ISGSR 2011 - Vogt, Schuppener, Straub & Bräu (eds) - © 2011 Bundesanstalt für Wasserbau ISBN 978-3-939230-01-4

How Reliable Are Reliability-Based Multiple Factor Code Formats?

K. K. Phoon
National University of Singapore, Singapore

J. Ching
National Taiwan University, Taipei, Taiwan

J. R. Chen
National Chi Nan University, Puli, Taiwan

ABSTRACT: The objective of this paper is to investigate the degree of deviation from the target reliability index produced when LRFD/MRFD equations are applied to a database of forty-two actual drilled shafts installed in soil profiles underlying the city of Taipei, which contain clay, sand, gravel and rock layers or some partial combination thereof. In general, for soil profiles with multiple layers, conventional formats containing resistance and load factors are unable to achieve the prescribed target reliability index with the same consistency as that reported for homogeneous soil profiles. For the drilled shaft examples considered in this study, the direct application of quantiles in the RBD equation (uniform quantile – η approach), rather than converting the quantiles to conventional resistance and load factors (uniform quantile – standard approach), appears to deliver the most consistent and most robust performance. There is a practical drawback associated with the application of the uniform quantile – η approach. The engineer is required to perform Monte Carlo simulation to estimate the η quantile of lumped random variables such as the total side resistance. This drawback is not present if the uniform quantile – η approach is applied to appropriate parameters where the probability distribution is known analytically or empirically.

Keywords: drilled shafts, axial compression, LRFD, MRFD, reliability code calibration, FORM design point method, uniform quantile method, calibration domain

1 INTRODUCTION

The objective of reliability-based design (RBD) is to adjust a set of design parameters such that a prescribed target probability of failure is achieved or at least not exceeded. For example, the depth of a drilled shaft is a practical design parameter that can be adjusted readily. In principle, it is possible to adjust the shaft diameter but it is less practical to constantly change the diameter of a rotary auger within a single site. These practical considerations apply to the current working stress design (WSD) method. In fact, from a mechanical calculation perspective, there is no difference between RBD and WSD. The former considers a design to be satisfactory if a target probability of failure, say one in a thousand, is achieved. The latter considers a design to be satisfactory if a target global factor of safety, say three, is achieved. The substantive advantage of using the probability of failure (or an equivalent reliability index) in place of the global factor of safety has been discussed elsewhere (Phoon et al. 2003a).

Using the classical example of a drilled shaft under axial compression, the objective of RBD can be stated formally as follows:

$$\mathrm{Prob}(Q < L) \le p_T \tag{1}$$

in which Q = shaft capacity, L = axial load, and p_T = target probability of failure. EN1990:2002 (British Standards Institute, 2002) prescribes $p_T = 7.2 \times 10^{-5}$ (or reliability index, $\beta = 3.8$) for a reliability class 2 (RC2) structure (ultimate limit state). Note that it is straightforward to convert β to p_T and vice-versa using the following convenient EXCEL functions: p_T = NORMSDIST(-β) and β = NORMSINV(1−p_T). It is worthy to observe in passing that Q and L are typically modeled as independent lognormal variables in a number of geotechnical RBD code calibration exercises. This assumption is largely a matter of computational convenience because the left hand side of Eq. (1) can be evaluated in closed-form using the following classical lognormal formula:

$$\beta = \frac{\ln\left[\mu_Q / \mu_L \sqrt{(1+\theta_L^2)/(1+\theta_Q^2)}\right]}{\sqrt{\ln\left[(1+\theta_L^2)(1+\theta_Q^2)\right]}} \tag{2}$$

in which μ_Q, μ_L = mean shaft capacity and mean axial load, respectively and θ_Q, θ_L = coefficient of variation of shaft capacity and coefficient of variation of axial load, respectively. This lumped capacity assumption is convenient from a reliability calculation perspective, but it is rarely emphasized that it could be inconvenient from a physics perspective. The shaft capacity is typically related to side resistance and tip resistance. These resistances are related to geotechnical parameters that can be measured in the laboratory or in the field for obvious practical reasons. The statistics of these geotechnical parameters can be estimated directly from the measured data. Based on this physics perspective, the shaft capacity is a function of more basic random variables (geotechnical parameters). This function is generally nonlinear and the statistics of Q can only be estimated using Monte Carlo simulation. More fundamentally, Q is *not* a lognormal random variable even if all basic random variables are lognormally distributed and Eq. (2) is no longer valid. It could be argued that there are insufficient data to decide which approach is more correct. Nevertheless, it is the position of the authors that one should conform with the best physical model available to date, assign the simplest probability models consistent with measured laboratory/field data and known physical bounds, and live with the resulting complexity in reliability calculations. In short, a physics-centered approach is better than a reliability-centered approach.

For geotechnical problems where simple models are adequate, which is indeed the case for shaft capacity, it is relatively simple to evaluate the left hand side of Eq. (2) using Monte Carlo simulation. More complex problems requiring numerical solution models such as FEM can also be analyzed probabilistically using Monte Carlo simulation, but the computational cost is onerous for common PC platforms. Monte Carlo simulation is a completely general technique. The main disadvantage is tedium, because tens of thousands of design checks (i.e. $Q < L$?) are needed. In contrast, WSD only requires a single check per trial design. There are clever mathematical short-cuts such as the First-Order Reliability Method (FORM) that can reduce tens of thousands of design checks to less than ten design checks at the cost of loss of generality, more complex calculation steps, and occasionally hard-to-detect erroneous solutions. An alternate method called subset simulation (Au & Beck 2001) is gaining popularity, because it is almost as general as Monte Carlo simulation, but requires only about two thousands design checks to achieve a reasonably accurate estimate of the probability of failure. It is accurate to say that very few practitioners are comfortable to perform reliability analysis beyond Monte Carlo simulation which is physically appealing and requires very limited knowledge of probability theory as long as random number generators are available (it is available under “Data Analysis” > “Random Number Generation” in EXCEL). In fact, most practitioners do not find it worthwhile to perform Monte Carlo simulation even when it is available in commercial geotechnical softwares.

Simplified RBD equations in the form of Load and Resistance Factor Design (LRFD), Multiple Resistance Factor Design (MRFD), and partial factor approach (PFA) are popular because practitioners can comply with Eq. (1), albeit approximately, while retaining the simplicity of performing one check per trial design. To the authors’ knowledge, this simplified RBD approach is adopted in all geotechnical RBD codes to date. The practical challenge is to calibrate a set of resistance factors or soil partial factors that would produce designs that comply with Eq. (1) approximately over a range of representative design scenarios. Needless to say, one would prefer the smallest possible set of factors (generating a humungous list like a phonebook would be impractical) covering the widest possible design scenarios that would produce the least deviation from the target reliability index. Phoon et al. (2003a) explicitly recognized this challenge and proposed the following RBD calibration approach to balance pragmatism and compliance with Eq. (1):

1. Perform a parametric study on the variation of the reliability level with respect to each deterministic and statistical parameter in the design problem. Examples of deterministic parameters that control the design of foundations include the diameter (width) and depth to diameter (width) ratio. Examples of statistical parameters for foundations loaded under undrained conditions include the mean and coefficient of variation (COV) of the undrained shear strength.

2. Partition the parameter space into several smaller domains. An example of a simple parameter space is shown in Fig. 1. The reason for partitioning is to achieve greater uniformity in reliability over the full range of deterministic and statistical parameters. For those parameters identified in Step (1) as having a significant influence on the reliability level, the size of the partition clearly should be smaller. In addition, partitioning ideally should conform to existing geotechnical conventions.
3. Select a set of representative points from each domain. Note that each point in the parameter space denotes a specific set of parameter values (Fig. 1). Ideally, the set of representative points should capture the full range of variation in the reliability level over the whole domain.
4. Determine an acceptable foundation design for each point and evaluate the reliability levels in the designs. Foundation design is performed using the set of parameter values associated with each point, along with a simplified RBD format and a set of trial resistance factors. The reliability of the resulting foundation design then is evaluated using Monte Carlo simulation, FORM, or other algorithms.
5. Quantify the deviations of the reliability levels from a prescribed target reliability index, β_T. The following simple objective function can be used:

   $$H(\lambda_1,\lambda_2,\lambda_3\cdots)=\sum_{i=1}^{n}(\beta_i-\beta_T)^2 \tag{3}$$

   in which $H(\cdot)$ = objective function to be minimized, λ_i = partial/resistance factors that are being calibrated, n = number of points in the calibration domain, and β_i = reliability index for the i^{th} point in the domain.
6. Adjust the resistance factors and repeat Steps (4) and (5) until the objective function is minimized. The set of partial/resistance factors that minimizes the objective function (H) is the most desirable because the degree of uniformity in the reliability levels of all the designs in the domain is maximized. The following measure can be used to quantify the degree of uniformity that has been achieved:

   $$\Delta\beta=\sqrt{H/n} \tag{4}$$

   in which $\Delta\beta$ = average deviation from the target reliability index in the calibration domain.
7. Repeat Steps (3) to (6) for the other domains.

Comparable calibration methods have been adopted elsewhere (e.g., CIRIA 1977, Ellingwood et al. 1980, Moses and Larrabee 1988). The effectiveness of applying these simplified RBD equations to more realistic ground conditions containing multiple strata has not been studied, despite its obvious practical importance. The objective of this paper is to investigate the degree of deviation from the target reliability index produced when LRFD/MRFD equations are applied to a database of forty-two actual drilled shafts installed in the city of Taipei. The effect of the RBD calibration method (design point method, quantile-based method) and number of calibration points are also studied.

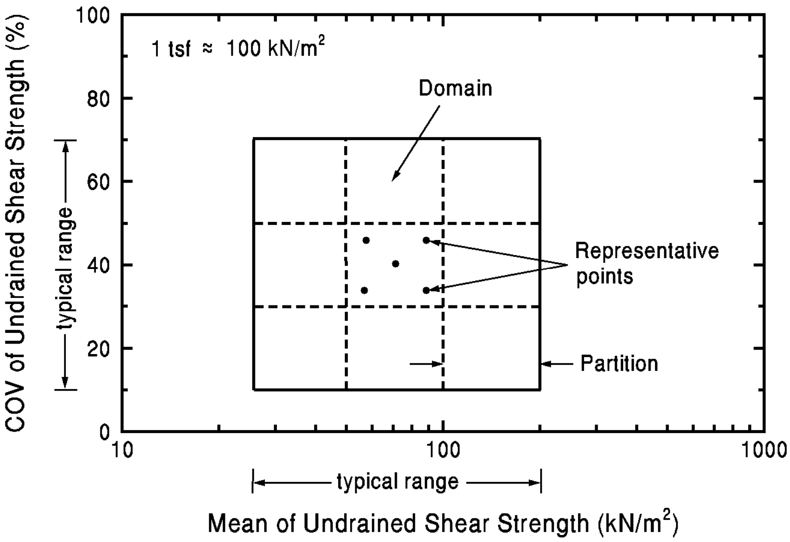

Figure 1. Partitioning of parameter space for calibration of resistance factors.

2 DRILLED SHAFT DATABASE

2.1 *Overview of database*

Table 1 summarizes the basic shaft and geotechnical information for forty-two drilled shafts installed in the city of Taipei. The shaft diameter, B, varies between 0.8 m and 2.5 m with an average of 1.24 m. The shaft length, D, varies between 20.7 m and 76 m with an average of 48.8 m. The D/B ratio varies between 18.8 and 63.3 with an average of 40.3. The compression capacity interpreted using the slope-tangent method varies between 6172 kN an 15372 kN with an average of 10772 kN. More details are reported elsewhere (Ching et al. 2011).

Table 1. Basic shaft and geotechnical information for drilled shafts installed in the city of Taipei.

Shaft No.	Site Location	Diameter B (m)	Depth D (m)	Water Table (m)	D/B	Soil Description	Q_{ST}* (kN)	Q_{L2}* (kN)	Group
CT-02	Xinyi District	1.2	29.2	-	24.3	Silty clay over sandstone	6172	11237	CR
CT-04	Da'an District	2.0	37.5	-	18.8	Interbeded silty clay and silty sand over gravel	-	-	CSG
CT-05	Xinyi District	1.0	53.5	-	53.5	Silty clay over gravel & sandstone	10069	13047	CGR
CT-07	Xinyi District	1.2	59.0	-	49.2	Interbeded silty clay and silty sand over gravel & mudstone	17638	17658	ALL
CT-09	Xinyi District	1.2	47.7	-	39.8	Interbeded silty clay and silty sand over sandstone	15990	17168	CSR
CT-10	Zhongshan District	1.0	55.5	-	55.5	Interbeded silty clay and silty sand over gravel	17040	21680	CS
CT-11	Beitou District	1.0	44.5	-	44.5	Silty clay over sandstone	6293	7456	CSR
CT-13	Taipei County	1.2	43.0	-	35.8	Interbeded silty sand and silty clay over gravel	13337	17291	CSG
CT-14-1	Xinyi District	1.2	76.0	-	63.3	Interbeded silty clay and silty sand over gravel & mudstone	32226	35542	ALL
CT-14-2	Xinyi District	1.5	66.0	-	44.0		36572	42948	ALL
CT-14-3	Xinyi District	1.5	65.0	-	43.3		37769	46499	ALL
CT-14-4	Xinyi District	1.5	56.0	-	37.3		25790	34266	ALL
CT-14-5	Xinyi District	1.2	59.0	-	49.2		26634	34835	ALL
CT-14-6	Xinyi District	2.5	70.3	-	28.1		66414	75213	ALL
CT-15	Shilin District	1.0	29.0	-	29.0	Interbeded silty sand and silty clay over sandstone	6959	7652	CSR
CT-16	Xinyi District	1.5	66.0	-	44.0	Interbeded silty clay and silty sand over gravel & mudstone	15892	36336	ALL
CT-17	Da'an District	1.5	48.0	3.6	32.0	Interbeded silty sand and silty clay over gravel	19483	25280	CSG
CT-18	Zhongzheng District	1.2	29.4	-	24.5	Interbeded silty sand and silty clay over sandstone	13214	16304	CSR
CT-19	Xinyi District	1.2	59.0	-	49.2		28917	36336	CSR
CT-20	Zhongzheng District	1.2	64.4	3.6	53.6		15127	17550	CSR
CT-21	Xinyi District	1.2	52.0	-	43.3	Interbeded silty clay and silty sand over gravel & mudstone	-	-	ALL
CT-22	Xinyi District	1.2	54.0	-	45.0	Interbeded silty clay and silty sand over gravel or mudstone	-	-	CSG
CT-23	Xinyi District	1.0	53.0	4.0	53.0		-	-	CSR
CT-24	Xinyi District	1.2	54.0	3.0	45.0		-	-	CSG
CT-25-1	Da'an District	1.5	45.2	5.5	30.1	Interbeded silty clay and silty sand over sandstone	19465	24280	CSR
CT-27	Xinyi District	1.2	76.0	-	63.3	Interbeded silty clay and silty sand over gravel & mudstone	27263	37327	ALL
CT-30	Zhongshan District	1.0	47.5	-	47.5	Interbeded silty clay and silty sand over sandstone	9609	11644	CSR
CT-31	Beitou District	1.0	46.6	-	46.6		6253	7720	ALL
CT-32	Beitou District	1.0	31.4	-	31.4		5960	5690	CSR
CT-33	Shilin District	1.0	31.4	-	31.4		7173	7917	CSR
CT-34-1	Zhongzheng District	0.9	42.0	-	46.7	Interbeded silty clay and silty sand over gravel	6355	7120	CS

Shaft No.	Site Location	Diameter B (m)	Depth D (m)	Water Table (m)	D/B	Soil Description	Q_{ST}* (kN)	Q_{L2}* (kN)	Group
CT-35-1	Zhongzheng District	1.2	46.4	-	38.6		11903	13381	CSG
CT-36-1	Zhongzheng District	0.9	47.6	-	52.9		6642	7161	CSG
CT-37-1	Taipei County	1.2	34.2	-	28.5		7808	8584	CSG
CT-38-1	Taipei County	1.8	48.5	3.8	26.9		32227	40643	CSG
CT-39-1	Songshan District	1.8	49.5	0.2	27.5	Interbeded silty sand and silty clay over gravel & clay layer	12724	23014	CSG
CT-40-1	Taipei County	0.8	29.1	6.0	36.4	Interbeded silty sand and silty clay over gravel	7624	9584	CS
CT-41	Zhongshan District	1.3	54.9	3.0	42.2		15733	18763	CSG
CT-42	Zhongzheng District	1.2	47.4	-	39.5		14056	17854	CSG
CT-43	Zhongshan District	1.0	46.2	0.2	46.2		11836	15629	CSG
CT-44	Beitou District	1.0	32.0	0.2	32.0	Silty clay over sandstone	6523	7014	CR
CT-45	Shilin District	1.0	20.7	0.2	20.7		15372	20012	CR

*Measured compression capacity from load test: Q_{ST} = capacity interpreted using the slope-tangent method, Q_{L2} = capacity interpreted using the L1-L2 method.

It is useful to observe that common drilled shaft diameters are covered in the database. However, based on D/B ratio, the database covers predominantly long friction shafts. More importantly, all the shafts are installed in non-homogeneous layered soils. Based on the strata that provide the side resistances, the shafts are classified into five groups: (a) Group ALL: the strata include clay, sand, gravel, and rock layers (11 shafts); (b) Group CSR: the strata include clay, sand, and rock layers (11 shafts); (c) Group CSG: the strata include clay, sand, and gravel layers (13 shafts); (d) Group CGR: the strata include clay, gravel, and rock layers (1 shaft); (e) Group CR: the strata include clay and rock layers (3 shafts); and (f) Group CS: the strata include clay and sand layers (3 shafts). It is apparent that this database covers a fairly comprehensive range of layered soil profiles.

2.2 *Axial compression capacity and its associated uncertainties*

The axial compression capacity of a drilled shaft is the sum of side resistances along the shaft and end bearing at the tip minus its own self-weight. For the long friction shafts shown in Table 1, it is adequate to assume that the shaft capacity (Q) is approximately equal to the total side resistance (S). For shafts installed in multiple strata with possible appearance of clay, sand, gravel, and rock layers, the total side resistance is expressed as:

$$S = S_c + S_s + S_g + S_r \tag{5}$$

in which S_c , S_s, S_g, and S_r = side resistances for the clay, sand, gravel, and rock layers, respectively. The side resistance in a given layer, denoted by S_x (the subscript 'x' denotes either 'c', 's', 'g', or 'r', depending on the stratum type of interest), can be computed as:

$$S_x = \pi B \sum_{i=1}^{N} f_{si} t_i \tag{6}$$

in which B = shaft diameter. For calculation purposes, each stratum is discretized into N layers, with f_{si} being the unit side resistance for the i^{th} layer and t_i being the thickness of the i^{th} layer. Note that Eq. (6) assumes that there is only one layer per geomaterial type (clay, sand, gravel, or rock). It is rather common to have interbeds consisting of different geomaterials, particularly clay, sand and gravel, in actual profiles. Eq. (6) can be easily generalized to more complex profiles. The models for unit side resistances in clay, sand, gravel, and rock are summarized in Table 2. They are developed from the α-method for clay and rock and β-method for sand and gravel. The model uncertainties, ε_{Sc}, ε_{Ss}, ε_{Sg}, ε_{Sr}, are described by zero-mean normal random variables with standard deviations given in Table 2. Details on calibration of these unit side resistance models and estimation of associated model statistics are given by Ching et al. (2011).

In addition, the measured soil parameters are modeled as the actual parameters contaminated with measurement errors:

$$\begin{aligned}
\ln(\sigma'_{vs,m}) &= \ln(\sigma'_{vs}) + e_{\sigma'_{vs}} \\
\ln(\sigma'_{vg,m}) &= \ln(\sigma'_{vg}) + e_{\sigma'_{vg}} \\
\ln(s_{u,m}) &= \ln(s_u) + e_{s_u} \\
\ln(q_{u,m}) &= \ln(q_u) + e_{q_u}
\end{aligned} \tag{7}$$

in which $\sigma'_{vs,m}$, $\sigma'_{vg,m}$, $s_{u,m}$, $q_{u,m}$ = measured values of average vertical effective stress of sand layer, average vertical effective stress of gravel layer, undrained shear strength of clay layer, unconfined compression strength for rock layer, respectively and e_{su}, $e_{\sigma'vs}$, $e_{\sigma'vg}$, e_{qu}. = measurement errors associated with the subscripted soil parameters. These measurement errors are modeled as zero-mean normal random variables with the following standard deviations: 0.20 for e_{su}, 0.10 for $e_{\sigma'vs}$ and $e_{\sigma'vg}$, and 0.47 for e_{qu}.

Table 2. Models for unit side resistances in clay, sand, gravel, and rock.

Geomaterial	Correlation model for unit side resistance (kN/m^2)	Standard deviation of model uncertainty ε
Clay	$f_c = \exp[2.70 + 0.30\ln(s_u) + \varepsilon_{s_c}]$	0.32
Sand	$f_s = \exp[1.08 - 0.66\ln(z) + \ln(\sigma'_v) + \varepsilon_{s_s}]$	0.54
Gravel	$f_g = \exp[2.18 - 0.75\ln(z) + \ln(\sigma'_v) + \varepsilon_{s_g}]$	0.67
Rock	$f_r = \exp[3.03 + 0.41\ln(q_u) + \varepsilon_{s_r}]$	0.72

Note: s_u = undrained shear strength, σ'_v = vertical effective stress, and q_u = unconfined compression strength.

3 RELIABILITY CALIBRATION

3.1 *Performance function*

The performance function, G, is an arbitrary function that is less than zero when its arguments result in a failure state. For drilled shafts considered in this study, it is natural to define the performance function as:

$$G = S_c + S_s + S_g + S_r - L_D - L_L \tag{8}$$

in which L_D = dead load and L_L = live load. It is clear that the ultimate limit state is exceeded when G < 0. The basic random variables describing the uncertainties in the side resistances are the unit side resistance model errors shown in Table 2 (ε_{Sc}, ε_{Ss}, ε_{Sg}, ε_{Sr}) and the soil parameter measurement errors (e_{su}, $e_{\sigma'vs}$, $e_{\sigma'vg}$, e_{qu}) mentioned in Section 2.2. If N = 1 in Eq. (6), the side resistance contributed by the clay layer is:

$$S_c = \pi B f_c t_c = \pi B \exp\{2.70 + 0.30[\ln(s_{u,m}) - e_{s_u}] + \varepsilon_{s_c}\} t_c \tag{9}$$

It is clear that S_c is a lognormal random variable in this special, because $\ln(S_c)$ = constant – $0.3e_{su} + \varepsilon_{Sc}$ is a normal variable by hypothesis. For the more general case in which N > 1, S_c is a sum of lognormal random variables and hence, it is *not* a lognormal random variable. The total side resistance $S = S_c + S_s + S_g + S_r$ is *not* a lognormal variable even if S_c, S_s, S_g, S_r are individually lognormal variables for the same reason. The dead load is modeled as a lognormal random variable with mean = μ_{LD} and coefficient of variation = θ_{LD} = 0.10. The live load is also modeled as a lognormal random variable with mean = μ_{LL} and coefficient of variation = θ_{LL} = 0.25. The ratio μ_{LL}/μ_{LD} = 0.5 unless stated otherwise.

Given that the capacity (S) and the load ($L_D + L_L$) are not lognormal variables, Eq. (2) cannot be applied. The probability of failure, Prob(G < 0), is computed using Monte Carlo simulation in this study.

3.2 *Simplified RBD equations*

Four simplified RBD equations are considered: LRFD, MRFD2, MRFD3, and MRFD4. They are defined in Table 3 below.

Table 3. Simplified RBD equations.

RBD equation	Definition	Comment
LRFD	$\gamma_{total}\left(S_c^* + S_s^* + S_g^* + S_r^*\right) \geq \gamma_D L_D^* + \gamma_L L_L^*$	Calibrate 1 resistance factor: γ_{total}
MRFD2	$\gamma_{csg}\left(S_c^* + S_s^* + S_g^*\right) + \gamma_r S_r^* \geq \gamma_D L_D^* + \gamma_L L_L^*$	Calibrate 2 resistance factors: γ_{csg}, γ_r
MRFD3	$\gamma_{cs}\left(S_c^* + S_s^*\right) + \gamma_g S_g^* + \gamma_r S_r^* \geq \gamma_D L_D^* + \gamma_L L_L^*$	Calibrate 3 resistance factors: γ_{cs}, γ_g, γ_r
MRFD4	$\gamma_c S_c^* + \gamma_s S_s^* + \gamma_g S_g^* + \gamma_r S_r^* \geq \gamma_D L_D^* + \gamma_L L_L^*$	Calibrate 4 resistance factors: γ_{cs}, γ_s, γ_g, γ_r

Note: Asterisk denotes nominal resistances or nominal loads. Nominal resistances, S_c^*, S_s^*, S_g^*, and S_r^*, are computed by assuming the model errors and soil parameter measurement errors are zero. Nominal loads, L_D^* and L_L^*, are computed at their mean values.

3.3 *RBD calibration method 1: FORM design point*

Two RBD calibration methods are considered in this study: (1) FORM design point method and (2) uniform quantile method. The first method is described in this section.

The First-Order Reliability Method (FORM) involves seeking for a design point lying on the performance function G that is closest to the origin in standard normal space. An illustrative example containing two standard normal random variables, U_1 and U_2, is shown in Fig. 2. It is possible to transform a set of non-normal physical random variables to a set of standard normal random variables. The probability of failure estimated using FORM is Prob(G_L < 0) = $\Phi(-\beta)$, in which $\Phi(\cdot)$ = standard normal cumulative distribution function evaluated using say NORMSDIST in EXCEL. By definition, the reliability index of a design satisfying the equation below is approximately equal to β (error due to linearization in FORM, $G_L \approx G$):

$$G\left(u_1^d, u_2^d\right) = G\left(x_1^d, x_2^d\right) = 0 \tag{10}$$

in which X_1, X_2 = physical random variables. Examples of physical random variables are given below.

This FORM design point method is described in Ang & Tang (1984). It is rarely emphasized that Eq. (10) forces the performance function to be coupled to the simplified RBD equation. For example, the LRFD format shown in Table 3 can only be calibrated using this design point method by stating the performance function in the following form:

$$G = S - L_D - L_L \tag{11}$$

in which S is the total side resistance. There are three physical random variables: X_1 = S, $X_2 = L_D$ and $X_3 = L_L$. Applying Eq. (10), it can be seen that:

$$x_1^d - x_2^d - x_3^d = 0 \tag{12}$$

Eq. (12) can be re-written in the LRFD format as follows:

$$\gamma_{total} x_1^* - \gamma_D x_2^* - \gamma_L x_3^* = 0 \tag{13}$$

in which x_1^*, x_2^*, x_3^* = nominal values of S, L_D, and L_L, respectively. The resistance and load factors are calibrated using: $\gamma_{total} = x_1^d / x_1^*$, $\gamma_D = x_2^d / x_2^*$ and $\gamma_L = x_3^d / x_3^*$.

For MRFD2, the performance function is stated in the following form:

$$G = S_{csg} + S_r - L_D - L_L \tag{14}$$

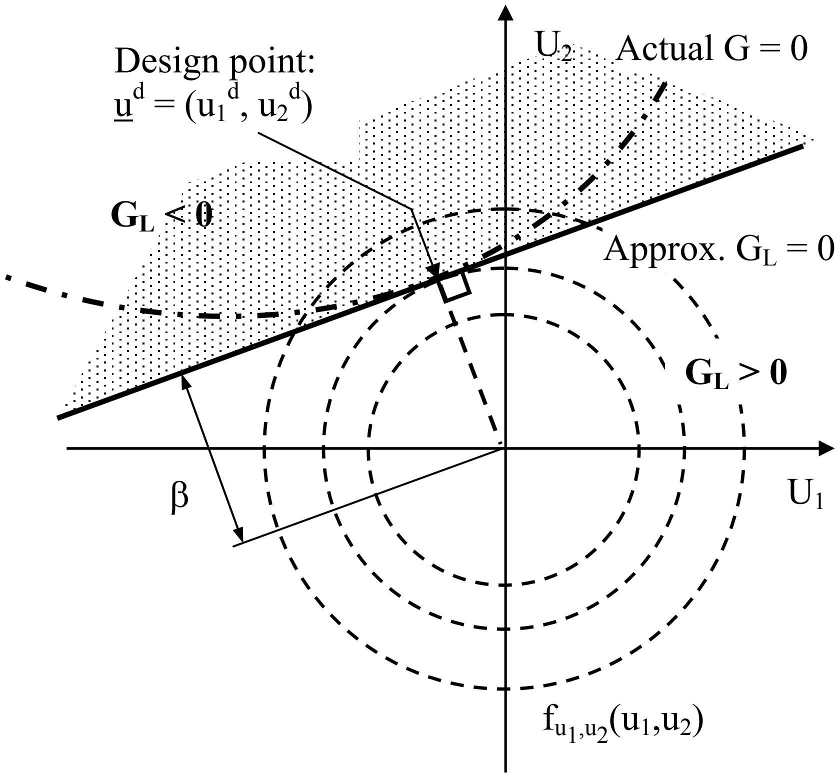

Figure 2. Definition of design point in First-Order Reliability Method (FORM).

in which $S_{csg} = S_c + S_s + S_g$. There are four physical random variables: $X_1 = S_{csg}$, $X_2 = S_r$, $X_3 = L_D$ and $X_4 = L_L$. Applying Eq. (10), it can be seen that:

$$x_1^d - x_2^d - x_3^d - x_4^d = 0 \tag{15}$$

Eq. (15) can be re-written in the MRFD2 format as follows:

$$\gamma_{csg} x_1^* + \gamma_r x_2^* - \gamma_D x_3^* - \gamma_L x_4^* = 0 \tag{16}$$

in which $\gamma_{csg} = x_1^d / x_1^*$, $\gamma_r = x_2^d / x_2^*$, $\gamma_D = x_3^d / x_3^*$ and $\gamma_L = x_4^d / x_4^*$.

Based on the above LRFD and MRFD2 examples, the nature of the coupling is clear. The physical random variables must be defined such that the resistance/load factor can appear as $\gamma_i = x_i^d / x_i^*$. For LRFD/MRFD, it is apparent that this is only possible when the resistances and loads are separable. A simple example where the resistance and the load cannot be separated is the bearing capacity of a shallow foundation subjected to inclined loading. The inclined load factor in the bearing capacity equation is a function of both the vertical and horizontal loads for drained loading (e.g., Annex D, BS EN1997-1:2004). It could be possible to circumvent this problem by assuming that the bearing capacity is correlated to the vertical load. Nonetheless, a second ad-hoc assumption that the non-normal bearing capacity and non-normal vertical load are correlated using a translation procedure is practically necessary at present (Phoon 2006). The adequacy of this ad-hoc assumption as applied to code calibration has not been examinued thus far.

For the drilled shaft example considered in this study, it is possible calibrate LRFD and MRFD formats using the FORM design point method. However, there is an important practical difficulty that is not highlighted in previous studies. Although Eq. (8), Eq. (11) and Eq. (14) are mathematically equivalent, the probability distributions of S and S_{csg} are *not* lognormals as explained in Section 3.1 and cannot be derived analytically even for the relatively simple drilled shaft example in this study where the model and parametric errors are assumed to be lognormally distributed. However, the probability distributions of S and S_{csg} can be derived *empirically* using Monte Carlo simulation. In general, these empirical probability distributions will not fit classical closed-form probability distributions commonly found in standard texts. The authors found that the FORM algorithm is not stable when the physical random variables such as S and S_{csg} are defined using empirical distributions (known only at discrete sample points and likely to be inaccurate at probability tails where the design point is located). This FORM computational difficulty is currently being studied.

It is possible to take the pragmatic approach of assuming the resistances S and S_{csg} as lognormals, rather than assuming that the model and parametric errors are lognormals as in Section 3.1. Nonetheless, it has been pointed out in Section 1 that it is not judicious to make probabilistic assumptions for the convenience of reliability calculations and in the same vein, certainly not for the convenience of code calibrations. Probability distributions should be fitted to measured laboratory/field data, comply with known physical upper and/or lower bounds, and respect the accumulated knowledge base on correlations between various soil properties. It may not be possible to identify an appropriate probability distribution for a basic soil parameter exactly, because of insufficient data and/or imperfect knowledge. However, a probability distribution that fits known data and state-of-the-art knowledge is "best" at a particular point in time. It can be revised when more data and/or state of knowledge improves – this is true for all aspects of scientific pursuit; not merely probabilistic analysis. In some past critiques of RBD, the inability to identify "correct" probability distributions has been used as a reason for doubting the practical relevance of RBD. While it may be fair to critique undue probabilistic simplifications made to suit computational convenience, it appears unreasonable to demand perfect knowledge of probability distributions (which is merely a mathematical model of reality) and it is against the grain of evolving scientific progress.

In this study, the basic physical random variables characterized in Section 3.1 are assumed to be "correct", because they are based on actual measured data. To apply the FORM design point method to LRFD/MRFD, it is necessary to use *lumped* variables such as S and S_{csg}. To circumvent potential numerical instability of FORM associated with the use of empirical distributions for S and S_{csg}, it is assumed that these lumped variables are lognormally distributed. However, the second-moment statistics (mean and standard deviation) of these lumped variables are correctly estimated using Monte Carlo simulation. Note that this ad-hoc lognormal assumption is only applied to calibrate the various resistance and load factors in the LRFD/MRFD formats. These calibrated formats are validated by evaluating the reliability indices of drilled shafts outside the calibration domain in Section 4. The reliability indices computed during validation in Section 4 are based on the "correct" probability models given in Section 3.1.

Finally, it is of interest to observe that the calibration of resistance and load factors using Eq. (12) and Eq. (13) for LRFD is carried out to conform to the historical practice of producing a simplified RBD design equation with the same "look and feel" as existing working stress design equation. This calibration method is termed "FORM-standard" in this study. It is possible to consider an alternate FORM-based calibration approach using the design point in standard normal space (u_1^d, u_2^d, u_3^d):

$$u_1^d - u_2^d - u_3^d = 0 \tag{17}$$

In this approach, the design point (u_1^d, u_2^d, u_3^d) determined from a single calibration shaft is assumed to apply to all other shafts. The LRFD equation is now written as:

$$x_1\left(u_1^d\right) - x_2\left(u_2^d\right) - x_3\left(u_3^d\right) = 0 \tag{18}$$

in which $x_i(u_i^d)$ = value of i^{th} physical variable for a validation shaft calculated using the value of the i^{th} standard normal variable at the *design point of the calibration shaft*. This calibration method is termed "FORM-u" in this study.

3.4 *RBD calibration method 2: uniform quantile*

The uniform quantile calibration method was proposed by Ching & Phoon (2011). The procedure is illustrated below using the LRFD format. Theoretical details are given in the above cited paper.

The η quantile of the total side resistance, S^η, is defined by:

$$\mathrm{Prob}\left(S < S^\eta\right) = \eta \tag{19}$$

in which η = number between 0 and 1, but typically order of 0.01 for practical problems. S^{η} is also called the 100η% exclusion limit. For example, for a normal random variable, the 5% quantile, $S^{0.05} = \mu_S (1 - 1.645\theta_S)$, in which μ_S and θ_S are the mean and coefficient of variation of S, respectively. This definition is sensible for resistances, because S^{η} is a conservative value less than the mean value. For loads, it is natural to consider the (1-η) quantile. For example, the (1-η) quantile of the live load, $L_L^{1-\eta}$, is defined by:

$$\begin{aligned} \mathrm{Prob}\left(L_L < L_L^{1-\eta}\right) &= 1-\eta \\ \mathrm{Prob}\left(L_L > L_L^{1-\eta}\right) &= \eta \end{aligned} \tag{20}$$

Hence, the probability of L_L *exceeding* $L_L^{1-\eta}$ is η. $L_L^{1-\eta}$ is also called the $1/\eta$ return period load if L_L is defined as the annual maximum load and L_L varies independently from year to year. For a normal random variable, the (100-5)% = 95% quantile, $L_L^{0.95} = \mu_{LL} (1 + 1.645\theta_{LL})$, in which μ_{LL} and θ_{LL} are the mean and coefficient of variation of L_L, respectively. This definition is sensible for loads, because $L_L^{1-\eta}$ or $L_D^{1-\eta}$ is a conservative value greater than the mean value.

For LRFD, the uniform quantile calibration method produces the following simplified RBD equation:

$$S^{\eta} \geq L_D^{1-\eta} + L_L^{1-\eta} \tag{21}$$

Comparing with the LRFD format shown in Table 3, it is clear that $\gamma_{total} = S^{\eta}/S^*$, $\gamma_D = L_D^{1-\eta}/L_D^*$, and $\gamma_L = L_L^{1-\eta}/L_L^*$. The resistance and load factors for the MRFD formats in Table 3 can be calibrated in a similar way. For example, the resistance and load factors for the MRFD2 format can be derived from the following quantile equation:

$$S_{csg}^{\eta} + S_r^{\eta} \geq L_D^{1-\eta} + L_L^{1-\eta} \tag{22}$$

in which S_{csg}^{η}, S_r^{η} = η quantile of S_{csg} and S_r, respectively.

The distinctive feature of the uniform quantile approach is that the *same* quantile, η, is applied to all resistance and load components, regardless of the number of components. In fact, it is also possible to apply the uniform quantile approach to the partial factor format, although this format is not included in the present study. The quantile is simply applied on more basic soil parameters such as the undrained shear strength in this instance. Examples of partial factors calibrated using the uniform quantile approach are given in Ching & Phoon (2011). Ching & Phoon (2011) also demonstrated theoretically that a *single* quantile applied in the manner illustrated by Eq. (21) or Eq. (22) can be found such that the probabilistic RBD objective given in Eq. (1) is achieved. This unique relationship between η and p_T exists for a single set of design parameters. For this relationship to be useful for RBD code calibration, it should be insensitive to changes in the design parameters over a range of practical values. For example, almost the same relationship should apply for undrained shear strength varying from 25 kPa (soft clay) to 200 kPa (very stiff clay). Ching & Phoon (2011) presented four common geotechnical examples to demonstrate the relative stability of this η and p_T relationship empirically. They did not study the effect of changes in soil profiles. This important variation in the design scenario is studied in Section 4. It is worthy to note in passing that the change in soil profile must not result in a change of the failure mechanism (or performance function). For example, the failure mechanism for a shallow foundation resting on a thin layer of dense sand overlying soft clay is not the same as the classical Buisman-Terzaghi mechanism for homogeneous soils. It is obvious that the $\eta - p_T$ relationships are distinctively different for different performance functions.

Note that the probability distributions of S and S_{csg} are not available in analytical forms as mentioned in Section 3.3. They can only be characterized empirically using Monte Carlo simulation. Nonetheless, there is an important computational difference between using empirical distributions in FORM or using empirical distributions to compute quantiles. The former creates potential numerical instabilities while the latter can be carried out in a very robust non-parametric way using ranks. In this study, S^{η}, S_{csg}^{η}, and similar statistics are estimated correctly using Monte Carlo simulation. The ad-hoc lognormal assumption adopted in Section 3.3 is not applied in the uniform quantile approach. Similar to the FORM design point calibration approach, two variations are considered. The first variation is termed "uniform quantile –

standard", which calibrates the usual resistance and load factors that would be applied to drilled shafts outside the calibration domain for validation. The second variation is termed "uniform quantile – η", which applies the calibrated quantile η directly to drilled shafts outside the calibration domain for validation.

It is worth emphasizing here that the uniform quantile approach bears no theoretical resemblance to the application of quantile in characteristic/nominal values. The latter refers to Eq. (19) or Eq. (20). The quantile is prescribed by design codes without reference to the target probability of failure. For example, a quantile between 5% and 10% is typically prescribed for the concrete compressive strength, f_{cu}, in structural design codes. The main purpose of this definition is to produce a suitably conservative compressive strength that varies consistently with the coefficient of variation of f_{cu}. The same quantile is applied to different performance functions, for example moment/shear capacity of a beam or compression capacity of a column. The quantile appearing in Eq. (21) or Eq. (22) is fundamentally different. It is calibrated rather than prescribed to achieve a specific target probability of failure. It decreases in a relatively unique way with the target probability of failure for a given performance function. It is intrinsically related to the performance function. Hence, the quantile for a given soil property, say undrained shear strength, will vary when the property is applied within the context of different performance functions, even for the same target probability of failure.

4 VALIDATION STUDIES

4.1 *FORM-standard versus uniform quantile-standard*

In this section, resistance and load factors are calibrated using the "FORM – standard" approach and the "uniform quantile – standard" approach. These approaches have been presented in Section 3.3 and Section 3.4, respectively. A single drilled shaft from Group ALL is selected for calibration. Note that Group ALL contains drilled shafts installed in soil profiles with clay, sand, gravel, and rock layers. The calibrated resistance and load factors are applied to determine the mean dead load corresponding to each drilled shaft in Table 1 in the following way for LRFD:

$$\mu_{LD} = \frac{\gamma_{total}\mu_S}{\left(\gamma_D + \gamma_L \dfrac{\mu_{LL}}{\mu_{LD}}\right)} = \frac{\gamma_{total}\mu_S}{(\gamma_D + 0.5\gamma_L)} \tag{23}$$

in which γ_{TOTAL}, γ_D, γ_L = resistance and load factors computed from a *single* shaft in Group ALL using the FORM – standard/uniform quantile – standard approach, μ_S = mean total side resistance for any validation shaft in Table 1, and μ_{LD} = mean dead load required to satisfy the LRFD format corresponding to an assumed mean live load to mean dead load ratio of 0.5. It is easy to generalize Eq. (23) to the MRFD formats shown in Table 3.

It is more common to compute the shaft depth for a given set of loads in actual foundation engineering practice. However, the foundation depths are already given in Table 1. Hence, the mean dead load implied by the LRFD format is computed. This "design" approach is rather unorthodox, but it has no impact on the evaluation of the performance of the LRFD/MRFD formats presented in Table 4 below. In other words, the ability of the LRFD/MRFD formats to achieve a uniform target reliability index of 3 for the validation shafts can be evaluated by computing the foundation depth from a given set of loads or vice-versa. The latter approach has been applied by Phoon et al. (2003b) as well. The effect of varying the coefficients of variation in the basic random variables is studied in Section 4.4. The effect of varying the mean live load to mean dead load ratio is studied in Section 4.5. It is of interest to note that there are only 11 shafts in Group ALL. The rest of the shafts (31 shafts) are installed in soil profiles with 3 or less soil layers. In short, 31/42 = 74% of the validation shafts are installed in soil profiles that are distinctively different from the 4-layer profile in Group ALL.

Once the mean dead load is computed using Eq. (23), the "actual" reliability index of each validation shaft can be estimated using Monte Carlo simulation. Note that the same nonlinear performance function and the same set of basic random variables (model and measurement errors) presented in Section 3.1 are applied to estimate the "actual" reliability indices for all validation shafts, regardless of the code formats under study. It has been emphasized in Section 3.3 that the performance function and basic random variables presented in Section 3.1 constitute our current best understanding of "reality". Hence, the reliability indices estimated using these realistic physical and probabilistic models are described as "actual" in this sense. For brevity, the term "actual" is dropped from hereon, because all reliability indices reported in

Section 4 are "actual". The ad-hoc lognormal assumption occasionally used in the FORM design point method is purely applied at the calibration stage. Once the resistance and load factors are calibrated, it is no longer relevant to the validation studies presented herein.

For each calibration shaft in GROUP ALL, 42 reliability indices can be determined. This calibration/validation exercise is carried out for all the 11 shafts in GROUP ALL, resulting in 42 × 11 = 462 reliability indices. The mean, coefficient of variation, highest value and lowest value of these reliability indices are reported in Table 4 under the column heading "1 shaft". It is worthy to clarify here that Eq. (23) is the same regardless of the RBD calibration approach (FORM – standard or uniform quantile – standard) used. The RBD calibration approach only affects the specific numerical values of the resistance and load factors used in Eq. (23). Note that the reliability index = 4.75 (corresponding to a probability of failure = 10^{-6}) appearing in Table 4 is an *error flag* indicating that the probability of failure is too small and cannot be estimated using the Monte Carlo simulation sample size = 10^6 adopted in this study. This right censorship will affect the mean and coefficient of variation reported in Table 4. In addition, statistical errors associated with Monte Carlo simulation increase with increasing reliability index. For example, probabilities of failure smaller than 10^{-5} are considered unreliable for a sample size = 10^6. Nonetheless, the performance data presented in Table 4 are useful in a qualitative sense to evaluate the performance of LRFD and MRFD formats. The LRFD format calibrated using FORM – standard is commonly adopted in numerous RBD codes in North America. The associated data are shaded in grey, because they provide useful benchmarks to measure the performance of other code formats and other calibration approaches.

Focusing on the column headings "1 shaft" in Table 4, it is apparent that the FORM approach is inferior to the uniform quantile approach. For LRFD, the mean reliability index is larger than 3 because of the ad-hoc lognormal assumption imposed on lumped random variable, S, during RBD calibration. This systematic bias is not present in the uniform quantile calibrated LRFD, because the ad-hoc lognormal assumption is not necessary during RBD calibration. For the FORM approach, the mean reliability index also decreases monotonically in the order LRFD, MRFD2, MRFD3, and MRFD4. It is postulated that this effect is caused by the increasing random dimension in the performance function. This undesirable effect is not present in the uniform quantile approach. The coefficient of variation (c.o.v.) of β is generally higher for the FORM approach as well. It is interesting to observe that the c.o.v. of β reduces more significantly with the size of the calibration domain when the MRFD format is applied.

Table 4. Comparison between FORM – standard and uniform quantile – standard RBD calibration method.

RBD Eq.	FORM – standard			Uniform quantile – standard		
	1 shaft	14 shafts	41 shafts	1 shaft	14 shafts	41 shafts
LRFD						
mean β	3.29	3.38	3.46	2.77	3.03	3.03
c.o.v. β	0.23	0.16	0.16	0.18	0.16	0.15
highest β	4.75	4.75	4.26	3.85	4.01	3.69
Lowest β	1.41	1.80	1.99	1.41	1.62	1.74
MRFD2						
mean β	3.17	3.23	3.25	2.80	3.05	3.01
c.o.v. β	0.25	0.18	0.13	0.20	0.14	0.12
highest β	4.75	4.75	4.01	4.26	4.26	3.54
Lowest β	0.97	0.30	1.90	1.12	1.59	1.72
MRFD3						
mean β	2.49	3.19	3.19	2.50	3.03	3.02
c.o.v. β	0.21	0.21	0.13	0.21	0.15	0.11
highest β	3.62	4.75	4.26	3.63	4.75	3.55
Lowest β	0.92	-0.32	2.06	1.00	1.53	1.86
MRFD4						
mean β	2.41	3.03	3.13	2.63	3.01	3.01
c.o.v. β	0.21	0.21	0.08	0.14	0.09	0.04
highest β	3.28	4.75	3.78	3.45	4.75	3.25
Lowest β	0.70	-0.45	2.19	1.53	2.10	2.72

*Note: Target reliability index = 3; β = 4.75 (corresponding to a probability of failure of 10^{-6}) is just an error flag indicating that the probability of failure is too small and cannot be estimated using the Monte Carlo simulation sample size = 10^6 adopted in this study.

From a practical engineering perspective, the most important index in Table 4 is the lowest reliability index produced by the population of validation shafts. This index describes the departure from the desired target reliability index for the most unconservative design. Both RBD calibration approaches are comparable based on this index. The clear exception is MRFD4, where the uniform quantile approach performs significantly better than the FORM approach. Overall, the degree of reliability control may be deemed unsatisfactory, but this is hardly surprising given that there is only one shaft in the calibration domain and the soil profiles in the validation domain are diverse. It is possible to view the performance under "1 shaft" as worst case, given the rather unreasonable demand of using one shaft to capture the range of diverse shaft and soil conditions.

4.2 *Effect of number of shafts in calibration domain*

It is more realistic to evaluate the performance of the LRFD and MRFD formats using more than one shaft in the calibration domain. Two calibration domains are studied in this section: (1) 14 shafts are selected randomly from the population of 42 shafts for calibration and the resulting resistance and load factors are validated using the remaining 42-14 = 28 shafts and (2) 41 shafts are selected randomly from the population of 42 shafts for calibration and the resulting resistance and load factors are validated using the remaining 42-41= 1 shaft. The former calibration domain ("14 shafts") can be viewed as a practical case, while the latter calibration domain ("41 shafts") is probably a "best case". In contrast, the "1 shaft" example discussed in Section 4.1 is a "worst case".

In general, the resistance and load factors are functions of the shaft and soil conditions. When there is more than one shaft in the calibration domain, it is necessary to deal with the variations in the resistance and load factors arising from *individual* calibration of different shafts. In this study, the resistance and load factors produced by the calibration shafts are predicted via linear regression using the relative side resistance contribution (S_c/S, S_s/S and S_g/S) as explanatory variables. The coefficients of determination of these regression equations are typically higher than 0.9. These regression equations are then applied to estimate the appropriate resistance and load factors for the validation shafts. Note that Phoon et al. (2003b) apply a different strategy of calibrating all the shafts in a group using optimization, rather than calibrating each shaft individually and then applying regression or other methods to reduce the resistance/load factors to a practical form that can be applied to validation shafts.

For each calibration group consisting of 14 shafts, 28 reliability indices can be determined from the validation shafts. This calibration/validation exercise is carried out 20 times by drawing 14 shafts from the population of 42 shafts repeatedly in a random way, resulting in 28 × 20 = 560 reliability indices. The mean, coefficient of variation, highest value and lowest value of these reliability indices are reported in Table 4 under the column heading "14 shafts". For each calibration group consisting of 41 shafts, 1 reliability index can be determined from the remaining validation shaft. This calibration/validation exercise can only be carried out 42 times, resulting in 1 × 42 = 42 reliability indices. The mean, coefficient of variation, highest value and lowest value of these reliability indices are reported in Table 4 under the column heading "41 shafts".

For the FORM approach, the systematic bias in the mean reliability index cannot be mitigated by increasing the number of shafts in the calibration domain. This is rather obvious as the bias is caused by the ad-hoc lognormal assumption in the case of LRFD and the decreasing mean reliability index from LRFD to MRFD4 is caused by the increasing random dimension. It is also rather obvious that the performance of the LRFD/MRFD formats improves with the number of shafts in the calibration domain. For the uniform quantile approach, the mean, highest β and lowest β converge almost monotonically to the target β with the calibration domain size. This is a desirable result. It provides an assurance that the departures from the target β can be diminished if one is willing to spend efforts to populate the calibration domain. In contrast, the lowest β for the "14 shafts" domain calibrated using FORM can become negative, indicating a probability of failure larger than 50% for the most unconservative validation shaft! This is worse than the lowest β produced by the "1 shaft" calibration domain.

4.3 *FORM – u versus uniform quantile – η*

It has been pointed out that resistance and load factors were developed purely for the practical reason of producing a simplified RBD design equation with the same "look and feel" as existing working stress design equation. For FORM, it is possible to apply the following alternate LRFD approach as given in Eq. (12):

$$\exp\left(\lambda_S + \xi_S u_1^d\right) = \exp\left(\lambda_D + \xi_D u_2^d\right) + \exp\left(\lambda_L + \xi_L u_3^d\right) \tag{24}$$

in which the lognormal parameters (λ_i and ξ_i) are related to the mean (μ_i) and standard deviation (σ_i) of the physical variable as:

$$\lambda_i = \ln(\mu_i) - \frac{1}{2}\xi_i^2$$
$$\xi_i^2 = \ln\left(1 + \frac{\sigma_i^2}{\mu_i^2}\right) \tag{25}$$

The design point in standard space (u_1^d, u_2^d, u_3^d) is determined from a single shaft in Group ALL. The mean dead load for each validation shaft is then computed as:

$$\mu_{LD} = \frac{\mu_S \exp\left(\xi_S u_1^d - 0.5\xi_s^2\right)}{\exp\left(\xi_D u_2^d - 0.5\xi_D^2\right) + \frac{\mu_{LL}}{\mu_{LD}}\exp\left(\xi_L u_3^d - 0.5\xi_L^2\right)} = \frac{\mu_S \exp\left(\xi_S u_1^d - 0.5\xi_s^2\right)}{\exp(\xi_D u_2^d - 0.5\xi_D^2) + 0.5\exp(\xi_L u_3^d - 0.5\xi_L^2)} \tag{26}$$

in which μ_S = mean total side resistance of the *validation* shaft. The design point in standard space (u_1^d, u_2^d, u_3^d) is a vector quantity. At present, there is no simple method of applying the FORM-u calibration approach to more than 1 shaft in the general case. Hence, the results in Table 5 for FORM – u are restricted to the "1 shaft" case.

For the uniform quantile approach, it is possible to apply the following alternate LRFD approach as given in Eq. (21):

$$\exp\left(\lambda_S + \xi_S k_\eta\right) = \exp\left(\lambda_D + \xi_D k_{1-\eta}\right) + \exp\left(\lambda_L + \xi_L k_{1-\eta}\right) \tag{27}$$

in which $k_\eta = \Phi^{-1}(\eta)$ and $k_{1-\eta} = \Phi^{-1}(1-\eta)$. The quantile η can be calibrated from a single shaft. For a calibration domain containing more than one shafts, η can also be predicted via linear regression using the relative side resistance contribution (S_c/S, S_s/S and S_g/S) as explanatory variables. The coefficients of determination of these regression equations are typically higher than 0.9. The mean dead load for each validation shaft is then computed as:

$$\mu_{LD} = \frac{\mu_S \exp\left(\xi_S k_\eta - 0.5\xi_s^2\right)}{\exp\left(\xi_D k_{1-\eta} - 0.5\xi_D^2\right) + \frac{\mu_{LL}}{\mu_{LD}}\exp\left(\xi_L k_{1-\eta} - 0.5\xi_L^2\right)} = \frac{\mu_S \exp\left(\xi_S k_\eta - 0.5\xi_s^2\right)}{\exp(\xi_D k_{1-\eta} - 0.5\xi_D^2) + 0.5\exp(\xi_L k_{1-\eta} - 0.5\xi_L^2)} \tag{28}$$

Eq. (27) and Eq. (28) are presented for conceptual clarity only and for comparison with Eq. (26). In practice, the η quantile of S is estimated directly from Monte Carlo simulation in this study, rather than using the ad-hoc lognormal fit in the above equations.

From the performance data presented in Table 5 below, it is clear that the FORM – u method is slightly better than the FORM – standard method. The margin of improvement is probably not practically significant. On the other hand, the uniform quantile – η is significantly better than the uniform quantile – standard method. It is of practical interest to examine the performance of the "14 shafts" calibration domain. The mean β is almost equal to the target β, the c.o.v is small, and perhaps most importantly, the lowest β is consistently above two. This level of lowest β cannot be achieved consistently even with the best case "41 shafts" calibration domain in Table 4. In Table 5, the performance of the "41 shafts" calibration domain is close to perfect.

There is a practical cost associated with the application of the uniform quantile – η approach that should be highlighted. The engineer is required to perform Monte Carlo simulation to estimate the η quantile of lumped random variables such as S and S_{csg}. The sample size for quantile estimate is typically smaller than that for the probability of failure estimate. For example, if η = 0.05, a sample size of $10/\eta$ = 200 is quite adequate. Hence, the computational cost is not beyond the reach of a PC platform. Nonetheless, the engineer is expected to be comfortable with Monte Carlo simulation. The uniform quantile – standard approach does not require the engineer to perform any Monte Carlo simulation. The code writer must perform Monte Carlo simulation to produce the resistance and load factors, but once these factors are available, the user only need to calculate a single set of nominal resistances and loads. It is possible to use the uniform quantile – η approach by making an ad-doc lognormal assumption for the lumped resistance, S, as shown in Eq. (28). This obviates the need for the user to carry out Monte Carlo simulation, but the resulting mean β will *not* be unbiased such as that shown in Table 5. The optimum code format associated with the uniform quantile – η approach is possibly the partial factor approach in which the quantile is applied on measured soil parameters, rather than lumped resistance components such as the total side resistance. The quantile for a soil parameter can be estimated directly from a set of measurements without the need to perform Monte Carlo simulation.

Table 5. Comparison between FORM – u and uniform quantile – η RBD calibration method.

RBD Eq.	FORM – u			Uniform quantile – η		
	1 shaft	14 shafts	41 shafts	1 shaft	14 shafts	41 shafts
LRFD						
mean β	3.50	–	–	2.95	3.00	3.00
c.o.v. β	0.12	–	–	0.03	0.03	0.03
highest β	4.75	–	–	3.15	3.18	3.15
Lowest β	2.86	–	–	2.60	2.67	2.68
MRFD2						
mean β	3.01	–	–	2.84	3.02	3.00
c.o.v. β	0.14	–	–	0.14	0.06	0.03
highest β	4.11	–	–	3.67	4.75	3.22
Lowest β	1.04	–	–	2.02	2.65	2.65
MRFD3						
mean β	2.62	–	–	2.59	3.00	3.01
c.o.v. β	0.15	–	–	0.15	0.06	0.05
highest β	3.54	–	–	3.55	4.11	3.32
Lowest β	1.02	–	–	1.68	2.18	2.66
MRFD4						
mean β	2.41	–	–	2.63	3.01	3.01
c.o.v. β	0.21	–	–	0.14	0.06	0.04
highest β	3.31	–	–	3.53	4.75	3.25
Lowest β	0.69	–	–	1.53	2.35	2.74

*Note: Target reliability index = 3; β = 4.75 (corresponding to a probability of failure of 10^{-6}) is just an error flag indicating that the probability of failure is too small and cannot be estimated using the Monte Carlo simulation sample size = 10^6 adopted in this study.

4.4 *"Unexpected" change in the coefficients of variation*

The validation studies conducted in Section 4.1, 4.2, and 4.3 are based on a single set of coefficients of variation for the model errors (ε_{Sc}, ε_{Ss}, ε_{Sg}, ε_{Sr}) and the measurement errors (e_{su}, $e_{\sigma'vs}$, $e_{\sigma'vg}$, e_{qu}). It is of practical interest to evaluate the performance of the code formats calibrated using one set of coefficients of variation when they are applied on validation shafts associated with lower/higher coefficients of variation (c.o.v.s). The c.o.v.s for (ε_{Sc}, ε_{Ss}, ε_{Sg}, ε_{Sr}) and (e_{su}, $e_{\sigma'vs}$, $e_{\sigma'vg}$, e_{qu}) are modified as follows: (1) uniformly reduce all c.ov.s by a factor of 0.5 and (2) uniformly increase all c.o.v.s by a factor of 1.5. The modified c.o.v.s are related to the calculation of side resistances. The c.o.v.s of the dead and live load remain unchanged. The mean live load to mean dead load ratio remains unchanged at 0.5.

It is important to point out that the performance shown in Table 6 (reduce c.o.v.s by 50%) and Table 7 (increase c.o.v.s by 150%) refers to a worst case calibration scenario in which the variations in the c.o.v.s are not included in the calibration shafts. In other words, Table 6 and Table 7 illustrates the performance of LRFD/MRFD formats when they are applied to design scenarios that are "unexpected" and hence, not considered by the code writer. With this observation in mind, it is not surprising that the performance shown in Table 6 and Table 7 are worse than that shown in Table 4 and Table 5. The FORM – standard approach is not robust against unexpected design scenarios, even when the calibration domain contains "41 shafts". It is rather obvious that it is not the total number of calibration shafts that is important per se. In the extreme, one cannot expect the LRFD/MRFD formats to perform adequately if they have been calibrated using say 100 near identical calibration shafts. The outcome is entirely different if the 100 calibration shafts are carefully selected to cover all expected design scenarios. In some code calibration methods, more commonly encountered design scenarios are assigned more weightage in the calibration domain by using more calibration shafts for instance. For the case of "14 shafts", it is possible to produce bizarre results in which the lowest β = -1.98 for MRFD3 and lowest β = -1.59 for MRFD4 when the c.o.v.s are reduced in Table 6. In other words, the designs become even more unconservative, although the underlying uncertainties governing side resistances are smaller!

The uniform quantile – standard approach will produce designs that are safer when c.o.v.s are reduced or designs that are less safe when c.o.v.s are increased. Its behavior is stable in this sense, but it is unable to achieve the prescribed target reliability index under an unexpected change in the c.o.v. that is not considered in the calibration domain. The uniform quantile – η approach is able to accommodate an unexpected change in the c.o.v., particularly when the LRFD/MRFD4 format is adopted. Note the c.o.v. β is lower when LRFD is adopted, but the MRFD4 format produces a mean β closest to the target value.

Table 6. Performance of LRFD/MRFD formats when applied to validation shafts with coefficients of variation of model and measurement errors reduced by **a factor of 0.5**.

RBD Eq.	FORM – standard		Uniform quantile – standard		Uniform quantile – η	
	14 shafts	41 shafts	14 shafts	41 shafts	14 shafts	41 shafts
LRFD						
mean β	4.69	4.69	4.55	4.57	3.39	3.41
c.o.v. β	0.05	0.04	0.09	0.08	0.03	0.03
highest β	4.75	4.75	4.75	4.75	3.62	3.57
Lowest β	3.53	3.71	3.13	3.27	3.06	3.14
MRFD2						
mean β	4.59	4.69	4.60	4.62	3.40	3.19
c.o.v. β	0.12	0.04	0.06	0.06	0.03	0.06
highest β	4.75	4.75	4.75	4.75	3.65	3.58
Lowest β	0.59	3.59	3.13	3.26	3.04	3.00
MRFD3						
mean β	4.40	4.69	4.59	4.64	3.29	3.19
c.o.v. β	0.23	0.04	0.07	0.06	0.05	0.06
highest β	4.75	4.75	4.75	4.75	4.26	3.48
Lowest β	-1.98	3.92	3.00	3.53	2.91	2.72
MRFD4						
mean β	4.49	4.71	4.66	4.69	3.20	3.03
c.o.v. β	0.17	0.03	0.04	0.02	0.11	0.05
highest β	4.75	4.75	4.75	4.75	4.75	3.41
Lowest β	-1.59	4.21	2.57	4.47	1.78	2.68

*Note: Target reliability index = 3; β = 4.75 (corresponding to a probability of failure of 10^{-6}) is just an error flag indicating that the probability of failure is too small and cannot be estimated using the Monte Carlo simulation sample size = 10^6 adopted in this study.

Table 7. Performance of LRFD/MRFD formats when applied to validation shafts with coefficients of variation of model and measurement errors reduced by a **factor of 1.5**.

RBD Eq.	FORM – standard		Uniform quantile – standard		Uniform quantile – η	
	14 shafts	41 shafts	14 shafts	41 shafts	14 shafts	41 shafts
LRFD						
mean β	2.46	2.39	2.19	2.21	2.77	2.77
c.o.v. β	0.17	0.16	0.17	0.16	0.02	0.02
highest β	3.31	3.08	2.85	2.72	2.87	2.85
Lowest β	1.22	1.34	1.10	1.17	2.52	2.55
MRFD2						
mean β	2.34	2.37	2.17	2.20	2.83	2.81
c.o.v. β	0.16	0.14	0.14	0.13	0.08	0.04
highest β	3.06	2.86	2.75	2.63	4.75	3.38
Lowest β	0.77	1.28	1.05	1.16	2.48	2.52

RBD Eq.	FORM – standard		Uniform quantile – standard		Uniform quantile – η	
	14 shafts	41 shafts	14 shafts	41 shafts	14 shafts	41 shafts
MRFD3						
mean β	2.29	2.32	2.25	2.20	2.87	2.87
c.o.v. β	0.19	0.13	0.17	0.12	0.07	0.05
highest β	3.35	2.77	4.75	2.60	4.11	3.28
Lowest β	-0.54	1.39	1.09	1.24	2.33	2.51
MRFD4						
mean β	2.25	2.28	2.22	2.19	2.98	2.97
c.o.v. β	0.22	0.09	0.12	0.05	0.09	0.05
highest β	4.75	2.73	4.75	2.42	4.75	3.32
Lowest β	-0.15	1.47	1.31	1.97	2.22	2.63

*Note: Target reliability index = 3; β = 4.75 (corresponding to a probability of failure of 10^{-6}) is just an error flag indicating that the probability of failure is too small and cannot be estimated using the Monte Carlo simulation sample size = 10^6 adopted in this study.

4.5 *"Unexpected" change in mean live load to mean dead load ratio, μ_{LL}/μ_{LD}*

The validation studies conducted in Section 4.1, 4.2, and 4.3 are based on a single load ratio, μ_{LL}/μ_{LD} = 0.5. It is of practical interest to evaluate the performance of the code formats calibrated using μ_{LL}/μ_{LD} = 0.5 when they are applied on validation shafts associated with lower/higher load ratios. Two additional load ratios are considered: (1) μ_{LL}/μ_{LD} = 0.1 and (2) μ_{LL}/μ_{LD} = 1.0. The c.o.v.s of all random variables remain unchanged.

Similar to Section 4.4, the performance shown in Table 8 (μ_{LL}/μ_{LD} = 0.1) and Table 9 (μ_{LL}/μ_{LD} = 1.0) refers to a worst case calibration scenario in which variations in the load ratio are "unexpected" and hence, not considered by the code writer. The FORM – standard approach is generally inferior to the uniform quantile – standard approach in terms of robustness against unexpected change in the load ratio. MRFD3 calibrated using FORM – standard and "14 shafts" is the most inferior as it produces a lowest β = -3.43! The uniform quantile – standard approach typically produces a higher c.o.v. β in contrast to the uniform quantile – η approach. – It is of practical interest to note that the negative reliability indices associated with lowest β for some MRFD formats in Table 6 to Table 9 disappear when the calibration domain is enlarged from "14 shafts" to "41 shafts". It is postulated that the MRFD formats require a larger calibration domain than the LRFD format, because it has more degrees of freedom (more resistance factors). It is worthy to reiterate the obvious guideline that the calibration domain should be as large and as representative as possible. It is also judicious to avoid applying LRFD/MRFD formats to design scenarios not covered in the calibration domain.

The performance data shown in Table 6 to Table 9 appear to indicate that the LRFD/MRFD4 format calibrated using the uniform quantile – η approach can produce consistent designs. The MRFD4 format seems to produce the least departures from the target reliability index if the calibration domain is sufficiently large and representative. The LRFD is more stable for a smaller calibration domain, but it is slightly inferior in achieving the target reliability index on the average.

It has been highlighted in Section 4.3 that the uniform quantile – η approach requires the user to perform Monte Carlo simulation to estimate the quantiles of lumped variables. This is practically inconvenient for the user, but given the significantly better performance of the uniform quantile – η approach, it is worth pondering if this approach is a good compromise between conventional multiple factor formats and full probabilistic analysis.

When the uniform quantile – η approach is applied at the level of soil parameters, rather than lumped resistance components, it has been pointed out previously that the user can estimate the required quantiles for design from measured data without performing Monte Carlo simulation in this special case.

Table 8. Performance of LRFD/MRFD formats when applied to validation shafts with mean live load to mean dead load ratio = **0.1**.

RBD Eq.	FORM – standard		Uniform quantile – standard		Uniform quantile – η	
	14 shafts	41 shafts	14 shafts	41 shafts	14 shafts	41 shafts
LRFD						
mean β	3.49	3.43	2.79	2.82	2.84	2.80
c.o.v. β	0.16	0.16	0.16	0.16	0.05	0.02
highest β	4.75	4.75	3.57	3.51	3.85	2.91
Lowest β	1.91	1.95	1.50	1.59	2.51	2.56
MRFD2						
mean β	3.19	3.23	2.89	2.85	2.88	2.84
c.o.v. β	0.18	0.14	0.15	0.13	0.10	0.04
highest β	4.75	4.01	4.75	3.36	4.75	3.17
Lowest β	-0.17	1.85	1.44	1.59	2.46	2.54
MRFD3						
mean β	3.07	3.13	2.89	2.87	2.82	2.87
c.o.v. β	0.30	0.12	0.16	0.12	0.13	0.05
highest β	4.75	3.81	4.75	3.41	4.75	3.17
Lowest β	-3.43	2.01	1.50	1.71	0.72	2.52
MRFD4						
mean β	3.04	3.06	2.94	2.94	2.90	2.89
c.o.v. β	0.22	0.08	0.14	0.14	0.08	0.04
highest β	4.75	3.63	4.75	4.75	4.34	3.12
Lowest β	0.06	2.15	0.56	0.56	2.32	2.63

*Note: Target reliability index = 3; β = 4.75 (corresponding to a probability of failure of 10^{-6}) is just an error flag indicating that the probability of failure is too small and cannot be estimated using the Monte Carlo simulation sample size = 10^6 adopted in this study.

Table 9. Performance of LRFD/MRFD formats when applied to validation shafts with mean live load to mean dead load ratio = **1.0**.

RBD Eq.	FORM – standard		Uniform quantile – standard		Uniform quantile – η	
	14 shafts	41 shafts	14 shafts	41 shafts	14 shafts	41 shafts
LRFD						
mean β	3.41	3.36	3.02	3.07	3.06	3.06
c.o.v. β	0.15	0.14	0.15	0.15	0.03	0.03
highest β	4.75	3.96	3.85	3.75	3.22	3.19
Lowest β	1.86	2.00	1.69	1.82	2.73	2.78
MRFD2						
mean β	3.14	3.21	3.08	3.06	3.03	3.04
c.o.v. β	0.16	0.12	0.13	0.12	0.05	0.04
highest β	4.75	3.84	4.75	3.62	4.26	3.37
Lowest β	0.72	1.93	1.67	1.81	2.70	2.75
MRFD3						
mean β	3.14	3.13	3.01	3.04	3.04	3.02
c.o.v. β	0.22	0.11	0.18	0.11	0.07	0.05
highest β	4.75	3.72	4.75	3.69	4.75	3.35
Lowest β	0.31	2.07	0.55	1.92	2.50	2.69
MRFD4						
mean β	3.07	3.09	3.05	3.01	3.00	3.00
c.o.v. β	0.20	0.08	0.11	0.04	0.06	0.04
highest β	4.75	3.62	4.75	3.31	4.05	3.24
Lowest β	-0.51	2.19	1.15	2.75	2.33	2.76

*Note: Target reliability index = 3; β = 4.75 (corresponding to a probability of failure of 10^{-6}) is just an error flag indicating that the probability of failure is too small and cannot be estimated using the Monte Carlo simulation sample size = 10^6 adopted in this study.

CONCLUSIONS

Simplified RBD equations in the form of LRFD and MRFD formats are increasingly being adopted in geotechnical engineering design codes worldwide. For example, the LRFD format calibrated using the First-Order Reliability Method (FORM) is adopted by AASHTO. The effectiveness of applying these simplified RBD equations to more realistic ground conditions containing multiple strata has not been studied, despite its obvious practical importance. The objective of this paper is to investigate the degree of deviation from the target reliability index produced when LRFD/MRFD equations are applied to a database of forty-two actual drilled shafts installed in soil profiles underlying the city of Taipei, which contain clay, sand, gravel and rock layers or some partial combination thereof.

Two RBD calibration approaches are studied. They are the FORM design point method and the more recently proposed uniform quantile method (Ching & Phoon 2011). The performance of the LRFD/MRFD formats is measured by computing the actual reliability indices produced by validation shafts designed using the code format under evaluation. These reliability indices are summarized using the following statistics: mean, coefficient of variation, highest value, and lowest value. From a practical engineering perspective, the most important statistic is the lowest reliability index produced by the population of validation shafts. This index describes the departure from the desired target reliability index for the most unconservative design.

In general, for soil profiles with multiple layers, conventional formats containing resistance and load factors are unable to achieve the prescribed target reliability index with the same consistency as that reported for homogeneous soil profiles. This is true regardless of the code format (LRFD/MRFD), the RBD calibration approach (FORM or uniform quantile), and the number of values associated with each resistance factor (one value or regression function). For the drilled shaft examples considered in this study, the direct application of quantiles in the RBD equation (uniform quantile – η approach), rather than converting the quantiles to conventional resistance and load factors (uniform quantile – standard approach), appears to deliver the most consistent and most robust performance. Consistency is measured by the ability to achieve the target reliability index on the average with minimum deviation. Robustness is measured by the ability to cater to unexpected design scenarios not covered in the calibration domain.

There is a practical cost associated with the application of the uniform quantile – η approach that should be highlighted. The engineer is required to perform Monte Carlo simulation to estimate the η quantile of lumped random variables such as S and S_{csg}. Some engineers may not be comfortable with Monte Carlo simulation or find it too tedious to perform. The uniform quantile – standard approach does not require the engineer to perform any Monte Carlo simulation. The code writer must perform Monte Carlo simulation to produce the resistance and load factors, but once these factors are available, the user only need to calculate a single set of nominal resistances and loads. This practical cost does not exist if the uniform quantile – η approach is applied to appropriate parameters where the probability distribution is known analytically or empirically. For the former, the MRFD4 format is feasible because the side resistances for each geomaterial type (S_c, S_s, S_g or S_r) happen to be lognormally distributed when the underlying model and measurement errors are normally distributed. For the latter, the partial factor approach in which the quantile is applied on measured soil parameters, rather than lumped resistance components, is feasible. The quantile for a soil parameter can be estimated directly from a set of measurements without the need to perform Monte Carlo simulation.

ACKNOWLEDGMENTS

The lead author would like to thank the Department of Civil Engineering, National Taiwan University for hosting his sabbatical visit.

REFERENCES

Ang, A. H-S. & Tang, W. H. (1984). Probability concepts in engineering planning and design. Vol. 2 (Decision, risk, and reliability), John Wiley & Sons, New York.

Au S. K. & Beck, J. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. Probabilistic Engineering Mechanics, 16(4), 263–77.

British Standards Institute (2002). Eurocode: Basis of structural design. EN 1990:2002, London.

British Standards Institute (2004). Eurocode 7: Geotechnical design – Part 1: General Rules. EN 1997-1:2004, London.

Ching, J.Y. & Phoon, K. K. (2011). A quantile-based approach for calibrating reliability-based partial factors. Structural Safety (in press).
Ching, J. Y., Chen, J. R. & Phoon, K. K. (2011). Multiple resistance factor design for axial compression of drilled shafts. Journal of Geotechnical and Geoenvironmental Engineering, ASCE (under review).
Construction Industry Research & Information Association (1977). Rationalization of safety and serviceability factors in structural codes. Report 63, CIRIA, London.
Ellingwood, B. R., Galambos, T. V., MacGregor, J. G., and Cornell, C. A. (1980). Development of probability-based load criterion for American National Standard A58. Special Publication 577, National Bureau of Standards, Washington.
Moses, F. and Larrabee, R. D. (1988). Calibration of draft RP2A-LRFD for fixed platforms. Proc. 20th Offshore Tech. Conf. (2), Houston, 171-180.
Phoon, K. K., Kulhawy, F. H. & Grigoriu, M. D. (2003a). Development of a reliability-based design framework for transmission line structure foundations. Journal of Geotechnical and Geoenvironmental Engineering, ASCE, 129(9), 798 – 806.
Phoon, K. K., Kulhawy, F. H. & Grigoriu, M. D. (2003b). Multiple resistance factor design (MRFD) for spread foundations. Journal of Geotechnical and Geoenvironmental Engineering, ASCE, 129(9), 807 – 818.
Phoon K. K. (2006). Modeling and simulation of stochastic data", Proc. GeoCongress, ASCE, Atlanta, Feb 26 – Mar 1, 2006.