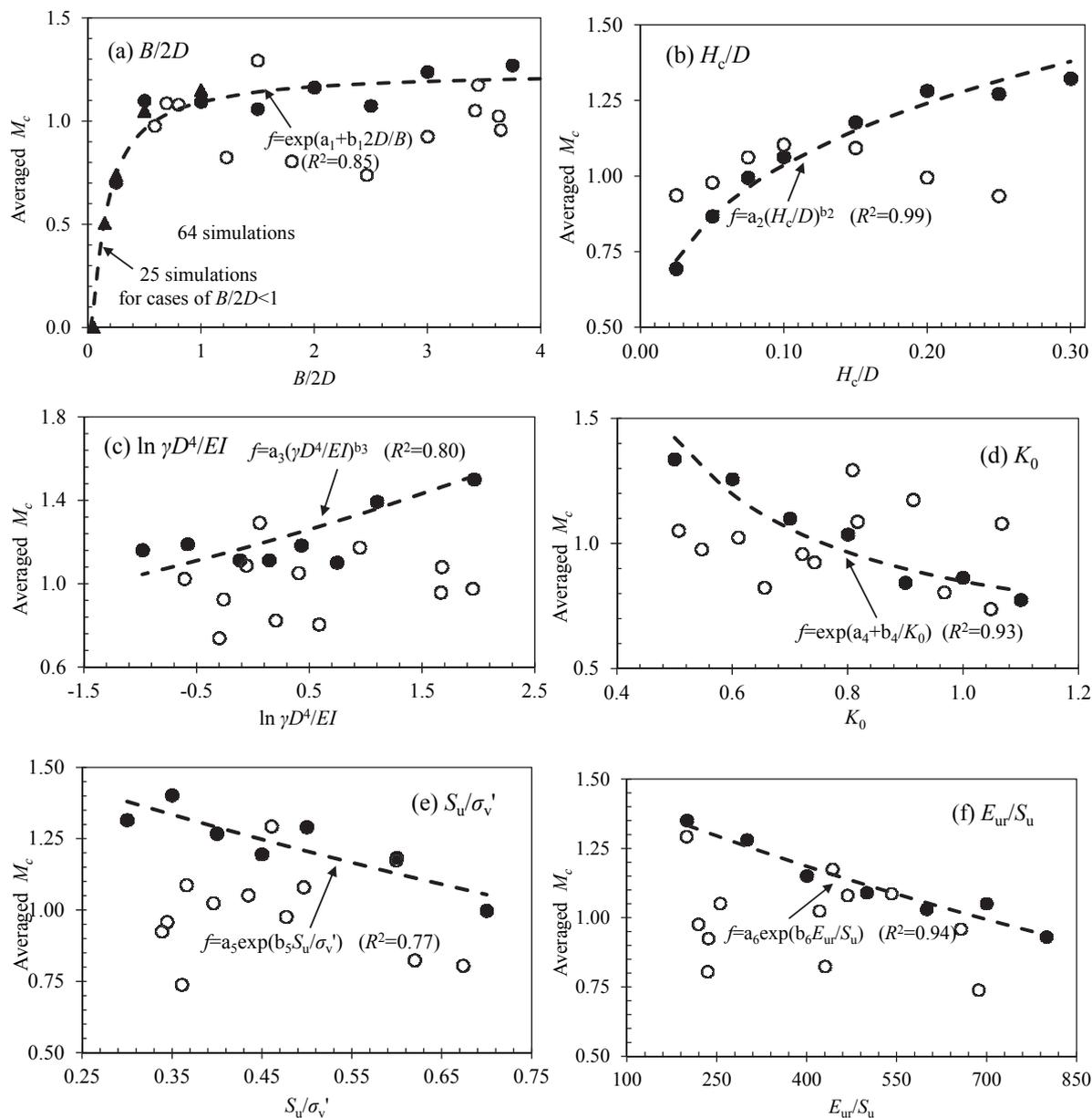


Figure 2-2 Variation of model factor M for MSD calculating the top deflection of a cantilever wall with input parameters (Zhang et al. 2015).

2.3 OVERVIEW OF EXISTING WORK

2.3.1 Shallow foundations

Appendices 2A and 2B summarize some work on model uncertainty prediction for bearing capacity calculation of shallow foundations.

Model uncertainty was defined in Appendix 2A by $M = X_{meas}/X_{cal}$ as in Eq. 2-1 and by $1/M = X_{cal}/X_{meas}$ in Appendix 2B. The work presented in Appendix 2B deals with vertically loaded foundations only, whereas the work in Appendix 2A considers combined loading as well. Both used extensive databases including field tests only (Appendix 2B) and laboratory and field tests (controlled and natural soil conditions – Appendix 2A).

Both studies show an influence of the footing size on the model factor or bias (possible effect of model scale), but this trend is not clear. In Appendix 2B the bias according to Eq. 2-1 decreases for greater footing width. In Appendix 2A a similar trend has been found for footings in natural soil conditions whereas the trend in controlled soil conditions is different (increase of bias for greater footings), see Figure 2-4. Hence, this confirms the earlier statement that the model uncertainty is affected among others by the variability of the soil especially under natural conditions and/or the derivation of the design parameters (here shear strength).

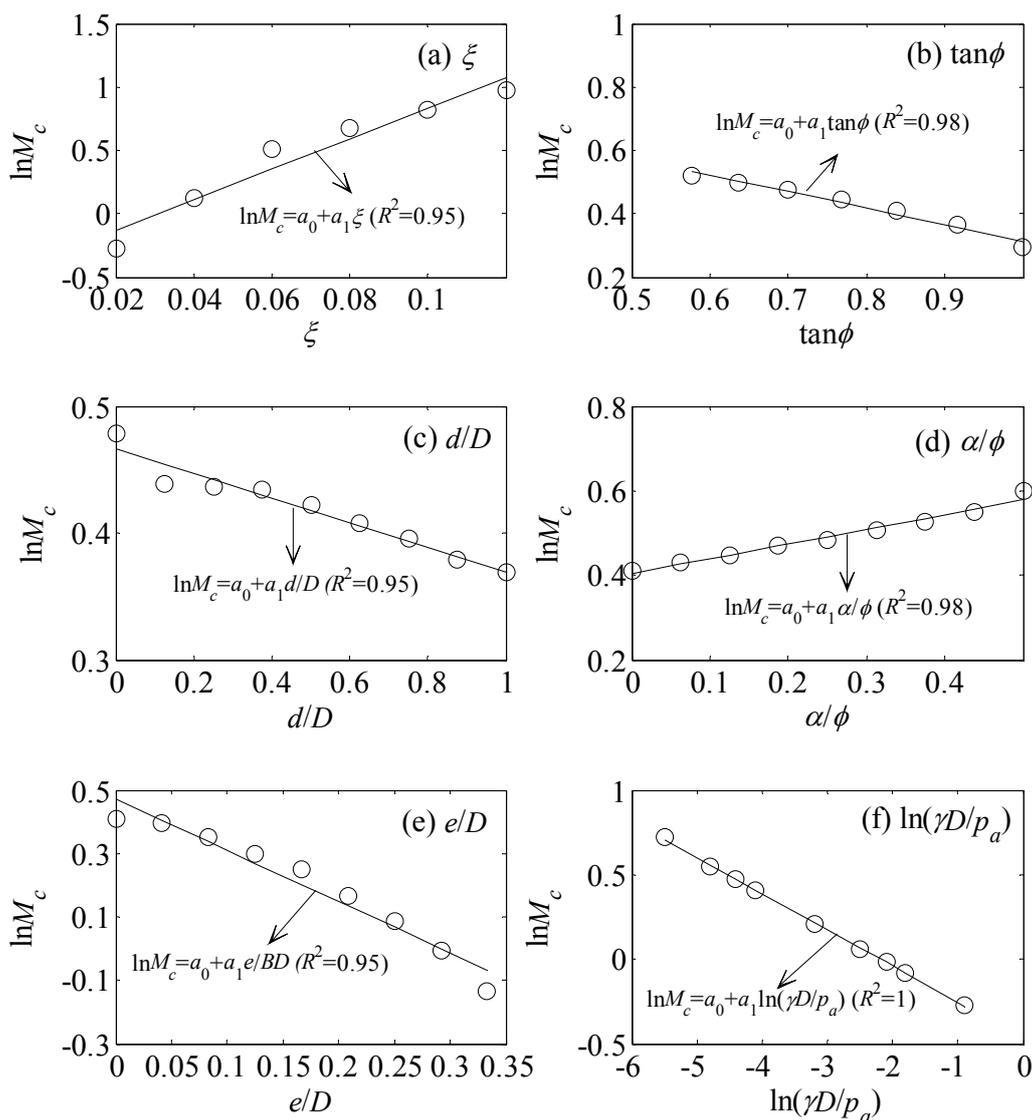


Figure 2-3 Variation of model factor M for Eurocode 7 approach calculating the bearing capacity of strip footings on sand under positive combined loading with input parameters (Phoon and Tang 2017).

Appendix 2A also reveals that the bias is not unique but depends on the type of loading. The research presented in Appendix 2A also discusses the importance of defining an appropriate failure criterion to evaluate the failure load from load-settlement curves.

Appendices 2B and 2C summarize some work on model uncertainty evaluation of settlement predictions.

The work presented in Appendix 2B is based on an extensive database of 426 case histories to

assess the model factor $1/M = X_{cal}/X_{meas}$ (where X is the settlement) for various calculation methods. The model uncertainty itself was considered as a random variable. For those cases where a complete load-settlement curve was available instead of a single response point, the load and settlement values at the elastic limit were used to estimate the accuracy of the settlement estimation methods (see reference in Appendix 2B for definition of elastic limit). The study reveals that the overall accuracy of all methods is very poor with a very high COV and a high percentage of cases in which the calculated settlements exceeds the measured ones. It is also stated that a log-normal distribution can be assumed for the model factor $1/M$.

Appendix 2B also includes results of another study which was performed using a selection of methods for predicting settlements based on SPT test results. The analysis procedure and the main conclusions are similar to the first study. Both studies investigated the influence of the inherent variability of soil properties, construction variabilities, and measurement errors on the model factor statistics related to instrumented structures and load tests.

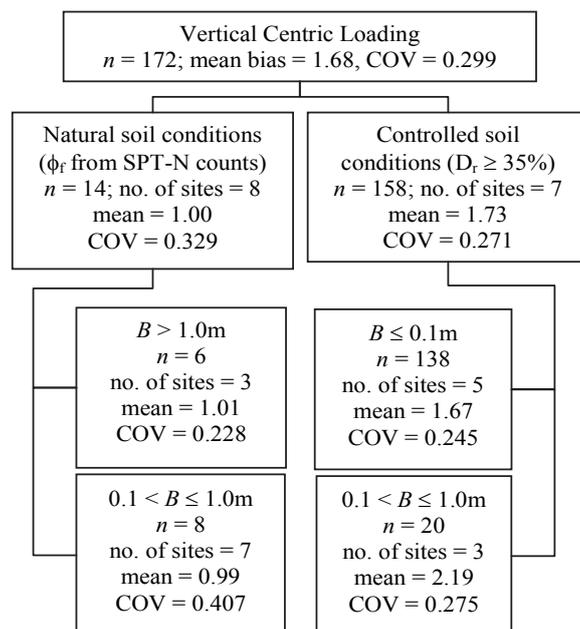


Figure 2-4 Biases of bearing capacity prediction for footings under vertical-centric loading differentiated according to soil conditions and model scale (Paikowsky et al. 2010).

Appendix 2C uses the same definition of the model factor as in Appendix 2B, using the variable $R = 1/M = X_{cal}/X_{meas}$ (where X is the settlement), i.e. the reciprocal of Eq. 2-1. In the presented study a rigid footing under pure vertical loading was considered. To interpret the model uncertainty the following quantities were introduced:

- (1) Accuracy: average value of R for all the cases in a database
- (2) Reliability: percentage of the cases for which $R \geq 1$

These quantities were combined in the so-called ranking distance RD:

$$RD = \sqrt{[1 - \text{mean}(R)]^2 + [\text{SD}(R)]^2} \quad (2-3)$$

The ranking distance considers the mean and the standard deviation (SD) of the model and can be visualized as a beam from the origin of the safe region, i.e. from the ideal value defined by a mean of 1.0 and a standard deviation of 0.0 as displayed in Figure 2-5.

In this figure, the deviation (bias) of the calculated mean of a dataset (shown as bullets) to its ideal value of 1.0 describes the accuracy of the applied design method whereas a standard deviation greater than zero describes its scatter around the mean value, thus its precision. This provides an assessment of the uncertainties related to the derived model factor itself which is of special

importance also if constant model factors finally need to be established for application in LRFD as outlined before in section 2.2.1 (see also Figure 2-1). Further on, entropy and the relative entropy (like the variance) are introduced to measure the dispersion of the R value.

Appendix 2C then presents a procedure which is based on a study of nine settlement prediction methods that all used SPT data to derive soil parameters. In this procedure, curves of different levels of probability for settlement target values can be drawn for the different formulas. From these curves the engineer may derive the exceedance probability of a certain value of the measured settlement, provided that the settlement value is known.

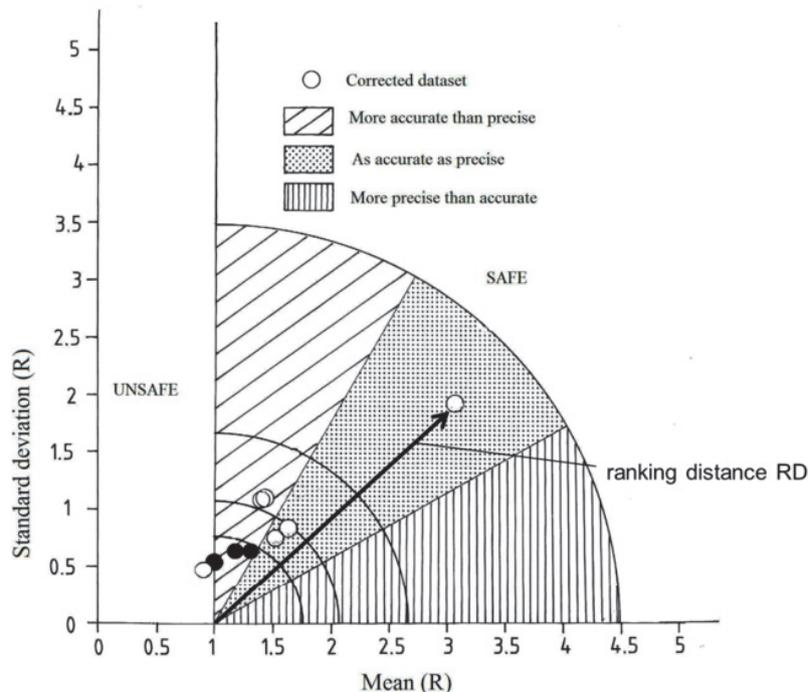


Figure 2-5 Mean versus standard deviation plot to be used in reliability based shallow foundation design with respect to settlement (see Appendix 2C).

2.3.2 Axially loaded pile foundations

Appendix 2D addresses model uncertainties of the bearing capacity prediction of axially loaded piles. Based on Bauduin (2002, 2003) it summarizes typical procedures to evaluate model uncertainties which are commonly expressed by Eq. 2-1. In this regard, Figure 2-6 highlights significant uncertainty sources within pile design.

Appendix 2D states that model uncertainty prediction for piles shall be founded on databases of static pile load tests which are especially useful also for analysing new pile types or calculation models. It is emphasized that the quality of the input data is of significant importance. In this regard the quality of site investigations (field or laboratory tests) and the test performance (testing depth, test load, measured quantities, e.g. separate measurement of shaft friction and end bearing resistance) are discussed. Appendix 2D also addresses the problem of load tests not performed to failure and the possibilities and limitations of applying extrapolation techniques to determine the failure load in the ultimate limit state (ULS) from a load-settlement curve of non-failed tests.

2.3.3 Retaining structures

Hsiao et al. (2008) evaluated the ground settlement induced by neighbouring excavations. Similar to the study of Teixeira et al. (2012) the model uncertainty was considered by a model factor according to Eq. 2-1 applied to the settlement. The model factor was introduced as a random variable in a reliability analysis using FORM with statistical parameters adopted from literature. The analysis

showed that the calculated settlement was highly sensitive to the model factor.

In this study, the model factor was later updated by back-calculation from on-site measurements during several construction stages using the Bayesian updating technique, so that the calculated ground settlement finally approached the measured one. By doing so, the updated model factor reflects all other influences from the respective site, i.e. it is site-dependent and cannot (necessarily) be transferred to other design situations.

A similar study was performed by Juang et al. (2012) who evaluated the damage potential of neighbouring buildings due to excavation activities. The so-called damage potential index (DPI) was derived from the angular distortion and the lateral strain in the building due to excavation (DPI load) which was compared to a limiting value (DPI resistance). Model factors were applied to both, DPI load and DPI resistance, and were assumed to be log-normally distributed. The DPIs were derived from on-site measurements. Hence, model uncertainties only reflect the current status of information for the specific case. Consequently, Juang et al. (2012) used the term apparent model factor. This study also showed a high sensitivity of the reliability analysis for model uncertainties.

A lot of other work could be analysed in order to give additional information. For example, the syntheses provided by Long (2001) or Moormann (2004) include many information about the behaviour of retaining walls. As outlined before the main difficulty remains the representative quantity to be considered, i.e. the maximum wall deflection, the wall deflection at the top of the retaining wall or any other.

2.4 USE OF DATABASES

In general, definition and evaluation of model uncertainties must be seen within the context of the applied procedure and the related complexity (RBD, LRFD – research work or practical application). In this regard, comprehensive data bases with well documented field and laboratory test data can be a good tool to evaluate model uncertainties if it is possible to express them by a model factor or bias as in Eq. 2-1. Therefore, this procedure is relatively popular and has been frequently used before. Phoon and Kulhawy (2005) give a short overview about using databases for model uncertainty evaluation.

Bauduin (2002) distinguishes three approaches to use databases in model uncertainty prediction with reference to pile design (see also Appendix 2D) leading to different quality levels in the derived model factors:

- (1) very high level of detail in the analysis, each type of pile can be calibrated for specific soil conditions
- (2) lower level of detail grouping similar piles and soil conditions together
- (3) larger generalization, treating all data as one sample, resulting in one calibration factor for the calculation method

Databases typically include model and/or prototype field or laboratory tests under natural or controlled soil conditions for a specific design problem. Phoon and Kulhawy (2005) emphasized the use of scaled laboratory tests under controlled conditions (i.e. controlled loading, controlled preparation of uniform soil bed etc.) as the uncertainties resulting from the soil characteristics can be minimized. However, they further stated that such kind of tests possibly do not lead to representative mean model factors as they are not free of other extraneous uncertainties as well.

Field tests such as pile load tests have the advantage that they are conducted in real and diverse site conditions in full scale (no scale effects). But as mentioned above they are more or less affected especially by the spatial soil variability. Hence, only a combination of field and laboratory model tests allows reliable predictions of model uncertainties.

The main problem of databases often is the limited number of tests coming from very different sources each covering only a limited range of possible design situations. One needs to keep in mind that the application of the derived model uncertainty beyond the boundaries of the given database must be verified (ideally by more appropriate tests).

On the other hand, this requirement also depends on the complexity of the design and the degree of simplification implied in the respective design method, which determines the transferability to other boundary conditions. According to Phoon and Ching (2015) a sufficient number of tests is especially important also if parameters show influential statistics. This requires a subdivision of the database into segments with available data preferably being evenly distributed among each segment.

Another concern is the validity, applicability and uniqueness of failure criteria for the definition of the ULS (or the serviceability limit state (SLS)) from the test result (e.g. the load-displacement

curve). Especially field tests are often not run to failure. Extrapolation techniques may be considered, but their application is difficult (e.g. Paikowsky and Tolosko 1999; Phoon and Kulhawy 2005). Also in field tests the measured load-displacement curves often do not show a clear peak, i.e. failure is difficult to be interpreted and not all failure criteria lead to consistent results (see e.g. Paikowsky et al. 2010).

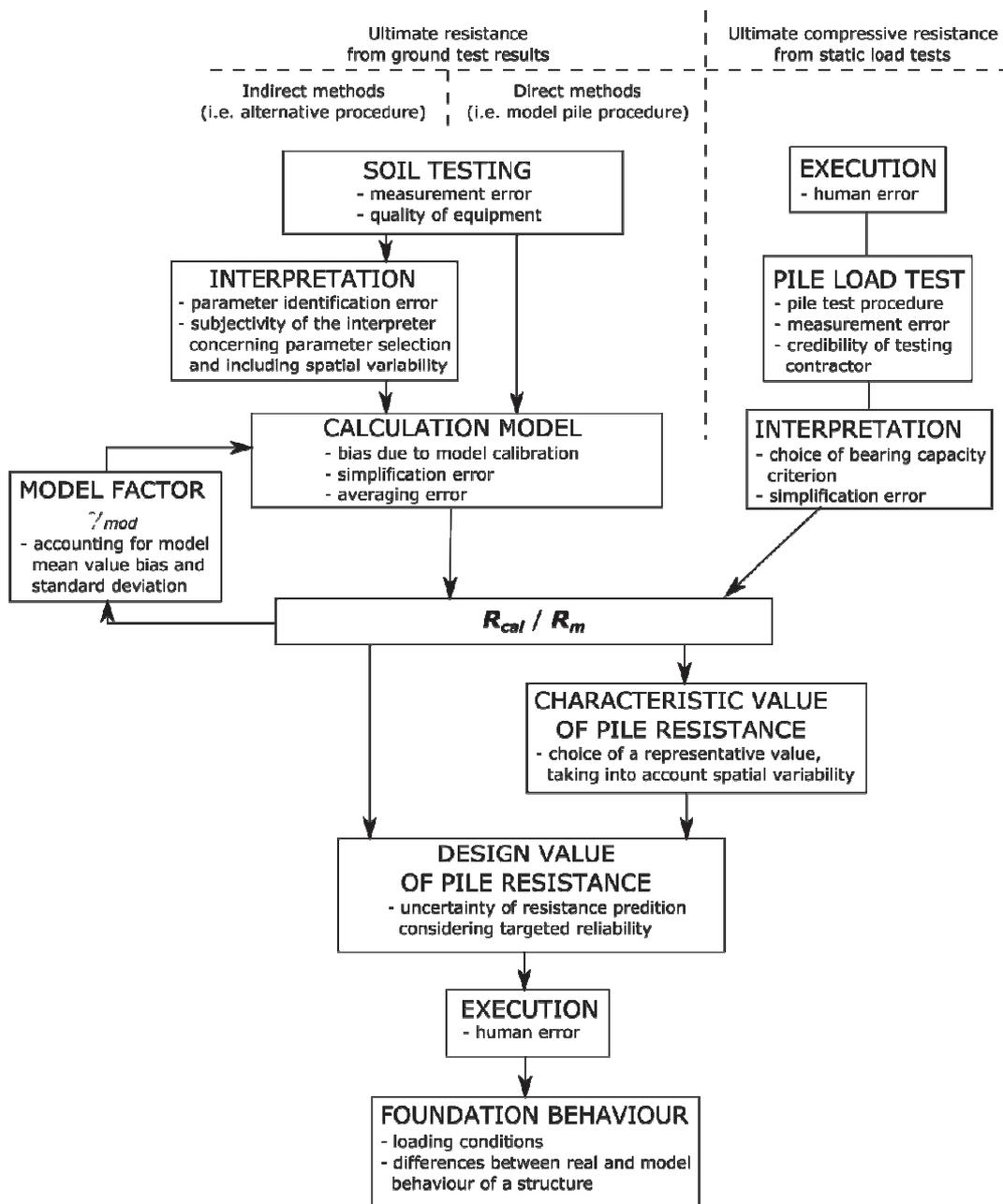


Figure 2-6 Pile design procedure with most significant uncertainty sources (see Appendix 2D).

In summary, the following aspects have been identified to be important for establishing databases that can properly be used for evaluating model uncertainties:

- (1) Availability of data (often propriety concerns)
- (2) Minimum representative number of tests for a specific design constellation
- (3) All geometric and geotechnical ranges according to possible design situations covered
- (4) Test scale (prototype tests and scaled model tests to evaluate scale effects)
- (5) Detailed information on geometry and installation procedure
- (6) Detailed information on applied loading

- (7) Detailed information on soil variability at the test site (especially related to field tests); quality and quantity of available soil information; test type for determination of soil parameters including test analysis and interpretation of results; influence of local or regional soil geology
- (8) Detailed information on performance, processing and evaluation of load tests (in the field, in the laboratory)
- (9) Detailed information on the failure criteria applied to define the ULS/SLS in the test

Other problems affecting the quality of the database input and its use are:

- (1) Effect of nonlinearities in the limit state equations for extrapolation of derived model factors to other design situations
- (2) Uncertainties in the prediction of loading

Furthermore, personal aspects related to the people involved shall be kept in mind. This includes the engineering judgement combined with human error and subjectivity and refers to the experience of the engineer who puts information into a database and later uses the database for evaluation. Here, especially the selection or filtration, the interpretation and analysis of the test results depend on experience and also on the applied accuracy.

2.5 TWO-STEP PROCEDURE TO REMOVE STATISTICAL DEPENDENCIES

Although one may suspect various input parameters to be the explanatory variables behind statistical dependencies between the model factor and the calculated quantity, it is not easy to remove these dependencies in a more physical way by regressing the model factor against each input parameter. This could be primarily attributed to the following two reasons:

- (1) Load tests in a database are usually limited, as stated in section 2.4 titled by “Use of databases”;
- (2) The values of input parameters cannot be varied systematically in the database for regression analysis.

Phoon et al. (2003) proposed a generalized model factor that involves regressing the measured capacity against the calculated capacity. This first-order approach is practical. However, the underlying influence of each input parameter is not explicitly considered, hence it is judicious to apply the generalized model factor only to foundations that are represented in the load test database.

To solve the deficiency in the generalized model factor approach, Zhang et al. (2015) presented a two-step procedure of using a mechanically consistent numerical method such as the finite element method (FEM) to remove the statistical dependencies and then revise the model factor to account for field effects using limited full-scale test data. This framework was recently applied to characterize the model uncertainty in the capacity prediction (Phoon and Tang 2015, 2017; Tang and Phoon 2016, 2017a; Tang et al. 2017a). The difference between these cited work and Zhang et al. (2015) is that FEM was replaced by finite element limit analysis (FELA). The procedure can be summarized as follows:

- (1) Define a correction factor M_c (see Eq. 2-4) as the ratio of the calculated quantity X_p from a more advanced numerical method and the calculated value using a simplified model. Perform regression analysis to remove the statistical dependencies between M_c and input parameters (see Eq. 2-5).

$$X_p = M_c \times X_{cal} \quad (2-4)$$

$$M_c = f \times \eta \quad (2-5)$$

where f is regression equation which is a function of input parameters and η is regression residual.

- (2) Use a calibration database to characterize the model factor M_p (see Eq. 2-6) of the numerical method where fewer idealizations are made and thus model uncertainty always exists. Thus, M_p is usually random (Zhang et al. 2015; Phoon and Tang 2015, 2017; Tang and Phoon 2016, 2017a; and Tang et al. 2017a).

$$X_{meas} = M_p \times X_p \quad (2-6)$$

- (3) Revise the model statistics according to Eq. 2-1 and Eq. 2-4 through Eq. 2-6, namely

$$X_{\text{meas}} = (M_p \times \eta) \times (f \times X_{\text{cal}}) = M' \times X'_{\text{cal}} \quad (2-7)$$

where $M' = M_p \times \eta$ is the modified model factor and $X'_{\text{cal}} = f \times X_{\text{cal}}$ is the modified calculated value. The practical significance of Eq. 2-7 is that the existing calculated models can be improved by multiplying with the regression equation f to remove the dependency on input parameters.

- (4) Verify the statistics of M' using a validation database.

Accordingly, it can be observed that the above procedure provides a practical strategy to derive model statistics, particularly for limited full-scale load tests, which is commonly encountered in geotechnical engineering. Numerical studies have been conducted to demonstrate the effectiveness of this approach (see Appendix 2E):

- (1) Serviceability limit state: calculation of the top deflection of a cantilever wall in undrained clay using MSD (see Zhang et al. 2015);
- (2) Ultimate limit state: calculation of the bearing capacity of strip footings on sand under general combined loading using Eurocode 7 (see Phoon and Tang 2015, 2017);
- (3) Ultimate limit state: calculation of the uplift capacity of helical anchors in undrained clay using the cylindrical shear method (see Tang and Phoon 2016);
- (4) Ultimate limit state: calculation of the bearing capacity of circular footings on dense sand (see Tang and Phoon 2017a);
- (5) Ultimate limit state: calculation of the bearing capacity of large circular and conical foundations on sand overlying clay (see Tang et al. 2017a).

2.6 IMPLEMENTATION IN DESIGN

Phoon and Kulhawy (2005) stated that RBD is a helpful tool in geotechnical design as it ensures self-consistency from physical and probabilistic requirements. They argued that it is a philosophical question if there is the willingness to accept RBD as a necessary basis for calibrating LRFD factors.

In this regard, one should keep in mind that the assessment of model uncertainties related to a particular design method is not only important for the design of the single structure. Moreover, depending on the relevance of this single component for the whole construction it is important for assessing the overall uncertainties of the global design in the decision process.

Once the general concept of RBD has been accepted, the quantification of model uncertainties can be implemented in geotechnical design on different levels of complexity:

- (1) Direct use in RBD as a random variable (e.g. FORM, MCS)
- (2) Use as a constant value in LRFD (e.g. model factor)

For geotechnical engineers the acceptance of such concepts for the daily design practice is strongly related to their practical applicability. This probably supports the use of model factors in LRFD design despite the disadvantages outlined above. In this regard, constant value means that a model factor is at least constant for a certain range of possible design situations.

During the revision process of Eurocode 7 the application of model factors within the limit state equation is currently being discussed. According to this, model factors shall be used to correct calculation models against a reference value, so that the model is either accurate or conservative. They shall be applied to actions, effects of actions, material properties or resistances.

Following the discussion in this report it needs to be emphasized here, that this reference value should be derived from appropriate test results preferably from databases fulfilling the requirements in section 2.4. The reference value shall not (only) be derived from another design method, just because it is assumed to be more sophisticated, thus representing reality, as this method itself may have (other) uncertainties.

As outlined before, a model factor usually cannot be unique in order to cover as many design situations as possible. This will also affect the derivation of the partial factors. Hence, geotechnical engineers must accept a set of factors for different design situations. Nevertheless, one should keep in

mind that the global aim of a constant reliability level for all designs still seems to be difficult if not impossible to reach. This especially applies to complex designs where model uncertainties cannot be reduced to single model factors.

2.7 CONCLUSIONS

The discussion illustrated the importance of evaluating and quantifying model uncertainties in geotechnical design due to the high sensitivity of many design models for this source of uncertainty. Databases of very well documented, high quality field and laboratory tests under diverse site conditions and under controlled laboratory conditions are a good tool for assessing model uncertainties. The main concern related to databases is the availability of a sufficiently large number of tests among the whole design range to be considered. This is especially important as model uncertainties are difficult to separate from other inherent uncertainties such as the soil variability or uncertainties related to the test performance.

The implementation of model factors in LRFD can provide a practical tool for considering model uncertainties in daily design practice. However, one should keep in mind that model factors are strongly related not only to the design method but also to the design situation. The use of single model factors is therefore (often) not correct.

The uncertainties of complex designs are difficult to be reduced to model factors in LRFD designs. In such cases model uncertainty assessment shall be performed by higher levels of RBD using e.g. various case studies for comparison.

2.8 REFERENCES

- Abchir, Z., Burlon, S., Frank, R., Habert, J. and Legrand, S. (2016). t-z curves for piles from pressuremeter test results. *Géotechnique*, 66(2), 137-148.
- Akbas, S. O. and Kulhawy, F. H. (2009). Axial compression of footings in cohesionless soil. I: load-settlement behaviour. *Journal of Geotechnical and Geoenvironmental Engineering*, 135(11), 1562-1574.
- Akbas, S. and Kulhawy, F. H. (2010). Model uncertainties in “Terzaghi and Peck” methods for estimating settlement of footings on sand. *Proceedings of GeoFlorida 2010: Advances in Analysis, Modeling and Design (GSP 199)*, ASCE, Orlando.
- Akbas, S. and Kulhawy, F. H. (2013). Model uncertainties in elasticity-based settlement estimation methods for footings on cohesionless soils. *Proceedings of the International Symposium on Advances in Foundation Engineering*, Singapore.
- Bauduin, C. (2002). Design of axially loaded piles according to Eurocode 7. *Proceedings of the International Conference on Piling and Deep Foundations, DFI 2002*, Nice, Presses de l’ENPC, Paris.
- Bauduin, C. (2003). Assessment of model factors and reliability index for ULS design of pile foundations. *Proceedings of 4th International Geotechnical Seminar on Deep Foundations on Bored and Auger Piles*, 2-4 June, Ghent.
- Burlon, S., Frank, R., Baguelin, F., Habert, J., and Legrand, S. (2014). Model factor for the bearing capacity of piles from pressuremeter test results - Eurocode 7 approach. *Géotechnique*, 64(7), 513-525.
- Cai, G., Liu, S., and Puppala, A. J. (2012). Reliability assessment of CPTU-based pile capacity predictions in soft clay deposits. *Engineering Geology*, 141-142, 84–91.
- Cherubini, C. and Vessia, G. (2005). The bearing capacity of piles evaluated by means of load tests according to reliability calculations. *ICOSSAR 2005*, G. Augusti, G.I. Schuëller, M. Ciampoli (eds), Millpress, Rotterdam.
- Cherubini, C. and Vessia, G. (2009). Probabilistic charts for shallow foundation settlements on granular soil. *Geotechnical Risk and Safety*, eds. Y. Honjo, M. Suzuki, Takashi H., Feng Z., CRC Press, Boca Raton, FL.
- Cherubini, C. and Vessia, G. (2010). Reliability-based pile design in sandy soils by CPT measurements. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 4(1), 2-12.
- Dithinde, M., Phoon, K. K., De Wet, M., & Retief, J. V. (2011). Characterization of model uncertainty in the static pile design formula. *Journal of Geotechnical and Geoenvironmental Engineering*,

137(1), 70-85.

- Dithinde, M. and Retief, J. V. (2013). Pile design practice in southern Africa Part I: Resistance statistics. *Journal of the South African institution of Civil Engineering*, 55(1), 60–71.
- Dithinde, M., Phoon, K. K., Ching, J. Y., Zhang, L. M. & Retief, J. V. (2016). Statistical characterization of model uncertainty. Chapter 5, *Reliability of Geotechnical Structures in ISO2394*, Eds. K. K. Phoon & J. V. Retief, CRC Press/Balkema, 127-158.
- Fenton, G. A., Griffith, D. V., and Williams, M. B. (2005). Reliability of traditional retaining wall design. *Géotechnique*, 55(1), pp. 55-62.
- Hsiao, E. C. L., Schuster, M., Juang, C. H., and Kung, T. C. (2008). Reliability analysis of excavation-induced ground settlement. *Journal of Geotechnical and Geoenvironmental Engineering*, 134(10), 1448-1458.
- Huffman, J. C. and Stuedlein, A.W. (2014). Reliability-based serviceability limit state design of spread footings on aggregate pier reinforced clay. *Journal of Geotechnical and Geoenvironmental Engineering*, 140(10): 04014055, 1-11.
- Huffman, J. C., Strahler, A. W. and Stuedlein, A. W. (2015). Reliability-based serviceability limit state design for immediate settlement of spread footings on clay. *Soils and Foundations*; 55(4), 798–812.
- Juang, C. H., Schuster, M., Ou, C. Y., and Phoon, K. K. (2012). Fully probabilistic framework for evaluating excavation-induced damage potential of adjacent buildings. *Journal of Geotechnical and Geoenvironmental Engineering*, 137(2), 130-139.
- Karlsrud, K. (2014). Ultimate shaft friction and load-displacement response of axially loaded piles in clay based on instrumented pile tests. *Journal of Geotechnical and Geoenvironmental Engineering*, 140(12): 04014074, 1-16.
- Lesny, K. and Paikowsky, S. G. (2011). Developing a LRFD procedure for shallow foundations. *Proceedings of the 3rd International Symposium on Geotechnical Safety and Risk, ISGSR 2011*, eds. N. Vogt, B. Schuppener, D. Straub, G. Bräu, pp. 47-65, Federal Waterways Engineering and Research Institute, Karlsruhe.
- Long, M. (2001). Database for retaining wall and ground movements due to deep excavations. *Journal of Geotechnical and Geoenvironmental Engineering*, 127(3), pp. 203-224.
- Moormann, C. (2004). Analysis of wall and ground movements due to deep excavations in soft soil based on a new worldwide database. *Soils and Foundations*, 44(1), pp. 87-98.
- Nadim, F. (2015). Accounting for uncertainty and variability in geotechnical characterization of offshore sites. In: *Geotechnical Safety and Risk V (ISGSR 2015)*, ed. Schweckendiek, T. et al., pp. 23-35, IOS Press.
- Osman, A. S. and Bolton, M. D. (2004). A new design method for retaining walls in clay. *Canadian Geotechnical Journal*, 41(3), 451–466.
- Paikowsky, S. and Tolosko, T. (1999). Extrapolation of pile capacity from non-failed load tests. *US DOT FHWA-RD-99-170*.
- Paikowsky, S. G., Canniff, M. C., Lesny, K., Kisse, A., Amatya, S., and Muganga, R. (2010). LRFD Design and construction of shallow foundations for highway bridge structures. *NCHRP Report 651*, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C.
- Phoon, K. K. and Kulhawy, F. H. (2003). Evaluation of model uncertainties for reliability-based foundation design. *Proceedings of 9th International Conference on Applications of Statistics and Probability in Civil Engineering*, San Francisco, July 6-9 2003, Vol. 2, 1351-1356.
- Phoon, K. K. and Kulhawy, F. H. (2005). Characterization of model uncertainties for laterally loaded rigid drilled shafts. *Géotechnique*, 55(1), 45-54.
- Phoon, K. K., Chen, J. R., and Kulhawy, F. H. (2006). Characterization of model uncertainties for augered cast-in-place (ACIP) piles under axial compression. *Foundation Analysis & Design: Innovative Methods (GSP 153)*, ASCE, Reston, pp. 82-89.
- Phoon, K. K., Chen, J. R. and Kulhawy, F. H. (2007). Probabilistic hyperbolic models for foundation uplift movements. *Probabilistic Applications in Geotechnical Engineering (GSP 170)*, ASCE, Reston, CDROM.
- Phoon, K. K. and Kulhawy, F. H. (2008). Serviceability Limit State Reliability-based Design. *Reliability-Based Design in Geotechnical Engineering: Computations and Applications*, Taylor & Francis, pp. 344-383.
- Phoon, K. K. and Ching, J. (2015). Is there anything better than LRFD for simplified geotechnical RBD? In: *Geotechnical Safety and Risk V (ISGSR 2015)*, ed. Schweckendiek, T. et al., pp. 3-15,

IOS Press.

- Phoon, K. K., and Tang, C. (2015b). Model uncertainty for the capacity of strip footings under negative and general combined loading. Proc. 12th International Conference on Applications of Statistics and Probability in Civil Engineering, Vancouver, Canada, July 12-15.
- Phoon, K. K., Retief, J. V., Dithinde, M., Schweckendiek, T., Wang, Y., and Zhang, L. M. (2016). Some observations on ISO2394:2015 Annex D (Reliability of Geotechnical Structures). *Structural Safety*, 62, 23-33.
- Phoon, K. K., and Tang, C. (2017). Model uncertainty for the capacity of strip footings under positive combined loading. *Geotechnical Safety and Reliability: Honoring Wilson H. Tang (Geotechnical Special Publication 286)*, edited by C. H. Juang, R. B. Gilbert, L. M. Zhang, J. Zhang, and L. L. Zhang, June 4-7, 2017, Denver, Colorado, pp. 40-60.
- Ronald, K. and Bjerager, P. (1992). Model uncertainty representation in geotechnical reliability. *Journal of Geotechnical Engineering*, 118(3), 363-376.
- Stuedlein, A.W., Reddy, S.C., 2013. Factors affecting the reliability of augered cast-in-place piles in granular soils at the serviceability limit state. *DFI Journal-The Journal of the Deep Foundations Institute*, 7 (2), 46–57.
- Tang, C. and Phoon, K. K. (2016). Model uncertainty of cylindrical shear method for calculating the uplift capacity of helical anchors in clay. *Engineering Geology*, 207, 14-23.
- Tang, C. and Phoon, K. K. (2017a). Model uncertainty of Eurocode 7 approach for bearing capacity of circular footings on dense sand. *International Journal of Geomechanics*, 17(3): 04016069, 1-9.
- Tang, C., Phoon, K. K., Zhang, L., and Li, D.-Q. (2017a). Model uncertainty for predicting the bearing capacity of sand overlying clay. *International Journal of Geomechanics*, 17(7), 04017015: 1-14.
- Tang, C., Phoon, K. K., and Akbas, S. O. (2017b). Model uncertainties for the static design of square foundations on sand under axial compression. *Geo-Risk 2017: Reliability-based design and code developments (Geotechnical Special Publication 283)*, edited by J. S. Huang, G. A. Fenton, L. M. Zhang, and D. V. Griffiths, June 4-7, 2017, Denver, Colorado, pp. 151-160.
- Tang, C., and Phoon, K. K. (2017b). Evaluation and consideration of model uncertainties in reliability-based design of axially loaded helical piles. *Journal of Geotechnical and Geoenvironmental Engineering*, under review.
- Tang, C., and Phoon, K. K. (2017c). Evaluation and consideration of model uncertainties in reliability-based serviceability limit state design of steel H-piles in axial compression. *Canadian Geotechnical Journal*, under revision.
- Teixeira, A., Honjo, Y., Gomes Correia, A., and Abel Henriques, A. (2012). Sensitivity analysis of vertically loaded pile reliability. *Soils and Foundations*, 52(6), pp. 1118-1129.
- Zhang, J. (2009). Characterizing geotechnical model uncertainty. PhD Thesis, The Hong Kong University of Science and Technology, Hong Kong.
- Zhang, D.M., Phoon, K.K., Huang, H.W., and Hu, Q.F. (2015). Characterization of model uncertainty for cantilever deflections in undrained clay. *Journal of Geotechnical and Geoenvironmental Engineering*, 141(1): 04014088, 1-14.

Appendix 2A: Model uncertainty of bearing capacity analysis for shallow foundations (K. Lesny)

The following summary is based on the references given below.

The model uncertainty of ULS bearing capacity analysis has been evaluated on the base of a comprehensive database of load tests on shallow foundations. The database contained 549 cases of load tests compiled from various publications and from own model test data, for details see references below. Table 2A-1 summarized the test data.

Most cases in the database are related to foundations subjected to vertical-centric loading in or on granular soils. Tests of foundations subjected to combined loadings (vertical-eccentric, inclined-centric and inclined-eccentric) were mainly small scale model tests performed in controlled soil conditions (in laboratories using soils of known particle size and controlled compaction).

Table 2A-1 Summary of model tests in the database.

Foundation type	Predominant soil type				Total
	Sand	Gravel	Mix	Others	
Plate load tests $B \leq 1\text{m}$	346	46	2	72	466
Small footings $1 < B \leq 3\text{m}$	26	2	4	1	33
Large footings $3 < B \leq 6\text{m}$	30	--	1	--	31
Rafts & Mats $B > 6\text{m}$	13	--	5	1	19
Total	415	48	12	74	549

Notes:

“Mix”: alternating layers of sand or gravel and clay or silt

“Others”: either unknown soil types or other granular materials like loamy Scoria

To define the ULS failure load of each test several failure criteria were examined:

- (1) Minimum slope failure criterion by Vesić (1975)
- (2) Log-log load-settlement curve method by de Beer (1967)
- (3) Two-slope criterion described in NAVFAC (1986)

For most test results the corresponding load-settlement curves did not show a clear peak, hence interpretation of the failure loads was difficult. In addition, many load tests were not carried out to failure, making them unsuitable for the analysis. This was especially the case for larger foundations for which failure would be associated with very large loads and excessive displacements. It was finally found that the minimum slope failure criterion provided the most consistent interpretation when establishing the measured bearing capacity from the load tests.

For establishing the calculated bearing capacity the basic bearing capacity equation by Vesić (1975) was used. Numerous analytical expressions for the different factors (bearing capacity, shape, and depth and inclination factors) were analyzed to find the most consistent expressions.

Uncertainties of this design method were expressed as a bias, i.e. the ratio of measured over calculated bearing capacity, including all sources of uncertainties such as scale effects, variation of soil properties and their interpretation, capacity interpretation etc. Biases were studied according to loading types (vertical-centric, vertical-eccentric, inclined-centric and inclined-eccentric) and soil conditions (natural (field) and controlled (laboratory) soil conditions). Figure 2A-1 shows the biases for vertical-centric loading.

Figure 2A-1 reveals that there might be a relation between the footing size and the bias. Laboratory small scale model tests (very small footings tested in controlled soil conditions, general failure) show a larger bias than larger footings especially in natural conditions where general shear failure is not always reached. However, this is superposed by the test conditions (natural versus

controlled soil conditions, i.e. soil variability, measurement errors etc.), so the differences in the bias cannot only be attributed to scale effects.

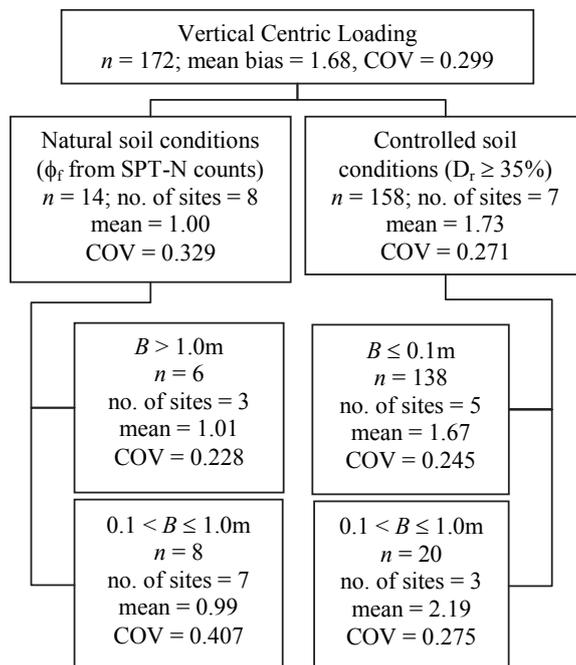


Figure 2A-1 Summary of biases for footings under vertical-centric loading differentiated according to soil conditions and model scale (Paikowsky et al. 2010).

For vertical-eccentric loading cases, the bias had a mean of 1.81 and a COV of 0.349. For inclined-centric loading cases, the bias had a mean of 1.43 and a COV of 0.295. Inclined-eccentric loading cases were distinguished into positive or reversible and negative moment, which also affect the bias (positive moment: mean = 1.41, COV = 0.278, negative moment: mean = 2.03, COV = 0.094).

The results show the conservatism in the design method, the influence of the site (soil conditions), some influence of the model scale and a very strong influence of the loading situation.

Based on the statistical parameters for the bias in different design situations, different resistance factors for different loading conditions were calibrated for a target reliability of $\beta_T = 3.0$ and lognormal distributions for loads and resistance from calculations using the first order second moment method (FOSM) and MCS.

References

- Lesny, K. and Paikowsky, S.G. (2011). Developing a LRFD procedure for shallow foundations. Proceedings of the 3rd International Symposium on Geotechnical Safety and Risk, ISGSR 2011, eds. N. Vogt, B. Schuppener, D. Straub, G. Bräu, pp. 47-65, Federal Waterways Engineering and Research Institute, Karlsruhe.
- Paikowsky, S.G., Canniff, M.C., Lesny, K., Kisse, A., Amatya, S, and Muganga, R. (2010). LRFD design and construction of shallow foundations for highway bridge structures. NCHRP Report 651, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C.
- Paikowsky, S.G., Amatya, S, Lesny, K., and Kisse, A. (2009). LRFD design specifications for bridge shallow foundations. Proceedings of the 2nd International Symposium on Geotechnical Safety and Risk - IS-GIFU2009 – June 11-12, Gifu, Japan.

Appendix 2B: Model uncertainty of bearing capacity and settlement analysis for shallow foundations on cohesionless soils (S. O. Akbas)

The following summary is based on several published studies:

Bearing capacity of shallow foundations on cohesionless soils

Summary from: Akbas, S.O. and Kulhawy, F.H. (2009). *Axial compression of footings in cohesionless soils. II: Bearing capacity. Journal of Geotechnical and Geoenvironmental Engineering, 135(11), pp. 1575-1582.*

An extensive database of full-scale field load tests was used to examine the bearing capacity of footings in cohesionless soils. This database summarizes published case histories from 37 sites with 167 axial compression field load tests on footings conducted in cohesionless soils ranging from silt to gravel. The case histories were categorized into three groups based on the quality of the load test data. The range and mean of the geometrical properties and effective stress friction angles, and the available in situ test results, are summarized in Table 2B-1.

Table 2B-1 Summary of some geometrical and geotechnical properties and availability of in situ test results.

Data group	No. load tests	Footing width B (m)		Footing length L (m)		Footing depth D (m)		$\bar{\phi}_{tc}$		In situ tests			
		Mean	Range	Mean	Range	Mean	Range	Mean	Range	SPT only	CPT only	SPT and CPT	Neither
1	97	0.70	0.25–2.49	0.74	0.25–2.50	0.22	0.00–1.00	37.6	28.0–55.0	35	3	40	19
2	28	0.86	0.30–3.02	0.93	0.30–3.02	0.19	0.00–1.04	38.1	32.0–50.0	13	—	15	—
3	42	1.05	0.30–4.00	1.06	0.30–4.00	0.20	0.00–2.50	39.9	30.0–53.0	11	—	6	25
All	167	0.81	0.25–4.00	0.85	0.25–4.00	0.18	0.00–2.50	38.1	28.0–55.0	59	3	61	44

Each load test curve was evaluated consistently to determine the interpreted failure load, i.e., bearing capacity, using the L_1 - L_2 method. This test value then was compared with the theoretical bearing capacity, computed primarily using the basic Vesic model. The predicted bearing capacity values (Q_{tcp}) are plotted versus the measured Q_{L2} in Figure 2B-1 for all of the data.

This figure shows that the predicted bearing capacity generally is underestimated for Q_{L2} less than about 1,000 kN, and the difference between Q_{tcp} and Q_{L2} generally increases with decreasing Q_{L2} . This trend suggests a possible relationship between the footing width B and Q_{tcp}/Q_{L2} , or bias, i.e. the ratio of calculated over measured bearing capacity, since the bearing capacity is expected to increase with B , for soils of comparable strength.

yll of the further analyses of the data showed an increase in the ratio of predicted-to-measured bearing capacity with increasing B , up to about $B = 1$ m. The comparisons show that, for footing widths $B > 1$ m, the field results agree quite well with Vesic’ predictions.

Settlements of shallow foundations on cohesionless soils - elasticity-based methods

Summary from: Akbas, S.O. and Kulhawy, F.H. (2013). *Model uncertainties in elasticity-based Settlement Estimation Methods for Footings on Cohesionless Soils. International Symposium on Advances in Foundation Engineering (ISAFE 2013).*

The empirical nature of these various elasticity-based methods suggests a significant uncertainty in the estimated settlements. To assess this uncertainty, an extensive database of 426 case histories was used to assess the model factor (i.e. the ratio of calculated to measured settlement) for the methods developed by D’Appolonia et al. (1970), Parry (1971), Schultze and Sherif (1973), Schmertmann et al. (1978), and Berardi and Lancellotta (1991). Considering the model uncertainty as a random variable, the uncertainty in the model factors is also characterized using the coefficient of variation (COV).

The database includes foundations of various sizes, from small test plates 0.25 m wide to mat

foundations up to 135 m wide. The structures corresponding to these foundations include bridges, test footings, buildings, tanks, embankments, chimneys, nuclear reactors, and silos. The details of all available geometric and geotechnical data are provided in Akbas (2007). For those cases where a complete load-settlement curve is available instead of a single response point, the load and settlement values at the elastic limit (L_1), Q_{L1} and ρ_{L1} , were used to estimate the accuracy of the settlement estimation methods. The elastic limit is described elsewhere (e.g., Akbas 2007, Akbas and Kulhawy 2009).

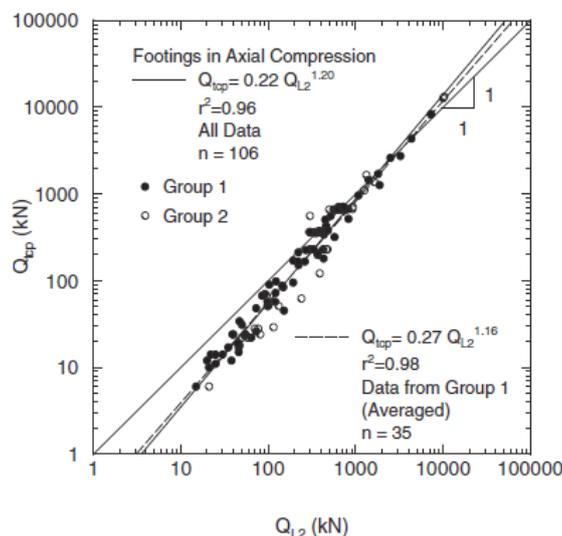


Figure 2B-1 Predicted (Q_{top}) versus measured (Q_{L2}) bearing capacity for footings in drained axial compression (all data and Group 1 data averaged per site).

Results for the five selected methods are summarized in Table 2B-2 and include the mean, maximum, minimum, COV, and percent exceedance. It is clear that there is much uncertainty for all of the methods. The mean ρ_c/ρ_m varies from 1.05 to 2.00, which is a much smaller range than that obtained for Terzaghi and Peck (1948) based methods (Akbas and Kulhawy 2010). Schultze and Sherif (1973) is the most accurate (lowest mean) and the least conservative (underestimates most, in 61% of the cases), while Schmertmann et al. (1978) is the least accurate (highest mean) or most conservative (overestimates most, in 71% of the cases). In general, the results suggest there is a trade-off between accuracy and conservatism of the settlement estimation methods, i.e., as the mean ρ_c/ρ_m approaches one, the number of underestimated cases increases. The COVs, which are indications for statistical precision, range between 70 and 94%, and these high values are consistent with previous research (e.g. Akbas and Kulhawy 2010). These values correspond to a narrow range of standard deviation values between 0.99 and 1.09 for D’Appolonia et al. (1970), Schultze and Sherif (1973), and Parry (1971), while Schmertmann et al. (1978) has the highest standard deviation of 1.40.

The following definition is used often to address model uncertainties (e.g. Phoon and Kulhawy 2003):

$$\rho_c = M\rho_m \quad (2B-1)$$

in which M is model factor, which usually is assumed to be a lognormally distributed random variable.

The ratios of calculated to measured settlements, or the model factors, were shown in Table 2B-2 for the statistical analysis of the 426 case histories. Using the same database, it was shown by Akbas and Kulhawy (2010) that very slight reduction of model factor COVs are obtained when data from 183 load tests only are used, excluding the remaining 243 observations on instrumented structures. A similar observation was also made by Phoon and Kulhawy (2005) for the model factors that correspond to the lateral capacity of free-head drilled shafts. Therefore, it was assumed that the extraneous uncertainties involved in the estimated model factors would be equal to the corresponding lower end values for measurement errors and inherent variabilities that were specified in Phoon et al.

(1995). To incorporate model uncertainties in reliability-based design (RBD), the assumption of lognormality of ρ_c/ρ_m or M must be established. For the five settlement methods, $\ln M$ is plotted on normal probability plots. From inspection, it was seen that lognormality is a reasonable assumption for the distribution of M . Also, the obtained P values indicate that the null hypothesis of normality for $\ln M$ cannot be rejected at a 5% level of significance, except for the D’Appolonia et al. (1970) method.

Table 2B-2 Relationship between calculated and measured settlements.

Method	(ρ_c/ρ_m)			COV	Exceedance ($\rho_c > \rho_m$)
	Mean	Max.	Min.		
D’Appolonia et al. (1970)	1.45	12.8	0.23	73%	66%
Schultze & Sherif (1973)	1.05	8.34	0.04	94%	39%
Parry (1971)	1.21	8.04	0.09	90%	51%
Schmertmann et al. (1978)	2.00	11.2	0.17	70%	71%
Berardi & Lancelotta (1991)	1.43	6.34	0.12	89%	52%

Further analyses indicate that, the overall accuracy of all methods is low, with no method being capable of estimating settlements within 5 mm of the measured settlements with a probability higher than about 10%.

Settlements of shallow foundations on cohesionless soils - Terzaghi and Peck-based methods

Summary from: Akbas, S.O. and Kulhawy, F.H. (2010). *Model Uncertainties in “Terzaghi and Peck” Methods for Estimating Settlement of Footings on Sand. GeoFlorida: Advances in Analysis, Modeling and Design (GSP 199)*.

The Terzaghi and Peck (1948) method was the first for predicting the settlement of footings on sand using standard penetration test blow counts (N values). Over the following years, various modifications to this basic method were suggested. An extensive database of 426 settlement case histories is used to assess the model factors for this family of methods.

As in many previous studies, the methods of estimating settlements are evaluated first by comparing the calculated (ρ_c) and measured (ρ_m) settlements. Statistically, an accurate method would be one that yields a mean close to or equal to 1.0 for a set of values of the ratio of measured to calculated settlements or vice-versa. The relationships between the calculated and measured settlements for the six selected methods are summarized in Table 2B-3 and include the mean, maximum, minimum, coefficient of variation (COV), and percent exceedance. It is clear that there is much uncertainty for all of the methods. The mean calculated to measured settlement (ρ_c/ρ_m) varies from 1.36 to 3.67. Gibbs and Holtz (1957) is the most accurate (lowest mean), while Terzaghi and Peck (1948) is the least accurate (highest mean) or most conservative (overestimates most, in 93% of the cases). The least conservative methods are those of Gibbs and Holtz (1957) and Alpan (1964), which both overestimated 58% of the cases. In general, the results indicate that there is a trade-off between the accuracy and the conservatism of the settlement estimation methods, i.e., as the mean (ρ_c/ρ_m) approaches one, the number of underestimated cases increases. The same observation was made by Tan and Duncan (1991).

As shown, the COVs range from 69 to 103%. These high values are consistent with previous research (e.g., Jeyapalan and Boehm 1986). Note that the Gibbs and Holtz (1957) and Terzaghi and Peck (1948) methods have the lowest COV (most precise), with a value of 69%. The Peck et al. (1974) and Alpan (1964) methods have the highest COVs (least precise), at 103 and 102%, respectively.

Because of the inherent variability of soil properties, construction variabilities, and measurement errors, it is found reasonable to assume that the model factor statistics are lumped statistics that include the model uncertainties and some extraneous uncertainties, especially for the instrumented structures and, to a lesser extent, for the load tests. To test this argument, a simple comparison was made between the model factors from the load tests only, which includes 183 tests, and from the whole database. The results are shown in Table 2B-4. Except for the Alpan (1964) method, the uncertainties decreased slightly when only data from the load tests were considered.

Table 2B-3 Relationships between calculated and measured settlements.

Method	(ρ_c/ρ_m)				Exceedance ($\rho_c > \rho_m$)
	Mean	Max.	Min.	COV	
Terzaghi & Peck (1948)	3.67	19.4	0.13	69%	93%
Gibbs & Holtz (1957)	1.36	9.78	0.11	69%	58%
Alpan (1964)	1.67	16.1	0.13	102%	58%
Meyerhof (1965)	2.34	13.9	0.14	95%	71%
Peck & Bazaraa (1969)	1.53	8.06	0.10	76%	62%
Peck et al. (1974)	2.70	14.3	0.22	103%	86%

Table 2B-4 Comparison of model factors for load tests and whole database.

Method	Load Tests Only (n = 183)		All Data (n = 426)	
	M = ρ_c/ρ_m	COV _M	M = ρ_c/ρ_m	COV _M
Terzaghi & Peck (1948)	3.70	66%	3.67	69%
Gibbs & Holtz (1957)	1.40	62%	1.36	69%
Alpan (1964)	1.65	103%	1.67	102%
Meyerhof (1965)	2.51	88%	2.34	95%
Peck & Bazaraa (1969)	1.54	73%	1.53	76%
Peck et al. (1974)	2.62	96%	2.70	103%

The decrease in COV ranges from 3 to 7%. Therefore, when using a database that includes load tests (field and lab) and field measurements on actual structures, extraneous uncertainties should be considered in estimating model uncertainties for RBD calculations. However, considering the Phoon and Kulhawy (2005) results, the effect of normalization, and the likelihood that the equipment and procedural control in documented and monitored cases would be higher than those in average construction, it can be assumed that the effect of extraneous uncertainties on the estimated model factors should be minimal, especially when only the highest quality data are used. This assumption also is supported by the relatively small differences in the COV values in Table 2. Therefore, it was deduced that the extraneous uncertainties involved in the estimated model factors would be equal to the corresponding lower end values for measurement errors and inherent variabilities that were specified in Phoon et al. (1995). It was also determined that for each method, model factors can be effectively modeled as lognormal random variables.

Based on the properties of the lognormal distribution, an attempt was made to estimate model factors that are free from extraneous uncertainties. Then, for a design settlement of 25 mm, the model factors were used to calculate the probability of failure, the reliability index (β), and the probability that the measured settlement will be between 0 and 10 mm, 10 and 25 mm, and 20 and 25 mm for the six methods considered, using the best-case and worst-case scenarios, respectively. The results indicate that the probability that the settlements would be within 5 mm of the targeted settlement ranges between only about 4 and 12%, even for the best-case scenario. A comparison of the estimated probabilities and reliability indices for the best-case and worst-case scenarios shows that the uncertainties from the N value and loading have only minor effect on the results compared to the model uncertainties.

References

- Akbas, S.O. (2007). Deterministic and probabilistic assessment of settlements of shallow foundations in cohesionless soils. PhD Thesis, Cornell Univ, Ithaca, NY.
- Akbas, S.O. and Kulhawy, F.H. (2009). Axial compression of footings in cohesionless soils. II: bearing capacity. *Journal of Geotechnical and Geoenvironmental Engineering*, 135(11), pp. 1575-1582.
- Akbas, S.O. and Kulhawy, F.H. (2010). Model uncertainties in “Terzaghi and Peck” methods for estimating settlement of footings on sand. *GeoFlorida: Advances in Analysis, Modeling and*

- Design (GSP 199), pp. 2093-2102.
- Akbas, S.O. and Kulhawy, F.H. (2013). Model uncertainties in elasticity-based settlement estimation methods for footings on cohesionless soils. International Symposium on Advances in Foundation Engineering (ISAFE 2013).
- Alpan, I. (1964). Estimating settlements of foundations on sands. Civil Engineering and Public Works Rev. 59(700), pp. 1415-1418.
- Berardi, B. and Lancellotta, R. (1991). Stiffness of granular soils from field performance. *Géotechnique*, 41(1), pp. 149-157.
- D'Appolonia, D.J., D'Appolonia, E.D., and Brissette, R.F. (1970). Closure to: Settlement of spread footings on sand. *Journal of the Soil Mechanics and Foundation Division of the ASCE*, 94(SM3), pp. 1011-1053.
- Gibbs, H. J. and Holtz, W. H. (1957). Research on determining the density of sands by spoon penetration testing. *Proceedings of the 4th International Conference on Soil Mechanics and Foundation Engineering*, London, Vol. 1, pp. 35-39.
- Jeyapalan, J.K. and Boehm, R. (1986). Procedures for predicting settlements in sands. *Proceedings of Settlement of Shallow Foundations on Cohesionless Soils (GSP 5)*, Ed. W. O. Martin, ASCE, New York, pp. 1-22.
- Meyerhof, G.G. (1965). Shallow foundations. *Journal of the Soil Mechanics and Foundation Division of the ASCE*, 91(SM2), pp. 21-31.
- Oweis, I.S. (1979). Equivalent linear model for predicting settlements of sand bases. *Journal of the Geotechnical Engineering Division of the ASCE*, 105(GT12), pp. 1525-1544.
- Parry, R.H.G. (1971). A direct method of estimating settlements in sands from SPT values. *Proceedings of the Symposium on Interaction of Structure and Foundation*, Birmingham, pp. 29-37.
- Peck, R.B. and Bazaraa A.R.S.S. (1969). Discussion on settlement of spread footings on sand. *Journal of the Soil Mechanics and Foundation Division of the ASCE*, 95(SM3), pp. 905-909.
- Phoon, K. K., Kulhawy, F. H., and Grigoriu, M. D. (1995). Reliability-based foundation design for transmission line structures. Rpt TR-105000, EPRI, Palo Alto, 412 p.
- Phoon, K. K. and Kulhawy, F. H. (2003). Evaluation of model uncertainties for reliability-based foundation design. *Proceedings of the 9th International Conference on Applied Statistics and Probability in Civil Engineering*, San Francisco, pp. 1351-1356.
- Phoon, K. K. and Kulhawy, F. H. (2005). Characterization of model uncertainties for laterally loaded rigid drilled shafts. *Géotechnique*, 55(1), pp. 45-54.
- Schmertmann, J. H., Hartman, J. P., and Brown, P. R. (1978). Improved strain influence factor diagrams. *Journal of the Geotechnical Engineering Division of the ASCE*, 104(GT8), pp. 1131-1135.
- Schultze, F. and Sherif, G. (1973). Prediction of settlements from evaluated settlement observations for sand. *Proceedings of the 8th International Conference on Soil Mechanics and Foundation Engineering*, Moscow, pp. 225-230.
- Tan, C. K. and Duncan, J. M. (1991). Settlement of footings on sands: accuracy and reliability. *Geotechnical Engineering Congress, 1991 (GSP27)*, ASCE, New York, pp. 446-455.
- Terzaghi, K. and Peck, R. B. (1948). *Soil Mechanics in Engineering Practice*. Wiley, New York, 566 p

Appendix 2C: Existing work on model uncertainty evaluation for shallow foundation settlement prediction on sandy soils and a new proposal (G. Vessia)

The estimation of the model uncertainty in the calculation of shallow foundation settlements on sandy soils was first addressed by Terzaghi and Peck (1967) that posed that density and compressibility of sandy deposits is often that erratic to be unrealistic to imagine that any method would be capable of estimating actual settlements of footings. Terzaghi and Peck (1967) suggested that, if several identical footings, all equally loaded, were built on the same sand deposit, the ratio between the upper bound and the lower bound of the estimated settlements will be likely to get to 2 or higher values.

Starting from the preceding considerations, the model uncertainty is considered as one value including other sources of uncertainties like (Uzielli and Mayne 2012):

- (1) inherent complexity and uniqueness of geomaterials, foundations and their functional interaction;
- (2) the epistemic uncertainty related to imperfect measurement of the soil-foundation system;
- (3) the transformation uncertainty in the load-displacement model itself.

A few authors devoted explicit efforts to calculate the model uncertainty in settlement predictions on sandy soils.

Tan and Duncan (1988) introduced the ratio R obtained as calculated settlements S_{calc} divided by measured settlements S_{mis} :

$$R = S_{calc} / S_{mis} \quad (2C-1)$$

Thus, databases of R values can be collected. Nevertheless, some statistics are needed to rank the magnitude of model uncertainty of several settlement prediction equations for sandy soils S_{calc} . Then et al. Duncan (1988) introduced two quantities, namely accuracy and reliability. They defined the accuracy as the average value of this ratio for all the cases in the database. A value of this average equal to unity represents the best possible accuracy. The reliability is defined as the percentage of the cases for which R is equal or greater than 1. Perfect reliability is equal to 100. Finally, they used adjustment factors to rank the model uncertainty as a list of ranks from the most to the less efficient settlement prediction equations. In the same period, Briaud and Tucker (1988) introduced a concise index that took into account accuracy and precision, namely “Ranking Index” RI:

$$RI = \mu [\ln(R)] + s [\ln(R)] \quad (2C-2)$$

where μ and s are the mean and the standard deviation of R. The precision represents the dispersion of R values about its mean and the accuracy measures the greater or smaller closeness of a set of measures to the real value. This latter is an indicator of the central trend, as the mean. Later on, Cherubini and Orr (2000) suggested using the “Ranking Distance” RD instead of the RI:

The calibrated bias and COV of a transformation model can be adopted to develop the point estimate and 95% CI, described as follows. Consider again the $LI-(s_u^{re}/P_a)$ model, let the site-specific LI value for the new design site be denoted by LI_{new} , the point estimate for $(s_u^{re}/P_a)_{new}$ is simply $b \times (\text{predicted target value}) = b \times (0.0144 \times LI_{new}^{-2.44})$. By assuming ε to be lognormal, the 95% CI for $(s_u^{re}/P_a)_{new}$ can be expressed as

$$RD = \sqrt{[1 - \mu(R)]^2 + [s(R)]^2} \quad (2C-3)$$

Plotting RD on a Cartesian plane it provides with a geometrical interpretation of the best combination between accuracy and precision among many possible R values (as shown in Figure 2C-1).

Afterwards, Cherubini and Vessia (2009) studied the model uncertainty R of 9 settlement prediction equations for sandy soils that used N_{SPT} measures. The S_{mis} measured settlements used for calculating R values (Eq. 2C-1) come from Burland and Burbidge (1985) database of 192 values from

full-scale shallow foundations of several dimensions set in several sandy soils. Thus, for each equation 192 R values have been calculated and 9 R samples have been collected whose statistics have been calculated.

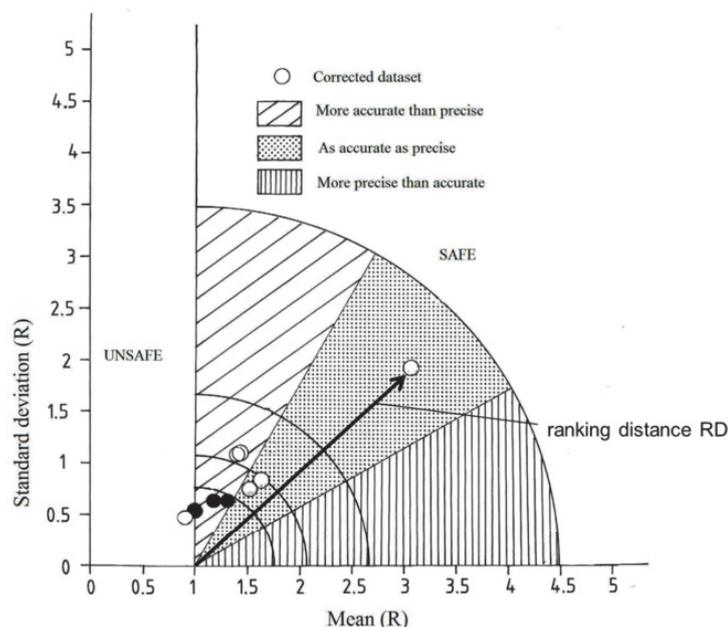


Figure 2C-1 Plane showing both mean values (on the x axis) and standard deviations (on the y axis) of R for the nine settlement formulas (modified from Cherubini and Vessia, 2009).

As the settlement prediction equations are concerned, the chosen nine equations are introduced below (all the references can be found in Cherubini and Vessia 2009), according to the following notations:

- B (m) least width of a rectangular foundation or diameter of a circular foundation;
- D (m) depth of footing embedment below ground surface;
- E (kPa) modulus of soil stiffness;
- H (m) depth of the incompressible soil;
- H_s (m) thickness of the compressible layer below the foundation;
- I influence factor for computing settlement from elasticity theory;
- q (kPa) net increase of the effective pressure at foundation level;
- S_c (mm) calculated settlement;
- N_{SPT} number of blow counts in an SPT;
- σ'_{v0} (kPa) overburden effective pressure at depth z.

- Terzaghi and Peck (1948)

$$S_c = 3.18 \frac{q}{N_{SPT}} \left(\frac{B}{B + B_0} \right)^2 C_d C_w \quad (2C-4)$$

where the depth coefficient C_d is assumed to be 1 and for $D = 0$, 0.75 for $D \geq B$ (Jordan, 1977), for $0 < D < B$ it is linearly interpolated as $C_d = 1 - 0.25D/B$ (Pasqualini, 1983). The water table coefficient, C_w , is assumed equal to 1 for $D_w \geq 2B$, to 2 for $D_w \leq 0$, for $0 < D_w < 2B$ it is linearly interpolated as $C_d = 2 - 0.5D_w/B$.

- Meyerhof (1965)

$$S_c = 2.12 \frac{q}{N_{SPT}} \left(\frac{B}{B+B_0} \right)^2 C_d \quad (2C-5)$$

where C_d has the same meaning as in Terzaghi and Peck’s method (1948).

- Meigh and Hobbs (1975)

$$S_c = 3.18 \frac{q}{N_{SPT}} \left(\frac{B}{B+B_0} \right)^2 C_d \frac{4}{(q_c/N_{SPT})} \quad (2C-6)$$

where for C_d the authors refer to Terzaghi and Peck’s method and the ratio q_c/N_{SPT} depends on the grain-size distribution of soil.

- Arnold (1980)

$$S_c = \frac{43.065 B \alpha}{[1 + (3.28 B)^m]^2} H_s \ln \left(\frac{Q}{Q - 0.5q} \right) \quad (2C-7)$$

where $H_s = \min(2B, H-D)$ which is the depth of the compressible layer below the foundation (in m), $\alpha = 0.032766 - 0.0002134D_R$, $Q = 19.63D_R - 263.3$, $m = 0.788 + 0.0025D_R$, and $D_R =$ relative density is given by $D_R(\%) = 25.6 + 20.37 \times [1.26 \times (N_{SPT} - 2.4) / (0.0208 \sigma'_{v0, D+B/2} + 1.36) - 1]^{0.5}$ for $N_{SPT} \geq 6$ and $45\% \leq D_R \leq 100\%$.

- Burland and Burbidge (1985)

$$S_c = 1.71 \cdot \left(q_{gross} - \frac{2}{3} \sigma'_{v0,D} \right) \frac{B^{0.7}}{(N_{SPT})_{AV}^{1.4}} \left(\frac{1.25L}{L + 0.25b} \right)^2 \frac{H_s}{Z} \left(2 - \frac{H_s}{Z} \right) \quad (2C-8)$$

where q_{gross} is the gross bearing pressure at foundation level (kPa), Z is the depth of influence for N_{SPT} averaging that is $(N_{SPT})_{AV}$. If N_{SPT} values decrease with depth, Z is equal to $2B$. Otherwise, it is calculated by the following interpolating equation: $Z = 0.933B^{0.779}$.

- Anagnostopoulos et al. (1991)

These authors suggest two settlement formulas:
Formula 1:

$$S_c \begin{cases} 0.57q^{0.94} B^{0.90} N_{SPT}^{-0.87} & \text{for } 0 < N_{SPT} \leq 10 \\ 0.35q^{1.01} B^{0.69} N_{SPT}^{-0.94} & \text{for } 10 < N_{SPT} \leq 30 \\ 604q^{0.9} B^{0.76} N_{SPT}^{-0.82} & \text{for } N_{SPT} > 30 \end{cases} \quad (2C-9)$$

Formula 2:

$$S_c \begin{cases} 1.9 \frac{q^{0.77} B^{0.45}}{N^{1.08}} & \text{for } B \leq 3m \\ 1.64 \frac{q^{1.02} B^{0.59}}{N^{1.37}} & \text{for } B > 3m \end{cases} \quad (2C-10)$$

- Schultze and Sherif (1973)

Starting from the general formula for predicted settlement by means of the elasticity equation:

$$S_c = qBI/E \quad (2C-11)$$

This expression is obtained by numerical integration of Boussinesq solution for circular footing whereas for rectangular footing the influence factor I by Steinbrenner’s method (1934) is introduced. The stiffness modulus E (in kPa) is given by $E = 1678(N_{SPT})^{0.87} \cdot B^{0.5} \cdot (1+0.4D/B)$.

- Berardi and Lancellotta (1991)

As the preceding authors do, Berardi and Lancellotta start from the elasticity equation for settlement:

$$S_c = qBI/E \quad (2C-12)$$

but calculate each term differently from the previous authors. As a matter of fact, the influence factor I has been calculated by numerical integration of Boussinesq solution considering rigid footing. The stiffness modulus, E, is given by $E = K_E \times p_a \times [(\sigma'_{v0, D+B/2} + \Delta \sigma'_{v0, D+B/2})/p_a]$, where p_a is the reference pressure and K_E is initially evaluated as $K_E = 100+900D_R$, $D_R = \min[1, (N_1/60)^{0.5}]$, and $N_1 = 2N_{SPT}/(1+\sigma'_{v0, D+B/2}/98.1)$.

Once the settlement has been calculated, K_E is corrected as $K_{E, corr} = K_E \cdot 0.1912 \cdot (S_c/B)^{-0.6248}$ and the predicted settlement is recalculated using $K_{E, corr}$ instead of K_E . Cherubini and Vessia (2009) suggested a procedure to rank the nine R samples of model uncertainty by the use of four indexes: (1) precision, (2) accuracy, (3) entropy, and (4) relative entropy. The first two indexes can be combined into the Ranking distance and plotted as shown in Figure 2C-1.

Entropy (Shannon, 1948) and relative entropy (Kinsley, 1983), like the variance, are measures of the dispersion according to the following expressions, respectively:

$$H(x) = -\sum_{i=1}^n P_i \ln(P_i) \quad (2C-13)$$

$$H_r(x) = H(x)/H_{max} \quad (2C-14)$$

where P_i is the probability value associated with the calculated R value.

The procedure consists of (1) calculating the RD, whose ideal value is 0, (2) calculating the relative entropy by Eq. 2C-14 whose ideal value is about 0, and (3) ranking the equations according to the calculated values of RD and H_r .

The final aim of the proposed procedure that can be extended to a larger number of equations is to provide a quick tool for designing. In fact, curves of different levels of probability for settlement target values can be drawn through several formulas. Hereafter just to provide with an example, three curves have been calculated for three formulas that are: two by Anagnostopoulos et al. (1991), named APK1 and APK2 and one by Schultze and Sherif (1973), named SS. The preceding equations are the most precise formulas among the nine considered by Cherubini and Vessia (2009) according to the available full-scale data by Burland and Burbidge’s settlement database.

As far as the values of the shallow foundation settlements are concerned, Eurocode 7 (BS 2004, updated 2013) suggests an upper limit equal to 50mm to be accepted. Nevertheless, the preceding value refers to common buildings with no complex structures. In practice, settlements can be accepted if they do not cause both static and functional problems to structures and services. Thus, five values of settlements are hereafter considered as acceptable: 5, 15, 25, 35 and 45 mm. The exceeding probability curves are related to three samples of R that are normally distributed. The curves are drawn according to the following expression:

$$p(R < \bar{R}) = p(R < \bar{S}_{calc}/\bar{S}_{mis}) = p(\bar{S}_{calc}/R > \bar{S}_{mis}) \quad (2C-15)$$

This expression shows that the probability that a settlement value is higher than the corresponding measured one is equal to the probability that any value of R is lower than a fixed value of R (5, 15, 25, 35 and 45). Figures 2C-2 to 2C-4 show the probability curves calculated through APK1, APK2 and SS.

The values corresponding to the probability of exceedance can be obtained from these settlement probability curves. For example, considering that a 30mm settlement has been estimated by means of the SS formula, Figure 2C-4 shows that there is a 30% probability that the actual settlement under footing will be more than 45mm, that is to say the probability of calculating settlements below 45mm will be 70%. Similarly, there will be a 15% probability that the settlement is below 15mm, i.e. an 85% probability that the settlement is more than 15mm. For all three formulas the exceeding probability curves for values of settlements occurring between the curves plotted on the graph will be obtained by interpolating the values of probability for the two curves which mark the range.

These probability charts let the engineer know the probability that a certain value of the settlement measured is exceeded provided that the settlement value is known, i.e. they provide the conditioned probability of the settlements measured:

$$P(S_{\text{mis}} > \bar{S} | S_{\text{calc}} = \bar{s}) \quad (2C-16)$$

where S_{calc} and S_{mis} are the calculated and measured settlements, respectively; \bar{s} and \bar{S} are the values of their respective settlements. So, the conditioned probability in terms of not-exceedance will be the following:

$$P(S_{\text{mis}} > \bar{S} | S_{\text{calc}} = \bar{s}) = 1 - P(S_{\text{mis}} < \bar{S} | S_{\text{calc}} = \bar{s}) \quad (2C-17)$$

Therefore, as a result of a reliability-based design in terms of settlements (calculated by means of one of the three formulas APK1, APK2 and SS), the probability that the actual settlement of the footing does not exceed a given value can be easily got by the conditioned probability theorem:

$$P(S_{\text{mis}} < \bar{S} | S_{\text{calc}} \leq \bar{s}) = P(S_{\text{mis}} < \bar{S} | S_{\text{calc}} = \bar{s}) \times P(S_{\text{calc}} \leq \bar{s}) \quad (2C-18)$$

where the first term comes from the probability charts (Figures 2C-2 to 2C-4) while the second one derives from the reliability-based design.

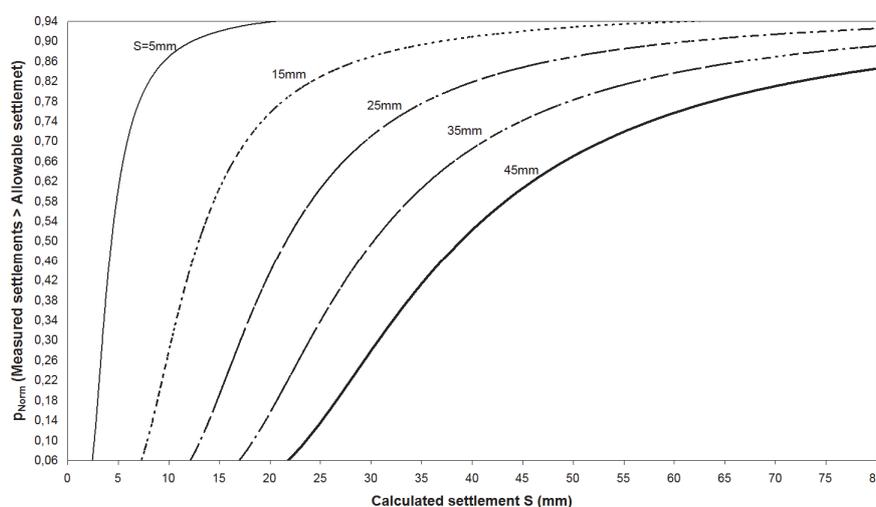


Figure 2C-2 Normal-distribution probability curves corresponding to different values of acceptable settlements calculated by APK1 formula.

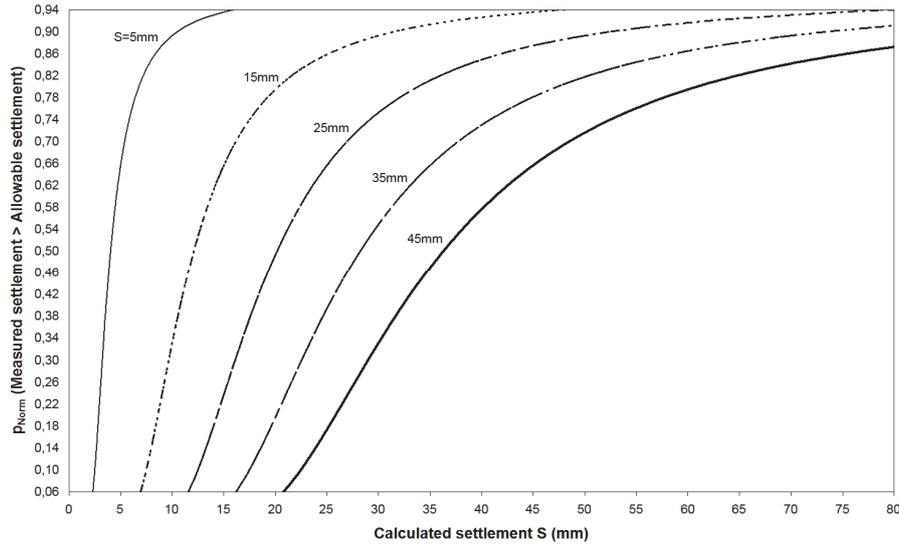


Figure 2C-3 Normal-distribution probability curves corresponding to different values of acceptable settlements calculated by APK2 formula.

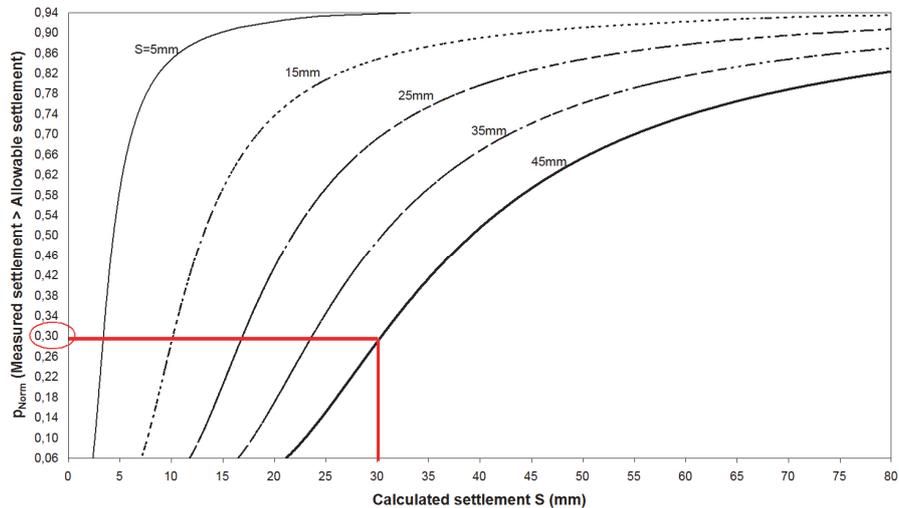


Figure 2C-4 Normal-distribution probability curves corresponding to different values of acceptable settlements calculated by SS formula.

In fact, if the settlements are estimated by the SS formula, and the probability of not exceedance of a 20 mm settlement is equal to 10^{-4} , the probability that the settlement actually measured is below 25 mm will be easily got from Figure 2C-4:

$$P(S_{\text{mis}} < 25 \text{ mm} | S_{\text{calc}} \leq 20 \text{ mm})_{\text{SS}} = (1 - 0.46) \cdot 10^{-4} = 0.54 \cdot 10^{-4} \quad (2C-19)$$

If the two other formulas are used to estimate settlements the following results will be calculated:

$$P(S_{\text{mis}} < 25 \text{ mm} | S_{\text{calc}} \leq 20 \text{ mm})_{\text{APK}_1} = (1 - 0.42) \cdot 10^{-4} = 0.58 \cdot 10^{-4} \quad (2C-20)$$

$$P(S_{\text{mis}} < 25 \text{ mm} | S_{\text{calc}} \leq 20 \text{ mm})_{\text{APK}_2} = (1 - 0.50) \cdot 10^{-4} = 0.50 \cdot 10^{-4} \quad (2C-21)$$

References

Burland, J. B. and Burbidge, M. C. (1985). Settlement of foundations on sand and gravel. Proceedings

- of the Institution of Civil Engineers, part 1, 78, pp. 1325-1381.
- Briaud, J. and Tucker, L. M. (1988). Measured and predicted axial load response of 98 piles. *Journal of Geotechnical Engineering*, 114(9), pp. 984-1001.
- Cherubini, C. and Orr, T. L. L. (2000). A rational procedure for comparing measured and calculated values in geotechnics. In: *Proceedings of the International Symposium on Coastal Geotechnical Engineering in Practice*, Yokohama, Balkema, Rotterdam pp. 261-265.
- Cherubini, C. and Vessia, G. (2005). Studi di Affidabilità nel calcolo di cedimenti in terreni sabbiosi. *Atti del III Convegno Scientifico Nazionale Sicurezza nei Sistemi Complessi*, 19-21 ottobre, Bari.
- Cherubini, C. and Vessia, G. (2009). Probabilistic charts for shallow foundation settlements on granular soil. *Proceedings of the 2nd International Symposium on Geotechnical Safety and Risk IS-GIFU, GIFU, Japan*, 11-12 June.
- Kinsley, H. W. (1983). Some geotechnical applications of entropy. In: Augusti, Borri and Vannucchi (ed.), *4th International Conference on Applications of Statistics and Probability in Soil and Structural Engineering*, Florence, 13-17 June, Italy.
- Shannon, C. E. (1948). A mathematical theory of communication. *The bell system technical journal* 27, pp. 379-423.
- Tan, C. K. and Duncan, J. M. (1988). Settlement of footings on sands – accuracy and reliability. *Journal of Geotechnical Engineering*, 114(9), pp. 446-455.
- Terzaghi, K. and Peck, R. B. (1967). *Soil mechanics in engineering practice*. New York: J. Wiley & Sons.
- Uzielli, M. and Mayne, P. (2012). Load-displacement uncertainty of vertically loaded shallow footings on sands and effects on probabilistic settlement estimation. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 6(1), pp. 50-69.

Appendix 2D: Model uncertainty evaluation for pile foundations (W. Bogusz)

While using the LRFD method in accordance with Eurocode 7, the choice of the calculation model used to correlate predicted bearing capacity with soil test results is the responsibility of the geotechnical designer, who has to guarantee its validity. Vast majority of currently used calculation models, some of which developed decades ago, were based on empirical data gathered by performing static load tests on full-scale piles. To ensure the safety of the design, these models were provided with associated safety factors, most commonly, single global factors of safety covering all the uncertainties concerning actions and resistances. For example, Bustamante and Gianselli (1981) proposed using a factor of 3 for the base and a factor of 2 for the shaft resistances of axially loaded piles, respectively, in association with their method. Moreover, those models are also often associated with additional rules (i.e. on embedment depth in bearing stratum) and limitations (i.e. maximum unit shaft friction or base resistance), which differentiate the reliability of different methods.

Model factor for piles

As different degrees of conservatism and calculation rules are associated with different calculation models, Bauduin (2003) states that it may be difficult to reach the required safety level linking these models with a set of partial factors provided in standards. These partial factors should be independent of the calculation model in use, and they should cover other sources of uncertainties. The main reason for the use of a model factor is to provide an expected level of reliability of the prediction of the calculated resistance value for a specific model, to either provide accurate results or err on the side of safety. According to Bauduin (2003), a model factor modifying the calculation results should be used to address the bias of $R_{\text{measured}}/R_{\text{calc}}$ presented by the model and its variability. The main point of using this factor is to ensure, with given probability, that the resistance of the pile will be larger than the predicted value.

Two commonly applied calculation model types are used in practice for pile foundations when prediction is based on soil test results. Firstly, semi-empirical methods (i.e. model pile procedure in Eurocode 7), where pile shaft and base resistance are derived from the measured ground parameter directly for a specific location, where tests were conducted. The tests most commonly used to directly derive pile capacities are CPT, SPT and PMT. Secondly, analytical models (indirect methods; i.e. alternative procedure in Eurocode 7 if unit resistance for a stratum is not directly correlated to measured values of a specific profile) may be used, which are often based on soil strength parameters derived from aforementioned in-situ tests or laboratory tests, introducing additional uncertainty due to parameter identification error and subjectivity of the parameter selection. According to Bauduin (2003), those uncertainties are included into the model uncertainty.

According to Bauduin (2002, 2003), the model factor introduced in the design aims to provide a certain reliability of prediction using a specific calculation model; thus, ensuring that there is only a p% probability that the real value is lower than calculated one. In order to do that, the model factor has to integrate both, the bias and the coefficient of variation.

$$\gamma_{\text{mod}} = \frac{1}{\left(\frac{R_{\text{measured}}}{R_{\text{calc}}} \right)_{p\%}} \quad (2D-1)$$

It is possible to distinguish three main approaches of database analysis in order to derive the model factor for a calculation method (Bauduin, 2002). The first one considers a very high level of detail in the analysis, where each type of pile can be calibrated for specific soil conditions. The second one allows a lower level of detail grouping similar piles and soil conditions together. The third and last approach allows for larger generalization, treating all data as one sample, resulting in one calibration factor for the calculation method.

The choice of the approach should depend on the availability of data and the possibility of modification of other factors affecting the reliability of the design, namely installation factors and partial safety factors. In the best-case scenario, when a large and very detailed database is available, it may be more advisable to modify the installation factors instead of introducing an additional model

factor as a calibration method. On the other hand, if the database is considered as a single sample due to limited amount of data, modification of the partial safety factor for resistance may be required to differentiate e.g. between bored and driven piles.

Bauduin (2002) argues that simplified global calibration of the resistance is justified; however, having separate data concerning shaft and base resistance, their separate calibration is also possible.

$$R_d = \frac{R_{\text{calc}}}{\gamma_{\text{mod}} \cdot \gamma_R} = \left(\frac{R_{b;\text{calc}}}{\gamma_b} + \frac{R_{s;\text{calc}}}{\gamma_s} \right) \cdot \frac{1}{\gamma_{\text{mod}}} \quad (2D-2)$$

Effects of installation, as well as small variations in pile geometry, are taken into account by the calculation rule or are included in the scatter of $R_{\text{measured}}/R_{\text{calc}}$ (Bauduin, 2003).

Existing databases with static load test results can be used to calibrate new calculation models. Additionally, to introduce a new type of pile, only five static load tests are necessary to establish its model factor (Bauduin, 2003). However, increasing their number lowers the uncertainty, allowing to lower its value. Appending new data of sufficient quality, gathered over time, would enable researchers and code drafters to regularly revise and possibly lower selected model factor values; this approach may give contractors an incentive to gather and share data as it might offer a possible cost reduction on future contracts.

Bauduin (2002) stated that an assessment of a calculation model for pile foundation should consider: soil type; method of pile installation, geometrical data of the pile, and the method of ground testing. It may be assumed that these data should also be included as basic information in any database of pile load tests.

Most of the older calculation models based on CPT results, used for bearing capacity prediction, utilized only cone resistance as an input value, while methods utilizing sleeve friction and pore pressure measurement (EN-ISO 22476-1) were developed relatively recently. However, it might be argued that the simplification of input data for the calculation model is a positive aspect, as it introduces some level of robustness to the design.

One of the important factors affecting the predicted bearing capacity is the filtration and averaging of the soil test results in order to eliminate extreme values, often done by the engineer responsible for prediction of the pile capacity. It introduces additional bias that is not always directly related to the uncertainty of the method itself. Secondly, limiting the maximum unit bearing capacity to a specific value, used in some of calculation models, imposes an additional margin of safety. Such uncertainties and limitations are a result of limited knowledge about the behaviour of a specific model in certain conditions; potentially, it could be eliminated if additional data becomes available.

Database of static load tests

Although static load tests are considered as the most definite way of assessing pile capacity, they are not free of uncertainties. As the load measurement is done directly, the procedure used for the test (maintained load test, maintained rate of penetration, or creep test), measurement technique and the interpretation introduce some degree of uncertainty.

Current assumptions about the ULS failure criterion provided in Eurocode 7 are that it corresponds to the settlement equal to 10% of the pile diameter. Although some calculation models assume even lower values (i.e. 5%), in fact, full mobilisation of end bearing capacity of a bored compression pile in sands can occur at settlement of approx. 20% of its diameter. Furthermore, due to the high non-linearity of pile behaviour, often not exhibiting clear plunging failure, different criteria are available for determining the ultimate bearing capacity. However, this is beyond the scope of this report.

Compiling databases of static load tests is necessary in order to upgrade existing and develop new calculation models for the bearing capacity prediction of piles. Such databases are limited by the current technological advancement of both, ground and pile testing techniques.

As the best approach for model factor derivation, as well as verification of existing and development of new calculation methods, is to use statistical methods; it is advisable to create databases of static load tests serving as reference values for calibration. However, the quality of input data is of significant importance. Firstly, the quality of the site investigation has to be considered. Optimally, in-situ testing should limit the influence of spatial variability, measurement errors, and

comprise of various techniques if possible. Additionally, the test should be performed to sufficient depth to take into account any possible weak strata that may be located below the pile base in its influence zone. Secondly, sufficient quality of test loads should be provided by performing only static load tests to failure, desirably, with distinction between shaft and end resistances. However, due to the costs and time consumption, in many cases these types of tests are avoided by contractors. Moreover, such tests performed to failure on working piles are quite seldom. Usually, if performed only as a proof test after pile foundation execution, the maximum load of a static test rarely exceeds 150% of the design load value to avoid unnecessary costs; these are associated mostly with the loading frame and anchoring required for higher loads. If the pile design is based on a conservative calculation model, it will not exhibit sufficient settlement which would correspond to the ultimate load criterion. Performing proof tests in limited loading range instead of up-to-failure tests gives very little knowledge about pile behaviour (Paikowsky and Tolosko 1999). Although there are methods of interpretation of non-failed static load tests, their use, especially as a basis for model calibration, is not recommended. Paikowsky and Tolosko (1999) analysed 63 load tests carried out to failure, disregarding the final part of the load-settlement data to simulate non-failed tests; then after applying two already existing and one proposed extrapolation methods, they compared the extrapolated pile behaviour with the real one. Over-prediction of bearing capacity based on extrapolation could be as high as 50%.

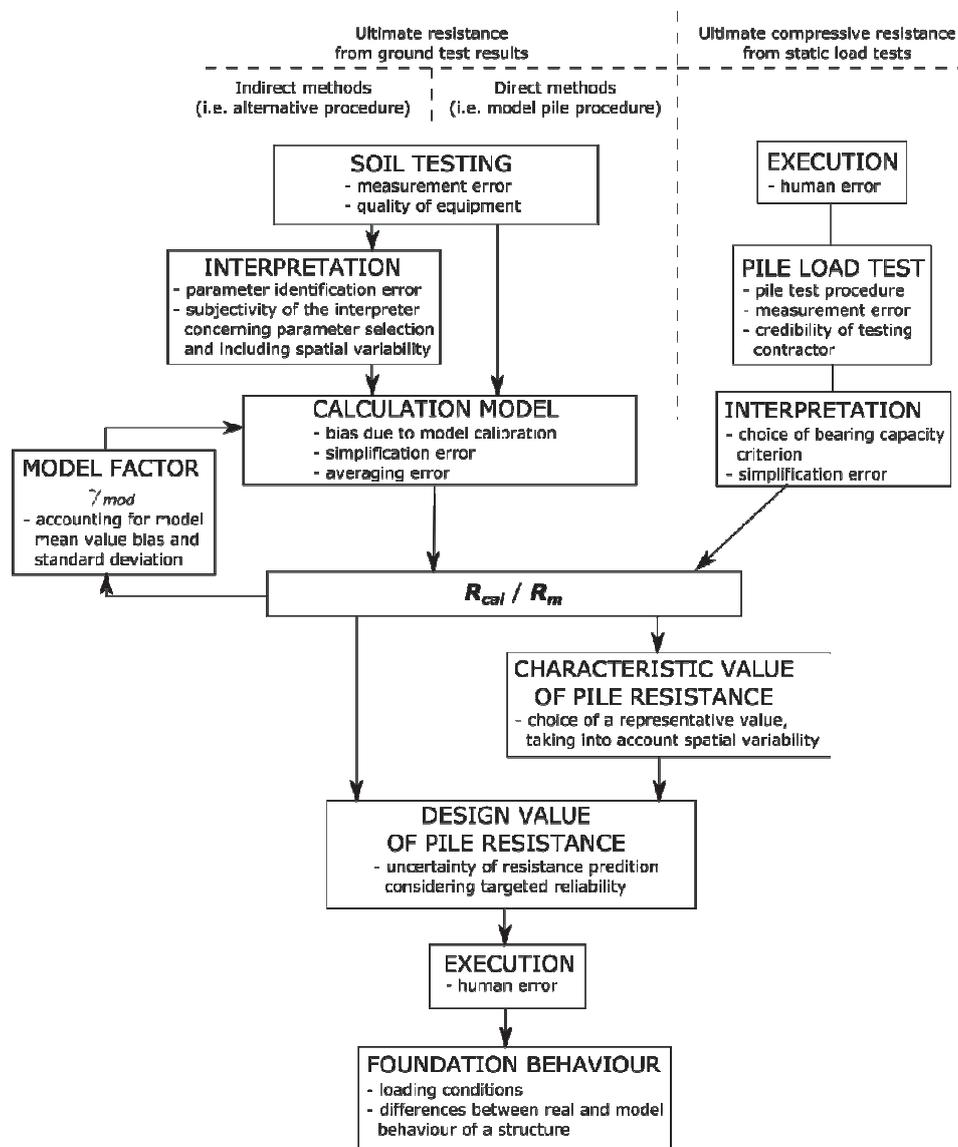


Figure 2D-1 Pile design procedure with most significant uncertainty sources (courtesy of Witold Bogusz).

Advancement in data mining and data analysis techniques may provide more suitable tools, i.e. based on artificial neural networks, to eliminate model uncertainties or to take into the account additional factors. However, the use of these tools will still be limited by the amount and quality of available data.

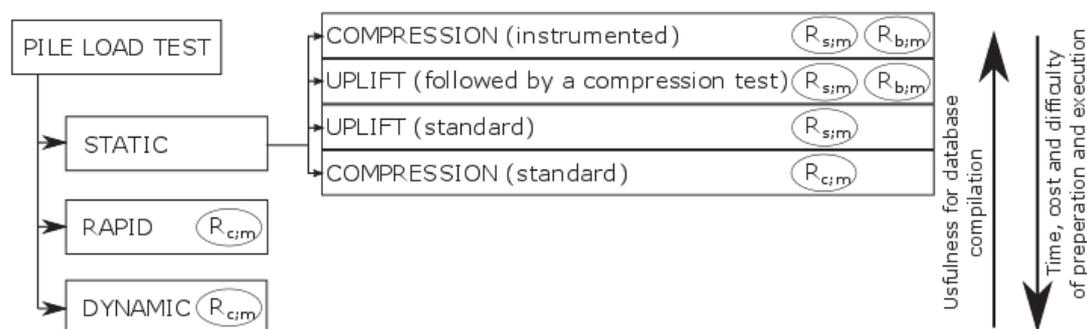


Figure 2D-2 Pile load test types used for verification of ultimate bearing capacity of pile foundations (courtesy of Witold Bogusz).

Practical application of calculation model calibration

An example of the development of a new calculation model, as an improvement of the Bustamante method (Bustamante and Gianceselli, 1981), with its calibration to derive model factors was presented by Burlon et al. (2014). It was based on the database of 174 full-scale, mostly instrumented, static pile load tests, which had been performed over a period of 40 years in France. The model factor was introduced in the French standard for pile design. Spatial variability was disregarded due to the location of the test piles at the exact locations of corresponding soil tests.

Burlon et al. (2014) suggested two possible approaches for derivation of the model factor. First by comparison of the dispersion of the former and new calculation models, assuming a sufficient safety level of the former model, and the second approach as a direct determination using statistical analysis. It is worth noting that for a group of piles, mostly including micropiles and injected piles, a high value of the model factor is used due to insufficient samples in the database and to encourage full-static load testing of such piles.

References

- Bauduin, C. (2002). Design of axially loaded piles according to Eurocode 7. Proceedings of the International Conference on Piling and Deep Foundations, DFI 2002, Nice, Presses de l'ENPC, Paris.
- Bauduin, C. (2003). Assessment of model factors and reliability index for ULS design of pile foundations. Proceedings of the 4th International Geotechnical Seminar on Deep Foundations on Bored and Auger Piles, 2-4 June, Ghent.
- Burlon, S., Frank, R., Baguelin, F., Habert, J., and Legrand, S. (2014). Model factor for the bearing capacity of piles from pressuremeter test results – Eurocode 7 approach. *Géotechnique*, 64(7), pp. 513-525.
- Bustamante, M. and Gianceselli, L. (1981). Prevision de la capacite portante des pieux isolés sous charge verticale – Regles pressiometriques et penetrometriques. *Bull. Lab. Ponts Chaussees*, 113, pp. 83-108.
- EN 1997-1: 2008 Eurocode 7: Geotechnical design – Part 1: General rules.
- EN-ISO 22476-1: 2012 Geotechnical investigation and testing - Field testing - Part 1: Electrical cone and piezocone penetration test.
- Paikowsky, S. and Tolosko, T. (1999). Extrapolation of pile capacity from non-failed load tests. US DOT FHWA-RD-99-170.

Appendix 2E: Model factor as a function of input parameters and revised model statistics (C. Tang)

The following summary is based on several published studies.

Top deflection of a cantilever wall in undrained clay

Summary from: Zhang, D.M., Phoon, K.K., Huang, H.W., and Hu, Q.F. (2015). *Characterization of model uncertainty for cantilever deflections in undrained clay. Journal of Geotechnical and Geoenvironmental Engineering*, 141(1): 04014088, pp. 1-14.

The model factor for MSD calculating the top deflection of a cantilever wall in undrained clay was evaluated using FEM and a database (see Table 2E-1).

Table 2E-1 Summary of database for top-deflection of a cantilever wall.

Reference	Variables	Range of values	#N
<u>Field case histories</u>			
Zhang et al. (2015)	Wall depth D (m)	12~40	
	Excavation depth H _c (m)	1.5~6.9	
	Excavation width B/2 (m)	6.3~75	
	Wall stiffness EI (MPa/m)	119~4388	
	Lateral earth pressure coefficient at rest K ₀	1.5~6.9	45
	Relative undrained shear strength s _u /σ' _v	0.32~1.5	
	Soil stiffness ratio E _{ur} /s _u	167~756	

#N=number of tests, which will be used throughout this Appendix.

The database consists of 45 deep excavations in soft to medium-stiff clays. The clay within the excavation depth was regarded as homogeneous. All case histories were internally braced deep excavations where the cantilever-type wall deflection was observed at the initial stage of construction. The available soil data consists of borelog descriptions and in-situ tests [e.g. standard penetration test/piezcone (SPT/CPTU) for undrained shear strength s_u] and advanced laboratory test results (e.g. K₀ test for K₀ and triaxial test for soil stiffness E_{ur}).

Bearing capacity of strip footings on sand under general combined loading

Summary from:

Phoon, K. K. and Tang, C. (2015a). *Model uncertainty for the capacity of strip footings under positive combined loading. Geotechnical Special Publication in honor of Wilson. H. Tang (in press).*

Phoon, K. K. and Tang, C. (2015b). *Model uncertainty for the capacity of strip footings under negative and general combined loading. Proc. 12th International Conference on Applications of Statistics and Probability in Civil Engineering, Vancouver, Canada, July 12-15.*

The model factor for Eurocode 7 approach calculating the bearing capacity of strip footings on sand under positive and negative combined loading was evaluated using FELA and laboratory small-scale load tests conducted on poorly graded sand. The database includes 120 load tests (60 for dense sand D_R=69% and 60 for medium-dense sand D_R=51%) for positive combined loading and 72 load tests (36 for dense sand and 36 for medium-dense sand) for negative combined loading (see Table 2E-2). In the database, foundation width D is 0.1 m. The embedment ratio d/D was varied from 0 to 1. The load eccentricity e ranges from 0 to 0.15D, while the load inclination α is from 0 to 20°.

Table 2E-2 Summary of database for strip footings on sand under combined loading.

Reference	Variables	Range of values	N
<u>Laboratory small-scale load tests</u>			
<u>Positive combined loading</u>			
Phoon and Tang (2017)	Foundation width D (m)	0.1	
	Foundation embedment depth d/D	0, 0.5, 1	
	Friction angle ϕ (°)	37.5, 40.8	120
	Load inclination α (°)	0, 5, 10, 15, 20	
	Load eccentricity e/D	0, 0.05, 0.1, 0.15	
<u>Negative combined loading</u>			
Phoon and Tang (2015)	Foundation width D (m)	0.1	
	Foundation embedment depth d/D	0, 0.5, 1	
	Friction angle ϕ (°)	37.5, 40.8	72
	Load inclination α (°)	5, 10, 15, 20	
	Load eccentricity e/D	0.05, 0.1, 0.15	

Table 2E-3 Summary of database for helical anchors in clay under tension loading.

Reference	Variables	Range of values	N
<u>Laboratory small-scale load tests</u>			
Tang and Phoon (2016)	Helical anchor diameter D (m)	0.033~0.15	
	Relative embedment depth of the uppermost helical plate H/D	1~10	
	Undrained shear strength s_u (kPa)	3~13.5	78
	Number of helical plates n	2~5	
	Relative spacing ratio S/D	0.83~2.3	
	<u>Field full-scale tests</u>		
	Helical anchor diameter D (m)	0.2~0.345	
	Relative embedment depth of the uppermost helical plate H/D	4~28	
	Undrained shear strength s_u (kPa)	31~99	25
	Number of helical plates n	3~5	
	Relative spacing ratio S/D	0.75~3	

Uplift capacity of helical anchors in clay

Summary from: Tang, C. and Phoon, K. K. (2016). *Model uncertainty of cylindrical shear method for calculating the uplift capacity of helical anchors in clay*. *Engineering Geology*, 207, pp. 14-23.

The model uncertainty of the cylindrical shear method predicting the uplift capacity of helical anchors in clay was characterized using FELA and a load test database (Table 2E-3). The database consists of 78 laboratory small-scale and 25 field load tests in soft to stiff clays.

Bearing capacity of circular footings on dense sand

Summary from: Tang, C. and Phoon, K. K. (2017). *Model uncertainty of Eurocode 7 approach for bearing capacity of circular footings on dense sand. International Journal of Geomechanics, 17(3): 04016069, pp. 1-9.*

The model factor for Eurocode 7 approach estimating bearing capacity of circular footings on dense sand (DR=70%~95%) was evaluated using FELA and 26 centrifuge tests (Table 2E-4) compiled from literature, where foundation diameter D varies from 0.3 m to 10 m.

Table 2E-4 Summary of database of circular footings on dense sand.

Reference	Variables	Range of values	N
<u>Centrifuge tests</u>			
Tang and Phoon (2017a)	Foundation diameter D (m)	0.3~10	
	Critical state friction angle ϕ_{cv} (°)	31~39.3	26
	Relative density of sand D_R	70%~95%	

Bearing capacity of large circular and conical foundations on sand overlying clay

Summary from: Tang, C., Phoon, K. K., Zhang, L., and Li, D.-Q. (2017). *Model uncertainty for predicting the bearing capacity of sand overlying clay. International Journal of Geomechanics (just released), 04017015, pp. 1-14.*

The model factor for conventional approaches (the load spread method and the punching-shear method) calculating the bearing capacity of dense sand overlying clay was statistically evaluated using FELA and 62 centrifuge tests. The database consists of two parts (Table 2E-5). The first part is associated with four centrifuge tests for onshore foundations [foundation diameter (D) =1.5~3 m] on Toyoura sand ($D_R=88\%$ and $\phi_{cv}=32^\circ$) overlying normally consolidated clay. The second part of the database was associated with 58 centrifuge tests for the peak resistance of large flat-circular footings (26 cases with D=6~16 m) and spudcan (32 cases with D=3~16 m) penetration in dense sand ($D_R=74\% \sim 99\%$ and $\phi_{cv}=31^\circ, 32^\circ$) overlying normally consolidated clay. In addition, 27 centrifuge tests (D=6~20 m) on dense to very loose sand ($D_R=24\% \sim 89\%$, $\phi_{cv}=31^\circ$ and 32°) overlying clay were used for verification purposes.

The results for the correction factor M_c (including the regression equation f and the residual part η), the model factor M_p for numerical methods (FEM or FELA), and the revised model factor M' for modified calculation models are summarized in Table 2E-6. Besides, comparisons between load tests and simplified calculation models with its modification multiplying by regression equation f are graphically presented in Figures 2E-1 to 2E-6.

According to the results presented cited above, it is reasonable to make the following conclusions:

- (1) The model factor $M=X_{meas}/X_{cal}$ for a simplified model predicting deflection (e.g. MSD) or foundation capacity (e.g. Eurocode 7 or punching shear method) is generally a function of input parameters such as problem geometries and soil mechanical properties, which cannot be treated as a random variable directly.
- (2) The model factor $M_p=X_{meas}/X_p$ for a mechanically consistent numerical method (e.g. FEM or FELA) is usually independent of input parameters, which can be characterized as a lognormal random variable with a mean around 1. The COV value of M_p could be about 0.1 at the ultimate limit state (ULS), while the corresponding value may be around 0.2 at the serviceability limit state (SLS).
- (3) On this basis, such a numerical method can be used to complement load tests to remove the statistical dependency, which is expressed as an exponential function (f) of some influential parameters. The performance of a simplified model can be improved by multiplying the established regression equation f (see Figures 2E-1 to 2E-6). The residual part η of the correction factor $M_c=X_p/X_{cal}=f \times \eta$ is then modelled as a lognormal random variable. It is of practical significance, as load tests are usually limited in geotechnical engineering.
- (4) Finally, the mean of the revised model factor M' is approximately equal to 1. However, the

COV value of M' at the SLS (e.g. Zhang et al. 2015) is larger than that for ULS (e.g. Phoon and Tang 2015, 2017; Tang and Phoon 2016, 2017a; and Tang et al. 2017a). In addition, when the problem geometry becomes complicated or soil is non-homogeneous, the COV value of M' increases (Tang et al. 2017a).

Table 2E-5 Summary of database for large flat-circular footing and spudcan penetration in sand overlying clay.

Reference	Variables	Range of values	N
Centrifuge tests			
<u>Flat-circular footing</u>			
Tang et al. (2017a)	Foundation diameter D (m)	1.5~16	30
	Thickness of sand layer H (m)	1.5~6.7	
	Critical state friction angle ϕ_{cv} (°)	31, 32	
	Relative density of sand D_R	74%~92%	
	Undrained shear strength at sand-clay interface s_{u0} (kPa)	8.71~24.5	
	Gradient of undrained shear strength ρ (kPa/m)	0.94~3.46	
	<u>Spudcan</u>		
Tang et al. (2017a)	Foundation diameter D (m)	3~16	32
	Cone angle α (°)	7~21	
	Thickness of sand layer H (m)	3~10.5	
	Critical state friction angle ϕ_{cv} (°)	31, 32	
	Relative density of sand D_R	74%~99%	
	Undrained shear strength at sand-clay interface s_{u0} (kPa)	7.2~27.2	
	Gradient of undrained shear strength ρ (kPa/m)	1.2~2.13	

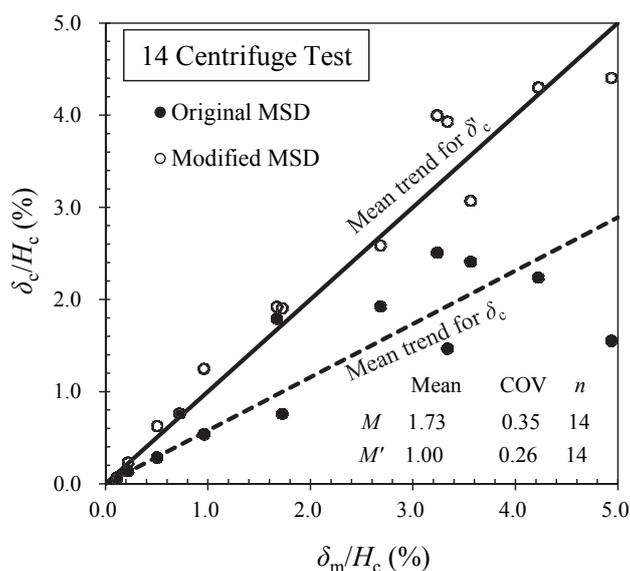


Figure 2E-1 Comparison of deflection from centrifuge tests and MSD with its modification.

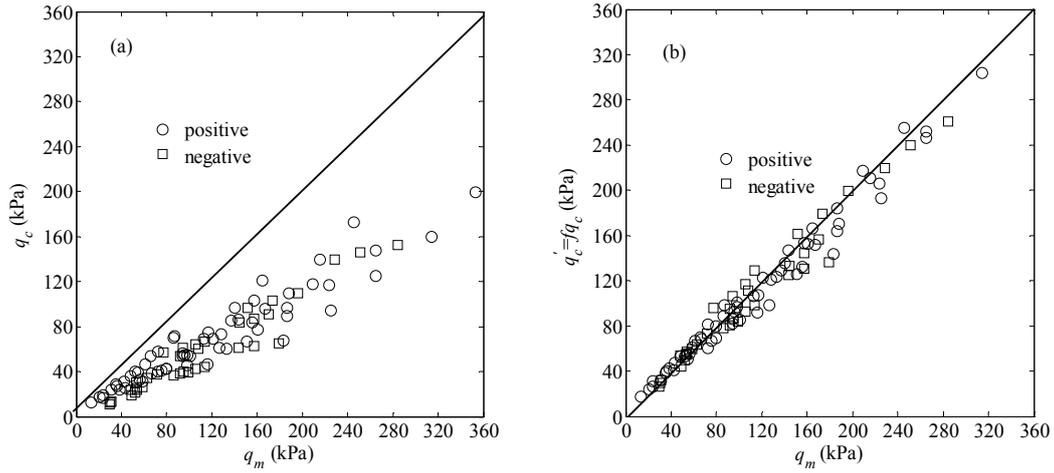


Figure 2E- 2 Comparison of bearing capacity of strip footings on sand under combined loading from laboratory tests and Eurocode 7 with its modification.

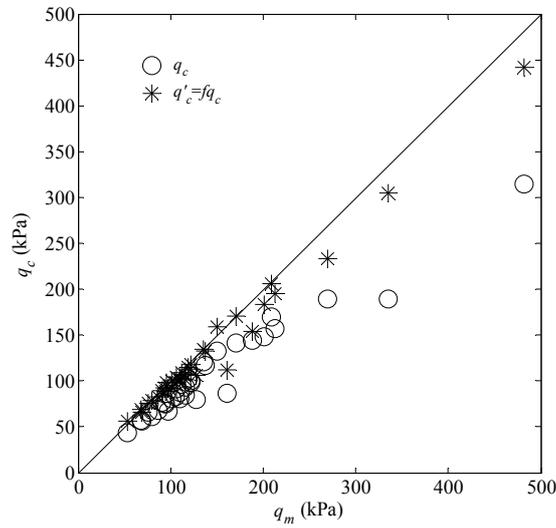


Figure 2E-3 Comparison of uplift capacity of helical anchors in clay under tension loading from database and cylindrical shear method with its modification.

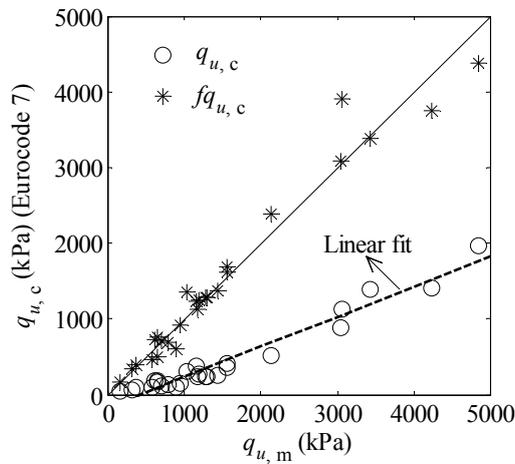


Figure 2E-4 Comparison of bearing capacity of circular footings on dense sand from database and Eurocode 7 with its modification.

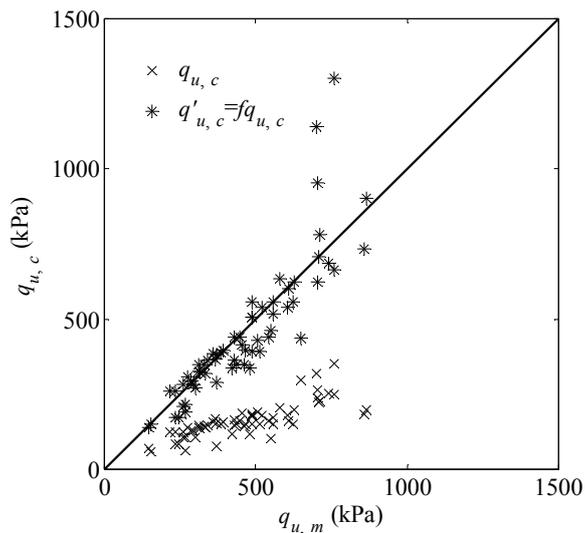


Figure 2E-5 Comparison of bearing capacity of sand overlying clay from calibration database and punching shear method with its modification.

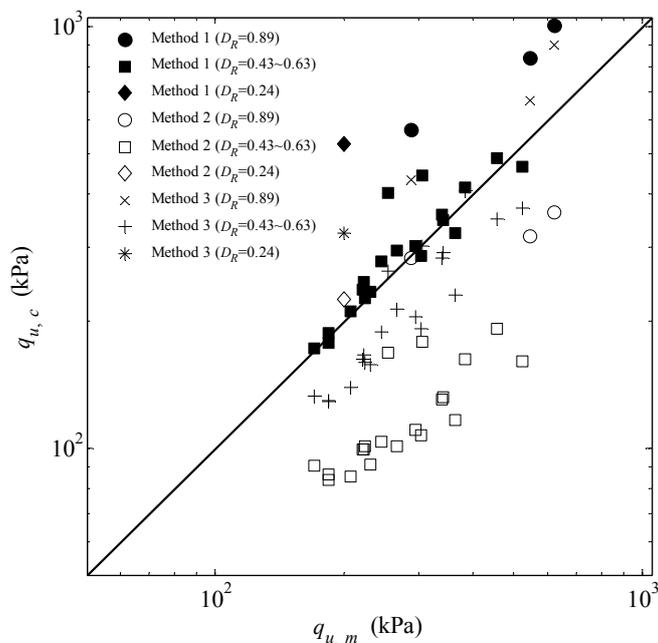


Figure 2E-6 Verification of modified punching shear method (Method 3), while Method 1 is failure-stress-dependent method (Hu et al. 2014) and Method 2 is original punching shear method.

References

- Hu, P., Stanier, S. A., Cassidy, M. J., and Wang, D. (2014). Predicting peak resistance of spudcan penetrating sand overlying clay. *Journal of Geotechnical Geoenvironmental Engineering*, 140(2): 04013009, pp. 1-13.
- Phoon, K. K., and Tang, C. (2015a). Model uncertainty for the capacity of strip footings under positive combined loading. *Geotechnical Special Publication in honour of Wilson. H. Tang* (in press).
- Phoon, K. K., and Tang, C. (2015b). Model uncertainty for the capacity of strip footings under negative and general combined loading. *Proc. 12th International Conference on Applications of Statistics and Probability in Civil Engineering*, Vancouver, Canada, July 12-15.
- Tang, C. and Phoon, K. K. (2016). Model uncertainty of cylindrical shear method for calculating the

- uplift capacity of helical anchors in clay. *Engineering Geology*, 207, pp. 14-23.
- Tang, C. and Phoon, K. K. (2017). Model uncertainty of Eurocode 7 approach for bearing capacity of circular footings on dense sand. *International Journal of Geomechanics*, 17(3): 04016069, pp. 1-9.
- Tang, C., Phoon, K. K., Zhang, L., and Li, D.-Q. (2017). Model uncertainty for predicting the bearing capacity of sand overlying clay. *International Journal of Geomechanics* (just released), 04017015, pp. 1-14.
- Zhang, D.M., Phoon, K.K., Huang, H.W., and Hu, Q.F. (2015). Characterization of model uncertainty for cantilever deflections in undrained clay. *Journal of Geotechnical and Geoenvironmental Engineering*, 141(1): 04014088, pp. 1-14.

Table 2E-6 Summary of regression equations and model factors statistics.

Problem and reference	Correction factor $M_c = X_p / X_{cal} = f \times \eta$		Model factor for numerical methods		Revised model factor				
	mean	COV	$\# X_i$	b_i	mean	COV	mean	COV	
Strip footing on sand under positive combined load (Phoon and Tang 2017)			η						
			$M_p = X_{meas} / X_p$				$M' = X_{meas} / (f X_{cal}) = \eta \times M_p$		
			η						
			mean	COV	mean	COV	mean	COV	
			$b_0 = 0.28$						
			$x_1 = \gamma D / p_a$						
			$b_1 = -5.05$						
		$x_2 = \xi$							
		$b_2 = 11.4$							
		$x_3 = \tan \phi$							
	1.65	0.22	$x_4 = d/D$						
			$b_3 = -0.26$						
			$b_4 = -0.09$	1.03	0.09	1.01	0.06	1.04	0.11
			$x_5 = \alpha / \phi$						
			$b_5 = 0.21$						
			$x_6 = e/D$						
			$b_6 = -1.12$						
			$x_7 = (e/D)(\alpha / \phi)$						
			$b_7 = -0.98$						
Strip footing on sand under negative combined load (Phoon and Tang 2015)			$b_0 = 0.1$						
			$x_1 = \gamma D / p_a$						
			$b_1 = -4.5$						
			$x_2 = \xi$						
			$b_2 = 10.4$						
		2	0.24	$x_3 = \tan \phi$					
				$b_3 = -0.25$	1.06	0.09	1.01	0.06	1.06
			$x_4 = d/D$						
			$b_4 = -0.12$						
			$x_5 = \alpha / \phi$						
			$b_5 = -1.03$						
			$x_6 = e/D$						
			$b_6 = -0.45$						
			$x_7 = (e/D)(\alpha / \phi)$						
			$b_7 = -1.81$						

M_c is not a random variable.

p_a =atmosphere pressure (101 kPa); d =foundation embedment depth; ξ =empirical factor=0.02~0.12; and γ =unit weight of soil.

Table 2E-6 (Continued.)

Problem and reference	Correction factor $M_c = X_p / X_{cal} = f \times \eta$			Model factor for numerical methods			Revised model factor			
	mean	COV	# x_i	f	mean	COV	η	mean	COV	COV
Strip footing on sand under general combined load (Phoon and Tang 2015)				$b_0 = 0.1$						
			$x_1 = \gamma D / p_a$	$b_1 = -4.5$						
			$x_2 = \xi$	$b_2 = 10.25$						
	2	0.25	$x_3 = \tan \phi$	$b_3 = -0.15$	1.04	0.09		1.06	0.09	0.13
			$x_4 = d/D$	$b_4 = 0.05$						
			$x_5 = \alpha / \phi$	$b_5 = -0.93$						
			$x_6 = e/D$	$b_6 = -0.05$						
		$x_7 = (e/D)(\alpha / \phi)$	$b_7 = -2.53$							
Helical anchors in clay under tension loading (Tang and Phoon 2016)				$b_0 = 0.78$						
			$x_1 = n$	$b_1 = -0.07$						
	1.37	0.26	$x_2 = S/D$	$b_2 = -0.1$	0.98	0.14		0.95	0.08	0.16
			$x_3 = H/D$	$b_3 = -0.02$						
		$x_4 = \gamma H / s_u$	$b_4 = 0.04$							
Circular footings on dense sand (Tang and Phoon 2017a)				$b_0 = 1.97$						
			$x_1 = \tan \phi_{ev}$	$b_1 = -3.12$						
	1.18	0.32	$x_2 = D_R$	$b_2 = 2.23$	1.02	0.11		1	0.1	0.11
		$x_3 = \gamma D / p_a$	$b_3 = -0.68$							

Table 2E-6 (Continued.)

Problem and reference	Correction factor $\eta M_c = X_p / X_{cal} = f \times \eta$				Model factor for numerical methods			Revised model factor	
	mean	COV	$\ln f = b_0 + \sum b_i x_i$	η	mean	COV	mean	COV	
Flat-circular and conical footings on sand overlying clay (punching shear method) (Tang et al. 2017a)			$b_0 = -0.79$						
			$x_1 = D_R$						
	2.63	0.31	$x_2 = H/D$	1.01	1.01	0.11	1.07	0.17	
			$x_3 = \tan \phi_{ev}$						
			$x_4 = S_{u0} N_c / \gamma D$						
Flat-circular and conical footings on sand overlying clay (load spread method $\tan \alpha_p = 1/3$) (Tang et al. 2017a)			$b_0 = -1.61$						
			$x_1 = D_R$						
	1.83	0.3	$x_2 = H/D$	1.01	1.01	0.11	1.06	0.17	
			$x_3 = \tan \phi_{ev}$						
			$x_4 = S_{u0} N_c / \gamma D$						
Flat-circular and conical footings on sand overlying clay (load spread method $\tan \alpha_p = 1/5$) (Tang et al. 2017a)			$b_0 = -1.48$						
			$x_1 = D_R$						
	2.32	0.33	$x_2 = H/D$	1.02	1.01	0.11	1.07	0.17	
			$x_3 = \tan \phi_{ev}$						
			$x_4 = S_{u0} N_c / \gamma D$						

ηM_c is not a random variable.

$\#n$ =number of helical plates; S =spacing between two adjacent helical plates; H =embedment depth of the uppermost helical plate or the thickness of the upper sand layer; and N_c =bearing capacity factor.

Chapter 3 Practical Geotechnical Design using Reliability-Based Design Methods

Lead discussor:

Peter Day

day@jaws.co.za

Discussors (alphabetical order):

Witold Bogusz, Nico de Koker, Bak Kong Low, Kok-Kwang Phoon, Brian Simpson

3.1 INTRODUCTION

Reliability analysis in structural and geotechnical engineering has been around for many years and provides the basis for the formulation and calibration of limit states design codes, such as the Eurocodes. In the past 10 years, reliability based design methods have become sufficiently accessible to serve as practical tools in a geotechnical design office, rather than being confined to research applications.

Many countries are not bound by any regional economic grouping or other influences to adopt a particular design method or set of codes, as is the case with member states within the European Union and the Eurocodes. Even though Limit States Design is now the international norm, the perceived obscurity of the method, multiplicity of approaches and openness to interpretation are impediments to widespread adoption of the method. As a result, some practitioners are electing to continue using working stress design methods while, at the other end of the spectrum, others are advocating the use of reliability based design methods.

There are several factors that need to be considered before decisions can be made regarding the selection of a preferred method of geotechnical design. Some such factors are suitability for routine design, breadth of application, design data requirements, uniform (adequate yet not excessive) levels of reliability, comparable output to current practice and clarity of performance requirements.

This chapter examines the reliability of solutions to common geotechnical problems obtained using limit states design methods and the influence of parameter values on this reliability. It then discusses the suitability of reliability-based methods for routine design purposes and the extent to which these methods address the perceived shortcomings of limit states design methods.

3.2 COMPARISON OF DESIGN APPROACHES BY SIMPLE EXAMPLES

Shortly after the release of Eurocode 7 (EN1997-1, 2004), an international workshop on the implementation of Eurocode 7 was held at Trinity College in Dublin (Orr, 2005). One of the purposes of this workshop was to compare the solutions to common geotechnical problems as prepared by representatives of various countries in Europe. Its legacy was an appreciation for the diversity of approaches and interpretations of Eurocode 7 and a set of model solutions which have been of considerable value to those seeking to implement the code.

In much the same way, there is benefit in applying reliability based design to simple problems and comparing the solutions obtained with those from other design methods. The examples presented in this chapter have deliberately been kept simple as this facilitates an intuitive interpretation of the outcomes, unobscured by any complexities in the analysis.

The comparisons presented here are based on an interpretation of Eurocode 7 and methods of selection of parameters that can be expected from a typical geotechnical practitioner for routine design.

3.2.1 Problems Analyzed

The problems analysed were based on those selected for the Dublin Workshop (Orr, 2005) and are shown in Figure 3-1. Only problems with closed-form solutions were used. In each case, the solution required is a dimension that determines the adequacy of the structure, such as the width of a footing or the length of a pile. The nominal dimensions shown in Figure 3-1 were taken as characteristic values with no adjustment.

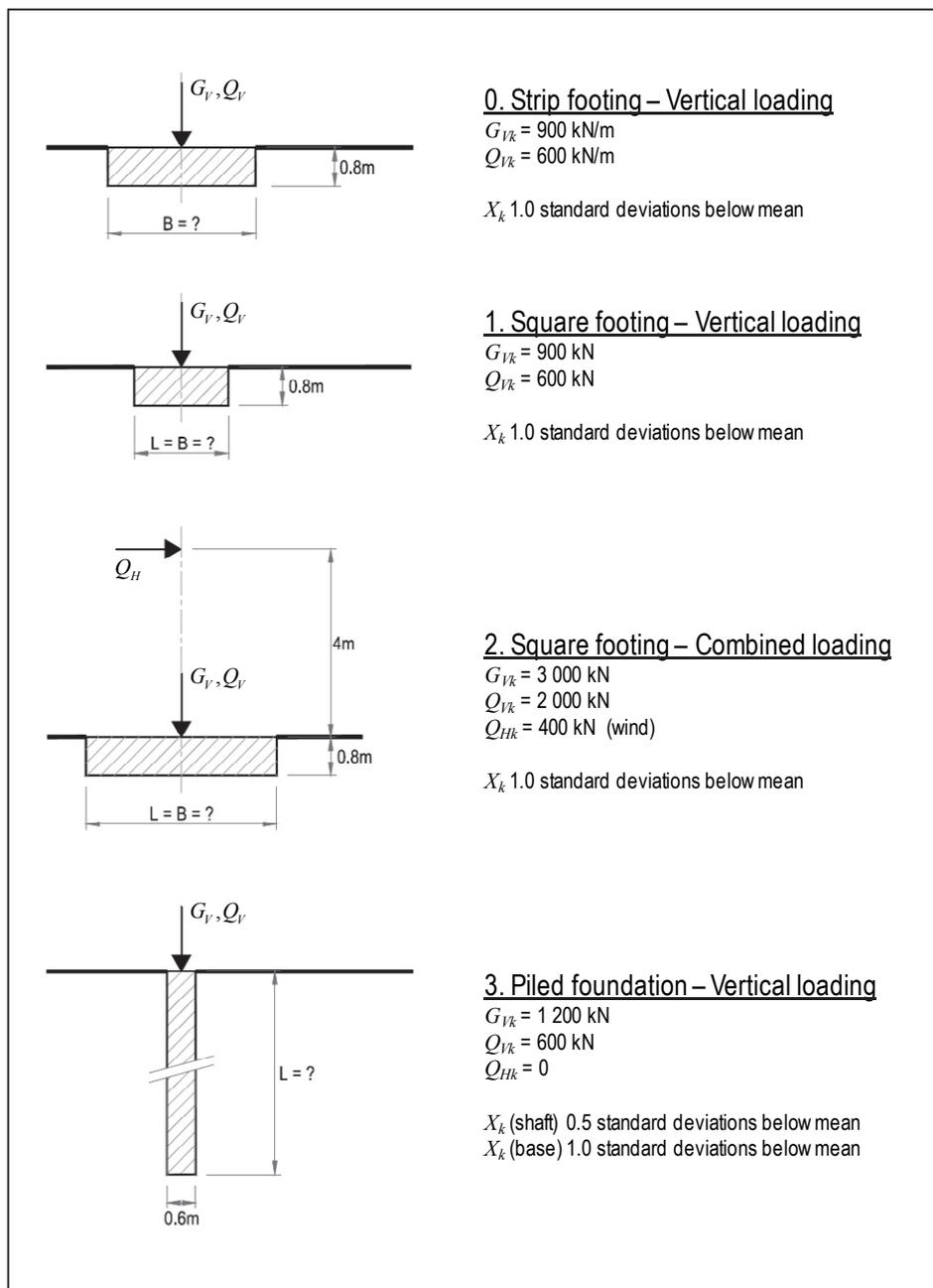


Figure 3-1a Examples 0 to 3 (Foundations)

3.2.2 Soil Type

For the purposes of this study, a single soil type was chosen, namely a cohesionless sand with a deep water table. As such, there are only two material (soil) parameters to be considered, namely friction angle and density. The selected properties are given in Table 3-1.

Table 3-1 Assumed Soil Properties

Soil Property Parameter	Friction angle ϕ'_k	Bulk density γ_k
Characteristic value	32°	20 kN/m ³
Distribution	log-normal	normal
Coefficient of variation	0.10	0.05
Correlation coefficient	0.2	

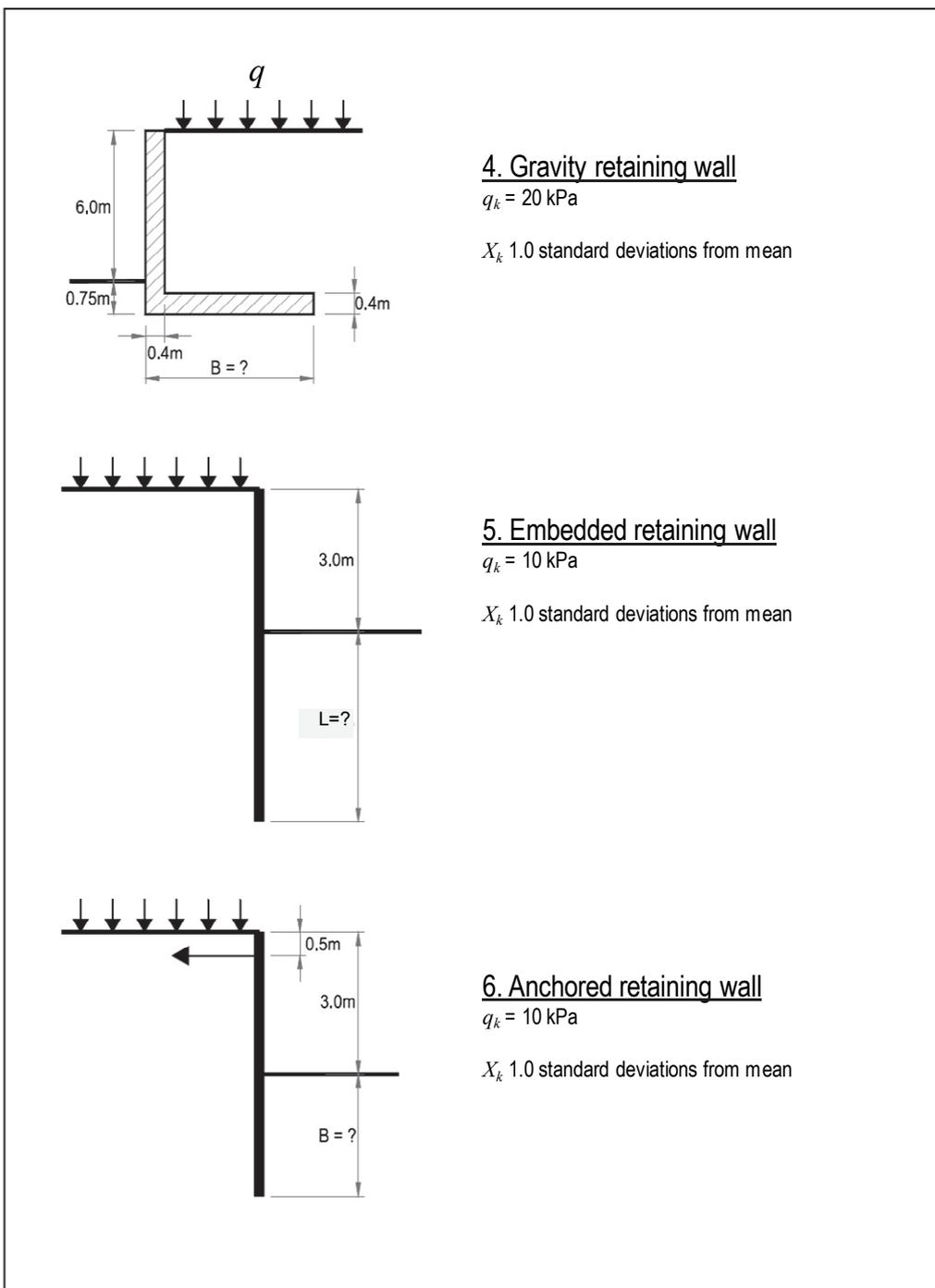


Figure 3-1b Examples 4 to 6 (Retaining structures)

In practice, the characteristic value of a material property is determined from the mean value minus (or plus) a number (n) of standard deviations as selected by the designer (Schneider, 1997; Bond & Harris, 2008). The multiplier n depends on the number of test results available and the extent to which the occurrence of the limit state (i.e. failure) is dependent on the average rather than minimum/maximum value of the parameter. In this study, n was taken as 1.0 in all cases except Example 3 where $n = 0.5$ was used for the pile shaft resistance. This difference recognises that the bearing capacity of a footing could be governed by local soil conditions whereas the shaft friction resistance of a pile is likely to “average out” local variations in soil properties along the length of the pile. For the purposes of the reliability analysis, the corresponding mean value is back-figured from the characteristic value by applying the process in reverse.

3.2.3 Loading

The following loads (actions) have been considered in the examples.

Table 3-2 Assumed Actions

Action	Distribution
Permanent action (G_v)	Fixed value, mean value = characteristic value
Variable action vertical (Q_v)	Log-normal distribution, coefficient of variation 0.25
Variable action horiz. (Q_h)	Gumbel distribution, coefficient of variation 0.5 (wind)

The vertical and horizontal actions are assumed to be independent. An action combination factor of 0.7 has been applied to the accompanying variable action. The vertical variable action may be favourable or unfavourable. The statistical distributions, coefficients of variation and ratio of characteristic to mean loading are based on Retief & Dunaiski (2009) and Phoon & Kulhawy (1999).

3.2.4 Analysis

The following approach was followed in this investigation:

- (1) Find the solution to the problem that satisfies the ultimate limit state verification requirements of EN1990 and EN1997 (Design approach 1, Combination 2) using the characteristic values of loads and material properties.
- (2) Determine the mean values of loads and material properties corresponding to the given characteristic values.
- (3) Using the corresponding mean values and coefficients of variation, determine the reliability index (β) of the solution using Monte Carlo and First Order Reliability Methods (FORM).
- (4) Determine the global factor of safety (FoS) using working stress design methods.

In the case of Example 2 (square footing under vertical and horizontal loads), the following additional analyses were performed, both of which required a re-evaluation of the limit state design compliant solution for each new set of material properties:

- (5) Repeat of steps (1)-(4) for a range of material properties to examine the variation of β and FoS .
- (6) Repeat of steps (1)-(3) for a range of coefficients of variation of the material properties to explore the sensitivity of β to uncertainties in material properties.

3.2.5 Example 2 – Vertically and Horizontally Loaded Square Footing

This example, which is based on Example 2 from Orr (2005), will be discussed in detail.

The same methodology was followed for the other examples, for which only a summary of the results is reported in Section 3.2.6 of this discussion document.

3.2.5.1 Problem Setup and Input Data

The problem setup and input data for Example 2 is shown in Figure 3-2.

3.2.5.2 LSD and WSD solutions

The limit states and working stress design solutions for this example are given below.

Table 3-3 Solutions – Example 2

LSD Solution	WSD (average values)	WSD (characteristic values)
$B = L = 3.99\text{m}$	$B = L = 3.99\text{m}$	$B = L = 3.99\text{m}$
$e_B = 0.49\text{m}$	$e_B = 0.21\text{m}$	$e_B = 0.36\text{m}$
$B' = 3.02\text{m}$	$B' = 3.57\text{m}$	$B' = 3.27\text{m}$
$R_d = E_d = 5\,126\text{kN}$	$q_f = 2\,067\text{kPa}$	$q_f = 1\,031\text{kPa}$
	FoS = 6.58	FoS = 2.60

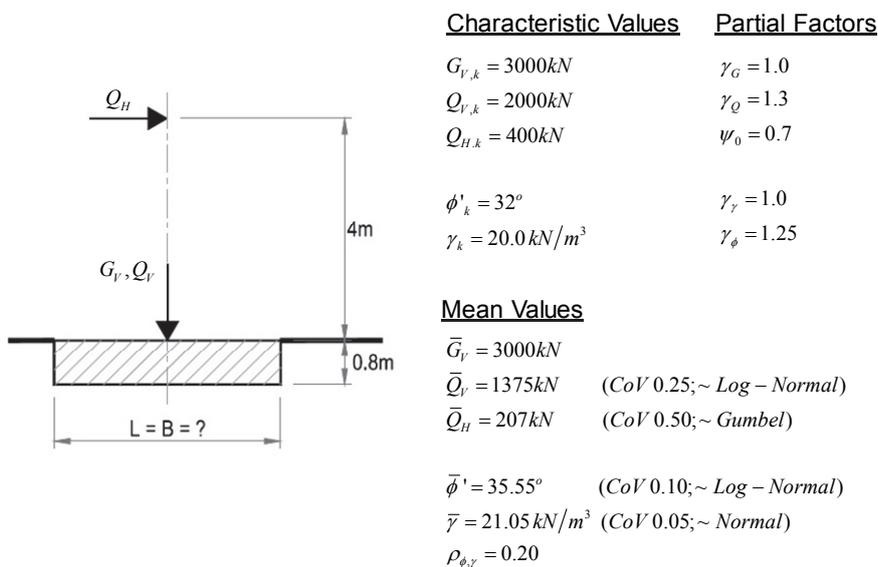


Figure 3-2 Input Data for Example 2

3.2.5.3 Reliability Analysis

$\phi' \sim \text{Log-normal}$

The results of the Monte Carlo and FORM analyses of the Eurocode-compliant solution are summarised in Figure 3-3.

Effect of ϕ' distribution

The effect of different assumptions regarding the statistical distribution of the friction angle of the soil is summarised in Table 3-4.

Table 3-4 Effect of Assumed Statistical Distribution for Friction Angle

	$\phi' \sim \text{Normal}$	$\phi' \sim \text{Log-normal}$	$\tan \phi' \sim \text{Log-normal}^*$
<u>Monte Carlo - with correlation ($\rho_{\phi',\gamma} = 0.20$)</u>			
Reliability Index (β)	3.26	3.59	3.59
No. iterations	10^6	10^6	10^6
No. failures	559	162	168
<u>FORM</u>			
$\beta, \rho_{\phi',\gamma} = 0.20$	3.40	3.69	3.69
$\beta, \rho_{\phi',\gamma} = 0.0$	3.46	3.73	3.73
Design point, $\rho_{\phi',\gamma} = 0.20$			
Q_V	1406	1305	1312
Q_H	358	638	622
γ	20.07	20.22	20.02
ϕ'	24.61	27.85	27.63

* $\bar{\phi}'$ adjusted to match FORM $\beta=3.69$. ($\bar{\phi}' = 35.41^\circ$, $CoV = 0.121$)

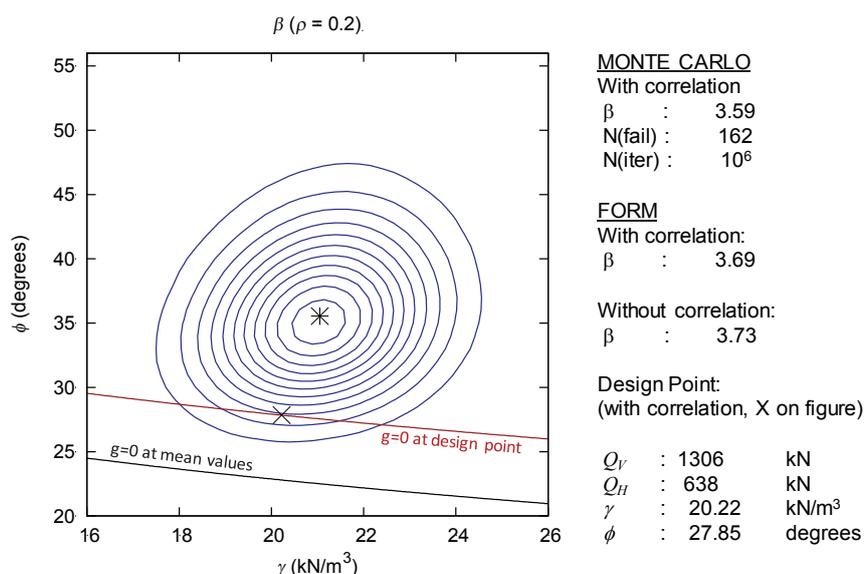


Figure 3-3 Reliability Analysis of Eurocode-compliant Solution to Example 2

Effect of range of material properties

The above analysis of Example 2 considered only a single set of ground properties. For partial factor limit states design to be an acceptable method of design, the target level of reliability should be achieved across the range of material properties likely to be encountered in practice.

The effect of variation in the values of the soil properties is shown in Figure 3-4, in which the limit states design solution is evaluated for a range of ϕ' - γ values, and the corresponding level of reliability determined using FORM.

Effect of variance in material properties

In the same way, the target level of reliability should be achieved for the range in variance (coefficient of variation) of geotechnical parameters likely to occur in practice. Changing the variance of the

material properties affects the ratio of the characteristic value to the mean value as described in Section 3.2.2.

The effect of changes in the variance of the soil properties is shown in Figure 3-5.

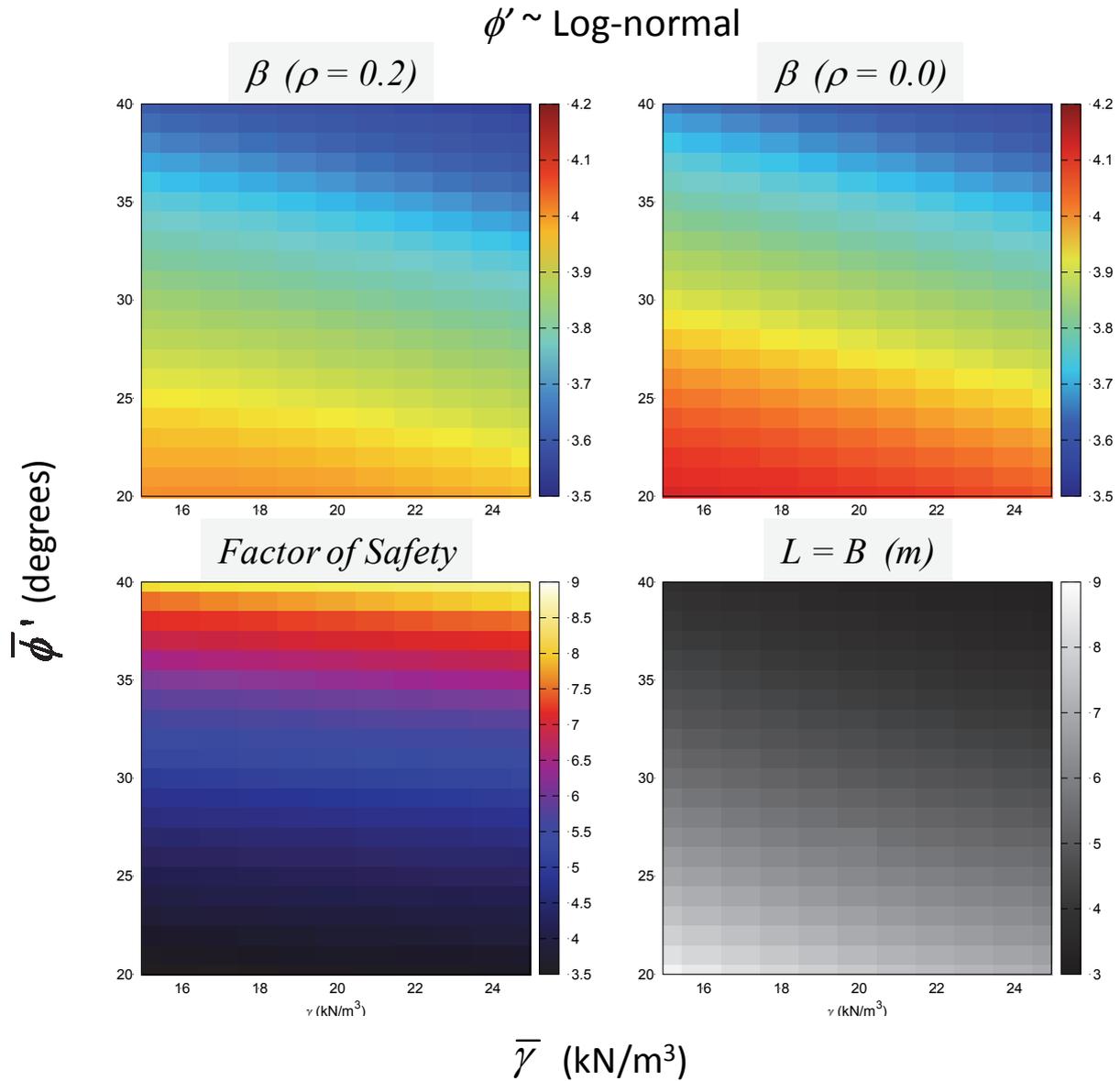


Figure 3-4 Effect of Range of Material Properties on Reliability Index and FoS

3.2.6 Summary of Results for Other Examples

A similar examination of the reliability of Eurocode-compliant design and the corresponding factors of safety has been carried out for the remaining examples. Only a single set of ground parameters was used. Analyses to assess the effect of the range and variance of ground properties are still underway.

The global factor of safety achieved by the Eurocode-compliant design has been calculated using the characteristic and mean values of the loads and material properties. The results are summarised in Table 3.5.

Table 3-5 Reliability Indices and Factors of Safety for Remaining Examples

Example	No. Variables	Solution B or L	Global Factor of Safety		Reliability Index β	
			Mean	Characteristic	Monte Carlo	FORM
0	3	3.10m	5.18	2.50	3.45	3.49
1	3	1.97m	4.86	2.40	3.46	3.51
2	4	3.99m	6.58	2.60	3.58	3.69
3	3	8.58m	2.76	1.73	3.35	3.36
4	3	3.52m	6.88	3.12	3.30	3.33
5	3	4.00m	2.34	1.63	3.39	3.40
6	3	2.57m	1.43	1.25	3.23	3.24

$\phi' \sim \text{Log-normal}$

$$\bar{\phi}' = 35.55^\circ$$

$$\bar{\gamma} = 21.05 \text{ kN/m}^3$$

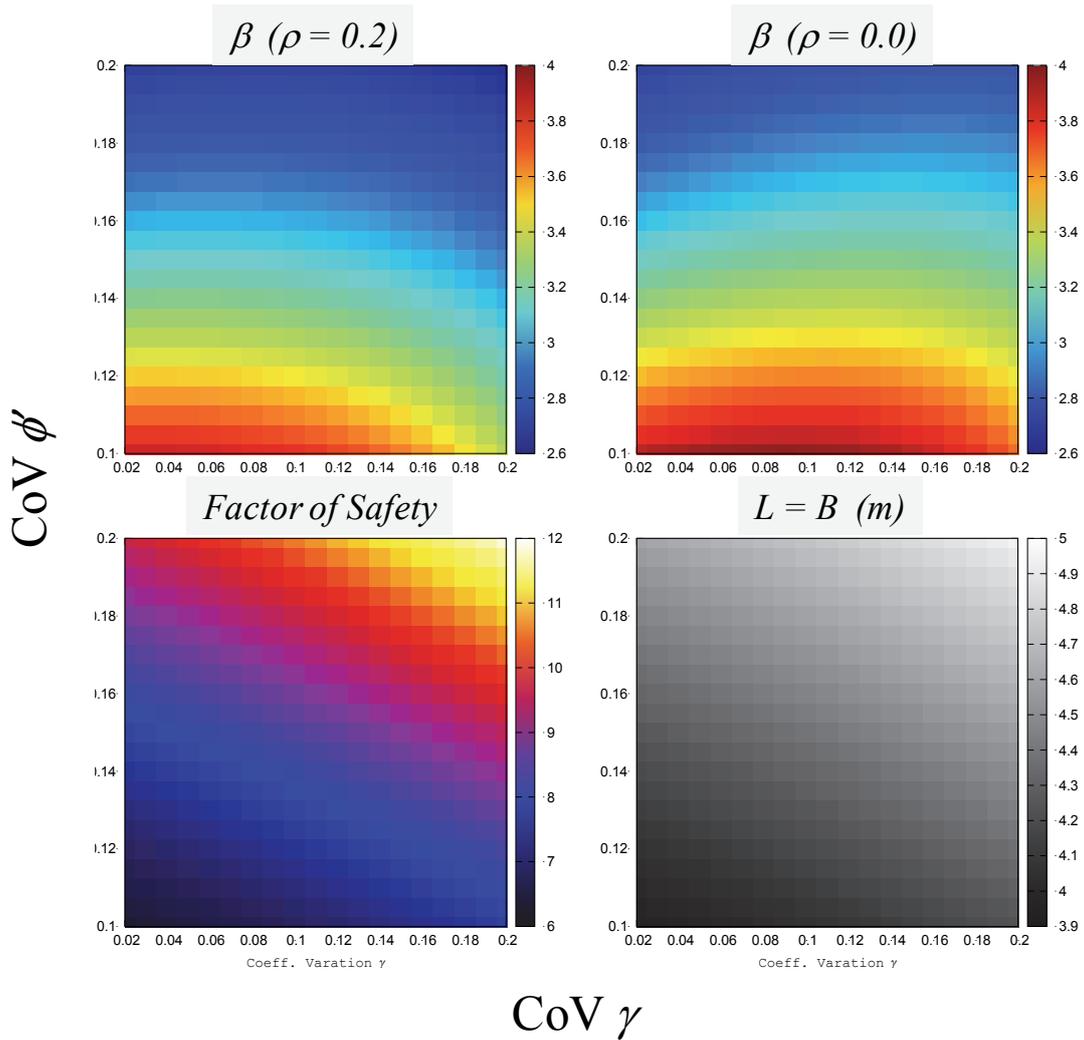


Figure 3-5 Effect of Variance of the Material Properties on Reliability Index and FoS

3.2.7 Conclusions from Examples

3.2.7.1 Limit States Design

The starting point for all the above analyses was an EN1997-1 compliant, partial factor limit states design solution to each of the seven common geotechnical problems shown in Figure 1. Four significant conclusions can be drawn from the results.

- (1) The reliability indices are fairly constant for the wide range of problems considered (footings, piles and retaining structures). For the chosen soil properties, the reliability index varied from 3.2 to 3.7.
- (2) Even with a wide range in material properties, the variation in reliability indices for Example 2 was not excessive. For friction angles ranging from 20° to 40° and densities from 15kN/m³ to 25kN/m³, combined with typical values for the coefficient of variation, the reliability indices varied from 3.5 to 4.2.
- (3) The reliability indices obtained are generally lower than the default target value of $\beta = 3.8$ from the Eurocodes, but not significantly so. The way the characteristic value is chosen will affect the level of reliability.
- (4) In all the above analyses, the relationship between the mean value and the characteristic value of the material properties takes account of the expected variance of the parameter. This dependence comes about by selecting the characteristic value to be a multiplier of standard deviations from the mean value with the multiplier being dependent on the degree to which the occurrence of the limit state is affected by the average or the minimum properties of the material. The change in the selected characteristic value due to the change in the coefficient of variation of the material properties is, however, insufficient to compensate for the effect which higher variance in soil properties has on the reliability index, which is seen in Figure 3-5 to decrease significantly with increases in the variance in soil properties. Caution should be exercised when using limit states design methods in ground where the friction angle in particular shows a high degree of variance.

3.2.7.2 Working Stress Design

Three significant observations are made regarding the factor of safety:

- (1) In contrast to the relatively modest variation in reliability index, the global factors of safety obtained from working stress design analyses vary widely for different problem types and across the range of soil properties considered. This supports the now well-established realisation that the global factor of safety is a poor means of assessing the reliability of a structure.
- (2) For the range of material properties considered in Example 2, an increase in the global factor of safety corresponds with a decrease in the reliability index. Thus, for example, the determination of allowable bearing pressure for a given global factor of safety using working stress methods will result in reduced levels of reliability as the friction angle of the material is increased.
- (3) The global factors of safety obtained when the working stress analysis is carried out using characteristic values of loads and material properties are closer to those traditionally used in practice than those obtained using mean values. The global factors of safety calculated using the mean values are significantly higher. This supports the view expressed by Krebs Ovesen and Simpson in the mid-80's that there is not a significant difference between the characteristic values of material properties and the values that would typically be chosen for working stress design methods.

3.2.7.3 Reliability Based Design

The analyses performed demonstrate that reliability analysis can be effectively and practically applied to common geotechnical problems. Furthermore, FORM analyses gave results that compared well with those obtained using Monte Carlo simulation. Spreadsheet applications of FORM (e.g. Low & Tang, 2007) make this a practical tool for use in the design office, particularly for problems with closed-form solutions.

One of the limitations of the reliability analyses described in this chapter is the assumption that a single value of a material parameter applies at all points in the soil mass and along the full length of

any failure surface. Consider, for example, the effect of variations in the friction angle of the soil in the case of the pile foundation in Example 3. The assumption made in both the FORM and Monte Carlo analyses is that the same friction angle applies throughout, i.e. for the calculation of both end bearing and shaft resistance. Thus, although variations in the strength of the ground are taken into account, the spatial variations and the degree to which these variations will be “averaged out” along the length of the pile shaft are ignored. In this respect, the approach used in limit states design is more appropriate in that different characteristic values are used for end bearing and shaft resistance, taking account of the degree to which the occurrence of the limit state is affected by the average or local properties of the ground.

3.3 RELIABILITY BASED DESIGN—A VIABLE ALTERNATIVE TO LIMIT STATES DESIGN?

Few informed geotechnical designers would doubt that Limit States Design methods are an improvement of working stress design using a global factor of safety. However, there are many designers who remain unconvinced that limit states design as expounded in the Eurocodes is the best design method available. The criticisms of the method include:

- (1) The multiplicity of design approaches and calculation models, often leading to very different outcomes. Although the intention of the Eurocodes was to create a set of technical rules for the design of construction works across CEN member states, the solutions received to design examples included in the 2005 International Workshop on the Evaluation of Eurocode 7 varied significantly from country to country (Orr, 2005b).
- (2) There is a perception that the partial factors used are somewhat arbitrary and that their selection appears to have been based more on replicating results from the past rather than being calibrated against the observed performance of constructed works. This remains a problem as, except for elements such as piles and ground anchors which are subjected to routine testing, insufficient data is available for rigorous calibration of partial factors.
- (3) The perceived complexity of the method which requires the consideration of numerous limit states and load combinations, and the application of multiple partial factors at different stages of the calculation.
- (4) The subjectivity involved in the selection of characteristic values.

Two questions arise. Firstly, does reliability based design address the perceived shortcomings of limit states design listed above? Secondly, are reliability based design techniques suitable for everyday design problems?

Reliability based design methods do not overcome the problems associated with the selection of different calculation models and design assumptions which were cited by Orr (2005b) as the major contributor to differences in solutions from various CEN countries. These models and assumptions still form part of the performance function to be evaluated in the reliability analysis. Reliability based methods still require the consideration of all possible modes of failure (limit states). The use of load combination factors is not required as the method will select the most critical combination of loads based on the statistical data specified by the designer. Nevertheless, the designer must still assess whether favourable loads will or will not be present. The method does not rely on any partial factors.

The problem with the selection of characteristic values is replaced by the equally onerous challenge of selecting statistical distributions for input parameters and the spatial variation of these parameters within the soil mass. Designers may, however, take comfort in the fact that this challenge is broken down into a number of individual selections including the type of distribution, the mean, the variance and the covariance of the variables, for which guidance is available in the literature.

The examples considered in Section 3.2 of this report show that FORM produces reliability indices very similar to those from Monte Carlo simulations, with considerably less computational effort. Efficient algorithms for carrying out FORM analyses are now readily available in spreadsheet format (Low & Tang, 2007). It is a simple matter to link the performance function used by these algorithms to existing spreadsheet-based calculation models which most designers have at their disposal, provided these have a closed-form solution. The problem becomes a little more complicated in cases where closed-form solutions are not available. Slope stability is an example.

The most common problem faced by the geotechnical designer is the quality and sufficiency of the available geotechnical data. No increase in the sophistication of the design method can compensate for poor data. Thus, irrespective of the design method used, the designer is still required

to exercise judgement and the outcome of the analysis is only as good as that judgement and the data on which it is based.

3.4 CONCLUSIONS

Reliability analyses of Eurocode-compliant designs for a number of common geotechnical structures have shown remarkable consistency in the reliability index for different types of structures and ranges in material parameters. For the assumptions made in the analyses, the reliability indices achieved agree reasonably with the target reliability index given in the Eurocodes. The same cannot be said of working stress design methods where global factors of safety vary widely across the range of structures and material properties. In some instances, changes in material parameters led to an increase in the global factor of safety while the reliability index decreased.

Reliability based design methods have reached the stage where they can be readily implemented for routine design of problems with closed-form solutions. These methods overcome some of the perceived shortcomings of limit states design such as the use of partial factors and identification of critical load combinations. They are still, however, reliant on the selection of an appropriate computational model and the assumptions made in the analysis. The problem with selection of material properties is simply shifted from that of the choice of characteristic values for use in limit states design to the statistical characterization of parameters for reliability based design. Problems with adequacy and quality of geotechnical data apply to both methods and the need for engineering judgement remains.

Until such time a new design codes are developed that recognise reliability based methods and provide guidance on their use, these methods are likely to remain a valuable tool to be used in parallel to currently accepted methods including limit states design.

3.5 REFERENCES

- Bond, A. & Harris, A. (2008). *Decoding Eurocode 7*. Taylor & Francis, London.
- EN1997-1:2004. *Eurocode 7: Geotechnical design–Part 1: General Rules*. European Committee for Standardisation, Brussels.
- Low, B.K and Tang, W.H. (2007). Efficient spreadsheet algorithm for First-Order Reliability Method. *Journal of Engineering Mechanics*, 133(12), 1378-1387.
- Orr, T.L.L. (2005). *Proceedings of the International Workshop on the Evaluation of Eurocode 7*. 21 March and 1 April 2005, Trinity College, Dublin.
- Orr, T.L.L. (2005b). *Review of workshop on the evaluation of Eurocode 7*. *Proceedings of the International Workshop on the Evaluation of Eurocode 7*, p 1-10. Trinity College, Dublin.
- Phoon, K. K. & Kulhawy, F. H. (1999). Characterization of geotechnical variability. *Canadian Geotechnical Journal* 36, 612–624.
- Retief, J.V & Dunaiski, P.E. (2009). *The limit states basis of structural design for SANS 10160-1*. Chapter 1-2. Background to SANS 10160. SUN MeDIA, Stellenbosch, S. Africa.
- Schneider, H.R. (1997). *Panel Discussion: Definition and determination of soil properties*. 14th International Conference on Soil Mechanics and Geotechnical Engineering, Hamburg, 6-12 Sept. Vol. 4. 2271-2274. Balkema, Rotterdam.

Discussion 3A – Impact of model factors and selection of statistical inputs

KK Phoon (National University of Singapore, Singapore)

I would like to offer additional views on an observation made in Section 3.3 pertaining to the selection of different calculation methods. Indeed, different calculation methods produce different solutions. Reliability analysis is just a tool to calculate the reliability index for a selected limit state function and the set of input random variables affecting the limit state function. The selection of an appropriate limit state function (which includes the capacity or movement calculation model) is not related to reliability analysis. It is the role of an adequately trained and qualified engineer to select the appropriate limit state function(s) governing the problem and the accompanying calculation models/input random variables that are realistic for his/her site conditions. It is also the role of an engineer to be mindful of methodological limitations and gaps between a model and reality when interpreting the solutions produced by an analysis. This is identical to the selection of appropriate constitutive models/input parameters in finite element analysis. The solutions produced by any analysis (reliability analysis or finite element analysis) are only sensible when the analysis is guided by sound understanding and judgment.

The issue of different calculation models producing different answers is related to the model factor (or model bias in the AASHTO literature, e.g. Paikowsky et al. 2004; 2010). The model factor can be represented as a lognormal random variable, M with mean = μ_M with coefficient of variation = COV_M . The model factor in its most basic form describes the ratio between a measured response and a calculated response, for example, $M = \text{measured capacity}/\text{calculated capacity}$ is widely used for foundations. More details are given in Chapter 2 of this report and Dithinde et al. (2016). If $Q_n = \text{nominal capacity}$ and $F_n = \text{nominal load}$, then the global factor of safety $FS = Q_n/F_n$ clearly changes depending on how Q_n is calculated. Different calculation models are expected to be associated with different degrees of conservatism or different model biases. However, a corrected $FS = (\mu_M \times Q_n)/F_n$ is relatively insensitive to the choice of calculation model for Q_n because μ_M partially corrects for the average model bias. In the same vein, the reliability index based on the performance function, $G = Q - F$ depends on how Q is calculated but the reliability index based on $G = (M \times Q) - F$ is almost insensitive to the choice of calculation model for Q . The only difference between an allowable stress design (FS) and reliability-based design is that the former only considers average model bias while the latter method can consider COV_M or the probability distribution of M . However, COV_M is typically less important than μ_M when we consider correction for model bias. Hence, in my opinion, the observation made in Section 3.3 that “reliability based design methods do not overcome the problems associated with the selection of different calculation models ...” is correct if we do not consider model factor, i.e. FS and reliability index are both affected by the choice of the calculation model.

For illustration, consider a simple design example of a laterally loaded rigid bored pile: $B = 1$ m, $e = 0.5$ m, s_u is lognormally distributed with mean 50 kPa and $COV = 30$ to 50% , and F is lognormally distributed with mean = 200 kN and $COV = 15\%$. Table 3A-1 presents the effect of capacity models on FS . It is evident that FS can vary between 1.7 and 3.4 for the same design. The corrected FS is more consistent, regardless of how the measured capacity is defined. Table 3A-2 shows that the required D/B to achieve a target reliability index of 3 is different even though the COV of s_u is the same for different capacity models. The application of a model factor in reliability will reduce these differences. Theoretically, some differences would remain, because the COV of M is different even though the COV of s_u is the same. Reliability analysis has the advantage of responding consistently to different uncertainties in both model factors and soil parameters.

Table 3A-1. Impact of capacity model on factor of safety with and without average model factor correction

Capacity model	Factor of safety, $FS =$		
	H_w/F	$\mu_M H_w/F$	$\mu_{MH} H_w/F$
Reese (1958)	3.1	2.8	4.3
Broms (1964)	1.7	2.6	4.0
Randolph & Houlsby (1984)	3.4	2.9	4.5

Notes: H_u = ultimate lateral capacity computed using limit equilibrium analysis, μ_{ML} = model factor for lateral or moment limit defined as measured capacity; μ_{MH} = model factor for hyperbolic limit defined as measured capacity; capacity models and average model factors for different capacity models given in Phoon and Kulhawy (2005).

Table 3A-2. Depth to diameter ratios (D/B) and nominal (uncorrected) factors of safety ($FS = H_u/F$) required to achieve a target reliability index of $\beta = 3$ for different capacity models.

	Reese (1958)		Broms (1964)		Randolph & Houlsby (1984)	
	<i>D/B</i>	<i>FS</i>	<i>D/B</i>	<i>FS</i>	<i>D/B</i>	<i>FS</i>
COV of $s_u = 30\%$						
w/o model factor	4.7	2.8	6.3	2.7	4.3	2.7
Model factor	4.7	2.8	5.5	2.1	4.6	3.0
COV of $s_u = 50\%$						
w/o model factor	6.7	4.9	8.9	4.9	6.4	4.9
Model factor	6.4	4.5	6.9	3.2	6.3	4.8

The second issue concerns the difficulties associated with the selection of statistical inputs for reliability analysis. From the perspective of a practitioner, I would agree with the remark that “the problem with the selection of characteristic values is replaced by the equally onerous challenge of selecting statistical distributions for input parameters and the spatial variation in these parameters throughout the soil mass”. Nonetheless, I would like to offer the following additional views for consideration:

- (1) Both selection of characteristic values and statistical inputs may be burdensome on the practitioner, but the latter can exploit the power of probability theory to do useful stuff such as updating data from comparable sites with site specific data using the Bayesian approach. It is useful to ponder the balance of cost and benefits.
- (2) There is no guideline on how to choose characteristic values for say a set of five input parameters that are correlated. Multiple input parameters are common in finite element analysis. Two negatively correlated parameters will mandate the choice of characteristic value lying below the mean to be associated with another characteristic value lying above the mean to maintain physical consistency (consistency with the scatter plot which is obviously “real” since it is simply a visual representation of measured bivariate data points – this has nothing to do with statistics). You cannot do this by judgment alone in the absence of any information on correlation. Hence, for a general problem, you cannot really avoid some statistical knowledge such as correlation, be it through a statistical construct such as a correlation coefficient or just an empirical scatter plot with no theoretical interpretation. In my opinion, the term “cautious estimate” is not referring to being “cautious” in the choice of the parameter (this implies taking value below the mean for all strength parameters), but “cautious” in its consideration of the effect on the limit state function (any value that moves design further away from limit state function). In short, there is no method to make selection of characteristic value less burdensome if one were to rely on judgment alone.
- (3) The challenge in selection of characteristic value is due to the attempt to sort out many complex and inter-related issues (uncertainties and mechanics) using judgment alone. EN 1997-1:2004, 2.4.5.2(2) recommends that the “characteristic value of a geotechnical parameter shall be selected as a cautious estimate of the value affecting the occurrence of the limit state.” The term “cautious estimate” is related to information uncertainties. The phrase “value affecting the occurrence of the limit state” is related to mechanics – location and shape of the critical failure mechanism.

One can ponder if judgment is meant to replace science or supplement limitations of science (which exists in all rational methods/models). In my opinion, the challenge in selection of statistical inputs is purely practical, e.g. insufficient site data to characterize a distribution, difficult for engineer to perform Bayesian updating etc. We can gradually reduce this

challenge by collecting more global data (which has been done for clays and sands, refer to Ching et al. 2016; Phoon et al. 2016), update limited site-specific data by the Bayesian method (which has been done – refer to Chapters 1, 5 and 6 of this report), provide engineer with a software to perform Bayesian updating automatically (some attempts have been made as described in Chapter 5 but more work is needed). A frequent critique that statistics is not applicable in geotechnical engineering due to small sample size is not quite valid if one adopts the Bayesian interpretation of probability. Bayes’ theorem merely ensures consistent adjustment in a subjective degree of belief in the presence of new evidence. I believe Bayesian informed reasoning is a useful complement to judgment.

References

- Broms, B. B. (1964). Lateral resistance of piles in cohesive soils. *Journal of Soil Mechanics and Foundations Division*, 90(2), 27-63.
- Ching, J. Y., Li, D. Q. & Phoon, K. K. (2016). Statistical characterization of multivariate geotechnical data. Chapter 4, *Reliability of Geotechnical Structures in ISO2394*, Eds. K. K. Phoon & J. V. Retief, CRC Press/Balkema, 2016, 89-126.
- Dithinde, M., Phoon, K. K., Ching, J., Zhang, L. M. & Retief, J. V. (2016). Statistical characterisation of model uncertainty. Chapter 5, *Reliability of Geotechnical Structures in ISO2394*, CRC Press/Balkema, Leiden, The Netherlands.
- EN 1997-1:2004. Eurocode 7: Geotechnical design—Part 1: General rules. European Committee for Standardization (CEN), Brussels, Belgium.
- Paikowsky, S. G., Birgisson, B., McVay, M., Nguyen, T., Kuo, C., Baecher, G. B., Ayyub, B., Stenersen, K., O’Malley, K., Chernauskas, L., and O’Neill, M. (2004). Load and resistance factors design for deep foundations. NCHRP Report 707, Transportation Research Board of the National Academies, Washington DC.
- Paikowsky, S. G., Canniff, M. C., Lesny, K., Kisse, A., Amatya, S., and Muganga, R. (2010). LRFD Design and Construction of Shallow Foundations for Highway Bridge Structures. NCHRP Report 651, Transportation Research Board of the National Academies, Washington DC.
- Phoon K. K. & Kulhawy, F. H. (2005). Characterization of Model Uncertainties for Laterally Loaded Rigid Drilled Shafts. *Géotechnique*, 55(1), 45-54.
- Phoon, K. K., Prakoso, W. A., Wang, Y., Ching, J. Y. (2016). Uncertainty representation of geotechnical design parameters. Chapter 3, *Reliability of Geotechnical Structures in ISO2394*, Eds. K. K. Phoon & J. V. Retief, CRC Press/Balkema, 2016, 49-87.
- Randolph, M. F. & Houlsby, G. T. (1984). Limiting pressure on a circular pile loaded laterally in cohesive soil. *Géotechnique*, 34(4), 613-623.
- Reese, L. C. (1958). Discussion of “Soil modulus for laterally loaded piles”. *Transactions, ASCE* 123, 1071-10.

Chapter 4 EXCEL-Based Direct Reliability Analysis and Its Potential Role to Complement Eurocodes

Lead discussor:

Bak Kong Low

CBKLOW@ntu.edu.sg

Discussors (alphabetical order):

Gregory Baecher, Richard Bathurst, Malcolm Bolton, Philip Boothroyd, Zuyu Chen, Jianye Ching, Peter Day, Nico De Koker, Mike Duncan, Herbert Einstein, Robert Gilbert, D.V. Griffiths, Karim Karam, Suzanne Lacasse, Sónia H. Marques, Lars Olsson, Trevor Orr, Kok-Kwang Phoon, Brian Simpson, Yu Wang, Tien Wu

4.1 INTRODUCTION

This chapter studies the similarities and differences between the design points of the first-order reliability method (FORM) and the Eurocode 7 (EC7). Nine geotechnical examples of reliability analysis and reliability-based design (RBD) are discussed with respect to parametric correlations, sensitivity information from RBD, ultimate limit states (ULS) and serviceability limit states (SLS), system reliability, spatially autocorrelated properties, characteristic values, and partial factors. Focus is on insights from RBD and how RBD can complement EC7 design approach in some situations, but limitations of RBD will also be mentioned. The reliability approach used here is the first-order reliability method (FORM), which extends the Hasofer-Lind index to deal with correlated non-Gaussian random variables. An intuitive perspective of the Hasofer-Lind index, FORM and design point is explained next, so that the symbols and discussions in later sections can be understood. The FORM can be done on the EXCEL platform. With respect to ease of application, the only key distinction between direct reliability and partial factors is the need for engineers to provide realistic statistical inputs describing the uncertainties affecting the limit state function. Section 4.10 describes the recommended practice for determination of these statistical inputs. This practice is in line with the current practice of estimating soil properties based on available site investigation data and data from comparable sites. The main limitation of FORM is that it is less suitable for more complex system reliability problems. Section 4.9 describes a practical subset simulation method (again available in EXCEL) that can mitigate this limitation.

The nine design examples presented in this Chapter show how reliability calculations could relieve engineering judgment from the unsuitable task of performance verification in the presence of uncertainties so that the engineer can focus on setting up the right lines of scientific investigation, selecting the appropriate models and parameters for calculations, and verifying the reasonableness of the results (Peck 1980). In this regard, the role of engineering judgment in reliability-based design is sharpened rather than diminished. By introducing greater realism into reliability analysis that caters to the distinctive needs of geotechnical engineering practice, focusing on how it can add genuine value to the profession, its clients, and the public, and be mindful of its limits, the discussion group believes that reliability analysis could play a useful complementary role in geotechnical design.

4.1.1 Intuitive expanding ellipsoid perspective for Hasofer-Lind index and FORM

The matrix formulation of the Hasofer-Lind (1974) index β is:

$$\beta = \min_{\mathbf{x} \in F} \sqrt{(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{C}^{-1} (\mathbf{x} - \boldsymbol{\mu})} = \min_{\mathbf{x} \in F} \sqrt{\left[\frac{x_i - \mu_i}{\sigma_i} \right]^T \mathbf{R}^{-1} \left[\frac{x_i - \mu_i}{\sigma_i} \right]} \quad (4-1)$$

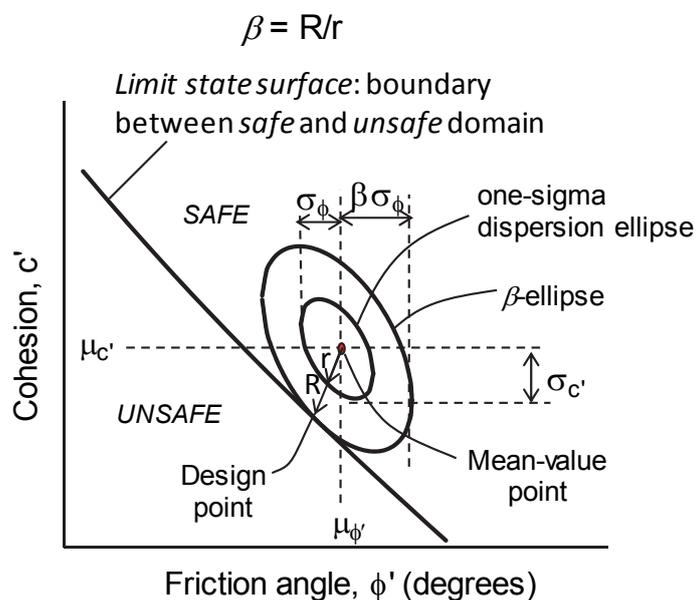
where \mathbf{x} is a vector representing the set of random variables x_i , $\boldsymbol{\mu}$ the vector of mean values μ_i , \mathbf{C} the covariance matrix, F the failure domain, \mathbf{R} the correlation matrix, and σ_i the standard deviations. The notations “T” and “-1” denote transpose and inverse, respectively.

The point denoted by the x_i values, which minimize Eq. (4.1) and satisfies $\mathbf{x} \in F$, is the most probable failure combination of parametric values (also known as the design point). It is the point of tangency of an expanding dispersion ellipsoid with the limit state surface (LSS), which separates safe combinations of parametric values from unsafe combinations (Fig. 4-1). The one-standard-deviation dispersion ellipse and the β -ellipse in Fig. 4-1 are tilted due to negative correlation between c' and ϕ' . The quadratic form in Eq. (4-1) appears also in the negative exponent of the established probability density function of the multivariate normal distribution. As a multivariate normal dispersion ellipsoid expands from the mean-value point, its expanding surfaces are contours of decreasing probability values. Hence, to obtain β by Eq. (4-1) means finding the smallest ellipsoid tangent to the LSS at the most probable point of failure (the design point). More details in Low (2015).

FORM extends the Hasofer-Lind index to deal with correlated non-Gaussian random variable, as explained in Ang and Tang (1984), Melchers (1999), and Baecher and Christian (2003), for example. In FORM, one can rewrite Eq. (4-1b) as follows (Low and Tang, 2004):

$$\beta = \min_{\mathbf{x} \in F} \sqrt{\left[\frac{x_i - \mu_i^N}{\sigma_i^N} \right]^T \mathbf{R}^{-1} \left[\frac{x_i - \mu_i^N}{\sigma_i^N} \right]} \quad (4-2)$$

where μ_i^N and σ_i^N are calculated by the Rackwitz-Fiessler (1978) transformation. Hence, for correlated nonnormals, the ellipsoid perspective still applies in the original coordinate system, except that the nonnormal distributions are replaced by an equivalent normal ellipsoid, centered not at the original mean values of the nonnormal distributions, but at the equivalent normal mean $\boldsymbol{\mu}^N$.



$\tan\phi'$ instead of ϕ' can be used in the figure above in line with EC7

Figure 4-1 Illustration of the reliability index β in the plane where c' and ϕ' are negatively correlated. This perspective is also valid for non-normal distributions, when viewed as “equivalent ellipsoids”.

Eq. (4-2) and the Rackwitz-Fiessler equations for μ_i^N and σ_i^N were used in the spreadsheet-automated constrained optimization FORM computational approach in Low & Tang (2004). An alternative to the 2004 FORM procedure is given in Low & Tang (2007), which uses the following equation for the reliability index β :

discuss insights. The mean value of spring stiffness k_3 at point 3 is 10 N/mm, and that of the rotational restraint λ_1 at point 1 is 500 Nmm/rad. Design point values from reliability analysis are indicated under the column labelled " x_i^* ".

FORM analysis reveals that the design values of λ_1 and k_3 hardly deviate from their mean values. This means that the buckling load P_{cr} is insensitive to these two parameters at their mean values of 500 and 10. In this case, the strut becomes a full sine curve, i.e. arching upwards on one side of point 3, and downwards on the other side, when $k_3 = 4$. Higher stiffness of k_3 serves no purpose once the full sine wave is formed. The same conclusion of insensitivity of λ_1 and k_3 was reached in Coates et al. (1994) after much plotting via deterministic analysis. For the case in hand, sensitivity of k_3 increases at lower values ($k_3 < 4$) when the strut has not gone into a full sine wave yet.

Distributions	Para1	Para2	Para3	Para4	x_i^*	units	n_i
Lognormal	P	700	140		973.95	N	1.7667
Triangular	L	800	1000	1200	1109.50	mm	1.2681
Lognormal	a	500	50		539.42	mm	0.8107
BetaDist	E	3	3	150000 250000	174306	N/mm ²	-1.3027
Lognormal	I	200	20		174.50	mm ⁴	-1.3173
PERTDist	λ_1	350	500	650	499.99	Nmm/rad	-0.0002
Gamma	k_3	100	0.1		9.97	N/mm	-0.0002

β	$\Phi(-\beta)$	Monte Carlo: $P_f = 0.36\%$ (250,000 trials)	P_{cr}	PerFunc
2.6513	0.40%		973.95	0.000

1	0	0	0	0	0	0
0	1	0.7	0	0	0	0
0	0.7	1	0	0	0	0
0	0	0	1	0.5	0	0
0	0	0	0.5	1	0	0
0	0	0	0	0	1	0.6
0	0	0	0	0	0.6	1

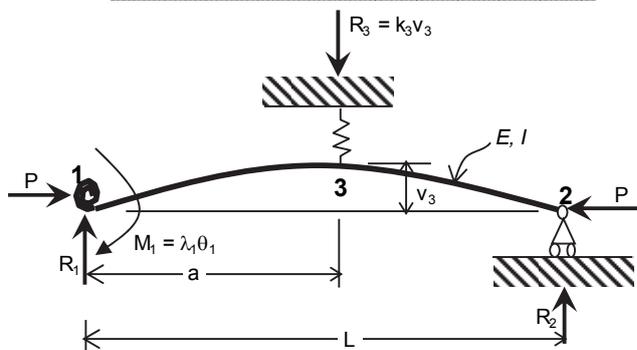


Figure 4.3 Excel-Solver reliability analysis of a strut with complex supports. Performance function is implicit.

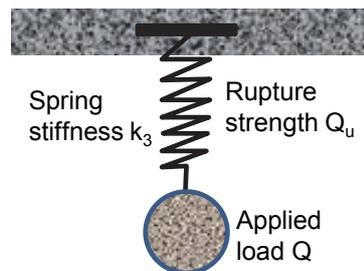


Figure 4.4 Spring suspending a vertical load.

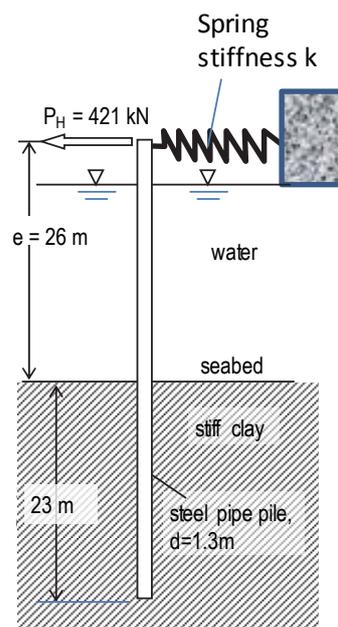


Figure 4.5. Horizontal spring at the head of a cantilever pile.

However, similar restraints in the simple system of Fig. 4.4 and near the cantilever pile head of the laterally loaded pile of Fig. 4.5, would be important and sensitive parameters to the SLS of vertical displacement and ULS of spring rupture of Fig. 4.4, and the SLS of pile head deflection and the ULS of spring rupture and pile bending failure of Fig. 4.5. Partial factors of spring stiffness k back-calculated from reliability analysis are of different values within the same problem and across different problems. Hence direct FORM reliability analysis and reliability-based design (RBD) are preferred. Partial factors or characteristic values back-calculated from FORM will not be pursued in this study, except when discussing the limitations of partial factors.

4.3 A GRAVITY RETAINING WALL AND AN ANCHORED SHEET PILE WALL

4.3.1 Reliability-based design of a gravity retaining wall

In Fig. 4.6 it is assumed that rotating mode is one of the ULS to be checked. $\tan\phi'$, $\tan\delta$ and c_a are normally distributed, with mean values (μ) and standard deviations (σ) as shown, and with correlation coefficient 0.8 between $\tan\phi'$ and $\tan\delta$. The x^* values are the design point values obtained in FORM reliability analysis. The column labelled **n** shows the values of $(x^* - \mu)/\sigma$.

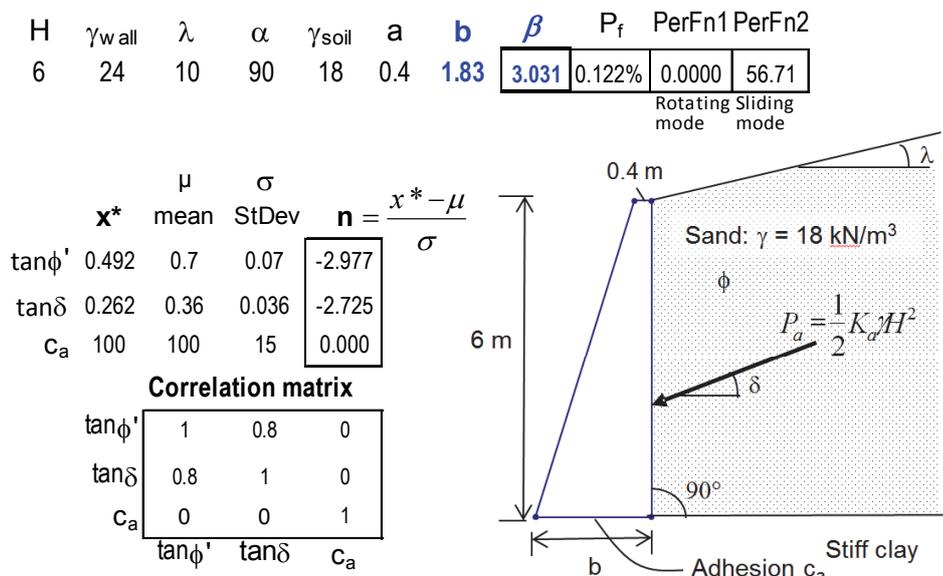


Figure 4.6 Reliability-based design of the base width b of a gravity retaining wall

For the statistical inputs shown, RBD obtains a design base width $b = 1.83$ m for a target β index of about 3.0, corresponding to a probability of rotation failure $P_f \approx \Phi(-\beta) = 0.122\%$. For comparison, Monte Carlo simulation with 200,000 realizations using @RISK (www.Palisade.com) yields a P_f of 0.120%. The n value of 0.0 for c_a means the design value of c_a (under the x^* column) stays put at its mean value, because rotating mode is not affected by c_a at all. Reliability analysis reveals input sensitivities. Reliability analysis with respect to the ultimate limit states of sliding and bearing capacity failure can be done, and system reliability for multiple limit states can be evaluated readily, for example using the Low et al. (2011) system reliability procedure.

4.3.2 Comparison with EC7 DA1b design of base width b for rotation ULS

Figure 4.7 shows EC7 Design Approach 1 Combination 2 (referred to as DA1b in this chapter) for the base width b with respect to the overturning ULS, via characteristic values and partial factors, starting from the same statistical inputs of mean values and standard deviations, but without considering correlations in EC7. Even though partial factors are specified, EC7 does not produce a unique design, but depends on how conservative the characteristic values are determined. This is not objectionable, for it allows flexibility in design to match the consequence of failure; in the same way that target reliability index can be higher or lower depending on the consequence of failure. Analogous situation exists for LRFD’s nominal values and load and resistance factors.

For a design width b obtained via EC7, the value of the corresponding reliability index β is not unique, but depends on whether parametric correlations (if any) are modelled. To compare with the target β of 3.0 in RBD (Fig. 4.6), correlations should be modelled.

EC7 Design Approach 1 Combination 1 (referred to as DA1a in this chapter) requires characteristic values of resistance and actions, on which partial factors are applied. If characteristic values are based on percentiles, one needs to know the probability distributions of actions and

resistance in order to estimate the upper tail (e.g. 70 percentile) characteristic value of actions and lower tail characteristic values of resistance (e.g. 30 percentile). For the case in hand, whether based on 5%/95% or 30%/70%, DA1a is satisfied; DA1b governs.

	Normal		EC7 DA1b, 5/95 percentiles for x_k			EC7 DA1b, 30/70 percentiles for x_k		
	μ	σ	x_k	Partial F	x^*	x_k	Partial F	x^*
$\tan\phi'$	0.7	0.07	$\tan\phi'$ 0.5849	1.25	0.4679	$\tan\phi'$ 0.6633	1.25	0.5307
$\tan\delta$	0.36	0.036	$\tan\delta$ 0.3008	1.25	0.2406	$\tan\delta$ 0.3411	1.25	0.2729
c_a	100	15	c_a 75.33	1.4	53.80	c_a 92.14	1.4	65.81

x_k = Characteristic value x^* = Design value Computed EC7 design $b = 1.90$ m for which $\beta = 3.47$ if ϕ - δ correlation is modelled, and $\beta = 4.36$ if not.	Computed EC7 design $b = 1.755$ m for which $\beta = 2.53$ if ϕ - δ correlation is modelled, and $\beta = 3.18$ if not.
--	---

Figure 4.7 Design of the base width b for rotation ULS of a gravity retaining wall, based on Eurocode 7 Design Approach 1 Combination 2 (DA1b).

4.3.3 Reliability-based design of the total height H of anchored sheet pile wall

In Fig. 4.8, free-earth support method was used, with K_a based on Coulomb formula, and K_p based on Kerisel-Absi chart. For the statistical inputs shown, RBD for a target $\beta = 3.0$ results in design H ($= 6.4 + z^* + d^*$) of 12.31 m, and $P_f \approx \Phi(-\beta) = 0.13\%$. For comparison, Monte Carlo simulation with 200,000 realizations gives $P_f = 0.14\%$. For the statistical inputs shown, $\tan\phi'$ and z are sensitive parameters, as indicated by the n values. The n values of $\tan\delta$ and γ are due largely to correlations with $\tan\phi'$, revealed if uncorrelated analysis is done. The design value of γ , 16.13 kN/m³, is lower than its mean value of 17 kN/m³, an apparent paradox which can be understood due to the logical positive correlation of γ to γ_{sat} and $\tan\phi'$ which both have design values below their respective mean values. If all six parameters are uncorrelated, the design value γ^* will be bigger than mean γ . Soil on either side is assumed to be same source, hence the same γ_{sat} must be used, with action-resistance duality. Reliability analysis yields $\gamma_{sat}^* = 17.32$ kN/m³, which is less than the mean γ_{sat} of 19 kN/m³.

The mean embedment depth $d = 12.31 - 6.4 - 2.4 = 3.51$ m. The design embedment depth $d^* = 2.99$ m, i.e., “overdig” = 0.52m, which is determined automatically as a by-product of RBD. More discussions are available in Low (2005).

4.3.4 Comparison with EC7 DA1b design of sheet pile total height H

The EC7 DA1b design for the anchored sheet pile wall is shown in Fig. 4.9. EC7 has an “unforeseen overdig” allowance for z , to account for the uncertainty of the dredge level. The design value of z is obtained from $\mu_z + 0.5$ m = 2.4 + 0.5 = 2.9 m, where 0.5 m is the “overdig”. Although EC7 partial factor of soil unit weight is specified to be 1.0, conservative characteristic values of γ and γ_{sat} still need to be estimated, and if originating from the same source, it is not logical to increase the unit weight on the active side while decrease the unit weight on the passive side. Also, assuming 5/95 percentiles for characteristic values leads to $\gamma^* > \gamma_{sat}^*$, which violates soil physics.

With characteristic values at 30/70 percentiles and EC7 partial factors from DA1b, one obtains a design H of 12.87 m, closer to the RBD design H of 12.31 m for a target β of 3.0. A less critical design H of 12.64 m is obtained if one wrongly set the characteristic value of γ_{sat} at the 70 percentile value (19.52 kN/m³) instead of at the 30 percentile value (18.48 kN/m³).

FORM reliability analysis based on the H from EC7 design will give different β index depending on whether correlations are modeled (correlation matrix, Fig. 4.8) or not.

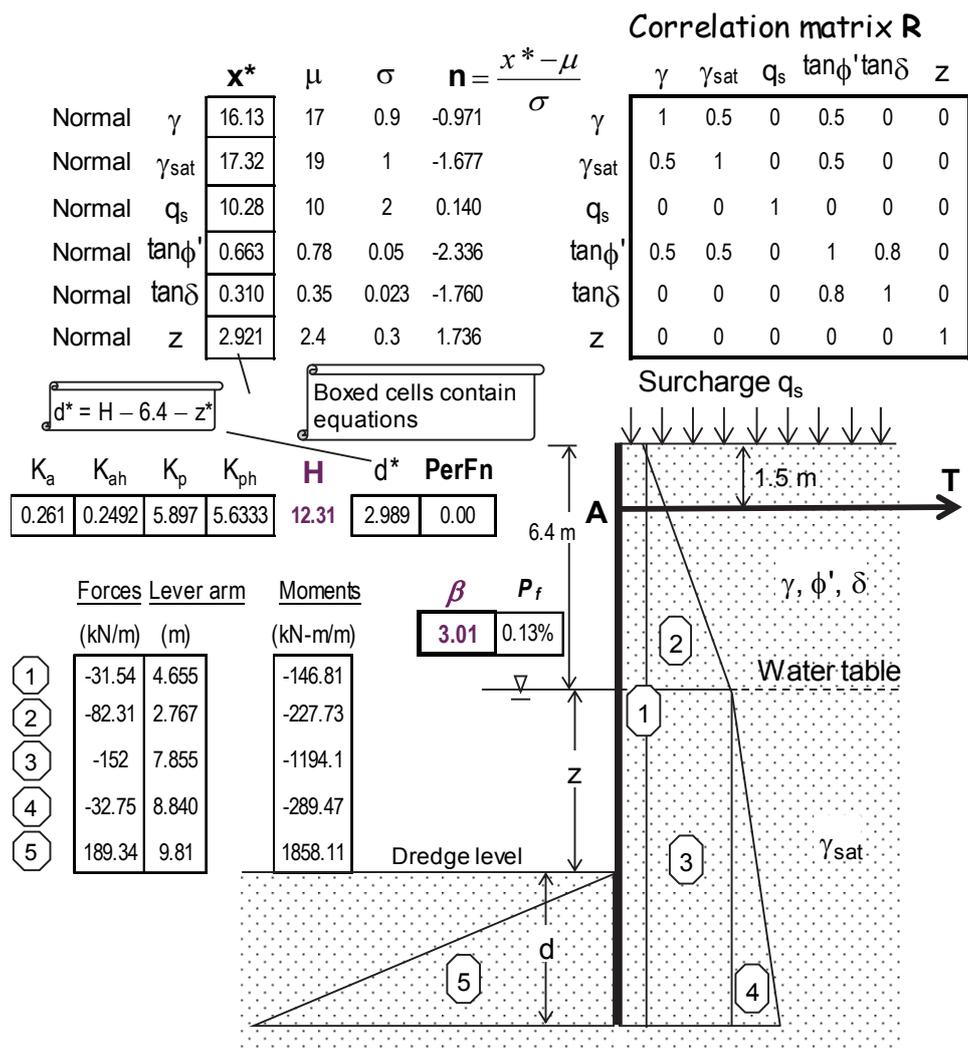


Figure 4.8 Reliability-based design of sheet pile total length H via FORM

Normal	EC7 DA1b, 5/95 percentiles for x_k			EC7 DA1b, 30/70 percentiles for x_k				
	μ	σ	x_k	Partial F	x^*	x_k	Partial F	x^*
γ	17	0.9	18.48	1.00	18.48	17.47	1.00	17.47
γ_{sat}	19	1	17.36	1.00	17.36	18.48	1.00	18.48
q_s	10	2	13.29	1.30	17.28	11.05	1.30	14.36
$\tan\phi'$	0.78	0.05	0.698	1.25	0.558	0.754	1.25	0.603
$\tan\delta$	0.35	0.023	0.312	1.25	0.250	0.338	1.25	0.270
z	2.4	0.3	NA	NA	2.90	NA	NA	2.90

x_k = Characteristic value
 x^* = Design value
 Computed EC7 design H = 13.95 m
 Note: Unrealistic that $\gamma^* > \gamma_{sat}^*$

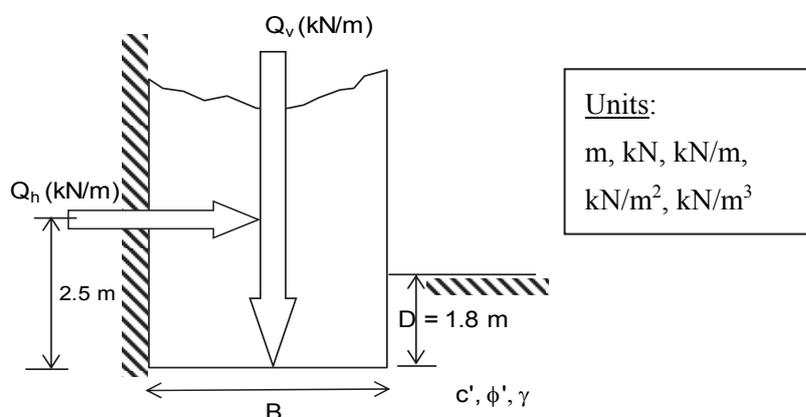
Computed EC7 design H = 12.87 m.
 Computed EC7 design H = 12.64 m if $\gamma_{sat,k}$ wrongly set at 70 percentile value (19.52) instead of the 30 percentile value (18.48).

Figure 4.9 Eurocode 7 DA1b design of sheet pile total length H

4.4 A FOOTING FOUNDATION AND A LATERALLY LOADED PILE

4.4.1 RBD of base width of a retaining wall against bearing capacity failure mode

Figure 4.10 is similar to a case in Low and Phoon (2015), except that Q_h and Q_v are uncorrelated, and the uncertainty in $\tan\phi'$ is modelled instead of ϕ' , in line with EC7 which applies partial factor to $\tan\phi'$. The mean values of c' , $\tan\phi'$, Q_h and Q_v are 15 kPa, 0.47, 300 kN/m and 1100 kN/m, respectively, based on deterministic Example 2.2 in Tomlinson (2001), which computed an F_s of 3.0 against general shear failure of the base of the wall when base width B is 5 m. For the statistical inputs shown, a base width B of 4.55 m is required to achieve a target reliability index of $\beta = 3.0$ against bearing capacity failure. The design value of c' , 15.03 kPa, is slightly above the c' mean value of 15 kPa, due to negative correlation coefficient of -0.5 between c' and $\tan\phi'$. For the case in hand, the design is much more sensitive to Q_h than to Q_v , with n values 2.49 versus -0.71, and much more sensitive to $\tan\phi'$ than c' , with n values -1.37 versus 0.11, where $n = (x^* - \mu^N)/\sigma^N$, in which superscript N denotes equivalent normal mean and equivalent normal standard deviation of lognormal distributions.



$$q_u = cN_c s_c d_c i_c + p_o N_q s_q d_q i_q + \frac{B'}{2} \gamma N_\gamma s_\gamma d_\gamma i_\gamma$$

Distr.Name	x^*	μ mean	σ StDev	n	Correlation matrix R			
Lognormal c'	15.03	15	3	0.11	1	-0.5	0	0
Lognormal $\tan\phi'$	0.41	0.47	0.047	-1.37	-0.5	1	0	0
Lognormal Q_h/m	429.82	300	45	2.49	0	0	1	0
Lognormal Q_v/m	1019.7	1100	110	-0.71	0	0	0	1

B	L	D	γ	e_B	B'	e_L	L'	q	$q_u(x^*)$
4.55	25	1.8	21	1.054	2.442	0	25	417.5	417.5

c_a	p_o	Ω	N_q	N_c	N_γ	s_q	s_c	s_γ	i_q	i_c	i_γ
12.0224	37.8	22.856	7.967	17.09	7.313	1.021	1.024	0.987	0.578	0.518	0.3235

PerFunc $g(x)$	β
0.00	3.00

d_q	d_c	d_γ
1.125	1.158	1

Figure 4.10 Reliability-based design of retaining wall base width B

The RBD in Fig. 4.10 assumes that the coefficients of variation of c' , $\tan\phi'$, Q_h and Q_v are 0.2, 0.1, 0.15 and 0.1, respectively. It is also assumed that c' and $\tan\phi'$ are negatively correlated (as shown in the correlation matrix), but Q_h and Q_v are uncorrelated (as befitting the horizontal earth thrust and the applied vertical load).

When $Q_h = 0$, the vertical load Q_v is an unfavorable action without ambiguity. However, when Q_h is acting and of comparable magnitude to Q_v , the latter possesses action-resistance duality, because load inclination and eccentricity decreases with increasing Q_v . RBD automatically takes this action-resistance (or unfavorable-favorable) duality into account in locating the design point. Interestingly, RBD reveals that the design value of Q_v (1019.7 kN/m) is about 7.3% lower than its mean value of 1100 kN/m, thereby revealing the action-resistance duality of Q_v when Q_h is acting.

Note that the bearing capacity equation is approximate even for idealized conditions. Also, several expressions for N_γ exist. The N_γ used here is attributed to Vesic in Bowles (1996). The nine factors s_j , d_j and i_j account for the shape and depth effects of foundation and the inclination effect of the applied load. The formulas for these factors are based on Tables 4.5a and 4.5b of Bowles (1996), which may differ from those in EC7.

RBD can be done for system reliability with multiple failure modes of ULS and/or SLS, as illustrated for a laterally loaded pile next.

4.4.2 Multi-criteria RBD of a laterally loaded pile in spatially autocorrelated clay

Figure 4.11 shows a steel tubular pile in a breasting dolphin, which was analysed deterministically in Tomlinson (1994), and probabilistically in Low et al. (2001). Soil-pile interaction was based on the nonlinear and strain-softening Matlock p-y curves. At the mean input values of P_H and undrained shear strength c_u , the pile deflection y is 0.06 m at seabed, and 1 m at pile head. For reliability analysis, the P_H was assumed to be normally distributed, with mean value 421 kN and a coefficient of variation of 25%. The mean c_u trend is $\mu_{c_u} = 150 + 2z$, kPa, with a coefficient of variation of 30%. Spatial autocorrelation was modelled for the c_u values at different depths below seabed. The β index obtained was 1.514 with respect to yielding at the outer edge of the annular steel cross section. The sensitivities of P_H and c_u change with the cantilever length e . The different sensitivities from case to case are automatically revealed in reliability analysis and RBD, but will be difficult to consider in codes based on partial factors.

A target β of 3.0 can be achieved in RBD for both ULS (bending) and SLS (assuming $y_{Limit} = 1.4$ m) using steel wall thickness $t = 32$ mm and external diameter $d = 1.42$ m, Fig. 4.11(c).

4.4.3 Questions and thoughts pertinent to sections 4.4.1 and 4.4.2

For the footing of Fig. 4.10, how would partial factor design approaches (e.g. EC7 and LRFD) deal with a parameter that possesses action-resistance duality (i.e., unfavorable-favorable duality), such as the vertical load Q_v in the presence of horizontal load Q_h ?

The laterally loaded pile example of Fig. 4.11 is one of a group of piles in a breasting dolphin, with 23 m embedment length below seabed and 26 m cantilever length in sea water. For both the bending ULS and the pile head deflection SLS, the design point in RBD shows decreasing sensitivity of c_u with depth, i.e., decreasing $(c_u^* - \mu_{c_u})/\sigma_{c_u}$ with depth, where c_u^* are the design undrained shear strength values at various depths obtained in RBD. How would partial factor design approaches determine the characteristic (or nominal) values of the undrained shear strength at different depths? Assuming uniform conservatism with depth in determining the characteristic c_u values do not accord well with the different sensitivities of c_u with depth as revealed by RBD, and may alter the behavior of the pile at ULS and SLS.

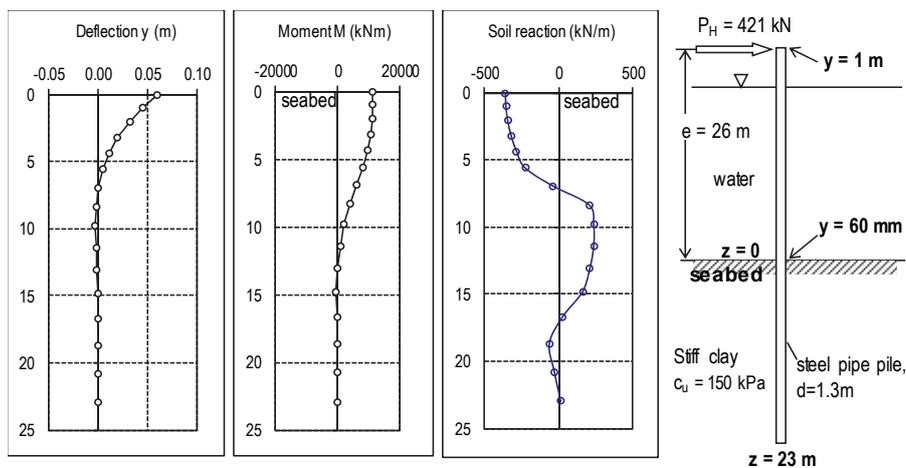
For ULS design (e.g. bending of pile), having obtained the conservative c_u characteristic values, should one apply the partial factor for c_u uniformly across the entire embedded portion of the pile despite different sensitivities revealed in RBD?

4.5 EXAMPLE RELIABILITY ANALYSIS OF SOIL SLOPES

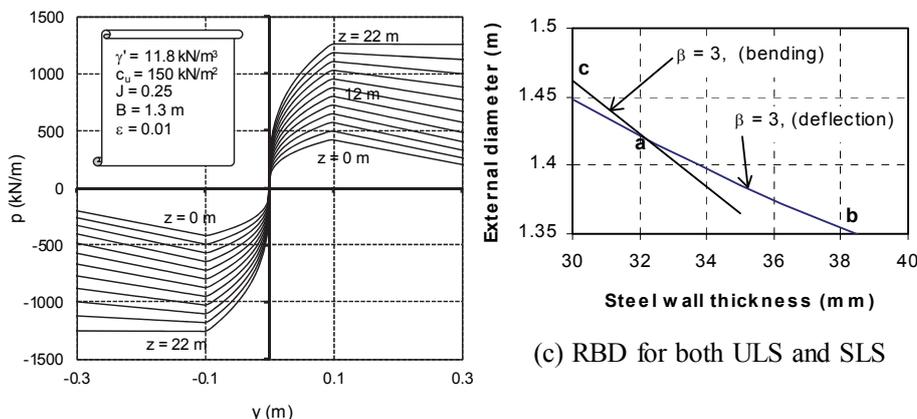
4.5.1 Underwater excavated slope failure in San Francisco Bay Mud

The 1970 failure of a slope excavated underwater in San Francisco Bay (Fig. 4.12) was part of a temporary excavation and was designed with an unusually low factor of safety to minimize construction costs. During construction a portion of the excavated slope failed. Low and Duncan (2013) analyzed it, first deterministically using data from field vane shear and laboratory triaxial tests,

then probabilistically, accounting for parametric uncertainty and positive correlation of the undrained shear strength and soil unit weight.



(a) Laterally loaded cantilever pile in soil with Matlock p-y curves.



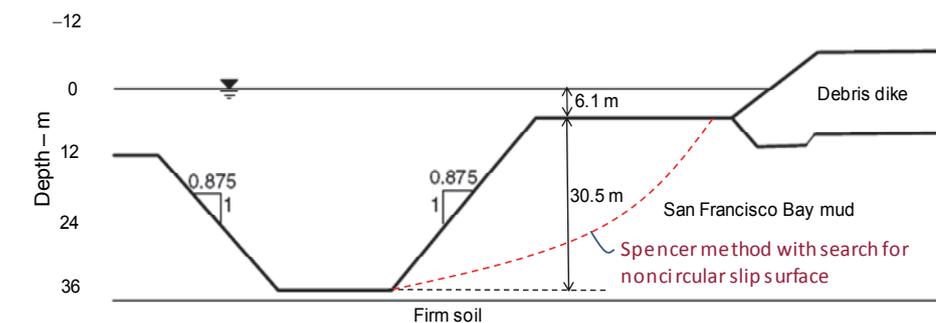
(b) Depth-dependent Matlock p-y curves.

Figure 4.11 Reliability-based design of a laterally loaded pile for ULS and SLS

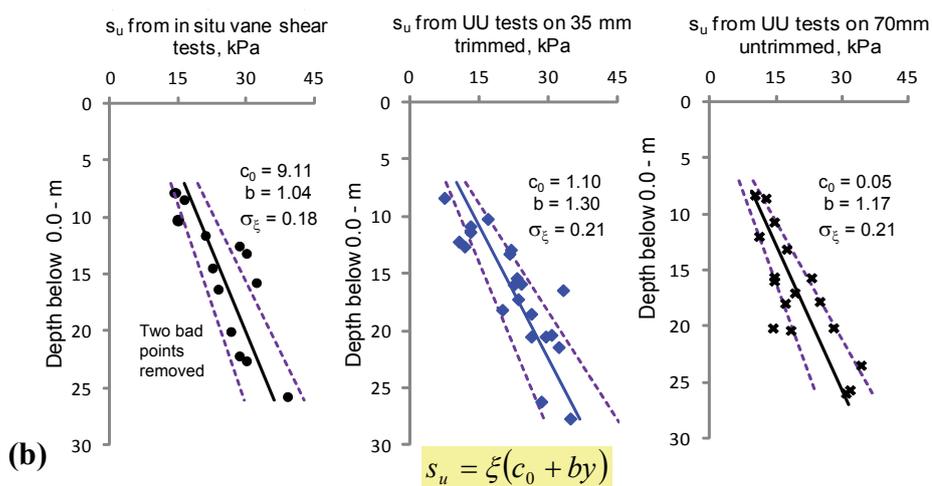
The computed lumped factors of safety with search for critical noncircular slip surface based on Spencer method are 1.20, 1.16 and 1.00, based on field vane test data, trimmed 35 mm diameter and untrimmed 70 mm diameter specimens in UU triaxial tests, respectively. The FORM analyses and Monte Carlo simulations for circular slip surfaces produce probabilities of failure of about 10%, 19% and 46%, respectively, all unacceptably high.

The results of both the deterministic and the probabilistic analyses are affected by biases in the strength measurements and interpretations. The measured strength values were affected by disturbance and rate of loading effects. Subtle errors were also caused by extrapolation of the undrained shear strength (in situ and lab tests data, available only for the upper 21 m of the Bay mud, from depth 6 m to depth 27 m) to the full depth of underwater excavation. Since the midpoint of a slip circular arc is at about the two-third depth, this means that in the slope stability analysis, half the slip surface was based on extrapolated strength. Nevertheless, the FORM P_f values (10%, 19% and 46%) are much higher than the P_f of about 0.6% for the commonly required β of 2.5, or P_f of 0.14% for a target β of 3.0. Hence a failure was not unlikely, and did happen.

The probability of failure from FORM should be regarded as nominal rather than precise. Nevertheless, RBD via FORM can detect unacceptably high P_f on the one hand, and achieve design aiming at a sufficiently low P_f on the other hand.



(a) A 75 m long section of the 600 m long trench failed during excavation 1970.



(b)

Tests	c_0 (kPa)	Slope of trend line, b	F_ξ , critical circular surface	F_ξ , critical noncircular surface
All data combined	2.23	1.21	1.15	1.12
In situ vane shear including two bad low-values	2.71	1.35	1.29	1.26
In situ vane shear excluding two bad tests	9.11	1.04	1.23	1.20
35 mm trimmed triaxial test specimens	1.10	1.30	1.19	1.16
70 mm untrimmed triaxial test specimens	0.05	1.17	1.03	1.00

(c)

Figure 4.12 (a) An underwater slope failure in San Francisco Bay mud, (b) Mean trend and variations of undrained shear strength values, and (c) Deterministic analysis based on mean values of s_u , using Spencer method for circular and noncircular slip surfaces.

4.5.2 Reliability analysis of a Norwegian slope accounting for spatial autocorrelation

Spatial autocorrelation arises in geological material by virtue of its formation by natural processes acting over unimaginably long time. This endows geomaterial with some unique statistical features (e.g. spatial autocorrelation) not commonly found in structural material manufactured under strict quality control.

A clay slope in southern Norway was analyzed in Low et al. (2007) using Spencer method and FORM, Fig. 4.13. Reliability analysis revealed that the slope is less safe when the unit weights near the toe are lower. This implication can be verified by deterministic runs using higher γ values near the toe, with resulting higher factors of safety. The design point (of 24 spatially correlated c_u values and 24 spatially correlated soil unit weight values) is located automatically in reliability analysis, and reflects parametric sensitivity from case to case in a way specified partial factors cannot.

The results of reliability analysis are only as good as the statistical input and reliability method used (e.g., FORM or SORM), in the same way that the results of deterministic analysis are only as good as the deterministic input and method used (e.g. Spencer method or other methods). A reliability analysis requires additional statistical input information which is not required in a deterministic

factor-of-safety approach, but results in richer information pertaining to the performance function and the design point that is missed in a deterministic analysis.

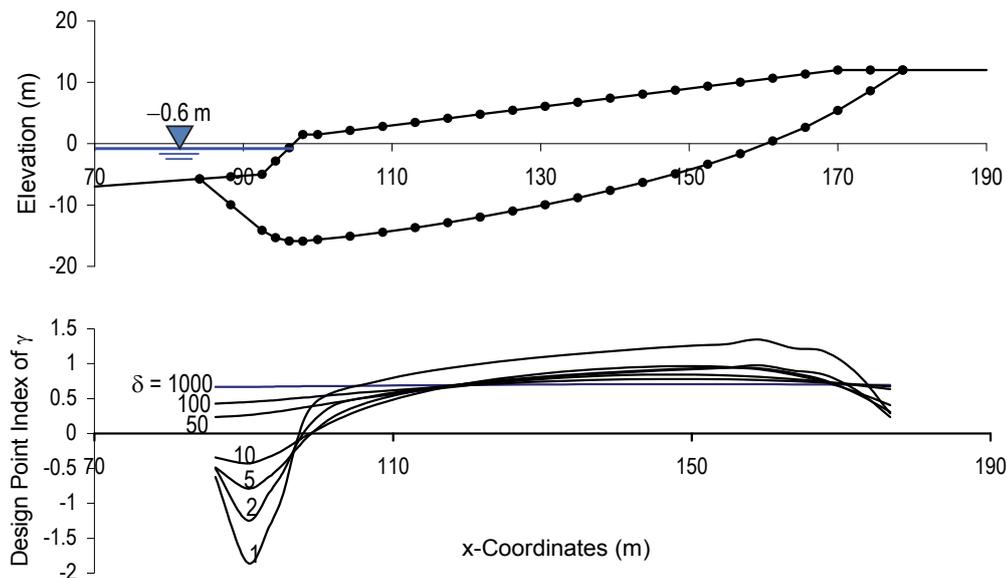


Figure 4.13 Slope reliability analysis accounting for spatially autocorrelated anisotropic undrained shear strength and soil unit weight.

4.6 ROCK SLOPES AND TUNNELS IN ROCK

4.6.1 Reliability-based design of Sau Mau Ping rock slope of Hong Kong

Figure 4.14 shows the FORM reliability-based design of a two-dimensional rock slope with five correlated random variables, two of which obey the highly asymmetric truncated exponentials. The statistical inputs follow those in Hoek (2007). For zero reinforcing force T and uncorrelated parameters, the FORM reliability index is $\beta = 1.556$, and $P_f \approx 1 - \Phi(\beta) = 6\%$, in good agreement with the Monte Carlo P_f of 6.4% in Hoek (2007).

With negative correlations between c and ϕ and between z and z_w/z , as shown in the correlation matrix R , a reinforcing force T of 257 tons (per m length of slope) inclined at $\theta = 55^\circ$ is needed to achieve a target reliability index β of 3.0. The most sensitive parameters for the case in hand, based on the values under the column labelled “ n ”, are the coefficient of horizontal earthquake acceleration α and the ratio z_w/z , followed by friction angle ϕ and cohesion c of the rock joint. The design point values of resistant parameters c and ϕ , at 8.11 t/m^2 and 29.65° respectively, are lower than their mean values of 10 t/m^2 and 35° respectively.

The tension crack depth z and the extent to which it is filled with water (z_w/z) are negatively correlated. This means that shallower crack depths tend to be water-filled more readily (i.e., z_w/z ratio will be higher) than deeper crack depths, consistent with the scenario suggested in Hoek (2007) that the water which would fill the tension crack in this Hong Kong slope would come from direct surface run-off during heavy rains. For illustrative purposes, a negative correlation coefficient of -0.5 is assumed between z and z_w/z .

For the reinforced rock slope of Fig. 4.14, the design point is where the 5D expanding ellipsoid (or equivalent dispersion ellipsoid when nonnormal distributions are involved) is tangent to the limit state surface, similar to the 2D case shown in Fig. 4.1. A reliability-based approach like the one presented here is able to locate the design point case by case and in the process reflect parametric sensitivities (related to the limit state surface and hence is application specific) and correlation structure in a way that design based on prescribed partial factors cannot. More discussions are given in Low and Phoon (2015).

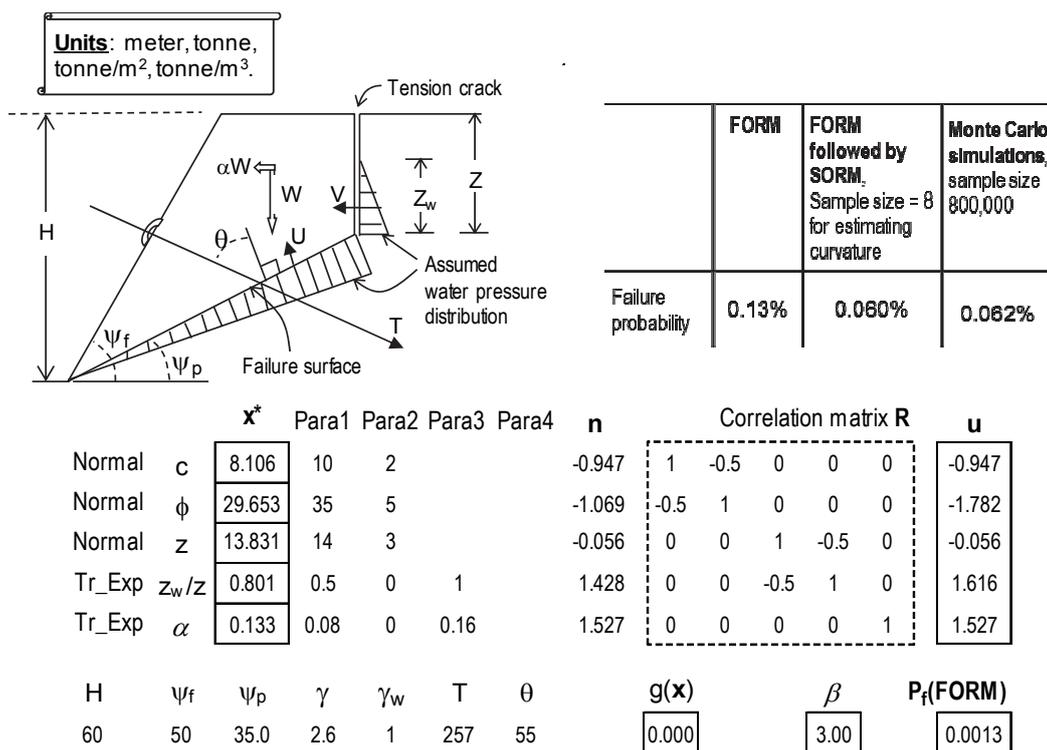


Figure 4.14 A reinforcing force of 257 tons/m is required to provide the rock slope with a reliability index β of 3.0

4.6.2 Reliability analysis of 3-D tetrahedral wedge mechanism in rock slope

The analysis of the stability of wedges in rock slopes requires resolution of forces in three-dimensional space. The problem has been extensively treated, for example in Hoek and Bray (1977). The methods used include stereographic projection technique, engineering graphics, and vector analysis.

Low (1989, 2007) presented compact closed form equations for analyzing the stability of two-joint tetrahedral wedges. In Fig. 4.15, the uncertainties of discontinuity orientations ($\beta_1, \delta_1, \beta_2, \delta_2$), shear strength of joints ($\tan\phi$ and $c/\gamma h$), and water pressure in joints (dimensionless parameter G_w) are modelled by the versatile beta general distributions which can assume non-symmetrical bounded pdf.

The reliability analysis here assumes the means and standard deviations of $\tan\phi$, $c/\gamma h$ and G_w on joint plane 1 are identical to those on joint plane 2. These assumptions are for simplicity, not compulsory. Reliability analysis yielded $\beta = 1.924$ against sliding on both planes, $\beta = 1.389$ against sliding on plane 1, and $\beta > 5$ for other modes.

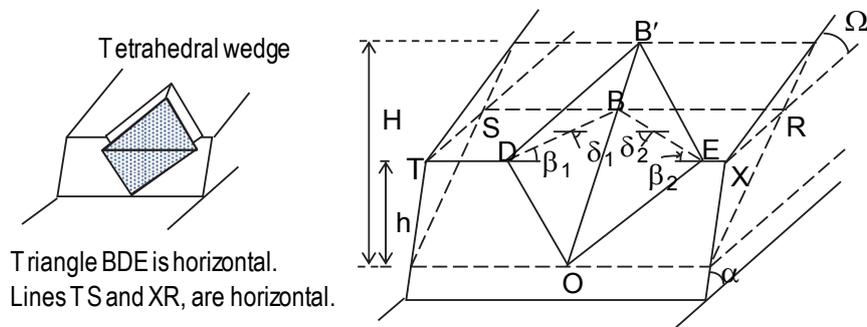
Although the governing failure mode at mean values is sliding on both planes, the reliability index β against sliding on plane 1 is—in the presence of uncertainty in discontinuity orientations ($\beta_1, \delta_1, \beta_2, \delta_2$)—more critical than that against sliding on both planes. This information would not be revealed in a deterministic analysis, or in a reliability analysis that considers only one failure mode.

4.6.3 Reliability analysis of tunnels in rocks

Low and Einstein (2013) discussed the ambiguous nature of the factor of safety of a tunnel with a roof wedge, Fig. 4.16, where two different definitions of the F_s are shown to be reconcilable via the first-order reliability method (FORM). RBD via FORM was then applied to a circular tunnel supported with elastic rockbolts in elasto-plastic ground with the Coulomb failure criterion (Fig. 4.16, top right). The spacings and length of rock bolts were designed so as to achieve a target reliability index. The similarities and differences between the ratios of FORM design-point values to mean values, on the one hand, and the partial factors of limit state design, on the other hand, are discussed. Unlike design point based on partial factors, the design point in FORM is obtained as a by-product of target reliability index

(and associated P_f), and reflects input uncertainties, sensitivities, and correlations from case to case in ways that design point based on rigid partial factors cannot. However, more statistical input information is required in RBD than in EC7.

In its current version, EC7 covers little on the characteristic values and partial factors of rock engineering parameters like orientations of discontinuities, in situ stresses, and properties of joints and rock material. RBD is a more flexible approach in dealing with case-specific uncertainties of input values and can potentially complement EC7 (and LRFD).



α is the slope face inclination, Ω the upper ground inclination.
 Slope face dip direction coincides with upper ground dip direction.
 H and h are related by equation; either can be input

$$F_s = \left(a_1 - \frac{b_1 G_{w1}}{s_\gamma} \right) \times \tan \phi_1 + \left(a_2 - \frac{b_2 G_{w2}}{s_\gamma} \right) \times \tan \phi_2 + 3b_1 \frac{c_1}{\gamma h} + 3b_2 \frac{c_2}{\gamma h}$$

Sliding on both planes, one of the three modes.

in which dimensionless coefficients a_1, a_2, b_1 and b_2 are functions of $\beta_1, \delta_1, \beta_2, \delta_2, \alpha$ and Ω , and s_γ is the specific gravity of rock (γ/γ_w), c_1 and c_2 cohesions, $G_{w1} = 3u_1/\gamma_w h$, $G_{w2} = 3u_2/\gamma_w h$, in which u_1 and u_2 are average water pressures on joint planes 1 and 2.

α	Ω	h	s_γ
70	0	16	2.6

		BetaDist parameters								Correlation matrix							
		min	max	\underline{x}^*	$\underline{\mu}^N$	$\underline{\sigma}^N$	\underline{nx}										
BetaDist	β_1	4	4	53	71	61.21	62.00	3.263	-0.243	1	0	0	0	0	0	0	0
BetaDist	δ_1	4	4	44	56	50.44	50.00	2.179	0.204	0	1	0	0	0	0	0	0
BetaDist	β_2	4	4	11	29	19.26	20.00	3.266	-0.226	0	0	1	0	0	0	0	0
BetaDist	δ_2	4	4	42	54	48.35	48.00	2.183	0.158	0	0	0	1	0	0	0	0
BetaDist	G_w	4	4	0.14	0.86	0.61	0.51	0.122	0.841	0	0	0	0	1	0	0	0
BetaDist	$\tan \phi$	4	4	0.25	1.15	0.59	0.70	0.157	-0.649	0	0	0	0	0	1	0	0
BetaDist	$c/\gamma h$	4	4	0.04	0.16	0.07	0.10	0.017	-1.548	0	0	0	0	0	0	0	1

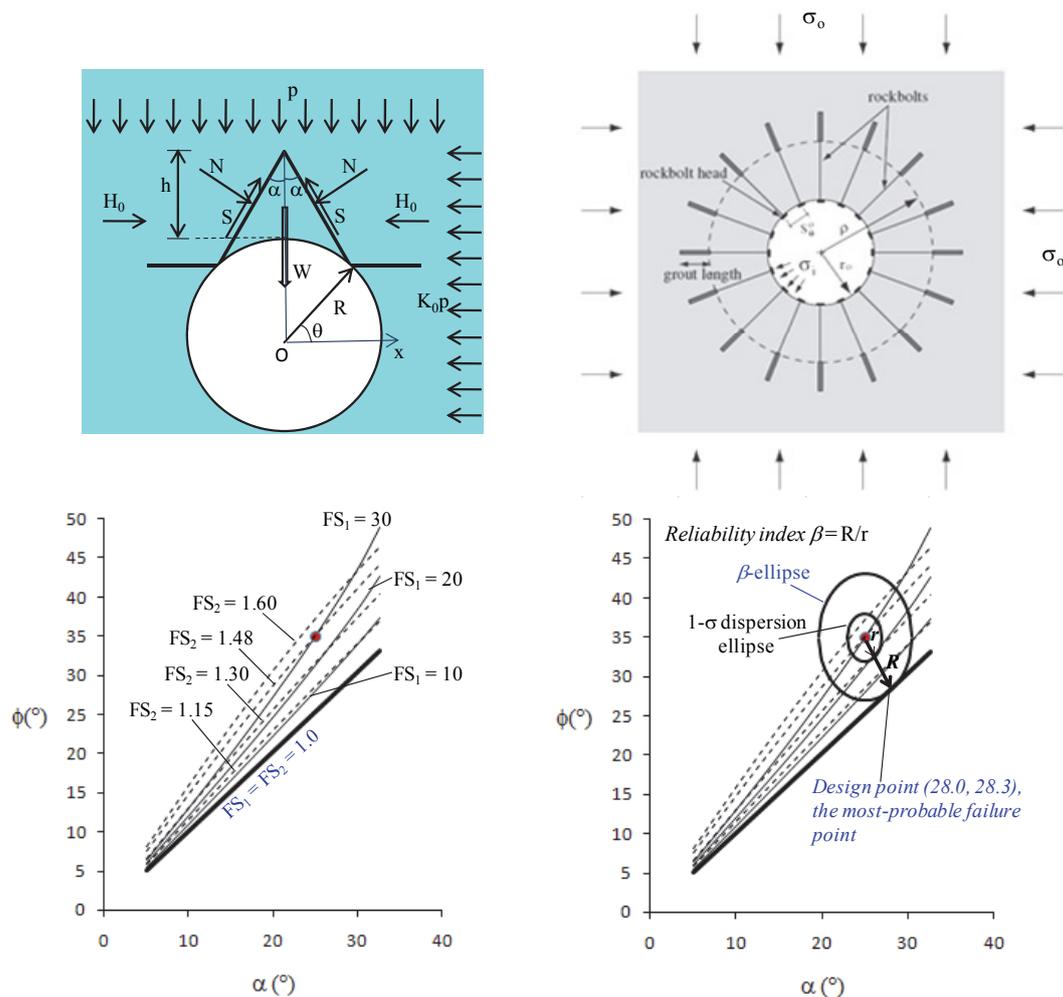
$g(x)$	β	P_f
0.00	1.924	0.0272

Figure 4.15 Reliability analysis of 3-D rock wedge with uncertain discontinuity orientations

4.7 POSITIVE RELIABILITY INDEX ONLY IF THE MEAN-VALUE POINT IS IN THE SAFE DOMAIN

In reliability analysis and reliability-based design one needs to distinguish negative from positive reliability index. The computed β index can be regarded as positive only if the performance function value is positive at the mean value point. This provides a simple check. For example, in the reliability-based design of the embedment depth of an anchored sheet pile wall in Fig. 4.8, the mean value point (prior to Excel Solver optimisation) yields a positive performance function value (cell $PerFn > 0$) for $H > 10.7$ m. The computed β index increases from about 0 (equivalent to a factor of

safety of 1.0) when H is 10.7 m to 3.0 when H is 12.3 m for the statistical inputs of Fig. 4.8. Another example for rock slope is given in Low (2014).



A tale of two factors of safety, and reconciliation via FORM

Figure 4.16 FORM analysis and RBD of tunnels in rocks.

4.8 SYSTEM FORM RELIABILITY ANALYSIS AND RBD INVOLVING MULTIPLE LIMIT STATES

Performance of engineering systems often involves multiple failure modes (various ULS and SLS). For instance, the geotechnical failure modes to be considered in the design of a semi-gravity retaining wall may include rotation about the toe of the wall, horizontal sliding along the base of the wall, and bearing capacity failure of the soil beneath the wall. Methods for estimating the bounds of system reliability are available. Low et al. (2011) presented a practical procedure for estimating system reliability based on the FORM reliability indices for individual modes and associated design points, illustrated by a semi-gravity retaining wall with two failure modes, and a soil slope with eight failure modes.

4.9 EXCEL-BASED SUBSET SIMULATION, APPLICATION EXAMPLE, AND ITS MERITS IN EC7

Subset simulation (Au and Beck 2001) is an advanced Monte Carlo Simulation (MCS) that aims to improve MCS’s computational efficiency, particularly at probability tails, while maintaining its

robustness. Subset Simulation stems from the idea that a small failure probability can be expressed as a product of larger conditional failure probabilities for some intermediate failure events, thereby converting a rare event (small probability levels) simulation problem into a sequence of more frequent ones. Subset Simulation is performed level by level (Fig. 4.17). The first level is direct MCS, and the subsequent levels utilize Markov chain Monte Carlo to generate conditional samples of interest. Details on Subset simulation are referred to Au and Beck (2001) and Au and Wang (2014). An Excel VBA Add-in called “Uncertainty Propagation using Subset Simulation” (UPSS) has been developed and can be obtain from <https://sites.google.com/site/upssvba> (Au et al. 2010, Au and Wang 2014).

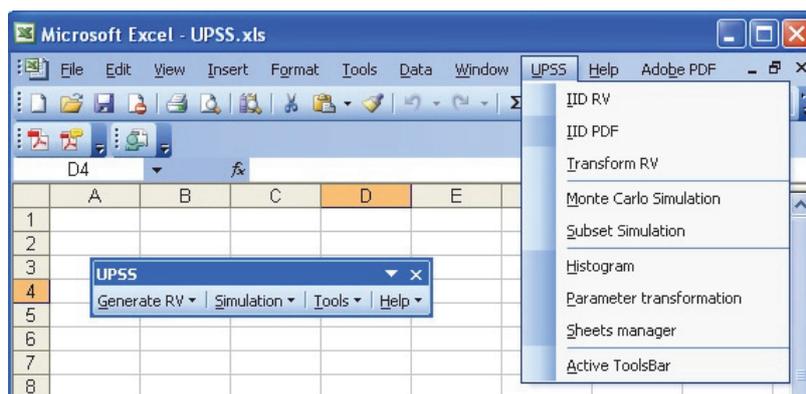
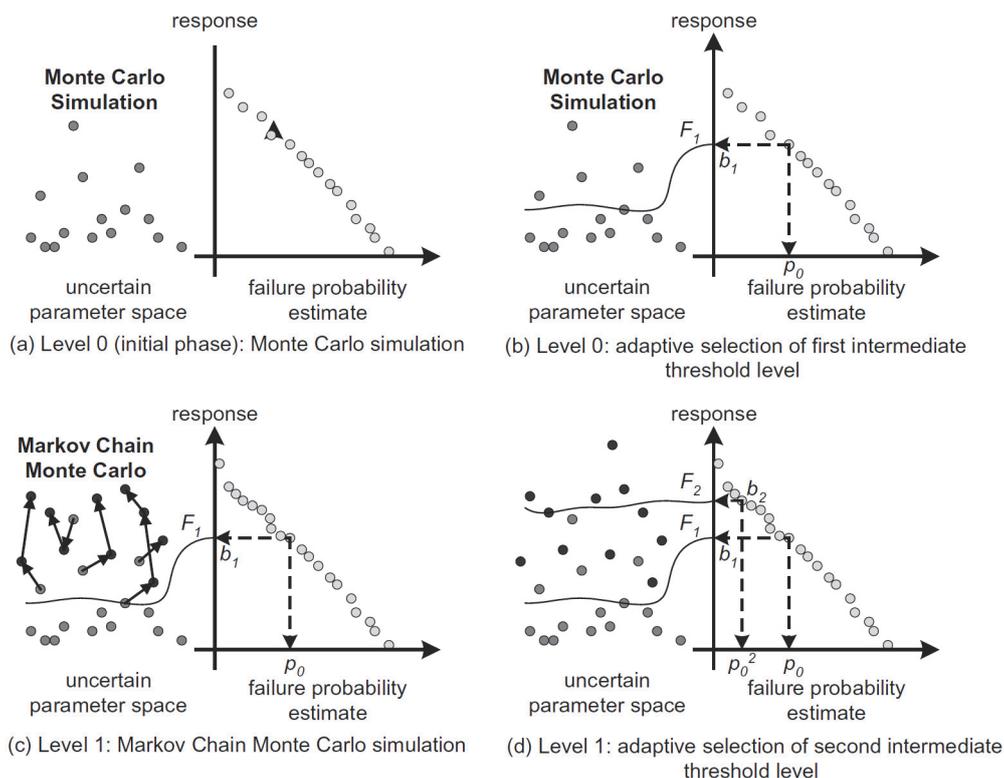


Figure 4.17 Excel-based subset simulation.

UPSS divides the reliability analysis or design into three separate processes: (1) deterministic modeling, (2) uncertainty modeling, and (3) uncertainty propagation by Subset simulation. The deterministic modeling is deliberately decoupled from uncertainty modeling and propagation. This allows three separate processes mentioned above to proceed in a parallel fashion. The uncertainty modeling and propagation are performed in a non-intrusive manner, and the robustness of MCS is well maintained. This removes the mathematical hurdles for engineering practitioners when performing reliability analyses or designs.

4.9.1 Subset Simulation Application Example

Excel-based Subset simulation has been used in reliability analysis of slope stability (Wang et al. 2010&2011, Wang and Cao 2015) and reliability-based design of foundation (Wang and Cao 2013&2015). A slope stability example is illustrated in this section (Figure 4.18). Details of the example are referred to Wang et al. (2011).

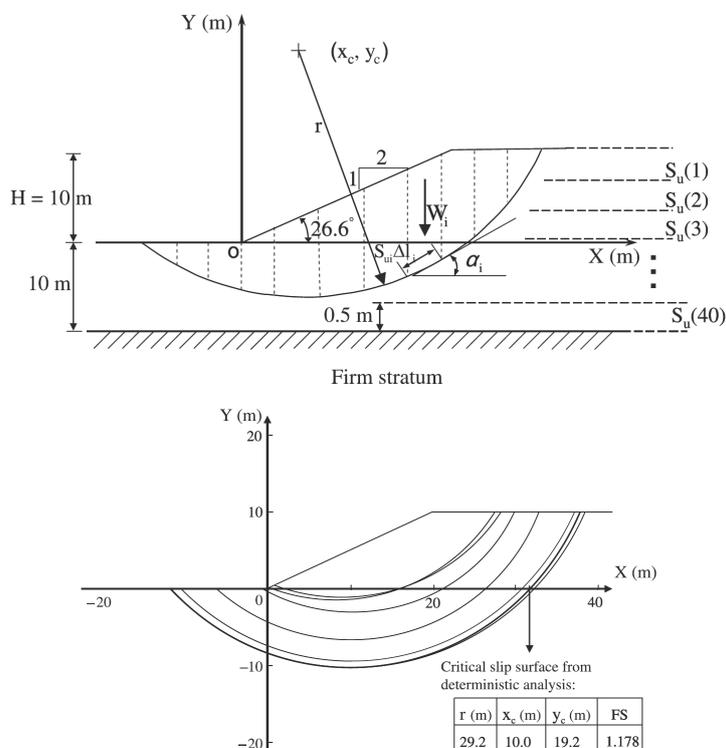


Figure 4.18 Subset simulation application example.

When the spatial variability of soil properties is considered in probabilistic slope stability analysis, the critical slip surface varies spatially. Using only one given critical slip surface significantly underestimates failure probability, and it is unconservative. Thus, it is necessary to properly model the spatial variability of the critical slip surface when the soil property spatial variability is considered, leading a dramatic increase in computational efforts. Subset simulation significantly improves computational efficiency and resolution, particularly at small probability levels.

4.9.2 Potential merits of MCS/Subset Simulation in EC7

EC7 adopts design formats similar to the traditional allowable stress design (ASD) methods. The factor of safety in ASD methods is replaced by a combination of partial factors in EC7, which are provided after some code calibration processes. Because design engineers are not involved in the calibration processes, many assumptions and simplifications adopted in the calibration processes are frequently unknown to the design engineers. This situation can lead to potential misuse of the partial factors that are only valid for the assumptions and simplifications adopted in the calibration processes. Design engineers may feel uncomfortable to accept these “black box” calibration processes blindly. In addition, design engineers have little flexibility in changing any of these assumptions/simplifications or making their own judgment because recalibrations are necessary when any assumption or simplification is changed.

MCS/Subset Simulation has potential merits in the aforementioned aspects. Because MCS/Subset Simulation can be treated as repeated computer (Excel) executions of the traditional ASD calculations, good geotechnical sense and sound engineering judgment that have been accumulated over many

years of ASD practice are well maintained during the development of the deterministic (ASD) model for MCS/Subset simulation. The MCS/Subset simulation – based design can be conceptualized as a systematic sensitivity study which is common in geotechnical practice and familiar to design engineers. MCS/Subset simulation is particularly beneficial in the following situations (Wang et al. 2016):

- (1) When the design scenario is out of the calibration domain (e.g. range of pile diameters, pile lengths, and statistics of geotechnical parameters) for semi-probabilistic RBD codes. In this case, it is inappropriate to use the load and resistance factors from semi-probabilistic RBD codes.
- (2) When the design model is different from the model selected during the calibration and development of semi-probabilistic RBD codes. In this situation, recalibration is needed and the resulting load and resistance factors are probably different.
- (3) When the uncertainty model [e.g. (i) decision on which variables are considered as uncertain; (ii) probabilistic modelling of the uncertain variables as random variables; and (iii) auto- and cross-correlation structures] which is an integral part of the calibration of semi-probabilistic RBD codes changes.
- (4) When the target failure probability needed in the design is different from the target failure probability pre-specified in EC7.
- (5) When the exact value of failure probability is needed in engineering applications, such as quantitative risk assessment and risk based decision making.
- (6) When the load and resistance are correlated. For example, the load and resistance for earth retaining structures and slopes are usually originated from the same sources (e.g., effective stress of soil) and correlated with each other. It is therefore difficult to decide whether the effective stress of soil or earth pressure should be regarded as a load or resistance.
- (7) When the reliability-based serviceability limit state design is required (most existing semi-probabilistic RBD codes only deal with ultimate limit state design of geotechnical structures).
- (8) When dealing with geometric uncertainties, such as orientation of joints in rock engineering. The geometric uncertainties cannot be easily considered by conventional partial factors.

4.10 PROBABILISTIC MODELS FOR GEOTECHNICAL DATA

This section provides a guide to the common question: “How to determine the statistical inputs for a design example?” The recommended practice is to combine all available data, both global data from comparable sites in the literature and local data from site investigation, using the Bayesian approach. A key input to the Bayesian approach is prior information (prior probability distribution). Simple probability models for describing single and multiple soil parameters suitable for Bayesian updating are discussed. It is useful to note that all the examples presented in Chapter 4 involve multiple soil parameters, although not all of them are strongly correlated. Extensive statistics have been compiled in the literature for soils (Phoon et al. 2016a, Ching et al. 2016) and model factors (Dithinde et al. 2016). In the absence of site-specific information, these generic statistics (Section 4.10.2) together with the models presented in Section 4.10.1 can be adopted as the prior probability distribution. However, it is common practice to complement data obtained from the literature with site investigation data. The prior distribution (from literature) can be updated systematically by site-specific data (from site investigation) using the Bayesian approach. This powerful Bayesian approach is only applicable within a probabilistic framework.

For example, in the absence of site-specific data but in the presence of data from comparable sites, the engineer may assess the effective stress friction angle ϕ' to fall between 28° and 51° based on the scatter of “x” markers in Fig. 4.19. This implies very loose to very dense sands, which is hardly informative for design. However, if site-specific SPT N-values are available and they fall in the vicinity of 25 blows, it is possible to reduce the uncertainty in ϕ' because the “⊗” markers fall within a more restrictive range of 36° and 46° . In many cases, this scatter is not uniformly distributed – it is more reasonable to restrict the range further using a 95% confidence interval from a normal distribution. Updating in the presence of new test data can be performed systematically and consistently within a powerful Bayesian framework. Ching and Phoon (2015) provide guidelines on how to fit geotechnical data (soil parameters and model factors) to practical probabilistic models in the Excel platform. Phoon (2006) provides reasons to consider the normal distribution as a default

distribution, particularly when the COV is “small”. One concrete way of checking suitability of normal distribution is to check the design point produced by FORM. If the design point is negative for a positive-valued parameter, then the normal distribution is not suitable. The lognormal distribution is often considered as the second default option.

Chapter 1 and BEST EXCEL Add-in described in Section 4.10.3 provide more details on Bayesian updating. Whether one derives a single cautious estimate or a probability distribution from a transformation model such as Fig. 4.19, the role of engineering judgment in selecting the appropriate transformation model and weeding out unreasonable estimates is obviously integral to this practice and needs no further emphasis.

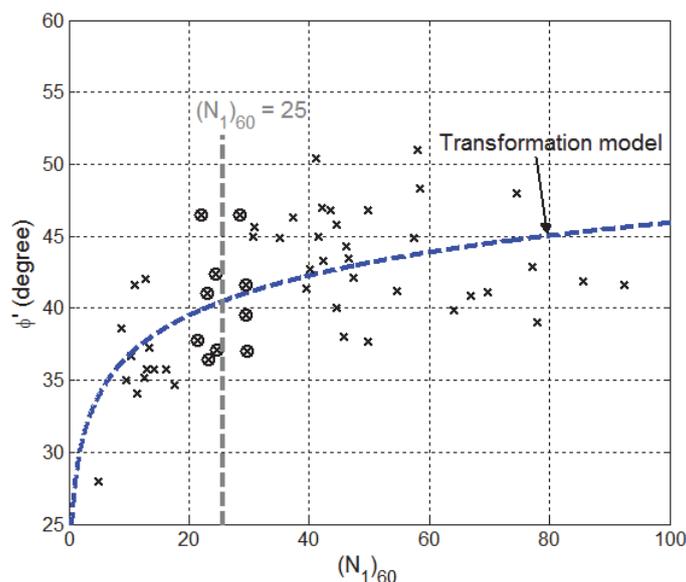


Figure 4.19 Relationship between effective stress friction angle and SPT blowcount.

4.10.1 Random vector – normal

For concreteness, assume that you have an EXCEL spreadsheet containing 3 columns. Column A contains data for the cone tip resistance, column B contains data for the sleeve friction, and column C contains data for pore pressure. We further assume that these 3 measurements were taken at 100 points in the depth direction. Therefore, the data is contained within the block of cells from A1 to C100. If the data are normally distributed, we can build a 3-dimensional normal random vector which consists of the following collection of random variables (Z_1, Z_2, Z_3). The random variable Z_1 is for cone tip resistance and so forth for Z_2 and Z_3 . It is easy to calculate the mean (μ_1) and standard deviation (σ_1) for Z_1 using the EXCEL “average” and “stdev.s” functions on each column of data. The means and standard deviations for Z_2 and Z_3 are obtained in the same way.

The key difference between a random variable and a random vector is a “correlation matrix” (Fig. 4.20), containing the correlation between cone tip resistance and sleeve friction (δ_{12}), the correlation between cone tip resistance and pore pressure (δ_{13}), and the correlation between sleeve friction and pore pressure (δ_{23}). You can get this correlation matrix directly from the data block A1:C100 using “Data Analysis > Correlation” under the Data tab in EXCEL. Once you obtained this correlation matrix, you can refer to the following sections in Ching & Phoon (2015) for simulation (Section 1.4.4) and Bayesian updating (Section 1.4.5).

Computational details involving geotechnical data which are multivariate and non-normal are given in Section 1.6 and 1.7 of Ching & Phoon (2015) and applications to actual soil databases are given in Ching et al. (2016).

	Z_1	Z_2	Z_3
Z_1	1	δ_{12}	δ_{13}
Z_2	δ_{12}	1	δ_{23}
Z_3	δ_{13}	δ_{23}	1

Figure 4.20 Correlation matrix for (Z_1, Z_2, Z_3)

4.10.2 Statistical guidelines

Extensive statistics have been compiled in the literature. These statistics are summarized by Phoon et al. (2016a) and Ching et al. (2016) for soils and Aladejare & Wang (2017) for intact rocks. The coefficient of variation (COV) is defined as the ratio of the standard deviation to the mean. Guidelines for COV for soil and rock parameters are given in Phoon et al. (2016a). It is important to note that COV of a soil or rock parameter can be small or large, depending on the site condition, the measurement method, and the transformation model. Resistance factors should be calibrated using the three-tier COV classification scheme shown in Table 4-1 to provide some room for the engineer to select the resistance factor that suits a particular site and other localized aspects of geotechnical practice (e.g. property estimation procedure) (Phoon et al. 2016b). A single resistance/partial factor ignores site-specific issues and it shares the same issues as the factor of safety approach where the nominal resistance has to be adjusted to handle site-specific considerations in the presence of a relatively constant factor of safety. For comparison, the COV for unit weight of soil, yield strength of steel, and compressive strength of concrete are less than 5%, 10%, and 20%, respectively.

Table 4-1. Three-tier classification scheme of soil property variability for reliability calibration
 (Source: Table 9.7, Phoon & Kulhawy 2008)

Geotechnical parameter	Property variability	COV (%)
Undrained shear strength	Low ^a	10 - 30
	Medium ^b	30 - 50
	High ^c	50 - 70
Effective stress friction angle	Low ^a	5 - 10
	Medium ^b	10 - 15
	High ^c	15 - 20
Horizontal stress coefficient	Low ^a	30 - 50
	Medium ^b	50 - 70
	High ^c	70 - 90

a - typical of good quality direct lab or field measurements

b - typical of indirect correlations with good field data, except for the standard penetration test (SPT)

c - typical of indirect correlations with SPT field data and with strictly empirical correlations

4.10.3 EXCEL Add-in for Bayesian Equivalent Sample Toolkit (BEST)

To deal with the issue of small sample size, Bayesian methods may be used to integrate limited measurement data in a specific site with prior knowledge (e.g., engineering experience and judgment, existing data from similar project sites) to provide updated knowledge on the soil parameter of interest (e.g., Wang et al 2016a). Because the updated knowledge might be complicated and difficult to express explicitly or analytically, Markov chain Monte Carlo (MCMC) simulation has been used to transform the updated knowledge into a large number of simulated samples of the soil parameter of interest, which collectively represent the soil parameter as a random variable (Wang and Cao 2013). An EXCEL add-in, called Bayesian Equivalent Sample Toolkit (BEST), has been developed for implementing the Bayesian method and MCMC simulation in a spreadsheet platform (Wang et al.

2016b). The BEST Add-in can be obtained without charge from <https://sites.google.com/site/yuwangcityu/best/1>. Engineering practitioners only need to provide input to the BEST Add-in, such as site-specific measurement data (e.g., several SPT N values) and typical ranges of soil parameters of interest (e.g., effective friction angle of soil) as prior knowledge. Then, the BEST Add-in may be executed to generate a large number of numerical samples of the soil parameters. Subsequently, conventional statistical analysis can be performed on these simulated samples using EXCEL’ built-in functions (e.g., “average” and “stdev.s”). The BEST Add-in can also be used for estimating soil parameters (e.g., undrained shear strength of clay) from “multivariate data” (e.g., SPT, CPT, and Liquidity index data) in a sequential manner using a Bayesian sequential updating method (Cao et al. 2016).

4.10.4 Model factors

The model factor for the capacity of a foundation is commonly defined as the ratio of the measured (or interpreted) capacity (Q_m) to the calculated capacity (Q_c), i.e. $M = Q_m/Q_c$. The value $M = 1$ implies that calculated capacity matches the measured capacity, which is unlikely for all design scenarios. Intuition would lead us to think that M takes different values depending on the design scenario. This intuitive observation is supported by a large number of model factor studies (Dithinde et al. 2016). Hence, it is reasonable to represent M as a random variable. It is straightforward to apply this simple definition to other responses beyond foundation capacity. For some simplified calculation models, M can depend on input parameters (i.e., M is not random) and additional efforts are required to remove this dependency (Zhang et al. 2015). A comprehensive summary of model factor statistics is presented by Dithinde et al. (2016). Multivariate model factors are not available at present.

4.11 CONCLUSIONS

The differences and similarities of the design point in RBD and EC7 were explained, and the insights and merits of RBD were illustrated for a strut with complex supports, a gravity retaining wall, an anchored sheet pile wall, a footing with inclined and eccentric loadings, a laterally loaded pile, soil slopes, 2D and 3D rock slopes, and tunnels in rocks. The ability of RBD to provide interesting information in its design point and to automatically reflect parametric uncertainties, correlations, loads with favourable-unfavourable duality, and case-specific sensitivities are demonstrated.

The limitations of imprecise and/or incomplete/non-exhaustive statistical inputs on FORM results are similar to the limitations of approximate inputs in deterministic analysis. Statistical inputs are approximate and often involve judgment, due to insufficient data. Further, one may have overlooked some factors (e.g. human factors). Besides input data, output of RBD (or any other design approach) also depends on the idealized mechanical model, the failure modes considered, etc. Hence the probability of failure based on RBD is not exact. The P_f from RBD is at best approximate (and sometimes way off), and hence the P_f must be regarded as nominal rather than precise. Nevertheless, the examples in this chapter demonstrate that FORM analysis and RBD via FORM may be very useful in the following ways:

- (i) Giving warnings when the computed P_f are unacceptably high;
- (ii) Sufficiently safe designs aiming at a target reliability index (and low nominal P_f);
- (iii) Comparative assessment of the relative reliability of different designs;
- (iv) Incorporating parametric correlations and spatial auto-correlations in design;
- (v) Complement EC7 design for parameters not yet covered in the design code;
- (vi) Complement EC7 design when the sensitivities of parameters vary from case to case;
- (vii) Complement EC7 design when reality warrants correlation among parameters;
- (viii) Complement EC7 design when a parameter possesses stabilizing-destabilising duality;
- (ix) Complement EC7 when uncertainty in unit weight of soil needs to be modelled.

That the P_f associated with a target reliability index in a RBD is more indicative and nominal than real should not deter the Geotech profession from appreciating the merits of RBD (like those listed above) and its potential complementary role to design approaches like EC7 and LRFD. The same limitations with respect to approximate inputs, idealizations and non-exhaustive factors also apply to the outputs of deterministic analysis to some extent (e.g. displacement prediction). One is reminded of Terzaghi’s pragmatic approach of aiming at designs such that unsatisfactory performance is not likely, instead of aiming at designs which would behave precisely (e.g. footing settlement of

exactly 25mm). It is in the same spirit that RBD aims to achieve sufficiently safe design, not at a precise probability of failure. For example, in a RBD for target reliability index of 3.0, the resulting design is not to be regarded as having exactly a probability of failure equal to $\Phi(-\beta) = 0.135\%$, but as a design aiming at a sufficiently small probability of failure (e.g. $<1\%$). One may note that a EC7 design (or LRFD design) via conservative characteristic (nominal) values and code-specified partial factors and different ULS and SLS also aims at sufficiently safe design by implicit considerations of parametric uncertainties and sensitivities. In comparison, the statistical data and correlations are open to view in RBD. Case-specific scrutiny and counter-suggestions for more reasonable statistical inputs and mechanical model in RBD are more likely to result in advancements and improvements.

4.12 REFERENCES

- Aladejare, A. E., and Wang, Y. (2017). Evaluation of rock property variability. *Georisk*, 11(1), 22-41.
- Ang, H. S., and Tang, W. H. (1984). Probability concepts in engineering planning and design, vol. 2-Decision, risk, and reliability. John Wiley, New York.
- Au, S. K., and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4), 263-77.
- Au, S. K., and Wang, Y. (2014). *Engineering Risk Assessment with Subset Simulation*, Wiley, ISBN 978-1118398043, 300p.
- Au, S. K., Cao, Z., and Wang, Y. (2010). Implementing advanced Monte Carlo Simulation under spreadsheet environment. *Structural Safety*, 32(5), 281-292.
- Baecher, G. B. and Christian, J. T. (2003). Reliability and statistics in geotechnical engineering. Chichester, West Sussex, England; Hoboken, NJ: J. Wiley.
- Bowles JE. Foundation analysis and design. 5th ed. McGraw-Hill, 1996.
- Cao, Z., Wang, Y., and Li, D. (2016). Site-specific characterization of soil properties using multiple measurements from different test procedures at different locations – A Bayesian sequential updating approach. *Engineering Geology*, 211, 150-161.
- Ching, J., and Phoon, K. K. (2015). Chapter 1 “Constructing Multivariate Distribution for Soil Parameters”. Risk and Reliability in Geotechnical Engineering, Eds. K.K. Phoon & J. Ching, CRC, Press, 3-76.
- Ching, J. Y., Li, D. Q., and Phoon, K. K. (2016), “Statistical characterization of multivariate geotechnical data”, Chapter 4, Reliability of Geotechnical Structures in ISO2394, Eds. K.K. Phoon & J.V. Retief, CRC Press/Balkema, 89-126.
- Coates, R. C., Coutie, M. G., and Kong, F. K., 1994. Structural analysis, 3rd Ed., Chapman and Hall, London.
- Dithinde, M., Phoon, K. K., Ching, J. Y., Zhang, L. M., and Retief, J. V. (2016). “Statistical characterization of model uncertainty”, Chapter 5, Reliability of Geotechnical Structures in ISO2394, Eds. K. K. Phoon & J. V. Retief, CRC Press/Balkema, 2016, 127-158.
- Hasofer A.M. and Lind N.C. (1974). Exact and invariant second-moment code format. *Journal of Engineering Mechanics*, 100, 111-21.
- Hoek, E. (2007). Practical rock engineering. http://www.rocscience.com/education/hoek_corner.
- Hoek, E., and Bray, J. (1977). Rock slope engineering. London: Inst Mining Metallurgy.
- Low, B. K. (1997). Reliability analysis of rock wedges. *J Geotechnical Geoenvironmental Engineering*, 123(6), 498–505.
- Low, B. K. (2005). Reliability-based design applied to retaining walls. *Géotechnique*, 55(1), 63-75.
- Low, B. K. (2007). Reliability analysis of rock slopes involving correlated nonnormals. *International Journal of Rock Mechanics and Mining Sciences*, 44(6), 922-935.
- Low, B. K. (2014). FORM, SORM, and spatial modeling in geotechnical engineering. *Structural Safety*, 49, 56–64.
- Low, B. K. (2015). Chapter 9: Reliability-based design: Practical procedures, geotechnical examples, and insights, (pages 355-393) of the book Risk and Reliability in Geotechnical Engineering, CRC Press, Taylor & Francis group, 624 pages, edited by Kok-Kwang Phoon, Jianye Ching.
- Low, B. K. and Duncan, J. M. (2013). Testing bias and parametric uncertainty in analyses of a slope failure in San Francisco Bay mud. *Proceedings of Geo-Congress 2013*, ASCE, March 3-6, San Diego, 937-951.
- Low, B. K., and Einstein, H. H. (2013). Reliability analysis of roof wedges and rockbolt forces in tunnels. *Tunnelling and Underground Space Technology*, 38, 1-10.

- Low, B. K., and Phoon, K. K. (2015). Reliability-based design and its complementary role to Eurocode 7 design approach. *Computers and Geotechnics*, 65, 30-44.
- Low, B. K., and Tang, W. H. (2004). Reliability analysis using object-oriented constrained optimization. *Structural Safety*, 26(1), 69-89.
- Low, B. K., and Tang, W. H. (2007). Efficient spreadsheet algorithm for first-order reliability method. *Journal of Engineering Mechanics*, 133(12), 1378-1387.
- Low, B. K., Lacasse, S., and Nadim, F. (2007). Slope reliability analysis accounting for spatial variation. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 1(4), 177-189.
- Low, B. K., Teh, C. I., and Tang, W. H. (2001). “Stochastic nonlinear p-y analysis of laterally loaded piles.” *Proceedings of the Eight International Conference on Structural Safety and Reliability, ICOSSAR '01, Newport Beach, California, 17-22 June 2001*, 8 pages, A.A.Balkema Publishers. (<http://www.ntu.edu.sg/home/cbklow/conference.htm>)
- Low, B. K., Zhang J. and Tang, W. H. (2011). Efficient system reliability analysis illustrated for a retaining wall and a soil slope. *Computers and Geotechnics*, 38(2), 196-204.
- Melchers, R. E. (1999). *Structural reliability analysis and prediction*. 2nd ed. New York: John Wiley.
- Peck, R. B. (1980). ‘Where has all the judgment gone?’ The fifth Laurits Bjerrum memorial lecture. *Canadian Geotechnical Journal*, 17(4), 584–590.
- Phoon K. K. (2006). “Modeling and Simulation of Stochastic Data”, *GeoCongress, ASCE, Atlanta*.
- Phoon, K. K., and Kulhawy, F. H. (2008). Serviceability limit state reliability-based design. Chapter 9, *Reliability-Based Design in Geotechnical Engineering: Computations and Applications*, Taylor & Francis, UK, 344-383.
- Phoon, K. K., Prakoso, W. A., Wang, Y., Ching, J. Y. (2016a). “Uncertainty representation of geotechnical design parameters”, Chapter 3, *Reliability of Geotechnical Structures in ISO2394*, Eds. K. K. Phoon & J. V. Retief, CRC Press/Balkema, 49-87.
- Phoon, K. K., Retief, J. V., Ching, J., Dithinde, M., Schweckendiek, T., Wang, Y., and Zhang, L. M. (2016b). Some Observations on ISO2394:2015 Annex D (Reliability of Geotechnical Structures), *Structural Safety*, 62, 24-33.
- Rackwitz, R., and Fiessler, B. (1978). “Structural reliability under combined random load sequences”, *Comput. Struct.* 9(5), 484–494.
- Tomlinson M. J. (1994). *Pile design and construction practice*. 4th ed. London: E & FN, Spon.
- Tomlinson M. J. (2001). *Foundation design and construction*. 7th ed. Longman Scientific.
- Wang, Y. and Cao, Z. (2013). Probabilistic characterization of Young's modulus of soil using equivalent samples. *Engineering Geology*, 159, 106-118.
- Wang, Y. and Cao, Z. (2013). Expanded reliability-based design of piles in spatially variable soil using efficient Monte Carlo simulations. *Soils and Foundations*, 53(6), 820–834.
- Wang, Y. and Cao, Z. (2015). Practical reliability analysis and design by Monte Carlo Simulations in spreadsheet. Chapter 7 in *Risk and Reliability in Geotechnical Engineering*, 301-335, edited by K.K. Phoon and J. Ching, CRC Press.
- Wang, Y., Akeju, O. V., and Cao, Z. (2016b). Bayesian Equivalent Sample Toolkit (BEST): an Excel VBA program for probabilistic characterization of geotechnical properties from limited observation data. *Georisk*, DOI: 10.1080/17499518.2016.1180399.
- Wang Y., Cao, Z. and Au, S. K. (2010). Efficient Monte Carlo Simulation of parameter sensitivity in probabilistic slope stability analysis. *Computers and Geotechnics*, 37(7-8), 1015-1022.
- Wang, Y., Cao, Z. and Au, S. K. (2011). Practical reliability analysis of slope stability by advanced Monte Carlo Simulations in spreadsheet. *Canadian Geotechnical Journal*, 48(1), 162-172.
- Wang, Y., Cao, Z., and Li, D. (2016a). Bayesian perspective on geotechnical variability and site characterization. *Engineering Geology*, 203, 117-125.
- Wang, Y., Schweckendiek T., Gong W., Zhao T., and Phoon K. K. (2016). “Direct probability-based design methods”. Chapter 7 in *Reliability of Geotechnical Structures in ISO2394*, 193-228, edited by K.K. Phoon and J. V. Retief, CRC Press.
- Zhang, D. M., Phoon, K. K., Huang, H. W., and Hu, Q. F. (2015). Characterization of model uncertainty for cantilever deflections in undrained clay. *Journal of Geotechnical and Geoenvironmental Engineering*, 141(1), 04014088.

Chapter 5 Selection of characteristic values for rock and soil properties using Bayesian statistics and prior knowledge

Lead discussor:

Yu Wang

yuwang@cityu.edu.hk

Discussors (alphabetical order):

Marcos Arroyo, Zijun Cao, Jianye Ching, Tim Länsivaara, Trevor Orr, Kok-Kwang Phoon, Hansruedi Schneider, Brian Simpson

5.1 INTRODUCTION

Selection of design values for material properties is indispensable in engineering analyses and designs. In many engineering disciplines (e.g., structural engineering), design values are derived from characteristic values of material properties (e.g., concrete strength) which are often determined using statistical methods. In geotechnical engineering, however, it has been a challenging task to use statistical methods for determination of soil or rock property characteristic values, because the number of soil/rock property data obtained during site investigation is generally too sparse to generate meaningful statistics, i.e., the so-called “curse of small sample size” (Phoon 2017). Engineering experience and judgment are often used to supplement the limited measurement data during the selection of characteristic values for soil or rock properties. A Bayesian framework may be used to integrate the limited measurement data with engineering experience and judgment (as prior knowledge) in a rational and quantitative manner. This has been recognized in Eurocode 7 (EC7 (e.g., CEN 2004)) and referenced in the clause 2.4.5.2(10): “If statistical methods are employed in the selection of characteristic values for ground properties, such methods should differentiate between local and regional sampling and should allow the use of a priori knowledge of comparable ground properties.” This preliminary report aims to provide a start-of-the-art review on the use of Bayesian statistics and prior knowledge for selection of ground property characteristic values. The report focuses on routine engineering practices on conventional types of geotechnical structures with no exceptional risk or difficult ground or loading conditions (e.g., the Geotechnical Category 2 in EC7). Some design examples of such category are given by Orr (2005). Random field modeling of variability and uncertainty is not covered in this report.

5.2 DEFINITION OF CHARACTERISTIC VALUE

The definition of characteristic value for ground properties itself might be an intriguing issue, although detailed discussion on this issue is beyond the scope of this report and is referred to Schneider and Schneider (2013) and Orr (2017) for the discussion associated with EC7. There are different definitions of the characteristic value for ground properties in various geotechnical design codes around the world. For example, a mean value is generally used as nominal value (i.e., characteristic value) in several reliability-based design codes in North America (e.g., Phoon et al. 2003a&b, Paikowsky et al. 2004&2010, Fenton et al. 2016). EC7 recommends that (see clause 2.4.5.2(2)) “characteristic value of a geotechnical parameter shall be selected as a cautious estimate of the value affecting the occurrence of the limit state.” It further notes that (see clause 2.4.5.2(11)) “If statistical methods are used, the characteristic value should be derived such that the calculated probability of a worse value governing the occurrence of the limit state under consideration is not greater than 5%.” A note to this clause clarifies that “In this respect, a cautious estimate of the mean value is a selection of the mean value of the limited set of geotechnical parameter values, with a confidence level of 95%; where local failure is concerned, a cautious estimate of the low value is a

5% fractile.” Although the above general statement describing the characteristic value in EC7 is sensible, there is an under-stated difficulty in making this statement sufficiently concrete for codification. Orr (2017) suggested that more guidance on the selection of characteristic values is needed for reduce the spread of the selected characteristic values and achieve designs with more consistent reliability.

Using 5% fractile, mean value or other percentage fractile as the characteristic value has their respective pros and cons. Although using the 5% fractile has the advantages of reflecting both the mean value and uncertainty (e.g., 5% fractile is equal to the mean value minus 1.65 standard deviations for a normal distribution) and being in harmony with the definition of the characteristics value for other engineering materials (e.g., concrete in structural engineering), the 5% fractile is difficult to quantify due to the limited ground property data obtained during a site investigation and often the large extent of the failure zone compared to the number of test results. In contrast, although the mean value may be estimated from limited ground property data with better accuracy than the 5% fractile, it does not provide any indication of the variability and hence the uncertainty in the ground properties. Therefore, the impact of different levels of ground property uncertainty on geotechnical design should be considered by other means within the design codes. For example, a three-tier system of different resistance factors for different levels of ground property uncertainty is developed in some design codes (e.g., Phoon et al. 2003a&b, Fenton et al. 2016, Phoon et al. 2016). Indeed, the definition of characteristic values for ground properties and the calibration of load and resistance (or partial) factors are intrinsically linked with each other. They should be compatible with each other and act together to properly account for the impact of different levels of ground property uncertainty on geotechnical design.

No matter how the characteristic value is defined for ground properties, quantification of ground property uncertainty is essential. Bayesian methods described in the next section not only effectively tackle the difficulty in dealing with limited site-specific ground property data, but are also consistent with existing geotechnical practice (i.e., using engineering experience and judgment together with limited site-specific measurement data).

5.3 BAYESIAN METHODS

Under a Bayesian framework, site information available prior to a project (e.g., existing data in literature, engineering experience, and engineers’ expertise) may be used as “prior” knowledge and integrated with limited project/site-specific measurement data in a rational and quantitative manner (e.g., Wang et al. 2016a). Starting from the Bayes Theorem, Eq. (5-1) can be derived to update statistical parameters Θ_p (e.g., mean μ and standard deviation σ) of a design ground property X_D (e.g., soil effective friction angle ϕ'), which is treated as a random variable, given a set of site-specific test data as (e.g., Ang and Tang 2007):

$$P(\Theta_p | \text{Data, Prior}) = K \times P(\text{Data} | \Theta_p) \times P(\Theta_p) \quad (5-1)$$

where K is a normalizing constant independent of the statistical parameters Θ_p of X_D ; $\text{Data} = \underline{X}_M$ is the site-specific measurement data (e.g., a set of standard penetration test SPT-N values); $P(\Theta_p)$ is the prior distribution of the statistical parameters in the absence of site-specific measurement data; and $P(\text{Data} | \Theta_p) = P(\underline{X}_M | \Theta_p)$ is the likelihood function. Two critical elements in the Bayesian framework are the formulations of prior distribution (see Section 5.4) and the likelihood function described in this section.

The likelihood function $P(\underline{X}_M | \Theta_p)$ is a probability density function, PDF, of site-specific measurement data \underline{X}_M for a given set of statistical parameters Θ_p . It quantifies probabilistically the Θ_p information provided by \underline{X}_M . Formulation of the likelihood function [i.e., $P(\underline{X}_M | \Theta_p)$] requires a likelihood model that probabilistically describes the relationship between the statistical parameters Θ_p of a design property X_D and project-specific test data \underline{X}_M . Generally speaking, the likelihood model shall reflect sound physical insights into the relationship between the design property X_D and the measurement data \underline{X}_M and the propagation of various uncertainties that occurred during site characterization (e.g., Wang et al. 2016a). As much as possible insights from soil or rock mechanics should be incorporated in the likelihood model. For example, insights from soil mechanics suggest

that undrained shear strength, S_u , of clay is not a fundamental soil property, but depends on the vertical effective stress, σ_v' . It is therefore a better likelihood model to consider S_u/σ_v' than S_u as a random variable (Cao and Wang 2014). In addition, the design property X_D might not be measured directly, and a transformation or regression model is needed to relate the measurement data X_M to X_D . The uncertainty (i.e., transformation uncertainty) associated with the transformation model should also be incorporated in the likelihood model. Based on the likelihood model, it can be derived that the measurement data X_M (e.g., SPT-N value) is a random variable that has a (e.g., normal or lognormal) PDF (Wang et al. 2016a). Statistical parameters for the random variable X_M are a function of the statistical parameters Θ_P for the random variable X_D and the transformation uncertainty. This establishes a link between the site-specific measurement data X_M and the statistical parameters Θ_P for the design property X_D and allows the likelihood function to be formulated mathematically. Therefore, the statistical parameters Θ_P for the design property X_D (e.g., mean μ and standard deviation σ for the soil effective friction angle ϕ') can be updated from X_M (e.g., SPT-N values), as shown in Eq. (5-1).

Using the theorem of total probability (e.g., Ang and Tang 2007), the posterior PDF of the design property X_D can be further expressed as (Wang and Cao 2013, Wang et al. 2016a):

$$P(X_D | \text{Data, Prior}) = \int P(X_D | \Theta_P) P(\Theta_P | \text{Data, Prior}) d\Theta_P \quad (5-2)$$

where $P(X_D | \Theta_P)$ is the conditional (e.g., normal or lognormal) PDF of X_D for a given set of statistical parameters Θ_P (e.g., μ and σ); and $P(\Theta_P | \text{Data, Prior})$ is obtained from Eq. (5-1).

When the prior knowledge and likelihood function in geotechnical practice are sophisticated, the X_D PDF might be complicated or difficult to express analytically or explicitly. To remove this mathematical hurdle in engineering practice, Markov chain Monte Carlo simulation (MCMCS, e.g., Robert and Casella 2004) is used to depict the X_D PDF numerically. The generated MCMCS samples collectively reflect the posterior PDF of X_D [i.e., $P(X_D | \text{Data, Prior})$ in Eq. (5-2)], and they are referred to as Bayesian equivalent samples of the design property X_D (Wang and Cao 2013).

It is worthwhile noting that Eq. (5-2) can also be interpreted as using the concept of the mixture model (e.g., McLachlan and Peel 2000, Wang et al. 2015), which considers $P(X_D | \text{Data, Prior})$ as a weighted summation of the various component density functions with different distribution parameters. Under the concept of the mixture model, $P(X_D | \Theta_P)$ in Eq. (5-2) is the component density function and $P(\Theta_P | \text{Data, Prior}) d\Theta_P$ is the weighting function. Because $P(X_D | \text{Data, Prior})$ is a weighted summation of various component density functions (e.g., normal or lognormal PDF) with different combinations of statistical parameters (e.g., means and standard deviations), it does not necessarily follow the same distribution type as the component density function, such as a normal or lognormal PDF (e.g., McLachlan and Peel, 2000; Wang et al., 2015). In other words, although a normal or lognormal PDF is often used for $P(X_D | \Theta_P)$, $P(X_D | \text{Data, Prior})$ in Eq. (5-2) may turn out to be another distribution.

5.4 SOURCES AND QUANTIFICATION OF PRIOR KNOWLEDGE

Geotechnical characterization of a project site often starts with a desk-study and site reconnaissance to collect prior knowledge (e.g., geological information, geotechnical problems and properties, groundwater conditions) of the project site from various sources (e.g., Trautmann and Kulhawy 1983, Clayton et al. 1995, Mayne et al. 2002, Cao et al. 2016). Geology information (e.g., bedrock geology, surficial geology, landform history) is available from existing geological records (e.g., geological maps, reports, and publications), regional guides, air photographs, soil survey maps and records, textbooks, etc. The information about geotechnical problems and parameters (e.g., soil classification and properties, and stratigraphy) can be collected from existing geotechnical reports (e.g., Kulhawy and Mayne 1990), peer-reviewed academic journals (e.g., geotechnical journals, engineering geology journals, and civil engineering journals), and previous ground investigation reports on similar sites. Information about groundwater conditions of the site (e.g., groundwater level) can be obtained from well records, previous ground investigation reports, topographical maps, and air photographs. In addition to these sources of existing information, the engineer's expertise (i.e., domain knowledge of engineers obtained from education, professional training, and experience from deliberate practice (Vick 2002, Cao et al. 2016)) provides useful information for geotechnical site characterization.

Based on its quality and quantity, prior knowledge can be divided into two categories: non-informative and informative prior knowledge. When only limited information is obtained during a desk-study and site reconnaissance, the prior knowledge is relatively non-informative, such as typical ranges and statistics of soil and rock properties (e.g., Aladejare and Wang 2017) summarized in the literature or from previous engineering experience. For example, Table 5-1 summarizes typical μ and σ ranges for soil properties (Cao et al. 2016). A uniform prior distribution can be used to represent the non-informative prior distribution quantitatively. In general, a uniform prior distribution indicates that there is no preference to any value within the typical range of ground property statistics (e.g., μ and σ) according to prior knowledge (e.g., Baecher and Christian 2003, Cao et al. 2016).

Table 5-1 Typical ranges of mean and standard deviation of soil properties (Cao et al. 2016)

Soil property	Soil type	Reference	Range of mean	Range of COV (%)	Range of the prior estimate of mean	Range of the prior estimate of standard deviation
Total unit weight, γ (kN/m ³)	Fine grained	12, 13	13–24	3–20	13–24	0.4–4.8
	Coarse grained	13	17–26	4–6	17–26	0.7–1.6
Dry unit weight, γ_d (kN/m ³)	Fine grained	1, 12, 15	9–18	2–21	9–18	0.2–3.8
Submerged unit weight, γ_s (kN/m ³)	All soils	7	5–11	0–10	5–11	0–1.1
Relative density, D_r (%)	Coarse grained	12	30–70	11–36	30–70	3.3–25.2
Natural water content, w_n (%)	Fine grained	1, 12, 15	13–105	7–46	13–105	1–48
Liquid limit, w_L (%)	Fine grained	1, 7, 12, 15	27–89	3–39	27–89	1–35
Plastic limit, w_p (%)	Fine grained	1, 7, 12	13–35	3–34	13–35	0.4–12
Plasticity index, PI (%)	Fine grained	1, 5, 10, 11, 12, 15	11–52	9–57	11–52	1–30
Liquidity index, LI	Fine grained	6, 12	0.5–2.5	5–88	0.5–2.5	0.025–2.2
Undrained shear strength, S_u (kPa)	Fine grained	1, 5, 6, 7, 9, 10, 11, 12, 14,	6–713	4–84	6–713	0.2–600
Undrained shear strength ratio, $r = S_u/\sigma'_{v0}$	Fine grained	5, 6, 7, 9, 10, 11	0.23–1.4	5–90	0.23–1.4	0.01–1.26
Effective stress friction angle, ϕ' (°)	Fine grained	12	9–41	4–50	9–41	0.4–20.5
	Coarse grained	2, 7, 12	30–42	2–17	30–42	0.6–7.1
Tangent of effective stress friction angle, $\tan\phi'$	Fine grained	12, 13	0.24–0.69	6–46	0.24–0.69	0.01–0.32
	Coarse grained	8, 12, 13	0.57–0.92	5–18	0.57–0.92	0.03–0.17
Young's modulus, E (MPa)	Coarse grained	12	5.2–15.6	26–68	5.2–15.6	1.35–10.6
Compression index, C_c	Fine grained	1, 3, 15, 16	0.18–0.996	14–47	0.18–0.996	0.03–0.47
Permeability, k (cm/s)	Fine grained	4, 8	2.9×10^{-9} – 1.0×10^{-5}	27–767	2.9×10^{-9} – 1.0×10^{-5}	7.8×10^{-10} – 7.7×10^{-6}

Non-informative prior knowledge can be treated as the baseline uncertainty for ground properties in the absence of sufficient site-specific data or informative prior knowledge. It can also be used as a starting point for developing informative prior knowledge, and they can be used together with other sources of local prior knowledge, including, but not limited to, local engineering experience, information from previous projects in similar geological settings, and various soil and rock properties reported elsewhere locally. As the quality and quantity of prior knowledge improve, the prior knowledge becomes more and more informative and sophisticated. For informative and sophisticated prior knowledge from various sources, a subjective probability assessment framework (SPAF) may be used to facilitate synthesis and a quantitative representation of the informative prior knowledge by a proper prior PDF and to assist geotechnical engineers in formulating and expressing their engineering judgments in a quantifiable and transparent manner. Details of the SPAF steps and suggestions on each SPAF step to assist engineers in reducing the effects of cognitive biases and limitations during subjective probability assessment are referred to Cao et al. (2016).

5.5 SOFTWARE

Although the Bayesian framework described above is general and applicable for various soil or rock properties, its formulations vary for different properties when they are estimated from various in-situ and laboratory tests. For example, the formulation of the Bayesian equivalent sample method for probabilistic characterization of uniaxial compressive strength (UCS) of a rock using point load test [$I_{s(50)}$] data (e.g., Wang and Aladejare 2015) is different from the formulation for characterizing effective friction angle of soil using SPT-N values (e.g., Wang et al. 2015). Therefore, extensive

backgrounds in probability, statistics, and simulation algorithms are needed to formulate the method for various properties. To remove this mathematical hurdle for geotechnical practitioners, a user-friendly Microsoft Excel-based toolkit, called Bayesian Equivalent Sample Toolkit (BEST), has been developed for implementing the Bayesian method and providing a convenient way of estimating reasonable statistics of different ground properties from prior knowledge and limited site-specific measurement data (Wang et al. 2016b).

The BEST is developed using the Visual Basic for Applications (VBA) in a commonly available spreadsheet platform (i.e., Microsoft Excel), and it is compiled as an Excel Add-in for easy distribution and installation. The Excel-based BEST Add-in can be obtained without charge from <https://sites.google.com/site/yuwangcityu/best/1>. Step by step procedure for installing Add-in in Excel is provided in the Microsoft Office webpage below: <https://support.office.com/en-us/article/Add-or-remove-add-ins-0af570c4-5cf3-4fa9-9b88-403625a0b460>. After installation of BEST, four menus (i.e., Clay Property, Sand Property, User-defined Model and Help) appear in the “Custom Toolbars” of “ADD-INS” in Microsoft Excel. Figure 5-1 shows the BEST menus in Excel 2013. The BEST Excel Add-in can be used in the same way as the Excel built-in functions. Details of the Excel-based BEST Add-in are given by Wang et al. (2016b). Moreover, an Android mobile phone APP version of the BEST has been recently developed, as shown Figure 5-2. A β -version of the BEST APP can be obtained by scanning the QR code shown in Figure 5-2(c) or visiting the website: http://server.m.pp.cn/download/apk?appId=7626085&custom=0&ch_src=pp_dev&ch=default.

Selecting either the “Clay Property” or “Sand Property” menu in Figure 5-1 prompts the “Built-in Model” window. Twelve clay or sand property models reported in literature have been implemented as “built-in Models” in the current version of BEST, such as estimating effective friction angle of sand or undrained Young’s modulus of clay from SPT-N values (e.g., Kulhawy and Mayne 1990, Ching et al. 2012). The input data required for the “Built-in Model” include site-specific measurement data and prior knowledge. An example of using the “Built-in Model” will be shown in Subsection 5.6.1.

Selection of the “User-defined Model” menu prompts the “User-defined Model” window, which allows users to specify their own transformation/regression model and model uncertainty. In addition to site-specific measurement data and prior knowledge, users are asked to input model parameters that define the transformation model and model uncertainty. An example of using the “User-defined Model” will be shown in Subsection 5.6.2.

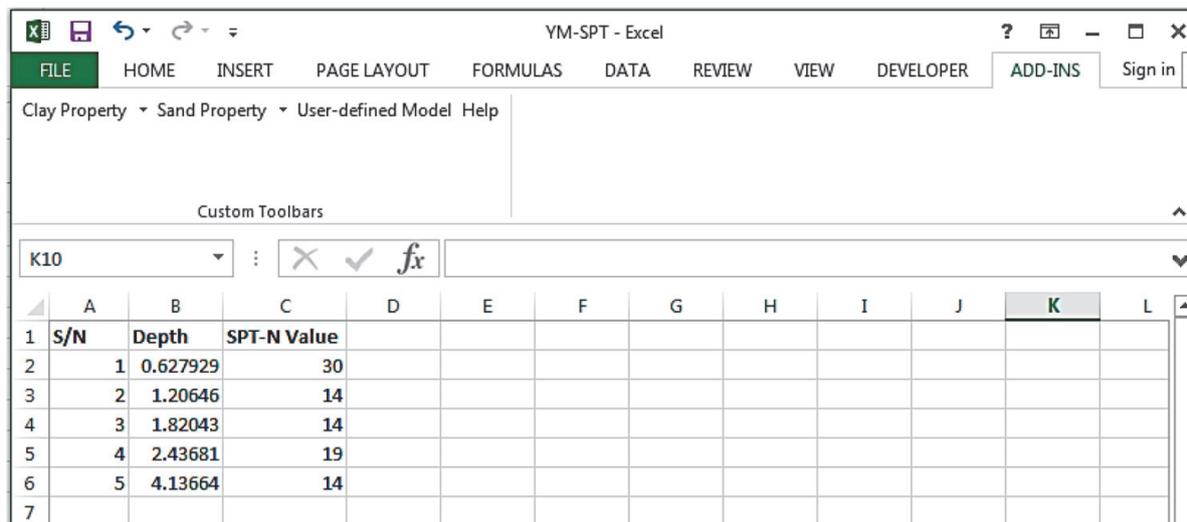


Figure 5-1 BEST Add-in menus after installation

After the required input data have been specified, both windows lead to the “Equivalent Sample Generation” window for generating a large number of equivalent samples of the design property X_D as output. The generated equivalent samples will be recorded in a newly created Excel worksheet, together with their statistics, such as mean, standard deviation, 5% and 95% fractile values. Characteristic values of soil or rock properties of interest may be determined from these statistics.



Geotechnical Parameter Uncertainty Quantification

Bayesian Equivalent Sample Toolkit
 Version 1.0

(a) Welcome Page



- Introduction
- Inputs and outputs
- Symbols and unit
- Prior information
- Transformation model
- Markov Chain Monte Carlo Simulation
- References

(b) Help Page



(c) QR code

Figure 5-2 BEST Android APP (β -version)

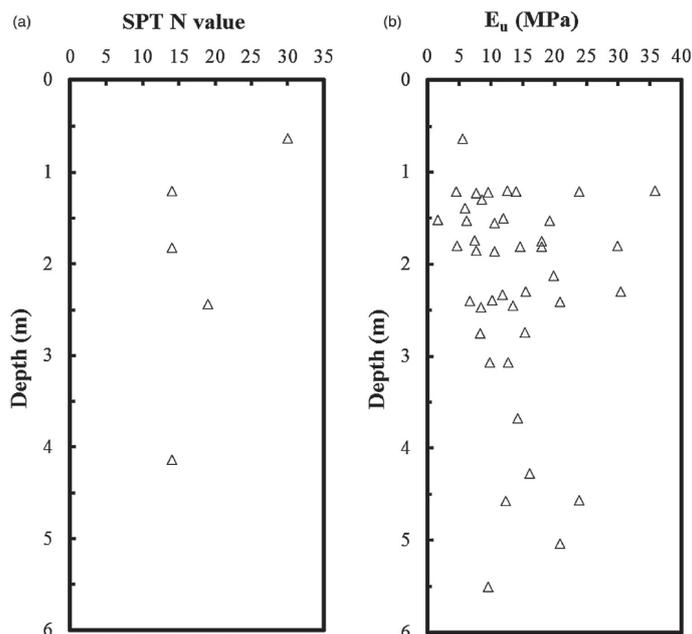


Figure 5-3 SPT-N values and undrained Young’s modulus, E_u , measured by pressuremeter tests at the clay site of the NGES at Texas A&M University (after Briaud 2000)

5.6 APPLICATION EXAMPLES

Two examples of soil and rock properties, respectively, are presented in this section to illustrate the Bayesian method, the BEST Excel Add-in, and how to obtain reasonable statistics for the selection of ground property characteristic values from limited site-specific measurement data and prior knowledge.

5.6.1 Soil property characteristic value

Consider, for example, characterization of the undrained Young’s modulus E_u of clay using SPT-N value data obtained from the clay site of the United States National Geotechnical Experimentation Sites (NGES) at Texas A&M University (Briaud 2000). A limited number of SPT-N values (i.e., 5 SPT-N values) were obtained within top stiff clay layer of the clay site, as illustrated in Figure 5-3(a). Figure 5-3(b) also shows the results of 42 pressuremeter tests performed in the same top clay layer at different depths (Briaud 2000) which are used for validating the BEST results.

The required design property in this example is the E_u of clay, and its corresponding measured data are the SPT-N values. Since BEST has a “Built-in Model” that relates the SPT-N values to the E_u of clay (Kulhawy and Mayne 1990, Phoon and Kulhawy 1999a), this “Built-in Model” under the “Clay Property” menu is used in this example. This example is performed in an Excel worksheet as shown in Figure 5-1, which contains 5 SPT-N values in Column “C” that correspond to those in Figure 5-3(a). A set of non-informative prior knowledge is used in this example, and it is taken as a joint uniform distribution with a mean of E_u varying between 5 MPa and 15 MPa and a standard deviation of E_u ranging from 0.5 MPa to 13.5 MPa. This set of prior knowledge is consistent with the typical ranges of E_u of clay reported in the literature (e.g., Kulhawy and Mayne 1990, Phoon and Kulhawy 1999a&b). Using this set of prior knowledge and the 5 SPT-N values shown in Figure 5-1, BEST is executed to generate 30,000 equivalent samples of E_u . It takes less than 2 minutes for BEST to generate 30,000 equivalent samples of E_u using a personal computer with an Intel® Core i7-4790 3.60GHz CPU and 8.0 GB RAM in the 64-bit Windows 8 operating system. Conventional statistical analysis, such as calculation of mean and standard deviation and plotting histogram for PDF or cumulative distribution function (CDF), can be easily performed on the 30,000 equivalent samples using built-in functions in Excel.

Table 5-2 shows statistics of the E_u samples obtained from BEST and their comparison with those obtained directly from the pressuremeter tests. The E_u PDF estimated from the BEST equivalent samples is shown in Figure 5-4 by a solid line with triangle markers. For validation, the E_u PDF generated by Matlab (Wang and Cao 2013) is also included in Figure 5-4 by a dashed line with circle markers. The solid line with triangle markers and dashed line with circle markers are both plotted on the primary vertical axis which represents the PDF of E_u . The dashed line virtually overlaps with the solid line, indicating that the equivalent samples from BEST are in good agreement those from Matlab. In addition, the results from the pressuremeter tests are included in Figure 5-4 and they are plotted on the secondary vertical axis which represents the frequency of the pressuremeter test results. About 36 out of the 42 pressuremeter tests results fall within the 90% inter-percentile range (3.95 MPa, 20.89 MPa) of the equivalent E_u samples from BEST. Figure 5-5 displays the CDFs of E_u estimated from the cumulative frequency diagrams of the BEST equivalent samples and the 42 pressuremeter test results by a solid line with triangle markers and open squares, respectively. The open squares plot close to the solid line, indicating that the E_u CDF obtained from BEST compares favorably with that obtained from the 42 pressuremeter tests. This agreement suggests that the information contained in the BEST equivalent samples is consistent with that obtained from the pressuremeter tests.

The E_u characteristic value may be selected from these statistics. For example, if the characteristic value is taken as the mean or 5% fractile value, it is about 11.5 MPa or 3.9 MPa, respectively, at this specific site.

Table 5-2 Summary of the E_u statistics

Statistics (MPa)	BEST Excel Add-in	MATLAB (Wang and Cao 2013)	Pressuremeter Tests	Difference between BEST and Pressuremeter Tests
Mean	11.46	11.60	13.50	2.04
Standard deviation	6.00	6.00	7.50	1.50

5.6.2 Rock property characteristic value

Consider, for example, characterization of the uniaxial compressive strength (UCS) of a granite deposit from point load test [$Is_{(50)}$] data. Table 5-3 summarizes laboratory test results of granite at the Malanjkhanda Copper Project in the State of Madhya, Pradesh, India (Mishra and Basu 2012).

The “User-defined Model” menu in BEST is used in this example. The design property of interest is the UCS, and the measurement data (i.e., the input data for BEST) are the Point load, $I_{S(50)}$, data. Note that the UCS data in the third Column of Table 5-3 are NOT the input to the BEST add-in, but are only used for comparing and validating the results obtained from the BEST add-in. A set of non-informative prior knowledge of the UCS statistical parameters is used in this study, and it is taken as a joint uniform distribution with a mean UCS varying between 121 MPa to 337 MPa and a standard deviation UCS ranging from 0 MPa to 36 MPa (Wang and Aladejare 2015).

Table 5-3 Laboratory test results of granite collected from Malanjkhanda Copper Project in the State of Madhya, Pradesh, India (after Mishra and Basu 2012)

Specimen number	$I_{S(50)}$ data (MPa)	UCS (MPa)
G1	8.35	139.04
G2	10.85	177.37
G3	10.02	167.17
G4	9.92	176.75
G5	11.73	160.82
G6	14.13	198.15
G7	10.63	148.34
G8	6.93	117.95
G9	8.49	134.76
G10	7.87	124.89
G11	8.41	138.22
G12	7.85	130.06
G13	5.99	122.74
G14	Invalid	201.73
G15	7.29	153.55
G16	11.36	182.33
G17	9.23	150.42
G18	6.92	127.47
G19	9.72	158.69
G20	5.66	91.48

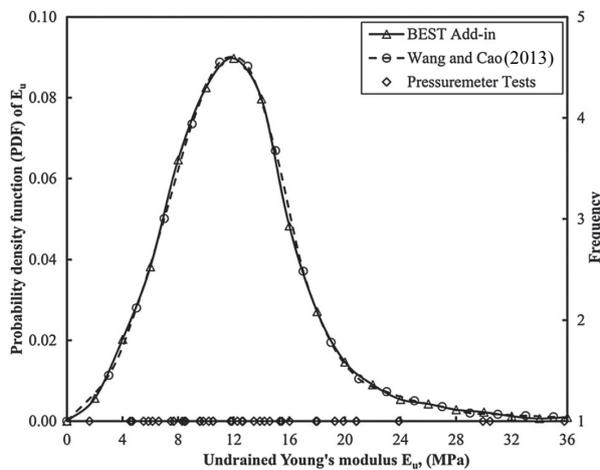


Figure 5-4 E_u PDF and frequency plots

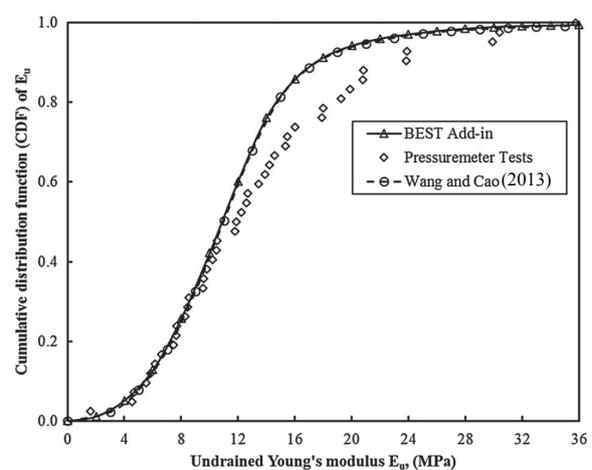


Figure 5-5 E_u CDF plot

Table 5-4 Summary of the UCS statistics

Statistics (MPa)	BEST Add-in	MATLAB (Wang and Aladejare 2015)	Compression Test	Difference between BEST and Compression Test
Mean	148.65	147.80	150.10	2.07
Standard deviation	17.13	18.70	28.30	9.25

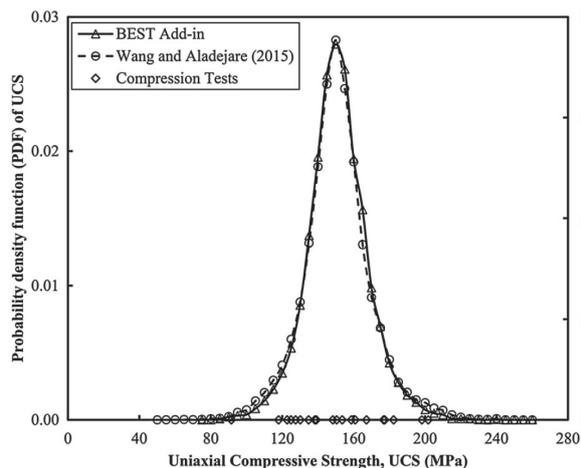


Figure 5-6 UCS PDF and frequency plots

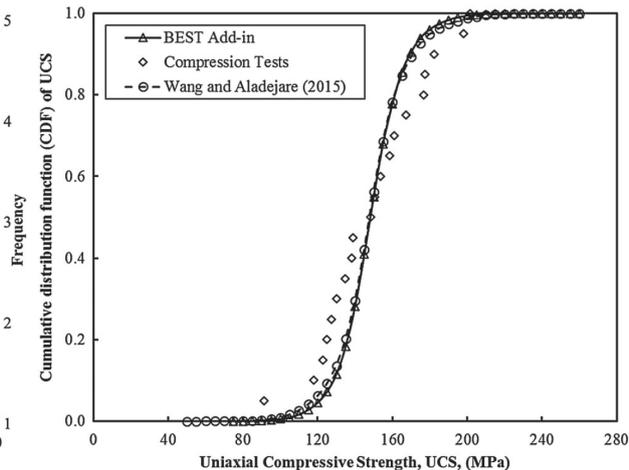


Figure 5-7 UCS CDF plot

In addition to the $Is_{(50)}$ data and prior distribution, a transformation model also needs to be defined in BEST. Wang and Aladejare (2015) have performed a model selection study using the $Is_{(50)}$ data and prior knowledge and found that the regression developed by Chau and Wong (1996) is a suitable model for this specific site. Their regression is expressed as:

$$Is_{(50)} = 0.061UCS + \varepsilon \quad (5-3)$$

where ε represents the model uncertainty and follows a normal distribution with a zero mean and standard deviation $\sigma_{\varepsilon} = 2.073$.

Using the prior knowledge, $Is_{(50)}$ data and Eq. (5-3), BEST is executed to generate 30,000 equivalent samples of UCS. The estimated mean and standard deviation of the UCS equivalent samples from BEST Add-in is shown in Table 5-4. These statistics are compared with those obtained from laboratory compression tests (Mishra and Basu 2012) or Matlab (Wang and Aladejare 2015). The UCS PDF and CDF are constructed and displayed in Figures 5-6 and 5-7, respectively, by a solid line with triangle markers. The UCS PDF and CDF obtained from Matlab (Wang and Aladejare 2015) are also included in Figures 5-6 and 5-7 by a dashed line with circle markers. In Figure 5-6, the solid line with triangle markers and dashed line with circle markers are both plotted on the primary vertical axis which represents the PDF of UCS. In both figures, the dashed line virtually overlaps with the solid line, indicating that the equivalent samples from BEST are in good agreement with those from Matlab (Wang and Aladejare 2015). In addition, the 20 compression test results obtained from the same site (i.e., UCS values from Table 5-3) are included in Figure 5-6 and they are plotted on the secondary vertical axis which represents the frequency of the test results. Figure 5-6 shows that 16 out of 20 UCS values from laboratory compression tests fall within the 90% inter-quartile range of the equivalent samples (121 MPa and 177 MPa) generated from BEST. The CDF of the 20 compression tests results is also included in the CDF plot (see Figure 5-7) and it shows that the CDF estimated from the equivalent samples is consistent with that from laboratory compression tests. This suggests that the BEST Add-in performs satisfactorily in obtaining reasonable statistics and probability distributions of ground properties using prior knowledge and limited measurement data.

The UCS characteristic value may be selected from the statistics above. For example, if the UCS characteristic value is taken as the mean or 5% fractile value, it is about 148 MPa and 121 MPa, respectively, at this specific site.

5.7 CONCLUDING REMARKS

This report summarized some recent developments on Bayesian statistics and prior knowledge in geotechnical site characterization, and particularly focuses on selection of characteristic values for soil

or rock properties in engineering practice. It covers the Bayesian equivalent sample algorithm, quantification of prior knowledge, user-friendly software in Excel, and application examples. The Bayesian method and software are developed for facilitating selection of ground property characteristic values in routine engineering practices, such as designing a foundation or retaining wall based on a limited number of SPT-N values obtained from a project site.

It is also worthwhile to note that the “User-defined Model” in BEST Add-in is applicable for the case of direct measurement of a design property. When there are a limited number of direct measurements of a design property (e.g., UCS), these direct measurement data can also be used as input to BEST and integrated with prior knowledge to generate a large number of equivalent samples of the design property.

5.8 REFERENCES

- Aladejare, A. E. and Wang, Y. (2017). Evaluation of rock property variability. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 11(1), 22-41.
- Ang, A. H.-S. and Tang, W. H. (2007). *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*. John Wiley and Sons, New York.
- Baecher, G. B. and Christian, J. T. (2003). *Reliability and Statistics in Geotechnical Engineering*, John Wiley & Sons, Hoboken, New Jersey, 605p.
- Briaud, J. L. (2000). The national geotechnical experimentation sites at Texas A&M University: clay and sand. A summary. *National Geotechnical Experimentation Sites*, ASCE GSP 93, 26–51.
- Cao, Z. and Wang, Y. (2014). Bayesian model comparison and characterization of undrained shear strength. *Journal of Geotechnical & Geoenvironmental Engineering*, 140(6), 04014018, 1-9.
- Cao, Z., Wang, Y., and Li, D. (2016). Quantification of prior knowledge in geotechnical site characterization. *Engineering Geology*, 203, 107-116.
- CEN (2004) EN 1997-1:2004 *Geotechnical Design – Part 1: General Rules*.
- Chau, K. T. and Wong, R. H. C. (1996) Uniaxial compressive strength and point load strength of rocks, *International Journal of Rock Mechanics & Mining Sciences*, 33, 183–188.
- Ching, J., Chen, J. R., Yeh, J. Y., and Phoon, K. K. (2012). Updating uncertainties in friction angles of clean sands. *Journal of Geotechnical and Geoenvironmental Engineering*, 138(2), 217-229.
- Clayton, C. R. I., Matthews. M. C., Simons, N. E. (1995). *Site investigation*. Blackwell Science, Cambridge, Mass., USA.
- Fenton, G. A., Naghibi, F., Dundas, D., Bathurst, R. J., and Griffiths, D. V. (2016). Reliability-based geotechnical design in 2014 Canadian Highway Bridge Design Code. *Canadian Geotechnical Journal*, 53(2), 236-251.
- Kulhawy, F. H. and Mayne, P. W. (1990). *Manual on Estimating Soil Properties for Foundation Design*, Report EL 6800, Electric Power Research Inst., Palo Alto, 306p.
- Mayne, P. W., Christopher, B. R., and DeJong, J. (2002). *Subsurface Investigations – Geotechnical Site Characterization*, No. FHWA NHI-01-031, Federal Highway Administration, U. S. Department of Transportation, Washington D. C.
- McLachlan, G. and Peel, D. (2000). *Finite mixture models*. New York, John Wiley & Sons.
- Mishra, D. A. and Basu, A. (2012). Use of the block punch test to predict the compressive and tensile strengths of rocks. *International Journal of Rock Mechanics and Mining Sciences*, 51, 119-27.
- Orr, T. L. L. (2005). Design examples for the Eurocode 7 Workshop. *Proceedings of the international workshop on the evaluation of Eurocode 7*, Trinity College, Dublin, 67–74.
- Orr, T. L. L. (2017). Defining and selecting characteristic values of geotechnical parameters for designs to Eurocode 7, *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 11(1), 103-115.
- Paikowsky, S. G., Birgisson, B., McVay, M., Nguyen, T., Kuo, C., Baecher, G., Ayyub, B., Stenersen, K., O’Malley, K., Chernauskas, L., and O’Neill, M. (2004). *Load and resistance factor design (LRFD) for deep foundations*. NCHRP Report 507, Transportation Research Board, Washington, DC.
- Paikowsky, S. G., Canniff, M. C., Lesny, K., Kisse, A., Amatya, S., and Muganga, R. (2010). *LRFD design and construction of shallow foundations for highway bridge structures*. NCHRP Report 651, Transportation Research Board, Washington, DC.
- Phoon K. K. (2017). Role of reliability calculations in geotechnical design, *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 11(1), 4-21.

- Phoon, K. K. and Kulhawy, F. H. (1999a). Characterization of geotechnical variability. *Canadian Geotechnical Journal*, 36(4), 612-624.
- Phoon, K. K. and Kulhawy, F. H. (1999b). Evaluation of geotechnical property variability. *Canadian Geotechnical Journal*, 36(4), 625-639.
- Phoon, K. K., Kulhawy, F. H., and Grigoriu, M. D. (2003a). Development of a reliability-based design framework for transmission line structure foundations. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(9), 798-806.
- Phoon, K. K., Kulhawy, F. H., and Grigoriu, M. D. (2003b). Multiple resistance factor design (MRFD) for shallow transmission line structure foundations. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(9), 807-818.
- Phoon, K. K., Retief, J. V., Ching, J., Dithinde, M., Schweckendiek, T., Wang, Y., and Zhang L. M. (2016). Some observations on ISO2394: 2015 Annex D (Reliability of Geotechnical Structures). *Structural Safety*, 62, 24-33.
- Robert, C. and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer.
- Schneider, H. R. and Schneider, M. A. (2013). Dealing with uncertainties in EC7 with emphasis on determination of characteristic soil properties. *Modern Geotechnical Design Codes of Practice*, P. Arnold et al. (Eds.), ISO Press, 87-101.
- Trautmann, C.H. and Kulhawy, F.H. (1983). Data sources for engineering geologic studies. *Bull. Assn. Eng. Geologists*, 20(4), 439-454.
- Vick, S. G. (2002). *Degrees of Belief: Subjective Probability and Engineering Judgment*. ASCE Press, Reston, Virginia.
- Wang, Y., Akeju, O. V., and Cao, Z. (2016b). Bayesian Equivalent Sample Toolkit (BEST): an Excel VBA program for probabilistic characterisation of geotechnical properties from limited observation data. *Georisk*, 10(4), 251-268.
- Wang, Y. and Aladejare, A. E. (2015). Selection of site-specific regression model for characterization of uniaxial compressive strength of rock. *International Journal of Rock Mechanics & Mining Sciences*, 75, 73-81.
- Wang, Y. and Cao, Z. (2013). Probabilistic characterization of Young's modulus of soil using equivalent samples. *Engineering Geology*, 159, 106-118.
- Wang, Y., Cao, Z., and Li, D. (2016a). Bayesian perspective on geotechnical variability and site characterization. *Engineering Geology*, 203, 117-125.
- Wang, Y., Zhao, T., and Cao, Z. (2015). Site-specific probability distribution of geotechnical properties. *Computers and Geotechnics*, 70, 159-168.

Discussions and Replies

During the preparation of this report, many valuable and insightful comments and suggestions have been gratefully received. Some of them have been incorporated in this final draft report, including those made by Marcos Arroya, Zijun Cao and Trevor Orr, and hence, they do not appear in this section. The other comments and suggestions are listed below by the date they were received.

Discussion by Tim Lämsivaara (Tampere University of Technology, Finland)

I think this is a very important issue and it would be great to have some progress in the form of guidelines in the determination of characteristic values. I think it would be good to first also discuss what is the definition of a characteristic value. Personally I don't think that the definition in Eurocodes as a 5% fractile value is a very good one.

Reply by Yu Wang (City University of Hong Kong, Hong Kong)

This draft final report contains a section (i.e., Section 5.2) on the definition of characteristic value. However, the definition of characteristic value for ground properties itself might be an intriguing issue, and detailed discussion on this issue is beyond the scope of this report, which focus on using Bayesian statistics and prior knowledge to facilitate the proper selection of characteristic value **for a given definition of characteristic value**. Detailed discussion on definition of characteristic value in EC7 is referred to Schneider and Schneider (2013) and Orr (2017).

Discussion by KK Phoon (National University of Singapore, Singapore)

Some observations on characteristic value

EN 1997-1:2004, 2.4.5.2(2) recommends that the “characteristic value of a geotechnical parameter shall be selected as a cautious estimate of the value affecting the occurrence of the limit state.” Much attention has been focused on how to obtain a “cautious estimate”. For example, EN 1997-1:2004, 2.4.5.2(11) notes that “If statistical methods are used, the characteristic value should be derived such that the calculated probability of a worse value governing the occurrence of the limit state under consideration is not greater than 5%.” A note to this clause clarifies that “In this respect, a cautious estimate of the mean value is a selection of the mean value of the limited set of geotechnical parameter values, with a confidence level of 95%; where local failure is concerned; a cautious estimate of the low value is a 5% fractile.” Less attention is focused on the “value affecting the occurrence of the limit state”.

In my opinion, the general statement describing the characteristic value “as a cautious estimate of the value affecting the occurrence of the limit state” is sensible.

However, there is an under-stated difficulty in making this statement sufficiently concrete *for codification*.

We acknowledge the critical role of engineering judgment. Sensibility and reality checks on all design aspects are clearly dependent on informed judgment. This discussion focuses only on those numerical aspects that probabilistic methods can add value to the estimation of the characteristic value.

In my opinion, it is crucial to examine the following 2 elements: (1) “value affecting the occurrence of the limit state” and (2) “cautious estimate” separately.

Value affecting the occurrence of the limit state

One can visualize the first element “value affecting the occurrence of the limit state” more clearly by assuming there is no uncertainty. In other words, we only look at *one* realization of a random field. It can be spatially heterogeneous or even spatially homogeneous in horizontal and/or vertical directions when the scales of fluctuation are large in those directions. We can assume we have sufficient direct measurements to safely ignore transformation and statistical uncertainties. Under this ideal *deterministic* condition, the first element is a question in *mechanics*.

Consider a bored pile as a concrete example. One can argue that the side resistance depends on the average strength in each layer supporting the pile. The tip resistance can conceivably be seen as depending on the average strength below one diameter of the tip. The mobilized strength “value” is related to the average over an influential volume of soil.

Consider a slope stability problem as a second example. For a homogeneous slope, the mobilized strength along the *critical* slip surface can be viewed as another average of a homogeneous soil mass. For a heterogeneous slope, the same argument applies, but an intriguing difficulty arises in this instance. If an engineer analyses this slope with characteristic values in each layer estimated from borehole/field test data (however they are selected), the relative magnitude of the characteristic values can affect the location of the critical slip surface emerging from a mechanical analysis, say limit equilibrium or finite element analysis. This slip surface may or may not be the same as the one emerging from a finite element analysis adopting the strength reduction approach.

The central difficulty is that the “occurrence of the limit state” is an *output* of a mechanical analysis; it is not linked to borehole/field test data (*input*) in a straightforward way although an experienced engineer could make an informed guess. An inexperienced engineer may judge incorrectly without guidance from a mechanical analysis.

Cautious estimate

The second element “cautious estimate” only arises in the presence of uncertainty. Spatial variability and a range of uncertainties (transformation, statistical, etc.) allow a range of possible values and possible scenarios to exist, because measured data are too limited to restrict a property to a single value and a profile to a single scenario. However the “value affecting the occurrence of the limit state” is defined, it is clear that it will also take a range of values in the presence of uncertainty, i.e. it is a random variable.

The key point is that this random variable is not necessarily the same as the random variable describing a soil property at a “point”.

If one accepts the average along the pile shaft as the “value affecting the occurrence of the limit state”, then the relevant random variable is the spatial average over the length of the shaft. If one further accepts that a customarily 95% confidence level is sufficient, that is select a threshold value so that the actual value will exceed this threshold value with 95% probability, then this definition is partially consistent with the statement “a cautious estimate of the mean value is a selection of the mean value of the limited set of geotechnical parameter values, with a confidence level of 95%”. This definition is also consistent with a 5% quantile (or fractile) of the spatial average (in the reliability literature).

The key difference is that the spatial average is a function of the length of the shaft while this length effect is not explicitly stated with reference to the “mean” in EN 1997–1:2004, 2.4.5.2(11). It goes without saying that the spatial average and the mean are affected by statistical uncertainty and this is related to the number of measurements.

For the statement “where local failure is concerned, a cautious estimate of the low value is a 5% fractile”, it is somewhat ambiguous in probabilistic terms, but my interpretation is that EN 1997–1:2004, 2.4.5.2(11) is referring to 5% quantile (or fractile) of the soil property at a “point”.

ISO2394:2015 (Annex D) Section D.5.5 discusses the need to clarify the mechanical and probabilistic aspects of the characteristic value (refer to excerpt below).

Closing thoughts for discussion

1. Statistical analysis of site information typically produces the statistics (mean, coefficient of variation) for a soil property at a “point”. The “point” random variable described by these statistics is not the same as the “mobilized” random variable “affecting the occurrence of the limit state”. To be specific, if the limit state involves bearing capacity, the “mobilized” random variable appearing in the bearing capacity equation is not the same as the “point” random variable. In other words, the “mobilized” random variable is the one relevant to the resistance/response calculation model.
2. The “occurrence of the limit state” is an output of a mechanical analysis. It is not straightforward to define this “mobilized” random variable from input soil data *alone*. This definition can depend on the limit state. EN 1997–1:2004, 2.4.5.2(11) already noted that the distinction between non- local and local failures. In my opinion, more research is needed. The best one can hope for is to define a “mobilized” random variable so that it approximates the probabilistic solution from a mechanical analysis (say random finite element analysis) with reasonable accuracy.
3. In my opinion, there are merits to replace term “mean” stated in EN 1997–1:2004, 2.4.5.2(11) by the “spatial average”:
 - a. “Mean” is a statistics for a set of measurements. It does not focus the attention of the engineer on the limit state. The spatial average depends on the averaging domain

- (line/surface/volume) where the limit state is expected to take place.
- b. One concrete improvement is that the characteristic value will depend on the size of the averaging domain (e.g. length of the pile shaft) if we refer to “spatial average”.
 - c. The uncertainty in the mean only depends on the number of measurements (statistical uncertainty). Spatial average can include other sources of uncertainties, including statistical uncertainty (see #7 below). In this sense, it is a more general and more physically meaningful concept.
4. Even the classical spatial average (Vanmarcke 1977) is an approximate solution, because the average along the critical slip path/surface is not the same as the average along a *fixed* prescribed path/surface. The former path/surface is partially affected by the distribution of “weak zones” in a spatially varying soil mass, which changes from realization to realization. The literature says that the classical spatial average is reasonable if the scale of fluctuation does not take a “critical” value (Ching & Phoon 2013; Ching et al. 2014, 2016a). Hence, the spatial average can be retained as a first-order approximation of the mobilized random variable for the time being. The 5% fractile of this spatial average can be used as a more concrete definition of the characteristic value for the time being, when it is suitably qualified.
 5. For limit states not governed by spatial averages, e.g. local failure, seepage, etc., more research is needed to clarify if the “mobilized” and “point” random variables are the same.
 6. From the perspective of a “mobilized” random variable, the terms “confidence level of 95%” and “5% fractile” are the same. I would recommend harmonizing these terms in EN 1997-1:2004, 2.4.5.2(11) to “5% fractile of the mobilized random variable”. The mobilized random variable can be approximated by the spatial average or other random quantities depending on the occurrence of the limit state.
 7. The coefficient of variation of the mobilized random variable is affected by statistical uncertainty (due to limited measurements), spatial variability (due to spatial extent of limit state), measurement error (due to measurement), and transformation uncertainty (due to conversion from measurements to desired properties). Statistical uncertainty is already covered in EN 1997-1:2004, 2.4.5.2, but extensive research has shown that other sources are present (Ching et al. 2016b; Phoon et al. 2016).

References

1. Ching, J. & Phoon, K. K. (2013b). Probability distribution for mobilized shear strengths of spatially variable soils under uniform stress states. *Georisk*, 7(3), 209-224.
2. Ching, J., Phoon, K. K. & Kao, P. H. (2014). Mean and variance of the mobilized shear strengths for spatially variable soils under uniform stress states. *ASCE Journal of Engineering Mechanics*, 140(3), 487-501
3. Ching, J., Lee, S. W. & Phoon, K. K. (2016a). Undrained strength for a 3D spatially variable clay column subjected to compression or shear. *Probabilistic Engineering Mechanics*, 45, 127-139.
4. Ching, J. Y., Li, D. Q. & Phoon, K. K. (2016b). Statistical characterization of multivariate geotechnical data, Chapter 4, *Reliability of Geotechnical Structures in ISO2394*, Eds. K. K. Phoon & J. V. Retief, CRC Press/Balkema, 89-126.
5. Phoon, K. K., Prakoso, W. A., Wang, Y., Ching, J. Y. (2016). Uncertainty representation of geotechnical design parameters, Chapter 3, *Reliability of Geotechnical Structures in ISO2394*, Eds. K. K. Phoon & J. V. Retief, CRC Press/Balkema, 2016, 49-87.
6. Vanmarcke, E. H. (1977). “Probabilistic modeling of soil profiles”, *Journal of Geotechnical Engineering Division, ASCE*, 103(GT11), 1227 – 1246.

Excerpt from ISO2394:2015 General principles on reliability for structures, Annex D Reliability of Geotechnical Structures

D.5.5 Characteristic value

The concept of a “characteristic value” is intrinsically linked to semi-probabilistic formats, particularly the partial factor approach. In this approach, using the ultimate limit state as an example, a “characteristic value” for a soil parameter (e.g., undrained shear strength), is divided by a strength partial factor to produce a “design” value and the geotechnical capacity based on this design value should be larger than the design load (characteristic load multiplied by a load factor).

The soil parameter must be defined such that it is relevant to the limit state equation. For

example, if a single undrained shear strength parameter appears in a slope stability equation, then the relevant physical definition is the spatial average along the most critical failure path. It is neither the undrained shear strength at a point in the soil mass nor a spatial average along a prescribed line in the soil mass. The emphasis in the geotechnical literature is on clarifying this physical aspect of the characteristic value, which is justifiably so.

It is necessary to make clear the physical meaning of the characteristic soil parameter before the uncertainty aspect could be rationally considered. For illustration, the characteristic undrained shear strength for the shaft friction of a pile is the spatial average along the length of the pile, while the characteristic undrained shear strength for the end bearing of a pile is the spatial average within a bulb of soil below the pile tip. When reliability analysis is carried out, the performance function will contain two random variables following two distinct probability distributions for these spatial averages.

When semi-probabilistic design is carried out, it would be necessary to select a single value characteristic of each probability distribution. This value may refer to the mean or to the lower 5% quantile. The statistical estimation of these characteristic values is subject to the same statistical uncertainties underlying the probability distributions appearing in reliability analysis. Clearly, this statistical aspect of the characteristic value is distinctive from the physical aspect of the characteristic value.

In principle, partial factors can be calibrated to achieve a prescribed target reliability index for any statistical definition of the characteristic value. In practice, it is known that a partial factor calibrated for the mean value could change significantly if the COV of the input random variable changes. This limitation is less severe for a partial factor calibrated using say the lower 5%-quantile. Hence, if the COV of an input random variable varies over a wide range within the scope of design scenarios covered in a design code and if there is a practical need to simplify presentation of a partial factor as a single number rather than as a function of COV, the lower 5% quantile definition is preferred (except in the case considering non-linear responses where special considerations must be made). It is useful to reiterate that the key function of a RBD code is to achieve a prescribed target reliability index (typically a function of limit state and importance of structure) over a range of commonly encountered design scenarios and not for a specific design scenario. The statistical definition of the characteristic value should be viewed within this broader context of what a RBD code intends to achieve, rather than adherence to past practice or a component separate from reliability calibration. In other words, the ensuing design is produced by design values, which is the product of partial/resistance factors with characteristic values, not merely characteristic values alone. There are practical concerns regarding: (1) estimation of quantiles reliably from limited data and (2) quantiles falling below known lower bounds (e.g. residual friction angle) due to inappropriate choice of unbounded probability distribution functions. However, both concerns do not merely affect the characteristic value but fundamentally affect the reliability analyses underlying code calibration as well.

Reply by Yu Wang (City University of Hong Kong, Hong Kong)

Thank you very much for adding very informative and valuable insights to this discussion. Although the definition of ground property characteristic values is clearly given in the Eurocodes (both the head code and EC7), it seems that there are some practical difficulty in implementing this definition in geotechnical practice, due to, for example, site-specific nature of ground properties, generally small sample size of site-specific data, usage of engineering experience and judgment (e.g., previous data from similar project or site conditions). In addition, the characteristic values in EC7 are linked with the output of a design calculation, i.e., “a question in *mechanics*” as pointed out in the discussion above. This indeed is a dilemma of “Which came first: the chicken or the egg?” because the characteristic values are supposed to be defined first for the subsequent design calculation.

A possible way out of this dilemma is to revise the definition of characteristic values in such a way that it does not involve the design calculation (or the “*mechanics*” of a design) when selecting the characteristic values. In other words, the characteristic value may be defined to reflect **only** the existing information on the site, including site-specific test data and pre-existing engineering experience and judgment as shown in this report.

The advantage of such a definition is that it allows different practitioners to arrive at the same characteristic value from a given (i.e., the same) set of site-specific test data and pre-existing engineering experience and judgment, even for different design problems (e.g., foundations, slope

stability). Then, based on different design problems, different influence zones are identified, and the characteristic values within the corresponding influence zones are used in the design calculations. For example, the influence zone for a shallow foundation is about one diameter of the foundation width below the foundation, and that for the side resistance of a pile is the length of the pile. This is consistent with the conventional (or deterministic) practice in geotechnical engineering.

The disadvantages of such a definition is that many sophisticated and important issues raised in the discussion above will NOT be considered in the definition of characteristic value, such as the “mobilized” random variable for different design problems (e.g., foundations, slope stability) and spatial averaging along a fixed surface vs an unknown surface. All these important issues involve the “*mechanics*” of the design problem under consideration, and they should be properly considered by other means in design codes, such as partial factors. Calibration of the partial factors always involves the “*mechanics*” of the design problem under consideration, and it is problem specific. It might be logical and convenient to remove all the “*mechanics*” related issues from the definition of characteristic values and incorporate them systematically during the calibration of partial factors.

It is also worth noting that, as pointed out in the discussion above, “*The concept of a characteristic value is intrinsically linked to semi-probabilistic formats, particularly the partial factor approach.*” If using a single characteristic value in semi-probabilistic formats is NOT able to properly reflect the “*mechanics*” of the design problem under consideration (e.g., the occurrence of different failure modes for different characteristic values adopted), direct probability-based design methods may be used for these sophisticated design problems. Detailed discussion on the direct probability-based design methods in geotechnical engineering is given by Wang et al. (2016).

Reference:

Wang, Y., Schweckendiek, T., Gong, W., Zhao, T., and Phoon, K. K. (2016). Direct probability-based design methods, Chapter 7 in Reliability of Geotechnical Structures In ISO2394, Edited by K. K. Phoon, J. V. Retief, Pages 193–226, DOI: 10.1201/9781315364179-8.

Following-up discussion by Jianye Ching (National Taiwan University, Taiwan)

I think we all agree that mechanics must enter into the design process. The question is where. Mechanics can enter through the characteristic value - this is the current EC7. In your reply, you propose that mechanics can enter only through the partial factors. The characteristics value can be mechanics free. You mentioned the chicken-egg dilemma, which is quite an interesting point.

My view is the following:

1. The chicken-egg dilemma is indeed inconvenient because the engineer needs to think about the mechanics during the planning of site investigation and the determination of characteristics value. However, I think this is not really a dilemma but a healthy thought process. It forces design engineers to think about the mechanics when he/she plans for site investigation and interprets the characteristics value. Without this thought process for mechanics, there seems to be no general guideline for planning site investigation and choosing characteristics value.

For instance, suppose we ask an engineer to plan for site investigation without telling him/her the type of construction (a foundation or an excavation). Without the mechanics in mind, he/she may decide to investigate up to 20 m deep, but what if the real construction is a 50 m deep drilled shaft.

You mentioned the concept of influence zone that is mechanics free. For instance, suppose we tell the design engineer that the influence zone is 20 m deep without telling him/her the type of construction. Without the mechanics in mind, he/she may decide to take the average value over the entire 20 m as the characteristics value. This makes sense for the shaft resistance of a pile, but what if the real construction is a 20 m mat foundation with critical slip surface passing through a weak zone.

My view is that the chicken-egg process is a healthy thought process with at least two advantages: (a) it forces engineers to think about the mechanics and (b) it prevents engineers from producing irrelevant site investigation plans.

2. It is true that the calibration of partial factors involves mechanics, e.g., partial factors can be calibrated by FORM, and FORM requires the definition of the limit-state function, and mechanics is in the limit-state function. However, this does not imply that the issue of irrelevant site investigation planning in #1 can be resolved by careful partial factor calibration. In fact, if the site investigation planning is irrelevant, the design will be irrelevant as well even if the partial factors are calibrated by a procedure that considers mechanics.

In my opinion, partial factor calibration considers "mechanics" for a different purpose. The limit-state function is needed so that the sensitivity of various uncertainty will be correctly quantified and reflected. For instance, for a limit-state function more sensitive to friction angle than to unit weight, a more conservative partial factor should be calibrated for friction angle. However, doing this sensitivity correctly does not mean we no longer need a relevant site investigation plan.

Reply by Yu Wang (City University of Hong Kong, Hong Kong)

Thank you for following up this discussion. The comments and discussions are very much appreciated. Planning of site investigation is an important issue, but this discussion group (or Chapter 5) focuses on how to use Bayesian statistics and prior knowledge to facilitate selection of characteristic values, **for a given definition of characteristic value and a given set of reasonable site investigation data.** Without reasonable data and/or prior knowledge (e.g., engineering judgment and experience), it is impossible to properly select characteristic values for ground properties. Different influence zones for different geotechnical structures are one kind of prior knowledge that is commonly used in engineering practice, including planning of site investigation.

It is a tough decision for code developers to strike the balance among different competing, or even contradicting, issues, such as the ideal design process (including site investigation) that can be specified in the codes and the practicality when the design codes are implemented by practicing engineers with diverse background in engineering practice. What I mentioned in the reply above is just an idea for code developers to consider.

I personally believe that the definition of characteristic values and the calibration of partial factors must be compatible with each other, because they act together to achieve the target performance that the design codes aim to achieve. For example, if the spatial variability is considered in the definition of characteristic value, it might be tricky to consider the spatial variability **again** in the calibration of partial factors.

Discussion by Brian Simpson (Arup, UK)

I found this report very interesting. I do think that Bayesian methods could be very helpful in deriving characteristic values of parameters. Thanks to Yu Wang.

I have a concern, however, about the definition of characteristic value, as understood in EC7. In the report, it seems to be treated as a 5% fractile of test results, which is not the intention of EC7. The author might want to comment further on this.

The basic definition given in EC7 is that the characteristic value is a "a cautious estimate of the value affecting the occurrence of the limit state" (2.4.5.2(2)). The paragraphs that follow this are important, including (7):

The zone of ground governing the behaviour of a geotechnical structure at a limit state is usually much larger than a test sample or the zone of ground affected in an in situ test. Consequently the value of the governing parameter is often the mean of a range of values covering a large surface or volume of the ground. **The characteristic value should be a cautious estimate of this mean value.**

And (9):

When selecting the zone of ground governing the behaviour of a geotechnical structure at a limit state, it should be considered that this limit state may depend on the behaviour of the supported structure. For instance, when considering a bearing resistance ultimate limit state for a building resting on several footings, the governing parameter should be the mean strength over each individual zone of ground under a footing, if the building is unable to resist a local failure. If, however, the building is stiff and strong enough, the governing parameter should be the mean of these mean values over the entire zone or part of the zone of ground under the building.

Paragraph (11) says:

If statistical methods are used, the characteristic value should be derived such that the calculated probability of a worse value governing the occurrence of the limit state under consideration is not greater than 5%.

And the note attached to that is important:

In this respect, a cautious estimate of the mean value is a selection of the mean value of the limited set of geotechnical parameter values, with a confidence level of 95%; where local failure is concerned, a cautious estimate of the low value is a 5% fractile.

It should be clear from this that the characteristic value required by EC7 is not a 5% fractile of test results, but rather there is a 5% probability that a worse value could be representative of the whole body or surface of soil that governs the occurrence of the limit state.

Schneider (1997) suggested that given at least 10 test results, the 5% probability value for the mean of the population lies about half a standard deviation from the mean of the test results. His ideas have been developed further since then. The following figure, taken from Simpson et al (2009) shows that this is much nearer the mean than the 5% fractile of the test results (obviously). The following paragraph, taken from the same paper, suggests that this may not be very different from North American practice:

A similar proposal was made by Dahlberg and Ronold (1993) for design of offshore foundations and recommended by Becker (1996) for more general use. This involves the use of a “conservatively assessed mean” (CAM) as the characteristic value, also about 0.5 standard deviations from the mean of the test results. These authors state that for a normal distribution 75% of the measured values would be expected to exceed this value. (More accurately, this requires an offset of 0.69 standard deviations from the mean, for a normal distribution, as shown in Figure 2.2). Foye et al (2006b) take up the same idea proposing to use a CAM with 80% exceedance, equivalent to 0.84 standard deviations below the mean of the test results for a normal distribution. These proposals are made in the development of North American practice, though at present the AASHTO Specifications do not use the term “characteristic” but refer less specifically to nominal values related to “permissible stresses, deformations, or specified strength of materials”.

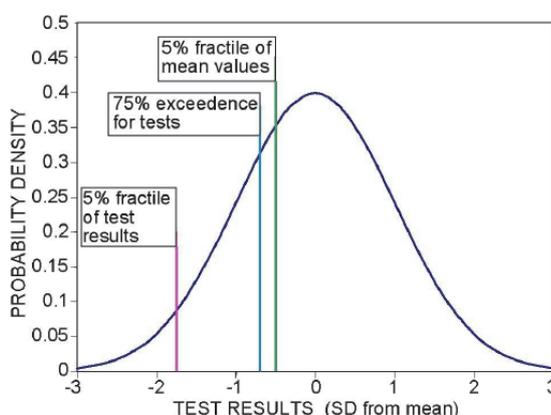


Figure 2.2: Alternative derivations of “characteristic values”

REFERENCES

- Becker, D.E. 1996. Eighteenth Canadian Geotechnical Colloquium: Limit states design of foundations. I: An overview of the foundation design process. *Canadian Geotech J*, 33(6), 956-983.
- Dahlberg, R and Ronold, KO (1993) Limit state design of offshore foundations. *Proc Int Symp Limit state design in geotechnical engineering*, Vol 2, pp491-500. Danish Geotechnical Society.
- Foye, K.C. Salgado, R. & Scott, B. 2006b. Resistance factors for use in shallow foundation LRFD. *J. Geotech. Geoenviron. Eng.*, 132(9), 1208–1218.
- Schneider H R (1997) Definition and determination of characteristic soil properties. Contribution to Discussion Session 2.3, XIV ICSMFE, Hamburg. Balkema.
- Simpson, B, Morrison, P, Yasuda, S, Townsend, B, and Gazetas, G (2009) State of the art report: Analysis and design. *Proc 17th Int Conf SMGE*, Alexandria, Vol 4, pp. 2873-2929.

Reply by Yu Wang (City University of Hong Kong, Hong Kong)

Thank you very much for the positive comment. This draft final report adds a section (i.e., Section 5.2) on the definition of characteristic value to clarify the comments raised in the discussion. However, the definition of characteristic value for ground properties itself might be an intriguing issue, and detailed discussion on this issue is beyond the scope of this report, which focus on using Bayesian statistics and prior knowledge to facilitate the proper selection of characteristic value **for a given definition of characteristic value**. Detailed discussion on definition of characteristic value in EC7 is referred to Schneider and Schneider (2013) and Orr (2017). The illustrative examples in this report have also been revised to highlight that, when the characteristic value is defined in different way, different numerical values will be obtained accordingly.

Chapter 6 Bayesian Method: A Natural Tool for Processing Geotechnical Information

Lead discussor:

Jie Zhang (contributor of 6.1, 6.3, 6.5, and 6.6)

cezhangjie@gmail.com

Discussors (alphabetical order):

Philip Boothroyd, Michele Calvello (contributor of 6.4.6), Malcolm Eddleston, Antonio Canavate Grimal (contributor of 6.3), Papaioannou Iason (contributor of 6.4.2), Zhe Luo (contributor of 6.2 and 6.5), Shadi Najjar (contributor of 6.4.1 and 6.4.3), Adrian Rodriguez-Marek (contributor of 6.4.4), Daniel Straub (contributor of 6.4.2), Marco Uzielli, Yu Wang (contributor of 6.4.5)

(All members have participated the discussion and revision of this report)

6.1 INTRODUCTION

In geotechnical analysis and prediction, it is a common practice for engineers to consider information from multiple sources. The combination of information is often made based on engineering judgement and experience. From a mathematical perspective, the Bayesian method is a formal and scientific tool to combine information from multiple sources for the purpose of updating prior knowledge given new information. The potential application of Bayesian methods in geotechnical engineering has been explored by many researchers. In practice, however, the application of Bayesian methods is still quite limited. The objective of this report is to illustrate the benefits of using Bayesian methods for processing geotechnical information. The structure of this report is as follows. First, the need for information combination in geotechnical engineering is analyzed, followed by a brief review of the theory and computational techniques for Bayesian analysis. Then, several examples are presented to show how the Bayesian method can be used to tackle different geotechnical problems. Finally, the potential application of the Bayesian method in the Eurocode is explored, and the challenges for application of the Bayesian method in geotechnical engineering are discussed. We hope that this report can provide a useful guide for practicing engineers to identify the merits of solving geotechnical problems from a Bayesian perspective, and thus encourage them to take advantage of the Bayesian method in practice. The computational details are not the emphasis of this report.

6.2 NEED FOR INFORMATION COMBINATION IN GEOTECHNICAL ENGINEERING

The practice of geotechnical engineering is always amenable to various uncertainties resulting from insufficient site investigation data, time and budget constraint, local experience, expected natural hazards, unpredictable environmental and social impact, construction disturbance, imperfect design models, etc. From the design perspective, uncertainties can be categorized into parameter uncertainty and model uncertainty. Each source of uncertainty can potentially result in unsatisfactory geotechnical performance with associated casualty and economic loss, so it is necessary to explicitly evaluate and quantify parameter uncertainty and model uncertainty in the context of geotechnical reliability-based design, including the commonly-known load and resistance factor design.

Given an established design method in a design code, a realistic challenge for practicing geotechnical engineers lies in how to determine design values for soil parameters. This challenge is caused by several factors including the uncertain depositional process and soil history and the limited

number of boreholes that the clients afford to characterize the soil profile and determine the design soil parameter values. To this end, either factor-of-safety-based design using conservatively estimated parameters or probabilistic analysis can be used. In either of the two approaches, the estimated design soil parameters are based on the limited and existing information gained through the initial geological survey, in-situ borehole drilling and laboratory testing, and local experience. Existing knowledge could lead to prior information that may not represent the actual soil conditions and there is a need to update the soil parameters using new information, such as the field observed responses (Peck 1969).

Variability in soil properties stems from various sources of uncertainty that can be grouped into two categories: aleatory uncertainty and epistemic uncertainty. The word aleatory is evolved from the Latin *alea*, which means the rolling of dice, and the aleatoric uncertainty refers to the intrinsic randomness of a phenomenon. In geotechnical engineering, the aleatory uncertainty includes spatial variability and random testing errors. The word epistemic is evolved from the Greek *episteme* (i.e., knowledge). The epistemic uncertainty is caused by lack of knowledge or data. In geotechnical engineering, the epistemic uncertainty is often related to measurement procedures and limited data availability (Whitman 1996).

Both aleatory uncertainty and epistemic uncertainty can be addressed using combined prior information and newly observed information in geotechnical engineering. For example, the aleatoric uncertainty that is reflected in a random field of a spatially varying soil parameter can be calibrated given sufficient field load tests. Given more data from case histories, the knowledge on soil parameters, the epistemic uncertainty, can be back-calculated as posterior information. As more observational data are introduced, the uncertainties can be reduced through updating the mean values and decreasing the variance of each parameter, which can benefit the subsequent stages of design and construction. Admittedly, the amount of data required depends on the number of uncertain variables. There is always a trade-off between the site investigation effort and the improved knowledge on design soil parameters. It is noted that the site investigation cost can be optimized using approaches such as the Bayesian method. Depending on the specific problem, the parameter uncertainty, or model uncertainty or both can be updated through back analysis.

6.3 BAYESIAN METHOD: CONCEPT AND COMPUTATIONAL TECHNIQUES

The uncertainties in the geotechnical design can be formally modelled through random variables. Those random variables are represented with probability distributions that quantify the knowledge in the model parameters and the model itself. The Bayesian approach provides a rigorous framework to reduce the uncertainties associated with the random variables when more information is available. Let vector θ denote uncertain variables to be updated with the observation data \mathbf{D} . The Bayesian method can be applied to both continuous variables and discrete random variables. As an example, suppose the elements of θ are all continuous variables, and the prior knowledge about θ can be denoted by a probability density function (PDF), $f(\theta)$. Let's denote $L(\theta|\mathbf{D})$ as a likelihood function, indicating the chance to observe \mathbf{D} given θ . One may refer to Juang et al. (2015) on how to construct the likelihood function using various types of geotechnical data. Based on Bayes' theorem, the prior knowledge about θ and the knowledge learned from the observed data can be combined as follows:

$$f(\theta|\mathbf{D}) = \frac{L(\theta|\mathbf{D})f(\theta)}{\int \dots \int L(\theta|\mathbf{D})f(\theta)d\theta} \quad (6-1)$$

where $f(\theta|\mathbf{D})$ = posterior PDF of θ representing the combined knowledge. In Eq. (6-1), both the likelihood function $L(\theta|\mathbf{D})$ and prior PDF $f(\theta)$ affect $f(\theta|\mathbf{D})$, and the role of the likelihood function will become more dominant as the amount of observed information increases.

Based on the law of total variance, it can be shown that the following inequality holds (Gelman et al. 2013):

$$\text{Var}(\theta) \geq E_{\theta}(\text{Var}(\theta|\mathbf{d})) \quad (6-2)$$

where $\text{Var}(\theta)$ denotes the prior variance of θ , $E_{\theta}(\text{Var}(\theta|\mathbf{D}))$ denotes the expected value of the posterior variance of θ . Eq. (6-2) indicates that the posterior variance is *on average smaller* than the prior variance, by an amount that depends on the variation in the posterior means over the distribution of

the possible data (Gelman et al. 2013). This inequality implies that if one consistently uses the Bayesian method, uncertainty reduction can be eventually achieved in the long run. However, there is no guarantee that the Bayesian method can always achieve uncertainty reduction in each individual application. The uncertainty in the posterior distribution may be larger than that in the prior distribution if the observed data significantly contradicts with the prior knowledge.

The posterior distribution $f(\boldsymbol{\theta}|\mathbf{D})$ in Eq. (6-1) is generally difficult to evaluate except for some special cases, such as the case where conjugate priors can be employed. In recent years, along with the advances in computational statistics, many methods have become practical for calculating $f(\boldsymbol{\theta}|\mathbf{D})$. Several commonly used computational methods will be described below.

6.3.1 Conjugate prior

If the posterior PDF $f(\boldsymbol{\theta}|\mathbf{D})$ is in the same family as the prior PDF $f(\boldsymbol{\theta})$, the prior and posterior are then called conjugate distributions. When the conjugate prior is adopted, the posterior distribution can be obtained analytically, and thus greatly simplifying the computational work involved. A comprehensive summary of conjugate priors can be found in the literature such as Ang and Tang (2007), Gelman et al. (2013), and Givens and Hoeting (2013).

6.3.2 Direct integration method

Based on the definition of the mean and the covariance matrix, the posterior statistics of $\boldsymbol{\theta}$ can be evaluated using the following equations:

$$\mu_{i|\mathbf{D}} = \int \theta_i f(\theta_i|\mathbf{D}) d\theta_i \quad (6-3)$$

$$\sigma_{i|\mathbf{D}}^2 = \int (\theta_i - \mu_{i|\mathbf{D}})^2 f(\theta_i|\mathbf{D}) d\theta_i \quad (6-4)$$

where $f(\theta_i|\mathbf{D})$ = posterior PDF of the i th element of $\boldsymbol{\theta}$, $\mu_{i|\mathbf{D}}$ = posterior mean of θ_i , and $\sigma_{i|\mathbf{D}}$ = posterior standard deviation of θ_i , respectively.

In principle, all existing methods such as Gaussian quadrature (e.g., Christian & Baecher 1999) for integration can be used to evaluate the above integrals. However, the computational work involved with the direction integration method may increase significantly with the dimension of $\boldsymbol{\theta}$. Thus, the direct integration method is often used for low dimensional problems.

6.3.3 Markov Chain Monte Carlo (MCMC) simulation

Albeit the apparent simplicity of Eq. (6-1), obtaining meaningful statistical information might require high dimensional integration which could be computationally challenging. A common approach to evaluate the posterior distribution is to use sampling methods such as Markov Chains Monte Carlo (MCMC) simulation (Brooks et al. 2011, Liu 2004, Robert and Casella 2004). The basic idea of MCMC simulation is to draw samples from a target distribution iteratively by means of a Markov chain that converges to the target distribution. When the Markov chain reaches its equilibrium state, the samples from the Markov chain are also those of the target distribution. These samples can then be used to infer the properties of the target distribution, and for subsequent geotechnical reliability analysis. In recent years, MCMC simulation is increasingly used in geotechnical engineering.

6.3.4 System identification (SI) method

In many cases, the parameters of a geotechnical model can hardly be determined accurately. The system identification (SI) method (e.g., Tarantola 2005), through embedding a deterministic model into the Bayesian formulation, is specifically developed for updating the parameters of a mechanical model based on the observed performance data. Let $g(\boldsymbol{\theta})$ be a geotechnical model with $\boldsymbol{\theta}$ denoting uncertain input parameters in this model. Let \mathbf{D} denote the observed performance data. Assume the prior knowledge about $\boldsymbol{\theta}$ can be denoted by a multivariate normal distribution with a mean of $\boldsymbol{\mu}_0$ and a covariance matrix of \mathbf{C}_0 . Assume that the observational uncertainty can be described by a multivariate

normal distribution with a mean vector of zero and a covariance matrix of \mathbf{C}_D . Assume further that the model uncertainty can be described by a multivariate normal distribution with a mean vector of zero and a covariance matrix of \mathbf{C}_m . With the above assumptions, $\boldsymbol{\theta}^*$ can be found by maximizing the posterior PDF, or equivalently, minimizing the following misfit function (Tarantola 2005)

$$2S(\boldsymbol{\theta}) = [\mathbf{g}(\boldsymbol{\theta}) - \mathbf{D}]^T \mathbf{C}_T^{-1} [\mathbf{g}(\boldsymbol{\theta}) - \mathbf{D}] + (\boldsymbol{\theta} - \boldsymbol{\mu}_0)^T \mathbf{C}_0^{-1} (\boldsymbol{\theta} - \boldsymbol{\mu}_0) \quad (6-5)$$

where $\mathbf{C}_T = \mathbf{C}_D + \mathbf{C}_m$.

Let $\boldsymbol{\theta}^*$ be the point where the misfit function is minimized. In the SI method, $f(\boldsymbol{\theta}|\mathbf{D})$ is approximated by a multivariate normal distribution with a mean of $\boldsymbol{\mu}_{\boldsymbol{\theta}|\mathbf{D}}$ and a covariance matrix of $\mathbf{C}_{\boldsymbol{\theta}|\mathbf{D}}$, where $\boldsymbol{\mu}_{\boldsymbol{\theta}|\mathbf{D}}$ and $\mathbf{C}_{\boldsymbol{\theta}|\mathbf{D}}$ are defined as (Tarantola 2005):

$$\boldsymbol{\mu}_{\boldsymbol{\theta}|\mathbf{D}} = \boldsymbol{\theta}^* \quad (6-6)$$

$$\mathbf{C}_{\boldsymbol{\theta}|\mathbf{D}} = (\mathbf{G}^T \mathbf{C}_T^{-1} \mathbf{G} + \mathbf{C}_0^{-1})^{-1} \quad (6-7)$$

$$\mathbf{G} = \frac{\partial \mathbf{g}(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}} \quad (6-8)$$

In Eq. (6-5), the effect of modeling uncertainty and observational uncertainty on the posterior distribution are represented by \mathbf{C}_0 and \mathbf{C}_m , and their overall effects are jointly represented by \mathbf{C}_T . If the model uncertainty is significantly larger than the observational uncertainty, \mathbf{C}_T will be dominated by \mathbf{C}_m , and in such a case the posterior distribution will be more affected by the model uncertainty. Similarly, if the observational uncertainty is significantly larger than the model uncertainty, the posterior distribution will be more affected by the observational uncertainty. On the other hand, the posterior distribution will be more affected by the prior distribution when the model and observational uncertainties increase. One can refer to Tarantola (2005) for a detailed discussion on the interplay between modeling, observational uncertainty, and prior information as well as their effect on the posterior distribution.

6.3.5 Other methods

In geotechnical engineering, several other techniques are also used for evaluating the posterior PDF, such as the extended Bayesian method (e.g., Honjo et al. 1994), the first order second moment Bayesian method (Gilbert 1999), and the importance sampling or the Latin Hypercube Sampling (Choi et al. 2006). Bayesian computational statistics is a field that is experiencing rapid progress. The interested readers may refer to the literature such as Gelman et al. (2013) for greater details about recent computational techniques for estimating posterior distributions. The Stochastic Finite Element Method (SFEM) includes a set of techniques which enable the propagation of parameter uncertainty through a deterministic Finite Element Method (FEM) (Ghanem & Spanos, 1991, Le Maître & Knio, 2010, Sudret 2008). As the SFEM and the Bayesian approach regard the model parameters as random variables, both can be used seamlessly. There has been some development in this direction (El Moselhy & Marzouk, 2012 and Cañavate et al. 2015).

In geotechnical engineering, Straub and Papaioannou (2015) illustrated how to perform Bayesian analysis for learning and updating geotechnical parameters and models with measurements. Juang and Zhang (2017) provided a tutorial on how the Bayesian methods can be formulated and used to solve different types of geotechnical problems. An in-depth discussion on how the Bayesian methods can be applied in different geotechnical problems can be found in Baecher (2017).

6.4 EXAMPLES

6.4.1. Back-analysis of soil properties from failed slopes

Slope stability problems are generally associated with uncertainties in the estimation of pore pressure regimes, soil/rock properties, and geometry of the failure surface. For failed slopes, Bayesian

probabilistic approaches allow for utilizing information on the location and geometry of the observed failure surface to update prior probability distributions of the shear strength parameters (ϕ' and c') and the pore pressure regime (r_u). The major strength of probabilistic back-analysis techniques for slopes is the recognition that there are numerous combinations of parameters that could result in the slope failure and the ability to quantify the relative likelihoods of these combinations. Zhang et al. (2010) presented an approach that uses a minimization procedure of a misfit function to refine/update the distributions of ϕ' , c' , and r_u . The main assumption is that all the parameters are normally distributed.

The example targets a case history involving the stability of a proposed highway in Algeria (Hasan and Najjar 2013), as shown in Figure 6-1. The prior mean values of ϕ' and c' were estimated as 21° and 15 kPa, respectively. By combining uncertainties due to spatial variability and statistical uncertainties, coefficients of variation (COVs) of 0.31 and 0.55 were determined for ϕ' and c' , respectively. The assumed prior pore water pressure regime was modelled with a mean r_u of 0.33 and an associated COV of 0.50. The high COV of 0.50 reflects the lack of site-specific piezometers.

Results of the updating process assuming statistically independent parameters indicate reductions in the updated mean ϕ' (from 21 to 17.3 degrees) and c' (from 15 to 12.1 kPa) and an increase in the mean r_u (from 0.33 to 0.42). These results are expected since the assumed prior mean values corresponded to a safety factor of ~ 1.5 . Reductions in ϕ' and c' along with an increase in r_u were required for failure conditions to prevail. Results also indicate that the standard deviations decreased for the updated case (5.5%, 27%, and 21.2% reductions in the standard deviations of c' , ϕ' and r_u). Finally, results indicated that although the prior parameters were assumed to be statistically independent, the updated parameters were found to be correlated with the following correlation coefficients: $\rho_{c,\phi} = -0.32$, $\rho_{\phi,r_u} = 0.78$, and $\rho_{c,r_u} = 0.29$.

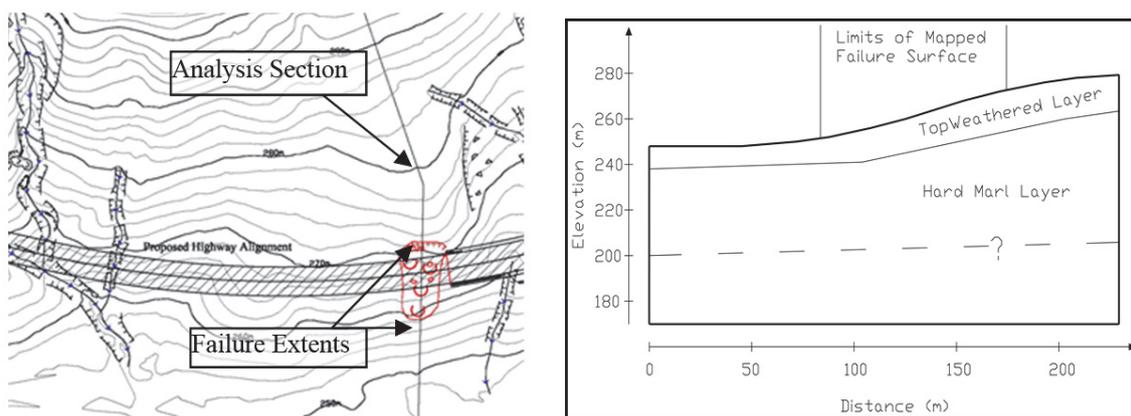


Figure 6-1 Plan and Section View for highway alignment and failure zone (Hasan and Najjar 2013).

To investigate the sensitivity of the results to the correlation between c' and ϕ' , the back analysis was repeated for $\rho_{c,\phi}$ values of -0.25, -0.5, -0.75, and -0.95. Results indicate that the assumption of negative correlation between c' and ϕ' results in an appreciable increase in the updated mean values of c' , ϕ' , and r_u with the effect being stronger as the correlation is assumed to be stronger (see Hasan and Najjar 2013). On the other hand, the updated standard deviations were found to be less sensitive to the assumed correlation coefficient.

Since uncertainties exist in the failure surface, it is important to determine whether that affects the results of the updating process. Results where the failure plane is varied indicate that the updated mean values of c' and ϕ' are sensitive to the assumed failure surface, indicating that accurate mapping of the failure surface is required for an accurate estimation of the soil properties.

6.4.2. Shallow foundation reliability based on spatially variable soil data

This example demonstrates the potential and benefits of using Bayesian analysis for incorporating information from spatially distributed data on soil into a reliability analysis. It is taken from (Papaioannou and Straub 2017). Exemplarily, the reliability of a centrally loaded rigid strip footing embedded in silty soil is evaluated. The bearing capacity of the foundation depends on the friction angle of the silty soil, which is a spatially variable property. To identify the friction angle, direct shear

tests of soil probes, taken at different depths in the area of the foundation are performed. The measurement outcomes are used to learn the spatial distribution of the soil property through application of Bayesian statistical analysis.

In reliability assessment, spatially variable properties are typically modeled by random variables with reduced variance to account for the spatial averaging effect. In this example, it is demonstrated that such an approach can be extended to a case in which spatial data are used to learn the distribution of spatially variable properties within a Bayesian context. This simplified random variable model is compared to a random field model that explicitly represents the spatial variability of the soil property, and provides the most accurate solution at the highest modeling and computational cost. One important detail: *in the random variable model*, the data are used to learn the statistics of the soil property and cannot be used to learn its posterior auto-covariance function. Therefore, spatial averaging is performed with the prior auto-correlation function.

Figure 6-2 illustrates the influence of the prior scale of fluctuation of the random field on the prior and posterior reliability estimates obtained with the RV (random variable) approach with spatial averaging and the RF (random field) approach. In the left panel of Figure 6-2, one can observe that the prior reliability estimates calculated with the two approaches agree well. However this is not the case for the posterior reliability estimates shown in the right panel of Figure 6-2. In the RF approach, the reliability increases fast when the scale of fluctuation becomes large. This is because the area of influence of the measurements increases as the prior spatial correlation increases. At low scales of fluctuation, the posterior statistics become almost uniform along the depth of the failure surface and the reliability increases again due to spatial averaging according to the prior correlation structure. The results obtained with the RV approach assume that the posterior correlation is the same as the prior independent of the prior scale of fluctuation. At low scales of fluctuation, this assumption is valid and the reliability estimates are close to the ones obtained with the RF approach. However, with increasing scale of fluctuation, the assumption that the spatial variability of the posterior is not influenced by the measurements is unrealistic and the reliability is significantly underestimated.

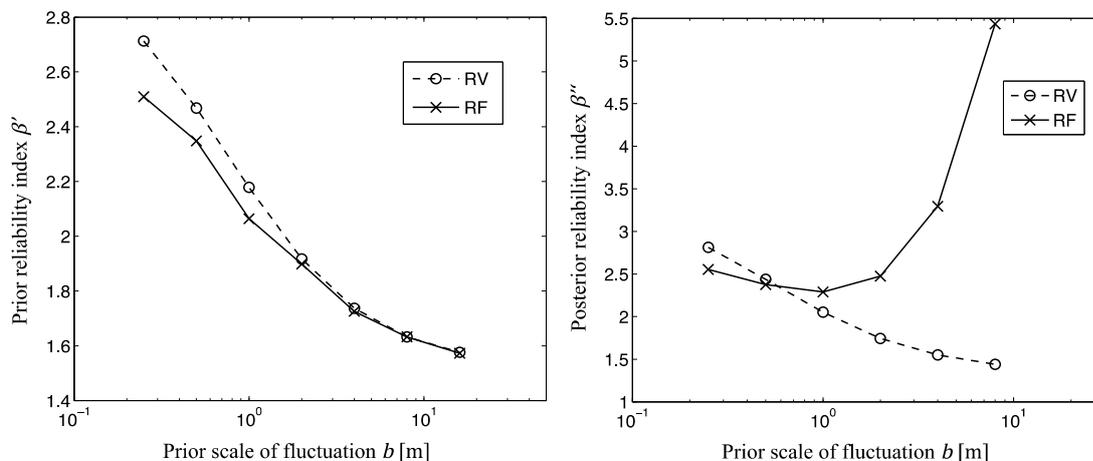


Figure 6-2 Prior and posterior reliability index vs. scale of fluctuation for the two modelling approaches: The random variable (RV) approach and random field (RF) approach.

6.4.3. Updating pile capacity at a site with load test results

In the last three decades, efforts have targeted analyzing the impact of pile load tests on the design of foundations in the framework of a reliability analysis. Examples include the work of Baecher and Rackwitz (1982), Zhang and Tang (2002), Zhang (2004), Najjar and Gilbert (2009), Kwak et al. (2010), Park et al. (2011, 2012), Zhang et al. (2014), Abdallah et al. (2015a, b), and Huang et al. (2016). In these studies, results of pile load tests are used to update the mean, median, lower bound, or actual capacity distributions of piles using Bayesian techniques. In this section, an example is presented to illustrate how the updating process is implemented. In the example, the uncertainty in the pile capacity is modeled by a lognormal distribution with (1) a coefficient of variation δ_r that represents the uncertainty due to spatial variability in pile capacity in a given site and (2) an uncertain

mean capacity that incorporates the model uncertainty of the pile capacity prediction method.

The mean of the pile capacity (r_{mean}) at a given site is generally a random variable. The mean and COV of r_{mean} are typically estimated from databases of pile load tests from the ratio of measured to predicted capacities, λ . As an example, Zhang (2004) reports COV values ranging from 0.21 to 0.57 for about 14 methods of pile capacity prediction. The COV of the pile capacity δ_r reflects the uncertainty due to within-site spatial variability (0.1 to 0.2) and is generally not updated in the analysis. In this example, the mean and COV of λ are assumed to be 1.30 and 0.50, respectively and δ_r is assumed to be equal to 0.2. Zhang (2004) presents the Bayesian formulation that allows for updating the pile capacity distribution for cases involving (1) pile load tests that are conducted to failure, and (2) proof load tests. The main difference is the form of the likelihood function.

For cases where piles carry the proof load, the mean value of the updated pile capacity increases with the number of tests while the COV value decreases. This is expected to result in increases in the reliability of the updated pile design at the site. For cases where some of the tested piles fail at carrying the proof load, the updated mean of the pile capacity decreases significantly as the number of positive tests decreases. For such cases the updated reliability could be lower than the prior.

Figure 6-3 shows the reliability index β for single piles designed with a factor of safety (FOS) of 2.0 and verified by several proof tests. The β value for the prior case is 1.49. If one positive proof test is conducted, then β will be updated to 2.23. The updated β will continue to increase if more positive proof tests are conducted. In the cases when not all tests are positive, the reliability of the piles will decrease with the number of tests that are not positive. The shaded zone in Figure 6-3 indicates that target reliability levels could be achieved with 1 to 3 positive tests with proof levels of twice the design load, provided that the piles act as part of a system. For a structure supported by four or fewer piles where the system effect may not be dependable, the target reliability should be slightly larger (ex. $\beta=3.0$) than those presented in Figure 6-3. Zhang (2004) shows that one or two successful proof load tests that are conducted at three times the design load are required to achieve a target reliability index of 3.0, as necessitated by non-redundant piles. Based on the above-mentioned methodology, a rational decision making framework (Najjar et al. 2016, 2017) could be envisaged to facilitate the choice of a load test program that has the maximum expected benefit to the project.

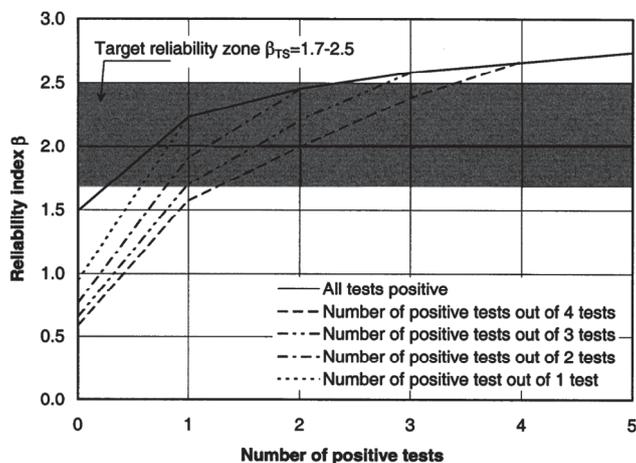


Figure 6-3 Variation of the reliability index with proof load test results (Zhang 2004).

6.4.4. Back-Analysis to determine the undrained strength of liquefied soil

Lateral spreading is the finite, lateral movement of gently to steeply sloping, saturated soil deposits caused by earthquake-induced liquefaction (Kramer 2016). A key parameter controlling the extent of lateral movements is the undrained strength of the liquefied ground. In this example we present a case history where a well-documented lateral spread that occurred during the M8.0 Pisco Earthquake is used along with Bayesian updating to estimate the undrained strength of the liquefied ground. The lateral spread occurred near the community of Canchamaná, on a marine terrace where pervasive liquefaction was observed (GEER 2007). Post-earthquake investigations as well as a comparison of pre- and post-earthquake satellite images were used to obtain a detailed displacement field over the entire 3km by 1 km where the lateral spread was observed (Cox et al. 2010). In addition, a

comprehensive field characterization study that included a series of Standard Penetration Tests (SPT) that covered the complete extent of the marine terrace was completed about 2.5 years after the earthquake. Details of this case history are presented in Gangrade et al. (2015).

The displacement field across the Canchamaná area was non-uniform, with displacements concentrated along certain cross sections. On the other hand, the SPT measurements had a more or less uniform distribution throughout the terrace. This is compatible with the distribution of liquefaction features over the area. Slope stability analysis indicated that the difference in displacements were likely due to differences in surface topography, which in turn imposed different driving forces on the liquefied soil. A key parameter in the stability analyses was the undrained strength of the liquefied soil. Current methodologies, such as Olson and Stark (2002), correlate this strength to corrected SPT N values:

$$\frac{s_u}{\sigma_v} = mN_{1,60} + c \quad (6-9)$$

where c and m are fitting parameters. The relationship applies for $SPT \leq 12$; a different slope is applied for $SPT > 12$ (Davies and Campanella 1994).

For the purpose of this study, we assume that m is given and we let c be a random variable. The objective of the Bayesian updating analysis is to determine the value of c in Equation (6-9) that is compatible with the observations of failure/no-failure at four cross-sections in the Canchamaná complex. We select a uniform prior for c within the bounds given by Olson and Stark (2002). The posterior distribution for c , given that some bins have failed and some have not failed, is given by:

$$f(c|F_i\bar{F}_j) = \frac{P(F_i|c)P(\bar{F}_j|c)}{P(F_i\bar{F}_j)} f(c) \quad (6-10)$$

where $f(c)$ is the prior distribution, F_i denotes *failure* of bin i , and \bar{F}_j denotes *no-failure* in bin j . The conditional failure probabilities are computed from pseudo-static slope stability analyses assuming a lognormal distribution for the input PGA. Posterior distributions for c were obtained for pair-wise combinations of a failed and no-failed bin. Modal values of c from these distributions are plotted in Figure 6-4 for different values of median and standard deviation of the PGA. Recorded values of PGA at nearby stations, as well as estimated intra-event standard deviation ($\sigma_{\ln PGA}$) from existing ground motion prediction equations can be used to obtain best-estimate values of the parameter c . In this example, the use of Bayesian updating allowed for a formal way of incorporating previous knowledge (e.g., the Olson and Stark model) with observations to improve the predictive ability of the model. The analyses could be improved through more rigorous determination of the prior distributions and a more formal inclusion of model uncertainties.

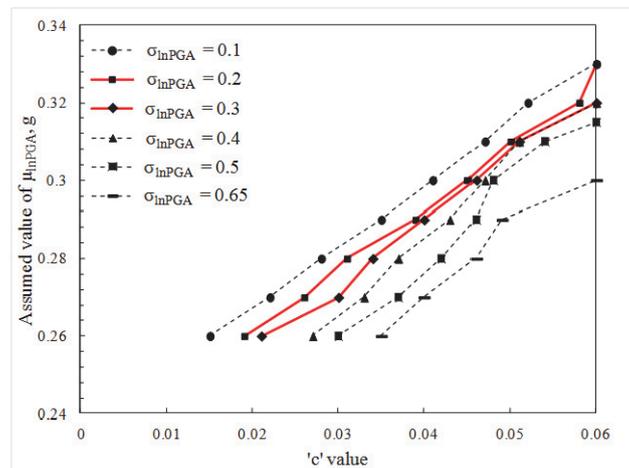


Figure 6-4 Correlation between ‘ c ’ values and the estimated mean PGA ($\mu_{\ln PGA}$).

6.4.5. Identification of underground soil stratification

Identification of underground soil stratification is an important aspect in geotechnical site characterization, even within a single type of soil. Consider, for example, London Clay Formation (LCF). It is well recognized that the LCF contains five depositional cycles (i.e. lithological units) and each cycle records an initial marine transgression followed by gradual shallowing of the sea. Because geotechnical properties (e.g. strength, stiffness and consolidation characteristics) of London Clay vary significantly in different lithological units or soil strata, it is of practical significance to identify the soil strata in London Clay so that the geotechnical property data within the same soil strata can be compared or used effectively.

Bayesian model class selection (BMCS) method has been developed to properly identify underground soil strata (Wang et al. 2014&2016b). In BMCS, a model class is referred to a family of stratification models that share the same number of soil strata but have different model parameters (Cao and Wang 2013, Wang et al. 2013). The number of soil strata is considered as a variable k , which is a positive integer varying from 1 to a maximum possible number N_{Lmax} . Therefore, there are N_{Lmax} candidate model classes M_k , $k=1, 2, \dots, N_{Lmax}$, and the k -th model class M_k has k soil strata. For a given set of site-specific observation data X_M , the plausibility of each model class is quantified by conditional probability $P(M_k|X_M)$, $k = 1, 2, \dots, N_{Lmax}$. Then, the most probable model class M_{k^*} is determined by comparing the conditional probabilities $P(M_k|X_M)$ of all N_{Lmax} candidate model classes and selecting the one with the maximum value of $P(M_k|X_M)$. The number of soil strata corresponding to M_{k^*} is taken as the most probable number k^* of soil strata.

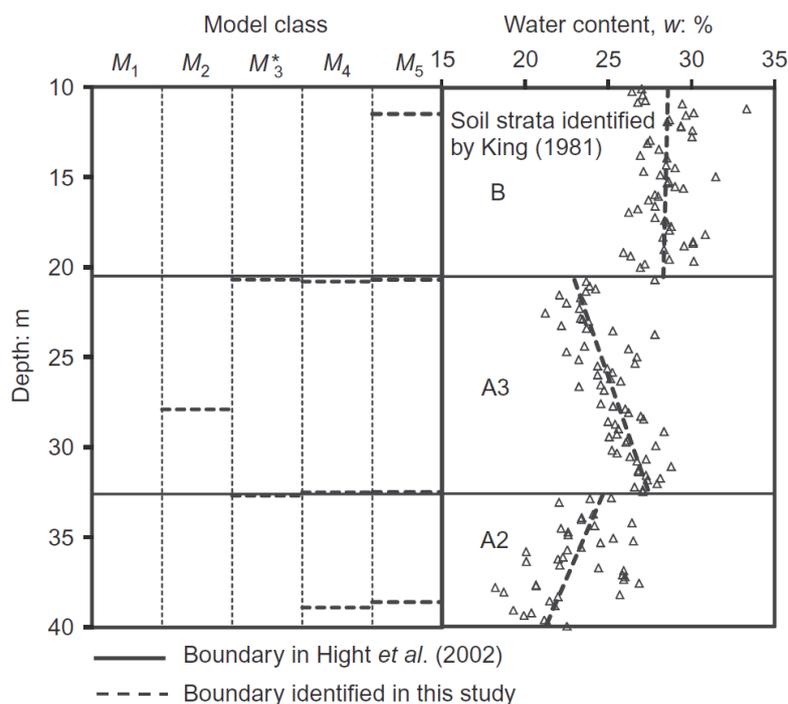


Figure 6-5 Identification of soil strata in London Clay (Wang et al. 2014).

As shown in Figure 6-5, BMCS method successfully identifies different lithological units of LCF using water content data (Wang et al. 2014). In addition, BMCS method can also be used together with CPT data to classify soil behavior types (SBT) and identify soil layers (Wang et al. 2013) or to identify statistically homogeneous soil layers and their associated spatial variability parameters in each statistically homogeneous soil layer (Cao and Wang 2013).

6.4.6 Other applications

Embankments. In one of the earliest applications of Bayesian analysis to geotechnical problems, Honjo et al. (1994) present the case study of an embankment on soft clay to illustrate the effectiveness

of what they call a “new type of indirect inverse analysis procedure”, which they call extended Bayesian method, EBM. The proposed method is based on a Bayesian model proposed by Akaike (1978) and combines objective information and subjective information. Calle et al. (2005) apply a Bayesian updating concept to develop a method to predict expected mean values and standard deviations of embankment settlement, as function of time. The method is based on both prior assumptions regarding expected means and standard deviations of settlement parameters and computation model uncertainty, as well as actually observed settlement behavior, e.g. during the construction stage. In Wu et al. (2007), the Bayesian method is used to update the material properties used in the prediction for a test embankment, and the updated properties are then used to update the prediction for the test embankment and to predict the performance of the full-scale embankment. Schweckendiek & Vrouwenvelder (2013) demonstrate how the Bayesian method can be used to reduce uncertainties in piping reliability assessment of an embankment. Huang et al. (2014) presents two examples where Bayesian statistical methods can be used for the prediction of future performance. The second example is to update embankment settlement predictions when field settlement monitoring data are available. More recently, Kelly & Huang (2015) present a proof of concept study to assess the potential for Bayesian updating to be combined with the observational method to allow timely and accurate decision-making during construction of embankments on soft soils.

Tunnels. Lee & Kim (1999) adopt EBM for a finite element analysis implemented to predict the ground response. In particular, they determine various geotechnical parameters of a FE implementation of an actual tunnel site in Pusan, Korea, including the elastic modulus, the initial horizontal stress coefficient at rest, the cohesion and the internal friction angle. Cho et al. (2006) combine EBM with a 3-dimensional finite element analysis to predict ground motion by using relative convergence as observation data. The proposed back-analysis technique is applied and validated by using the measured data from two tunnel sites in Korea. Camos et al. (2016) present a Bayesian method for updating the predicted tunneling-induced settlements when measurements are available. They also show how maximum allowable settlements, which are used as threshold values for monitoring of the construction process, can be determined based on reliability-based criteria in combination with measurements. The proposed methodology is applied to a group of masonry buildings affected by the construction of a metro line tunnel in Barcelona, Spain.

Piles. Goh et al. (2005) use a Bayesian neural network algorithm to model the relationship between the soil undrained shear strength, the effective overburden stress, and the undrained side resistance alpha factor for drilled shafts. The proposed approach provides information on the characteristic error of the prediction that arises from the uncertainty associated with interpolating noisy data. Kerstens (2006) deals with the prediction of the ultimate limit state (bearing capacity) of a single foundation pile. The proposed Bayesian statistical method, which combines information on pile capacity with the results of full scale tests, is applied to establish the probability of contending PDF's of the model uncertainty. Huang et al. (2014) present two examples where Bayesian statistical methods can be used for the prediction of future performance.

Deep excavations. The back analysis or inverse analysis of the field instrumentation data is a common technique to ascertain the design parameter validity in deep excavation projects. In Juang et al. (2013), a Bayesian framework using field observations for back-analysis and updating of soil parameters in a multistage braced excavation is presented. With the updated soil parameters, not only is the mean prediction improved, but also the variation of the prediction is reduced. In Wu et al. (2014), a novel method for updating the probability distribution of the maximum wall displacement at the i th excavation stage based on the measurements at earlier stages is proposed based on the concept of Bayesian updating. Canavate-Grimal et al. (2015) propose a Bayesian-type methodology to solve inverse problems which relies on the reduction of the numerical cost of the forward simulation through stochastic spectral surrogate models. The proposed methodology is validated with three calibration examples.

Soil Liquefaction. The soil liquefaction potential assessment is another field where the Bayesian method finds wide applications. Juang et al. (2000) suggest a Bayesian method to calibrate the liquefaction probability calculated from a reliability analysis. Cetin et al. (2002) develop a Bayesian method to calibrate a liquefaction potential assessment model based on the standard penetration test data. The method suggested by Cetin et al. (2002) was later used by Moss et al. (2006) and Boulanger & Idriss (2015) to develop liquefaction potential assessment models based on the cone penetration test data. Christian & Baecher (2015) discussed the merits of liquefaction potential assessment based on the Bayesian method. Juang et al. (2017) provide a comprehensive review on the application of probabilistic methods for soil liquefaction assessment.

6.5 POSSIBLE/SUGGESTED APPLICATION IN RELATION TO EUROCODES

The Eurocodes consist of ten European Standards that provide common approaches for the design of several civil engineering problems. Among the ten standards, the Eurocode 7: Geotechnical design (EN 1997) documents how to design geotechnical structures. This code covers a few important geotechnical design aspects including the basis of geotechnical design, geotechnical data, spread foundations, pile foundations, anchorages, retaining structures, hydraulic failure, overall stability and embankments. Eurocode 7 was approved by the European Committee for Standardization (CEN) in 2006 and has been mandatory in member countries since 2010. Based on the aforementioned review of Bayesian approaches, the following suggested applications are recommended in relation to Eurocodes:

- (1) **Estimation of ultimate resistance of pile foundation based on static load tests.** In Clause 7.6.2.2 of Eurocode 7, it provides guidelines on how to determine the ultimate compressive resistance based on results from static load tests. As illustrated in the pile foundation example (Zhang 2004), the estimation of the pile capacity can be potentially improved by combining information from regional database and the site-specific load test results. This application of Bayesian approaches (e.g., Zhang 2004; Najjar and Gilbert 2009) is recommended to improve Eurocode 7 for estimation of the ultimate resistance of pile foundations.
- (2) **Implementation of observational method.** In Clause 2.7 of Eurocode 7, it encourages the use of the observational method when prediction of geotechnical behavior is difficult. Nevertheless, how to formally implement it is not described in detail. The slope back analysis example (Hasan and Najjar 2013) and the liquefaction back analysis example (Gangrade et al. 2015) demonstrates how the Bayesian method can be used to reduce uncertainty associated with the properties of a geotechnical system based on the observed performance with explicit consideration of useful information of other sources. In this regard, the Bayesian method is recommended to implement the observational method as described in Eurocode 7. For example, the method suggested in Zhang et al. (2010) for back analysis of slope failure can be conveniently implemented and easily incorporated into Eurocode 7.
- (3) **Estimation of ground properties.** In Clause 2.4.3 of Eurocode 7, it provides guidelines on how to estimate the properties of soil and rock masses base on test results, either directly or through correlation, theory or empiricism, and from other relevant data. As illustrated in the shallow foundation example (Papaioannou and Straub 2017) and the soil stratification identification example (Wang et al 2014), the Bayesian method can be used to combine information from multiple sources for improved estimation of ground properties.
- (4) **Bayesian calibration of partial factors for consistent reliability level.** Eurocode 7 adopts partial factors in the limit state design. However, a consistent reliability level may not be achieved for a specific design due to local site variability. The Bayesian method may be used to calibrate of the partial factors to achieve consistent reliability levels (e.g., Ching et al. 2013; Juang et al. 2013).
- (5) **Use of Bayesian statistics and prior knowledge for selection of characteristic values for soil or rock properties.** In many cases, determination of the characteristic values of geotechnical parameters is a key step for application of Eurocode 7. In geotechnical engineering, a transformation model that relates the design soil parameter to the site investigation result (e.g., SPT N versus ϕ') is typically established by regional data or general data in the literature, which can serve as “prior” information for correlation behaviors among various soil parameters. Given the site-specific measurement (e.g., SPT N), one can adopt the Bayesian method to obtain a more accurate PDF of a soil parameter (e.g., ϕ') than that estimated directly based on transformation model, which can then be used to derive a point estimate as well as 95% confidence interval for the soil parameter. Details of the above technique are described in a separate report for the discussion group entitled “*Transformation models and multivariate soil databases*”.

On the other hand, Bayesian equivalent sample method has been developed to integrate prior knowledge with the often limited site-specific measurement data and transform the integrated knowledge into a large number of equivalent samples using MCMC (Wang and Cao 2013; Cao and Wang 2014; Wang and Aladejare 2015; Wang et al. 2016a). Excel-based user-friendly software has also been developed for the Bayesian equivalent sample method (Wang et al. 2016b). Details of the method, software and application examples are referred to

a separate report for the discussion group entitled “*Selection of characteristic values for rock and soil properties using Bayesian statistics and prior knowledge*”.

6.6 CHALLENGES AND LIMITATIONS

Despite the usefulness of the Bayesian methods as described above, the real application of the Bayesian method in geotechnical engineering is quite limited. Several possible causes are identified for such a dilemma, which may provide future directions for better use of the Bayesian methods.

First, most engineers are not fully aware of the benefits of Bayesian methods. As a result, there is a lack of willingness for common engineers to use the Bayesian method in practice. On the other hand, the research on the application of Bayesian methods in geotechnical engineering is quite active. To narrow such a gap, researchers in Bayesian geotechnics are encouraged to outreach to the industry to improve the communication with the practicing profession.

Second, most geotechnical engineers do not have special training in Bayesian statistics, which indeed requires advanced statistical concepts and in some cases knowledge of programming. This challenge may be tackled from two directions. Since reliability-based design courses have been incorporated into the curriculum of civil engineering programs in many institutions, instructors in these courses may incorporate Bayesian statistics as a main topic that is covered in the course. Second, TC304 may also consider organizing short courses on Bayesian statistics at different occasions to help interested geotechnical engineers to develop knowledge and capability to solve geotechnical problems using the Bayesian method. The Bayesian short course offered at the Georisk2017 conference is a good example of such efforts.

Third, in many occasions, the application of Bayesian methods involves quite some computational effort wherein specialized software may be needed. Currently, however, very few geotechnical engineering software has such capability. Researchers in the Bayesian geotechnics may consider developing easy-to-use procedures for implementing the Bayesian method for geotechnical applications. For instance, The Solver in Excel is a powerful and convenient optimization tool which may significantly facilitate the application of Bayesian methods in geotechnical engineering.

Last but not least, how to specify the prior distribution could be challenging. Research can be conducted to recommend rational/practical practices on how to derive the prior information in a more objective and more defensive way. For instance, the regional experience is often one important source of information for deriving the prior distribution. Recent studies have shown that the prior information can be derived quantitatively through calibration of the global database (Ching and Phoon 2014), the Bayesian equivalent sample method (Wang et al 2016b), or through the multi-level Bayesian modeling approach (Zhang et al 2016). The calibration of prior information may significantly facilitate the application of Bayesian methods in geotechnical engineering.

6.7 CONCLUSIONS

In this report, a systematic investigation has been conducted on the usefulness of the Bayesian method in geotechnical engineering. The investigation showed that the Bayesian method may have a wide range of applications whenever information combination is needed. Nevertheless, a knowledge gap is observed to exist between the academic research and practical application. Recommendations are made to leverage the power of the Bayesian method into practice.

6.8 REFERENCES

- Abdallah, Y., Najjar, S. S., and Saad, G. (2005). Reliability-based design of proof load test programs for foundations. *Geotechnical Engineering*, 46(2), 63-80.
- Abdallah, Y., Najjar, S. S., and Saad, G. (2015). Impact of proof load test programs on the reliability of foundations. *Proceedings of International Foundations Exposition and Equipment Exposition*, San Antonio, Texas, March 17-21, 2015.
- Akaike, H. (1978). On the likelihood of a time series model. *The Statistician*, 27, 217-235
- Ang, A.H.-S. AND Tang, W.H. (2007). *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*, Vol. 1, 2nd Ed., Wiley, New York, NY.
- Baecher, G. B., and Rackwitz, R. (1982). Factors of safety and pile load tests. *International Journal of*

- Numerical and Analytical Methods in Geomechanics, 6(4), 409–424.
- Baecher, G.B. (2017). Bayesian thinking in geotechnical applications. Proceedings of Georisk 2017 on Geotechnical Risk from Theory to Practice, June 4-7 2017, Denver, Colorado, USA.
- Boulanger, R., and Idriss, I. (2015). CPT-based liquefaction triggering procedure. *Journal of Geotechnical and Geoenvironmental Engineering*, 10.1061/(ASCE)GT.1943-5606.0001388, 04015065.
- Brooks, S., Gelman, A., Jones, G., and Meng, X.-L. (2011). *Handbook of Markov Chain Monte Carlo*, Chapman and Hall.
- Cao, Z. J. and Wang, Y. (2013). Bayesian approach for probabilistic site characterization using cone penetration tests. *Journal of Geotechnical & Geoenvironmental Engineering*, 139(2), 267-276.
- Cao, Z. J. and Wang, Y. (2014). Bayesian model comparison and characterization of undrained shear strength. *Journal of Geotechnical & Geoenvironmental Engineering*, 140(6), 04014018, 1-9.
- Calle, E., Sellmeijer, H., and Visschedijk, M. (2005). Reliability of settlement prediction based on monitoring. Proceedings of the 16th International Conference on Soil Mechanics and Geotechnical Engineering: Geotechnology in Harmony with the Global Environment, 3, 1681-1684.
- Camós, C., Špačková, O., Straub, D., & Molins, C. (2016). Probabilistic approach to assessing and monitoring settlements caused by tunneling. *Tunnelling and Underground Space Technology*, 51, 313-325.
- Cañavate-Grimal, A., Falcó, A., Calderón, P., and Payá-Zaforteza, I. (2015). On the use of stochastic spectral methods in deep excavation inverse problems. *Computers and Structures*, 159, 41-60.
- Cetin, K. O., Der Kiureghian, A., and Seed, R. B. (2002). Probabilistic models for the initiation of seismic soil liquefaction. *Structural Safety*, 24, 67-82.
- Ching, J., Phoon, K., Chen, J., and Park, J. (2013). Robustness of constant load and resistance factor design factors for drilled shafts in multiple strata. *Journal of Geotechnical and Geoenvironmental Engineering*, 139(7), 1104-1114.
- Ching, J. and Phoon, K.K. (2014). Transformations and correlations among some clay parameters – the global database, *Canadian Geotechnical Journal*, 51(6), 663-685.
- Cho, K. H., Choi, M. K., Nam, S. W., and Lee, I. M. (2006). Geotechnical parameter estimation in tunnelling using relative convergence measurement. *International Journal for Numerical and Analytical Methods in Geomechanics*, 30(2), 137-155.
- Christian, J. and Baecher, G. (1999). Point-estimate method as numerical quadrature. *Journal of Geotechnical and Geoenvironmental Engineering*, 125(9), 779-786.
- Christian, J. H., and Baecher, B. G. (2015). Bayesian methods and liquefaction. Proceedings of the 12th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP12), The University of British Columbia, Vancouver, Canada.
- Choi, S.K., Grandhi, R.V., and Canfield, R.A. (2006). *Reliability-based structural design*. Springer, Berlin.
- Cox, B., Cothren, J., Barnes, J., Wartman, J., Meneses, J., and Rodriguez-Marek, A. (2010) Towards quantifying movement of a massive lateral spread using high-resolution satellite image processing. Proceedings, 9th U.S. National and 10th Canadian Conference on Earthquake Engineering, July 25-29, Toronto, Canada.
- Davies, M.P., and Campanella, R.G. 1994. Selecting design values of undrained strength for cohesionless soils, Proceedings of the 47th Canadian Geotechnical Conference, 176–186, Halifax, Nova Scotia, Canada, 21-23 September 1994.
- El Moselhy, T.A. and Marzouk, Y.M. (2012). Bayesian inference with optimal maps. *Journal of Computational Physics*, 231, 7815-50.
- Gangrade, R.M., Rodriguez-Marek, A., Cox, B.R. and Wartman, J. (2015). Analysis of lateral spreading case histories from the 2007 Pisco, Peru Earthquake. Proceedings of the 5th International Symposium on Geotechnical Safety and Risk (ISGSR), Rotterdam, The Netherlands, 13-16 October.
- GEER Association Report No. GEER-012 (2007). Preliminary reconnaissance report on the geotechnical engineering aspects of the August 15, 2007 Pisco, Peru Earthquake. Report of the National Science Foundation-sponsored Geotechnical Earthquake Engineering Reconnaissance (GEER) team. http://www.geerassociation.org/GEER_Post%20EQ
- Gelman, B.A., Carlin, B.P., Stern, H.S. and Rubin, D.B. (2013). *Bayesian Data Analysis*, 3rd Ed., Chapman & Hall, London, United Kingdom.
- Ghanem, R.G. and Spanos, P.D. (1991). *Stochastic finite elements: a spectral approach*. New York:

- Springer.
- Gilbert, R.B. (1999). First-order, Second-moment Bayesian Method for Data Analysis in Decision Making, Geotechnical Engineering Center, Dept. of Civil Engineering, Univ. of Texas at Austin, Austin, TX.
- Givens, G.H. and Hoeting, J.A. (2013). Computational Statistics, 2nd Ed., John Wiley & Sons, New York, NY.
- Goh, A. T. C., Kulhawy, F. H., & Chua, C. G. (2005). Bayesian neural network analysis of undrained side resistance of drilled shafts. *Journal of Geotechnical and Geoenvironmental Engineering*, 131(1), 84-93.
- Hasan, S. and Najjar, S.S. (2013). Probabilistic back analysis of failed slopes using Bayesian techniques. Proceedings, GeoCongress, Stability and Performance of Slopes and Embankments III, ASCE, 2013, San Diego, USA
- Honjo, Y., Liu Wen-Tsung, and Guha, S. (1994). Inverse analysis of an embankment on soft clay by extended Bayesian method. *International Journal for Numerical & Analytical Methods in Geomechanics*, 18(10), 709-734.
- Huang, J., Kelly, R., Li, D., Zhou, C. and Sloan, S. (2016). Updating reliability of single Piles and pile groups by load tests," *Computers and Geotechnics*, 73, 221-230.
- Huang, J., Kelly, R., Li, L., Cassidy, M., and Sloan, S. (2014). Use of Bayesian statistics with the observational method. *Australian Geomechanics Journal*, 49(4), 191-198.
- Juang, C. H., Chen, C. J., Rosowsky, D. V., and Tang, W. H. (2000). CPT-based liquefaction analysis, Part 2: Reliability for design. *Géotechnique*, 50 (5), 593-599.
- Juang, C.H., Khoshnevisan, S. and Zhang, J. (2015). Maximum likelihood principle and its application in soil liquefaction assessment. In *Risk and Reliability in Geotechnical Engineering*, Ed. K. K. Phoon & J. Ching, Taylor & Francis, London, United Kingdom.
- Juang, C.H., Luo, Z., Atamturktur, s., and Huang, H.W. (2013). Bayesian updating of soil parameters for braced excavations using field observations. *Journal of Geotechnical and Geoenvironmental Engineering*, 139(3), 395-406.
- Juang, C.H., Wang, L., Liu, Z., Ravichandran, N., Huang, H., and Zhang, J. (2013). Robust geotechnical design of drilled shafts in sand: new design perspective. *Journal of Geotechnical and Geoenvironmental Engineering*, 139(12), 2007-2019.
- Juang, C.H., and Zhang, J. (2017). Bayesian Methods for Geotechnical Applications - A Practical Guide. In *Geotechnical Safety and Reliability: Honoring Wilson H. Tang (GSP 286)*, ASCE, Reston.
- Juang, C.H., Zhang, J., Khoshnevisan, S., and Gong, W. (2017). Probabilistic methods for assessing soil liquefaction potential and effect. In *Geo-Risk 2017: Keynote Lectures (GSP 282)*, ASCE, Reston.
- Kelly, R. and Huang, J. (2015). Bayesian updating for one-dimensional consolidation measurements. *Canadian Geotechnical Journal*, 52(9), 1318-1330.
- Kerstens, J. G. M. (2006). A Bayesian way to solve the pdf selection problem: An application in geotechnical analysis. *Heron*, 51(4), 233-250.
- Kwak, K., Kim, K. J., Huh, J., Lee J. H., and Park, J. H. (2010). Reliability-based calibration of resistance factors for static bearing capacity of driven steel pipe piles. *Canadian Geotechnical Journal*, 47(5), 528-538.
- Kramer (2016). Lateral Spreading. In *Encyclopedia of Natural Hazards*, p. 623. DOI: 10.1007/978-1-4020-4399-4_215.
- Lee, I. M. and Kim, D. H. (1999). Parameter estimation using extended Bayesian method in tunnelling. *Computers and Geotechnics*, 24(2), 109-124.
- Le Maître, O.P. and Knio, O.M. (2010). Spectral methods for uncertainty quantification with applications to computational fluid dynamics. New York: Springer.
- Liu, J.S. (2004). Monte Carlo Strategies in Scientific Computing, Springer, New York, USA.
- Moss, R. E. S., Seed, R. B., Kayen, R. E., Stewart, J. P., Der Kiureghian, A., and Cetin, K. O. (2006). CPT-based probabilistic and deterministic assessment of in situ seismic soil liquefaction potential. *Journal of Geotechnical and Geoenvironmental Engineering*, 132(8), 1032-1051.
- Papaoannou, I. and Straub, D. (2017) Learning soil parameters and updating geotechnical reliability estimates under spatial variability – theory and application to shallow foundations, *Georisk*, 11(1), 116-128.
- Najjar, S. S. and Gilbert, R. B. (2009a). Importance of proof-load tests in foundation reliability. Geotechnical Special Publication No. 186, Proceedings of IFCEE 2009, Contemporary Topics in

- In-Situ Testing, Analysis, and Reliability of Foundations, 2009a, ASCE, Orlando, Florida, 340-347.
- Najjar, S.S., Saad, G., and Abdallah, Y. (2016). Sustainable design of pile load test programs using reliability-based principles. Proceedings of the 7th Civil Engineering Conference in the Asian Region, CECAR7, Aug. 30 - Sept. 2 2016, Waikiki, Honolulu, USA.
- Najjar, S.S., Saad, G., and Abdallah, Y. (2017). Rational decision framework for designing pile load test programs. *Geotechnical Testing Journal*, 40(2), 302-316.
- Olson, S.M. and Stark, T.D. (2002). Liquefied strength ratio from liquefaction flow failure case histories. *Canadian Geotechnical Journal*, 39(3), 629-647.
- Park, J. H., Kim, D., and Chung, C. K. (2011). Reliability index update for driven piles based on Bayesian theory using pile load tests results. *International Journal of Offshore and Polar Engineering*, 21(4), 330–336.
- Park, J. H., Kim, D., and Chung, C.K. (2012). Implementation of Bayesian theory on LRFD of axially loaded driven piles. *Computers and Geotechnics*, 42, 73–80.
- Robert, C. and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer, New York, USA.
- Schweckendiek, T., and Vrouwenvelder, A.C.W.M. (2013). Reliability updating and decision analysis for head monitoring of levees, *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 7(2), 110-121.
- Straub, D., and Papaioannou, I. (2015). Bayesian analysis for learning and updating geotechnical parameters and models with measurements. In *Risk and Reliability in Geotechnical Engineering* (KK Phoon and J Ching, editors), CRC Press, 221–264.
- Sudret B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety*, 93(7), 964–79.
- Tarantola A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia: SIAM.
- Wang, Y. and Aladejare, A. E. (2015). Selection of site-specific regression model for characterization of uniaxial compressive strength of rock. *International Journal of Rock Mechanics & Mining Sciences*, 75, 73-81.
- Wang, Y., Akeju, O. V., and Cao, Z. (2016b). Bayesian Equivalent Sample Toolkit (BEST): an Excel VBA program for probabilistic characterisation of geotechnical properties from limited observation data. *Georisk*, 10(4), 251-268.
- Wang, Y., Cao, Z. and Li, D. (2016a). Bayesian perspective on geotechnical variability and site characterization. *Engineering Geology*, 203, 117-125.
- Wang, Y., Huang, K. and Cao, Z. (2013). Probabilistic identification of underground soil stratification using cone penetration tests. *Canadian Geotechnical Journal*, 50(7), 766-776.
- Wang, Y., Huang, K. and Cao, Z. (2014). Bayesian identification of soil strata in London Clay. *Géotechnique*, 64(3), 239-246.
- Whitman, R.V. (1996). Organizing and evaluating uncertainty in geotechnical engineering. *Uncertainty in the Geologic Environment: from Theory to Practice, Proceedings of Uncertainty 96*, New York, Vol.1, 1-28.
- Wu, T.H., Zhou, S.Z., and Gale, S.M. (2007). Embankment on sludge: predicted and observed performances. *Canadian Geotechnical Journal*, 44, 545-563.
- Wu, S.H., Ching, J., Ou, C.Y. (2014). Probabilistic observational method for estimating wall displacements in excavations. *Canadian Geotechnical Journal*, 2014, 51(10): 1111-1122.
- Zhang, J., Li, J.P., Zhang, L.M., Huang, H.W. (2014). Calibrating cross-site variability for reliability-based design of pile foundations. *Computers and Geotechnics*, 62, 154-163.
- Zhang, J., Tang, W., and Zhang, L.M. (2010). Efficient probabilistic back-analysis of slope stability model parameters. *Journal of Geotechnical and Geoenvironmental Engineering*, 136(1), 99-109.
- Zhang, L. (2004). Reliability verifications using proof pile load tests. *Journal of Geotechnical and Geoenvironmental Engineering*, 130(11), 1203-1213.
- Zhang, L. and Tang, W. (2002). Use of load tests for reducing pile length. Proceedings of the International Deep Foundations Congress, February 14-16, 2002, 993-1005.
- Zhang, J., Juang, C.H., Martin, J.R., and Huang, H.W. (2016). Inter-region variability of Robertson and Wride method for liquefaction hazard analysis, *Engineering Geology*, 203, 191-203.

Chapter 7 Incorporating Spatial Variability into Geotechnical Reliability-based Design

Lead discussor:

Dian-Qing Li

dianqing@whu.edu.cn

Discussors (alphabetical order):

Zi-Jun Cao, Jianye Ching, Satyanarayana Murty Dasaka, Jinsong Huang, Mark Jaksa,
Shin-ichi Nishimura, Armin Stuedlein, Giovanna Vessia

7.1 INTRODUCTION

Geotechnical materials are natural materials. Their properties are affected by various factors in natural geological processes, such as parent materials, weathering and erosion processes, transportation agents, conditions of sedimentation, and other processes (e.g., Mitchell and Soga, 2005). These factors vary spatially from one location to another, which subsequently leads to inherent spatial variability (ISV) of geotechnical properties (Vanmarcke, 2010). ISV has been considered as a major source of uncertainties in soil properties (e.g., Christian et al., 1994; Kulhawy, 1996; Phoon and Kulhawy, 1999a; Baecher and Christian, 2003; Stuedlein et al. 2012a; Wang et al., 2016). It significantly affects the safety (measured by factor of safety, FS) and reliability (measured by probability of failure, P_f , or reliability index, β) of geotechnical structures, such as foundations (e.g., Fenton and Griffiths, 2002, 2003 and 2007; Stuedlein et al. 2012b; Wang and Cao, 2013; Li et al., 2015a; Stuedlein 2017), retaining structures (e.g., Fenton and Griffiths, 2005), and slopes (e.g., Griffith and Fenton, 2004, 2009; Huang et al., 2010; Wang et al., 2011; Li et al., 2014; Jiang et al., 2014; Li et al., 2016a; Xiao et al., 2016). ISV shall, hence, be rationally taken into account in geotechnical designs, which constitutes a major difference in reliability-based designs (RBD) of geotechnical structures and building structures.

ISV can be explicitly modeled in geotechnical RBD using random field theory (Vanmarcke, 2010). Figure 7-1 shows major steps for incorporating ISV in geotechnical RBD based on random field theory. In general, it starts with probabilistic *characterization* of ISV based on site investigation data (e.g., in-situ/laboratory test results) and site information available prior to the project (namely prior knowledge), which determines statistical information of geotechnical design parameters, including spatial trend, statistics (e.g., mean and standard deviation), and correlation functions. Such information is needed as input for *modeling* ISV in geotechnical RBD, which represents (or simulates) ISV of geotechnical design parameters using random fields with pre-defined statistical information. Here, the authors need to emphasize that Sections 7.2-7.5 of this report focus on *modeling* ISV in geotechnical RBD based on known/assumed statistical information of geotechnical parameters. The probabilistic *characterization* of ISV, i.e., how to derive statistical information from site investigation data, will be briefly discussed in Section 7.6. Relevant studies on probabilistic characterization of ISV of geotechnical parameters are referred to DeGroot and Baecher (1993), Jaksa (1995), Fenton et al. (1999a, b), Uzielli et al. (2005), Wang et al., (2010), Dasaka and Zhang (2012), Stuedlein (2011), Stuedlein et al. (2012a), Cao and Wang (2013, 2014), Firouziandbandpey et al. (2014, 2015), Ching et al., (2015), Cao et al., (2016), Wang et al. (2016), and Tian et al. (2016).

Based on random field theory, there are generally two ways to model ISV in geotechnical RBD, which are named as rigorous (R)-method and approximate(A)-method in this report. As shown in Figure 7-1, R-method directly simulates random fields of geotechnical design parameters based on their statistical information derived from site investigation without considering influence zones and/or critical slip surfaces that affect responses (e.g., resistance moment, bearing capacity, and settlement, etc.) of geotechnical structures concerned. With the R-method, a number of random field realizations of geotechnical parameters can be directly generated under a full-probabilistic RBD framework using

Monte Carlo simulation (MCS) (e.g., Cao et al., 2013; Wang and Cao, 2013; Cao and Wang, 2014; Li et al., 2016b). Then, each random field realization is used as input of deterministic geotechnical model to predict responses of geotechnical structures concerned in design. By this means, ISV modeling is deliberately decoupled from deterministic analyses of geotechnical structures. This provides flexibility in choosing different deterministic geotechnical models (for example, both limit equilibrium methods and finite element methods can be used as deterministic models to evaluate the safety margin of slope stability for each random field realization) and allows searching for critical slip surfaces and determining influences zones of soil masses affecting responses of geotechnical structures. For example, R-method is applied in random finite element method to model ISV for geotechnical probabilistic analysis and risk assessment (e.g., Griffith and Fenton, 2004, 2009; Huang et al., 2010, 2013; Stuedlein et al. 2012b; Li et al., 2016a, Xiao et al., 2016), in which ISV modeling does not involve information on geotechnical failure mechanisms or influence zones.

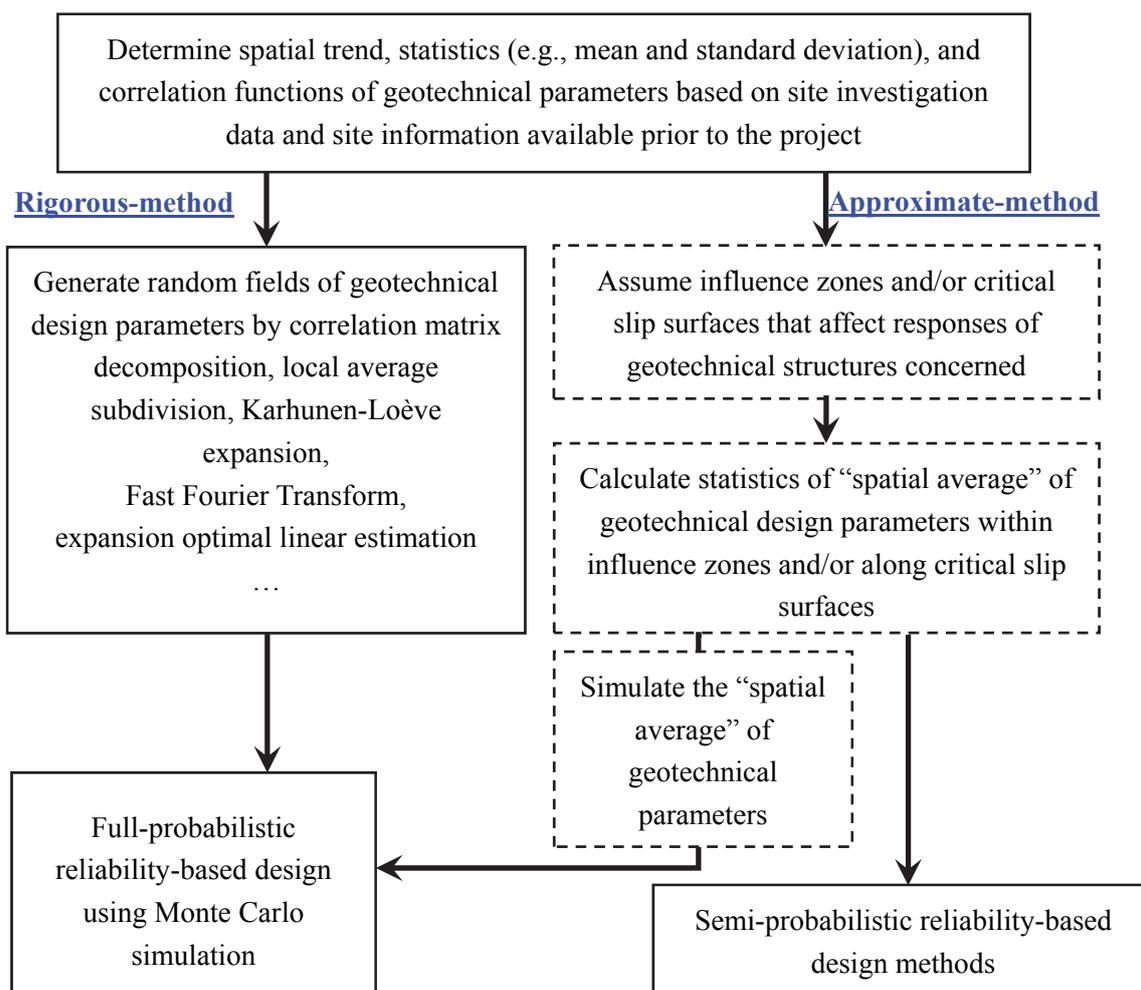


Figure 7-1 Incorporating spatial variability into geotechnical RBD based on random field theory.

In contrast, A-method uses random field theory to calculate statistics of spatial averages of geotechnical design parameters within influence zones and/or along critical slip surfaces affecting responses of geotechnical structures (Vanmarcke, 1977; El-Ramly et al., 2002, 2005; Stuedlein et al. 2012b; Zhang and Chen, 2012; Wang and Cao, 2013). The spatial average of geotechnical design parameters over a spatial curve (e.g., slip surfaces) or area (e.g., influence zones) has the same mean as geotechnical design parameters at a “point” that is directly simulated in R-method, but its variance is reduced due to spatial averaging. The extent of reduction in variance is quantified by variance reduction function Γ^2 that is defined as a ratio of the variance of the spatial average over the variance of geotechnical design parameters at a “point” (Vanmarcke, 2010). Calculating Γ^2 requires geometric information (e.g., location and length) of influence zones and/or critical slip surfaces. Such information is, however, unknown prior to geotechnical deterministic analyses and shall be assumed

for calculating statistics of spatial averages of geotechnical design parameters in A-method. Hence, using A-method in geotechnical RBD, ISV modeling is coupled with geotechnical deterministic analyses. After statistics of spatial averages of geotechnical design parameters are calculated, they can either be used to determine characteristic values of geotechnical design parameters for semi-probabilistic RBD approaches (e.g., Orr, 2015) or be applied to simulating spatial averages within pre-defined influence zones and/or along assumed critical failure surfaces under a full-probabilistic RBD framework (El-Ramly, 2002, 2006; Wang and Cao, 2013).

A-method uses spatial averages of geotechnical design parameters as their corresponding estimates along critical slip surfaces or within influence zones. However, the spatial average might not be the same as “mobilized” values of geotechnical parameters that control responses of geotechnical structures concerned in designs, which has been demonstrated under simple stress states (e.g., Ching and Phoon, 2012, 2013; Ching et al., 2014). Hence, compared with direct and rigorous modeling of ISV using R-method, A-method is an indirect and approximate way to model ISV in geotechnical RBD. How such an indirect and approximate modeling of ISV affects RBD of real geotechnical structures is unclear. This issue is systematically explored for different geotechnical structures, including drilled shaft, sheet pile wall, and soil slope, in this report. In addition, Γ^2 is needed for implementing A-method. It can be calculated exactly according to correlation functions of geotechnical design parameters, or be evaluated approximately by a simplified formulation to bypass the need of determining correlation functions. This report will also discuss effects of different ways to calculate Γ^2 on geotechnical RBD.

7.2 PROBABILISTIC MODELING OF SPATIAL VARIABILITY IN GEOTECHNICAL RBD

As shown in Figure 7-1, both R-method and A-method can be applied to modeling ISV in full-probabilistic RBD approach using MCS. To enable a consistent comparison, this report adopts a recently developed MCS-based RBD approach, so-called the expanded RBD approach, to perform RBD of geotechnical structures with R-method and A-method, respectively. The expanded RBD approach formulates the design process as a systematic sensitivity analysis on possible designs in design space (e.g., a possible range of drilled shaft length) defined by geotechnical engineers, in which P_f values of all the possible designs are calculated by a single run of MCS. Then, the final design is determined according to target reliability levels and economic requirement. Details of algorithms and implementation procedures of the expanded RBD approach are referred to Wang (2011), Wang et al. (2011), Wang and Cao (2013, 2015), and Li et al. (2016b).

The expanded RBD approach provides flexibility of modeling ISV in different ways for geotechnical RBD, such as R-method and A-method. Based on the expanded RBD approach, this report aims to preliminarily reveal effects of indirect and approximate ISV modeling through A-method on geotechnical RBD and probabilistic analysis by comparing respective RBD results and/or reliability estimates (e.g., P_f) that are obtained using R-method and A-method, and to demonstrate effects of different ways to calculate Γ^2 . For the illustration and simplification, only one-dimensional (1-D) ISV of geotechnical parameters is considered in this report. The following two subsections describe 1-D ISV modeling using R-method and A-method, respectively.

7.2.1 Rigorous modeling by random field simulation (R-method)

R-method models ISV of geotechnical parameters in a direct and explicit manner. Consider, for example, a geotechnical design parameter X (e.g., effective friction angle ϕ') in a statistically homogenous soil layer. As shown in Figure 7-2, the ISV of X with depth (or in some direction) can be characterized by a 1-D homogenous lognormal random field $\underline{X}(z_i)$, in which z_i is a spatial coordinate (e.g., depth in the vertical direction) of the i -th location and X is a lognormal random variable with a mean μ and standard deviation σ (or coefficient of variation $COV_X = \sigma/\mu$). In the context of random fields, the spatial correlation between variations of X at different locations is characterized by the scale of fluctuation and correlation function (Vanmarcke, 1977 and 2010). Here, the correlation function is taken as a single exponential correlation function, and the correlation coefficient ρ_{ij} between the logarithms [e.g., $\ln X(z_i)$ and $\ln X(z_j)$] of X at i -th and j -th locations is given by:

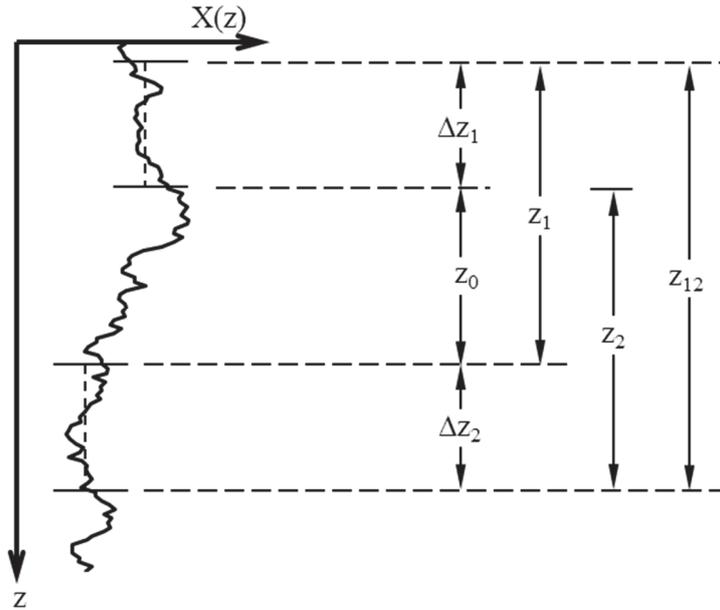


Figure 7-2 Illustration of 1-D spatial variability.

$$\rho_{ij} = \exp\left(-2|\Delta D_{i,j}|/\lambda\right) \quad (7-1)$$

where λ = scale of fluctuation; $|\Delta D_{i,j}|$ = the distance between i -th and j -th locations. For a given set of statistics (including μ , σ , and λ) and correlation function, the statistically homogenous random field \underline{X} of X can be generated using various simulation techniques, such as correlation matrix decomposition (e.g., Wang et al., 2011; Li et al., 2015b, 2016a), local average subdivision (e.g., Fenton and Vanmarcke 1990; Fenton and Griffiths, 2008), Karhunen-Loève expansion (e.g., Phoon et al. 2002; Cho 2010; Jiang et al. 2015), expansion optimal linear estimation (e.g., Li and Der Kiureghian, 1993; Xiao et al., 2016). Consider, for example, using the covariance matrix decomposition method to simulate \underline{X} in this report, by which X can be written as:

$$\ln X = \mu_{\ln X} \mathbf{I} + \sigma_{\ln X} \underline{L}^T \underline{\varepsilon} \quad (7-2)$$

where $\mu_{\ln X} = \ln \mu_X - (\sigma_{\ln X})^2/2$ and $\sigma_{\ln X} = \{\ln[1 + (\sigma_X/\mu_X)^2]\}^{0.5}$ are the mean and standard deviation of the logarithm (i.e., $\ln X$) of X , respectively; \mathbf{I} = a vector with N_t components that are all equal to one; $\underline{\varepsilon} = [\varepsilon_1, \dots, \varepsilon_{N_t}]^T$ = a standard Gaussian vector with N_t independent components; \underline{L} = a N_t -by- N_t upper-triangular matrix obtained by Cholesky decomposition of the correlation matrix \underline{R} satisfying

$$\underline{R} = \underline{L}^T \underline{L} \quad (7-3)$$

and the (i, j) -th entry of \underline{R} is given by the correlation function, e.g., Eq. (7-1). Using Eqs. (7-1) - (7-3), the ISV of X is explicitly simulated and is used as input in subsequent deterministic analysis of geotechnical structures to evaluate their responses (e.g., resistance moment, bearing capacity, and settlement, etc.) concerned in RBD. Note that little information on deterministic model of geotechnical structures is involved in Eqs. (7-1)-(7-3), making the ISV simulation using R-method be decoupled from the geotechnical deterministic analysis.

7.2.2 Approximate modeling by spatial average technique (A-method)

In A-method, the geotechnical design parameter X over a depth interval (e.g., influence zones) or along a spatial curve (e.g., critical slip surface of slope stability) is characterized by a single random variable $X_{\Delta z}$ that represents the spatial average of X over the depth interval or along the spatial curve and has a reduced variance due to spatial averaging (e.g., Vanmarcke, 1977; Griffiths and Fenton, 2004; Wang and Cao, 2013). Let Δz denote the length of the depth interval or the spatial curve. Due to the spatial averaging over Δz , the variance of the equivalent normal random variable $\ln X$ of X is reduced, and the variance reduction of $\ln X$ is described by a variance reduction function $\Gamma_{\Delta z}^2$ in A-method. For example, $\Gamma_{\Delta z}^2$ for the single exponential correlation function given by Eq. (7-1) is calculated as (Vanmarcke, 2010)

$$\Gamma_{\Delta z}^2 = \left[\lambda / (2\Delta z) \right]^2 \left\{ 2 \left[2\Delta z / \lambda - 1 + \exp(-2\Delta z / \lambda) \right] \right\} \quad (7-4)$$

Note that calculation of $\Gamma_{\Delta z}^2$ depends on the correlation function. Hence, for an exact evaluation of $\Gamma_{\Delta z}^2$, the correlation function is needed. This is a non-trivial task in geotechnical design practice because proper determination of the correlation function requires a large amount of geotechnical data that is usually not available at a specific site for routine geotechnical designs. Based on a limited number of geotechnical data obtained from site investigation, the most probable correlation function can be selected from a pool of candidates using Bayesian approaches (Cao and Wang, 2014, Tian et al., 2016). Alternatively, to avoid determining the correlation function, $\Gamma_{\Delta z}^2$ can be approximate as (e.g., Vanmarcke, 2010)

$$\Gamma_{\Delta z}^2 = \begin{cases} 1, & \Delta z < \lambda \\ \lambda / \Delta z, & \Delta z > \lambda \end{cases} \quad (7-5)$$

Eq. (7-5) gives a simplified form of the variance reduction function to conveniently calculate the variance reduction factor for various correlation structures, and it is valid for different correlation functions (Vanmarcke, 1977). Using Eq. (7-5) in A-method avoids determining the correlation function of geotechnical parameters. It is widely used in geotechnical literature (e.g., Vanmarcke, 1977; Phoon and Kulhawy, 1999b; El-Ramly et al., 2002; Stuedlein et al. 2012b; Wang and Cao, 2013). Figure 7-3 shows variance reduction functions given by Eqs. (7-4) and (7-5) by a solid line and a dashed line, respectively. The difference between the two variance functions is also plotted in Figure 7-3 by a dotted line. It is shown that the difference increases as $\Delta z / \lambda$ increases from 0 to 1 and then decreases as $\Delta z / \lambda$ increases further. The maximum difference occurs as $\Delta z / \lambda = 1$. Effects of using the approximate variance reduction function on geotechnical designs will be discussed later in this report.

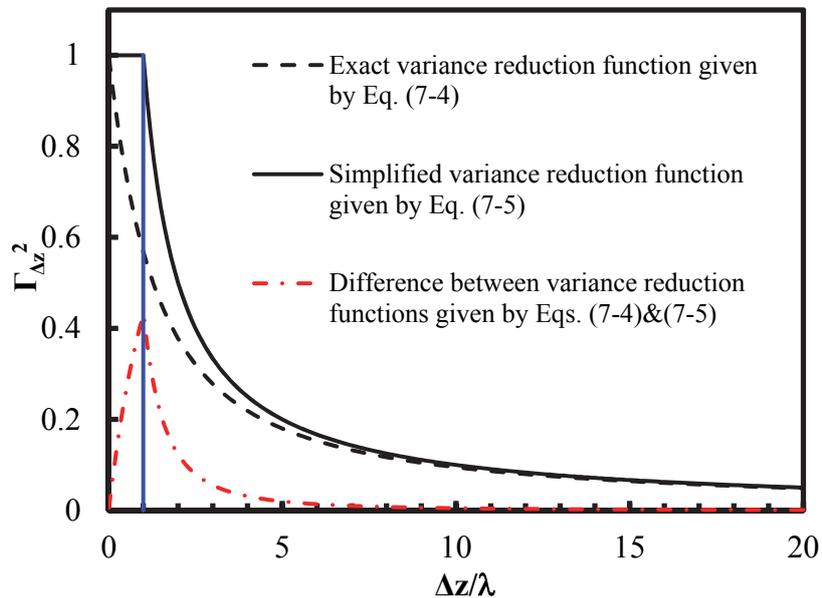


Figure 7-3 Comparison of variance reduction functions given by Eqs. (7-4) and (7-5).

Moreover, the geotechnical deterministic model may involve more than one spatial average of X over different sections (e.g., two depth intervals), which are spatially correlated. Let Δz_1 and Δz_2 denote the respective lengths of spatial average sections. When using A-method to model 1-D ISV, the spatial correlation between spatial averages of X over Δz_1 and Δz_2 is calculated as:

$$\rho_A = \frac{z_0^2 \Gamma_{z_0}^2 - z_1^2 \Gamma_{z_1}^2 - z_2^2 \Gamma_{z_2}^2 - z_{12}^2 \Gamma_{z_{12}}^2}{2\Delta z_1 \Delta z_2 \sqrt{\Gamma_{\Delta z_1}^2 \Gamma_{\Delta z_2}^2}} \quad (7-6)$$

where z_0 = separation distance between the two spatial average sections; z_1 = the distance from the beginning of the first section to the beginning of the second section; z_{12} = the distance from the beginning of the first section to the end of the second section; and z_2 = the distance from the end of the first section to the end of the second section. $\Gamma_{z_0}^2$, $\Gamma_{z_1}^2$, $\Gamma_{z_2}^2$, $\Gamma_{z_{12}}^2$, $\Gamma_{\Delta z_1}^2$, and $\Gamma_{\Delta z_2}^2$ = the respective variance reduction factors of $\ln X$ due to the spatial averaging over z_0 , z_1 , z_2 , z_{12} , Δz_1 and Δz_2 , which are illustrated in Figure 7-2.

Note that Eqs. (7-4)-(7-6) need the length of spatial average sections (e.g., influence zones for side resistance of drilled shafts and the critical slip surface of slope stability) as input information, which depends on geotechnical deterministic models. Determining proper spatial average sections is pivotal to calculating the variance reduction function in A-method. Hence, A-method couples the ISV modeling and geotechnical deterministic analyses, and it incorporates ISV into geotechnical design in an indirect and approximate manner. Effects of using A-method to model ISV on geotechnical designs can be evaluated by comparing the design results or reliability estimates (e.g., P_f) that are obtained using R-method and A-method, respectively. This is discussed using three geotechnical examples (including a drilled shaft example, a sheet pile wall example, and a soil slope example) in the following three sections.

7.3 Illustrative example I: Drilled Shaft

To explore effects of indirect and approximate modeling of ISV on foundation designs, this section redesigns a drilled shaft example using the expanded RBD approach together with R-method and A-method to model ISV, respectively. As shown in Figure 7-4, the drilled shaft is installed in loose sand with a total unit weight $\gamma = 20.0 \text{ kN/m}^3$ and mean effective stress friction angle $\mu_{\phi'} = 32^\circ$. The shaft is assumed to fail in drained general shear under a design compression load $F_{50} = 800 \text{ kN}$ with an allowable displacement $y_a = 25 \text{ mm}$. The key design parameters in this example are the drilled shaft diameter B and depth D , which are required to support the design compression load and to have a shaft displacement less than 25 mm.

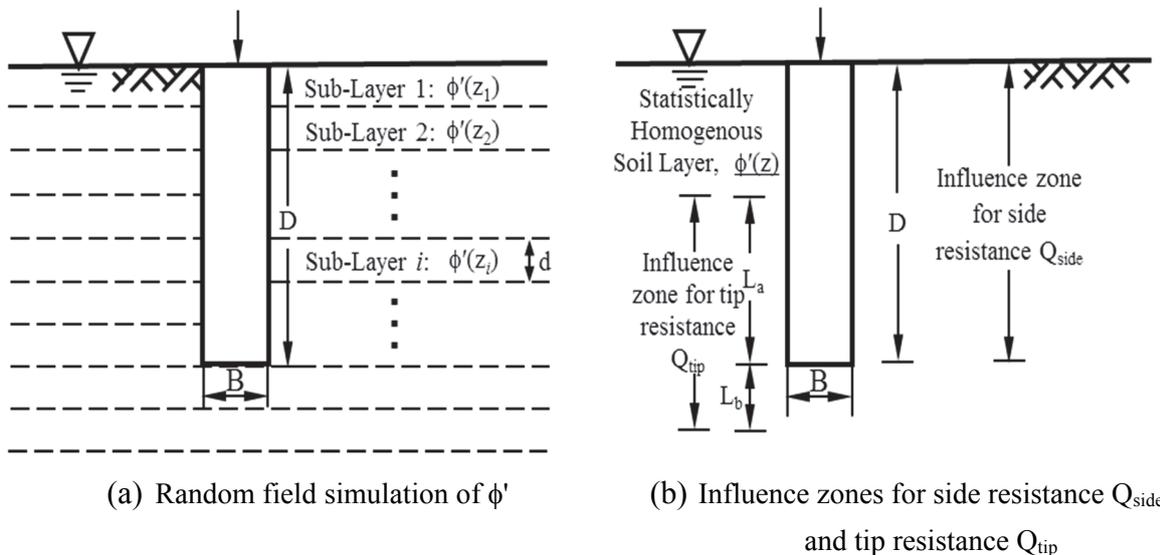


Figure 7-4 Illustration of spatial variability modeling using R-method [see 7-4(a)] and A-method [see 7-4(b)] in drilled shaft designs.

The expanded RBD approach is used to determine the minimum feasible design value (i.e., D_{min}) of D for a given B value. For comparison, R-method and A-method are applied to modeling ISV of ϕ' , leading to different design results through the expanded RBD approach. In R-method, Eqs. (7-1) -

(7-3) are used to directly simulate the random field of ϕ' in the sand layer (see Figure 7-4(a)), where λ varies from 0.2 to 1000m. In contrast, ϕ' in the sand layer surrounding the drilled shaft is modeled by ϕ'_{side} and ϕ'_{tip} in A-method, which represent the respective spatial averages of ϕ' over influence zones for evaluating side resistance Q_{side} and tip resistance Q_{tip} . As shown in Figure 7-4(b), the influence zone of Q_{side} is taken as the depth interval from ground surface to the tip and its length is equal to shaft depth D . The influence zone of Q_{tip} is taken as a depth interval from L_a (e.g., $\min\{8B, D\}$) above the tip to L_b (e.g., $3.5B$) below the tip, and its maximum length is $D_{max}+3.5B$ for a given B , where D_{max} is the maximum possible value of the shaft depth and is taken as 10m in this example. Note that the locations and lengths of influence zones for evaluating Q_{side} and Q_{tip} shall be specified in A-method (see Eqs. (7-4) - (7-6)) prior to the design, which depends on the deterministic analysis model used in design. More details of modeling and calculations of the drilled shaft example are referred to Wang and Cao (2013).

Figure 7-5 shows the variation of D_{min} for $B = 1.2\text{m}$ obtained using the expanded RBD approach with R-method as a function of normalized λ by circles. For each value of normalized λ , Figure 7-5 also includes D_{min} values obtained using A-method with the exact form (i.e., Eq. (7-4)) and the approximate form (i.e., Eq. (7-5)) of variance reduction function by squares and crosses, respectively. The D_{min} values here are determined according to the target failure probability $p_T = 0.0047$ for serviceability limit state, which has been shown to control the design in this example (Wang and Cao, 2013). As shown in Figure 7-5, for a given value of normalized λ , the circle generally plots closely to the squares. The results obtained using the A-method with the exact form of variance reduction function agree well with those obtained using R-method. This indicates that the spatial average of ϕ' represents “mobilized” value of ϕ' over the influence zones for drilled shaft resistance reasonably well in this example. Such an observation is further confirmed by comparing the side resistance estimated from realizations of ϕ' random field in R-method and their corresponding spatial averages of ϕ' over the shaft depth, as shown in Figure 7-6.

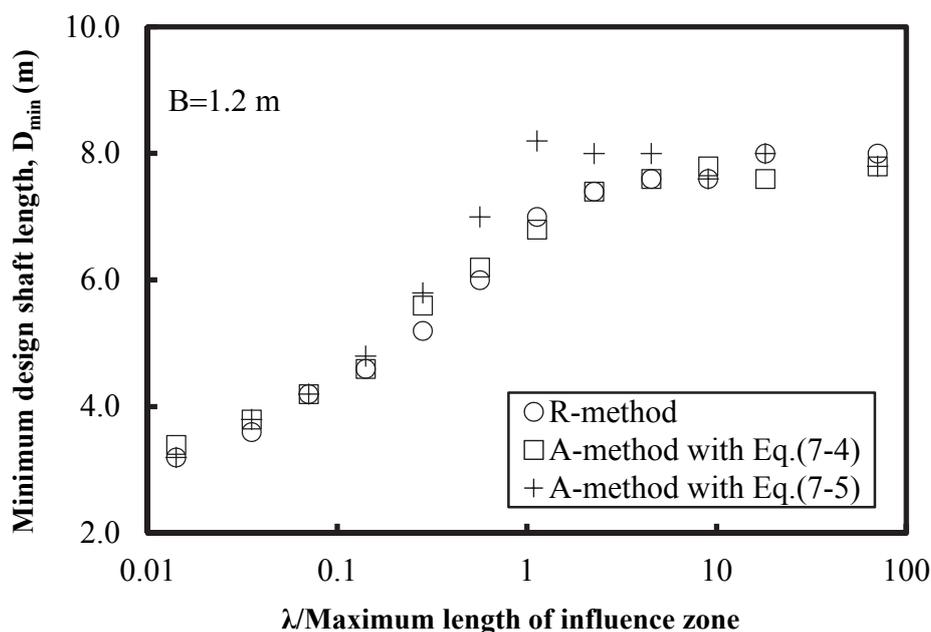


Figure 7-5 Comparison of drilled shaft design results using R-method and A-method in expanded RBD for $B=1.2\text{ m}$.

Figure 7-5 also compares design results obtained from A-method with the exact form (i.e., Eq. (7-4)) and the approximate form (i.e., Eq. (7-5)) of variance reduction function. As λ is smaller than one tenth of the maximum length L_{max} (i.e., $D_{max}+3.5B$ in this example) of influence zone for drilled shaft resistance or greater than ten times of L_{max} , using Eq. (7-4) and (7-5) gives similar design results. However, when λ is close to L_{max} , there is apparent difference between the two set of design results. Such a difference is attributed to approximation in variance reduction function. As shown in Figure 7-3, the maximum difference between variance reduction functions given by Eqs. (7-4) and (7-5)

occurs as λ is equal to the length of spatial average interval, e.g., L_{max} in this example. Hence, when the length of influence zone is close to λ adopted in design, the approximate variance reduction function given by Eq. (7-5) shall be used with caution.

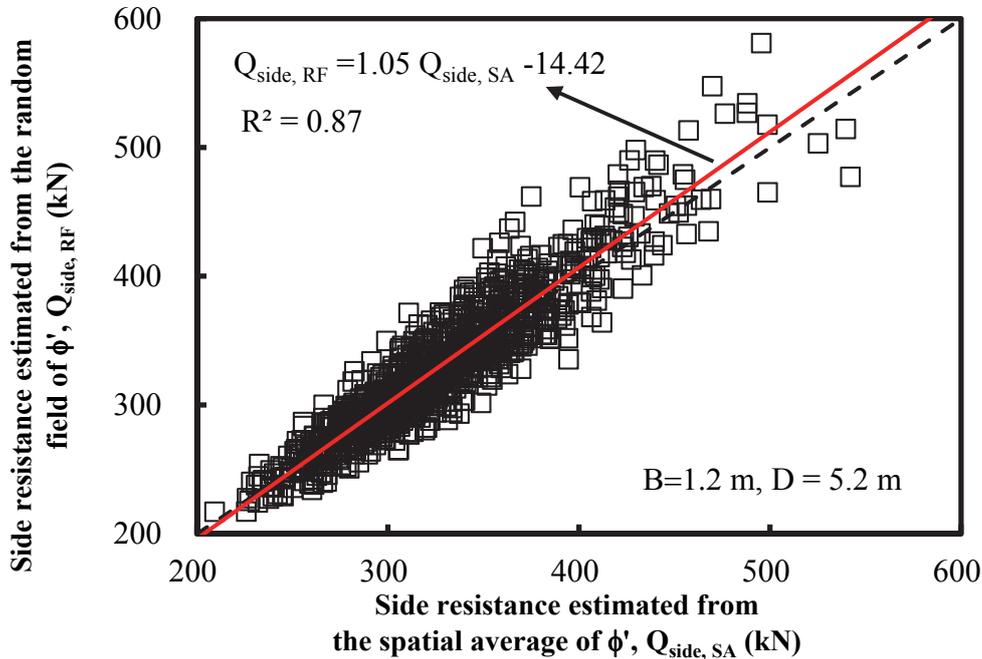


Figure 7-6 Comparison of side resistance calculated from realizations of ϕ' random fields and their corresponding spatial averages over influence zones for the design with $B = 1.2\text{m}$ and $D=5.2\text{m}$.

7.4 Illustrative example II: Sheet Pile Wall

For further illustration, this section redesigns an embedded sheet pile wall example using the expanded RBD approach together with R-method and A-method to model ISV in design, respectively. As shown in Figure 7-7, the embedded sheet pile wall is designed for a 3-m deep excavation, and is installed in a sand layer, where the total unit weight of sand is 20 kN/m^3 and effective friction angle ϕ' (i.e., $\ln X$ in Eq. (7-2)) of sand is normally distributed with a mean of 39° and a standard deviation of 3.9° . The ground water levels are at the ground surface in front of the wall and at the depth of 1.5m behind the wall. In addition, the surcharge q behind the wall is considered as a variable load, which is normally distributed and has a mean of 8.02 kPa and a standard deviation of 1.20 kPa . The aim of the sheet pile wall design example is to find an embedded depth d that satisfies the moment equilibrium about point O and to determine an additional embedded depth Δd by solving the horizontal force equilibrium equation (Wang, 2013). For simplification, Δd is commonly taken as $0.2d$, which often leads to conservative designs (e.g., Craig, 2004; Wang, 2013; Li et al., 2016b). Then, the required depth D_{spw} of the sheet pile wall example is equal as $1.2d$, and it ranges from 1m to 8m with an increment of 0.1m. For a given D_{spw} value, d (i.e., $D_{spw}/1.2$) and Δd (i.e., $0.2d$) are calculated, and the net resistance moment M_R about point O provided by passive earth pressure is evaluated, as well as the net overturning moment M_O resulted from the active pressure acting. After that, the FS is obtained, details of which are referred to Craig (2004) and Wang (2013).

The expanded RBD approach is used to determine the minimum feasible design value (i.e., D_{min}) of D_{spw} . Similar to the drilled shaft design; R-method and A-method are applied to modeling ISV of ϕ' in the sand layer in expanded RBD. In R-method, Eqs. (7-1) - (7-3) are used to directly simulate the random field of ϕ' in the sand layer (see Figure 7-7(a)), where λ varies from 0.2 m to 1000 m. In contrast, ϕ' in the sand layer surrounding the sheet pile wall is modeled by ϕ'_O and ϕ'_R in A-method, which represent the respective spatial averages of ϕ' over influence zones for evaluating M_O and M_R . As shown in Figure 7-7(b), the influence zone of M_R is taken as the depth interval from ground surface in front of the wall to point O, and its length is equal to d minus over-digging depth(0.3m); the influence zone of M_O is taken as a depth interval from ground surface behind the wall to point O, and

its maximum length is $d_{\max}+3.0\text{m}$ for a given B , where d_{\max} is the maximum possible value of the embedded depth and is taken as $8.0/1.2 = 6.67\text{m}$ in this example. More details of modeling and calculations of this drilled shaft design example are referred to Wang and Cao (2013) and Li et al. (2016b).

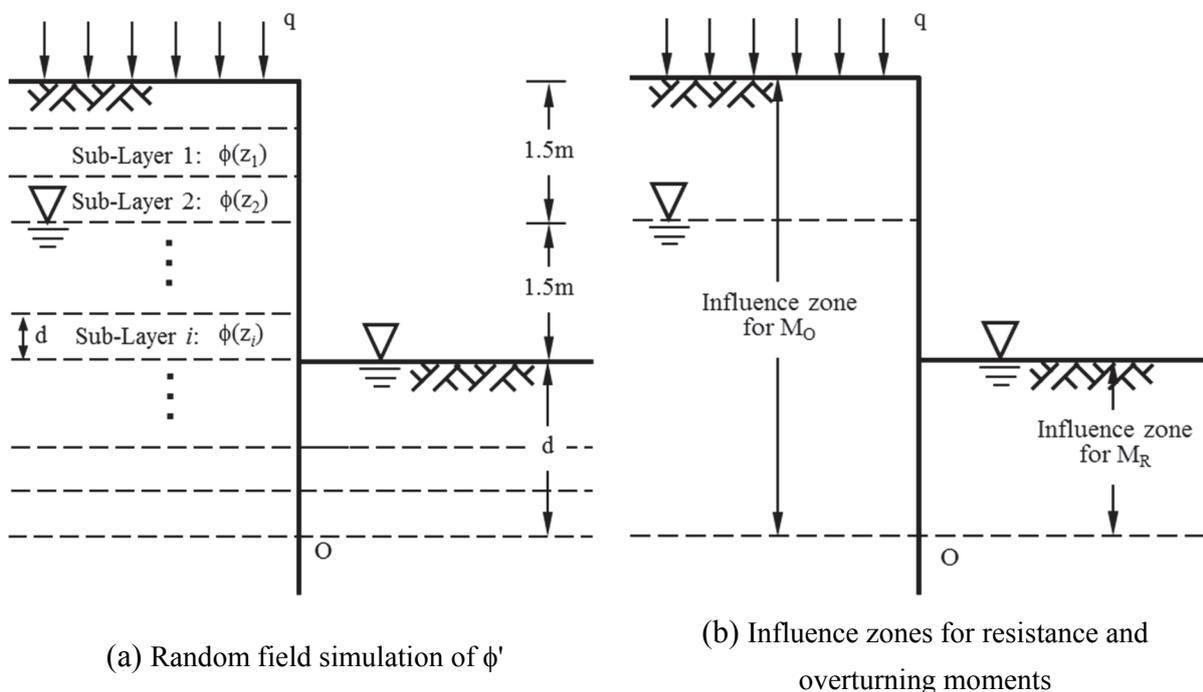


Figure 7-7 Illustration of spatial variability modeling using R-method [see 7-7(a)] and A-method [see 7-7(b)] in sheet pile wall designs.

Figure 7-8 shows the variation of D_{\min} obtained using the expanded RBD approach with R-method as a function of normalized λ by circles. For each value of normalized λ , Figure 7-8 also includes D_{\min} values obtained using A-method with the exact form (i.e., Eq. (7-4)) and the approximate form (i.e., Eq. (7-5)) of variance reduction function by squares and crosses, respectively. The D_{\min} values here are determined according to $p_T = 7.2 \times 10^{-5}$ adopted in Eurocode 7 (e.g., Orr and Breyse, 2008). As shown in Figure 7-8, for a given value of normalized λ , the circle generally plots closely to the squares. The results obtained using A-method with the exact form of the variance reduction function agree well with those obtained using R-method. This indicates that spatial averages of ϕ' represents “mobilized” values of ϕ' over the influence zones for M_R and M_O reasonably well in this example. Such an observation is further confirmed by comparing M_R and M_O values estimated from realizations of ϕ' random field in R-method with those calculated from their corresponding spatial averages of ϕ' over influence zones for M_R and M_O , as shown in Figures 7-9(a) and 7-9(b), respectively.

Figure 7-8 also compares design results obtained from A-method with the exact form (i.e., Eq. (7-4)) and the approximate form (i.e., Eq. (7-5)) of variance reduction function. Similar to the drilled shaft design example, when λ is close to L_{\max} , there is apparent difference between the two set of design results due to the obvious difference between variance reduction functions calculated from Eqs. (7-4) and (7-5) at $\lambda = L_{\max}$ (see Figure 7-3). This, again, indicates that, as the length of influence zone is close to λ adopted in design, using the approximate variance reduction function given by Eq. (7-5) in A-method may lead to inaccurate design results. Although using Eq. (7-5) in A-method gives conservative designs in the drilled shaft design example (see Figure 7-5) and the sheet pile design example (see Figure 7-8) as λ is close to L_{\max} , such an observation cannot be generalized, as illustrated using a soil slope example in the next section.

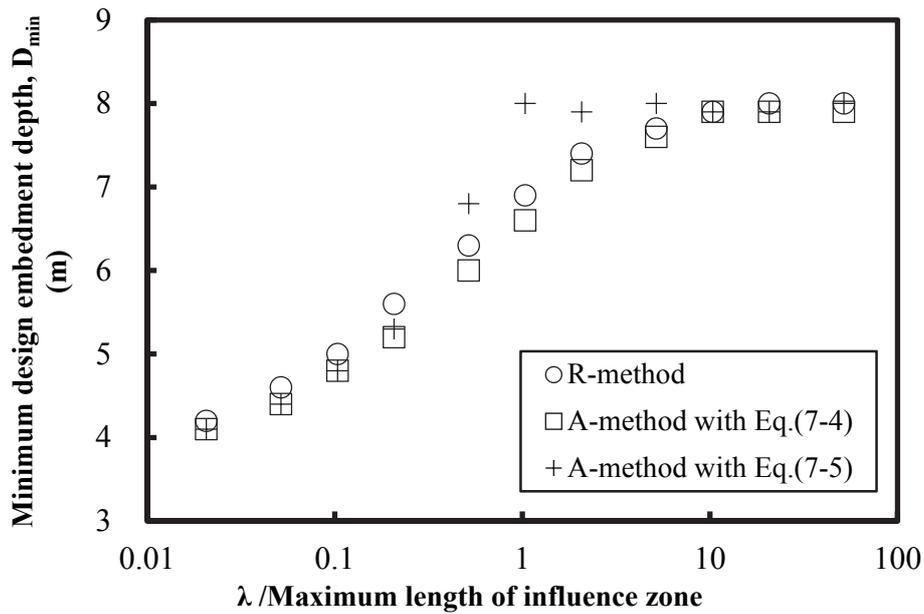


Figure 7-8 Comparison of sheet pile wall design results using R-method and A-method in expanded RBD.

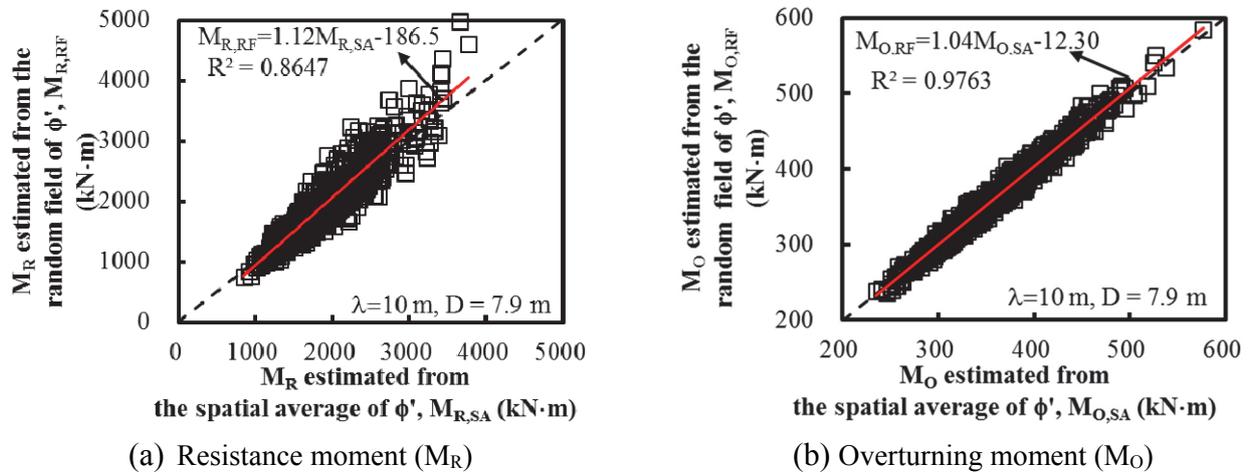


Figure 7-9 Comparison of resistance and overturning moments estimated from realizations of ϕ' random fields and their corresponding spatial averages over influence zones for the design with $D=7.9\text{m}$.

7.5 Illustrative example III: Lodalen Slide

This section illustrates effects of using different spatial variability modeling methods (i.e., R-method and A-method) on the “calculated” reliability (or probability of failure) of slope stability using Lodalen slide example. The Lodalen slide occurred in 1954 nearby the Oslo railway station, Norway. As shown in Figure 7-10, the slope has a height of 17m and a slope angle of 26° . The stratigraphy of slope is comprised of a marine clay layer underlying a 1 m thick clay crust at the top. The clay crust does not significantly affect the stability of the slope (e.g., El-Ramly et al., 2006) and is, hence, not shown in Figure 7-10. The spatial variability of effective cohesion c' , friction angle ϕ' and pore water pressure u in the marine clay layer is considered in this example, and they have respective mean values of 10 kPa, 27.1° , and 0m of water and respective standard deviations of 1.72kPa, 2.21° , and 0.34m of water (e.g., El-Ramly et al., 2006). The correlation structures of the three parameters are

considered as identical and are taken as an isotropic single exponential correlation function with a scale of fluctuation λ ranging from 2m to 5000m. More details of the Lodalen slide and its probabilistic assessment (including uncertainty characterization and propagation) are referred to El-Ramly et al. (2006).

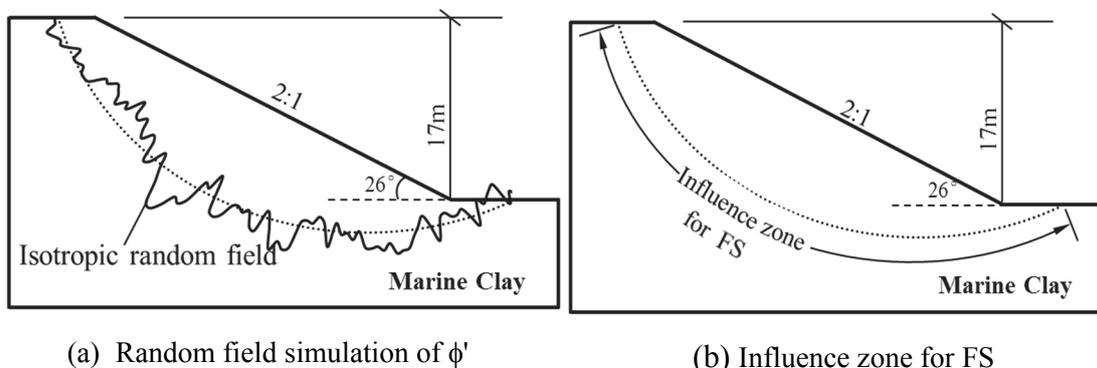


Figure 7-10 Illustration of spatial variability modeling for Lodalen slide using R-method [see 7-10(a)] and A-method [see 7-10(b)].

For illustration, R-method and A-method are applied to modeling spatial variability of effective shear strength parameters (i.e., c' and ϕ') in the marine clay layer. In this example, the spatial variability of u is always explicitly simulated as a random field by R-method no matter which method is used to modeling spatial variability of c' and ϕ' . Using R-method, random fields of c' and ϕ' can be directly simulated in the marine clay layer without needs of information on the slip surface of Lodalen slide. In contrast, such information is needed for determining the influence zone of sliding resistance of Lodalen slope in A-method. For simplicity, the critical slip surface adopted by El-Ramly et al. (2006) is considered in this report. Then, c' and ϕ' along the critical slip surface are modeled by c'_A and ϕ'_A in A-method, which represent the respective spatial averages of c' and ϕ' over the critical slip surface for evaluating its corresponding FS. The variances of c'_A and ϕ'_A are calculated using their respective variances at the “point” level and the variance reduction function given by Eq. (7-4) (exact form) or Eq. (7-5) (approximate form), in which the length of spatial average interval is taken as the length of the critical slip surface, i.e., about 52m in this example. Using R-method and A-method to model spatial variability of c' and ϕ' in the marine clay layer, the occurrence probability P_f of Lodalen slide along the prescribed critical slip surface is calculated for different values of λ varying from 2m to 5000m.

Figure 7-11 shows the variation of P_f values obtained using R-method as a function of normalized λ by circles. For each value of normalized λ , Figure 7-11 also includes P_f values obtained using A-method with the exact form (i.e., Eq. (7-4)) and the approximate form (i.e., Eq. (7-5)) of variance reduction function by squares and crosses, respectively. It is shown that the circles generally plot closely to the squares. The results obtained using the A-method with the exact form of the variance reduction function agree well with those obtained using R-method. This indicates that the spatial average of effective shear strength of the marine clay represents the “mobilized” value of effective shear strength over the critical slip surface reasonably well in this example. Similar to previous two examples, such an observation is further confirmed by comparing FS values estimated from realizations of c' and ϕ' random fields in R-method with those calculated from their corresponding spatial averages over the critical slip surface for a given λ value (e.g., 50m), as shown in Figure 7-12.

Figure 7-11 also compares P_f values obtained from A-method with the exact form (i.e., Eq. (7-4)) and the approximate form (i.e., Eq. (7-5)) of variance reduction function. When λ is close to the length of the critical slip surface, the crosses plot below the squares, indicating that using the approximate form (i.e., Eq. (7-5)) of variance reduction function leads to underestimation of P_f at relatively large failure probability levels, which is un-conservative. Such un-conservative results are attributed to overestimation of variance reduction function by Eq. (7-5) and, hence, variance of shear strength parameters. This, again, indicates that, as the length of influence zone is close to λ , Eq. (7-5) shall be used with caution because un-conservative reliability estimates might be obtained when it is applied.

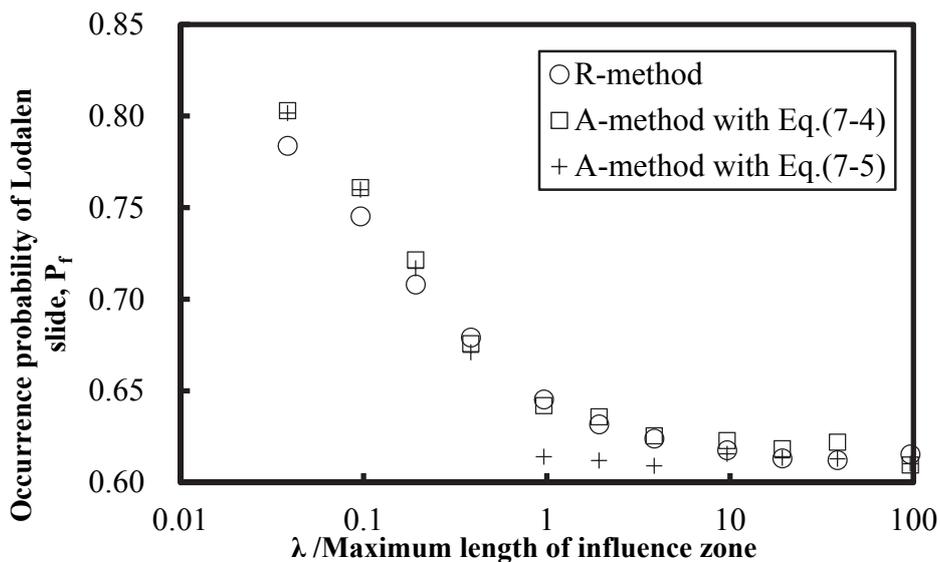


Figure 7-11 Comparison of occurrence probabilities of Lodalen slide using R-method and A-method in MCS.

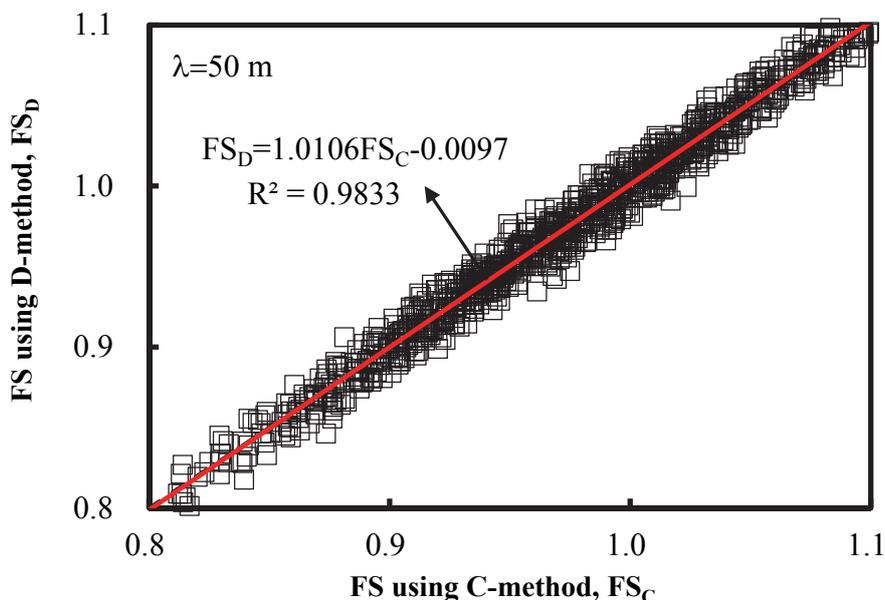


Figure 7-12 Comparison of safety factors of Lodalen slope calculated from random fields and spatial averages of effective shear strength parameters of the marine clays for $\lambda = 50\text{m}$.

7.6 Probabilistic characterization of inherent spatial variability

It is customary to decompose the spatial distribution of a soil property $y(z)$ into a spatial trend term $t(z)$ and a spatial variation term $w(z)$:

$$y(z) = t(z) + w(z) \quad (7-7)$$

where $w(z)$ is usually modeled as a zero-mean stationary random field with a standard deviation $= \sigma$ and with a finite-scale correlation structure characterized by an auto-correlation function (ACF). The key parameter for the ACF is the scale of fluctuation λ . Previous sections in this chapter have assumed $t(z)$, σ , and λ are known. In reality, they are unknown and should be estimated based on site-specific investigation data. Due to the limited site-specific data, $t(z)$, σ , and λ cannot be estimated with full certainty. This section considers the statistical uncertainty for $t(z)$, σ , and λ .

7.6.1 Statistical uncertainty in $t(z)$

The trend $t(z)$ for ISV is clearly site-specific. $t(z)$ can be estimated using regression based on the site investigation data. It is customary to assume that the estimated trend is the same as the actual trend. This is the underlying assumption for de-trending: data after de-trending (residuals) are treated as the zero-mean $w(z)$ data (e.g., Fenton 1999b; Uzielli et al. 2005). However, the de-trended data will not have zero mean if the estimated trend is not the actual trend. Past studies have recognized that de-trending deserves more rigorous attention (e.g., Kulatilake 1991; Li 1991; Jaksa et al. 1997; Fenton 1999a). The estimated trend is in principle not the same as the actual trend. Ching et al. (2017) addressed the identifiability of $t(z)$ in the presence of ISV using cone penetration test (CPT) data. They showed that if the CPT data is not “thick” enough, it is not always possible to discriminate between $t(z)$ and $w(z)$ because only their summation, i.e., $y(z)$, is measured. “Thick” means the CPT length is greater than $50 \times \lambda$ (Ching et al. 2017); many soil layers are not “thick” enough. There are two scenarios that can hinder the demarcation between $t(z)$ and $w(z)$: (a) a very flexible regression function so that $t(z)$ falsely over-fits the data; (b) a poor regression so that $t(z)$ does not fit the data well, but the residual is falsely interpreted as $w(z)$ with a large scale of fluctuation. Scenario (a) can be mitigated if a simple regression function (e.g., a linear function) is adopted or if the Sparse Bayesian Learning (Tipping 2001) framework is adopted. Moreover, it is hard to mitigate scenario (b). One possible solution discovered in Ching et al. (2017) is to adopt an auto-correlation model that produces “smooth” $w(z)$ realizations, such as the square exponential model. However, Ching et al. (2017) opined that the auto-correlation model should be chosen based on how well it captures the real data characteristics, rather than on its computational advantage.

7.6.2 Statistical uncertainty for σ and λ

The σ (or COV) and λ for ISV are also site-specific, because the COV and λ at one site are typically not the same as those at another site. The ranges summarized in previous tables represent past experiences in the literature. Although it is possible to assume conservative values for COV and λ based on these tables, as suggested by Honjo and Setiawan (2007), there are sometimes practical difficulties for doing so. First of all, COV and λ values in these tables vary in a wide range. For instance, the COV for ISV of the undrained shear strength of a clay varies from 6% to 80% (Phoon and Kulhawy 1999a). Its vertical λ can be as low as 0.1 m (Stuedlein 2011; Stuedlein et al. 2012a) to as great as 6.2 m (Phoon and Kulhawy 1999a). Its horizontal λ is known in a very limited way (one clay site in Stuedlein et al. (2012a) ranging from 4 to 8 m, 3 studies in Phoon and Kulhawy 1999a, ranging from 46 to 60 m). If the conservative COV value is taken to be 80% (the upper bound), this would be too conservative for most sites. If 80% is excessively conservative, which COV is reasonably conservative? It is not trivial to answer this question. The same question can be asked when λ is selected based on the ranges in Phoon and Kulhawy (1999a). Moreover, λ may depend on the problem scale (Fenton 1999b) and COV and λ may also depend on the adopted trend function and the sampling interval (Cafaro and Cherubini 2002). The scale considered in previous studies may not be similar to the scale applicable for the geotechnical project at hand. The trend function and sampling interval studied in the literature may not be applicable to the conditions in the project at hand.

Jaksa et al. (2005) took a different strategy: they suggested using a “worst case” λ , which for the example of a 3-storey, nine-pad footing building examined, is equal to the spacing between footings. This “worst case” strategy circumvents the need to estimate λ from past experiences. However, “the spacing between footing” is only applicable to the footing example. Table 7-1 shows the worse-case λ s reported in previous studies. The worst-case SOF is typically comparable to some multiple of the characteristic length of the structure (e.g., structure bay width or spacing between footings, height of retaining wall, diameter of tunnel, depth of excavation, height of slope). However, there is no universal way of determining the “worse-case” λ .

It is more prudent to implement the site investigation data to obtain the site-specific COV and λ than to assume their values using past experiences or to assume λ to be the “worse-case” λ . There are several techniques that can be employed to estimate COV and λ (in particular λ), such as the method of moments (Uzielli et al. 2005; Dasaka and Zhang 2012; Firouziandbandpey et al. 2014; Lloret-Cabot et al. 2014), the fluctuation function method (Wickremesinghe and Campanella 1993; Cafaro and Cherubini 2002), the maximum likelihood method (DeGroot and Baecher 1993), and the Bayesian method (Wang et al. 2010; Cao and Wang, 2013; Tian et al., 2016; Wang et al., 2016; Ching and

Phoon 2017). However, the main difficulty lies in the fact that the amount of available information in a typical site investigation program is not sufficient to accurately determine the site-specific COV and λ . Ching et al. (2015) showed that the vertical λ in a soil property cannot be estimated accurately if the total depth of the investigation data is less than 20 times of the actual vertical λ and that the COV cannot be estimated accurately if the total depth is less than 4 times of the vertical λ . They also found another requirement regarding the sampling interval. The vertical λ cannot be estimated accurately if the sampling interval is larger than 1/2 of the vertical λ . Furthermore, Ching et al. (2015) and Ching et al. (2017) showed that for soil layers that are not “thick” enough (thickness $< 50 \times \delta$), there can be a strong tradeoff between estimated COV and λ : there are numerous combinations of COV and λ that are all plausible with respect to the observed site-specific data. This tradeoff is closely related to the over-fit and poor-fit scenarios mentioned earlier (Ching et al. 2017). Accordingly, when insufficient data exists to determine the ISV parameters for a given site, the geological processes responsible for deposition and aging of the soil units of interest should be identified and previously reported ISV parameters for similar geological units should be used as a proxy for the site specific data, and with all due caution.

Table 7-1 Worse-case SOFs reported in previous studies.

Study	Problem type	“Worse case” definition	Characteristic length	Worse-case SOF
Jaksa et al. (2005)	Settlement of a nine-pad footing system	Under-design probability is maximal	Footing spacing (S)	1×S
Fenton and Griffiths (2003) Soubra et al. (2008)	Bearing capacity of a footing on a c- ϕ soil	Mean bearing capacity is minimal	Footing width (B)	1×B
Fenton et al. (2005)	Active lateral force for a retaining wall	Under-design probability is maximal	Wall height (H)	0.5~1×H
Fenton and Griffiths (2005)	Differential Settlement of Footings	Under-design probability is maximal	Footing spacing (S)	1×S
Breysse et al. (2005)	Settlement of a footing system	Footing rotation is maximal	Footing spacing (S)	0.5×S
		Different settlement between footings is maximal	Footing spacing (S) Footing width (B)	f(S,B) (no simple equation)
Griffiths et al. (2006)	Bearing capacity of footing(s) on a $\phi = 0$ soil	Mean bearing capacity is minimal	Footing width (B)	0.5~2×B
Ching and Phoon (2013) Ching et al. (2014)	Overall strength of a soil column	Mean strength is minimal	Column width (W)	1×W (compression) 0×W (simple shear)
Ahmed and Soubra (2014)	Differential Settlement of Footings	Under-design probability is maximal	Footing spacing (S)	1×S
Hu and Ching (2015)	Active lateral force for a retaining wall	Mean active lateral force is maximal	Wall height (H)	0.2×H
Stuedlein and Bong (2017)	Differential Settlement of Footings	Under-design probability is maximal	Footing spacing (S)	1×S
Ali et al. (2014)	Risk of infinite slope	Risk of rainfall induced slope failure is maximal	Slope height (H)	1×H

7.7 SUMMARY AND CONCLUDING REMARKS

This reports summarized major procedures for modeling spatial variability in geotechnical reliability-based design, based on which two methods, so-called rigorous(R)-method and approximate (A)-method are introduced. The probabilistic characterization of ISV, i.e., how to derive statistical information about geotechnical parameters such as spatial trend, standard deviation, and scale of fluctuation from site investigation data, is also briefly discussed. R-method directly simulates random fields of geotechnical design parameters without considering influence zones and/or critical slip surfaces that affect responses (e.g., resistance moment, bearing capacity, and settlement, etc.) of geotechnical structures. With R-method, ISV modeling is deliberately *decoupled* from geotechnical deterministic analyses. In contrast, A-method uses random field theory to calculate statistics of spatial averages of geotechnical design parameters within influence zones and/or along critical slip surfaces affecting responses of geotechnical structures. Information on influence zones and critical slip surfaces is needed for calculating statistics of the spatial average of geotechnical design parameters in A-method. Hence, using A-method in geotechnical RBD, ISV modeling is *coupled* with geotechnical deterministic analyses. Based on R-method and A-method, Monte Carlo simulation-based approaches (e.g., expanded RBD) is applied to explore effects of different spatial variability modeling methods on RBD and probabilistic analysis of geotechnical structures, including drilled shaft, sheet pile wall and soil slope. The major conclusions drawn from this study are given below:

- (1) Using A-method with exact variance reduction function gives design results and reliability estimates with satisfactory accuracy provided that reasonable influence zones or critical slip surfaces are assumed prior to the analysis. For a given influence zone or critical slip surface, the spatial average serves as a reasonable estimate of “mobilized” shear strength parameters for geotechnical RBD and probabilistic analysis when 1-D spatial variability is considered.
- (2) Compared with using exact form of variance reduction function (e.g., Eq. (7-4)), using approximate form of variance reduction function (e.g., Eq. (7-5)) in A-methods might lead to conservative or un-conservative reliability estimates and design results, depending on the failure probability level. As the length of spatial average interval is close to the scale of fluctuation, Eq. (7-5) shall be used with particular caution. Note that the effect of the simplified form of variance reduction function can be amplified when 2-D and 3-D spatial variability are considered because reductions in variances of different dimensions are basically “*multiplied*”.
- (3) Results obtained in this report are preliminary in the sense that failure mechanisms were prescribed prior to the analysis in the three examples and only 1-D spatial variability was taken into account. This is, however, a good starting point to explore the problems concerned in this report for real geotechnical structures, which is beneficial to the development of semi-probabilistic geotechnical RBD codes, such as Eurocode 7, and to communication with other research communities with limited background of geotechnical reliability and risk. More rigorous explorations are also warranted that account for effects of different failure mechanisms and 2-D (or 3-D) spatial variability. Some valuable attempts have been made in literature (e.g., Ching and Phoon, 2013; Ching et al., 2014, 2016).

7.8 REFERENCES

- Ahmed A. and Soubra A.-H. 2014. “Probabilistic analysis at the serviceability limit state of two neighboring strip footings resting on a spatially random soil,” *Structural Safety*, Vol. 49, pp. 2 – 9.
- Ali, A., Huang, J., Lyamin, A. V., Sloan, S.W, Griffiths DV, Cassidy, M.J. and Li, J.H., “Simplified quantitative risk assessment of rainfall-induced landslides modelled by infinite slopes.” *Engineering Geology*, 2014.179(0): 102-116.
- Baecher, G. B., Christian, J. T., 2003. Reliability and Statistics in Geotechnical Engineering. John Wiley and Sons, New Jersey.
- Breyse, D., Niandou, H., Elachachi, S., Houy, L., 2005. A generic approach to soil-structure interaction considering the effects of soil heterogeneity. *Géotechnique*, 55(2), 143-150.
- Cafaro, F., Cherubini, C., 2002. Large sample spacing in evaluation of vertical strength variability of clayey soil. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 128(7), 558-568.
- Cao, Z. J., Wang, Y., Li, D. Q., 2016. Quantification of prior knowledge in geotechnical site characterization. *Engineering Geology*, 203, 107-116.

- Cao, Z. J., Wang, Y., 2014. Bayesian model comparison and selection of spatial correlation functions for soil parameters. *Structural Safety*, 49, 10-17.
- Cao, Z. J., Wang, J. M., Wang, Y., 2013. Effects of spatial variability on reliability-based design of drilled shafts. *Geotechnical Special Publication: Foundation Engineering in the Face of Uncertainty: Honoring Fred H. Kulhawy*, 229, 602-616.
- Cao, Z. J., Wang, Y., 2013. Bayesian approach for probabilistic site characterization using cone penetration tests. *Journal of Geotechnical and Geoenvironmental Engineering*, 139(2), 267-276.
- Christian, J. T., Ladd, C. C., Baecher, G. B., 1994. Reliability applied to slope stability analysis. *Journal of Geotechnical Engineering*, 120 (12), 2180-2207.
- Ching, J., Phoon, K.K., Beck, J.L., Huang, Y., 2017. On the identification of geotechnical site-specific trend function, submitted to *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*.
- Ching, J., Phoon, K.K., 2017. Characterizing uncertain site-specific trend function by sparse Bayesian learning, to appear in *ASCE Journal of Engineering Mechanics*.
- Ching, J., Hu, Y. G., Phoon, K. K., 2016. On characterizing spatially variable soil shear strength using spatial average. *Probabilistic Engineering Mechanics*, 45, 31-43.
- Ching, J., Wu, S. S., Phoon, K. K., 2015. Statistical characterization of random field parameters using frequentist and Bayesian approaches. *Canada Geotechnical Journal*, 53(2), 285-298.
- Ching, J., Phoon, K. K., Kao, P. H., 2014. Mean and variance of mobilized shear strength for spatially variable soils under uniform stress states. *Journal of Engineering Mechanics*, 140(3), 487-501.
- Ching, J., Phoon, K. K., 2013. Mobilized shear strength of spatially variable soils under simple stress states. *Structural Safety*, 41(3), 20-28.
- Ching, J., Phoon, K. K., 2012. Probabilistic model for overall shear strengths of spatially variable soil masses. *GeoCongress 2012: State of the Art and Practice in Geotechnical Engineering*, 2866-2876.
- Cho, S. E., 2010. Probabilistic assessment of slope stability that considers the spatial variability of soil properties. *Journal of Geotechnical and Geoenvironmental Engineering*, 136 (7), 975-984.
- Craig, R. F., 2004. *Craig's Soil Mechanics*. Taylor & Francis, London.
- Dasaka, S. M., Zhang, L. M., 2012. Spatial variability of in situ weathered soil. *Géotechnique*, 62 (5), 375-384.
- DeGroot, D. J., Baecher, G. B., 1993. Estimating autocovariance of in situ soil properties. *Journal of Geotechnical Engineering*, 119(1), 147-166.
- El-Ramly, H., Morgenstern, N. R., Cruden, D. M., 2006. Lodalen slide: A probabilistic assessment. *Canada Geotechnical Journal*, 43(9), 956-968.
- El-Ramly, H., Morgenstern, N. R., Cruden, D. M., 2005. Probabilistic assessment of stability of a cut slope in residual soil. *Géotechnique*, 55(1), 77-84.
- El-Ramly, H., Morgenstern, N. R., Cruden, D. M., 2002. Probabilistic slope stability analysis for practice. *Canada Geotechnical Journal*, 39, 665-683.
- Fenton, G., Griffiths, D. V., 2008. *Risk Assessment in Geotechnical Engineering*. John Wiley and Sons, Hoboken, New Jersey.
- Fenton, G. A., Griffiths, D. V., 2007. Reliability-based Deep Foundation Design. *Geotechnical Special Publication: Probabilistic Applications in Geotechnical Engineering*, 170, 1-12.
- Fenton G.A., Griffiths D.V. 2005. Three-dimensional Probabilistic Foundation Settlement, *J. of Geotech. Geoenv. Engrg.*, 131(2), pp. 232–239.
- Fenton, G. A., Griffiths, D. V., Williams, M. B., 2005. Reliability of traditional retaining wall design. *Géotechnique*, 55(1), 55-62.
- Fenton, G. A., Griffiths, D. V., 2003. Bearing capacity prediction of spatially random $c-\phi$ soils. *Canada Geotechnical Journal*, 40 (1), 54-65.
- Fenton, G. A., Griffiths, D. V., 2002. Probabilistic foundation settlement on spatially random soil. *Journal of Geotechnical and Geoenvironmental Engineering*, 128 (5), 381-390.
- Fenton, G., 1999a. Estimation for stochastic soil models. *Journal of Geotechnical and Geoenvironmental Engineering*, 125(6), 470-485.
- Fenton, G., 1999b. Random field modeling of CPT data. *Journal of Geotechnical and Geoenvironmental Engineering*, 125(6), 486-498.
- Fenton, G., Vanmarcke, E. H., 1990. Simulation of random fields via local average subdivision. *Journal of Engineering Mechanics*, 116(8), 1733-1749.
- Firouziandbandpey, S., Griffiths, D. V., Ibsen, L. B., Andersen L. V., 2014. Spatial correlation length of normalized cone data in sand: case study in the north of Denmark. *Canada Geotechnical Journal*,

- 51(8), 844-857.
- Firouziandbandpey, S., Ibsen, L. B., Griffiths, D. V., Vahdatirad, M. J., Andersen L. V., Sørensen, J. D., 2015. Effect of spatial correlation length on the interpretation of normalized CPT data using a kriging approach. *Journal of Geotechnical and Geoenvironmental Engineering*, 141 (12), 04015052.
- Griffiths, D. V., Fenton, G. A., 2009. Probabilistic settlement analysis by stochastic and random finite-element methods. *Journal of Geotechnical and Geoenvironmental Engineering*, 135 (11), 1629-1637.
- Griffiths, D. V., Fenton, G. A., Manoharan, N., 2006. Undrained bearing capacity of two-strip footings on spatially random soil. *ASCE International Journal of Geomechanics*, 6(6), 421-427.
- Griffiths, D. V., Fenton, G. A., 2004. Probabilistic slope stability analysis by finite elements. *Journal of Geotechnical and Geoenvironmental Engineering*, 130 (5), 507-518.
- Honjo, Y., Setiawan, B., 2007. General and local estimation of local average and their application in geotechnical parameter estimations. *Georisk*, 1(3), 167-176.
- Hu, Y. G., Ching, J., 2015. Impact of spatial variability in soil shear strength on active lateral forces, *Structural Safety*, 52, 121-131.
- Huang, J., Lyamin, A.V., Griffiths, D.V., Krabbenhøft. K., Sloan, S.W., 2013. Quantitative risk assessment of landslide by limit analysis and random fields. *Computers and Geotechnics*, 53, 60–67.
- Huang, J. S., Griffiths, D. V., Fenton, G. A., 2010. System reliability of slopes by RFEM. *Soils and Foundations*, 50 (3), 343-353.
- Jaksa, M. B., Goldsworthy, J. S., Fenton, G. A., Kaggwa, W. S., Griffiths, D. V., Kuo, Y. L., Poulos, H.G., 2005. Towards reliable and effective site investigations. *Géotechnique*, 55(2), 109-121.
- Jaksa, M. B., Brooker, P. I., Kaggwa, W. S., 1997. Inaccuracies associated with estimating random measurement errors. *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, 123(5), 393-401.
- Jaksa, M. B., 1995. The influence of spatial variability on the geotechnical design properties of a stiff, overconsolidated clay. PHD Thesis of the University of Adelaide.
- Jiang, S. H., Li, D. Q., Cao, Z. J., Zhou, C. B., Phoon, K. K., 2015. Efficient system reliability analysis of slope stability in spatially variable soils using Monte Carlo simulation. *Journal of Geotechnical and Geoenvironmental Engineering*, 141(2), 04014096.
- Jiang, S. H., Li, D. Q., Zhang, L. M., Zhou, C. B., 2014. Slope reliability analysis considering spatially variable shear strength parameters using a non-intrusive stochastic finite element method. *Engineering Geology*, 168, 120-128.
- Kulatilake, P. H. S., 1991. Discussion on 'Probabilistic potentiometric surface mapping' by P.H.S. Kulatilake. *ASCE Journal of Geotechnical Engineering*, 117(9), 1458-1459.
- Kulhawy, F. H., 1996. From Casagrande's 'Calculated Risk' to reliability-based design in foundation engineering. *Civil Engineering Practice*, 11 (2), 45-56.
- Li, K. S., 1991. Discussion on 'Probabilistic potentiometric surface mapping' by P.H.S. Kulatilake. *ASCE Journal of Geotechnical Engineering*, 117(9), 1457-1458.
- Li, C. C., Kiureghian, A. D., 1993. Optimal discretization of random fields. *Journal of Engineering Mechanics*, 119(6), 1136-1154.
- Li, D. Q., Xiao, T., Cao, Z. J., Zhou, C. B., Zhang, L. M., 2016a. Enhancement of random finite element method in reliability analysis and risk assessment of soil slopes using Subset Simulation. *Landslides*, 13, 293-303.
- Li, D. Q., Shao K. B., Cao Z. J., Tang, X. S., Phoon, K. K., 2016b. A generalized surrogate response aided-subset simulation approach for efficient geotechnical reliability-based design. *Computers and Geotechnics*, 74, 88-101.
- Li, D. Q., Qi, X. H., Cao Z. J., Tang, X. S., Zhou, W., Phoon, K. K., Zhou, C. B., 2015a. Reliability analysis of strip footing considering spatially variable undrained shear strength that linearly increases with depth. *Soils and Foundations*, 55(4), 866-880.
- Li, D. Q., Jiang, S. H., Cao Z. J., Zhou, W., Zhou, C. B., Zhang, L. M., 2015b. A multiple response-surface method for slope reliability analysis considering spatial variability of soil properties. *Engineering Geology*, 187, 60-72.
- Li, D. Q., Qi, X. H., Phoon, K. K., Zhang, L. M., Zhou, C. B., 2014. Effect of spatially variable shear strength parameters with linearly increasing mean trend on reliability of infinite slopes. *Structural Safety*, 49, 45-55.
- Lloret-Cabot, M., Fenton, G.A., Hicks, M.A., 2014. On the estimation of scale of fluctuation in

- geostatistics. *Georisk*, 8(2), 129-140.
- Mitchell, J. K., Soga, K., 2005. *Fundamentals of Soil Behavior*. John Wiley and Sons, New Jersey.
- Orr, T. L., 2015. Risk and reliability in geotechnical engineering. *Managing Risk and Achieving Reliable Geotechnical Designs Using Eurocode 7*, Chap. 10. Edited by Phoon, K. K., Ching, J. Y., Taylor & Francis CRC Press, 395-433.
- Orr, T. L., Breyse, D., 2008. Eurocode 7 and reliability-based design. In *Reliability-Based Design in Geotechnical Engineering Computations and Applications*, Chap. 8. Edited by Phoon, K. K., London: Taylor & Francis, 298-343.
- Phoon, K. K., Huang, S. P., Quek, S.T., 2002. Implementation of Karhunen-Loève expansion for simulation using a wavelet-Galerkin scheme. *Probabilistic Engineering Mechanics*, 17(3), 293-303.
- Phoon, K. K., Kulhawy, F. H., 1999a. Characterization of geotechnical variability. *Canada Geotechnical Journal*, 36 (4), 612-624.
- Phoon, K. K., Kulhawy, F. H., 1999b. Evaluation of geotechnical property variability. *Canada Geotechnical Journal*, 36(4), 625-639.
- Soubra, A. H., Massih, Y. A., Kalfa, M., 2008. Bearing capacity of foundations resting on a spatially random soil. *GeoCongress 2008: Geosustainability and Geohazard Mitigation (GSP 178)*, New Orleans Louisiana USA, ASCE, 66-73.
- Stuedlein, A.W. 2017. Role of Lower Bound Capacity and Shear Strength Anisotropy on Probabilistic Bearing Capacity of Plastic Fine-grained Soils. *Geotechnical Special Publication Honoring Wilson Tang*, ASCE, Reston, VA, 13pp.
- Stuedlein, A.W. and Bong, T. 2017. Effect of Spatial Variability on Static and Liquefaction-induced Differential Settlements, *Georisk 2017*, ASCE, Denver, Co.
- Stuedlein, A.W., Kramer, S. L., Arduino, P., Holtz, R. D., 2012a. Geotechnical characterization and random field modeling of desiccated clay. *Journal of Geotechnical and Geoenvironmental Engineering*, 138(11), 1301-1313.
- Stuedlein, A.W., Kramer, S.L., Arduino, P., and Holtz, R.D. 2012b. Reliability of Spread Footing Performance in Desiccated Clay. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, Vol. 138(11), pp. 1314-1325.
- Stuedlein, A.W. 2011. Random Field Model Parameters for Columbia River Silt. *Proceedings, GeoRisk, GSP No. 224*, ASCE, Reston, VA. 9 pp.
- Tian, M., Li, D. Q., Cao, Z. J., Phoon K. K., Wang, Y., 2016. Bayesian identification of random field model using indirect test data. *Engineering Geology*, 210, 197-211.
- Tipping, M.E., 2001. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1, 211-244.
- Uzielli, M., Vannuchi, G., Phoon, K. K., 2005. Random field characterization of stress-normalized cone penetration testing parameters. *Géotechnique*, 55(1), 3-20.
- Vanmarcke, E. H., 2010 *Random fields: analysis and synthesis*. World Scientific.
- Vanmarcke, E. H., 1977. Probabilistic modeling of soil profiles. *Journal of the Geotechnical Engineering Division*, 103 (11), 1227-1246.
- Wang, Y., Cao, Z. J., Li, D. Q., 2016. Bayesian perspective on geotechnical variability and site characterization. *Engineering Geology*, 203, 117-125.
- Wang, Y., Cao, Z. J., 2015. Practical reliability analysis and design by Monte Carlo Simulation in spreadsheet. *Risk and Reliability in Geotechnical Engineering*, Chap. 7. Edited by Phoon, K. K., Ching, J. Y., Taylor & Francis CRC Press, 301-335.
- Wang, Y., Cao, Z. J., 2013. Expanded reliability-based design of piles in spatially variable soil using efficient Monte Carlo simulations. *Solis and Foundations*, 53(6), 820-834.
- Wang, Y., 2013. MCS-based probabilistic design of embedded sheet pile walls. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 7(3), 151-162.
- Wang, Y., Au, S. K., Kulhawy, F. H., 2011. Expanded reliability-based design approach for drilled shafts. *Journal of Geotechnical and Geoenvironmental Engineering*, 137(2), 140-149.
- Wang, Y., Cao, Z. J., Au, S. K., 2011. Practical reliability analysis of slope stability by advanced Monte Carlo simulations in spreadsheet. *Canada Geotechnical Journal*, 48 (1), 162-172.
- Wang, Y., 2011. Reliability-based design of spread foundations by Monte Carlo simulations. *Géotechnique*, 61(8), 677-685.
- Wang, Y., Au, S. K., Cao, Z. J., 2010. Bayesian approach for probabilistic characterization of sand friction angles. *Engineering Geology*, 114 (3-4), 354-363.
- Wickremesinghe, D., Campanella, R. G., 1993. Scale of fluctuation as a descriptor of soil variability.

Probabilistic Methods in Geotechnical Engineering (Eds.: Li and Lo), Balkema, Rotterdam, The Netherlands, 233-239.

Xiao, T., Li, D. Q., Cao, Z. J., Au, S. K., Phoon, K. K., 2016. Three-dimensional slope reliability and risk assessment using auxiliary random finite element method. *Computers and Geotechnics*, 79, 146-158.

Zhang, L. Y., Chen, J. J., 2012. Effect of spatial correlation of standard penetration test (SPT) data on bearing capacity of driven piles in sand. *Canada Geotechnical Journal*, 49 (4), 394-402.

Discussions and Replies

After the interim report was released, constructive comments and suggestions are gratefully received from group members (including Jin-Song Huang, Shin-ichi Nishimura and Armin Stuedlein) and Jianye Ching. Their comments have been incorporated into this final report. Some discussions are summarized in this section and are listed below by the date they were received.

Discussion by Jianye Ching (National Taiwan University, Taiwan)

I have the following questions and comments:

1. I understand your D-method is the rigorous method, such as random finite element. However, I do not understand why it is called "decoupled" method. Maybe it is because D-method does not need to worry about the critical slip surface (?). However, this does not imply D-method does not consider critical slip surface. In fact, it DOES consider the critical slip surface, but in an automatic way: D-method automatically finds it. The term "decoupled" can be misleading because it seems to suggest that the rigorous method decouples the random field simulation with the finding of the critical slip surface, which seems to be not true to me.
2. My understanding is that your C-method is an approximate method that treats the critical slip surface as a prescribed surface (?). Again, I do not understand why it is called "coupled". My feeling is that it should actually be called "decoupled" because the slip surface is prescribed. You can simulate random field along it without searching for the actual critical slip surface (which is not prescribed). Therefore, C-method decouples the random field simulation with the finding of the critical slip surface.

Maybe it will be simpler and also clearer to name them as rigorous and approximate methods, or non-prescribed and prescribed methods?

Reply by Dianqing Li & Zijun Cao (Wuhan University, China)

Thank you for the comments. We agree that the wording “decoupled” and “coupled” is misleading. Actually, we used “decoupled” and “coupled” methods in the interim report to highlight that:

- Spatial variability modeling using D-method does **NOT** require information on influence zones and/or critical slip surfaces which shall be determined from deterministic analyses of geotechnical structures. This allows spatial variability modeling to **be decoupled from** deterministic analyses of geotechnical structures.
- Spatial variability modeling using C-method needs information on influence zones and/or critical slip surfaces which is unknown prior to geotechnical deterministic analyses and shall be assumed for calculating statistics of spatial averages of geotechnical design parameters. Hence, using C-method in geotechnical RBD, spatial variability modeling **is coupled** with geotechnical deterministic analyses.

As pointed in your comment 1, spatial variability modeling will eventually affect the mechanical analysis of geotechnical structures, e.g., finding of critical slip surface in random finite element analysis, for each realization of random fields generated in D-method.

On the other hand, although the critical slip surface is prescribed in C-method and searching for critical slip surface is avoided, spatial variability is still coupled with deterministic analysis because information from deterministic analysis (e.g., geometry and locations of critical slip surface) is used in C-method to model spatial variability.

In summary, it is **NOT** accurate to say that spatial variability modeling is *decoupled* or *coupled* from deterministic analysis based on, only, whether information from deterministic analysis is used in spatial variability modeling or not. As suggested, we have renamed the two methods as “**rigorous**” and “**approximate**” methods in the final report.

Following-up discussion by Jinsong Huang (The University of Newcastle, Australia)

I agree with Jianye that the terminology of C-method and D-method are confusing. “A and B are coupled” means that the change in A will affect B and vice versa. Although the failure mechanism is

affected by the uncertain material properties, the uncertain material properties are independent of failure mechanism.

I have two comments:

1. The second line on page 2. I suggest removing “limit equilibrium method”. The Random Finite Element Method finds the failure mechanism automatically for a given realization of random field. The Limit Equilibrium Method predefines the shape of slip surfaces, which I believe, is inappropriate when the material properties are modelled as random fields. It is hard to believe that we can predefine the shape of slip surface if the strengths of soils are spatially nonhomogeneous.
2. For the Lodalen slide example, only one fixed slip surface is considered. In geotechnical reliability-based design, when the geometry of slope is varied, how can we find such a slip surface?

Reply by Dianqing Li & Zijun Cao (Wuhan University, China)

We agree that the wording “decoupled” and “coupled” is misleading and have renamed the two methods as “**rigorous**” and “**approximate**” methods in the final report. Please refer to responses to Jianye’s comments above.

Re the comment 1, “limit equilibrium method and finite element method” are taken as examples of different geotechnical models. We do **NOT** mean that limit equilibrium methods can allow automatic searching of critical slip surfaces with various shapes in this report although this might be possible for rigorous limit equilibrium procedures, such as Morgenstern-Price method. We have revised the report to clarify that limit equilibrium method and finite element method are examples of different deterministic analysis models of slope stability.

Re the comment 2, we agree that, when slope geometry is changed, the critical slip surface shall be determined by performing deterministic analysis, such as finite element analysis. When spatial variability is considered, the critical slip surface is hardly fixed. In this report, we aim to explore two problems:

- How does indirect and approximate modeling of inherent spatial variability using the A-method affect reliability-based design of real geotechnical structures?
- How does using different ways to calculate the variance reduction function in A-method affect geotechnical reliability-based design?

For the purpose of simplification, the critical slip surface adopted by El-Ramly et al. (2006) is used in illustrative example III. We consider this simplification as a starting point to explore the problems concerned in this report for real geotechnical structures although the results might be preliminary due to the simplification, as pointed out by point (3) in section 7.6 entitled “**Summary and concluding remarks**” of the report. We also agree that more rigorous explorations are warranted that account for effects of different failure mechanism/paths, such as attempts made by Ching and Phoon (2013) and Ching et al., (2014, 2016).

REFERENCE:

- Ching, J. Y., Phoon, K. K., 2013. Mobilized shear strength of spatially variable soils under simple stress states. *Structural Safety*, 41(3), 20-28.
- Ching, J. Y., Phoon, K. K., Kao, P. H., 2014. Mean and variance of mobilized shear strength for spatially variable soils under uniform stress states. *Journal of Engineering Mechanics*, 140(3), 487-501.
- Ching, J. Y., Hu, Y. G., Phoon, K. K., 2016. On characterizing spatially variable soil shear strength using spatial average. *Probabilistic Engineering Mechanics*, 45, 31-43.
- El-Ramly, H., Morgenstern, N. R., Cruden, D. M., 2006. Lodalen slide: A probabilistic assessment. *Canada Geotechnical Journal*, 43(9), 956-968.

Discussion by Shin-ichi Nishimura (Okayama University, Japan)

I believe one of most important points for the modeling of the spatial variability is separating the trend and the random variable. For this task, the information statistics such as AIC, BIC, and etc.

is very effective. In the report, this issue is not presented well. If possible, I hope it should be considered in report. I will attach my example.

Reply by Dianqing Li & Zijun Cao (Wuhan University, China)

Thank you for the comments. As noted by Jianye in his reply below, this report focuses on **modeling** inherent spatial variability in geotechnical reliability-based design based on known/assumed statistical information of geotechnical parameters, and does not discuss probabilistic **characterization** of spatial variability, i.e., how to derive statistical information from site investigation data and prior knowledge. Some relevant studies on probabilistic characterization of spatial variability of geotechnical parameters are provided in the report for reference. Please see Paragraph 2 of section 7.1 entitled “**Introduction**”.

Reply by Jianye Ching (National Taiwan University, Taiwan)

You are very right that the identification of trend is crucial. I think Dianqing's report focuses on the modeling of spatial variability, not in the characterization of spatial variability.

Chapter 8 Imprecise Probabilistic and Interval Approaches Applied to Partial Factor Design

Lead discussor:
Sónia H. Marques
sonia.h@fe.up.pt

Discussors:
Kerstin Lesny

8.1 CHAPTER AT A GLANCE

EN 1997, adopted as Eurocode 7, is intended to be applied to the geotechnical aspects of the design of civil engineering works. The limit state design concept adopted by Eurocode 7 is used in conjunction with a partial factor methodology. The selection of appropriate partial factors is important to ensure the reliability of geotechnical design to Eurocode 7, as design values are determined by applying partial factors to characteristic values. When the partial factor format is first introduced, it should preferably produce a design comparable to the resultant from the safety factor methodology, promoting the continuity of past experience. Actually, the partial factor format through the Design Approach DA.2* is based on a modified global safety concept, noted that different systems associated to the same factor may have a different probability of failure due to the fact that important variabilities are disregarded. At the present, the performance of a partial factor format is measured by the ability to produce a design achieving a desired target reliability within acceptable margin of error. To achieve the required target reliability, Eurocode 7 does not provide any variation in the partial factors but rather requires that greater attention is given to other accompanying measures related to design supervision and inspection differentiation by a system of failure control. The issue of adopting multiple resistance partial factors in geotechnical design is then on discussion.

At a glance, multiple resistance partial factors should be clearly related to the determination of characteristic values, primary cause of inconsistent reliability evaluations. Robustness is considered as one of the primary requirements in a design process accounting for uncertainty. Although there is no universal definition for robustness, the concept expresses the degree of independence among any changes in the whole set of parameters and the fluctuation of the response considered a global specification on a minimum variance with respect to input variations. The state of play on the assessment of geotechnical robustness has been recently extended to partial factor design. The issue of adopting multiple resistance partial factors in geotechnical design is then addressed. This discussion involves a comprehensive design example referred to a strip spread foundation designed by the Eurocode 7 methodology. A comparative study for bearing capacity safety assessment comprehending combinatory geotechnical characteristic values and variabilities is then presented wherein the influence of variability of soil shear strength parameters on resistance partial factor design is investigated. Principal conclusions detail at first the particular conditions on how it is possible to achieve a reasonable reliability level in geotechnical design for a set of scenarios. Thereafter, the shear strength parameters of the foundation soil are implemented as intervals based on which the characteristic values for design are derived. On this case study the shear strength parameter friction angle of the foundation soil is further separately implemented as interval in the format of a conditional analysis. Limit state imprecise probabilistic grid-based and fuzzy-based approaches applied to the Eurocode 7 partial factor design for bearing capacity safety assessment are then pursued. It is observed the format of a sensitivity analysis for a group of model cases, considered for demonstration only normal variabilities. Imprecise approaches to robust design are then discussed on the calculation of resistance factors capable to maintain a more uniform reliability level over a range of design parameters. At last, the safety margin is expressed in the interval format so that a nonprobabilistic concept of reliability is approached by bounds.

8.2 DESIGN EXAMPLE

The design example is referred to the strip spread foundation on a relatively homogeneous soil shown in Figure 8-1, wherein groundwater level is away. Considered the vertical noncentric loading problem and the calculation model for bearing capacity, the performance function may be described by the simplified Equation (8-1):

$$M = f(B, D, \gamma_s, c_f, \varphi_f, \gamma_f, P, Q) \quad (8-1)$$

if B is the foundation width; D is the soil height above the foundation base; γ_s is the unit weight of the soil above the foundation base; c_f is the cohesion of the foundation soil; φ_f is the friction angle of the foundation soil; γ_f is the unit weight of the foundation soil; P is the dead load; and Q is the live load.

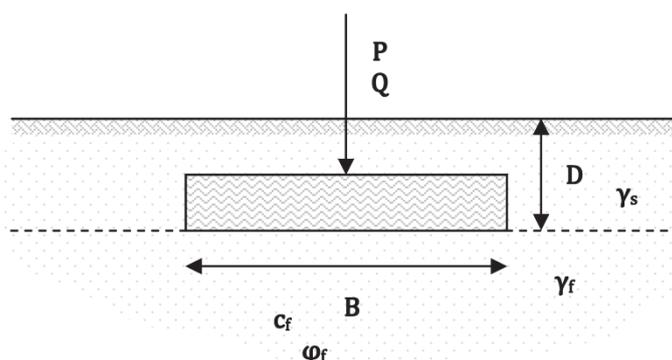


Figure 8-1.Strip spread foundation.

The strip spread foundation is designed by the Eurocode 7 methodology, Design Approach DA.2*. At first, a comparative study for bearing capacity safety assessment comprehending multiple resistance partial factors for a set of scenarios involving combinatory geotechnical characteristic values and variabilities is presented. For the purpose, the considered correlation coefficients between the basic input variables are presented in Table 8-1 and the description of basic input variables, with different types of distributions, is summarised in Table 8-2.

Table 8-1.Correlation coefficients between the basic input variables.

Correlation matrix												
ρ_{x1x1}	ρ_{x1x2}	ρ_{x1x3}	ρ_{x1x4}	ρ_{x1x5}	ρ_{x1x6}		1.0	0.0	0.5	0.9	0.0	0.0
ρ_{x2x1}	ρ_{x2x2}	ρ_{x2x3}	ρ_{x2x4}	ρ_{x2x5}	ρ_{x2x6}		0.0	1.0	0.0	0.0	0.0	0.0
ρ_{x3x1}	ρ_{x3x2}	ρ_{x3x3}	ρ_{x3x4}	ρ_{x3x5}	ρ_{x3x6}	=	0.5	0.0	1.0	0.5	0.0	0.0
ρ_{x4x1}	ρ_{x4x2}	ρ_{x4x3}	ρ_{x4x4}	ρ_{x4x5}	ρ_{x4x6}		0.9	0.0	0.5	1.0	0.0	0.0
ρ_{x5x1}	ρ_{x5x2}	ρ_{x5x3}	ρ_{x5x4}	ρ_{x5x5}	ρ_{x5x6}		0.0	0.0	0.0	0.0	1.0	0.0
ρ_{x6x1}	ρ_{x6x2}	ρ_{x6x3}	ρ_{x6x4}	ρ_{x6x5}	ρ_{x6x6}		0.0	0.0	0.0	0.0	0.0	1.0

x_1 - γ_s ; x_2 - c_f ; x_3 - φ_f ; x_4 - γ_f ; x_5 - P ; x_6 - Q ; ρ -coefficient of correlation.

Table 8-2.Summary description of basic input variables.

Basic input variables	Distributions	Mean value μ	Coefficient of variation C_v
B (m)	Deterministic	EC7 DA.2* results	0.00
D (m)	Deterministic	1.00	0.00
γ_s (kN/m ³)	Normal	16.80	0.05
c_f (kN/m ²)	Lognormal	14.00	0.20;0.40;0.60
φ_f (°)	Lognormal	32.00	0.05;0.10;0.15
γ_f (kN/m ³)	Normal	17.80	0.05
P (kN/m)	Normal	370.00	0.10
Q (kN/m)	Normal	70.00	0.25

Regarding the imprecise probabilistic grid-based analysis, the considered correlation coefficients between the basic input variables are either presented in Table 8-1 and the description of basic input variables, with different types of distributions, is summarised in Table 8-3. In addition, the imprecise probabilistic fuzzy-based analysis is performed on uncorrelatedness assumption and Table 8-4 summarises the description of basic input variables, including both random and interval variables. The coefficient of variation is in agreement with published values from the literature.

Table 8-3. Summary description of basic input variables on grid-based analysis.

Basic input variables	Distributions	Mean value μ	Coefficient of variation Cv
B (m)	Deterministic	EC7 DA.2* results	0.00
D (m)	Deterministic	1.00	0.00
γ_s (kN/m ³)	Normal	16.80	0.05
c_f (kN/m ²)	Lognormal	[0.00,40.00]	0.40
ϕ_f (°)	Lognormal	[25.00,40.00]	0.10
γ_f (kN/m ³)	Normal	17.80	0.05
P (kN/m)	Normal	370.00 \vee 1110.00	0.10
Q (kN/m)	Normal	70.00 \vee 210.00	0.25

Lower load combination: $\mu P=370.00$ [kN/m] \wedge $\mu Q=70.00$ [kN/m].

Higher load combination: $\mu P=1110.00$ [kN/m] \wedge $\mu Q=210.00$ [kN/m].

Table 8-4. Summary description of basic input variables on fuzzy-based analysis.

Basic input variables	Distributions	Mean value μ	Coefficient of variation Cv
B (m)	Deterministic	EC7 DA.2* results	0.00
D (m)	Deterministic	1.00	0.00
γ_s (kN/m ³)	Normal	16.80	0.05
c_f (kN/m ²)	Deterministic	0.00	0.00
ϕ_f (°)	Interval	[25.00,40.00]	0.00
γ_f (kN/m ³)	Normal	17.80	0.05
P (kN/m)	Normal	370.00 \vee 1110.00	0.10
Q (kN/m)	Normal	70.00 \vee 210.00	0.25

Lower load combination: $\mu P=370.00$ [kN/m] \wedge $\mu Q=70.00$ [kN/m].

Higher load combination: $\mu P=1110.00$ [kN/m] \wedge $\mu Q=210.00$ [kN/m].

The imprecise probabilistic grid-based analysis is performed in a set of scenarios wherein the shear strength parameters of the foundation soil are jointly implemented as intervals, based on which the characteristic values for design are derived and then combined as mean values in the reliability evaluation. The interval model is further combined with the other uncertain parameters, all of them characterised as random variables including dependencies. In contrast, the imprecise probabilistic fuzzy-based analysis is performed in a scenario wherein the shear strength parameter friction angle of the foundation soil is separately implemented as interval in the format of a conditional analysis and then combined with other probabilistic parameters, considered a null cohesion of the foundation soil. Thereby, mean values are assigned for the determination of characteristic values for each geotechnical parameter, noted that the characteristic load values are considered as 95% fractile values from the considered normal probability distribution and the remaining parameters are deterministic. Monte Carlo simulation (MCS) is further used on testing.

8.3 RESULTS AND DISCUSSION

A cautious estimate of the 95% reliable mean value, Equation (8-2), or a cautious estimate of the 5% fractile value, Equation (8-3), are considered for the determination of characteristic values for each geotechnical parameter by assuming an underlying normal probability distribution for a number of five test results, wherein the coefficient $k_{n,mean}$ is considered as 0.74 and the coefficient $k_{n,low}$ is considered as 1.80 regarding the mean values and known coefficients of variation on Table 8-2:

$$X_k = X_m \left(1 - k_{n,mean} \cdot Cv_x \right) \quad (8-2)$$

$$X_k = X_m (1 - k_{n,low} \cdot C_{v_x}) \quad (8-3)$$

if X_k is the characteristic value; X_m is the mean value; C_{v_x} is the coefficient of variation; and $k_{n,mean}$ and $k_{n,low}$ are statistical coefficients taking into account the sampling, the number of test results, the value affecting the occurrence of the limit state (the mean value or the lowest value, respectively), and the statistical level of confidence required for the assessed characteristic value, expressed by a considered t factor of the Student’s distribution; Equation (8-2) is referenced as case Cv known mean and Equation (8-3) is referenced as case Cv known low.

From Table 8-5 and corresponding Figure 8-2 it is possible to conclude that a resistance partial factor between 1.4 and 1.5 stands for a mean reliability index which satisfies the 3.8 target reliability index, noted that 1.4 is the resistance partial factor generally considered on Eurocode 7, approximately derived from a characteristic safety factor of 2.0. Total satisfactory performance for the combinatory group is only attained for a resistance partial factor of 2.2 and a characteristic safety factor of 3.0. It is demonstrated that under the high variability on the case Cv known mean, satisfactory performance is not attained in most part of the resistance partial factor interval.

From Table 8-6 and corresponding Figure 8-3 it is possible to conclude that total satisfactory performance for the combinatory group is attained in every situation for each resistance partial factor from a mean reliability index of about 5.0 to 6.0 and a characteristic safety factor of about 1.0 to 2.0. Therefore, the characteristic safety factor does not adequately reflect the actual design safety. It is further noted that the estimation of characteristic values is a determinant issue when analysing differences in geotechnical design, as the foundation width B [m] is variable from about 1.0 to 3.5 when considered the case Cv known low.

Table 8-5. Statistics for the foundation width B [m] and the reliability index β on case Cv known mean.

γ_R	F_{s_k}	Case Cv known mean on nine Cv combinations, see forthcoming graphs					
		Interval B [m]	μ B [m]	Cv B	Interval β	μ β	Cv β
1.0	1.3780	[0.8348,1.1831]	0.9983	0.1094	[2.0113,3.3198]	2.5115	0.1921
1.1	1.5158	[0.9042,1.2757]	1.0788	0.1083	[2.2150,3.7870]	2.8216	0.2033
1.2	1.6536	[0.9718,1.3655]	1.1571	0.1073	[2.4038,4.2174]	3.1092	0.2116
1.3	1.7914	[1.0379,1.4527]	1.2333	0.1063	[2.5799,4.6176]	3.3775	0.2179
1.4	1.9292	[1.1024,1.5376]	1.3077	0.1055	[2.7449,4.9903]	3.6288	0.2229
1.5	2.0670	[1.1656,1.6203]	1.3803	0.1046	[2.9003,5.3403]	3.8655	0.2270
1.6	2.2048	[1.2274,1.7011]	1.4512	0.1039	[3.0473,5.6691]	4.0890	0.2303
1.7	2.3426	[1.2881,1.7799]	1.5207	0.1032	[3.1865,5.9804]	4.3009	0.2331
1.8	2.4804	[1.3475,1.8570]	1.5886	0.1026	[3.3189,6.2745]	4.5020	0.2353
1.9	2.6182	[1.4059,1.9325]	1.6553	0.1020	[3.4452,6.5544]	4.6938	0.2373
2.0	2.7560	[1.4632,2.0064]	1.7206	0.1015	[3.5659,6.8207]	4.8768	0.2390
2.1	2.8938	[1.5195,2.0789]	1.7848	0.1009	[3.6815,7.0747]	5.0519	0.2404
2.2	3.0316	[1.5749,2.1501]	1.8478	0.1005	[3.7925,7.3178]	5.2198	0.2417

EC 7 DA.2* results for the determination of the foundation width B [m] statistics.

FORM results for the determination of the reliability index β statistics.

γ_R -resistance partial factor.

F_{s_k} -characteristic safety factor.

μ -mean value.

Cv-coefficient of variation.

$F_{s_k} = \gamma_E \cdot \gamma_R$, γ_E =effect actions partial factor=1.3780.

Table 8-5 and Table 8-6 summarise the calculation interval for the foundation width B [m] and the reliability index β by the application of the first order reliability method (FORM). On this parametric study it is investigated the influence of variability of soil shear strength parameters on resistance partial factor design for the various cases considered, which include lower and higher variability scenarios as showed on Figure 8-2 and Figure 8-3. A cautious estimate of the mean value

or of the lowest value affecting the occurrence of the limit state is considered for the determination of characteristic values for each geotechnical parameter, respectively case Cv known mean on Table 8-5 and Figure 8-2 or case Cv known low on Table 8-6 and Figure 8-3.

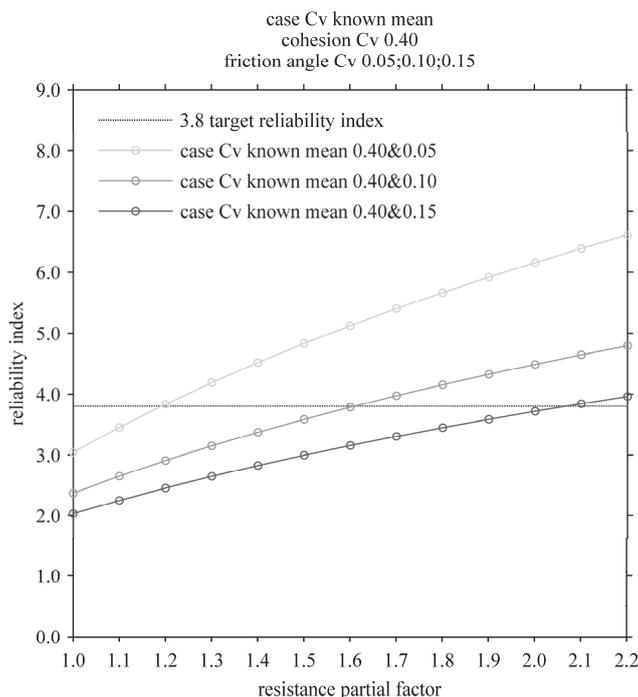


Figure 8-2. Influence of variability of soil shear strength parameters on resistance partial factor design for the cases Cv known mean 0.40&0.05 and 0.40&0.10 and 0.40&0.15.

Table 8-6. Statistics for the foundation width B [m] and the reliability index β on case Cv known low.

γ_R	F_{Sk}	Case Cv known low on nine Cv combinations, see forthcoming graphs					
		Interval B [m]	μ B [m]	Cv B	Interval β	μ β	Cv β
1.0	1.3780	[1.1019,2.5880]	1.7428	0.2664	[4.0131,5.6175]	4.6624	0.1129
1.1	1.5158	[1.1898,2.7554]	1.8675	0.2627	[4.3119,6.0149]	4.9956	0.1175
1.2	1.6536	[1.2752,2.9160]	1.9879	0.2594	[4.5891,6.3809]	5.3031	0.1221
1.3	1.7914	[1.3582,3.0707]	2.1043	0.2565	[4.8478,6.7204]	5.5885	0.1264
1.4	1.9292	[1.4392,3.2200]	2.2171	0.2539	[5.0904,7.0366]	5.8549	0.1304
1.5	2.0670	[1.5181,3.3644]	2.3267	0.2515	[5.3056,7.3331]	6.1047	0.1340
1.6	2.2048	[1.5952,3.5044]	2.4333	0.2494	[5.4827,7.6118]	6.3398	0.1373

EC 7 DA.2* results for the determination of the foundation width B [m] statistics.

FORM results for the determination of the reliability index β statistics.

γ_R -resistance partial factor.

F_{Sk} -characteristic safety factor.

μ -mean value.

Cv-coefficient of variation.

$F_{Sk}=\gamma_E \cdot \gamma_R$, γ_E =effect actions partial factor=1.3780.

Compared the case Cv known mean and the case Cv known low on equal characteristic safety factor, it is concluded that a higher foundation width B [m] stands for a higher reliability index. The interval derived for the reliability index is as well determined by the shear strength variability. In this way, robustness in the context of Eurocode 7 partial factor design is not achieved on a unique resistance partial factor.

Afterwards and considered the Eurocode 7 Design Approach DA.2*, results for the determination of the foundation width B [m] statistics are brought together. Regarding safety, a

minimum 0.6 foundation width B [m] is considered from the construction industry practice on the type of geotechnical engineering structure. Thereafter, FORM results are gathered within the specified grid of model cases for the determination of the reliability index statistics. As illustration, Figure 8-4 represents FORM results in the reliability index three-dimensional joint view to safety assessment for a resistance partial factor level of 3.0, considered the interval scenario [0.0,40.0] for cohesion [kN/m²] and [25.0,40.0] for friction angle [°] on lower load combination, median case selected among the forthcoming cases.

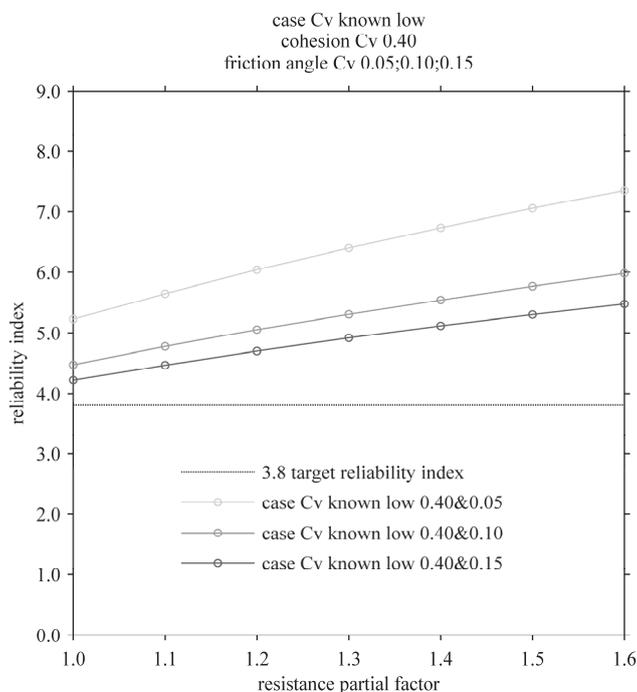


Figure 8-3. Influence of variability of soil shear strength parameters on resistance partial factor design for the cases Cv known low 0.40&0.05 and 0.40&0.10 and 0.40&0.15.

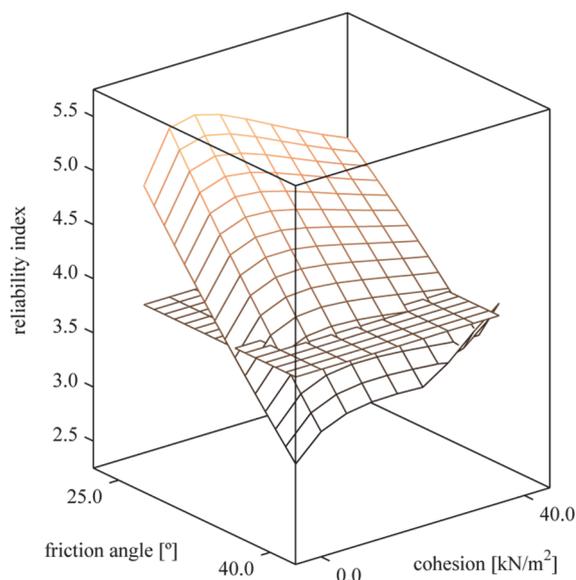


Figure 8-4. Reliability index three-dimensional joint view to safety assessment for a resistance partial factor level of 3.0, considered the interval scenario [0.0, 40.0] for cohesion [kN/m²] and [25.0, 40.0] for friction angle [°] on lower load combination.

In this sketch it is noted the horizontal position of the 3.8 target reliability level so that it is observable a curved surface wherein part is below the 3.8 target reliability level. The lower left corner appears as the critical and the lower right corner is shaped due to the minimum 0.6 foundation width B [m]. It is further noted that the maximum reliability index corresponds to a lower cohesion and to a minimum friction angle characteristic values. Thereby, Figure 8-5 and Figure 8-6 represent FORM results for the 3.8 target reliability index ISOLINES, considered the interval scenario $[0.0, 40.0]$ for cohesion $[\text{kN}/\text{m}^2]$ and $[25.0, 40.0]$ for friction angle $[\text{°}]$. The individual resistance partial factor is detailed in three cases, 2.5 and 3.0 and 3.5, for two load combinations, the higher on a 3.0 incremental ratio. The ISOLINES concept is detailed hereafter.

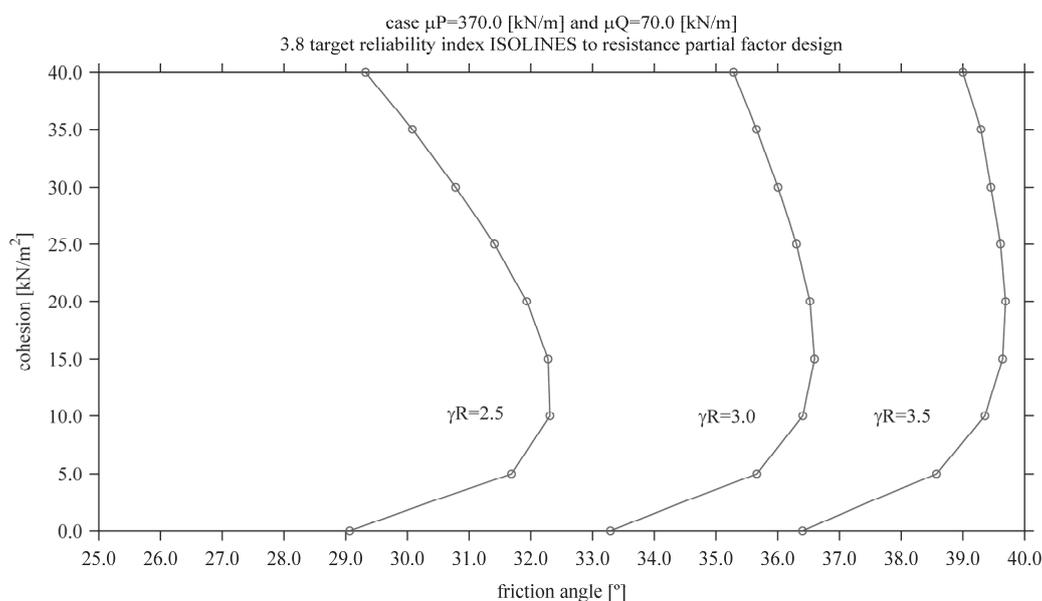


Figure 8-5. Imprecise probabilistic grid-based approach on variable foundation width [m] for each resistance partial factor level of 2.5, 3.0, 3.5, considered the interval scenario $[0.0, 40.0]$ for cohesion $[\text{kN}/\text{m}^2]$ and $[25.0, 40.0]$ for friction angle $[\text{°}]$ on lower load combination.

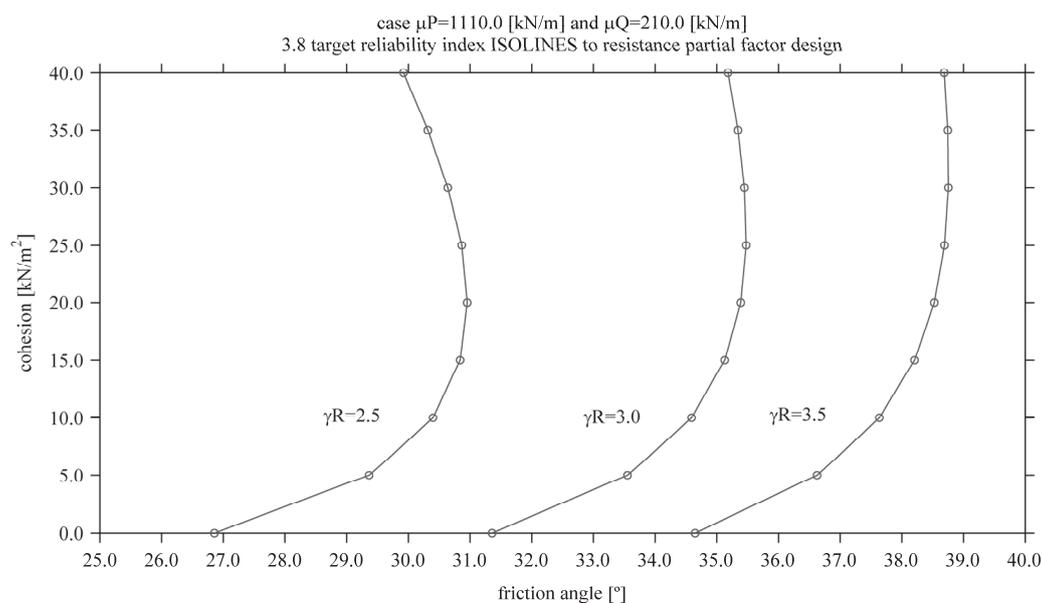


Figure 8-6. Imprecise probabilistic grid-based approach on variable foundation width [m] for each resistance partial factor level of 2.5, 3.0, 3.5, considered the interval scenario $[0.0, 40.0]$ for cohesion $[\text{kN}/\text{m}^2]$ and $[25.0, 40.0]$ for friction angle $[\text{°}]$ on higher load combination.

The ISOLINES in Figure 8-5 and Figure 8-6 represent the set of shear strength parameters that is capable to satisfy the required 3.8 target reliability index for each resistance partial factor in detail, noted that every combination of shear strength parameters on the left of each curve falls on the safe side. Thus, whenever a variety of possibilities instead of one clear model are advanced, a grid-based analysis may be pursued to identify a relative importance within a grid of model cases in a decision-making technique which reflects the sensitivity with respect to the load and resistance design. It is noted that this particular case study is based on the assumption that every grid element is eligible on uniform possibility on weighting. The shear strength parameters on the grid area may be combined and weighted for differentiation through the segmentation of sets into new subsets, analysed any background information on the ground nature and considered the risk tolerance of the geotechnical engineer, as well as any other economic issues related to design feasibility. Complementary knowledge may be provided by a number of references on foundation engineering and geotechnical site investigation or handbooks of design tables. It is noted that this particular case study is calculated for demonstration only for normal variabilities.

Table 8-7 and Table 8-8 summarise the statistics for the foundation width B [m] and reliability index β , respectively on lower and higher load combination. It is noted that the calculation of the weighted resistance partial factor is based on the safe cases in every cluster delimited by each curve drawn respectively at Figure 8-5 and Figure 8-6, and not on the safe cases for the calculation of the safe percentage. Furthermore, Table 8-9 details the resistance partial factor design corresponding to the 3.8 target reliability index calculated from MCS results on four cases corresponding to the lower and higher load combination on 0.0 cohesion [kN/m²] and 25.0 or 40.0 friction angle [°].

Table 8-7. Statistics for the foundation width B [m] and reliability index β , considered the interval scenario [0.0, 40.0] for cohesion [kN/m²] and [25.0, 40.0] for friction angle [°] on lower load combination.

$W\gamma_R$ vs F_{Sk}	$I\gamma_R$ vs F_{Sk}	Interval B [m]	Interval β	Safe cases	Safe percentage [%]
3.0 vs 4.1340	2.5 vs 3.4450	[0.6000,3.7821]	[2.6055,4.7294]	59/144	≈40
	3.0 vs 4.1340	[0.6000,4.2263]	[2.9966,5.4132]	102/144	≈70
	3.5 vs 4.8230	[0.6000,4.6364]	[3.3363,5.9990]	131/144	≈90
	4.3 vs 5.9254	[0.6651,5.2377]	[3.8023,6.7906]	144/144	≈100

EC7 DA.2* results for the determination of the foundation width B [m] statistics.

FORM results for the determination of the reliability index β statistics.

$W\gamma_R$ -weighted resistance partial factor; $I\gamma_R$ -individual resistance partial factor.

F_{Sk} -characteristic safety factor.

$W\gamma_R \approx [2.5 \cdot 59 + 3.0 \cdot 43 + 3.5 \cdot 29 + 4.3 \cdot 13] / [144] \approx 3.0$.

$F_{Sk} = \gamma_E \cdot \gamma_R$, γ_E =effect actions partial factor=1.3780.

Table 8-8. Statistics for the foundation width B [m] and reliability index β , considered the interval scenario [0.0, 40.0] for cohesion [kN/m²] and [25.0, 40.0] for friction angle [°] on higher load combination.

$W\gamma_R$ vs F_{Sk}	$I\gamma_R$ vs F_{Sk}	Interval B [m]	Interval β	Safe cases	Safe percentage [%]
3.2 vs 4.4096	2.5 vs 3.4450	[1.0726,7.2200]	[2.4658,4.4920]	48/144	≈35
	3.0 vs 4.1340	[1.2469,8.0020]	[2.8347,5.1364]	92/144	≈65
	3.5 vs 4.8230	[1.4127,8.7221]	[3.1551,5.6880]	119/144	≈85
	4.7 vs 6.4766	[1.7825,10.2659]	[3.7897,6.7727]	144/144	≈100

EC7 DA.2* results for the determination of the foundation width B [m] statistics.

FORM results for the determination of the reliability index β statistics.

$W\gamma_R$ -weighted resistance partial factor; $I\gamma_R$ -individual resistance partial factor.

F_{Sk} -characteristic safety factor.

$W\gamma_R \approx [2.5 \cdot 48 + 3.0 \cdot 44 + 3.5 \cdot 27 + 4.7 \cdot 25] / [144] \approx 3.2$.

$F_{Sk} = \gamma_E \cdot \gamma_R$, γ_E =effect actions partial factor=1.3780.

Additional FORM constraints to relate interval and probabilistic variables are expendable on uncorrelatedness assumption as it is not possible to attain precision on a verifiable basis in the big data analysis on the high reliability space. The foundation width B [m] is as well recalculated for the EC7 DA.2* resistance partial factor level of 1.4 on every case corresponding to the 0.0 cohesion [kN/m²] and fuzzy interval scenario [25.0, 40.0] for friction angle [°], see on Figure 8-7 FORM results on

lower and higher load combination. Thereafter, the fuzzy reliability calculation evinces now a 10.0 quasi uniform reliability level. Considered the Eurocode 7 Design Approach DA.2*, results for the determination of the foundation width B [m] for the critical case which considers the pair 0.0 cohesion [kN/m²] and 40.0 friction angle [°] are determined for resistance partial factors of 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0 on lower and higher load combination. Thereafter, FORM results are gathered within the interval scenario [25.0, 40.0] for friction angle [°] for the determination of the reliability index statistics. Thereby, Figure 8-8 and Figure 8-9 represent FORM results for the imprecise probabilistic fuzzy-based approach wherein the reliability index interval is related to the friction angle subsets for every resistance partial factor. On this fuzzy reliability calculation the friction angle is considered a deterministic parameter combined with the other uncorrelated probabilistic variables. The smooth curved lines at Figure 8-8 and Figure 8-9 denote the friction angle demand to attain the indicative reliability marks. They are sketched for every resistance partial factor, so that the assignment of the critical case which considers the pair 0.0 cohesion [kN/m²] and 40.0 friction angle [°] corresponds to the high reliability space. It is remarked that each line is developed for a unique foundation width B [m] design, the closeness of the case 1.0 at Figure 8-8 due to the minimum 0.6 foundation width B [m].

Table 8-9. Resistance partial factor design corresponding to the 3.8 target reliability index.

I_{γ_R} vs F_{S_k}			
0.0 cohesion [kN/m ²] and 25.0 friction angle [°]		0.0 cohesion [kN/m ²] and 40.0 friction angle [°]	
Lower load combination	Higher load combination	Lower load combination	Higher load combination
2.1514 vs 2.9646	2.3423 vs 3.2277	4.2985 vs 5.9233	4.7245 vs 6.5104

MCS results from 5e6 simulations.

I_{γ_R} -individual resistance partial factor.

F_{S_k} -characteristic safety factor.

$F_{S_k} = \gamma_E \cdot \gamma_R$, γ_E =effect actions partial factor=1.3780.

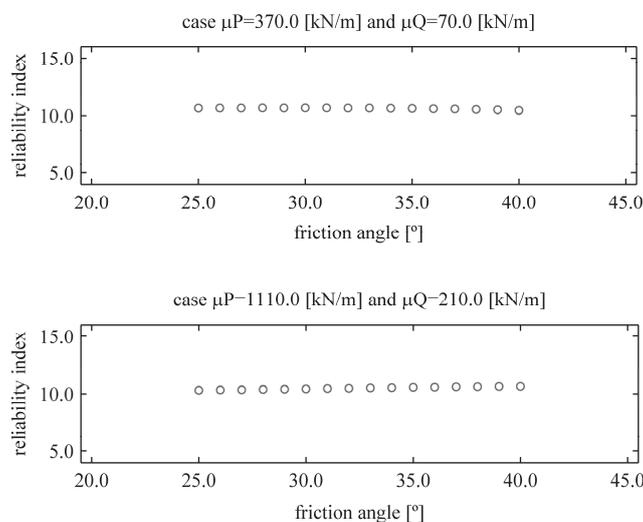


Figure 8-7. Imprecise probabilistic fuzzy-based approach on variable foundation width [m] for the EC7 DA.2* resistance partial factor level of 1.4, considered the interval scenario [25.0, 40.0] for friction angle [°] on lower and higher load combination.

The limit state imprecise probabilistic analysis is then interpreted altogether with a limit state imprecise interval analysis for bearing capacity. Thus, the limit state charts to safety assessment are separately sketched for the cases cohesion and friction angle interval scenario wherein the random variables are bounded on different levels of probability, see Figure 8-10 for the case friction angle interval scenario. Distinct levels of credibility are used to sketch the lines which express the limit state bounds. It is possible to search for a credibility level which ensures no failure regardless of the parameter value on the horizontal axis and conversely, it is possible to find the threshold parameter which ensures no failure for a given credibility level and then to proceed with proper ground

investigation and testing or improvement, see in the chart the circle crossing the zero limit state boundary and the 0.9900 credibility level line. In decision making, this valuable approach may be extended by numerical analysis for high dimensional cases with several indecision variables in simultaneous, see Figure 8-11 for the limit state three-dimensional joint view to safety assessment considered simultaneously the interval variables cohesion and friction angle. In the multivariate case the considered probability level prescribes a credible region in the hyperspace, then the imprecise interval analysis complies with a mixed set of probabilistic and nonprobabilistic interval models wherein different bounding measures may be applied in order to find the limit state lower and upper bounds in different scenarios.

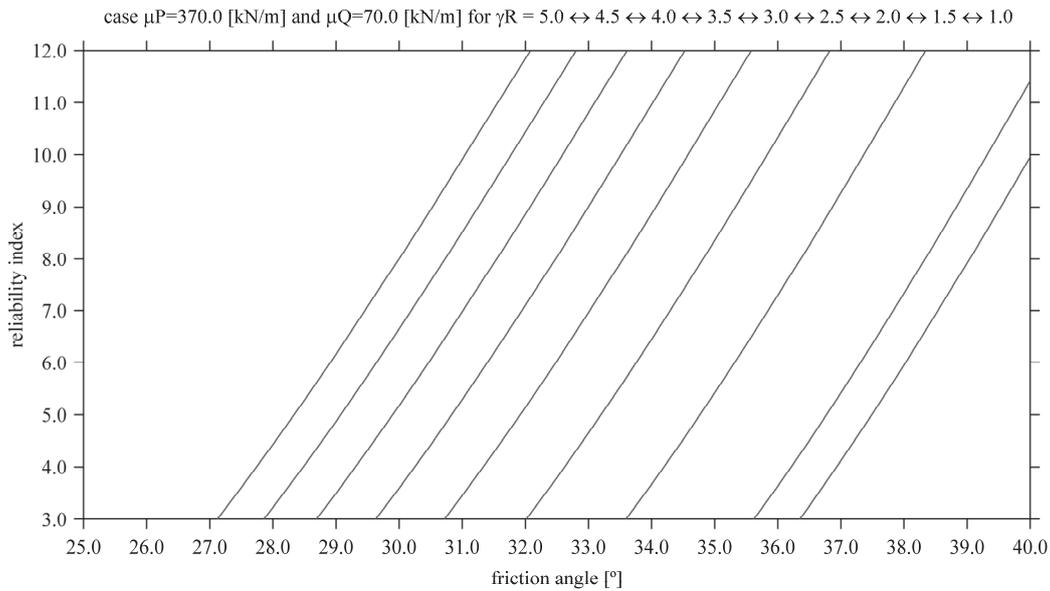


Figure 8-8. Imprecise probabilistic fuzzy-based approach on constant foundation width [m] for each resistance partial factor level of 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, considered the interval scenario [25.0, 40.0] for friction angle [°] on lower load combination.

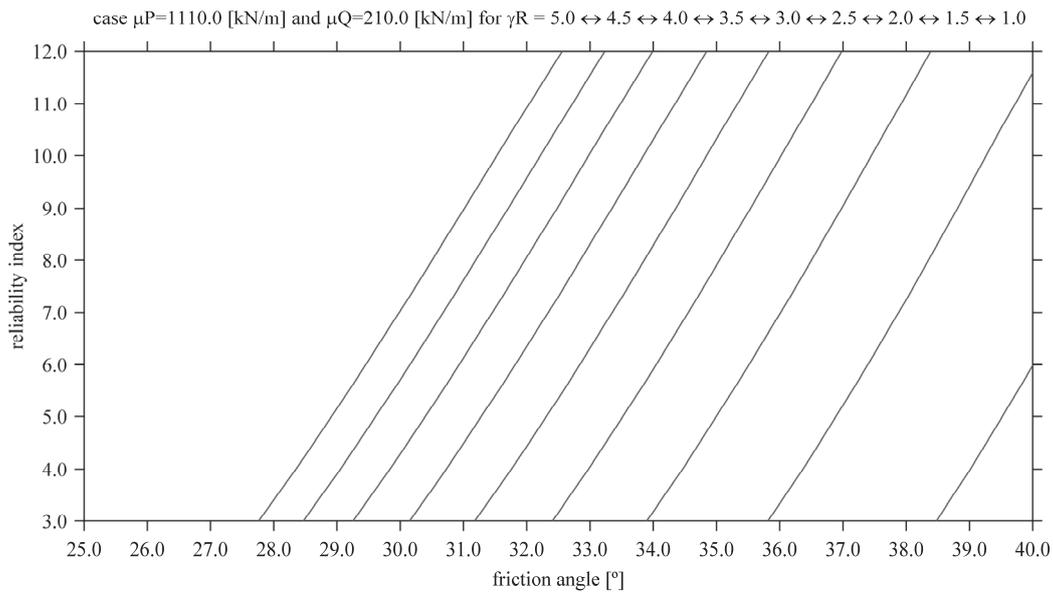


Figure 8-9. Imprecise probabilistic fuzzy-based approach on constant foundation width [m] for each resistance partial factor level of 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, considered the interval scenario [25.0, 40.0] for friction angle [°] on higher load combination.

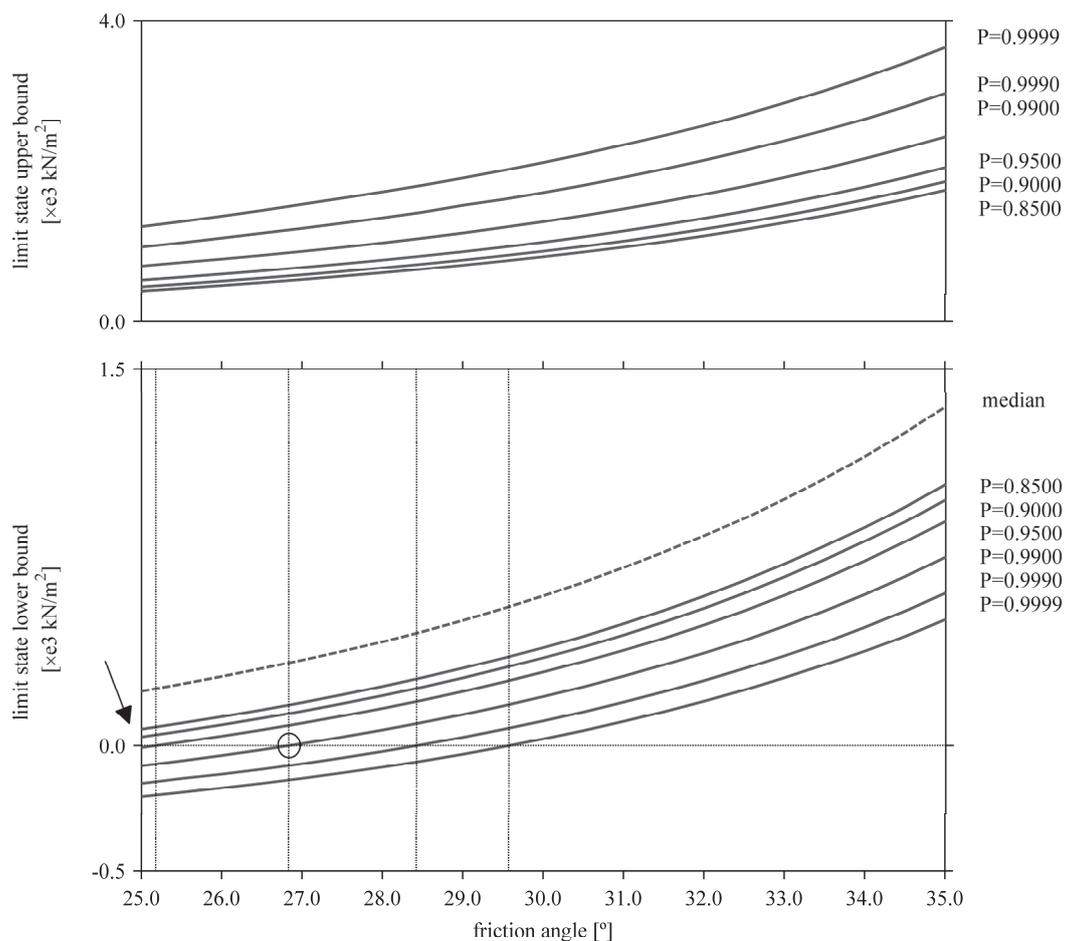


Figure 8-10. Limit state chart to safety assessment for the case friction angle interval scenario.

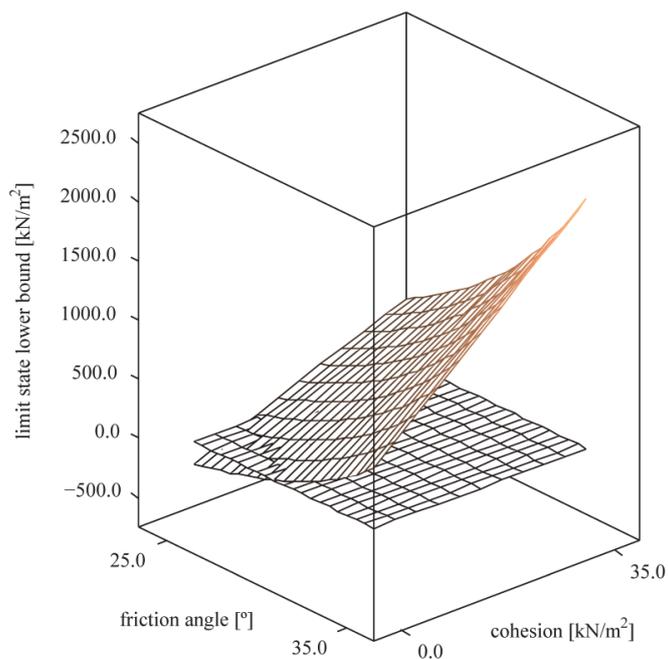


Figure 8-11. Limit state three-dimensional joint view to safety assessment for a 0.9900 probability level.

On this framework, Figure 8-11 displays the limit state three-dimensional joint view to safety assessment for a 0.9900 probability level corresponding to a satisfactory level of reliability. A three-dimensional representation of the zero limit state reference surface is crossed in one corner by the limit state lower bound surface constructed from a joint assumption of the values of the interval variables cohesion and friction angle. According to the results on Figure 8-11, unsafe coordinates are limited to a cohesion parameter on a value between 7.0 kN/m² and 10.5 kN/m² or to a friction angle parameter on a value between 29.0 ° and 30.0 °, considered the two worst combination cases corresponding to a friction angle of 25.0 ° or to a cohesion of 0.0 kN/m², respectively. On this particular case study it is clearly shown that small variations in the friction angle input are very influential in that the median and the imprecise lower bound of probability correspond to a boundless immeasurable reliability.

8.4 SUMMARY

Considered the underlying motivation on the endeavour for safety, approaches to robust design are nowadays discussed on the calculation of resistance factors capable to maintain a more uniform reliability level over a range of design parameters. As information is not certain but rather imprecise, a sensitivity analysis may be pursued through probabilistic grid-based and fuzzy-based approaches on multiple intervals. The imprecise probabilistic analysis is then a supplementary element which enriches the variety of models to be combined with the traditional overview in improved adaptability. The admissibility of imprecision in beliefs is the primary difference in motivation between imprecise probabilistic and robust Bayesian approaches, beliefs are imprecise or there exists a prior that captures the true beliefs, although it may be hard to identify this distribution. Imprecise probabilistic approaches are robust whenever insensitive to small deviations from the assumed probabilistic models.

From the imprecise probabilistic grid-based analysis applied to this particular case study on lower and higher load combination, a safe percentage between 70 and 75 is attained whenever considered a weighted resistance partial factor calculated on uniform possibility. Thus, a meaningful interpretation on a high dimensional space is based on the joint analysis of multiple cases instead of a lower and upper probabilistic evaluation, noted that the reliability index is variable within a limited expectation.

From the imprecise probabilistic fuzzy-based analysis applied to this particular case study on lower and higher load combination, the safety concept emerges realigned to a new vision on a quasi-uniform reliability level highly attained along the friction angle interval, whenever the shear strength parameters are truly assigned. Thus, it is evinced the capability to approach a uniform reliability level on the parametric range of interest, considered the extension of geotechnical robustness to partial factor design.

One motivation for adopting imprecise probabilistic approaches consists on the safety evaluation in the scope of the shear strength parametric change by weathering on real world problems such as the probabilistic interpretation of mine pillar capacity. On the imprecise probabilistic approaches multiple resistance partial factors may be clearly related to the determination of characteristic values. Expressed simply by intervals, geotechnical parameters on scarce probabilistic information are assigned based on experience. A meaningful interpretation on a high dimensional space is based on the joint analysis of multiple cases instead of a lower and upper probabilistic evaluation. The limit state imprecise probabilistic analysis may be interpreted altogether with a limit state imprecise interval analysis. This sensitivity analysis may provide meaningful results for safety-based decision in ground investigation and testing or improvement and the partial factor design may be discussed on the basis of distinct levels of credibility.

8.5 REFERENCES

- Beer, M. (2009). “Non-traditional prospects in the simultaneous treatment of uncertainty and imprecision.” In: *Soft computing in civil and structural engineering*. Edited by Topping, B.H.V., and Tsompanakis, Y., Saxe-Coburg Publications, 247-266.
- Ching, J., and Phoon, K.K. (2013). “Quantile value method versus design value method for calibration of reliability-based geotechnical codes.” *Structural Safety*, 44, 47-58.

- Ching, J., Phoon, K.K., Chen, J.R., and Park, J.H. (2013). "Robustness of constant load and resistance factor design: factors for drilled shafts in multiple strata." *Journal of Geotechnical and Geoenvironmental Engineering*, 139(7), 1104-1114.
- Fenton, G.A. (2013). "Geotechnical design code development in Canada." In: *Modern geotechnical design codes of practice: implementation, application and development*. Edited by Arnold, P., Fenton, G.A., Hicks, M.A., Schweckendiek, T., and Simpson, B., IOS Press, 277-294.
- Frank, R., Bauduin, C., Driscoll, R., Kavvas, M., Ovesen, N.K., Orr, T., and Schuppener, B. (2004). *Designers' guide to EN 1997-1: Eurocode 7: geotechnical design-general rules*. Thomas Telford.
- Lesny, K. (2013). "Implementation of Eurocode 7 in German geotechnical design practice." In: *Modern geotechnical design codes of practice: implementation, application and development*. Edited by Arnold, P., Fenton, G.A., Hicks, M.A., Schweckendiek, T., and Simpson, B., IOS Press, 46-59.
- Look, B. (2007). *Handbook of geotechnical investigation and design tables*. Taylor & Francis.
- Marques, S.H., Beer, M., Gomes, A.T., and Henriques, A.A. (2015a). "Limit state imprecise probabilistic analysis in geotechnical engineering." In: *ISGSR 2015, Geotechnical Safety and Risk, Proc. of the 5th International Symposium on Geotechnical Safety and Risk*, Rotterdam, The Netherlands, Kingdom of the Netherlands, IOS Press, 269-274.
- Marques, S.H., Beer, M., Gomes, A.T., and Henriques, A.A. (2015b). "Limit state imprecise interval analysis in geotechnical engineering." In: *ISGSR 2015, Geotechnical Safety and Risk, Proc. of the 5th International Symposium on Geotechnical Safety and Risk*, Rotterdam, The Netherlands, Kingdom of the Netherlands, IOS Press, 380-385.
- Orr, T. (2013). "Implementing Eurocode 7 to achieve reliable geotechnical designs." In: *Modern geotechnical design codes of practice: implementation, application and development*. Edited by Arnold, P., Fenton, G.A., Hicks, M.A., Schweckendiek, T., and Simpson, B., IOS Press, 72-86.
- Phoon, K.K., and Ching, J. (2013). "Can we do better than the constant partial factor design format?" In: *Modern geotechnical design codes of practice: implementation, application and development*. Edited by Arnold, P., Fenton, G.A., Hicks, M.A., Schweckendiek, T., and Simpson, B., IOS Press, 295-310.
- Phoon, K.K., Ching, J., and Chen, J.R. (2013). "Performance of reliability-based design code formats for foundations in layered soils." *Computers & Structures*, 126, 100-106.
- Schuppener, B. (2013). "The safety concept in German geotechnical design codes." In: *Modern geotechnical design codes of practice: implementation, application and development*. Edited by Arnold, P., Fenton, G.A., Hicks, M.A., Schweckendiek, T., and Simpson, B., IOS Press, 102-115.
- Simpson, B. (2013). "British choices of geotechnical design approach and partial factors for EC7." In: *Modern geotechnical design codes of practice: implementation, application and development*. Edited by Arnold, P., Fenton, G.A., Hicks, M.A., Schweckendiek, T., and Simpson, B., IOS Press, 116-127.

Chapter 9 Robustness in Geotechnical Design

Lead discussor:

Brian Simpson

brian.simpson@arup.com

Discussors (alphabetical order):

Hongwei Huang, Charnghein Juang, Sónia H. Marques, Bernd Schuppener,

Timo Schweckendiek, Paul Vardanega, Norbert Vogt

9.1 INTRODUCTION

The purpose of the Working Group was to examine how sufficient robustness can be ensured in geotechnical designs, using reliability analysis or other safety formats. The members of the group first set out to define the term “robustness” in a relevant way and then exchanged emails and papers to develop an understanding of how it can be provided in geotechnical design and codes of practice.

Two distinctly different types of robustness are considered in this chapter. Therefore, following the section on Definitions below, the chapter is divided into two parts, each considering one definition of robustness.

9.2 DEFINITIONS

9.2.1 Introduction

The term “robustness” can take several different meanings. The issue of concern to designers and to codes of practice is the robustness of a civil engineering construction, usually in its final form but also during the process of construction. This is therefore the subject of this chapter.

Two principal types of robustness have been identified by the group:

- (1) The ability of the final design to accommodate events and actions that were not foreseen or consciously included in design. This is considered in Part A of the chapter below.
- (2) The sensitivity of the final design to variations of the known parameters within their anticipated range of uncertainty. This is considered in Part B of the chapter below.

9.2.2 Accommodating what is unforeseen

This definition of robustness relates, in particular, to the ability of the construction to withstand without failure events and actions that were not foreseen or consciously included in design. Although the precise nature of such actions may be unknown to the designer, their magnitude can be considered: society expects that a construction will be able to withstand moderate unforeseen events and actions, but probably not extremely severe ones. A design that produces such a construction can be termed a “robust design”.

A concise definition is given by ISO 2394, which equates robustness to “damage insensitivity”. This will be taken to be the basic definition used in Part A of this report:

Ability of a structure to withstand adverse and unforeseen events (like fire, explosion, and impact) or consequences of human errors without being damaged to an extent disproportionate to the original cause (ISO 2394:2014, 2.1.46).

An alternative definition, with the same basic meaning, could help designers to understand the degree of robustness required:

Ability of a structure to withstand adverse events that are unforeseen but of a magnitude such that society will expect that our designs can accommodate them, having tolerance against

mistakes within the design process and during construction.

9.2.3 Local damage and progressive failure

The term robustness is often applied to a complete structure rather than to an individual element of it.

For example, CEN (2016) Practical definition of structural robustness vDraft, gives a definition of structural robustness:

Structural robustness is an attribute of a structural concept, which characterizes its ability to limit the follow-up indirect consequences caused by the direct damages (component damages and failures) associated with identifiable or unspecified hazard events (which include deviations from original design assumptions and human errors), to a level that is not disproportionate when compared to the direct consequences these events cause in isolation.

Robustness is often linked to the ability to prevent progressive failure, which could lead to damage disproportionate to cause (eg COST (2011) Structural robustness design for practising engineers). This is probably consistent with strict limit state definitions in which ultimate limit state (ULS) is a state of danger, but as a practical design expedient ULS is often considered as only localised failure, not necessarily dangerous in itself. EN 1990 3.3(3) is relevant to this: “States prior to structural collapse, which, for simplicity, are considered in place of the collapse itself, may be treated as ultimate limit states.”

Val (2006), discussing robustness of framed structures, provides a definition similar to that of ISO 2394, and then offers as an alternative:

The robustness of a structure can be defined as ability of the structure to withstand local damage without disproportionate collapse, with an appropriate level of reliability.

9.2.4 Resilience

Robustness can be distinguished from “resilience”, which refers to the ability of a structure to be recovered after it has failed. On the other hand, a complete structure, or a system such as a metro system, might be considered robust if its members are all resilient, so that local failures can be repaired without failing the complete system (Huang et al 2016).

9.3 PART A—ACCOMMODATING EVENTS AND ACTIONS THAT WERE NOT FORESEEN IN DESIGN

9.3.1 Events and actions relevant to robust design

In most design processes, “lead variables” are identified and the possibility that they might adopt extreme values, or occur in adverse combinations, is considered in some way. Lead variables are usually actions (loads), material strengths and component resistances. However, most designs are also affected by a large number of “secondary variables”, which the design is expected to accommodate.

Robustness relates to the ability of a construction to withstand events and actions that were not foreseen or consciously included in design, in effect because they were considered “secondary”. These have to be judged in their context. For example, in a building structure if a heating engineer puts a 150mm hole through a wall, it would be unacceptable for the wall to fail; however, if the same hole were put through a 250mm column the heating engineer, not the column designer, could be liable for the failure that ensued.

The definition of robustness given in ISO2394, in common with EN1990, mentions as examples fire, explosion, impact and human errors. Human errors occur both in design and construction, the latter often resulting in geometric inaccuracies in the construction. In a geotechnical context, other secondary variables could include sedimentation or erosion around a structure in water, excavation of small trenches etc., or of the ground above a structure relying on the weight of ground, disturbance caused by burrowing animals, unidentified loading above retaining walls, and vandalism of various kinds.

If these events are very large, it might be judged that the designer should have allowed for them, or they might lead to successful insurance claims or prosecution of the perpetrators. However, where

they are only moderate in magnitude, clients and society reasonably expect that they will not cause significant problems to constructions. In this respect, although the events themselves are unforeseen at the time of design, the magnitude that a design must be able to accommodate is understood, at least roughly. For example, whilst all structures may be expected to have reasonable robustness against vandalism, ability to resist more severe acts of terrorism is only required in the specifications of more exceptional structures.

In reliability work, the term “black swan” is used to describe something that was unforeseeable and that has an extreme impact - https://en.wikipedia.org/wiki/Black_swan_theory . The implication is that nobody could have prepared for the disaster that was caused, and society would accept that no designer could be blamed. Robustness relates to events that are also unforeseen but are of smaller magnitude, such that society will expect that robust designs can accommodate them. It might be helpful to think of these as grey swans – signets – they are neither black nor white and somewhat smaller.

9.3.2 Ensuring robustness in various design formats

9.3.2.1 Prescriptive measures relevant to robustness

Studies of robustness in structural design highlight two important prescriptive measures: provision of redundancy or “alternative paths”, and “tying the structure together”. In the “alternative path” method individual members are removed in the analysis to prove robustness of the structure. Val (2006) notes:

It is stressed that the removal of a single vertical load bearing element "is not intended to reproduce or replicate any specific abnormal load or assault on the structure". Rather, member removal is simply used as a "load initiator" and serves as means to introduce redundancy and resiliency into the structure.

As a geotechnical example of this, Simpson et al (2008) argued that the Nicoll Highway collapse in Singapore probably would not have occurred, despite human errors, if the design had included a check for loss of a single strut in the excavation; this was a requirement in the Singapore code at the time of design.

As with other issues related to safe design, checking, review and supervision of design and construction are extremely valuable. In some cases, these processes may suggest that some “unforeseen” events and actions should be classified as “foreseeable” and consciously included in the design process.

9.3.2.2 Use of partial factor methods

In this chapter, the term “partial factor methods” will be taken to include all safety formats in which factors of safety are spread among several variables. The variables may include actions (loads), effects of actions such as internal forces derived in calculations, material strengths, and resistances of structural components (such as bending capacity) or of bodies of ground (such as bearing resistance). Thus all the “Design Approaches” of Eurocode 7 and all LRFD formats are included as “partial factor methods”. Some of the partial factors may be “model factors”.

Many studies have been carried out to derive values for partial factors using reliability analysis (eg Foye et al 2006, Schweckendiek et al 2012). However, in practice, almost all values used in modern codes of practice have been derived by calibration against previous experience of successful design. Sometimes, further reliability studies have been used to provide additional justification. The disadvantage of calibration processes is that the “successful” designs demonstrated adequate success in terms of both ultimate and serviceability limit states and also with regard to robustness. So it is difficult, if not impossible, to determine which of these criteria actually required the factors used. However, calibration against existing experience shows that the factors adopted have provided, at least, a level of robustness that has been found to be adequate.

EC7 notes one particular aspect of robustness, without using that word: the accommodation of small geometric variations. For these it says:

The partial action and material factors (γ_F and γ_M) include an allowance for minor variations in geometrical data and, in such cases, no further safety margin on the geometrical data should be required. (EC7, 2.4.6.3(1))

CEN (2014) Robustness in Eurocodes notes that: “The national partial safety factors are also expected to cover a (part of) the effects of errors in design and execution. (Section 2, page 4).

It may be concluded, therefore, that the use of partial factor methods with values derived by calibration against existing successful experience, is a valid approach to provision of adequate robustness. Their values are roughly aligned with typical coefficients of variation of the lead parameters. As will be noted below, this is probably an optimal strategy.

9.3.2.3 Direct use of reliability methods

The potential benefit of reliability methods over partial factor methods is that they can take account directly of the real uncertainty of the lead variables, for which data may be available. This would allow the safety of designs to be gauged by a reliability index, β , which, in principle, is related to the probability of failure, intended to be very low. Reliability methods are generally more complicated to implement than partial factor methods, so designers and codes of practice are only likely to adopt them if they are shown to have clear advantages.

The Working Group has not been able to suggest practicable methods of accommodating robustness (type (a) in 2.1, as discussed in 2.2) in reliability based design. It is possible that a major study of civil engineering failures, of large and small magnitude, might provide a database that could be used as an input to reliability studies. This would give, for example, objective data on the occurrence and significance of human errors in design. However, an immediate problem arises that in many cases the detailed analysis of failures is confidential to legal proceedings, so accumulation of reliable data would be very difficult.

It might be possible to calibrate reliability methods against past experience in the same way that partial factor methods have been calibrated. This could mean that values of the reliability index β , which relates to the probability of the lead variables dominating the design, could be chosen so as to reproduce previous successful designs, which are considered to have sufficient robustness. Unfortunately, this would lose the logical connection between reliability index, probability of failure and the actual uncertainty of the lead variables.

It was noted above that while actions and events for which robustness is needed are not identified at the time of design, their magnitude is roughly determined by what is acceptable to society. Because they are independent of the lead variables, they are also independent of the range of uncertainty of those variables. This means that the magnitudes of unforeseen actions and events, for which robustness is required, cannot be measured on the same scale as the uncertainties of the lead variables. Hence, simply designing for larger β might not achieve what is required.

Consider, for example, a situation in which the coefficients of variation of the lead variables are considered to be very small. In that case, a large value for β could be achieved with little change to the design, and no significant robustness to meet unforeseen actions and events. In this respect, the use of partial factors with values roughly aligned to typical coefficients of variation of the lead variables, but not tuned specifically for individual designs, appears to be advantageous. This situation commonly occurs, for example, in the design of water reservoirs for with the maximum loading, when the reservoir is full, is expected to occur but cannot be exceeded. A simple reliability approach could lead to the conclusion, in effect, that no margin of safety need be added to the forces calculated from the water load, but this would give no allowance for robustness. Evolution Group 9 of Eurocode 7, reporting on design for water pressures, therefore concluded that partial factors should be applied to the action effects from water pressures calculated in structures; treatment of the forces and stresses in the ground caused by water pressures depends on how the ground strength or resistance is being factored.

9.3.2.4 Use of reliability methods to determine partial factors for inclusion in standards

Reliability methods can be used as a means of informing the choice of values for partial safety factors in standards. This avoids the need for skill in reliability theory on the part of designers. An example related to partial factors used in the design of flood defences in the Netherlands is discussed by Schweckendiek et al (2012).

The process of a rigorous reliability exercise as part of the design development of such major structures, requiring careful discussion among experts of several disciplines, is considered to have benefits in raising issues that might normally be overlooked and encouraging proper investigation of the parameters controlling the design. It could be that this process will, in itself, improve robustness against “unforeseen” events and actions by forcing more of them to be explicitly foreseen and accommodated in the design. This is usually to be expected when designs are critically reviewed by a multi-disciplinary team with a high level of expertise. One possible danger that must be avoided is

that the process becomes so dominated by probability expertise that clear thinking about the physical processes involved gets crowded out.

It seems likely that studies of this type will provide valuable insights to the process of setting values for partial factors. In relation to robustness, a key issue is to ensure that the eventual designs are able to accommodate, to a reasonable extent, events and actions beyond those normally included in conventional designs.

9.3.2.5 Direct assessment of design values

If design values are assessed directly, such as by using “worst credible values”, attention could be concentrated entirely on the lead variables, as tends to happen in reliability analyses, making no provision for robustness. Alternatively, directly assessed design values could be consciously chosen so as to make an allowance for robustness. Such an approach would have no calibration to past successful design, and it would be very difficult to standardise.

9.3.3 Concluding remarks for Part A

Part A of this chapter has concentrated on “type (a)” robustness identified in 2.1: the ability of the final design to accommodate events and actions that were not foreseen or consciously included in design. In Part B, an alternative form of robustness is discussed (type(b)): the sensitivity of the final design to variations of the known parameters within their anticipated range of uncertainty.

For type (a) robustness it is noted that the margins of safety required may relate more to the magnitudes of the lead variables, which govern the overall geometry and strength of the structure, than to their uncertainties. In this case, simply reducing the target probability of failure or increasing the reliability index β calculated for the lead variables may not provide the robustness required. A partial factor approach may more readily accommodate this requirement. Similarly, carrying out design for the “worst credible” values of the lead variables may not provide the required robustness.

For large projects, processes that involve critical reviews of designs or proposed design standards by multi-disciplinary teams of experts are likely to identify a larger range of situations and variables for which the designs should be checked. They will therefore increase robustness by transferring some events and actions from the category of “unforeseen” (and therefore not explicitly designed for) to “foreseen”. Rigorous study using reliability schemes and processes will probably be helpful in this respect, provided the concentration on reliability expertise is not allowed to eclipse the other skills needed in the critical review.

9.3.4 References for Part A

- CEN (2014) Robustness in Eurocodes. CEN document CEN/TC250/WG6/N10.
- CEN (2016) Practical definition of structural robustness vDraft. CEN/TC 250/WG 6, N042 WG6.PT1, NA 005-51-01 AA N 439.
- COST (2011) Structural robustness design for practising engineers. COST Action TU0601 - Robustness of Structures. NA005-51-01AA_N0132, Ed. T. D. Gerard Canisius. European Cooperation in Science and Technology.
- Foye, K.C. Salgado, R. & Scott, B. (2006b) Resistance factors for use in shallow foundation LRFD. J. Geotech. Geoenviron. Eng., 132(9), 1208–1218.
- Huang, Hongwei, Shao, Hua, Zhang, Dongming and Wang, Fei (2016) Deformational Responses of Operated Shield Tunnel to Extreme Surcharge: A Case Study. Structure and Infrastructure Engineering.
- Schweckendiek, T., Vrouwenvelder, A. C. W. M., Calle, E. O. F., Kanning, W., & Jongejan, R. B. (2012). Target Reliabilities and Partial Factors for Flood Defenses in the Netherlands. In P. Arnold, G. A. Fenton, M. A. Hicks, & T. Schweckendiek (Eds.), *Modern Geotechnical Codes of Practice - Code Development and Calibration* (pp. 311–328). Taylor and Francis. doi:10.3233/978-1-61499-163-2-311

9.4 PART B–ACCOMMODATING VARIATIONS OF THE KNOWN PARAMETERS WITHIN THEIR ANTICIPATED RANGE OF UNCERTAINTY

Part B of this report was drafted by Hongwei Huang, C. Hsein Juang, and Wenping Gong

9.4.1 Introduction

This report represents a short review on the robust geotechnical design (RGD) proposed by Dr. Juang and his colleagues. In the context RGD, an optimal design is sought with respect to design robustness and cost efficiency, while satisfying the safety requirements; and thus, RGD is generally implemented as multi-objective optimization problem. The safety requirements, in RGD, may be evaluated with either deterministic (i.e., factor of safety-based) or probabilistic (i.e., reliability-based) approach based on the characterization of the uncertain input parameters, this is consistent with the traditional geotechnical design approaches. A design, in RGD, is considered robust (i.e., having high degree of design robustness) if the system response of concern is insensitive to, or robust against, the variation in the uncertain input parameters. And, the optimal design, in RGD, is sought through carefully adjusting the “design parameters” (i.e., parameters that can be easily controlled by the engineer, such as the geometry) without reducing the uncertainty in the “noise factors” (i.e., uncertain input parameters that could not be characterized accurately). In this report, two main elements in RGD, namely, robustness measure and multi-objective optimization, are discussed. Next, the procedures for implementing the RGD are outlined. Finally, the RGD is illustrated with cases study, including braced excavation, shield tunnel, and retaining wall; the results of which demonstrate the versatility and effectiveness of the RGD.

9.4.2 Elements in Robust Geotechnical Design

Two fundamental elements in RGD, in terms of the robustness measure and the multi-objective optimization, are detailed in this section.

9.4.2.1 Robustness measure

According to the level of characterization of the uncertain input parameters (or noise factors), three levels of robustness measure could be employed in RGD: (1) site-specific data is quite limited and only the nominal values of the noise factors could be approximately estimated, the gradient-based sensitivity index (SI) (Gong et al. 2016b) could be employed; (2) site-specific data is limited and the upper bounds and lower bounds of the noise factors could be characterized, the fuzzy set-based signal-to-noise ratio (SNR) (Gong et al. 2014a&2015) could be employed; and (3) more site-specific data availability is achieved and the probability distributions of the noise factors could be characterized, however, the statistical information of the distributions (e.g., coefficient of variation) cannot be calibrated accurately, the reliability-based feasibility index (β_f) (Juang et al. 2012&2013; Juang and Wang 2013; Khoshnevisan et al. 2014; Huang et al. 2014a) could be adopted.

Gradient-based sensitivity index

In reference to Figure 9-1, two different designs (referred to herein as \mathbf{d}_1 and \mathbf{d}_2) are compared. Here, \mathbf{d}_2 is seen more robust than \mathbf{d}_1 against the variation of noise factors $\boldsymbol{\theta}$, as the gradient of the system response to the noise factors is lower in the case of \mathbf{d}_2 than in \mathbf{d}_1 . As such, the design robustness can be effectively evaluated using the gradient of the system response to the noise factors (Gong et al. 2014b). Here, the gradient of the system response to the noise factors, ∇g , at a check point of noise factors, $\boldsymbol{\theta}'$, can be expressed as follows:

$$\nabla g|_{\boldsymbol{\theta}=\boldsymbol{\theta}'} = \left\{ \left. \frac{\partial g(\mathbf{d}, \boldsymbol{\theta})}{\partial \theta_1} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}'}, \left. \frac{\partial g(\mathbf{d}, \boldsymbol{\theta})}{\partial \theta_2} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}'}, \dots, \left. \frac{\partial g(\mathbf{d}, \boldsymbol{\theta})}{\partial \theta_n} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}'} \right\} \quad (9-1)$$

where $g(\mathbf{d}, \boldsymbol{\theta})$ represents the system performance of concern, which is a function of the design parameters (\mathbf{d}) and noise factors ($\boldsymbol{\theta}$); and n represents the number of noise factors. In situations where only the nominal values of the noise factors, denoted as $\boldsymbol{\theta}_n$, could be characterized and available to the engineer, the nominal values of noise factors can be reasonably assigned as the checkpoint in Eq. (9-1): $\boldsymbol{\theta}' = \boldsymbol{\theta}_n$.

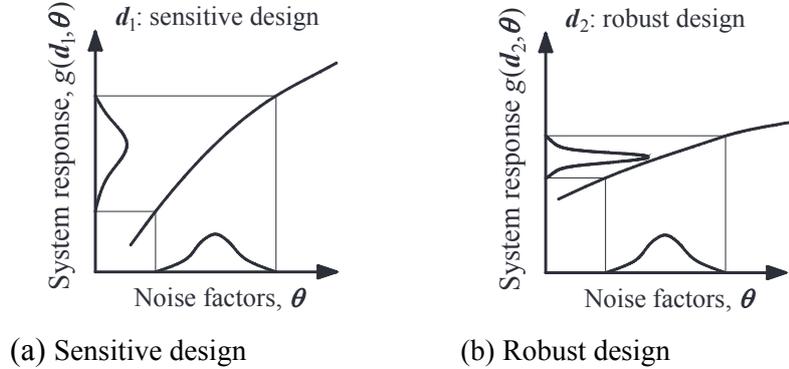


Figure 9-1. Illustration of the sensitivity of the system response to noise factors (Gong et al. 2014b)

While the gradient ∇g , defined in Eq. (9-1), is shown as an effective indicator of the design robustness, two problems need to be resolved before the robust design optimization could be implemented. First, the gradient is an n -dimensional vector; as the units of noise factors are different, the mathematical operation of this vector could be a problem. Second, the gradient is a vector rather than a scalar; it is not as convenient and effective as a scalar to use for screening candidate designs in the design pool.

To solve the first problem, each partial derivative in the gradient vector $\partial g(\mathbf{d}, \boldsymbol{\theta}) / \partial \theta_i$ for $\boldsymbol{\theta} = \boldsymbol{\theta}'$ is multiplied by a scaling factor of θ_i' so that the effect of the units of noise factors on the design robustness can be eliminated. Then, the gradient vector shown in Eq. (9-1) is re-written as follows, which is defined herein as the normalized gradient vector (\mathbf{J}):

$$\mathbf{J} = \left\{ \frac{\theta_1' \partial g(\mathbf{d}, \boldsymbol{\theta})}{\partial \theta_1} \Big|_{\boldsymbol{\theta} = \boldsymbol{\theta}'}, \frac{\theta_2' \partial g(\mathbf{d}, \boldsymbol{\theta})}{\partial \theta_2} \Big|_{\boldsymbol{\theta} = \boldsymbol{\theta}'}, \dots, \frac{\theta_n' \partial g(\mathbf{d}, \boldsymbol{\theta})}{\partial \theta_n} \Big|_{\boldsymbol{\theta} = \boldsymbol{\theta}'} \right\} \quad (9-2)$$

Note that a noise factor that exhibits higher variability could contribute more to the design robustness. Thus, a weighting factor, which indicates the contribution of the noise factor to the robustness, might be adopted in formulation of the normalized gradient vector (\mathbf{J}), which is detailed in Gong et al. (2016a). To solve the second problem, the Euclidean norm of the normalized gradient vector, which signals the length of the normalized gradient vector (\mathbf{J}), is adopted and defined herein as the sensitivity index (SI).

$$SI = \sqrt{\mathbf{J}\mathbf{J}^T} \quad (9-3)$$

The sensitivity index (SI) shown in Eq. (9-3) yields a single value representation of the normalized gradient vector. As can be seen, a higher SI value signals lower design robustness, as it indicates a greater variation of the system response in the face of the uncertainty in the noise factors.

Fuzzy set-based signal-to-noise ratio

A fuzzy set is a set of ordered pairs, $[\theta, \mu(\theta)]$, where a member θ belongs to the set with a certain confidence, called membership grade, $\mu(\theta)$. These ordered pairs collectively define a membership function that specifies a membership grade for each member. Note that although the membership function is not a probability density function (PDF), a membership grade does give a degree of

confidence that a member θ belongs to this set. If the highest membership grade in a fuzzy set is normalized to 1 and the shape of the membership function is unimodal, this fuzzy set becomes a fuzzy number. For a geotechnical parameter with known upper bound and lower bound, the membership function could be conveniently constructed by setting the membership grade at $\theta =$ upper bound or upper bound to 0, while the membership grade at $\theta =$ the average of the upper bound and the upper bound to 1, as shown in Figure 9-2(a). As such, the uncertain input parameters are modelled with triangular fuzzy numbers (i.e., the fuzzy numbers with a triangular shape membership function). Of course, other membership function, such as trapezoidal shape, may be used.

For a geotechnical system with fuzzy input data, the uncertainty propagation may be studied with vertex method (Dong and Wong 1987). In the context of vertex method, the corresponding interval of output, in terms of g_{ai}^- and g_{ai}^+ , for the α_i -cut level of input fuzzy data (see Figure 9-2b) is able to be obtained through 2^n deterministic analysis, where n represents the number of fuzzy input data. After finishing the analysis of all α -cut levels, the final fuzzy output could be easily constructed, which represents the final outcomes of the uncertainty propagation through the solution model. Detailed information of the system performance could be provided from which. For example, the mean and standard deviation of the system performance $g(\mathbf{d}, \boldsymbol{\theta})$, denoted as $E[g]$ and $\sigma[g]$, respectively, can readily be derived from the fuzzy out shown in Figure 9-2(b), using the formulation in Gong et al. (2014a&2015). Then, the signal-to-noise ratio (SNR), defined below, is constructed to measure the design robustness (Phadke 1989).

$$SNR = 10 \log_{10} \left(\frac{E^2[g(\mathbf{d}, \boldsymbol{\theta})]}{\sigma^2[g(\mathbf{d}, \boldsymbol{\theta})]} \right) \quad (9-4)$$

Here, a higher SNR signals lower variability of the system response, and thus higher design robustness.

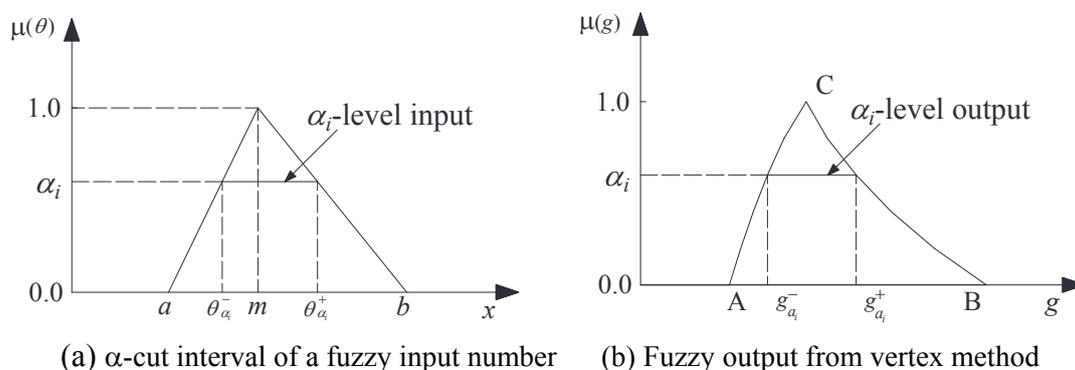


Figure 9-2. Geotechnical analysis with fuzzy input data.

Reliability-based feasibility index

Note that although the probabilistic distributions of the noise factors could be determined, the statistical (e.g., coefficient of variation) of the noise factors could not be characterized with certainty because of the limited availability of site-specific data. However, the failure probability estimate obtained from the probabilistic approach is often greatly dependent upon the adopted statistical information of the uncertain input parameters. In consideration of the uncertainty in the statistical characterization of the noise factors, the failure probability of a geotechnical system may not be able to be accurately derived and which will be uncertain. In such a circumstance, the variation of the failure probability, which could arise from the uncertainty in the statistical characterization of the noise factors, needs to be estimated and minimized in the context of RGD. That is to say, the variation of the failure probability may be adopted as the robustness measure (Juang et al. 2012&2013). A smaller variation of the failure probability signals lower variability of the system response (i.e., failure probability in the context of the probabilistic approach), and thus higher design robustness. It is noted that the target failure probability might be different for different geotechnical system and the

magnitude of the variation of the failure probability could vary in a significant range. Thus, the reliability-based feasibility index (β_β), defined below, and could be employed (Juang et al. 2012&2013; Huang et al. 2014a).

$$\beta_\beta = \frac{\ln \left[\frac{P_{fT}}{P_{fmean}} \sqrt{1 + (P_{fstdev}/P_{fmean})^2} \right]}{\sqrt{\ln \left[1 + (P_{fstdev}/P_{fmean})^2 \right]}} \quad (9-4)$$

where P_{fT} represents the target failure probability; and, P_{fmean} and P_{fstdev} represent the mean and standard deviation of the failure probability. As can be seen in Eq. (9-5), the failure probability is assumed to be lognormally distributed; here, the feasibility index (β_β) can be interpreted as the feasibility probability of the geotechnical system that the target failure probability is stratified:

$$\Phi(\beta_\beta) = \Pr[P_f < P_{fT}] \quad (9-5)$$

where $\Phi(\cdot)$ represents the cumulative distribution of the standard normal variable, and $\Pr[P_f < P_{fT}]$ represents that the target failure probability of this geotechnical system could be stratified in the face of the uncertainty in the statistical characterization of the uncertain input parameters.

For a given set of statistics of the uncertain input parameters, the failure probability (P_f) of the geotechnical system can readily be estimated with the probabilistic methods such as first order reliability method (FORM) (Low and Tang 2007), Monte Carlo simulation (MCS), and point estimate method (PEM) (Zhao and Ono 2000). In consideration of the uncertainty in the statistical characterization of the noise factors, the two-loop probabilistic analysis should be conducted. The inner loop is employed to estimate the failure probability for a given set of statistics of the noise factors, this is similar to the existing probabilistic analysis. The second loop is employed to derive the mean and standard deviation of the failure probability that arise from the uncertainty in the statistics of the noise factors. To this end, the PEM-FORM (Juang et al. 2013), PEM-MCS, and weighted MCS (Peng et al. 2016) may be employed.

9.4.2.2 Multi-objective optimization

The essence of RGD is to seek an optimal design with respect to design robustness and cost, while satisfying the safety requirements. Once the system response of concern is chosen, and the design robustness, cost, and safety are evaluated, the optimal design could be obtained through a multi-objective optimization formulated as follows:

$$\begin{aligned} \text{Find:} & \quad \text{design parameters } \mathbf{d} \\ \text{Subject to:} & \quad \mathbf{d} \in \text{design space } \mathbf{DS} \\ & \quad \text{satisfying safety requirements} \\ \text{Objectives:} & \quad \text{maximizing design robustness} \\ & \quad \text{minimizing cost} \end{aligned} \quad (9-6)$$

Based on the level of characterization of the uncertain input parameters, the safety requirements may be evaluated using either the deterministic (i.e., factor of safety-based) or probabilistic (i.e., reliability-based) approach; similarly, the design robustness can be evaluated using either the gradient-based sensitivity index (SI), fuzzy set-based signal-to-noise ratio (SNR), or reliability-based feasibility index (β_β).

In reference to the optimization setting shown in Eq. (9-7), a single best optimal design is generally unattainable since these two objectives (i.e. robustness and cost) are conflicting. The multi-objective optimization in this scenario yields a set of “non-dominated” designs, the collection of all these non-dominated designs is known as Pareto front (Deb et al. 2002). Among all the designs on

the Pareto front, none is superior or inferior to others on the Pareto front with respect to both objectives, but they are all superior to the designs in the feasible domain. Figure 9-3 shows a conceptual sketch of Pareto front in a bi-objective optimization problem. Note that the utopia point, shown in Figure 9-3, is an unattainable design, the concept of which is discussed later.

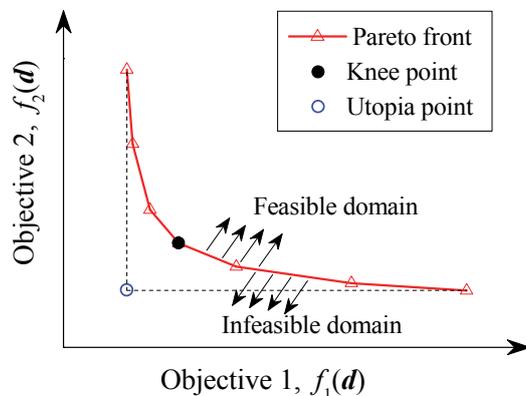


Figure 9-3. Conceptual sketch of Pareto front and knee point in a bi-objective optimization

The Pareto front in Figure 9-3 could be easily obtained with the multi-objective optimization algorithms such as “Non-dominated Sorting Genetic Algorithm” version II (NSGA-II) developed by Deb et al. (2002). The derived Pareto front is problem-specific, which could be employed as a design aid to assist in making an informed design decision. For example, at a preferred (pre-specified) cost level, the design with the highest robustness among all points on the Pareto front can be taken as the final design. On the other hand, at a pre-specified robustness level, the design with the least cost among all points on the Pareto front can be taken as the final design. The choice of an appropriate level of cost or robustness, however, is problem-specific. When no such a design preference is specified, the knee point on the Pareto Front, which yields the best compromise between robustness and cost efficiency, may be taken as the most preferred design in the design space. Interested readers are referred to Branke et al. (2004) and Deb and Gupta (2011) for detailed procedures for identifying the knee point on the Pareto Front.

Instead of the genetic algorithms such as NSGA-II, the Pareto front shown in Figure 9-3 could also be identified with the simplified procedure detailed in Khoshnevisan (2015), in which the bi-objective optimization is transformed into a series of single-objective optimizations. Further, the owner or client may be only interested in the most preferred design in the design space (i.e., the knee point on the Pareto front), and not the Pareto front per se. Thus, a simplified procedure is further developed in Gong et al. (2016b), in which the multi-objective optimization is solved through a series of single-objective optimizations and the knee point on the Pareto front could be identified directly (Khoshnevisan et al. 2014; Gong et al. 2016b).

9.4.3 Procedures for Implementing Robust Geotechnical Design

The procedures for implementing the proposed RGD could be summarized in the following main steps:

Step 1: Describe the geotechnical problem of concern with mathematical models. Here, the system response of concern, noise factors, and design parameters are identified; meanwhile, the design (safety) requirements, design robustness, cost, and design space are formulated.

Step 2: Carry out the robust design optimization considering design robustness, cost efficiency, and safety requirements using the optimization setting shown in Eq. (9-7), where the design robustness and safety requirements for each candidate design could be analyzed based on the level of characterization of the uncertain input parameters (or noise factors). The results of the optimization culminate in a Pareto front showing a tradeoff between design robustness and cost efficiency for all the non-dominated designs that satisfy the safety requirements. Here, the Pareto front can be identified using either the genetic algorithms such as NSGA-II (Deb et al. 2002) or simplified procedure in Khoshnevisan (2015).

Step 3: Select the most preferred design on the derived Pareto front. In principle, either the least cost design that is above a pre-specified level of design robustness or the most robust design that falls within a pre-specified cost level may be selected as the most preferred design in the design space. Alternatively, the knee point, which represents the best compromise solution in the design space, may be identified (Branke et al. 2004; Deb and Gupta 2011). It is worth noting that the most preferred design in the design space could also be identified directly with the simplified procedure in Gong et al. (2016b).

9.4.4 Cases Study

To demonstrate the versatility and effectiveness of the RGD, three cases, including braced excavation, shield tunnel, and retaining wall, are studied in this section.

9.4.4.1 Case 1: Robust design of braced excavation

The first case concerns the robust design of a shoring system, which consists of soldier piles (i.e., reinforced concrete piles) with timber laggings and tieback anchors, for an excavation in a sandy soil deposit, as shown in Figure 9-4. The robust design of this case is detailed in Gong et al. (2016b). In this case, the diameter of the concrete soldier pile (D), length of the concrete soldier pile (L), interval of concrete soldier piles (I), vertical spacing of tieback anchors (V), horizontal spacing of tieback anchors (H), and the angle of tieback anchors with respect to the horizontal direction (α) are taken as the design parameters. Whereas, the preload of tieback anchors is chosen at 20 ton per tieback, and the length of tieback anchors is set at 8.0 m based on local practice. For illustration purpose, a discrete design space is considered, which specifies the possible selections of the design parameters, as listed in Table 9-1, and 38,500 candidate designs are considered.

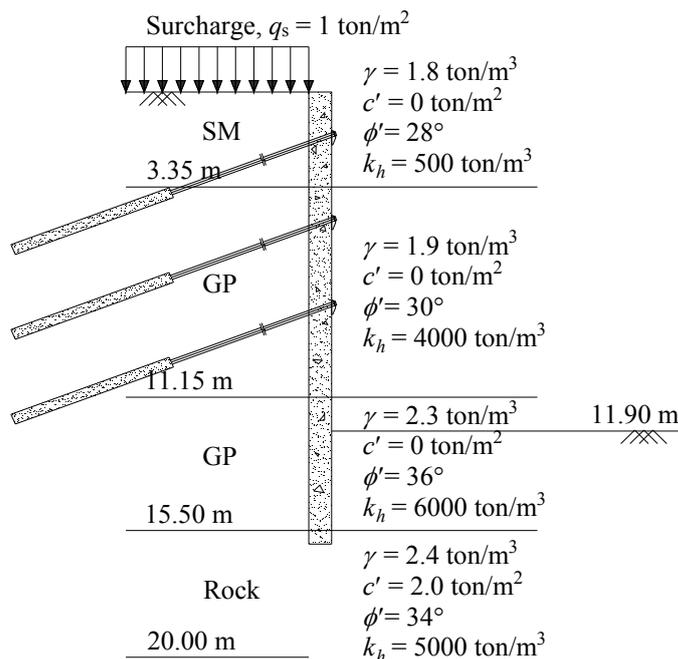


Figure 9-4. Excavation with a shoring system of soldier piles and anchor tiebacks

The drained cohesion (c'), drained friction angle (ϕ') and modulus of horizontal subgrade reaction (k_h), along with the surcharge behind the wall (q_s), are considered the noise factors (i.e., uncertain input parameters). Due to the limited availability of site specific-data, only the nominal values of the noise factors could be estimated; as such, the design robustness is measured herein by the gradient-based sensitivity index (SI). Through which, the variations in the noise factors are recognized but there is no need to perform a detailed statistical characterization of the noise factors, as the system response (i.e., stability and deformation) and its sensitivity with respect to the noise factors could be approximately evaluated with the nominal values of the noise factors.

In general, the safety requirement of a braced excavation is evaluated through the limiting factors of safety and limiting maximum wall and/or ground deformation (JSA 1988; TGS 2001; PSCG 2000). Here, TORSA, a commercially available FEM code based on the beam-on-elastic-foundation theory (Sino-Geotechnics 2010), is used to compute the system responses, including the factor of safety against push-in failure (F_{S1}), factor of safety against basal heave failure (F_{S2}), and the maximum wall deflection (y). In RGD, the maximum wall deflection is chosen as the system response of concern for the purpose of defining the design robustness; whereas, the safety requirement is evaluated with the computed factors of safety and wall deformation.

Table 9-1. Design space of the design parameters for Case 1

Design parameter	Design space
Diameter of the soldier pile, D (m)	{0.3 m, 0.4m, 0.5m, 0.6 m, 0.7 m}
Length of the soldier pile, L (m)	{14 m, 15 m, 16 m, 17 m, 18 m, 19 m, 20 m}
Horizontal interval of the soldier pile, I (m)	{ D , $D + 0.1$ m, $D + 0.2$ m, ..., $D + 1.0$ m}
Vertical spacing of tieback anchors, V (m)	{2.0 m, 2.5 m, 3.0 m, 3.5 m}
Horizontal spacing of tieback anchors, H (m)	{1.5 m, 2.0 m, 2.5 m, 3.0 m, 3.5 m}
Installed angle of the tieback anchor, α (°)	{10°, 15°, 20°, 25°, 30°}

For a shoring system project, the cost (C) should be the sum of the cost on excavation, cost on soldier pile wall and cost on tieback anchors. Because the site dimensions and excavation depth, in the specified project, are predefined based on the project’s requirements, the cost on excavation will not affect the optimization results, and only the cost on the shoring system is considered in the robust design optimization. The detailed formulation of the cost (C) could be found in Gong et al. (2016b). The robust design optimization setting of this case is depicted in Figure 9-5.

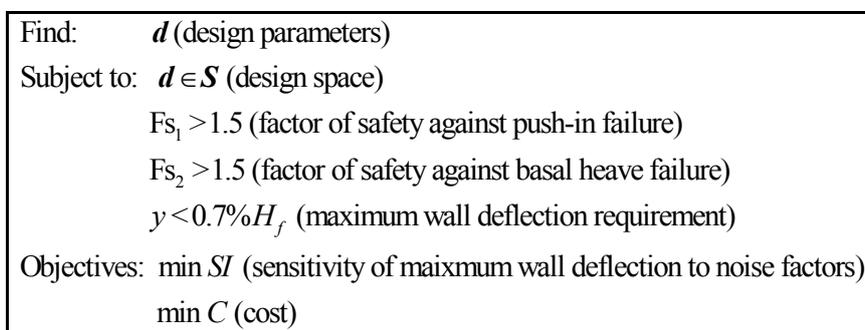


Figure 9-5. Robust optimization setting for Case 1 (where H_f is the final excavation depth)

Applying the genetic algorithm NSGA-II (Deb et al. 2002), the robust design optimization shown in Figure 9-5 yields a Pareto front, as plotted in Figure 9-6(a); then, the knee point is located, which is also plotted in Figure 9-6(a). Meanwhile, the robust design is carried out using the simplified procedure in Gong et al. (2016b); with which, the knee point is identified directly without constructing the Pareto front and the results are plotted in Figure 9-6(b). Note that the difference between the knee point obtained by the simplified procedure and that obtained by the multi-objective optimization algorithm NSGA-II is quite negligible. Next, a comparison with the original design that was selected by the engineering firm is made. The original design is the one designed by an experienced engineering firm (Hsui-Sheng Hsieh, personal communication 2013) without the knowledge of RGD. While the original design appears to be a sound engineering practice, offering a compromise between the least cost design and the most robust design, it is inferior to the knee point on the Pareto front, as the latter is more robust and cost less. Through this real-world application, the advantages of RGD are demonstrated.

The design parameters of these designs are tabulated in Table 9-2. Here, the knee point on the resulting Pareto front is obtained by the normal boundary intersection approach (Deb and Gupta 2011) and marginal utility function approach (Branke et al. 2004). These two approaches yield the same

design, denoted as \mathbf{d}_{2-1} . The difference between the design parameters of the most preferred design obtained by the simplified procedure, denoted as \mathbf{d}_{2-2} , and those of \mathbf{d}_{2-1} is relatively small and could be ignored. The results show that the most preferred design obtained by the simplified procedure is practically the same as the knee point on the Pareto front obtained by the multi-objective optimization method, which requires a two-step solution (developing a Pareto front by the multi-objective optimization using genetic algorithms such as NSGA-II, and then searching for knee point on the Pareto front). From there, the effectiveness of the simplified procedure is demonstrated.

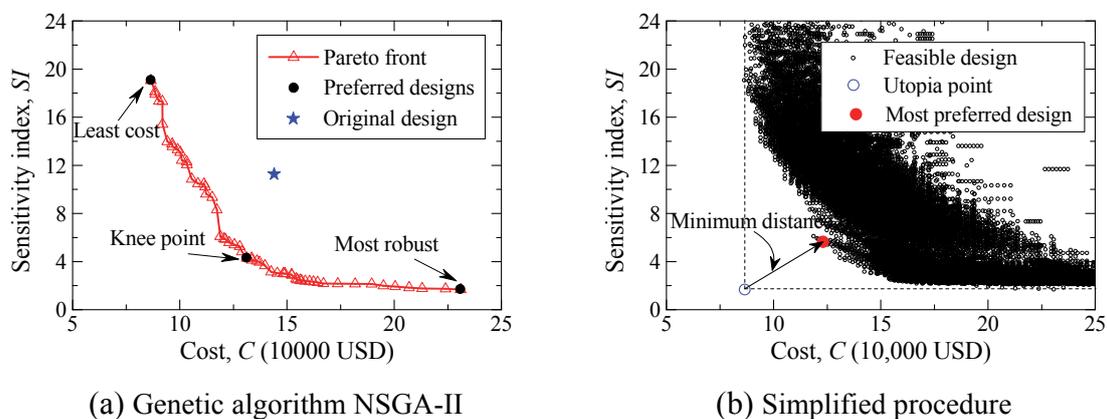


Figure 9-6. Results of the robust design for Case 1

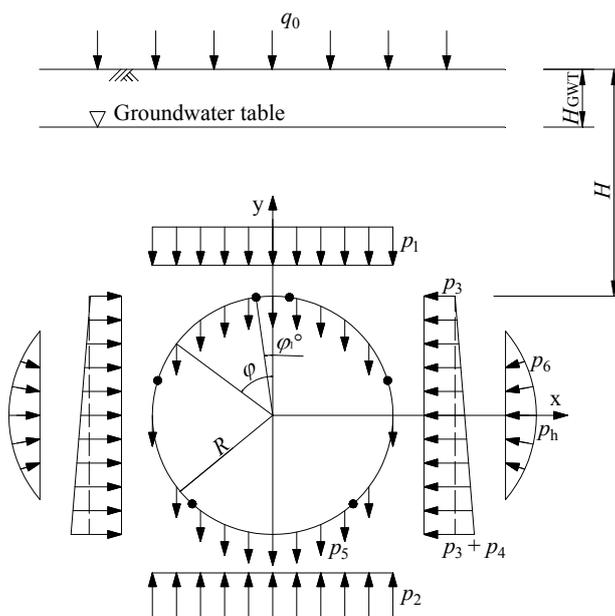


Figure 9-7. Analysis model of shield tunnels (Huang et al. 2014b)

9.4.4.2 Case 2: Robust design of shield tunnel

The second case considers the robust design of the cross section of a shield tunnel in Shanghai, as shown in Figure 9-7. The robust design of this case is detailed in Huang et al. (2014b). In this case, the segment thickness (t), steel reinforcement ratio (ρ), and diameter of joint bolt (D_j) are dealt as the design parameters and which are to be optimized in a pre-assigned design space. The soil resistance coefficient (K_s), soil cohesion strength (c), soil friction angle (ϕ), ground water table (H_{GWT}), and surcharge (q_0) are considered as the noise factors. Here, only the upper and lower bounds of the noise factors can be estimated and which are tabulated in Table 9-3. The other deterministic parameters to assess the tunnel performance, in terms of the structure safety (i.e., ULS performance) and serviceability (i.e., SLS performance), are tabulated in Table 9-4.

Table 9-2. Most preferred design obtained with different approaches for Case 1

Adopted approach	Design parameters						Design performances			Cost, C (10,000 USD)	Sensitivity index, SI
	D (m)	L (m)	I (m)	V (m)	H (m)	α (°)	F _{S1}	F _{S2}	y (cm)		
NSGA-II and normal boundary intersection approach, \mathbf{d}_{2-1}	0.6	18	1.6	3.0	2.0	10	5.67	2.96	3.48	13.12	4.28
NSGA-II and marginal utility function approach, \mathbf{d}_{2-1}	0.6	18	1.6	3.0	2.0	10	5.67	2.96	3.48	13.12	4.28
Simplified procedure, \mathbf{d}_{2-2}	0.5	18	1.4	3.0	2.0	10	5.67	2.96	4.94	12.31	5.58
Original design, \mathbf{d}_0	0.5	17	0.6	3.0	2.5	20	4.46	2.75	4.13	14.42	11.22

The design robustness, in this case, is evaluated using the fuzzy set-based signal-to-noise (SNR), the cost (C) is represented by the material cost of one tunnel ring that consists of segment concrete cost, steel reinforcement cost and joint bolts cost, and the safety requirements (i.e., ULS and SLS behaviour) are evaluated using the reliability indexes that are derived from the fuzzy outputs. The formulations of the design robustness, cost, and safety requirements are detailed in Gong et al. (2014a).

Table 9-3. Parameters characterizing membership functions of noise factors

Noise factors	Lower bound, a	Mode, $m = (a + b)/2$	Upper bound, b
Soil resistance coefficient, K _s (kN/m ³)	3500	9250	15000
Soil cohesion strength, c (kN/m ²)	0	7.5	15
Soil friction angle, ϕ (°)	30	32.65	35.3
Ground water table, H _{GWT} (m)	0.5	1.25	2
Ground surcharge, q ₀ (kN/m ²)	0	10	20

In this case, the design parameters (t, ρ, D_j) are to be optimized in the contiguous design space of [0.2 m, 0.5 m], [0.5%, 4.0%] and [10.0 mm, 50.0 mm] such that the design robustness and cost efficiency are maximized simultaneously. The robust design optimization setting of this case is set up as follows:

$$\begin{aligned}
 \text{Find:} & \quad (t, \rho, D_j) \\
 \text{Subjected to:} & \quad 0.2\text{m} \leq t \leq 0.5\text{m}; 0.5\% \leq \rho \leq 4.0\%; 10\text{mm} \leq D_j \leq 50\text{mm} \\
 & \quad |\beta_1 - 4.2| \leq 0.1\%; |\beta_2 - 2.7| \leq 0.1\% \\
 \text{Objectives:} & \quad \text{Maximizing the robustness index of ULS, } \text{SNR}_1 \\
 & \quad \text{Maximizing the robustness index of SLS, } \text{SNR}_2 \\
 & \quad \text{Minimizing the cost, } C(t, \rho, D_j)
 \end{aligned} \tag{9-8}$$

where $\beta_{\beta 1}$ and $\beta_{\beta 2}$ represent the reliability index of this shield tunnel with respect to the ULS and SLS behaviour, respectively; and, SNR_1 and SNR_2 represent the design robustness of this shield tunnel with respect to the ULS and SLS performance, respectively.

With the robust design optimization setting shown in Eq. (9-8), the RGD of this shield tunnel is readily conducted with NSGA-II (Deb et al. 2002). In this non-dominated optimization using NSGA-II, the population size is assigned as 50 while the generation number is set as 100. The

resulting non-dominated optimal designs are depicted in Figure 9-8, the tradeoff relationship between the robustness (i.e., SNR_1 and SNR_2) and cost is clearly illustrated: design robustness tends to increase with the cost. Thus, the desire to maximize the design robustness and the desire to minimize the cost are two conflicting objectives.

Note that while the obtained non-dominated optimal designs shown in Figure 9-8 concentrate in a relative narrow range due to the safety requirements adopted, no single best design could be screened out. In order to further ease the decision making in the RGD of shield-driven tunnels, the knee point on the Pareto front is identified. The resulting knee point is employed as the most preferred design and the best compromise among the conflicting design objectives. The design parameters of the identified knee point are: $t = 288.1$ mm, $\rho = 1.16$ %, $D_j = 49.2$ mm, and corresponding 3-D coordinate in Figure 9-8 is: $SNR_1 = 10.793$, $SNR_2 = 16.070$, $C = 1234.2$ USD.

Table 9-4. Constant parameters involved in the tunnel design

Category	Parameter	Value
Tunnel geometry parameters	Embedded depth, H (m)	15.0
	Tunnel inner radius, R_{in} (m)	2.75
	With of tunnel ring, b (m)	1.0
	Joint position of half structure, φ_i (°)	8, 73, 138
Tunnel segment	Unit weight of concrete, γ_c (kN/m ³)	25.0
	Elastic modulus of concrete, E_c (kN/m ²)	35×10^6
	Compression strength of concrete, f_c (kN/m ²)	39×10^3
	Ultimate plastic strain of concrete, ϵ_p	0.0033
Reinforcement steel	Elastic modulus of steel, E_s (kN/m ²)	210×10^6
	Yielding strength of steel bar, f_y (kN/m ²)	345×10^3
	Thickness of protective cover, a (m)	0.05
Joint bolts	Bolt length, l_b (m)	0.4
	Number of bolts at each joint	2
	Distance from joint bolts to tunnel inside surface, h	$t/3$

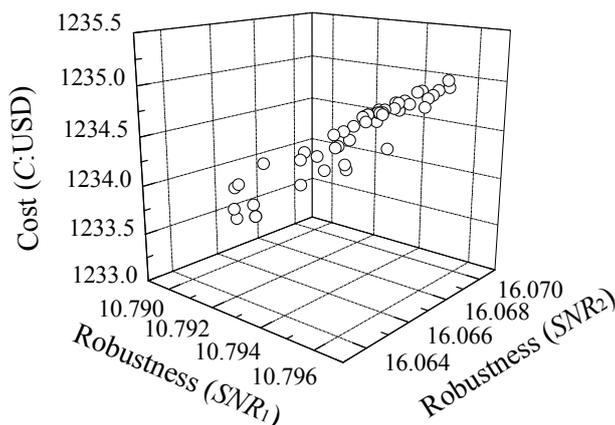


Figure 9-8. Resulting non-dominated optimal designs for Case 2

To demonstrate the significance of the RGD of shield tunnels, a comparison among the robust design, probabilistic design and current practice (i.e., design adopted in Shanghai) is conducted, and the results are listed in Table 9-5. Comparing with probabilistic design and current practice, the design parameters of robust design are notably adjusted: the segment thickness is decreased while the steel reinforcement ratio and joint bolts diameters are increased; that is to say, the joint stiffness is

increased while the stiffness of segment is decreased. This adjustment of the design parameters of the shield tunnel is quite reasonable. Though the resulting robustness indexes (SNR_1 and SNR_2) do not change much, the variation (i.e., COV) of tunnel performances do decrease significantly. For example, the variation of the system performance of the robust design is significantly reduced (as large as 30% for ULS) whereas the cost is only increased by 25%. Thus, the significance of the RGD is illustrated.

Table 9-5. Comparison among three design designs for Case 2

Category	Parameter	Robust design	Probabilistic design	Current practice
Design parameters	t (mm)	288.1	343.5	350.0
	ρ (%)	1.16	0.83	0.50
	D_j (mm)	49.2	28.5	30.0
Safety	β_1 of ULS	4.20	4.20	2.51
	β_2 of SLS	2.70	2.70	3.08
Robustness	SNR_1 of ULS	10.793	10.160	8.533
	SNR_2 of SLS	16.070	16.210	16.424
Cost	C (USD)	1234.2	1175.5	988.9
Coefficient of variation (COV)	F_{S1} of ULS	0.289	0.310	0.374
	F_{S2} of SLS	0.157	0.155	0.151

9.4.4.3 Case 3: Robust design of retaining wall

The third case considers the robust design of a retaining wall, as shown in Figure 9-9. The robust design of this case is detailed in Huang et al. (2014a). In this case, the top width (a) and base width (b) of the retaining wall are treated as the design parameters, and which are to be optimized in the discrete design space of $\{(a, b) \mid a = 0.2 \text{ m}, 0.4 \text{ m}, 0.6 \text{ m} \text{ and } b = 0.6 \text{ m}, 0.7 \text{ m}, 0.8 \text{ m}, \dots, 3.0 \text{ m}\}$. The unit weight of the backfill soil (γ), soil friction angle (ϕ), friction angle between the backfill and retaining wall (δ), and the adhesion (c_a) are considered as the noise factors. Here, the noise factors are characterized as uncertain variables, however, the statistics of which could not be estimated with certainty. The statistical information of the noise factors is tabulated in Table 9-6; note that the COVs of the noise factors are assumed to be lognormally distributed. Here, the performances regarding the overturning and sliding failure are studied.

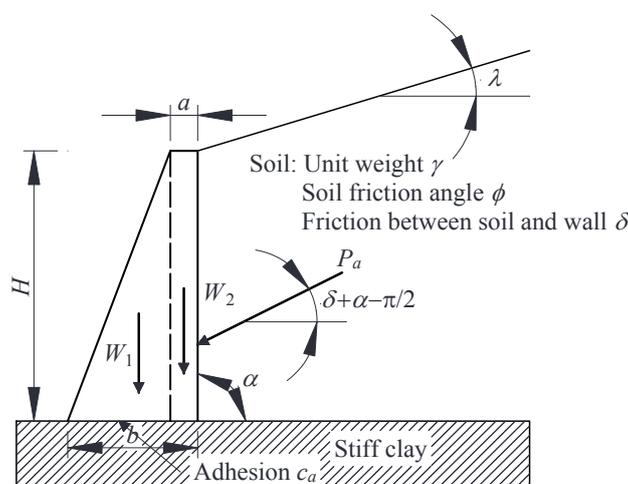


Figure 9-9. Schematic diagram of the retaining wall design for Case 3

The safety requirements are evaluated using the mean of the failure probabilities (i.e., P_{f1} and P_{f2} for the overturning and sliding failure, respectively), the design robustness is evaluated using the reliability-based feasibility index (i.e., $\beta_{\beta1}$ and $\beta_{\beta2}$ for the overturning and sliding failure, respectively), and the cost (C) is evaluated using the volume of the retaining wall. Detailed formulations of these factors could be found in Huang et al. (2014a).

Table 9-6. Statistical information of the noise factors for Case 3

Noise factors	Distribution type	Mean, μ	Coefficient of variation, COV	Mean of COV, μ_{cov}	Standard deviation of COV, σ_{cov}
Unit weight, γ_s	Normal	18kN/m ³	2~10%	6.5%	1.17%
Friction angle, ϕ	Normal	35°	5~20%	10%	2.50%
Friction between soil and retaining wall, δ	Normal	20°	5~20%	10%	2.50%
Adhesion between wall base and clay, c_a	Normal	100kPa	10~30%	15%	3.33%

Figure 9-10 shows the tradeoff relationship between the variation of the failure probability and the cost, the variation of the failure probability generally decreases with the increase of the cost. Next, the reliability-based feasibility indexes of these discrete candidate designs are studied and the results are plotted in Figure 9-11; as expected, the reliability-based feasibility index often increases with the cost. With the aid of Figure 9-11, the final design could be readily identified. For example, Table 9-7 illustrates the resulting robust designs that are identified for a series of target reliability-based feasibility indexes.

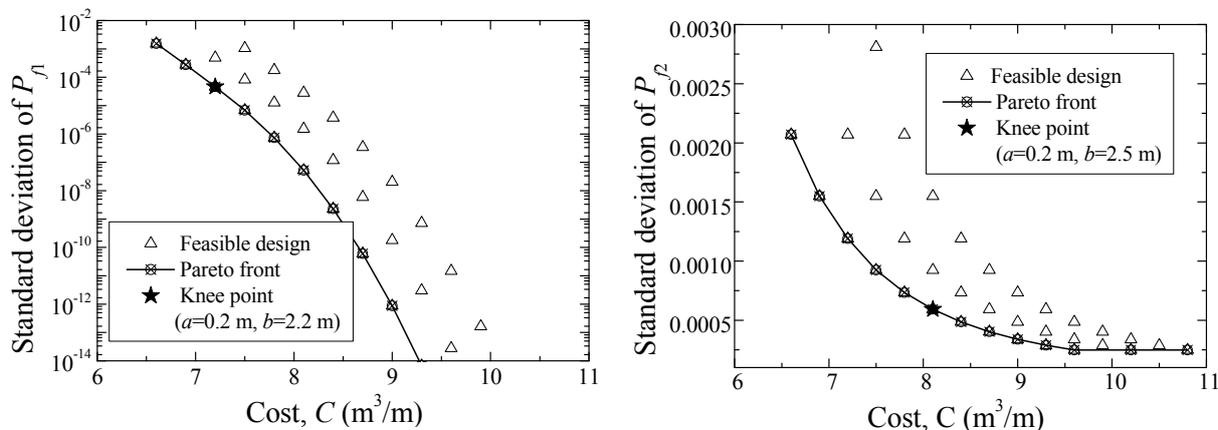
Table 9-7. Identified final designs for Case 3

Target reliability-based feasibility index, $\beta_{\beta T}$	Confidence level, $\Pr[P_f < P_{fT}]$	Identified final design	Reliability-based feasibility index, β_{β}	Cost, C
$\beta_{\beta T} = 1.5$	93.32%	a = 0.2 m b = 2.1 m	$\beta_{\beta1} = 2.36$ $\beta_{\beta2} = 1.88$	6.9m ³ /m
$\beta_{\beta T} = 2.0$	97.72%	a = 0.2 m b = 2.2 m	$\beta_{\beta1} = 3.41$ $\beta_{\beta2} = 2.11$	7.2 m ³ /m
$\beta_{\beta T} = 2.5$	99.38%	a = 0.2 m b = 2.5 m	$\beta_{\beta1} = 6.21$ $\beta_{\beta2} = 2.65$	8.1 m ³ /m
$\beta_{\beta T} = 3.0$	99.87%	a = 0.2 m b = 2.8 m	$\beta_{\beta1} = 10.21$ $\beta_{\beta2} = 3.03$	9.0 m ³ /m

9.4.5 Part B – Discussion and Conclusion

The uncertainties in soil parameters, solution model, applied loads, and those caused by the construction, often make it difficult to ascertain the performance of a geotechnical design. In traditional deterministic approaches, these uncertainties could not be explicitly characterized and included in the design analysis; rather, a conservative factor of safety (FS) is adopted based on the concept of “calculated risk”. This FS-based design approach often leads to an inefficient over-design with an unknown and/or inconsistent safety level, although under-design is also a possibility. To achieve a more rational and consistent assessment of the safety, the reliability-based design (RBD) approach has long been suggested as an alternative. The RBD approach for the design of a

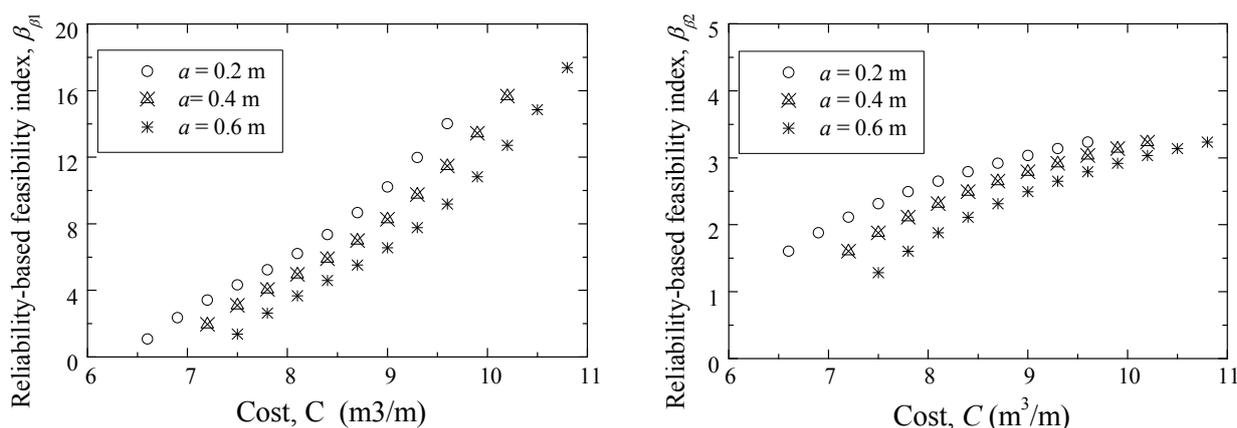
geotechnical structure is often implemented with a target reliability index, which is derived from a cost-benefit analysis that balances investment and risk considering the failure probability and consequence.



(a) Overturning failure mode

(b) Sliding failure mode

Figure 9-10. Tradeoff between the variation of the failure probability and the cost (Case 3)



(a) Overturning failure mode

(b) Sliding failure mode

Figure 9-11. Relationship between the reliability-based feasibility index and the cost (Case 3)

In the context of RBD approach, the performance of a geotechnical structure is analyzed using probabilistic methods that consider explicitly uncertainties in input parameters and/or solution models. It is noted that although various methods have been investigated to estimate the statistics of soil parameters and model errors, the statistics of soil parameters and those of model errors could not be characterized with certainty due to limited availability of site-specific data. Because of the difficulty in obtaining the accurate statistical characterization of soil parameters and model errors in practice, the RBD approach is not widely applied in geotechnical practice; rather, the load and resistance factor design (LRFD) approach, which is a simpler variant of the RBD approach by design, is more commonly used. The LRFD code employs partial factors (e.g., resistance factors and load factors), which have been calibrated to achieve a target reliability index approximately over a range of design scenarios covered by the code. The resulting design is a function of the specified partial factors and selected nominal values, with due consideration of cost. As is well recognized, LRFD is meant to be a simpler variant of the more demanding RBD; the ideal outcome is that the design obtained by LRFD could achieve the same target reliability index as that obtained by RBD. However, the standard LRFD approach that involves fixed partial factors cannot cover all design scenarios involving different levels

of variation of soil parameters and model errors. For a given design scenario, the use of the standard LRFD code may lead to a design that deviates from the target reliability index by an unknown amount, more likely on the conservative side but under-design is also a possibility..

In such circumstances, the robust geotechnical design (RGD) philosophy was advanced. With which, the uncertainty in the predicted performance of a given geotechnical design could be effectively reduced in the face of recognized but unquantified uncertainties (i.e., the uncertainties in soil parameters, solution model, applied loads, and those caused by construction). The purpose of robust design is to derive a design that effectively accounts for the effect of the variation in “noise factors” while simultaneously considers the safety and cost efficiency. In this report, the RGD, along with the fundamental issues of how the design robustness is measured, how the robust design optimization is conducted, and how the most preferred design in the designs space is selected, is presented and illustrated with cases studies. Based upon the results outlined in this report, the versatility and significance of the RGD are demonstrated.

9.4.6 References for Part B

- Branke, J., Deb, K., Dierolf, H., and Osswald, M. (2004). “Finding knees in multi-objective optimization.” *Parallel Problem Solving from Nature-PPSN VIII*, pp. 722-731.
- Deb K., Pratap A., Agarwal S., and Meyarivan T. (2002). “A fast and elitist multi-objective genetic algorithm: NSGA-II.” *IEEE Transactions on Evolutionary Computation*, 6(2), 182-197.
- Deb, K., Gupta, S. (2011). “Understanding knee points in bicriteria problems and their implications as preferred solution principles.” *Engineering Optimization* 43(11), 1175-1204.
- Dong, W. M., and Wong, F. S. (1987). “Fuzzy weighted averages and implementation of the extension principle.” *Fuzzy Sets and Systems*, 21(2), 183-199.
- Gong, W., Wang, L., Juang, C. H., Zhang, J., and Huang, H. (2014a). “Robust geotechnical design of shield-driven tunnels.” *Computers and Geotechnics*, 56, 191-201.
- Gong, W., Khoshnevisan, S., and Juang, C. H. (2014b). “Gradient-based design robustness measure for robust geotechnical design.” *Canadian Geotechnical Journal*, 51(11), 1331-1342.
- Gong, W., Wang, L., Khoshnevisan, S., Juang, C. H., Huang, H., and Zhang, J. (2015). “Robust geotechnical design of earth slopes using fuzzy sets.” *Journal of Geotechnical and Geoenvironmental Engineering*, 141(1), 04014084.
- Gong, W., Juang, C. H., Khoshnevisan, S., and Phoon, K. K. (2016a). “R-LRFD: Load and resistance factor design considering robustness.” *Computers and Geotechnics*, 74, 74-87.
- Gong, W., Huang, H., Juang, C. H., and Wang, L. (2016b). “Simplified robust geotechnical design of soldier pile-anchor tieback shoring system for deep excavation.” *Marine Georesources & Geotechnology*, online.
- Huang, H., Gong, W., Juang, C. H., and Zhang, J. (2014a). “Robust geotechnical design of gravity retaining wall.” *Journal of Tongji University (natural science)*, 42(3), 377-385.
- Huang, H., Gong, W., Juang, C. H., and Khoshnevisan, S. (2014b). “Robust geotechnical design of shield-driven tunnels using fuzzy sets.” In *American Society of Civil Engineers (ASCE) Tunnelling and Underground Construction*, p. 184-194.
- Japanese Society of Architecture (JSA). (1988). *Guidelines of Design and Construction of Deep Excavations*. Japanese Society of Architecture, Tokyo, Japan.
- Juang, C.H., Wang, L., Khoshnevisan, S., and Atamturktur, S. (2012). “Robust geotechnical design – methodology and applications.” *Journal of GeoEngineering*, 8(3), 71-81.
- Juang, C. H., Wang, L., Liu, Z., Ravichandran, N., Huang, H., and Zhang, J. (2013). “Robust geotechnical design of drilled shafts in sand: New design perspective.” *Journal of Geotechnical and Geoenvironmental Engineering*, 139(12), 2007-2019.
- Juang, C. H., and Wang, L. (2013). “Reliability-based robust geotechnical design of spread foundations using multi-objective genetic algorithm.” *Computers and Geotechnics*, 48, 96-106.
- Khoshnevisan, S., Gong, W., Wang, L., and Juang, C. H. (2014). “Robust design in geotechnical engineering—an update.” *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 8(4), 217-234.
- Khoshnevisan, S., Gong, W., Juang, C. H., and Atamturktur, S. (2015). “Efficient robust geotechnical design of drilled shafts in clay using a spreadsheet.” *Journal of Geotechnical and Geoenvironmental Engineering*, 141(2), 04014092.
- Low, B.K., and Tang, W.H. (2007). “Efficient spreadsheet algorithm for first-order reliability

- method.” *Journal of Engineering Mechanics*, 133(12), 1378-1387.
- Peng, X., Li, D. Q., Cao, Z. J., Gong, W., and Juang, C. H. (2016). “Reliability-based robust geotechnical design using Monte Carlo simulation.” *Bulletin of Engineering Geology and the Environment*, online.
- Phadke, M. S. (1989). *Quality Engineering Using Robust Design*, Prentice Hall. Englewood Cliffs, NJ.
- Professional Standards Compilation Group (PSCG). (2000). *Specification for Excavation in Shanghai Metro Construction*. Professional Standards Compilation Group, Shanghai, China.
- Sino-Geotechnics. (2010). *User Manual of Taiwan Originated Retaining Structure Analysis for Deep Excavation*. Sino-Geotechnics Research and Development Foundation, Taipei, Taiwan.
- Taiwan Geotechnical Society (TGS). (2001). *Design Specifications for the Foundation of Buildings*. Taiwan Geotechnical Society, Taipei, Taiwan.
- Zhao, Y.G., and Ono, T. (2000). “New point estimates for probability moments.” *Journal of Engineering Mechanics*, 126(4), 433-436.

Discussion

Petr Koudelka (Koudelka@itam.cas.cz, Czech Academy of Science, Czech Republic)

Robustness is a new and indefinite phenomenon in geotechnics. Chapter 9 contains an important and advanced well-done work in a range formerly marked as “Hic sunt leones”. Two robustness types and their definitions are given and analysed in Parts A and B. Comments and remarks relate to the Part A – accommodating events and actions that were not foreseen in design (par. 9.3):

1. **9.3.1 Human errors.** In my opinion, design should not take into account human errors either “grey swans” either “black swans”. Of course, a solid design should be so robust to accommodate **little** allowable geometric inaccuracies. An allowable level for human errors should be e.g. “splashy white swans”.
2. **9.3.2.1 Prescriptive measures.** Prescriptive measures referred to relevant robustness are very important and can provide considerable effects and increased robustness. Val (2006) notes on removal of a single vertical element do not appear to be entirely exact. The element and type of structure are not specified. Due to it, it is impossible to evaluate if the element removal would lead still to redundant or on the contrary, determinant structure or to a loss of structure stability.
3. **9.3.2.2 Partial factor methods.** Almost all structure design ranges except of geotechnical and plastic materials carry out analyses of structures in elastic state of the materials. All geotechnical tasks are very complexly non-linear due to extremely complex non-linear behaviour of soil shear strength (top and residual values). Models accommodating soils are also, of course, complexly non-linear. Those models may be analysed according to laws of mathematics and mechanics with the most probable material values which have behaviour near to a real structure. If we changed material values far from the most probable ones we analyse an entirely other model with entirely other behaviour.

All ULSD design approaches of EN 1997-1 (EC) are based on a definition of characteristic values and partial factors for material/soil properties those change a crucial soil property (shear strength) from the most probable value (50 % mean value) to almost improbable design value, e.g.:

Primarily consider soil variability coefficient $v=0.10$ and material partial factor $\gamma_m=1.25$ (corresponding to effective shear strength) then a probability of a worse design value occurrence is of 0.04 % which is of 125 times less than occurrence to the relevant characteristic value and of 2500 times less to the relevant most probable value.

Secondary consider variability coefficient $v=0.15$ and material partial factor $\gamma_m=1.4$ (corresponding to undrained shear strength) then a probability of a worse design value occurrence is of 0.11 % which is of 45 times less than occurrence of relevant characteristic value and of 900 times less to the relevant most probable value. Or consider $v=0.10$ and the probability is less of 0.01 % which is more than of 500 times less to relevant characteristic value probability and more than of 10, 000 times less to the relevant most probable value.

Therefore, in my opinion, it is possible it to state:

- a) Geotechnical designs according to the last EC are still too much conservative. It means, they should be very robust if site conditions agree with model. On the other hand this robustness would have not to be adequate if variation coefficient is higher or a model would not consider a significant event, e.g. existing erosion. From this point of view, prescriptive measures according to 9.3.2.1 (checking, review and supervision of design and construction) appear more important for an acceptable robustness than too conservative “usual” EC design.
 - b) Models according all Design Approaches of EC demonstrate different behaviours those are very dissimilar to behaviours of real soil masses. Hence, no model partial factors or factors close to of 1.0 would be necessary if the most probable material values would be considered.
 - c) More detailed comments on a concept of more applicable partial factor method see in following point 4 (to 9.3.2.5).
4. **9.3.2.5 Direct assessment of design values.** An apprehension on no provision for robustness of procedures using “worst credible values” or other similar procedures may not be true. A

concept of simpler and probably more suitable procedure to the EC approaches can be as follows:

- a) Definition of soil **material** design values: Design value is a cautious most probable value of the value set related to the structure. The term “cautious” could mean an elimination of favourable out layers or extremes, judgement of possibility of a properties change (e.g. water saturation) etc.
 - b) Design safety and reliability is ensured with a system of **partial factors for a resistance**. A base of the system could be taken according to former safety factor system that affords a lot of long-term experiences and could be improved using contemporary knowledge, e.g. probability based design.
 - c) Former system of partial factors for resistance also contains a stage of robustness that has been verified in long practice. New requirements for robustness or structure importance might to be taken simply into account by resistance factor increments.
 - d) The design value concept relates to **soil** material properties **only**. Others partial factors and characteristic values would not be changed.
 - e) The concept would be applicable for all methods and procedures, advanced including.
5. **9.3.3 Concluding remarks.** An apprehension on “worst credible”/most probable values may not provide the required robustness appears not well-founded. See point 4.

Conclusion:

A looking for more or less very improbable design material values for soils appears to be counter-productive. A clear and simple solution for all methods (FEM, BEM, simple methods, RBD) is application of the cautious most probable design values and relevant partial resistance factors.

Chapter 10 Future Directions & Challenges

10.1 INTRODUCTION

The ISSMGE Global Survey on the State of the Art (SoA) and the State of the Practice (SoP) in geotechnical engineering was conducted between 10 March and 30 April 2017. To quote the words of ISSMGE President Roger Frank in his message to the Society on 27 February, this survey is intended to support the main goal of ISSMGE, which pertains to “transferring to practice more results from the academic research and, reciprocally, help better feeding the academic research with the needs of practice”. The President noted that the “ISSMGE Technical Committees (TCs) are, obviously, the main tool for achieving this important goal.”

This Report constitutes a partial response from TC205 and TC304 to facilitate exchange of ideas between researchers and practitioners. It does not claim to be comprehensive in coverage or to offer a definitive resolution to any topics. It does assemble a fairly substantial portion of our recent literature that is regarded by various contributors as potentially useful to practice. The contributions collated in this Chapter are intended to inform activities for the next TC205 and TC304.

10.2 SUMMARY OF ISSMGE GLOBAL SURVEY OUTCOMES FOR TC304

Prepared by TC304 for ISSMGE Technical Oversight Committee

10.2.1 Background

TC304 identified the following ‘hot topics’ in the practice –

1. Value of reliability- and risk-informed decision making in practice
2. Geotechnical databases and probability models
3. Practical methods to manage geotechnical risk for real-world problems (semi-probabilistic, reliability-based, risk-informed)

Three questions related to the hot subjects are asked during the survey. The survey has received 192 responses, accounting for 14.83% of the total responses that ISSMEGE received during this survey. The three questions are as follows.

Q1: Have you adopted a reliability- or risk-informed method for design, assessment, and/or management of a geotechnical project?

Q1a: If YES to the first question, provide some project details and highlight how your method complements the factor of safety approach?

Q1b: If NO to the first question, what are your reservations or concerns?

Q2: In relation to bridging state-of-the-art and state-of-the-practice, which aspect(s) of reliability- or risk-informed decision making deserve the most attention (e.g. role of engineering judgment, selection of characteristic value, statistical interpretation of site data and other geotechnical data, target reliability/risk level, Bayesian observational approach, codification versus site-specific needs)?

Q3: What are the top three items in your 'wish list' that you feel would facilitate adoption of reliability or risk-informed decision making in geotechnical practice (e.g. statistical guidelines, databases, software, short courses, clearer design standards/codes, case studies, bibliography)

Q1 is used to identify whether the participants have experiences on reliability and risk informed methods. Q1a and Q1b are follow-up questions for Q1. Q1b is important, because it may identify possible causes for the gap between SOA and SOP. Q2 and Q3 are used to further identify possible solutions that can bridge the gap.

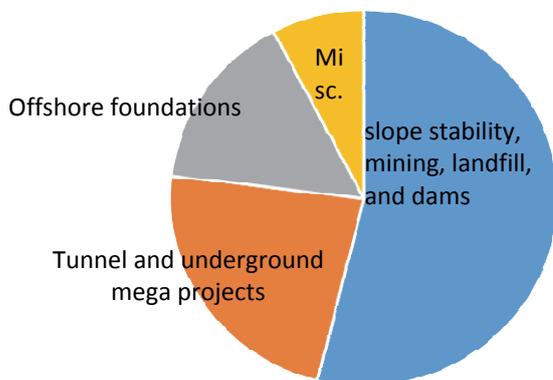
10.2.2 Summary of survey results

Q1 is about whether the participants have used reliability- and risk-informed methods in practice. 29 participants responded and about 65% are positive.



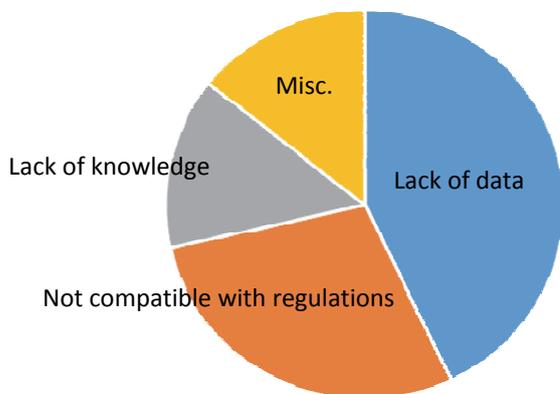
Q1: Have you adopted a reliability- or risk-informed method for design, assessment, and/or management of a geotechnical project?

Q1a received 15 replies. The following areas/subjects are among the most frequent ones where reliability- and risk-informed methods are applied: (a) slope stability, mining, landfill, and dams (7 replies); (b) tunnel and underground mega projects (3 replies); (c) offshore foundations (2 replies).



Q1a: If Yes, provide details and highlight how your method complements the FOS method?

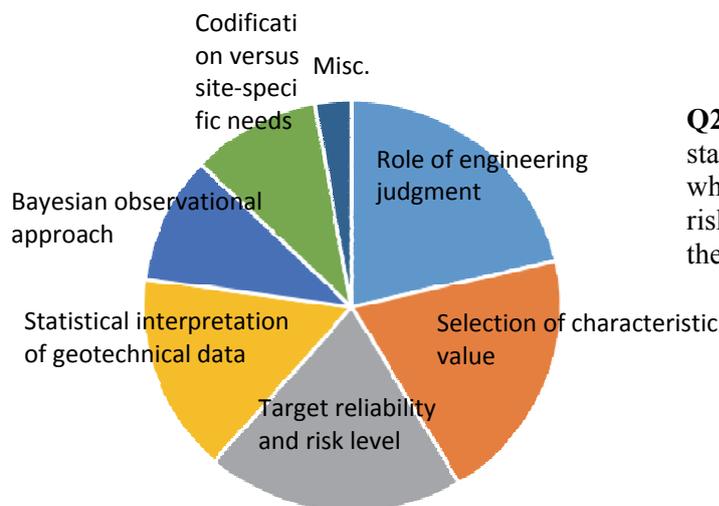
Q1b is about the possible causes for the gap between SOA and SOP. It received 7 replies. The following causes are among the most frequent ones: (a) lack of data (3 replies); (b) not compatible with regulations (2 replies); and (c) lack of knowledge (1 reply).



Q1b: If No, what are your reservations or concerns?

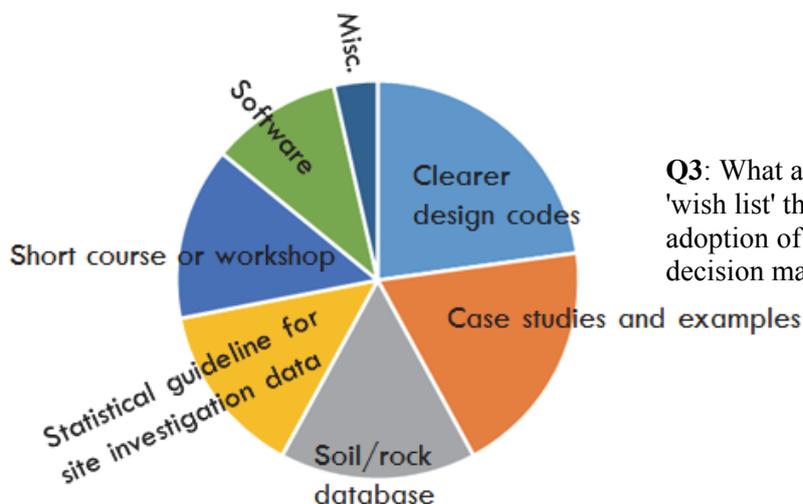
Q2 is about the aspect of reliability- or risk-informed methods that deserves the most attention, in order to bridge SOA and SOP. These aspects can be viewed as actions that can improve SOA. 70 participants responded. The following aspects are among the most frequent ones: (a) role of engineering judgement (15 replies); (b) selection of characteristic value (14 replies); (c) target

reliability/risk level (14 replies); (d) statistical interpretation of geotechnical data (11 replies); (e) Bayesian observational approach (7 replies); and (f) codification versus site-specific needs (7 replies).



Q2: In relation to bridging state-of-the-art and state-of-practice, which aspect(s) of reliability or risk-informed decision making deserve the most attention?

Q3 is about the top three items in wish list to facilitate adoption of reliability- or risk-informed methods. These items can be viewed as actions that can improve SOP. The following items are among the most frequent ones: (a) clearer design codes (13 replies); (b) case studies and examples (11 replies); (c) soil/rock database (9 replies); (d) statistical guideline for site investigation data (8 replies); (e) short course or workshop (8 replies); and (f) software (6 replies).



Q3: What are the top three items in your 'wish list' that you feel would facilitate adoption of reliability or risk-informed decision making in geotechnical practice

10.2.3 Analysis of survey results

Q1: use of reliability- and risk-informed methods in practice

There are about 65% of participants have actually applied reliability- and risk-informed methods. This may not represent the actual condition for two reasons. First, the percentage for participants with academic background in the survey may be higher than normal. Second, participants for the survey may be already interested in the reliability- and risk-informed methods.

Q1a: areas with most application

The survey results are largely consistent with existing practice: reliability- and risk-informed methods have been implemented to areas such as slope stability, mining, landfill, embankments, dams, and tunnels. These are areas that traditionally welcome reliability- and risk-informed methods. Areas where reliability-based design codes are available, e.g., foundation design, are

not in the list probably because practitioners only need to follow the codes. It is interesting to see emerging areas such as offshore engineering and mega city engineering also welcome reliability- and risk-informed methods.

Q1b: causes for the gap between SOA and SOP

It is very interesting to see that the lack of data is a main cause for the gap. This raises a red flag to the recent research trend that focuses on models and methodologies. Indeed, reliability- and risk-informed methods heavily rely on data. Without real data, models and methods have no use. Insufficient background knowledge is another cause for the gap. This indicates that education is important.

Q2: bridging SOA and SOP – actions to improve SOA

Engineering judgment is an important element for geotechnical engineering, no matter reliability methods are adopted or not. The participants for the survey acknowledge this importance. Selection of characteristic value and determination of reliability level are examples of exercising engineering judgment. However, researchers (SOA) tend to focus more on developing new models and methods but less on the role of engineering judgment. There seem to be needs to provide clearer guidelines to facilitate engineering judgment, such as clearer guidelines for selecting characteristic value and target reliability level. There also seem to be needs to develop guidelines for statistical interpretation of geotechnical data.

Q3: wish list – actions to improve SOP

Geotechnical engineering practice is deeply connected with design codes. The participants express the sentiment that current design codes are not clear enough. One possible factor is that the principle and steps for the selection of characteristic value are unclear. Another one is that the statistical guidelines for analyzing geotechnical data are unclear. Soil and rock databases are essential for reliability- and risk-informed methods. The participants express the sentiment that such databases may not be available. Education (in terms of application examples, short courses, workshops, etc.) is another item that needs more attention. Finally, there seem to be needs for software.

Lessons learned from survey results

The survey results are stimulating. Based on the results, TC304 plans to focus on the following activities during the next term (from 2017 to 2021):

1. Design codes. TC304 has been interacting with TC205 in the discussion of possible use of statistical methods in future Eurocode 7 in a series of activities. TC304 will continue this line of activities. The selection of characteristic value and statistical guideline for site investigation data will be possible subjects.
2. Soil/rock databases and softwares. Task forces will be established within TC304 to develop soil/rock databases. These databases will be made available to the public through web-based or app-based softwares. These softwares will also be able to conduct statistical analysis for the soil/rock data.
3. Education. TC034 will organize short courses and workshops to educate engineers about the following aspects: (a) reliability and risk methods; (b) statistical methods for site investigation data; (c) case studies and application examples.

10.3 SUMMARY OF ISSMGE GLOBAL SURVEY OUTCOMES FOR TC205

Prepared by TC205 for ISSMGE Technical Oversight Committee

10.3.1 Issues covered

TC205 took the opportunity to examine three areas, considered to be most critical in current debates:

- How safety is currently being prescribed in practical design
- What limit states are considered to be most important
- How input for design is derived from ground investigation and load testing.

10.3.2 Overview of the survey respondents, demographics etc

The responses were interesting and informative, although, with an average of 14 responses per question it is difficult to gauge whether they are representative of the profession, or of any sector or geographical region within it. All questions were answered. 55% of the responses to the survey as a whole came from Europe (Q2). It is perhaps surprising that most of the responses came from

“industry”, with only 28% of responses from academic institutions (Q3). These issues may have biased the balance of the answers, and it is unfortunate that the sources from which responses came to TC205 specifically are not known.

10.3.3 Summary of responses received

Safety formats

It was clear that a majority of respondents (80%) now consider use of limit state design with partial factors to be “normally used in practice” (Q86). A few considered older methods such as global factors or permissible stress design to be normal, while only one thought use of reliability calculations is normal practice. This changed very little when respondents were asked about “the next decade or so” (Q87). There was divided opinion about whether different safety formats are used for more critical design situations (Q88).

About two-thirds of respondents thought that very little attention is given to serviceability limit states, in comparison with that given to ultimate limit states (Q89). About two-thirds of these thought this was inappropriate, presumably indicating that more attention should be given to serviceability (Q90). The question whether this situation is likely to change over the next decade drew a wide range of responses, with an average of 50%, i.e. undecided overall (Q91).

Surprisingly, considering the answers to Q86 (above), the average view of percentage of large civil engineering projects for which reliability calculations are used in practical design was 23% (Q92), while about half the respondents said that reliability calculations are used for other purposes such as national calibration of codes of practice (Q94). Asked whether they expected this situation to change over the next decade, the average response was “37%”, indicating a lack of expectation that much will change (Q93).

25% of respondents considered that Performance Based Design is adopted in practice (Q95), with an average of “30%” indicating a lack of expectation that much will change over the next decade, (Q96). (Caution: within the last few months TC205 has discussed Performance Based Design at some length, and it is clear that there are considerably different understandings of the meaning of this term. This makes the response to this question difficult to interpret.)

Respondents were asked to rate “how often are national standards generally followed literally” (0) compared with “other calculation methods commonly used (100). The average response was 27, with a large scatter – standard deviation of 29 (Q97). 25% of respondents answered “0” and one answered “100”. We take this to mean that national standards are most often not followed very literally, but presumably used for guidance.

Ground investigations and characteristic values

Almost all respondents said that borehole sampling and testing is commonly used, with about two-thirds also selecting CPTs and two thirds selecting SPTs, including some who selected both of these (Q98). A third of respondents said that field tests such as pressuremeters are common, with one mention of shear wave testing.

Asked how often results of site investigation are used to generate design resistances (0), rather than to derive soil properties (100), the average answer was “65” (Q99), indicating that they are used to derive soil properties slightly more often than to derive resistances directly. A variety of different approaches was used to provide margins of safety when resistances are derived directly from test results (Q100). For calculations based on soil strengths, partial factors applied to material strength were most commonly applied; there were some other alternatives, but the use of factors on resistances was not mentioned in this context (Q101).

Both advantages and disadvantages were noted for direct derivation of resistances from ground tests and calculations based on soil parameters (Q102, Q103). It was noted that direct derivation of resistances removes one step of uncertainty, but generally respondents thought that use of soil parameters was more achievable and gave the engineer greater understanding and control.

Respondents unanimously agreed that insufficient site investigation is normally undertaken in terms of cost-benefit (Q104).

10.3.4 Significance of the survey for TC205

Overall, the results of the survey, including Q244 (asking for general feedback on narrowing the gap between SoA and SoP), support themes that are already being followed by TC205:

- Partial factoring is the dominant method of prescribing safety, and is likely to remain so for the foreseeable future. Refinement of this process therefore remains important.
- Cautious development of reliability methods, balancing interest in exploring benefits with a degree of scepticism.
- Greater clarity in the process of derivation of “characteristic values” is desired.
- There is a perceived need to develop and improve serviceability limit state design.

The results of the survey were interesting, but the number of respondents was small (average 14) and their demographics are unknown. A future survey could usefully overcome these two obstacles. While this might partly be achieved by the members of the TC being more proactive in encouraging responses, that would also introduce bias into the results, so it is desirable that CAPG somehow find a way of increasing the uptake of a survey of this type.

TC205 provides a point of contact between people currently working in both SoA and SoP, both of which are under active development. The biggest gap is probably between academics working on reliability methods and practitioners who, as indicated by the survey, are cautious about adopting such methods.

10.4 SOIL/ROCK PROPERTIES

Jianye Ching, jyching@gmail.com

10.4.1 Generic versus site-specific transformation model

An engineering practice exercised by practitioners for years is to estimate the design parameter (such as the friction angle of a sand) using transformation models. In spite of their popularity, transformation models are frequently criticized because they are constructed by non-site-specific data and may not be applicable to a specific site. However, if we narrow down to a specific site, there are insufficient data points to construct the transformation model. A dilemma exists here: a generic model is not plausible, yet a site-specific model is not feasible. Does there exist an intermediate transformation model that is neither 100% generic nor 100% site-specific, but is far more plausible than a generic model and far more feasible than a site-specific model? I believe that developing such intermediate transformation models is an important future research direction.

10.4.2 Rock engineering

We have not explored much about rock properties and the related transformation models. Rock properties seem to be more variable and less predictable than soil properties. I believe that developing rock databases and derive probabilistic transformation models for rocks are important future research directions.

10.4.3 Web-based software and app

It is important to get practitioners involved in the subject of “transformation models and soil/rock databases”. Without practitioners’ broader involvement and feedbacks, the true usefulness and true limitations of these models will never be explored. It is now timely to develop web-based software or app for practitioners to use. Researchers should communicate with practitioners during the development process to understand practitioners’ needs, and practitioners should feedback with their concerns. It will be great if the software or app can serve as an interface with which practitioners can upload their soil/rock data points to broaden the coverage of the soil/rock databases. Some screening mechanisms are needed for this purpose to prevent problematic data points.

10.5 CHARACTERISTIC VALUE

Yu Wang, yuwang@cityu.edu.hk

How to use statistical method to facilitate the selection of geotechnical property design profile (e.g.

characteristic value profile for soil properties) when a reasonable number of data points along depths are available from site characterization? The current engineering practice relies heavily on engineering judgement, which is subjective and may be different for different individuals. Because of this subjective selection and individual judgment, the characteristic value profile, derived by different geotechnical engineers even from the same dataset, may vary greatly, especially when the test data contains significant variability. For example, Bond and Harris (2008) asked different geotechnical engineers to assess the characteristic parameter values of London and Lambeth clays from the results of Standard Penetration Tests (SPTs) carried out in these soils (Orr 2017). The characteristic SPT value profiles selected by the different engineers are shown in Figure 10-1 by solid lines. The wide spread of solid lines in Figure 10-1 demonstrates effect of subjective judgment on the selected characteristic value profile and highlights the need of objective methods for selection of characteristic value profile.

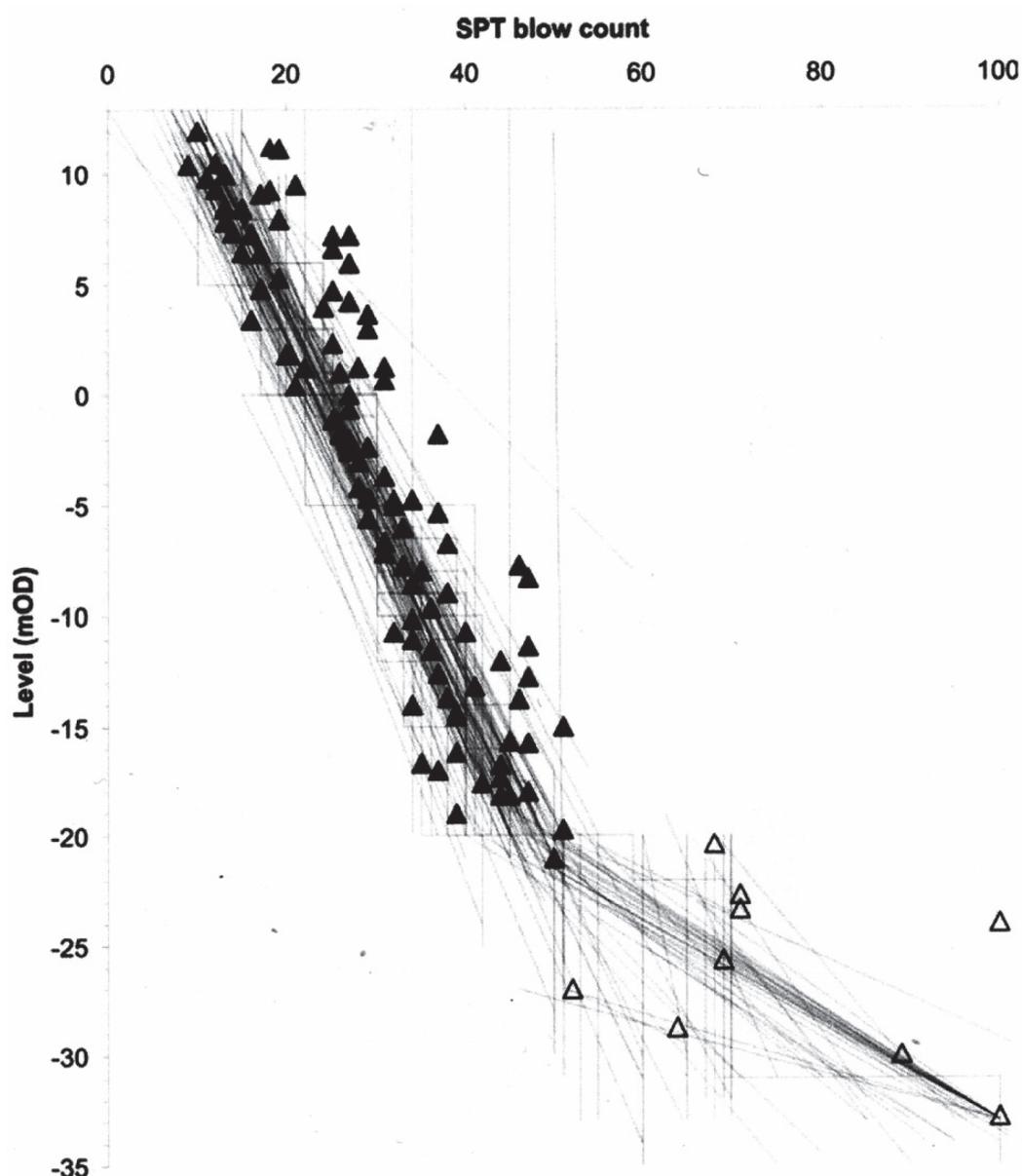


Figure 10-1 Illustrative example of engineers' interpretations of characteristic value profile (Bond and Harris 2008; Orr 2017)

Table 10-1 provides the original data for Figure 10-1 (Bond 2017). Interested readers are welcome to use the original data to interpret their own characteristic value profile.

Table 10-1 Original data for Figure 10-1 (Bond 2017)

Level (mOD)	SPT (N)	Level (mOD)	SPT (N)	Level (mOD)	SPT (N)
11.25	18	-5.55	29	-10.15	34
9.25	18	-8.55	34	-13.15	41
7.25	25	-11.55	36	-16.15	39
4.25	27	-14.55	39	-18.15	44
1.25	28	-17.55	42	8.45	15
-1.75	26	10.5	12	7.3	16
-4.75	32	8.5	13	4.7	25
-7.75	46	6.5	17	1.3	22
-10.75	40	4	24	-1.7	26
-13.75	46	1	26	-4.7	29
-16.75	44	-2	27	-7.7	33
9.5	21	-5	32	-10.7	32
7.25	27	-8	35	-13.7	38
4.25	27	-11	34	-16.7	35
1.25	31	-14	34	-19.7	51
-1.75	37	-17	37	9.85	11
-4.75	34	-19	39	7.85	13
-6.75	38	12	10	5.35	19
-12.75	47	10	13	2.35	25
-15.75	47	8	19	-0.65	27
11.15	19	6	27	-3.65	31
9.15	17	3	29	-6.65	31
6.65	25	0	27	-9.65	36
3.65	29	-3	28	-12.65	37
0.65	31	-6	33	-15.65	45
-2.35	29	-9	38	-18.15	45
-5.35	37	-12	44	-20.35	68
-8.35	47	-15	51	-23.35	71
-11.35	47	-18	47	-24	100
-14.35	39	-21	50	-27	52
-17.35	44	9.35	12	-30	89
10.45	9	7.35	14	-32.85	100
8.45	13	4.85	17	-22.7	71
6.45	15	1.85	20	-25.7	69
3.45	16	-1.15	25	-28.7	64
0.45	21	-4.15	28		
-2.55	27	-7.15	31		

References

- Bond, A. (2017). Personal communication.
 Bond, A., and A. Harris. (2008). Decoding Eurocode 7. London: Taylor and Francis.
 Orr, T. L. L. (2017). Defining and selecting characteristic values of geotechnical parameters for designs to Eurocode 7, Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 11(1), 103-115.

10.6 MONITORING AND DEALING WITH COMPLEX ROCK FORMATIONS

Giovanna Vessia, g.vessia@unich.it

10.6.1 Geotechnical structure performance against natural hazard

How to check the evolution in time of the geotechnical structure performances based on both monitoring data and failure predictions represents a challenging issue to be addressed in the future. Dams, infrastructures and buildings set in hazardous territories (threaten by Natural hazards) must be

monitored for public safety and the temporal datasets collected must be used to predict soil-structure interaction in order to prevent unpredicted failures. To this end, not only the structural integrity must be checked but also soil parameters and their spatial changes in values must be re-evaluated through deterministic and statistical approaches based on proper monitoring datasets (e.g. the strain gauge measures within an earth dam) or investigation campaigns performed before and after the natural event (e.g. earthquakes, floods).

10.6.2. Rock-soil engineering

The evaluation of the quality of rock masses and their stability (within the mass in tunnel engineering or on its front in rock slope instability) is nowadays based on classification criteria (like Bieniawski, Barton and GSI classification) and numerical simulations. They are developed for hard and weak rock masses but not for those complex formations (such as flysch rocky-soil deposits) that show a high spatial variability in soil and rock distribution and properties. Try to find out effective numerical and statistical tools to dealing with (1) mechanical characterization and (2) failure predictions of complex formations in addition to engineering judgement and monitoring activities should be tackled in the future.

10.7 CLOSING THE GAP BETWEEN RESEARCH AND PRACTICE

Kerstin Lesny, *kerstin.lesny@hcu-hamburg.de*

To my opinion, one of the biggest challenges still is to close the gap between theoretically based, complex approaches as provided by RBD and the design practice applied by the ‘normal’ geotechnical engineer who is usually not familiar with probabilistic methods of any kind. Maybe, we can also say to provide a bridge or a link between the different ‘languages speaking’ or different ‘ways of thinking’ of the different parties. One may argue that the application of RBD in practical design has moved a big step forward as there are meanwhile tools or methods available which can be easily applied. This may be correct, but the question remains for whom it is easy - for those, who are familiar with it anyhow or also for those which do not have any experience? People need to understand and they need to be convinced about the benefits of a new method compared to the usual one. They have to be provided with good arguments, realistic assessments, and illustrative examples (simple, but probably not too simple to be unrealistic).

I'm writing this with the background of the ongoing evolution of Eurocode 7. At least in Germany there is a quite strong opposition especially in the construction industry and the consultancy against any complication of the current EC7 version.

10.8 SENSITIVITY ANALYSIS BY USING IMPRECISE PROBABILISTIC AND INTERVAL APPROACHES

Sónia H. Marques, *sonia.h@fe.up.pt*

EN 1997, adopted as Eurocode 7, is intended to be applied to the geotechnical aspects of the design of civil engineering works. The limit state design concept adopted by Eurocode 7 is used in conjunction with a partial factor methodology. The selection of appropriate partial factors is important to ensure the reliability of geotechnical design to Eurocode 7, as design values are determined by applying partial factors to characteristic values. When the partial factor format is first introduced, it should preferably produce a design comparable to the resultant from the safety factor methodology, promoting the continuity of past experience. Actually, the partial factor format through the Design Approach DA.2* is based on a modified global safety concept, noted that different systems associated to the same factor may have a different probability of failure due to the fact that important variabilities are disregarded. At the present, the performance of a partial factor format is measured by the ability to produce a design achieving a desired target reliability within acceptable margin of error. To achieve the required target reliability, Eurocode 7 does not provide any variation in the partial factors but rather requires that greater attention is given to other accompanying measures related to design supervision and inspection differentiation by a system of failure control. The issue of adopting multiple resistance partial factors in geotechnical design is then on discussion. At a glance, multiple

resistance partial factors should be clearly related to the determination of characteristic values, primary cause of inconsistent reliability evaluations. Considered the underlying motivation on the endeavour for safety, approaches to robust design are nowadays discussed on the calculation of resistance factors capable to maintain a more uniform reliability level over a range of design parameters. As information is not certain but rather imprecise, a sensitivity analysis may be pursued through probabilistic grid-based and fuzzy-based approaches on multiple intervals. Thereby, the imprecise probability theory emerges as a basis for decision-making when providing the investigation of the most plausible models. The key feature concerns on the identification of probability bounds for scenarios of interest which reflect the uncertainty as the range between the limits. The imprecise probabilistic analysis is then a supplementary element which enriches the variety of models to be combined with the traditional overview in improved adaptability. These models may include interval and fuzzy probabilities in association to a probability box structure constructed from search amid the competing models. The admissibility of imprecision in beliefs is the primary difference in motivation between imprecise probabilistic and robust Bayesian approaches, beliefs are imprecise or there exists a prior that captures the true beliefs, although it may be hard to identify this distribution. Imprecise probabilistic approaches are robust whenever insensitive to small deviations from the assumed probabilistic models. One motivation for adopting imprecise probabilistic approaches consists on the safety evaluation in the scope of the shear strength parametric change by weathering on real world problems such as the probabilistic interpretation of mine pillar capacity. On the imprecise probabilistic approaches multiple resistance partial factors may be clearly related to the determination of characteristic values. Expressed simply by intervals, geotechnical parameters on scarce probabilistic information are assigned based on experience. A meaningful interpretation on a high dimensional space is based on the joint analysis of multiple cases instead of a lower and upper probabilistic evaluation. The limit state imprecise probabilistic analysis may be interpreted altogether with a limit state imprecise interval analysis. This sensitivity analysis may provide meaningful results for safety-based decision in ground investigation and testing or improvement and the partial factor design may be discussed on the basis of distinct levels of credibility.

10.9 ROBUSTNESS

Brian Simpson, *Brian.Simpson@Arup.com*

Actual probabilities of failure are of importance to society at large, including the general public, governments and insurers. A notional probability of failure or reliability index does not provide this function, especially if, as is widely believed, actual failures are most often caused by events and actions outside the range of the anticipated variations of leading variables – loads, strengths etc.

Robustness relates to the provision in design for unforeseen events and actions, including human error, of a magnitude that society would expect the construction to withstand without becoming dangerous. In the design process, the uncertainty of the leading variables is obvious, but over-concentration on these can distract from the need to provide a margin of safety that could accommodate the unexpected.

Correlation with past practice plays a significant role in determining the values given to partial factors, so it is likely that they provide a degree of robustness that is acceptable to society. It is important that this topic should be addressed in the development of reliability methods if they are to give a meaningful indication of actual probabilities of failure.

10.10 MACHINE LEARNING

He-Qing Mu, *cthqmu@scut.edu.cn*

Machine learning can be defined as a set of methods for data pattern recognition, future data prediction and decision-making under uncertainty [1]. The fields of machine learning and pattern recognition have undergone substantial development over the past decades. It is known that many problems in civil and geotechnical engineering are associated with significant levels of uncertainty. Thus, providing a rigorous solution for uncertainty quantification is necessary when applying the machine learning techniques to inference problems in civil and geotechnical engineering. In particular, Bayesian-based machine learning, which is a systematic inference, prediction and decision-making framework embedding the advantages of both probabilistic reasoning and pattern recognition, has become

increasingly popular. Recently, sparse Bayesian learning methods have been applied to civil [2] and geotechnical engineering [3] problems. However, it is accurate to say that the power of these methods have yet to be appreciated in the mainstream geotechnical reliability literature. One can imagine geotechnical engineering practice being transformed by the combination of machine learning and the Internet of Things to support decision making in the face of uncertainty in real time as discussed in the Minutes for 25th ISSMGE TC304 Meeting on 4 June 2017.

1. Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.
2. Mu, H. Q., & Yuen, K. V. (2017). Novel Sparse Bayesian Learning and Its Application to Ground Motion Pattern Recognition. *Journal of Computing in Civil Engineering*, 31(5), 04017031.
3. Ching, J., & Phoon, K. K. (2017). Characterizing uncertain site-specific trend function by sparse Bayesian learning. *Journal of Engineering Mechanics*, 143(7), 04017028.