Better correlations for geotechnical design

Kok-Kwang Phoon
National University of Singapore

Jianye Ching
National Taiwan University

ABSTRACT: One of the most important task in geotechnical design is the estimation of pertinent design soil parameters, but it may not be feasible to measure these parameters directly or there are too few direct measurements to get a sense of their variabilities (this is important for sensitivity analysis). The tradition of geotechnical engineering is steeped in empiricism and one notable aspect is arguably the widespread application of generic correlation models to estimate site-specific design parameters from indirect data, such as the results of laboratory index tests and field tests. These correlations (say between undrained shear strength and SPT-N) are invariably associated with significant transformation uncertainties, which are usually presented in the form of data scatter about the correlation line. This paper highlights the availability of extensive multivariate databases for clays, sands, and rocks that can be fruitfully exploited to develop global multivariate probability models. The improvements offered by these models (derived from a theoretically powerful multivariate normal framework) over existing bivariate correlations are: (1) *multiple* design parameters can be estimated *simultaneously* from *multiple* sources of indirect data, (2) precision of the estimates can be quantified using the 95% confidence region (non-trivial generalization of the well-known 95% confidence interval to multiple possibly correlated design parameters), (3) entire *multivariate* posterior distributions can be updated by combining these global models (treated as prior distributions) with site-specific data – this is significantly better than relying on pure judgment to combine site-specific data with prior experience, (4) serve as the calculation engine for SPM2 (Soil Properties Manual 2) – a software undergoing development that will obviate the need to learn multivariate probability theory and Bayesian updating, and (5) *monetizing* the value of information in site investigation by establishing an explicit link between type/number of tests conducted at a site, reduction of parameter estimation variability, and resulting design savings (if reliability-based design permitted in ISO2394:2015 were to be adopted).

1 INTRODUCTION

One of the most important task in geotechnical design is the estimation of pertinent soil parameters, particularly the values governing the behaviour of a geotechnical structure at a limit state. There are at least two important aspects that one should consider, explicitly or otherwise. First, it may not be feasible to measure the desired soil parameters directly or the budget set aside for site investigation does not allow sufficient undisturbed samples and direct laboratory tests to be conducted (ideally, matching in-situ stress state, stress path, strain rate, drainage, etc.). It is important to perform a sufficient number of tests to get a good sense of ground variation and precision of design parameters. For example, the BCA/IES/ACES advisory note 1/03 on site investigation and load tests provided the following guidelines:

Site investigation should be carried out to sufficient extent and depth to establish the significant soil strata and ground variation.

(a) The number of boreholes should be the greater of (i) one borehole per 300 sq m or (ii) one borehole at every interval between 10m to 30m, but no less than 3 boreholes in a project site.
(b) Boreholes should go more than 5 metres into hard stratum with SPT blow counts of 100 or more than 3 times the pile diameters beyond the intended founding level.

Second, a geotechnical structure at a limit state typically interacts with a "zone of ground" and it is the *mobilized* values in this zone that one should consider in design. EN 1997−1 (2004), Clause 2.4.5.2 - Characteristic values of geotechnical parameters describes this "zone of ground" in the following two application rules:

(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.

(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.

The above two application rules essentially require the engineer to have a sense of how a geotechnical structure will behave *prior* to analysis. Hence, a common caveat in this estimation exercise is that engineering judgment/experience is vital – this exercise is not a matter of selecting numbers from a bore log arbitrarily and/or relying on statistics as a proxy for correct understanding of the soil-structure interaction behaviour. At the least, it would require the engineer to appreciate qualitative issues such as parameter values outside the approximate zone (say SPT-N values below a depth of B or 2B for a footing of width B) are less influential, special zones (low strength, highly compressible, discontinuities, etc.) may exercise a disproportionate effect on behaviour, and soil-structure interaction as described in Clause 2.4.5.2(9) above. Clause 2.4.5.2 is cognizant of the close inter-connection between the ground, soil behaviour, and modelling in geotechnical engineering practice and the role of experience (supported by empiricism and precedent) and risk management in mediating these elements. Figure 1 shows a revised Burland Triangle (origin version attributed to Burland 1987) that illustrates the centrality of experience and risk management (Lee Barbour & Krahn 2004).

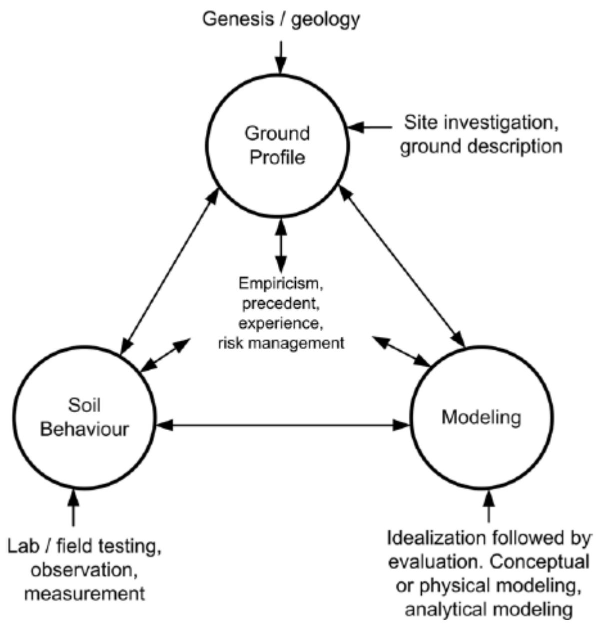


Figure 1. Revised Burland Triangle (Lee Barbour & Krahn 2004).

Although the notion of a "characteristic value" as described in Clause 2.4.5.2 is sensible, it has defied many attempts to clarify it to such an extent that engineers could roughly agree on the actual numerical value/profile when presented with the *same* bore log information (Bond & Harris 2008). The difficulty arises to a large extent from the spatial variability of the ground. The interaction between spatial variability and formation of critical slip surfaces is complex – the applicability of classical failure mechanisms and their solutions in standard texts is questionable given the standard homogeneous ground assumption underlying these solutions. Clause 2.4.5.2 attempts to cover this complex topic in application rule (11):

(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%.

NOTE 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.

Recent research has demonstrated that this application rule is inadequate for the complexity of the problem and the range of spatial variability encountered in practice (Ching & Phoon 2013a, 2013b; Ching et al. 2014a, 2016a, 2016b, 2016c, 2017a, 2017b; Hu & Ching, 2015). In addition, it is silent on how multiple correlated soil parameters in finite element analysis should be selected (Ching et al. 2017c). A realistic assessment of multiple characteristic values in the context of spatial variability where non-classical failure mechanisms can emerged is certainly beyond the reach of judgment/experience uninformed by analysis (Phoon 2017).

For the purpose of this paper, it suffices to note that the above complications mainly arise from the "zone of ground governing the behaviour of a geotechnical structure at a limit state". The consideration of this zone of ground in relation to a limit state is vital in design, but to keep the scope of this paper manageable, we will set this second aspect aside and focus solely on the first aspect, which is to estimate the pertinent soil parameter *at a "point"* in the ground (typically at a desired depth) using correlations. Hence, the spatial variation of a soil parameter is not considered in this paper. Because the estimation of a characteristic value relevant to design is beyond converting a field test result to a design parameter (say SPT N-value to undrained shear strength) in a depth-wise manner, it is not surprising to find the following Principle in Clause 2.4.5.2:

(1)P The selection of characteristic values for geotechnical parameters shall be based on results and derived values from laboratory and field tests, complemented by well-established experience.

The role of engineering judgment and experience, particularly how it complements statistical analysis, is explained in Phoon (2007) and would not be repeated herein. The purpose of this paper is to suggest an improvement to our ubiquitous correlation models. These correlation models are very useful in practice, because engineers can obtained an estimate of a soil parameter pertinent to design (called "design parameter") using more commonly available data that are indirectly related to this design parameter, say data from a laboratory index test or a field test. In fact, these correlation models grew in popularity as a result of applying more rational soil mechanical principles to engineering practice. It is well appreciated that a rational model is only better than an empirical model if its input parameters can be reliably estimated (Lambe 1973). An early example is the estimation of the compression ratio from natural water content (Figure 2). It is worth pointing out that the majority of these correlations is bivariate in the sense of estimating one desired design parameter (such as compression ratio) from one indirect source of data (such as natural water content). A minority is multivariate in the sense of estimating one design parameter from multiple sources of data (such as natural water content and plasticity index). This multivariate correlation can appear in an approximate form of a series of bivariate correlations indexed by a secondary parameter. To our knowledge, none has considered *simultaneous* estimation of two or more design parameters from two or more sources of data. This requires deriving a conditional distribution from a multivariate probability framework so that (loosely speaking) the dependencies between all sources of data are accounted for.

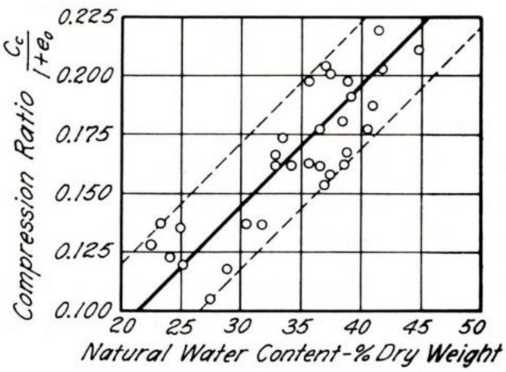


Figure 2. Correlation between compression ratio and natural water content [Source: Fig. 45.9, Terzaghi & Peck (1967) citing Fadum (1941)]

The authors are of the opinion that our existing bivariate correlation models can be significantly improved by extending them using a multivariate probability framework. Although this framework is abstract and unfamiliar to most engineers, it can offer the following practical advantages:

1. Multivariate transformation models can be obtained. For example, one can derive a relation between the compression ratio and multiple sources of indirect data, say liquid limit, plasticity index, initial void ratio, natural water content, cone tip resistance, and dilatometer modulus. Figure 2 is a "bivariate" model in the sense that the design parameter is estimated from one source of data (natural water content). The natural generalization from the existing bivariate case to the multivariate case is called a "transformation model" from hereon.

2. These transformation models not only can predict the mean value of the design parameter but also can predict its coefficient of variation (COV). It is clear from Figure 2 that estimating the mean value (bold line) is not sufficient. It is natural to expect all transformation models to contain transformation uncertainty (scatter of the open circles about the bold line) and this uncertainty must be quantified explicitly as it has an impact on design. For example, the mean estimate of the compression ratio for a natural water content of 35% is 0.17. However, it can take values between 0.14 and 0.20 (dashed lines). Adopting a compression ratio of 0.17 for design is not conservative, but adopting a value of 0.20 is arguably overly conservative as the dashed lines appear to indicate best and worst case estimates within the sample size of 30 measurements. The question of which estimate is appropriate in the face of uncertainty is best discussed in the context of reliability-based design (RBD) (Phoon & Retief 2016), which is outside the scope of this paper. It suffices to note that geotechnical RBD is permitted in Annex D of ISO2934:2015 (International Organization for Standardization 2015; Phoon et al. 2016) and the COV is needed for this

simplified risk-based design approach. It is also needed for sensitivity analysis even within the remit of our existing deterministic practice in the form of upper/lower bounds, e.g. mean plus/minus one standard deviation or 95% confidence interval of the compression ratio.

3. The multivariate probability model can be used as a proper and empirically supported prior distribution to derive the posterior distribution of design parameters based on limited but site-specific laboratory/field data. This prior is clearly much preferred over a pure judgment-based or an uninformative prior. Note that the entire multivariate distribution of multiple design parameters is derived, not simply means and COVs. When a multivariate distribution is available, multiple design parameters can be updated *simultaneously* from multiple laboratory/field measurements. This formal Bayesian updating approach is significantly more advantageous than relying on judgment alone to combine site-specific data with prior experience.

4. The practical significance of updating is that biases and variabilities are generally reduced in the presence of site-specific data and it is possible to link this improvement to design savings explicitly in the context of RBD. By *monetizing the value* of site information, it will be easier for engineers to present site investigation to clients as an investment, rather than allowing it to be unjustifiably stigmatized as an un-necessary cost (Phoon & Ching 2013a). The fact that many clients choose the minimum site investigation mandated by building regulations argues against the prevalent belief that geotechnical needs alone with an indirect link to safety are sufficient to convince clients to choose a more appropriate but more costly site investigation programme.

This paper presents a summary of global clay/sand/rock databases and a description of the parameters/sites covered to dispel the belief that there are insufficient data to support the practical construction of such multivariate probability models. In short, it is possible to derive better transformation models for practice that will further enhance and justify the cost of site investigation – this is not theorizing for theorizing sake. Correlation between two soil parameters and its generalization to a correlation matrix to capture correlation between more than two soil parameters are briefly explained as these concepts are key to the construction of multivariate probability models. More complete theoretical details are given elsewhere (Ching & Phoon 2015a; Ching et al. 2016d). The biases and COVs of existing transformation models for clays, sands, and rocks are summarized as one useful direct outcome of this multivariate database compilation exercise. An on-going software development effort, called SPM2 (Soil Properties Manual version 2) is described to provide a glimpse of useful tools that an engineer can gain access to on his/her mobile to make better decisions on the choice of design soil/rock parameters. The name "Soil Properties Manual" is inspired by the seminal work done by Kulhawy & Mayne (1990) which is considered as "version 1". It is not necessary for the engineer to learn multivariate probability theory and Bayesian updating to put SPM2 to good use in practice. After all, it is not necessary for the engineer to learn the Galerkin method and finite element theory to compute internal forces and displacements. As pointed out above with reference to BCA/IES/ACES advisory note 1/03, the basic role of the engineer is to "establish the significant soil strata and ground variation" from all available site investigation data, precedents, and experience. SPM2 is intended as a supporting tool to enable the engineer to better fulfill this role.

2 FROM EPRI EL-6800 TO SPM2

The report EL-6800 "Manual on Estimating Soil Properties for Foundation Design" was published by the Electric Power Research Institute (EPRI) in 1990 (Kulhawy & Mayne 1990). It remains one of the most widely used and comprehensive reference for estimating design parameters from laboratory or field test data for soils. Extensive correlations for in-situ stress, strength, elastic behaviour, time-dependent deformability, and permeability are presented, with commentaries on their historical origins and subsequent evolution, supporting data sources, and limitations. Examples are shown in Figure 3. In fact, one important limitation that is obscured when a correlation is plotted as a simple curve without the background data cloud is the lack of guidance on the range of applicability and the degree of transformation uncertainty. The danger of extrapolating beyond the calibration database is real when the range of applicability is not clearly shown. A lack of appreciation of the transformation uncertainty can mislead an inexperienced engineer to assign more precision to the estimate than what is warranted by the data scatter, which can be significant as shown in Figure 3. In addition, if the recommended curve is an average curve, it is not possible to derive a conservative curve for design without a qualitative appreciation of the data scatter. If the recommended curve is a "conservative" curve, the degree of conservatism is usually not provided. These problems are associated with the lack of information on the transformation uncertainty. It is clear that a characterization of the transformation uncertainty, be it visually through presentation of the data cloud or quantitatively through presentation of the statistics, is very useful even in the context of existing deterministic design. To our knowledge, EL-6800 is the first report to provide key statistics such as the sample size (n), coefficient of determination (r^2), and the standard error (S.D.) for all correlations in a systematic way. Some attempts are carried out to ensure that the data are homogeneous, by screening out more unusual geo-materials (e.g. fissured clays). The report also cautions against indiscriminate application of correlations, particularly where some geo-materials may exhibit different behaviours in the presence of cementation (clay, sand), sensitivity (clay), organic/diatom content (clay), aging (sand), plastic fines content (sand), particle crushing (sand), etc.

To the authors' knowledge, EL-6800 has remained a useful reference in design offices to this day. The goal of improving correlations should be considered with EL-6800 as the baseline and the availability of mobile technology in mind. For the time being, our software project SPM2 (Soil Properties Manual version 2) will target the following "wish list" that is developed based on the needs of the modern day geotechnical engineer:

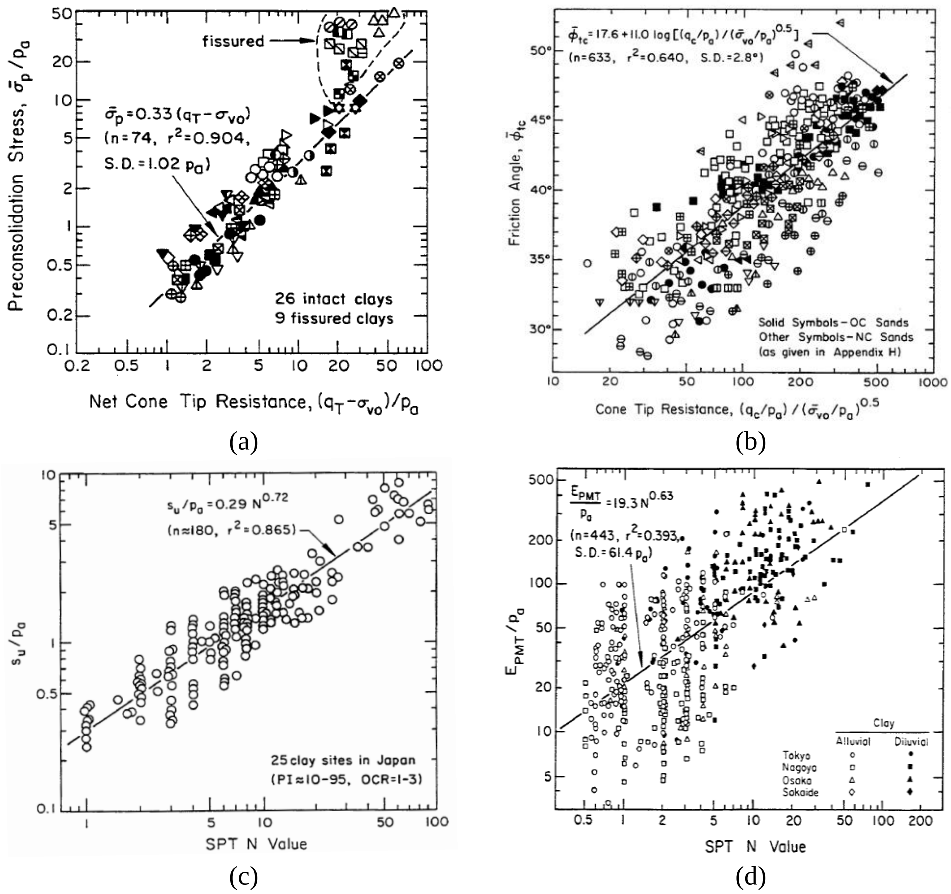


Figure 3. Examples of correlations in EPRI EL-6800 (Kulhawy & Mayne 1990).

1. Compile existing soil/rock data in digital form (some global databases are presented in Table 1). This is part of a long term project to update the data sources in EL-6800, which was published 18 years ago.

2. Develop data sharing and privacy protection protocols (ongoing).

3. Extend existing bivariate correlations to multivariate probability models (a brief discussion on how to do this is given in Section 4).

4. Outlier/error detection, including clustering to recognize potentially distinctive populations (e.g. intact versus fissured clays) (ongoing).

5. Selection of appropriate prior (related to #4) and determination of updated posterior based on site-specific data (ongoing). The prior and posterior distributions are general in the sense that they can be *multivariate* and *non-normal* in their marginal distributions.

6. Export precision of the estimates in the form of a 95% confidence region (non-trivial generalization of the well-known 95% confidence interval to multiple possibly correlated design parameters) for sensitivity analysis (examples given in Fig. 4).

7. Export generic or site-specific multivariate transformation equations for estimation of design soil parameters (examples given in Tables 2, 3, and 4).

8. Export the bias and variability (in the form of a coefficient of variation) of a transformation model as benchmarked against an intended range of soil/rock conditions (examples given in Tables 2, 3, and 4).

9. Export statistics and distributions of the input design parameters for reliability-based design (not discussed).

10. Make SPM2 engineer-friendly on mobile (ongoing).

The remaining sections of this paper present some of the outcomes achieved in this SPM2 project.

3 MULTIVARIATE SOIL/ROCK DATABASES

This section will review some multivariate soil/rock databases. Table 1 shows a summary of these databases, labelled as (geo-material type)/(number of parameters of interest)/(number of data points). This research has inspired comparable databases to be assembled in the literature recently (Müller et al. 2014, Liu et al. 2016). The availability of SPM2 as a freeware will hopefully encourage more data sharing and further enrichment of these databases to cover more parameters and/or more site conditions.

It is important to emphasize that the multivariate distributions constructed from the databases shown in Table 1 are generic in nature, because data are drawn from many sites rather than one single site. Nonetheless, the authors submit that it is reasonable to adopt these multivariate distributions as prior information for a specific site. The posterior probability distribution of a site-specific design parameter can be obtained from this prior information when it is updated by site-specific field data. There are occasional concerns expressed that only site-specific prior information is meaningful in this updating exercise. In other words, data gathered from the literature pertaining to comparable soils and/or sites cannot be used or more specifically, a generic multivariate distribution is not useful as prior information. This concern is understandable, but it is at odds with existing practice. The tradition of geotechnical engineering is steeped in empiricism and one notable aspect is arguably the widespread application of non-site specific generic transformation models to estimate site-specific design parameters. Report EL-6800 is one notable example. Whether one derives a single cautious estimate or a probability distribution from a transformation model, the role of engineering judgment and experience in selecting the appropriate transformation model and weeding out unreasonable estimates is obviously integral to this practice and needs no further emphasis.

3.1 CLAY/5/345 (Ching & Phoon 2012a)

There are 345 data points with complete {Y_1 = LI, $Y_2 = s_u$, $Y_3 = s_u^{re}$, $Y_4 = \sigma'_p$, $Y_5 = \sigma'_v$} information from 37 sites in the CLAY/5/345 database. Each data "point" consists of a set of values stored in one row in the Excel worksheet. This database is a genuine multivariate database, because the {Y_1, Y_2, …, Y_5} information is simultaneously known for each data point. The geographical regions cover Canada, United States, Sweden, Japan, Thailand, United Kingdom, Brazil and India. The clay parameters cover a wide range of sensitivity (1~several hundred; few sites >1000), OCR (1~4; one site up to 12) and LI (0.1~3.8). The clay types are also broad, covering marine clays, sandy/silty clays, Leda clays, etc. Most are quick clays with S_t > 8, and highly OC (fissured) and organic clays are nearly absent in this database.

For all cases, the s_u values in the literature were obtained from various types of tests, including CIUC (isotropically consolidated undrained compression), CIUE (isotropically consolidated undrained extension), CK_0UC (K_0-consolidated undrained compression), CK_0UE (K_0-consolidated undrained extension), DSS (direct simple shear), UU (unconsolidated undrained compression), UC (unconfined compression), and FV (field vane). These values cannot be directly compared because s_u depends on stress state, strain rate, and sampling disturbance. By following the recommendations made by Bjerrum (1972), Mesri & Huvaj (2007), and Kulhawy & Mayne (1990), these s_u values are all converted to the "mobilized" s_u values, denoted by s_u(mob), which is defined as the in-situ undrained shear strength mobilized in embankment and slope failures (Mesri & Huvaj 2007).

3.2 CLAY/6/535 (Ching et al. 2014b)

The CLAY/6/535 database consists of 535 lightly overconsolidated clay data points with complete measurement of {$Y_1 = s_u/\sigma'_v$, Y_2 = OCR, $Y_3 = (q_t-\sigma_v)/\sigma'_v$, $Y_4 = (q_t-u_2)/\sigma'_v$, $Y_5 = (u_2-u_0)/\sigma'_v$, $Y_6 = B_q$} from 40 sites. This database is a genuine multivariate database, because the {Y_1, Y_2, …, Y_6} information is simultaneously known for each data point. The geographical regions cover Brazil, Canada, Hong Kong, Italy, Malaysia, Norway, Singapore, Sweden, UK, USA, and Venezuela. The clay parameters cover a wide range of OCR (mostly 1~6 except for 5 sites) and wide range of plasticity index PI (10~168). Highly OC (fissured) and organic clays are nearly absent in this database. For all cases, the reported s_u values were obtained from various types of tests. All measured s_u is converted into the equivalent CIUC values.

3.3 CLAY/7/6310 (Ching & Phoon 2013c)

The CLAY/7/6310 database consists of a large number of s_u data points obtained from seven different test procedures (CIUC, CIUE, CK_0UC, CK_0UE, DSS, UU, and UC). This database is not a genuine multivariate database, because s_u is typically known for a small subset of the seven procedures. Many s_u data points are associated with a known test mode (6310 points), a known OCR (4584 points), and a known plasticity index (PI) (4541 points). There are some data points in which OCR is not known. CLAY/7/6310 consists of data points from 164 studies. The number of data points associated with each study varies from 1 to 167 with an average 38.5 data points per study. The geographical regions cover Australia, Austria, Brazil, Canada, China, England, Finland, France, Germany, Hong Kong, Iraq, Italy, Japan, Korea, Malaysia, Mexico, New Zealand, Norway, Northern Ireland, Poland, Singapore, South African, Spain, Sweden, Thailand, Taiwan, United Kingdom, United States, and Venezuela. The clay parameters cover a wide range of OCR (mostly 1~10, few studies OCR > 10, but nearly all studies are with OCR < 50) and a wide range of sensitivity S_t (sites with S_t = 1~ tens or hundreds are fairly typical).

3.4 CLAY/10/7490 (Ching & Phoon 2014a)

The CLAY/10/7490 database consists of 7490 data points for ten clay parameters from 251 studies in the literature. This database is not a genuine multivariate database, because the $\{Y_1, Y_2, \ldots, Y_{10}\}$ information is typically partially known for each data point. The number of data points associated with each study varies from 1 to 419 with an average 30 data points per study. 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 parameters 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). Most data points are classified as clays (some are sensitive or organic clays) on the Robertson's CPTU soil classification chart. Some data points are classified as clayey silts or silt mixtures, and few are classified as sand mixtures or sands.

CLAY/10/7490 contains ten dimensionless clay parameters categorized into three groups:

1. Index parameters, including liquid limit (LL), plasticity index (PI), and liquidity index (LI).

2. Stresses and strengths, including normalized vertical effective stress (σ'_v/P_a) (P_a is one atmosphere pressure = 101.3 kN/m^2), normalized preconsolidation stress (σ'_p/P_a), normalized undrained shear strength (s_u/σ'_v), and sensitivity ($S_t = s_u/s_u^{re}$) (s_u^{re} is the remoulded undrained shear strength). The s_u values in the literature

were obtained from various types of tests. These s_u values are all converted to s_u(mob).

3. Parameters from the piezocone test (CPTU), including pore pressure ratio $B_q = (u_2-u_0)/(q_t-\sigma_v)$ (u_2 is the pore pressure behind the cone; u_0 is the hydrostatic pore pressure; q_t is the corrected cone tip resistance; σ_v is the total effective stress), normalized cone tip resistance $(q_t-\sigma_v)/\sigma'_v$, and normalized effective cone tip resistance $(q_t-u_2)/\sigma'_v$.

3.5 F-CLAY/7/216 (D'Ignazio et al. 2016)

The F-CLAY/7/216 database consists of 216 genuine multivariate clay data points from 24 different test sites in Finland. Undrained shear strength from field vane (s_u^{FV}), vertical effective stress (σ'_v), vertical preconsolidation pressure (σ'_p), natural water content (w_n), liquid limit (LL), plastic limit (PL) and sensitivity ($S_t = s_u/s_u^{re}$) are simultaneously known for each data point.

3.6 SAND/7/2794 (Ching et al. 2017d)

The SAND/7/2794 consists of 2794 data points for seven parameters of cohesionless soils from 176 studies in the literature. The label "SAND" broadly denote cohesionless soils, siliceous sands, and gravels. This database is not a genuine multivariate database. The number of data points associated with each study varies from 1 to 295 with an average 9.3 data points per study. Unlike clay databases that are dominated by data from undisturbed in-situ clay samples, the SAND/7/2794 database is dominated by data from laboratory reconstituted soils such as Erksak, Hokksund, Monterey, Ottawa, Sacramento River, Ticino, and Tonegawa sands. Many of these reconstituted soils are clean sands. The remaining (about 15%) data points in the database are in-situ samples obtained from tube sampling, block sampling, or ground freezing techniques. The geographical regions for these in-situ samples cover Canada, Chile, Germany, Greek, India, Italy, Japan, Kuwait, Pakistan, Puerto Rico, Russia, Slovakia, Taiwan, United Kingdom, and United States. The parameters in SAND/7/2794 cover a wide range of median grain size (D_{50}) (0.1mm to more than 100mm), uniformity coefficient (C_u) (1 to more than 1000), relative density (D_r) (-0.1% to 117%), and overconsolidation ratio (OCR) (1 to 15, but mostly 1).

SAND/7/2794 contains seven sand parameters categorized into three groups:

1. Index parameters: the median grain size (D_{50}), coefficient of uniformity (C_u), and relative density (D_r).

2. Effective stress and strength: the normalized vertical effective stress (σ'_v/P_a) (σ'_v is the vertical effective stress, and P_a is one atmosphere pressure = 101.3 kN/m^2) and effective stress friction angle (ϕ'). The friction angle is the secant friction angle obtained in a triaxial compression test.

3. In-situ tests: for cone penetration test (CPT), the normalized cone tip resistance $q_{t1} = (q_t/P_a)\times C_N$ is recorded, where q_t is the cone tip resistance, and C_N is the correction factor for overburden stress. For standard penetration test (SPT), the normalized N value $(N_1)_{60} = N_{60}\times C_N$ is recorded, where N_{60} is the N value corrected for the energy ratio.

3.7 ROCK/9/4069 (Ching et al. 2017e)

The ROCK/9/4069 database consists of 4069 data points from 184 studies for nine parameters of intact rocks. Jointed rock masses are not covered by this database. This database is not a genuine multivariate database. The number of data points associated with each study varies from 1 to 163 with an average 23.6 data points per study. The database is dominated by igneous and sedimentary rocks (27.5% and 59.4%, respectively). The remaining (about 13.1%) data points are metamorphic rock. About 14% of the data points are for weathered rocks, and about 4% are foliated metamorphic rocks. There is no data point for saturated rocks. The geographical regions cover 44 countries/regions, including Afghanistan, Australia, Austria, Brazil, Canada, China, Egypt, France, Germany, Greece, Hong Kong, Hungary, India, Indonesia, Iran, Israel, Italy, Japan, Macao, Malaysia, Mexico, Morocco, Nepal, Netherlands, New Zealand, Nigeria, North Sea, Norway, Pakistan, Portugal, Russia, Saudi Arabia, Singapore, South Africa, South Korea, Spain, Sweden, Taiwan, Thailand, Turkey, United Kingdom, United States, Ukraine, and Uruguay. The properties of the data in ROCK/9/4069 cover a wide range of unit weight (γ) (15 to 35 kN/m^3), porosity (n) (0.01 to 55%), uniaxial compressive strength (σ_c) (0.7 to 380 MPa), Young's modulus (E) (0.03 to 120 GPa), and P-wave velocity (V_p) (0.4 to 8 km/sec).

ROCK/9/4069 contains nine intact rock parameters categorized into four groups:

1. Index properties: porosity (n), unit weight (γ), L-type Schmidt hammer hardness (R_L), and Shore scleroscope hardness (S_h).

2. Strengths: Brazilian tensile strength (σ_{bt}), point load strength index (I_{s50}), and uniaxial compressive strength (σ_c). Note that the point load strength I_s is corrected to a standard diameter of 50 mm, namely I_{s50}, because point load test can be conducted over a wide range of diameter.

3. Stiffness: Young's modulus (E). In ROCK/9/4069, 50.6% of the E data are E_{t50} (tangent modulus at 50% of the peak strength), 10.1% are E_{s50} (secant modulus at 50% of the peak strength), and 39.3% are E_{av} (average modulus for the linear portion of stress-strain curve). The E values for these three definitions do not differ much, compared with the transformation uncertainty in E. As a result, E data with different definitions are combined to form the entire E data.

4. Dynamic property: P-wave velocity (V_p).

4 MULTIVARIATE PROBABILITY FRAMEWORK

We broadly spoke of "correlations" in Section 2 without making this important concept precise. However, it is evident from Figure 3 that "correlations" are very useful as far as estimating design parameters is concerned. This section provides a brief introduction to how the dependency information between two parameters (say s_u and SPT-N) can be captured *numerically* using a single number called a correlation and how this concept can be extended in a simple but powerful way to capture the much more complicated dependency structure underlying multiple soil parameters. It is evident that no estimation is possible if two parameters vary independently. Section 3 clearly shows that we do not live in a two-parameter universe when the behaviour of geo-materials is characterized. Despite its abstract nature, this multivariate probability framework is the simplest possible to enable estimation and updating of multiple design parameters in a consistent way (Phoon 2006). It is sufficient for the engineer to appreciate the physical significance of a correlation explained below – the mathematical intricacies will be handled by SPM2. To illustrate the concept of correlation, let us consider the following simple transformation between two soil parameters (Y_1, Y_2):

$$Y_1 = a + bY_2 + \varepsilon \tag{1}$$

The transformation model is the functional relation $Y_1 = a + bY_2$, while ε is a normal random variable with mean = 0 and standard deviation = s_ε. An example is $Y_1 = \ln(s_u/p_a)$ and $Y_2 = \ln(N)$ shown in Figure 3c. The equation $\ln(s_u/p_a) = \ln(0.29)+0.72\times\ln(N)$ is shown as a straight line. The transformation uncertainty is ε – it can be visualized as the data scatter about the straight line. When $s_\varepsilon = 0$, there is no scatter. The product-moment (Pearson) correlation between Y_1 and Y_2 is defined as:

$$\rho_{12} = \frac{COV(Y_1, Y_2)}{\sqrt{Var(Y_1)}\sqrt{Var(Y_2)}} = \frac{b \times \sqrt{Var(Y_2)}}{\sqrt{b^2 \times Var(Y_2) + s_\varepsilon^2}} \tag{2}$$

where Var(Y) denotes the variance of Y, and $Cov(Y_1,Y_2)$ denotes the covariance between Y_1 and Y_2.

It is clear that if $s_\varepsilon = 0$ (no data scatter), $\rho_{12} = \pm 1$ and perfect correlation exists between Y_1 and Y_2. In this case, given the information of $Y_2 = y_2$, $Y_1 = a + b\times y_2$ is deterministic, and COV = 0 for Y_1 (i.e., Y_1 is no longer uncertain when Y_2 is known). For example, in the absence of data scatter, if $Y_2 = \ln(N) = \ln(10)$, then $\ln(s_u/p_a) \approx 0.42$ and the engineer could confidently estimate $s_u \approx 150$ kPa. In contrast, if s_ε is large (large data scatter), ρ_{12} is close to zero, and weak correlation exists between Y_1 and Y_2. In this case, given the information of $Y_2 = y_2$, $Y_1 = a + b\times y_2 + \varepsilon$ is almost the same as ε and COV is relatively larger for Y_1 (i.e., no point measuring Y_2 if the purpose is to estimate Y_1).

The above simple example shows that correlation the ρ_{ij} between (Y_i, Y_j) quantifies how effective one piece of information (Y_i) can be used to update a second piece of information (Y_j), and such effectiveness can be quantified by the *updated* COV of Y_j – the updated COV is small if ρ_{ij} is close to ±1 and is relatively larger if ρ_{ij} is close to zero. Evans (1996) labelled $|\rho_{ij}| \geq 0.8$ as "very strong", $0.6 \leq |\rho_{ij}| < 0.8$ as "strong", $0.4 \leq |\rho_{ij}| < 0.6$ as "moderate", $0.2 \leq |\rho_{ij}| < 0.4$ as "weak", and $|\rho_{ij}| < 0.2$ as "very weak".

Consider another example with three soil parameters: $Y_1 = \ln(s_u/\sigma'_v)$, Y_2 = LI, Y_3 = ln(OCR) (σ'_v is the vertical effective stress; LI is the liquidity index; OCR is the overconsolidation ratio), and consider the following two transformation models:

$$\begin{aligned} &\ln\left(s_u/\sigma'_v\right) = -0.87 + 0.24 \times \text{LI} + \varepsilon \\ &\ln\left(s_u/\sigma'_v\right) = -1.47 + 0.8 \times \ln\left(\text{OCR}\right) + e \end{aligned} \tag{3}$$

Note that the second equation is related to the SHANSEP concept (Ladd & Foott 1974). The question now is how to update $Y_1 = \ln(s_u/\sigma'_v)$ given the *bivariate* information of [LI, ln(OCR)]? The key observation here is that the knowledge of ρ_{12} and ρ_{13} is not sufficient for the updating – we also need to know ρ_{23}. If $\rho_{23} = 1$ (this can happen if $\varepsilon = e$), one piece of the information in (LI, OCR) is redundant, and we only need the information LI (or OCR) to update $\ln(s_u/\sigma'_v)$. In contrast, if ρ_{23} is relatively small (this may happen if ε and e are statistically independent), both pieces of information (LI, OCR) should be used to update $\ln(s_u/\sigma'_v)$. That is to say, updating Y_1 based on multivariate information ($Y_2 = y_2$, $Y_3 = y_3$, …, $Y_n = y_n$) requires pairwise (or bivariate) correlations (ρ_{ij}: i = 1, …, n-1; j = i+1, …, n). Note that only $n\times(n-1)/2$ correlations are needed, because $\rho_{ij} = \rho_{ji}$ by definition.

Generally speaking, for updating purposes, a multivariate probability distribution function should be characterized from multivariate information, e.g., (s_u, OCR, N) simultaneously measured at approximately the same spatial point in the soil mass. The collection of bivariate correlations (ρ_{ij}: i = 1, …, n-1; j = i+1, …, n) is not sufficient. However, complete multivariate information is rarely available. For example, only CLAY/5/345, CLAY/6/535, and F-CLAY/7/216 are "complete". Among multivariate probability distributions, the multivariate *normal* distribution is available analytically and can be easily constructed based on the collection of bivariate correlations alone. Because bivariate correlations between soil parameters are more commonly available (e.g., Figure 3) the multivariate normal distribution is a sensible and practical choice to capture the multivariate dependency among soil parameters in the presence of transformation uncertainties (Phoon 2006).

Many soil parameters are *not* normally distributed. Let Y denote a non-normally distributed soil parameter. One well known cumulative distribution function (CDF) transform approach can be applied to convert Y into a standard normal variable X: $X = \Phi^{-1}[F(Y)]$, where $\Phi(.)$ is the CDF of the standard normal random variable, F(.) is the CDF of Y, and $\Phi^{-1}(.)$ is the inverse function of $\Phi(.)$. The CDF transform approach does not admit an analytical solution, unless F(Y) is a special function belonging to the Johnson system of distributions (Phoon & Ching 2013b; Ching & Phoon 2015a). There is significant computational advantage in adopting the Johnson system of distributions for Bayesian updating as well.

A set of multivariate soil parameters $\underline{Y} = (Y_1, Y_2, \ldots, Y_n)'$ can be transformed into $\underline{X} = (X_1, X_2, \ldots, X_n)'$ by mapping Y_1 to X_1, Y_2 to X_2, and so forth. By construction, X_1, X_2, …, X_n are *individually* standard normal random variables. It is crucial to note here that *collectively* $(X_1, X_2, \ldots, X_n)'$ does not necessarily follow a multivariate normal distribution even if each component is normally distributed. Even so, applications on actual multivariate databases presented in Section 3 showed that the multivariate normal distribution is an acceptable approximation for clays, sands, and rocks.

The multivariate (standard) normal probability density function for $\underline{X} = (X_1, X_2, \ldots, X_n)'$ can be defined uniquely by a correlation *matrix*:

$$f\left(\underline{X}\right) = \left|\mathbf{C}\right|^{-\frac{1}{2}} \left(2\pi\right)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\underline{X}' \cdot \mathbf{C} \cdot \underline{X}\right) \tag{4}$$

where **C** is the correlation matrix. For n = 3, the correlation matrix is given by:

$$\mathbf{C} = \begin{bmatrix} 1 & \delta_{12} & \delta_{13} \\ \delta_{12} & 1 & \delta_{23} \\ \delta_{13} & \delta_{23} & 1 \end{bmatrix} \tag{5}$$

where δ_{ij} = product-moment (Pearson) correlation between X_i and X_j (not equal to the correlation ρ_{ij} between the original physical variable Y_i and Y_j). It is clear that the full multivariate dependency structure of a normal random vector only depends on a correlation matrix (**C**) containing bivariate correlations between all possible pairs of components, namely X_1 and X_2, X_1 and X_3, and X_2 and X_3. One may be tempted to say that it is not necessary to measure X_1, X_2, and X_3 *simultaneously*. In other words, information on X_1 and X_2, X_1 and X_3, and X_2 and X_3 can be collected at three separate borehole locations, rather than one single borehole location (which is a more restrictive condition). However, although the former collection strategy can produce three correlation coefficients to populate **C** fully, it does not guarantee that **C** is a positive definite matrix. This abstract but important matrix property is explained elsewhere (Ching & Phoon 2015a; Ching et al. 2016d). It suffices to note here that multivariate dependency is more complicated than the superficial simplicity of Eq. (5).

5 BIAS AND COV OF EXISTING TRANSFORMATION MODELS

Useful compilations of existing transformation models for soils and rocks are available in the literature (e.g., Djoenaidi 1985; Kulhawy & Mayne 1990, Mayne et al. 2001; Zhang 2016). Tables 2 to 4 show some examples. All existing transformation models are suitable for the range of conditions found in their own calibration databases. However, some existing transformation models are calibrated by local databases. These local transformation models have their merits. However, for scenarios where local experiences are not available, it may be desirable to adopt a generic transformation model calibrated by a global database such as CLAY/10/7490, SAND/7/2794, and ROCK/9/4069.

The bias and COV for existing clay transformation models can be calibrated using the CLAY/10/7490 database (Table 2, with the exception of M_r models), those for existing sand transformation models can be calibrated using the SAND/7/2794 database (Table 3), whereas those for existing intact rock transformation models can be calibrated using the ROCK/9/4069 database (Table 4). For comparison, M_r models in Table 2 are calibrated using a local database J-Clay/5/124. It is not surprising that the bias is close to 1 as the equations are developed from the same database. The COV is always smaller for a local database. To explain the significance of the bias and COV for a transformation model, consider the q_t-σ'_p model proposed by Kulhawy & Mayne (1990) in Table 2. The actual target value is σ'_p/P_a, and the predicted target value is $0.33\times(q_t-\sigma_v)/P_a$. For each data point in the database with simultaneous knowledge of (q_t-σ_v, σ'_p), the ratio (actual target value)/(predicted target value) = $(\sigma'_p/P_a)/\ [0.33\times(q_t-\sigma_v)/P_a]$ 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 of the transformation model. To be specific,

$$\text{Actual target value} = \text{predicted target value} \times \text{b} \times \varepsilon \tag{6}$$

where b is the bias factor (b = 1 means unbiased as measured against a specific global database such as CLAY/10/7490), and ε is random term with mean = 1 and COV = θ. If θ = 0, there is no data scatter about the transformation model, i.e. the prediction is single-valued or deterministic, rather than a distribution. The bias factors and COVs for some clay/sand/rock transformation models are shown in the last two columns of Tables 2 to 4, respectively. The number of data points used for each calibration is listed in the table ('N' in the fifth column).

The bias and COV of a transformation model can be adopted to derive the unbiased estimate and 95% CI, described as follows. Consider again the q_t-σ'_p model, the calibrated b = 0.97 and calibrated θ = 0.39. Let the site-specific (q_t-σ_v) value for the new design site be denoted by $(q_t-\sigma_v)_{new}$, the unbiased estimate for $(\sigma'_p)_{new}$ is simply b×(predicted target value) = $0.97\times[0.33\times(q_t-\sigma_v)]$. By assuming ε to be lognormal, the 95% CI for $(\sigma'_p)_{new}$ can be expressed as

$$\frac{\text{Unbiased estimate}}{\sqrt{1+\theta^2}} \times \exp\left[\pm 1.96 \times \sqrt{\ln\left(1+\theta^2\right)}\right] \tag{7}$$

Consider a case in the new design site with $q_{t,new}$ = 1000 kN/m^2, $B_{q,new}$ = 0.1, and $\sigma_{v,new}$ = 100 kN/m^2. The unbiased estimate for $\sigma'_{p,new}$ is equal to 0.97×[0.33×(1000-100)] = 288.1 kN/m^2, whereas Eq. (7) suggests that the 95% CI is 128.4 ≤ $\sigma'_{p,new}$ ≤ 561.1 kN/m^2.

It is worth noting that the 95% CI is genuine (i.e., the chance for the actual target value to be within the 95% CI is indeed close to 95%) only if the design site is a general site from the same "population" for the calibration database, e.g., CLAY/10/7490. If the design site is from a different population, the 95% CI may not be genuine. For instance, for a Finland site, the chance for the actual target value to be within the above nominal 95% CI (with b and θ calibrated by F-CLAY/7/216) may be greater than 95%. Also, the numbers of calibration data points (N) for some sand transformation models are quite limited (Table 3). For those transformation models, their nominal 95% CI may not be genuine, either.

Now consider the q_t-s_u model proposed by Ching & Phoon (2012b) in Table 2. The calibrated b = 0.95 and calibrated θ = 0.49. The unbiased estimate for $[(q_t-\sigma_v)/s_u]_{new}$ is 0.95×[29.1×exp(-0.51×$B_{q,new}$)] = 26.27. This suggests the unbiased estimate for $s_{u,new}$ is simply $(q_t-\sigma_v)_{new}$/26.27 = (1000-100)/26.27 = 24.3 kN/m^2. Equation (7) suggests that The 95% CI for $[(q_t-\sigma_v)/s_u]_{new}$ is 9.5 ≤ $[(q_t-\sigma_v)/s_u]_{new}$ ≤ 58.56. This suggests the 95% CI for $s_{u,new}$ is $(q_t-\sigma_v)_{new}$/58.56 ≤ $s_{u,new}$ ≤ $(q_t-\sigma_v)_{new}$/9.5, or simply 15.4 ≤ $s_{u,new}$ ≤ 94.7 kN/m^2.

As shown above, unbiased estimates and 95% CIs for σ'_p and s_u can be obtained from the existing transformation models in Tables 2-4 calibrated by soil/rock databases. Alternatively, generic transformation models can be directly produced by the multivariate probability models calibrated by CLAY/10/7490, SAND/7/2794, and ROCK/9/4069 (see Ching & Phoon 2014b, Ching et al. 2017f, and Ching et al. 2017g for such models). These transformation models have the following advantages: (a) multiple input information is allowed; (b) multiple outputs are allowed; moreover, the joint distribution of all output parameters can be updated. Ching & Phoon (2014b) considered an example involving updating the normalized preconsolidation stress (σ'_p/P_a) and the normalized undrained shear strength (s_u/σ'_v) based on the normalized effective vertical stress (σ'_v/P_a), the CPTU pore pressure ratio (B_q) and the CPTU normalized cone tip resistance $[(q_t-\sigma_v)/\sigma'_v]$. The left plot in Figure 4 shows the multivariate probability density function (PDF) based on the CLAY/10/7490 database. This multivariate PDF is referred to as the "prior" PDF that purely reflects the multivariate database. Given the site-specific site investigation data $(\sigma'_v/P_a)_{new}$, $B_{q,new}$, and $[(q_t-\sigma_v)/\sigma'_v]_{new}$, the multivariate PDF can be updated in the "posterior" PDF, as shown in the right plot in Figure 4. Note that the 95% confidence region, which can be interpreted as values falling within a probability density contour that captures 95% of the probability mass, is not a rectangle constituted by the individual confidence intervals (dashed lines in Fig. 4), because the design parameters, (σ'_p/P_a) and (s_u/σ'_v), are correlated.

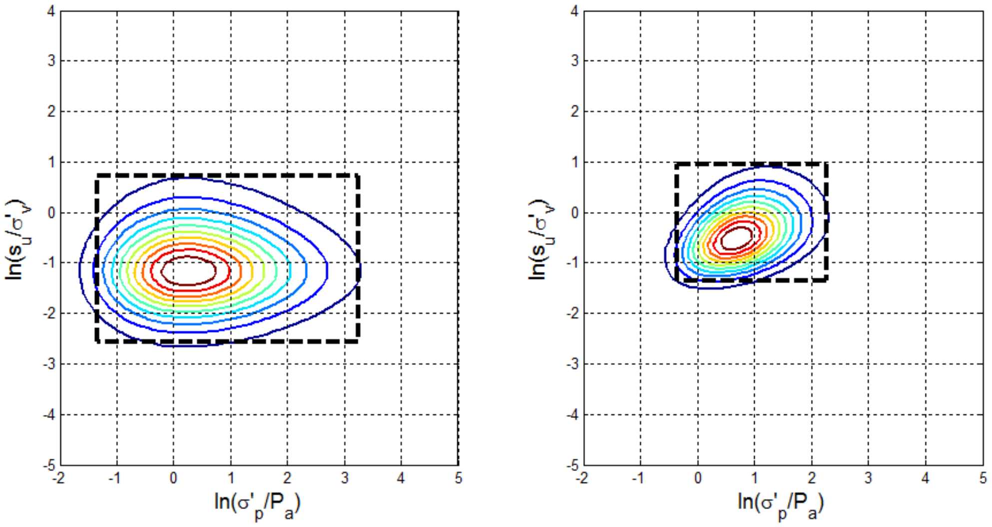


Figure 4. Prior (left) and posterior (right) PDFs for (σ'_p/P_a) and (s_u/σ'_v).

6 CONCLUDING REMARKS

It is accurate to say that the estimation of design parameters has not developed in tandem with significant advances in numerical analysis. The Burland Triangle clearly advocates that geotechnical practice requires knowledge of the ground, soil behaviour, and modelling and mediation of their inter-connections by experience (empiricism and precedent) and risk management. There is an important inter-connection between estimating a design parameter at a point in the ground and estimating the mobilized parameter for a zone of ground governing the behaviour of a geotechnical structure at a limit state. This paper focuses only on the former aspect.

One notable aspect that remains largely empirical is the widespread application of generic correlation models to estimate site-specific design parameters from indirect data, such as the results of laboratory index tests and field tests. These correlations (say between undrained shear strength and SPT-N) are invariably associated with significant transformation uncertainties, which are usually presented in the form of data scatter about the correlation line. While transformation uncertainties can be reduced by accruing better knowledge of soil/rock behaviour or better testing methods, they are unlikely to become negligible in the foreseeable future.

In the opinion of the authors, this critical design parameter estimation step can be improved by characterizing transformation uncertainties quantitatively and doing so in a multivariate probability framework. In practice, site investigation information always appears in a multivariate form. For example, it is not uncommon to find data on unit weight, plasticity index, liquid limit, natural water content, SPT-N, and undrained shear strength in a bore log. Section 3 presents five multivariate property databases for clays, one for sands, and one for rocks to showcase the availability of extensive real soil and rock data to undertake this improvement in practice (not merely in theory). One example is the CLAY/10/7490 database that consists of 7490 data points from 251 studies covering index parameters, stresses and strengths, and piezocone parameters.

The key obstacle to realizing this improvement in practice is the lack of a probabilistic property estimation software. An engineer may be able to construct existing bivariate correlations using regression analysis. Tools for regression analysis are commonly available, such as the Regression function in the Data Analysis Toolbox of EXCEL. The construction of a multivariate probability distribution from a multivariate soil/rock database is less straightforward, both in terms of demanding a more sophisticated understanding of multivariate dependency (this is more complex than a correlation coefficient describing the dependency between two parameters as explained in Section 4) and the lack of genuine multivariate databases. Although Bayesian updating is not discussed in this paper, it is a very useful tool to combine prior information with current site-specific data and it is one major practical application of the multivariate probability distribution. The engineer needs to be fairly computationally sophisticated to compute the desired posterior distributions. The outcomes from Bayesian updating are however far superior to what an engineer can size up from judgment alone. Details are given elsewhere (Ching & Phoon 2015a; Wang et al. 2016a).

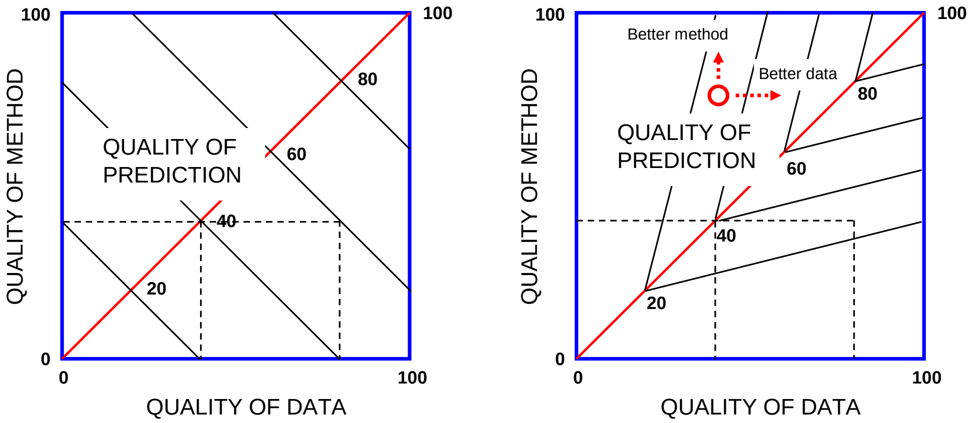
A probabilistic property estimation software called SPM2 (Soil Properties Manual version 2) is currently being developed to remove this obstacle. It is envisaged that an engineer can gain access to SPM2 on his/her mobile to make better decisions on the choice of design soil/rock parameters and to update the full posterior distributions of these parameters in real time based on site-specific data. If the characteristic value is defined as a 5% fractile, this value can be calculated by SPM2 from the posterior distribution. It is not necessary for the engineer to learn multivariate probability theory and Bayesian updating to put SPM2 to good use in practice. After all, it is not necessary for the engineer to learn the Galerkin method and finite element theory to compute internal forces and displacements. In the opinion of the authors, demonstrating the usefulness of SPM2 or similar softwares (see BEST; Wang et al. 2016b) in interpreting site information is an important intermediate step towards popularizing reliability-based or risk-informed design in practice.

The importance of SPM2 to geotechnical practice is illustrated in Figure 5. Lambe (1973) questioned the correctness of Figure 5a that shows the quality of prediction increasing with the quality of the method regardless of the quality of data. He opined that Figure 5b is closer to reality. The quality of prediction is optimal for a particular combination of method and data. Increasing the quality of method or data alone beyond this “best” combination may not result in an improvement of the quality of prediction.

It is safe to say that the quality of our method is currently much higher than the quality of our data (open circle in Figure 5b). We conclude with a question: do we need even better methods? Or better data?

7 ACKNOWLEDGMENTS

We appreciate Dr Chong Tang for his editorial assistance and for computing the bias (mean and COV) of the resilient modulus transformation models in Table 2.



(a) (b)

Figure 5. Quality of prediction as a function of quality of method and quality of data (modified from Lambe 1973)

Table 1. Summary of some soil/rock databases.

Database	Reference	Parameters of interest	# data points	# sites/studies	Range of parameters		
					OCR	PI	S_t
CLAY/5/345	Ching & Phoon (2012a)	LI, s_u, s_u^{re}, σ'_p, σ'_v	345	37 sites	1~4		Sensitive to quick clays
CLAY/6/535	Ching et al. (2014b)	s_u/σ'_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 & Phoon (2013c, 2015b)	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 & Phoon (2014a)	LL, PI, LI, σ'_v/P_a, S_t, B_q, σ'_p/P_a, s_u/σ'_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}, σ'_v, σ'_p, w_n, LL, PL, S_t	216	24 sites	1~7.5	Low to very high plasticity	Insensitive to quick clays
J-Clay/5/124	Liu et al. (2016)	M_r, q_c, f_s, w_n, γ_d	124	16	Soft to stiff clayey soils and silty clay soils with high variability of the strength and stiffness characteristics M_r = 12.54~95.82 MPa q_c = 0.22~3.93 MPa f_s = 0.03~0.14 MPa w_n (%) = 6.91~78.11 γ_d=10.47~19.92 kN/m^3		
SAND/7/2794	Ching et al. (2017d)	D_{50}, C_u, D_r, σ'_v/P_a, ϕ', q_{t1}, $(N_1)_{60}$	2794	176 studies	1~15	D_{50} = 0.1~40 mm C_u = 1~1000+ D_r = -0.1~117%	
ROCK/9/4069	Ching et al. (2017e)	n, γ, R_L, S_h, σ_{bt}, I_{s50}, V_p, σ_c, E	4069	184 studies	γ = 15~35 kN/m^3 n = 0.01~55% σ_c = 0.7~380 MPa E = 0.03~120 GPa		

Note: LL = liquid limit; PL = plastic limit; PI = plasticity index; LI = liquidity index; w_n = natural water content; M_r = resilient modulus; q_c = cone tip resistance; f_s = sleeve friction; γ_d = dry density; 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); n = porosity; γ = unit weight; R = Schmidt hammer hardness (R_L = L-type Schmidt hammer hardness); S_h = Shore scleroscope hardness; σ_{bt} = Brazilian tensile strength; I_s = point load strength index ($I_{s50} = I_s$ for diameter 50 mm); V_p = P-wave velocity; σ_c = uniaxial compressive strength; E = Young's modulus.

Table 2. Bias factor and variability for some existing clay transformation models (calibrated by CLAY/10/7490 except M_r models. M_r models are calibrated using a local database J-Clay/5/124).

Target parameter	Measured parameter(s)	Literature	Transformation model	Calibration results		
				N	Bias (b)	COV (θ)
σ'_p	LI, S_t	Stas & Kulhawy (1984)	$\sigma'_p/P_a \approx 10^{1.11-1.62\times LI}$	249	2.94	1.90
σ'_p	LI, S_t	Ching & Phoon (2012a)	$\sigma'_p/P_a \approx 0.235\times LI^{-1.319}\times S_t^{0.536}$	489	1.32	0.78
σ'_p	w_n, PL, LL	Kootahi & Mayne (2016)	If $5.512\times\log_{10}(\sigma'_v/P_a)-0.061LL-0.093PL+6.219e_n > 1.123$ $\Rightarrow \sigma'_p/P_a \approx 1.62\times(\sigma'_v/P_a)^{0.89}\times LL^{0.12}\times w_n^{-0.14}$ Otherwise $\Rightarrow \sigma'_p/P_a \approx 7.94\times(\sigma'_v/P_a)^{0.71}\times LL^{0.53}\times w_n^{-0.71}$	1242	1.10	0.67
σ'_p	q_t	Kulhawy & Mayne (1990)	$\sigma'_p \approx 0.33\times(q_t-\sigma_v)$	690	0.97	0.39
			$\sigma'_p \approx 0.54\times(u_2-u_0)$	690	1.18	0.75
σ'_p	q_t	Chen & Mayne (1996)	$\sigma'_p/P_a \approx 0.227\times[(q_t-\sigma_v)/P_a]^{1.200}$	690	0.99	0.42
			$\sigma'_p/P_a \approx 0.490\times[(q_t-u_2)/P_a]^{1.053}$	542	1.08	0.61
			$\sigma'_p/P_a \approx 1.274+0.761\times(u_2-u_0)/P_a$	690	0.49	0.59
OCR	q_t	Kulhawy & Mayne (1990)	$OCR \approx 0.32\times[(q_t-\sigma_v)/\sigma'_v]$	690	1.00	0.39

			$OCR \approx 0.259\times[(q_t-\sigma_v)/\sigma'_v]^{1.107}$	690	1.01	0.42
OCR	q_t	Chen & Mayne (1996)	$OCR \approx 0.545\times[(q_t-u_2)/\sigma'_v]^{0.969}$	542	1.06	0.57
			$OCR \approx 1.026\times B_q^{-1.077}$	779	1.28	0.86
s_u	PI	Mesri (1975)	$s_u/\sigma'_p \approx 0.22$	1155	1.04	0.55
s_u	OCR	Jamiolkowski et al. (1985)	$s_u/\sigma'_v \approx 0.23\times OCR^{0.8}$	1402	1.11	0.53
s_u	OCR, S_t	Ching & Phoon (2012a)	$s_u/\sigma'_v \approx 0.229\times OCR^{0.823}\times S_t^{0.121}$	395	0.84	0.34
			$(q_t-\sigma_v)/s_u \approx 29.1\times\exp(-0.513\times B_q)$	423	0.95	0.49
s_u	q_t	Ching & Phoon (2012b)	$(q_t-u_2)/s_u \approx 34.6\times\exp(-2.049\times B_q)$	428	1.11	0.57
			$(u_2-u_0)/s_u \approx 21.5\times B_q$	423	0.94	0.49
M_r	q_c		$M_r=(1.64q_c^{0.53}+2.58)^{2.44}$	124	1.02	0.24
M_r	f_s		$M_r=(26.11f_s^{1.4}+3.83)^{2.44}$	124	1	0.34
M_r	w_n		$M_r=(-1.07w_n^{0.34}+8.12)^{2.44}$	124	1.02	0.27
M_r	γ_d	Liu et al. (2016)	$M_r=(0.0019\gamma_d^{2.33}+3.51)^{2.44}$	124	1.03	0.33
M_r	q_c, f_s		$M_r=(1.46q_c^{0.53}+13.55f_s^{1.4}+2.36)^{2.44}$	124	0.99	0.23
M_r	w_n, γ_d		$M_r=(-0.94w_n^{0.34}+0.0011\gamma_d^{2.33}+7)^{2.44}$	124	1.02	0.25
M_r	q_c, f_s, w, γ_d		$M_r=(1.13q_c^{0.53}+13.06f_s^{1.4}-0.75w_n^{0.34}+0.0007\gamma_d^{2.33}+4.75)^{2.44}$	124	0.97	0.06

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

Table 3. Bias factor and variability for some existing sand transformation models (calibrated by SAND/7/2794).

Target parameter	Measured parameter(s)	Literature	Transformation model	Calibration results		
				N	Bias (b)	COV (θ)
D_r	$(N_1)_{60}$	Terzaghi & Peck (1967)	D_r (%) ≈ $100\times[(N_1)_{60}/60]^{0.5}$	198	1.05	0.231
D_r	N_{60}, OCR, C_u	Marcuson & Bieganousky (1977)	D_r (%) ≈ $100\times\{12.2+0.75\times[222\times N_{60}+2311-711\times OCR-779\times(\sigma'_v/P_a)-50\times C_u^2]^{0.5}\}$	132	1.00	0.211
D_r	$(N_1)_{60}$, OCR, D_{50}	Kulhawy & Mayne (1990)	D_r (%) ≈ $100\times\{(N_1)_{60}/[60+25\log_{10}(D_{50})]/OCR^{0.18}]\}^{0.5}$	199	1.01	0.205
D_r	q_{t1}	Jamiolkowski et al. (1985)	D_r (%) ≈ $68\times[\log_{10}(q_{t1})-1]$	681	0.84	0.327
D_r	q_{t1}, OCR	Kulhawy & Mayne (1990)	D_r (%) ≈ $100\times[q_{t1}/(305\times Q_c\times OCR^{0.18})]^{0.5}$	840	0.93	0.339
ϕ'	D_r, ϕ'_{cv}	Bolton (1986)	$\phi' \approx \phi'_{cv} + 3\times\{D_r\times[10-\ln(p_f')]-1\}$	391	1.03	0.052
ϕ'	D_r, ϕ'_{cv}	Salgado et al. (2000)	$\phi' \approx \phi'_{cv} + 3\times\{D_r\times[8.3-\ln(p_f')]-0.69\}$	127	1.08	0.054
ϕ'	$(N_1)_{60}$	Hatanaka & Uchida (1996)	$\phi' \approx [15.4\times(N_1)_{60}]^{0.5}+20$	28	1.04	0.095
ϕ'	$(N_1)_{60}$	Hatanaka et al. (1998)	If $(N_1)_{60} \le 26 \Rightarrow \phi' \approx [15.4\times(N_1)_{60}]^{0.5}+20$ Otherwise $\Rightarrow \phi' \approx 40$	58	1.07	0.090
ϕ'	$(N_1)_{60}$	Chen (2004)	$\phi' \approx 27.5 + 9.2\times\log_{10}[(N_1)_{60}]$	59	1.00	0.095
ϕ'	q_t	Robertson & Campanella (1983)	$\phi' \approx \tan^{-1}[0.1+0.38\times\log_{10}(q_t/\sigma'_v)]$	99	0.93	0.056
ϕ'	q_{t1}	Kulhawy & Mayne (1990)	$\phi' \approx 17.6 + 11\times\log_{10}(q_{t1})$	376	0.97	0.081

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

Table 4. Bias factor and variability for some existing intact rock transformation models (calibrated by ROCK/9/4069).

Target parameter	Measured parameter	Literature	Transformation model	Calibration results		
				N	Bias (b)	COV (θ)
σ_c	n	Kılıç & Teymen (2008)	$\sigma_c \approx 147.16\times e^{-0.0835n}$	911	0.91	0.747
σ_c	R_L	Karaman & Kesimal (2015)	$\sigma_c \approx 0.1383\times R_L^{1.743}$	664	0.76	0.560
σ_c	S_h	Altindag & Guney (2010)	$\sigma_c \approx 0.1821\times S_h^{1.5833}$	297	1.15	0.650
σ_c	σ_{bt}	Prakoso & Kulhawy (2011)	$\sigma_c \approx 7.8\times\sigma_{bt}$	525	1.31	0.496
σ_c	I_{s50}	Mishra & Basu (2013)	$\sigma_c \approx 14.63\times I_{s50}$	1074	1.18	0.445
σ_c	V_P	Kahraman (2001)	$\sigma_c \approx 9.95\times V_P^{1.21}$	1247	1.26	0.632
E	R_L	Katz et al. (2000)	$E \approx 0.00013\times R_L^{3.09074}$	289	1.29	0.997
E	S_h	Deere & Miller (1966)	$E \approx 0.739\times S_h+11.51$	197	0.61	0.712
E	σ_c	Deere & Miller (1966)	$E \approx 0.303\times\sigma_c-0.8745$	1152	1.23	0.941
E	V_P	Yaşar & Erdoğan (2004)	$E \approx 10.67\times V_P-18.71$	192	0.90	0.724

8 REFERENCES

Altindag, R. & Guney, A. 2010. Predicting the relationships between brittleness and mechanical properties (UCS, TS and SH) of rocks. *Scientific Research and Essays*, 5(16): 2107-2118.

Bjerrum, L. 1972. Embankments on soft ground. *Proceedings of Specialty Conference on Performance of Earth and Earth supported Structures*, 2: 1-54.

Bolton, M.D. 1986. The strength and dilatancy of sands. *Géotechnique*, 36(1): 65-78.

Bond, A., and A. Harris. 2008. *Decoding Eurocode 7*. Taylor and Francis, London.

Burland, J. B. 1987. Nash Lecture: The teaching of soil mechanics – a personal view. *Proceedings, 9th ECSMFE*, Dublin, Vol. 3: 1427-1447.

Chen, J.R. 2004. Axial Behavior of Drilled Shafts in Gravelly Soils. *Ph.D. Dissertation*, Cornell University, Ithaca, NY.

Chen, B.S.Y. & Mayne, P.W. 1996. Statistical relationships between piezocone measurements and stress history of clays. *Canadian Geotechnical Journal*, 33 (3): 488-498.

Ching, J. & Phoon, K.K. 2012a. Modeling parameters of structured clays as a multivariate normal distribution. *Canadian Geotechnical Journal*, 49(5): 522-545.

Ching, J. & Phoon, K.K. 2012b, Establishment of generic transformations for geotechnical design parameters. *Structural Safety*, 35: 52-62.

Ching, J. Y. & Phoon, K. K. 2013a. Mobilized shear strength of spatially variable soils under simple stress states. *Structural Safety*, 41: 20-28.

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.

Ching, J. & Phoon, K.K. 2013c. Multivariate distribution for undrained shear strengths under various test procedures. *Canadian Geotechnical Journal*, 50(9): 907-923.

Ching, J. and Phoon, K.K. 2014a. Transformations and correlations among some clay parameters – the global database. *Canadian Geotechnical Journal,* 51(6): 663-685.

Ching, J. and Phoon, K.K. 2014b. Correlations among some clay parameters – the multivariate distribution, Canadian Geotechnical Journal, 51(6), 686-704.

Ching, J., Phoon, K. K. & Kao, P. H. 2014a. Mean and variance of mobilized shear strength for spatially variable soils under uniform stress states. *Journal of Engineering Mechanics*, ASCE, 140(3): 487-501.

Ching, J., Phoon, K.K., and Chen, C.H. 2014b. Modeling CPTU parameters of clays as a multivariate normal distribution. *Canadian Geotechnical Journal*, 51(1): 77-91.

Ching, J. & Phoon, K. K. 2015a. Constructing multivariate distribution for soil parameters. *Chapter 1, Risk and Reliability in Geotechnical Engineering*. CRC Press, Leiden, 3-76.

Ching, J. & Phoon, K.K. 2015b. Reducing the transformation uncertainty for the mobilized undrained shear strength of clays. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, 141(2): 04014103.

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.

Ching, J., Hu, Y. G. & Phoon, K. K. 2016b. On characterizing spatially variable shear strength using spatial average. *Probabilistic Engineering Mechanics*, 45: 31-43.

Ching, J., Tong, X.W., and Hu, Y.G. 2016c. Effective Young's modulus for a spatially variable elementary soil mass subjected to a simple stress state. *Georisk*, 10(1): 11-26.

Ching, J., Li, D. Q. & Phoon, K. K. 2016d. Statistical characterization of multivariate geotechnical data. *Chapter 4, Reliability of Geotechnical Structures in ISO2394*. CRC Press/Balkema, Leiden, 89-126.

Ching, J., Phoon, K. K. & Pan, Y. K. 2017a. On characterizing spatially variable soil Young's Modulus using spatial average. Structural Safety, 66, May 2017, 106-117.

Ching, J., Sung, S. P. & Phoon, K. K. 2017b. Worst case scale of fluctuation in basal heave analysis involving spatially variable clays. *Structural Safety*, 68, Sep 2017: 28-42.

Ching, J., Phoon, K. K., Chen, K. F., Orr, T. L. L. & Schneider, H. R. 2017c. Statistical determination of multivariate characteristic values using quantile-value method. *Structural Safety*, under review.

Ching, J., Lin, G.H., Chen, J.R. & Phoon, K.K. 2017d. 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.

Ching, J., Li, K.H., Weng, M.C. & Phoon, K.K. 2017e. Generic transformation models for some intact rock properties. *Engineering Geology*, under review.

Ching, J., Lin, G.H., Phoon, K.K. & Chen, J.R. 2017f. Correlations among some parameters of coarse-grained soils – the multivariate probability distribution model. *Canadian Geotechnical Journal*, 54 (9): 1203-1220.

Ching, J., Phoon, K.K., Lin, K.H. & Weng, M.C. 2017g. Multivariate probability distribution for some intact rock properties. *Engineering Geology*, under review.

Deere, D.U. and Miller, R.P. 1966. Engineering Classification and Index Properties for Intact Rock. *Technical Report No. AFWL-TR-65-116*, Air Force Weapons Lab, Kirtland Air Force Base, Albuquerque, NM.

D'Ignazio, M., Phoon, K.K., Tan, S.A. & 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.

EN 1997-1 2004. *Eurocode 7: Geotechnical design-Part 1: General rules*. European Committee for Standardization (CEN), Brussels, Belgium.

Evans, J.D. 1996. *Straightforward Statistics for the Behavioral Sciences*. Brooks/Cole Publishing, Pacific Grove, California.

Fadum, R. E. 1941. Observations and analysis of building settlements in Boston. *Sc.D. Thesis*. Harvard University.

Hatanaka, M. & Uchida, A. 1996. Empirical correlation between penetration resistance and internal friction angle of sandy soils. Soils and Foundations, 36(4): 1-9.

Hatanaka, M, Uchida, A., Kakurai, M. & 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)

Hu, Y.G. and Ching, J. 2015. Impact of spatial variability in soil shear strength on active lateral forces. *Structural Safety*, 52: 121-131.

International Organization for Standardization 2015. General Principles on Reliability of Structures. ISO2394:2015, Geneva.

Jamiolkowski, M., Ladd, C.C., Germain, J.T. & 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.

Kahraman, S. 2001. Evaluation of simple methods for assessing the uniaxial compressive strength of rock. *International Journal of Rock Mechanics and Mining Sciences*, 38(7): 981-994.

Karaman, K. and Kesimal, A. 2015. A comparative study of Schmidt hammer test methods for estimating the uniaxial compressive strength of rocks. *Bulletin of Engineering Geology and the Environment*, 74(2): 507-520.

Katz, O., Reches, Z., and Roegiers, J.C. 2000. Evaluation of mechanical rock properties using a Schmidt hammer. *International Journal of Rock Mechanics and Mining Sciences*, 37(4): 723-728.

Kılıç, A. and Teymen, A. (2008). Determination of mechanical properties of rocks using simple methods. *Bulletin of Engineering Geology and the Environment*, 67(2): 237-244.

Kootahi, K. & Mayne, P.W. 2016. Index test method for estimating the effective preconsolidation stress in clay deposits. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, 04016049.

Kulhawy, F. H. and Mayne, P. W. 1990. Manual on estimating soil properties for foundation design. *Report EL-6800*, Electric Power Research Institute, Palo Alto, California.

Lambe, T.W. 1973. 13th Rankine Lecture: Predictions in soil engineering. *Géotechnique*, 23(2): 149-202.

Ladd, C.C., and Foott, R. 1974. New design procedure for stability in soft clays. *Journal of the Geotechnical Engineering Division*, ASCE, 100(7): 763-786.

Lee Barbour S. & Krahn, J. 2004. Numerical modelling – prediction or process? *Geotechnical News*, 44-52.

Liu, S., Zou, H. Cai, G., Bheemasetti, B. V., Puppala, A. J. & Lin J. 2016. Multivariate correlation among resilient modulus and cone penetration test parameters of cohesive subgrade soils, *Engineering Geology*, 209: 128–142.

Marcuson, W.F. III & Bieganousky, W.A. 1977. SPT and relative density in course sands. *Journal of the Geotechnical Engineering Division*, ASCE, 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. & Huvaj, N. 2007. Shear strength mobilized in undrained failure of soft clay and silt deposits. *Geotechnical Special Publication 173*, ASCE, Reston.

Mishra, D.A. and Basu, A. 2013. Estimation of uniaxial compressive strength of rock materials by index tests using regression analysis and fuzzy inference system. *Engineering Geology*, 160: 54-68.

Müller, R., Larsson, S. & Spross, J. 2014. Extended multivariate approach for uncertainty reduction in the assessment of undrained shear strength in clays. *Canadian Geotechnical Journal*, 2014, 51(3): 231-245.

Phoon K. K. 2006. Modeling and simulation of stochastic data. *GeoCongress 2006*, ASCE, Reston, CDROM.

Phoon, K. K. 2017. Role of reliability calculations in geotechnical design. *Georisk*, 11(1): 4-21.

Phoon, K. K., & Ching, J. 2013a. Is site investigation an investment or expense? - a reliability perspective. *Proceedings*, 18th Southeast Asian Geotechnical Conference (18SEAGC) & Inaugural AGSSEA Conference (1AGSSEA), 29-31 May 2013, Singapore, 25-43.

Phoon, K.K. & Ching, J. 2013b. Multivariate model for soil parameters based on Johnson distributions, Foundation Engineering in the Face of Uncertainty, *Geotechnical Special Publication honoring Professor F. H. Kulhawy*, 337-353.

Phoon, K. K. & Retief, J. V. 2016. *Reliability of Geotechnical Structures in ISO2394*. CRC Press/Balkema, Leiden.

Phoon, K. K., Retief, J. V., Ching, J., Dithinde, M., Schweckendiek, T., Wang, Y. & Zhang, L. M. 2016. Some observations on ISO2394:2015 Annex D (Reliability of Geotechnical Structures). *Structural Safety*, 62: 24-33.

Prakoso, W.A., and Kulhawy, F.H. 2011. Effects of testing conditions on intact rock strength and variability. *Geotechnical and Geological Engineering*, 29(1): 101-111.

Robertson, P.K. & Campanella, R.G. 1983. Interpretation of cone penetration tests: part I - sands. *Canadian Geotechnical Journal*, 20(4): 718-733.

Salgado, R., Bandini, P. & Karim, A. 2000. Shear strength and stiffness of silty sand. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, 126(5): 251-462.

Stas, C. V. & Kulhawy, F. H. 1984. Critical evaluation of design methods for foundations under axial uplift and compressive loading. *Report EL-3771*. Electric Power Research Institute, Palo Alto, California.

Terzaghi, K. & Peck, R. B. 1967. *Soil Mechanics in Engineering Practice*. Second Edition. John Wiley & Sons, New York.

Wang, Y., Cao, Z., and Li, D. 2016a. Bayesian perspective on geotechnical variability and site

characterization. *Engineering Geology*, 203, 117-125.

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

Yaşar, E. and Erdogan, Y. 2004. Correlating sound velocity with the density, compressive strength and Young's modulus of carbonate rocks. *International Journal of Rock Mechanics and Mining Sciences*, 41(5): 871-875.

Zhang, L. 2016. *Engineering Properties of Rocks*, 2nd Edition. Elsevier Ltd., Cambridge, MA, USA.