Future Directions in Reliability-Based Geotechnical Design

Gordon A. Fenton[1], M.ASCE, D.V. Griffiths[2], F.ASCE, and Farzaneh Naghibi[1]

[1]Department of Engineering Mathematics, Dalhousie University, Halifax, NS, Canada B3J 2X4.
[2]Department of Civil and Environmental Engineering, Colorado School of Mines, Golden, CO 80401, United States.

ABSTRACT: World-wide, geotechnical design codes-of-practice are increasingly targeting acceptable failure probabilities, rather than factors of safety, since the latter do not provide an accurate estimate of safety, despite their name. This trend requires an ever-increasing understanding of the probabilistic behaviour of geotechnical systems. As a result, probabilistic geotechnical models are becoming more complex, yet more realistic. In particular, models which consider the effects of the ground's spatial variability on failure probability of geotechnical systems are rapidly gaining popularity. This is because it is well known that spatial variability leads to weakest paths which are preferentially followed by geotechnical failure mechanisms. The paper begins by looking at the current state-of-the-art in probabilistic ground models. The effect of spatial variability on geotechnical system failure probability is discussed, followed by how the random finite element method (RFEM) has and can be used to aid in the calibration of geotechnical design codes-of-practice. The paper finally looks at what is needed in the future to further improve cost effective geotechnical design practices while increasing overall geotechnical system reliability.

INTRODUCTION

Most geotechnical design codes have been migrating towards reliability-based concepts for several decades now. From both the code development and the designer point of view, the problem becomes one of how to produce a design which achieves the code specified target reliability? In recent years, the solution to this design problem has been through the use of a methodology called Load and Resistance Factor Design (LRFD) embedded within a Limit States Design (LSD) framework. Essentially this means that load (α_i) and resistance (φ_g) factors are calibrated so that the satisfaction of a design equation of the general form

$$\varphi_g \hat{R} \geq \sum_i \alpha_i \hat{F}_i \tag{1}$$

leads to an acceptably safe geotechnical system for each limit state. In eq. 1, $\hat{R}$ is the characteristic (nominal) geotechnical resistance and $\hat{F}_i$ is the *i*'th characteristic (nominal) load effect. In most civil engineering design codes, the load factors are specified in the structural part of the code, and so the challenge on the geotechnical "resistance" side is to find the values of the resistance factor, φ_g which achieve the code specified target reliability.

There are many uncertainties that must be considered in order to calibrate the resistance factors. Perhaps one of the major questions that must be answered in order to develop a rational reliability-based geotechnical design code is how to properly account for the fact that the ground is a highly and spatially variable material.

This paper concentrates on this last question. It starts by looking at the current state of the art in modeling the ground. In particular, how to best include spatial variability in the ground's properties, and what effect spatial variability has on the ground response and failure probability of geotechnical systems? If the spatial variability of the ground can be realistically modeled and failure probabilities reasonably estimated, the paper then discusses how this information can used be to calibrate the LRFD to achieve target levels of safety (or, equivalently, sufficiently small failure probabilities). Finally, the paper discusses future directions and requirements in the further development of reliability-based geotechnical design.

MODELING THE GROUND

Random fields are convenient mathematical models of the ground, representing the ground's spatial variability in a simple way. A random field is basically a collection of random variables, one for each point in the field. If the field is one-dimensional, then the points are arranged along a line. A CPT sounding is an example of a one-dimensional random field, as illustrated in Figure 1. Every point in the z direction (depth) can be represented by a random variable, $q_1 = q_c(z_1)$, $q_2 = q_c(z_2), \ldots$. The set of random variables is characterized by a joint probability density function which, for a continuous random process, is infinite-dimensional. In practice, the depth direction in Figure 1 is discretized into n points and the random variables at the n points are used to approximate the continuous process. In this case the joint probability density function becomes n-dimensional. Unless n is very small (in which case the continuous process may be poorly approximated) specifying a complete n-dimensional probability density function can still be very cumbersome, and in most cases impractical. In practice, a number of simplifying assumptions are made and this section briefly summarizes the resulting simplified random field models.

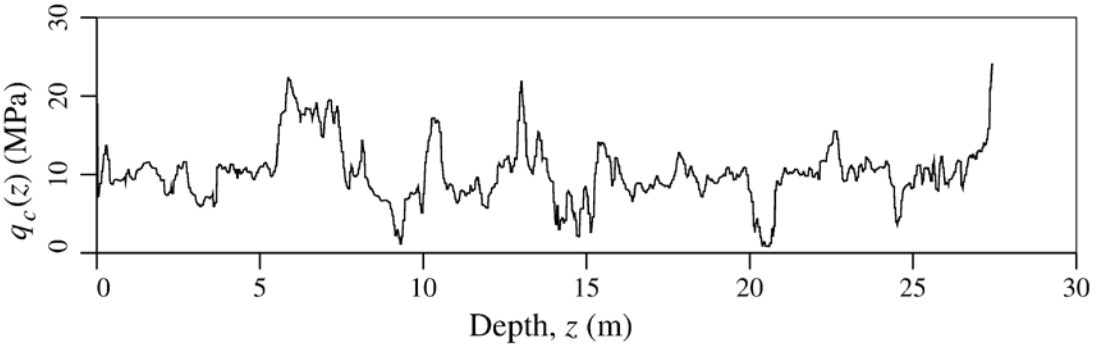

Fig. 1. Tip resistance $q_c(z)$ measured over depth z by a cone penetrometer.

Simple Stationary Random Fields

The random field will now be generically referred to as $X\left(\underset{\sim}{x}\right)$, which varies with spatial position $\underset{\sim}{x}$. One major simplification is to assume that *X* is normally distributed – $X\left(\underset{\sim}{x}\right)$ is then commonly referred to as a Gaussian random field. The main advantage to this assumption is that a jointly normally distributed random field is completely specified by its first two moments: its mean and covariance structure. If it is further assumed that the field is statistically *stationary*, then its mean and variance are everywhere the same (spatially constant) and the correlation coefficient between any two points is dependent only on the distance (and possibly orientation) between the points. The correlation coefficient between two points $X\left(\underset{\sim}{x}_1\right)$ and $X\left(\underset{\sim}{x}_2\right)$ is commonly expressed using a function such a

$$\rho\left(\underset{\sim}{\tau}\right) = \exp\left\{-\sqrt{\left(\frac{2\tau_x}{\theta_x}\right)^2 + \left(\frac{2\tau_y}{\theta_y}\right)^2 + \left(\frac{2\tau_z}{\theta_z}\right)^2}\right\} \tag{2}$$

where $\underset{\sim}{\tau} = \underset{\sim}{x}_1 - \underset{\sim}{x}_2$, the vector between the two points, has components $\underset{\sim}{\tau} = \left(\tau_x, \tau_y, \tau_z\right)$ in three-dimensional space. The parameters θ_x, θ_y, and θ_z are the directional *correlation lengths*, which basically govern how rapidly the random field varies. Small correlation lengths lead to rapidly varying random fields. In the limit, as the correlation lengths go to zero, all points in the field become independent – the field becomes infinitely rough (white noise). At the other extreme, as the correlation lengths go to infinity, the field becomes spatially constant – a single random variable.

If $\theta_z = \theta_x = \theta_y = \theta$, then the field is said to be *isotropic*, an assumption which might be made in non-site specific studies, i.e., where the actual relationship between horizontal and vertical correlation lengths is unknown, or when the correlation lengths are basically unknown and only the effect of their magnitude on probabilistic site response is being investigated. Figure 2 shows a possible realization of a two-dimensional isotropic random field (contrast this to Figure 3 which is anisotropic).

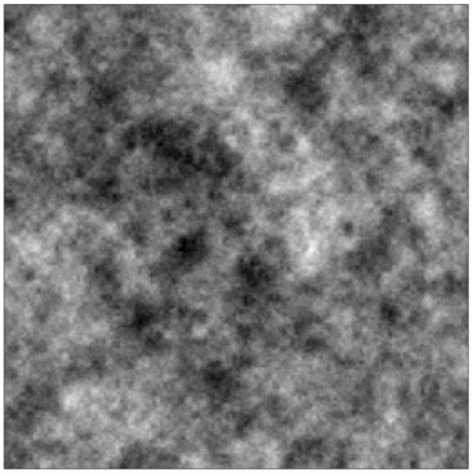

Fig. 2. Realization of a spatially variable isotropic two-dimensional random field with θ equal to 1/10 of the field dimension.

Anisotropic Random Fields

Most soil profiles exhibit some degree of layering, usually horizontally, so that the soil varies more rapidly in the vertical direction than in the horizontal directions. Assuming that *z* is the vertical direction, a more rapidly varying field in the vertical direction can be achieved simply by taking θ_z to be smaller than θ_x or θ_y, the latter of which are often taken to be equal. In this case, the random field is said to be anisotropic – the rate that the correlation function decays with distance now depends on direction.

Another way to achieve an anisotropic random field is simply to simulate an isotropic random field and then either stretch it in the direction(s) of the longer correlation length, or compress it in the direction(s) of the shorter correlation length. For example, if a random field of dimension $L_x \times L_y \times L_z$ is desired where $\theta_x = \theta_y = 1$ and $\theta_z = 0.25$, then simulating an isotropic random field, with $\theta = 1$, of dimension $L_x \times L_y \times 4L_z$ and then compressing it to dimension $L_x \times L_y \times L_z$ will yield a field with the proper statistics.

Figure 3 illustrates a realization of an anisotropic random field having a horizontal correlation length equal to 10 times the vertical correlation length.

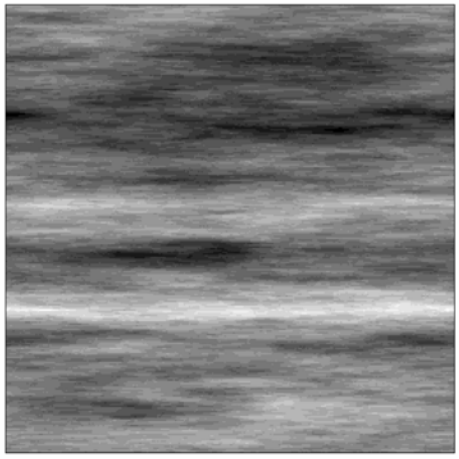

Fig. 3. Realization of a spatially variable two-dimensional random field with the horizontal correlation length equal to 10 times the vertical correlation length.

Non-Stationary Random Fields

On occasion the ground parameters do show a spatial trend in the mean, e.g., when the mean shear strength increases with depth. Very rarely would the standard deviation be assumed to vary with position, although it might be assumed that the coefficient of variation remains constant, so that if the mean changes, the standard deviation changes accordingly. In the event that the mean and/or standard deviation do change with position, the random field can be expressed in the form

$$X\left(\underset{\sim}{x}\right) = \mu\left(\underset{\sim}{x}\right) + \sigma\left(\underset{\sim}{x}\right) G\left(\underset{\sim}{x}\right) \tag{3}$$

where $G\left(\underset{\sim}{x}\right)$ is a stationary zero mean, unit variance, Gaussian random field. What Eq. 3 implies is that a random field which is non-stationary in the mean and/or standard deviation is easily derived from a stationary random field. If the correlation

function (e.g., Eq. 2) is non-stationary, then simulation of the random field becomes more complicated and direct techniques, such as Covariance Matrix Decomposition (see, e.g., Fenton and Griffiths, 2008), may be required.

Non-Gaussian Random Fields

Most ground properties are non-negative and so cannot be truly modeled by the normal (Gaussian) distribution whose range is $(-\infty, +\infty)$. These properties are better modeled by a non-negative distribution, and possibilities include the exponential, Weibull, Chi-square, Gamma, and lognormal. One significant advantage to using the lognormal distribution to model non-negative engineering properties is that it arises from a simple transformation of the normal distribution:

$$X(\underset{\sim}{x}) = \exp\left\{\mu(\underset{\sim}{x}) + \sigma(\underset{\sim}{x})G(\underset{\sim}{x})\right\} \tag{4}$$

and so is still fully characterized by only the first two moments: the mean and covariance structure of $G(\underset{\sim}{x})$. Because of this advantage, one rarely sees the other distributions (e.g., Weibull, etc) employed for random fields. Some ground properties are also bounded above. For example, porosity, degree of saturation, and friction angle are all bounded both below and above. Possible bounded distributions include the uniform (which assumes equilikely possible outcomes), Beta, and Tanh. The last is also a transformation of a Gaussian random field

$$X(\underset{\sim}{x}) = a + 0.5(b-a)\left[1 + \tanh\left(\frac{m + sG(\underset{\sim}{x})}{2\pi}\right)\right] \tag{5}$$

where *a* and *b* are the lower and upper bounds and *m* and *s* are location and scale parameters. See Fenton and Griffiths (2008) for more details.

Multiple Random Fields

When the ground is made up of distinct geologic units and the geometry of these units is relatively well known, then modeling the ground generally just involves using a separate random field for each geologic unit. For example, suppose that a site consists of a 5 m thick sand layer overlying a 20 m thick clay layer. Two random fields would probably be used to represent the site, one 5 m thick random field having the mean, variance, and correlation structure appropriate for the sand layer, and one 20 m thick random field having the clay layer statistics.

If the geometry of the geologic units is unknown and therefore random, then the model becomes more complicated. The simplest case would be when the number and type of layers is known, just their thicknesses are unknown. This then might involve simulating random layer thicknesses as a sequence of single random variables and then generating the random fields within each layer thickness realization. If, more realistically, the layer thickness varies randomly over the (x, y) plane, then each layer thickness can be simulated using a two-dimensional random field. The simulation of properties within each layer would then proceed as follows;

1. for each layer, simulate the layer thickness as a two-dimensional random field,
2. simulate a three-dimensional random field of the layer ground properties having thickness equal to the maximum thickness of the layer simulated in step 1,

3. extract the elements of the three-dimensional random field, simulated in step 2, which lie within the simulated layer thickness. These elements then form the final layer property field.

More complex variations are possible to simulate more complex ground stratifications, for example, clay lenses embedded in a sandy soil. In these cases, an "indicator" random field can be used to simulate the boundaries of the "lenses". For example, the following algorithm might be used;

1. simulate a standard Gaussian random field, $G(\underset{\sim}{x})$, having zero mean and unit variance, with some prescribed correlation structure (correlation lengths).
2. denote as "clay" all regions where $G(\underset{\sim}{x}) > c$, where c is some threshold, and as "non-clay" all regions where $G(\underset{\sim}{x}) \le c$. If the ground has more than two ground types, then multiple disjoint and collectively exhaustive ranges can be used to simulate the random boundaries of each material.
3. once the boundaries of each material have been simulated, simulate the material properties within each "lense" using an appropriately specified random field(s).

Multiple random fields in which the geometric aspects are unknown are rarely used in practice. This is because the distribution(s) of the geometric uncertainties can be very difficult to estimate. For example, even specifying the mean and variance of a layer's thickness implies that the layer thickness has been sampled at a reasonable number of locations. If that is the case, then it makes more sense to simply assume the layer thickness is known at the sampled locations. This motivates the final type of random field model to be considered here, as discussed next.

Conditioned Random Fields

When the objective is to simulate spatially varying ground conditions at a site where some samples have been taken, then it makes sense to assume that the ground conditions are (at least approximately) known at the locations where the ground was sampled. This can be done using a combination of random fields and best linear unbiased estimated fields. In particular

$$X_c(\underset{\sim}{x}) = X_u(\underset{\sim}{x}) + \left[X_k(\underset{\sim}{x}) - X_s(\underset{\sim}{x})\right] \tag{6}$$

where $X_c(\underset{\sim}{x}) =$ desired conditional simulation, $X_u(\underset{\sim}{x}) =$ unconditional simulation, $X_k(\underset{\sim}{x}) =$ best linear unbiased estimate of the field based on the known (measured) values at the sample locations, and $X_s(\underset{\sim}{x}) =$ best linear unbiased estimate of the field based on the values of $X_u(\underset{\sim}{x})$ at the sample locations. Further details can be found in Fenton and Griffiths (2008).

THE RANDOM FINITE ELEMENT METHOD

Once a more realistic model of the ground, including its spatial variability, has been developed, the next major challenge is to model the response of the ground to external or internal loads. When the ground is spatially variable, its failure mechanisms become more complex. For example, the traditional symmetric double log-spiral failure mechanism found in most textbooks to predict bearing failure under a spread footing assumes that the ground is spatially constant.

When the ground properties vary spatially, the bearing failure mechanism is no long symmetric and is attracted to weaker zones. Figure 4 shows what the failure mechanism might look like in a real soil. The lighter (weaker) region to the right of the footing attracts the failure mechanism, which is now non-symmetric. The failure mechanism is following the path of least resistance through the ground. What this means is that the traditionally assumed symmetric failure mechanism is unconservative -- it gives a higher strength than actually provided by the ground along its weakest path.

A natural approach to finding the weakest failure mechanism is to employ a finite element model of the ground (see, e.g., Smith and Griffiths, 2004). The basic idea is to simulate a random field of ground properties, map these properties to a finite element mesh and use the finite element method to predict the ground response. Figure 5 shows a cross-section through a finite element model of the ground under a stiff footing for a typical realization of the ground's effective elastic modulus field in a probabilistic settlement analysis.

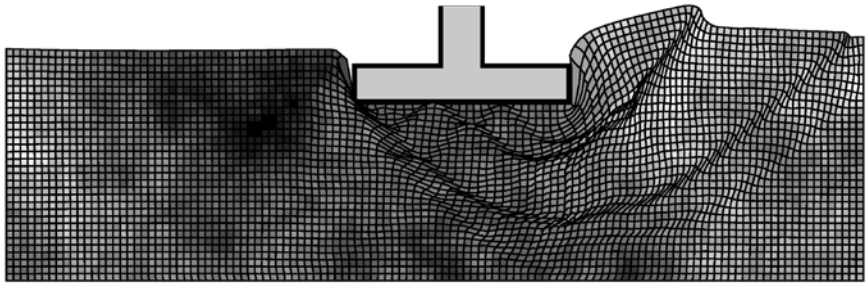

Fig. 4. Non-symmetric "weakest path" failure mechanism for spatially variable ground.

Some discussion of the relative merits of various methods of representing random fields in finite element analysis has been carried out over the years (see, for example, Li and Der Kiureghian, 1993). While using a spatially averaged discretization of the random field is just one approach to the problem, it is appealing in the sense that it reflects the simplest idea of the finite element representation of a continuum as well as the way that soil samples are typically taken and tested in practice, ie. as local averages. Regarding the discretization of random fields for use in finite element analysis, Matthies et al. (1997) makes the comment that "One way of making sure that the stochastic field has the required structure is to assume that it is a local averaging process.", referring to the conversion of a nondifferentiable to a differentiable (smooth) stochastic process. Matthie further goes on to say that the advantage of the local average representation of a random field is that it yields accurate results even for rather coarse meshes.

Following this reasoning, realizations of the ground property random field are produced using the Local Average Subdivision (LAS) method (Fenton and Vanmarcke, 1990). Specifically, LAS produces a discrete grid of local averages, $G_e(\underset{\sim}{x}_i)$, of a standard Gaussian random field, having correlation structure given by, for example, Eq. 2, where $\underset{\sim}{x}_i$ are the coordinates of the centroid of the *i*'th grid cell.

In detail, $G_e(\underset{\sim}{x}_i)$ is the local arithmetic average of a continuous standard Gaussian random field, $G(\underset{\sim}{x})$, over the element having centroid $\underset{\sim}{x}_i$.

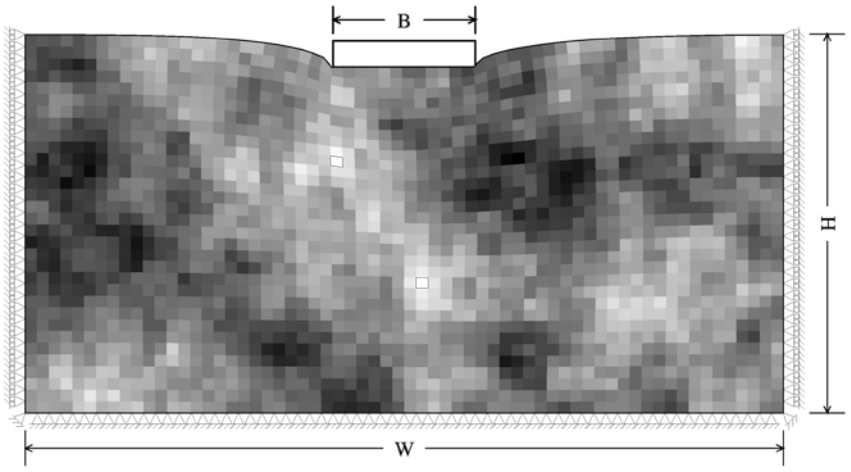

Fig. 5. Cross-section through a realization of the spatially random elastic modulus underlying a footing. Lighter soils are less stiff.

If the ground property in question is assumed to be lognormally distributed, as is commonly the case, these local averages are then mapped to finite element properties according to

$$X_e(\underset{\sim}{x}_i) = \exp\left\{\mu_{\ln X} + \sigma_{\ln X} G_e(\underset{\sim}{x}_i)\right\} \tag{7}$$

where X_e is the property assigned to the *i*'th finite element, $\mu_{\ln X}$ is the mean of $\ln X$, and $\sigma_{\ln X}$ is the *point* standard deviation of $\ln X$.

One of the features of using local arithmetic averaging is that the variance of the average reduces as the element size (averaging dimension) increases. Since a finite element generally employs low-order shape functions to approximate the behaviour of a continuum, the finite element is essentially modeling the average behaviour of the material within the domain of the element. Thus, it makes sense to use an average of the material properties within the element, which implies that the variance of the material property assigned to the element should reduce as the element becomes coarser (more averaging). In other words, the wedding of a local average random field with the (low-order shape function) finite element method is natural and consistent.

EFFECT OF SPATIAL VARIABILITY

To illustrate the effect that spatial variability has on the response of the ground to external or internal loads, two examples will be considered below.

Shallow Foundation Settlement

The RFEM can be used to estimate distribution of settlements of a single footing, as shown in Figure 5, and estimate the probability density function governing total settlement of the footing as a function of footing width for various statistics of the underlying soil. In this example, only the soil elasticity is considered to be spatially

random. In addition, the soil is assumed to be isotropic – that is, the correlation structure is assumed to be the same in both the horizontal and vertical directions. Although soils generally exhibit a stronger correlation in the horizontal direction, due to their layered nature, the degree of anisotropy is site specific. In that this example is demonstrating the basic probabilistic behaviour of settlement, anisotropy is left as a refinement for the reader. The program used to perform the study presented in this example is RSETL2D (Fenton and Griffiths 2002, Griffiths and Fenton 2007; see also http://www.engmath.dal.ca/rfem).

Assuming that the settlement, δ of a single footing is lognormally distributed, as was found to be reasonable by Fenton and Griffiths (2002), having probability density function

$$f_\delta(x) = \frac{1}{\sqrt{2\pi}\sigma_{\ln\delta} x} \exp\left\{ -\frac{1}{2}\left(\frac{\ln x - \mu_{\ln\delta}}{\sigma_{\ln\delta}} \right)^2 \right\}, \qquad 0 \le x < \infty \qquad (8)$$

the task is to estimate the parameters $\mu_{\ln\delta}$ and $\sigma_{\ln\delta}$ as functions of the footing width, *B*, elastic modulus standard deviation, σ_E, and correlation length $\theta_{\ln E}$. Figure 6 shows how the estimator of $\mu_{\ln\delta}$, denoted $m_{\ln\delta}$, varies with $\sigma^2_{\ln E}$ for $B = 0.1H$. All correlation lengths are drawn in the plot, but are not individually labeled since they lie so close together. This observation implies that the mean log-settlement is largely independent of the correlation length, $\theta_{\ln E}$. This is as expected since the correlation length does not affect the mean of a local average of a normally distributed process. Figure 6 suggests that the mean of log-settlement can be closely estimated by a straight line of the form,

$$\mu_{\ln\delta} = \ln(\delta_{det}) + \frac{1}{2}\sigma^2_{\ln E} \qquad (9)$$

where δ_{det} is the `deterministic' settlement obtained from a single finite element analysis (or appropriate approximate calculation) of the problem using $E = \mu_E$ everywhere. This equation is also shown in Figure 6 and it can be seen that the agreement is very good. Even closer results were found for other footing widths.

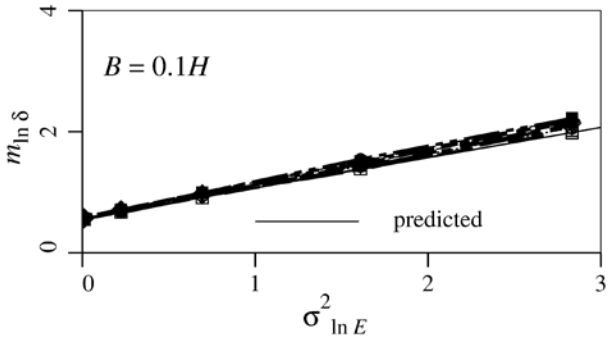

Fig. 6. Estimated mean of log-settlement along with that predicted by Eq. 9.

Estimates of the standard deviation of log-settlement, $s_{\ln\delta}$, are plotted in Figure 7 (as symbols) for two different footing widths. Intermediate footing widths give

similar results. In all cases, $s_{\ln\delta}$ increases to $\sigma_{\ln E}$ as $\theta_{\ln E}$ increases. The reduction in variance as $\theta_{\ln E}$ decreases is due to the local averaging variance reduction of the log-elastic modulus field under the footing (for smaller $\theta_{\ln E}$, there are more 'independent' random field values, so that the variance reduces faster under averaging).

Following this reasoning, and assuming that local averaging of the area under the footing accounts for all of the variance reduction seen in Figure 7, the standard deviation of log-settlement is

$$\sigma_{\ln\delta} = \sqrt{\gamma(B,H)}\,\sigma_{\ln E} \tag{10}$$

where $\gamma(B,H)$ is the variance reduction function, which depends on the averaging region, $B \times H$ as well as on the correlation length, $\theta_{\ln E}$. Since $\sigma_{\ln E}$ is constant for each value of σ_E / μ_E, Figure 7 is essentially a plot of the variance function, $\gamma(B,H)$, illustrating how the variance of a local average decreases as the correlation length decreases. Predictions of $\sigma_{\ln\delta}$ using Eq. 10 are superimposed on Figure 7 using lines. The agreement is remarkable.

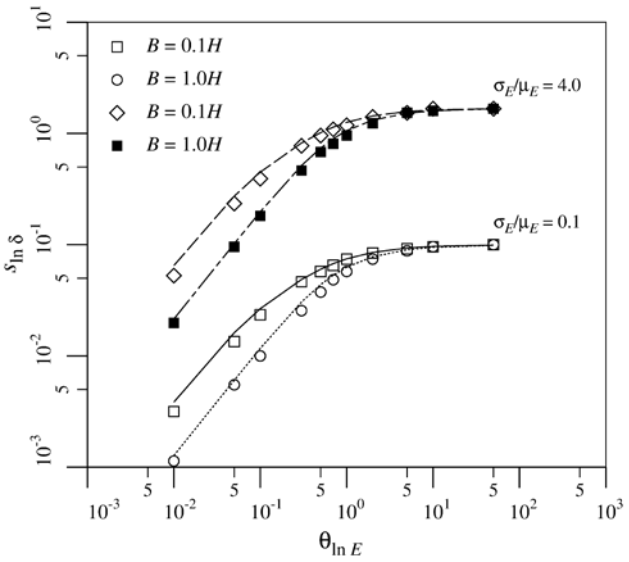

Fig. 7. Comparison of simulated sample standard deviation of log-settlement, shown with symbols, with theoretical estimate via Eq. 10, shown with lines.

An alternative physical interpretation of Eq's 9 and 10 comes by generalizing the settlement prediction to the form

$$\delta = \frac{\delta_{det}\mu_E}{E_g} \tag{11}$$

where E_g is the geometric average of the elastic modulus values over the region of influence,

$$E_g = \exp\left\{\frac{1}{BH}\int_0^H \int_0^B \ln E(x,y)\,dx\,dy\right\} \tag{12}$$

Taking the logarithm of Eq. 11 and then computing its mean and variance leads to Eq's 9 and 10. The geometric mean is dominated by small values of elastic modulus,

which means that the total settlement is dominated by low elastic modulus regions underlying the footing, as would be expected.

These results can be extended to the serviceability limit state design of a single footing. If a square footing of dimension $B \times B$ is considered, the design requirement is to find B and the ratio of the load to resistance factors, α / φ_g, such that

$$\delta_{max} = u_1\left(\frac{\alpha \hat{F}}{B\varphi_g \hat{E}}\right) \tag{13}$$

and

$$\mathrm{P}\left[u_1 \ \frac{F}{BE_{eff}} > u_1\left(\frac{\alpha \hat{F}}{B\varphi_g \hat{E}}\right)\right] = p_m \tag{14}$$

where δ_{max} is the maximum tolerable settlement (serviceability limit state), u_1 is an influence factor (see Fenton et al., 2005, for more details), F is the actual load, E_{eff} is the equivalent elastic modulus as seen by the footing, $\hat{F}$ is the characteristic (nominal) load, $\hat{E}$ is the characteristic (nominal) elastic modulus, and p_m is the maximum tolerable failure probability. In the above, we are assuming that the soil's elastic modulus is the 'resistance' to the load and that it is to be factored due to its high uncertainty.

Five different sampling schemes will be considered in this example, as illustrated in Figure 8. The outer solid line denotes the edge of the soil model, which is 9.6 x 9.6 m in plan and 4.8 m in depth as in Figure 5, and the interior dashed line the location of the footing. The small black squares show the plan locations where the site is virtually sampled. It is expected that the quality of the estimate of E_{eff} will improve for higher numbered sampling schemes. That is, the probability of design failure will decrease for higher numbered sampling schemes, everything else being held constant.

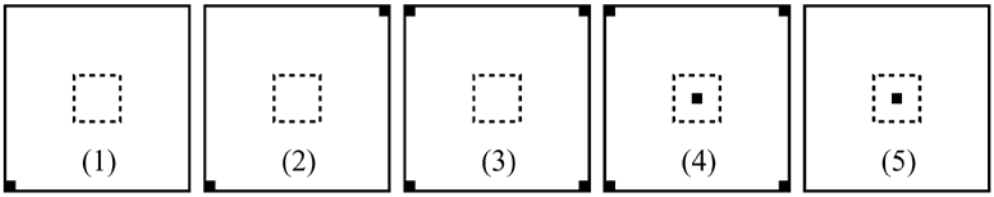

Fig. 8. Sampling schemes considered in this example.

For fixed resistance factor, φ_g, the soil samples allow an estimate of the characteristic elastic modulus, $\hat{E}$ and Eq. 13 can then be used to design the footing. Repeating the design for many realizations of the soil allows the probability that a footing design using φ_g will result in excessive settlement to be estimated. Figure 9 illustrates the effect of correlation length on the probability of excessive settlement, p_f, for sampling scheme #1. It is evident that a) spatial variability of the ground has a strong influence on p_f, and b) that there is a worst case correlation length, in this case around 10 m – which is of the order of the distance from the footing to the sampling point (6.8 m).

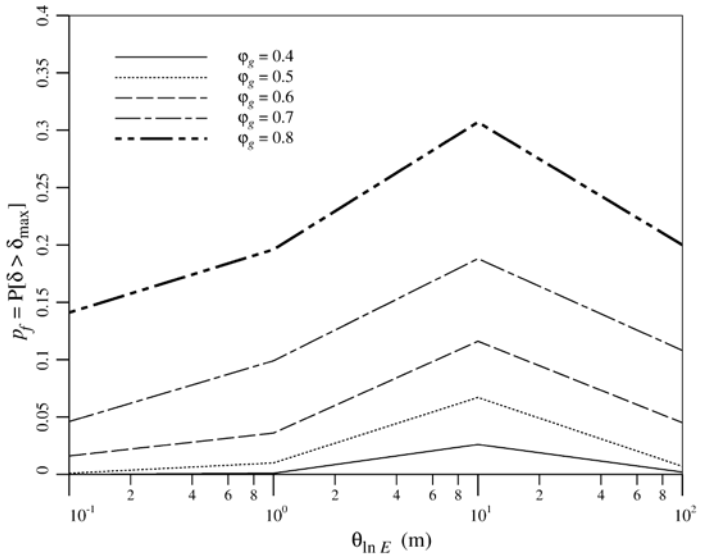

Fig. 9. Effect of correlation length $\theta_{\ln E}$ on probability of excessive settlement $p_f = \mathrm{P}[\delta > \delta_{max}]$.

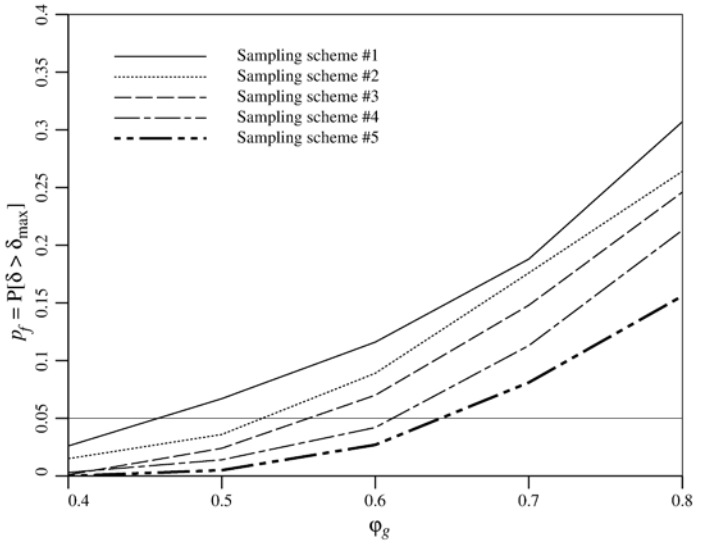

Fig. 10. Effect of resistance factor φ_g on probability of failure $p_f = \mathrm{P}[\delta > \delta_{max}]$ for $v_E = 0.5$ and $\theta_{\ln E} = 10$ m.

Figure 10 shows the failure probability for the various sampling schemes at a coefficient of variation, $v_E = 0.5$, and $\theta_{\ln E} = 10$ m. Improved sampling (i.e. improved understanding of the site) makes a significant difference to the required value of φ_g, which ranges from $\varphi_g \approx 0.46$ for sampling scheme #1 to $\varphi_g \approx 0.65$ for sampling scheme #5, assuming a target probability of $p_m = 0.05$. Note that if a distance-weighted or trend estimate were used, sampling scheme #4 would have been better than #5. In general, more samples are preferable – however, only a simple average was used in this study to estimate the soil properties so that the four samples not taken directly under the footing in sampling scheme #4 actually just "muddy the

waters", decreasing the accuracy of the sample taken under the footing. The overall implications of Figure 10 are that when soil variability is significant, considerable design/construction savings can be achieved when the sampling scheme is improved.

Bearing Capacity

The design of a shallow footing typically begins with a site investigation aimed at determining the strength of the founding soil or rock. Once this information has been gathered, the geotechnical engineer is in a position to determine the footing dimensions required to avoid entering various limit states. In so doing, it will be assumed here that the geotechnical engineer is in close communication with the structural engineer(s) and is aware of the loads that the footings are being designed to support. The limit states that are usually considered in the footing design are serviceability limit states (typically deformation – see example above) and ultimate limit states. The latter is concerned with safety and includes the load-carrying capacity, or *bearing capacity*, of the footing.

This example illustrates an LRFD approach for shallow foundations designed against bearing capacity failure. The design goal is to determine the footing dimensions such that the *ultimate geotechnical resistance* based on characteristic soil properties, $\hat{R}_u$, satisfies

$$\varphi_g \hat{R}_u \geq \sum_i \alpha_i \hat{F}_i \tag{15}$$

where φ_g is the *geotechnical resistance factor*, α_i is the *i*'th *load factor*, and $\hat{F}_i$ is the *i*'th *characteristic load effect*. The relationship between φ_g and the probability that the designed footing will experience a bearing capacity failure will be summarized below (from Fenton et al., 2007) followed by some results on resistance factors required to achieve certain target maximum acceptable failure probabilities for the particular case of a strip footing (from Fenton et al., 2008).

The characteristic ultimate geotechnical resistance $\hat{R}_u$ is determined using characteristic soil properties, in this case characteristic values of the soil's cohesion, *c*, and friction angle, ϕ (note that although the primes are omitted from these quantities it should be recognized that the theoretical developments described in this example are applicable to either total or effective strength parameters).

The characteristic value of the cohesion, $\hat{c}$, is defined here as the median of the sampled observations, c_i^o, which, assuming *c* is lognormally distributed, can be computed using the geometric average,

$$\hat{c} = \left[\prod_{i=1}^{m} c_i^o \right]^{1/m} = \exp\left\{ \frac{1}{m} \sum_{i=1}^{m} \ln c_i^o \right\} \tag{16}$$

The geometric average is used here because if *c* is lognormally distributed, as assumed, then $\hat{c}$ will also be lognormally distributed. The characteristic value of the friction angle is computed as an arithmetic average

$$\hat{\phi} = \frac{1}{m} \sum_{i=1}^{m} \phi_i^o \tag{17}$$

The arithmetic average is used here because ϕ is assumed to follow a symmetric bounded distribution and the arithmetic average preserves the mean. That is, the mean of $\hat{\phi}$ is the same as the mean of ϕ.

To determine the characteristic ultimate geotechnical resistance $\hat{R}_u$, it will first be assumed that the soil is weightless (and thus cohesive). This simplifies the calculation of the ultimate bearing stress q_u to

$$q_u = c\, N_c \tag{18}$$

The assumption of weightlessness is conservative since the soil weight contributes to the overall bearing capacity. This assumption also allows the analysis to explicitly concentrate on the role of $c\, N_c$ on ultimate bearing capacity, since this is the only term that includes the effects of spatial variability relating to *both* shear strength parameters *c* and ϕ.

Bearing capacity predictions, involving specification of the N_c factor in this case, are generally based on plasticity theories (see, e.g., Prandtl, 1921; Terzaghi, 1943; and Sokolovski, 1965) in which a rigid base is punched into a softer material. These theories assume that the soil underlying the footing has properties which are spatially constant (everywhere the same). This type of ideal soil will be referred to as a *uniform soil* henceforth. Under this assumption, most bearing capacity theories (e.g., Prandtl, 1921; Meyerhof, 1951, 1963) assume that the failure slip surface takes on a logarithmic spiral shape to give

$$N_c = \frac{e^{\pi \tan\phi} \tan^2\left(\frac{\pi}{4}+\frac{\phi}{2}\right) - 1}{\tan\phi} \tag{19}$$

The theory is derived for the general case of a $c-\phi$ soil. One can always set $\phi = 0$ to obtain results for an undrained clay.

Consistent with the theoretical results presented by Fenton et al. (2008), this example will concentrate on the design of a strip footing. In this case, the characteristic ultimate geotechnical resistance $\hat{R}_u$ becomes

$$\hat{R}_u = B\hat{q}_u \tag{20}$$

where *B* is the footing width and $\hat{R}_u$ has units of load per unit length out-of-plane, that is, in the direction of the strip footing. The characteristic ultimate bearing stress $\hat{q}_u$ is defined by

$$\hat{q}_u = \hat{c}\hat{N}_c \tag{21}$$

where the characteristic N_c factor is determined using the characteristic friction angle in Eq. 19,

$$\hat{N}_c = \frac{e^{\pi \tan\hat{\phi}} \tan^2\left(\frac{\pi}{4}+\frac{\hat{\phi}}{2}\right) - 1}{\tan\hat{\phi}} \tag{22}$$

For the strip footing and just the dead and live load combination, the LRFD equation becomes

$$\varphi_g B\hat{q}_u = \left[\alpha_L \hat{F}_L + \alpha_D \hat{F}_D\right] \quad \Rightarrow \quad B = \frac{\left[\alpha_L \hat{F}_L + \alpha_D \hat{F}_D\right]}{\varphi_g \hat{q}_u} \tag{23}$$

To determine the resistance factor φ_g required to achieve a certain acceptable reliability of the constructed footing, it is necessary to estimate the probability of bearing capacity failure of a footing designed using Eq. 23. Once the probability of failure p_f for a certain design using a specific value for φ_g is known, this probability can be compared to the maximum acceptable failure probability p_m . If p_f exceeds p_m, then the resistance factor must be reduced and the footing redesigned. Similarly, if p_f is less than p_m, then the design is overconservative and the value of φ_g can be increased. Using either simulation or theory, design curves can then be developed from which the value of φ_g required to achieve a maximum acceptable failure probability can be determined.

Figure 11 shows the resistance factors required for the case where the soil is sampled at a distance of $r = 4.5$ m from the footing centerline for the target failure probability, $p_m = 0.001$. In the figure, v_c is the coefficient of variation of cohesion.

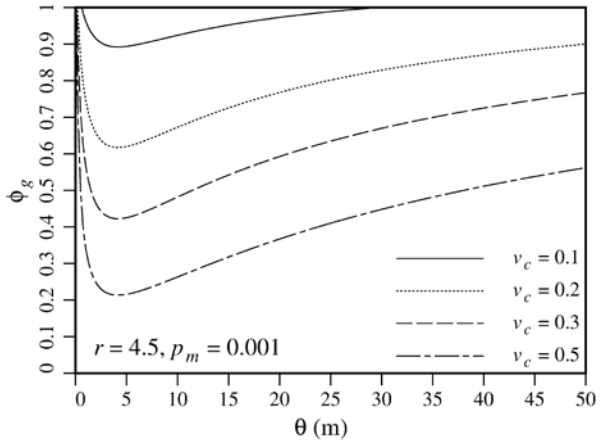

Fig. 11. Resistance factors required to achieve acceptable failure probability p_m when soil is sampled at $r = 4.5$ m from footing centerline and $p_m = 0.001$.

The worst-case correlation length is evidently about 5 m. This worst-case correlation length is of the same magnitude as the mean footing width which can be explained as follows: If the random soil fields are stationary, then soil samples yield perfect information, regardless of their location, if the correlation length is either zero (assuming soil sampling involves some local averaging) or infinity. When the information is perfect, the probability of a bearing capacity failure goes to zero and $\varphi_g \to 1.0$ (or possibly greater than 1.0 to compensate for the load bias factors). When the correlation length is zero, the soil sample will consist of an infinite number of independent "observations" whose average is equal to the true mean (or true median, if the average is a geometric average). Since the footing also averages the soil properties, the footing 'sees' the same true mean (or true median) value predicted by the soil sample. When the correlation length goes to infinity, the soil becomes

uniform, having the same value everywhere. In this case, any soil sample also perfectly predicts conditions under the footing.

At intermediate correlation lengths soil samples become imperfect estimators of conditions under the footing, and so the probability of bearing capacity failure increases, or equivalently, the required resistance factor decreases. Thus, the minimum required resistance factor will occur at some correlation length between 0 and infinity. The precise value depends on the geometric characteristics of the problem under consideration, such as the footing width, depth to bedrock, length of soil sample, and/or the distance to the sample point.

RELIABILITY-BASED GEOTECHNICAL DESIGN CODE DEVELOPMENT

A report by Littlejohn (1991) entitled *Inadequate Site Investigation* made the statement "You pay for a site investigation whether you have one or not", which quite clearly points out the cost of geotechnical uncertainty. Essentially, if one does not bother with a sufficient geotechnical investigation, one either pays the cost immediately by requiring a more conservative design or is going to pay the cost later due to some level of performance failure of the designed system. Since performance failure at some future date is generally very expensive, there is a real desire in the geotechnical community to account for the level of uncertainty during the design phase. That is, the level of site and modeling understanding should be balanced against the conservatism of the design – the greater the understanding, the less conservative, and thus less expensive, the design. Site understanding refers to how well the ground providing the geotechnical resistance is known and model understanding means the degree of confidence that a designer has in the (usually mathematical) model used to predict the geotechnical resistance.

To provide for designs that account for degree of understanding, it makes sense to have a resistance factor which is adjusted as a function of site and model understanding. There are at least two advantages to such an approach: 1) overall safety can be maintained at a common target maximum failure probability, and 2) the direct economic advantage related to increasing site and model understanding can be demonstrated. For example, the pre-2014 Canadian design codes specify a single resistance factor for bearing capacity design (0.5). It doesn't matter how confident one is in one's prediction of the bearing capacity of a foundation, the same resistance factor must be used. Thus, there is no direct advantage to improving the geotechnical response prediction. If only a single resistance factor can be used, one might as well spend the least amount of time one can on the site investigation and modeling.

The resulting desire for a resistance factor which depends on site and model understanding is not new. The Australian Standard for *Bridge Design, Part 3: Foundations and Soil-Supporting Structures* (AS 5100.3, Standards Australia, 2004) provides a range in "geotechnical strength reduction factors" accompanied by guidance as to which end of the scale should be used. For example, AS 5100.3 suggests that the lower end of the resistance factor range (more conservative) should be used for limited site investigations, simple methods of calculation, severe failure consequences, and so on. It is of interest to note that the Australian Standard

recommendations for the resistance factor considers both site and model understanding along with failure consequence in their single factor.

As is well known, the overall safety level of any design should depend on at least three things: 1) uncertainty in the loads, 2) uncertainty in the resistance, and 3) the severity of the failure consequences. These three items are all usually deemed to be independent of one another and in most modern codes are thus treated separately. Uncertainties in the loads are handled by load and load combination factors, failure consequences are handled by applying a multiplicative importance factor to the more site-specific and highly uncertain loads (e.g. earthquake, snow, and wind), and uncertainties in resistance are handled by material specific resistance factors (e.g. φ_c for concrete, φ_s for steel, etc).

Because the ground is also site specific and highly uncertain, it makes sense to apply a partial safety factor that depends on both the resistance uncertainty and consequence of failure of the ground. This would be analogous to how wind load, for example, in the NBCC (NRC, 2010) has both a load factor associated with wind speed uncertainty as well as an importance factor associated with failure consequences. Figure 12 illustrates the basic idea, where the overall partial factor applied to the geotechnical resistance varies with both site and model understanding and failure consequence level. The numbers in the figure are relative to the default central partial factor (i.e. relative to 1.0) and it is assumed that current geotechnical design approaches in Canada lead to typical or default levels of site and model understanding so that, for typical failure consequence geotechnical systems, the central value is what is currently used in design. From this value, increased site investigation and/or modeling effort leads to higher understanding and a higher geotechnical resistance factor (and so a more economical design). Similarly, for geotechnical systems with high failure consequences, e.g. failure of the foundation of a major multi-lane highway bridge in a large city, the resistance factor is decreased to provide a decreased maximum acceptable failure probability. Of particular note in Figure 12 is the fact that if a geotechnical system with high failure consequence is designed with low site and model understanding, the designer is penalized by a low geotechnical resistance factor.

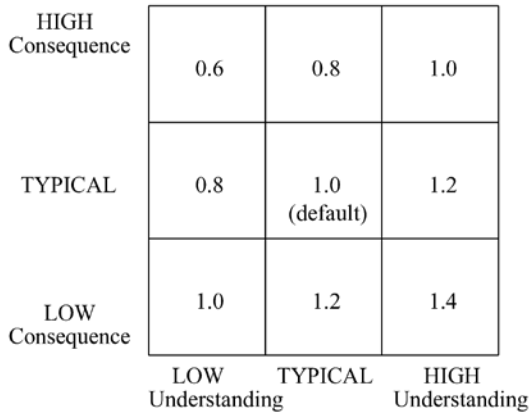

Fig. 12. Floating partial safety factor, relative to the default, applied to geotechnical resistance (numbers are for illustration only).

Rather than introducing a 3 x 3 matrix of resistance factors for each limit state, the multiplicative approach taken in structural engineering (where the load is multiplied by both a load factor and an importance factor) is adopted for geotechnical resistance as well in the 2014 CHBDC (CSA, 2014).

In other words, the overall safety factor applied to geotechnical resistance is broken into two parts;

1. a resistance factor, φ_{gu} or φ_{gs}, which accounts for resistance uncertainty. This factor basically aims to achieve a target maximum acceptable failure probability equal to that used currently for geotechnical designs for typical failure consequences (e.g. lifetime failure probability of 1/5,000 or less). The subscript *g* refers to 'geotechnical' (or 'ground'), while the subscripts *u* and *s* refer to ultimate and serviceability limit states, respectively.
2. a consequence factor, Ψ, which accounts for failure consequences. Essentially, $\Psi > 1$ if failure consequences are low and $\Psi < 1$ if failure consequence exceeds that of typical geotechnical systems. The basic idea of the consequence factor is to adjust the maximum acceptable failure probability of the design down (e.g. to 1/10,000) for high failure consequences, or up (e.g. to 1/1,000) for low failure consequences.

The geotechnical design would then proceed by ensuring that the factored characteristic geotechnical resistance at least equals the effect of factored characteristic loads. For example, for ultimate limit states, this means that in the 2014 CHBDC the geotechnical design will need to satisfy an equation of the form

$$\Psi \varphi_{gu} \hat{R} \geq \sum_i \alpha_{ui} \hat{F}_{ui} \tag{24}$$

which is almost identical to Eq. 1, with the exception that the overall geotechnical resistance factor is expressed as the product of the consequence factor, Ψ, and the ultimate geotechnical resistance factor, φ_{gu}, and the loads and load factors appearing on the right-hand-side are also those specific for the ultimate limit state under consideration (and, hence, the subscript *u*). An entirely similar equation must be satisfied for serviceability limit states, with the subscript *u* replaced by *s*. The serviceability geotechnical resistance factors, φ_{gs}, will be closer to 1.0 than φ_{gu}, since serviceability limit states can have larger maximum acceptable probabilities of occurrence.

The geotechnical resistance factor, φ_{gu} or φ_{gs}, depends on the degree of site and prediction model understanding. Three levels are considered in the 2014 CHBDC;

- *High understanding*: extensive project-specific investigation procedures and/or knowledge are combined with prediction models of demonstrated quality to achieve a high level of confidence with performance predictions,
- *Typical understanding*: typical project-specific investigation procedures and/or knowledge are combined with conventional prediction models to achieve a typical level of confidence with performance predictions,
- *Low understanding*: limited representative information (e.g. previous experience, extrapolation from nearby and/or similar sites, etc.) combined with conventional prediction models to achieve a lower level of confidence with performance predictions.

The resulting table for ULS and SLS geotechnical resistance factors appearing in the 2014 CHBDC is shown in Table 1. How the geotechnical resistance factor values appearing in Table 1 were obtained is explained in the following sections on calibration.

The consequence factor, Ψ, appearing in Eq. 24, adjusts the maximum acceptable failure probability of the geotechnical system being designed to a value which is appropriate for the magnitude of the failure consequences. Three failure consequence levels are considered in the 2014 CHBDC;

- *High consequence*: the foundations and/or geotechnical systems are designed for applications, including bridges, essential to post-disaster recovery (e.g. lifeline) and/or having large societal or economic impacts.
- *Typical consequence*: the foundations and/or geotechnical systems are designed for applications, including bridges, carrying medium to large volumes of traffic and/or having potential impacts on alternative transportation corridors or structures.
- *Low consequence*: the foundations and/or geotechnical systems are designed for applications carrying low volumes of traffic and having limited impacts on alternative transportation corridors.

Table 1. Some of the geotechnical resistance factors for ULS and SLS appearing in Table 6.2 of the 2014 CHBDC (numbers are for illustration only).

Application	Limit State	Test Method/Model	Degree of understanding		
			Low	Typical	High
Shallow foundations	Bearing, φ_{gu}	Analysis	0.45	0.50	0.60
		Scale model test	0.50	0.55	0.65
	Sliding frictional, φ_{gu}	Analysis	0.70	0.80	0.90
		Scale model test	0.75	0.85	0.95
	Sliding cohesive, φ_{gu}	Analysis	0.55	0.60	0.65
		Scale model test	0.60	0.65	0.70
	Passive resistance, φ_{gu}	Analysis	0.40	0.50	0.55
	Settlement or lateral movement, φ_{gs}	Analysis	0.7	0.8	0.9
		Scale model test	0.8	0.9	1.0

Table 2. ULS and SLS consequence factors, Ψ, appearing in Table 6.1 of the 2014 CHBDC.

Consequence level	Consequence factor, Ψ
High	0.9
Typical	1.0
Low	1.15

The consequence factors specified in the 2014 CHBDC for the three consequence levels are shown in Table 2. This table is very similar to Table B-3 in Eurocode 0 (British Standard BS EN 1990, 2002) which specifies three multiplicative factors, 0.9, 1.0, and 1.1, to be applied to loads (actions) for low, medium, and high failure consequences, respectively (these factors are approximately the inverse of the factors seen in Table 2 because they appear on the load side of the LRFD equation). In other

words, the concept of shifting the target failure probability to account for severity of failure consequences is not new, although the application of the consequence factor to the resistance side, rather than the load side, of the LRFD equation appears to be new.

Calibration of Geotechnical Resistance Factors

The geotechnical resistance factor calibration must start with a review of the factors currently used in Canadian geotechnical design codes, as well as those used in other codes from around the world. Table 3 is a small subset of a much more extensive table that was prepared to compare the load and geotechnical resistance factors between a variety of codes, reports, and manuals from various jurisdictions (Fenton et al., 2016). In the calibration process, Table 3, and its more extensive counterpart, is used to suggest the 'best' currently acceptable estimates of 'typical' resistance factors. These are the φ_{gu} factors that have been found to lead to societally acceptable failure probabilities under current design practice. The factor $R_{D/L}$ is the dead to live load ratio which was assumed in the code calibration process.

Table 3. Table of design factors used for geotechnical design as specified in various codes of practice.

Source	$R_{D/L}$	α_L	α_D	α_T	φ_{gu}	F_s
NBCC (2010)	3.0	1.5	1.25	1.31	0.5	2.88
CHBDC (2006)	3.0	1.7	1.2	1.33	0.5	2.92
CFEM (1992)	3.0	1.5	1.25	1.31	• Cohesion (foundations) 0.5 • Cohesion (stability, earth pressure) 0.65 • Friction 0.8	2.88 2.22 1.80
AASHTO (2002)	3.7	2.86	1.3	1.65	a. Sand • Semi-empirical procedure using SPT data 0.45 • Semi-empirical procedure using CPT data 0.55 • Rational method • using φ_f estimated from SPT data 0.35 • using φ_f estimated from CPT data 0.45 b. Clay • Semi-empirical procedure using CPT 0.50 • Rational method • using shear strength measured in lab tests0.60 • using shear strength measured in field vane tests 0.60 • using shear strength estimated from CPT data 0.50 c. Rock • Semi-empirical procedure(Carter and Kulhawy) 0.60	 4.04 3.30 5.19 4.04 3.63 3.03 3.03 3.63 3.03
AASHTO (2007 and 2012)	3.7	1.75	1.25	1.35	• Theoretical method(Munfakh et al., 2001), in clay 0.50 • Theoretical method(Munfakh et al., 2001), in sand, using CPT 0.50 • Theoretical method(Munfakh et al., 2001), in sand, using SPT 0.45 • Semi-empirical methods(Meyerhof, 1957), all soils 0.45 • Footings on rock 0.45 • Plate Load Test 0.55	2.97 2.97 3.30 3.30 3.30 2.70

For example, Table 3 suggests that the geotechnical resistance factor for the bearing capacity of a shallow foundation ranges from about 0.35 to about 0.60,

depending on the confidence of the geotechnical performance prediction, with a typical value of about 0.50. This 'typical' value forms the starting point for the 'typical understanding' values appearing in Table 1. The range suggested in Table 3 provides some insight into the range that might be appropriate for the three levels of site and model understanding considered in the 2014 CHBDC.

Once the typical geotechnical resistance factor values have been established, the next two steps are to look at how these values should change as a result of changes in the level of site and model understanding, and how the consequence factor should be set to reflect changes in the failure consequence severity.

The question of how the geotechnical resistance factor should be adjusted as the level of site and model understanding changes brings up the question of how the reliability of a geotechnical design can be estimated in the first place, for any given level of site and model understanding. The approach used here is essentially to use Monte Carlo simulations, modeling the ground as a spatially varying random field, and carry out a *virtual* site investigation, design, and construction of the geotechnical system. The geotechnical system is then subjected to random maximum lifetime loads and checked to see if the particular limit state under investigation is exceeded. If so, a failure is recorded and the process is repeated. The failure probability of the design is then estimated as the number of failures divided by the number of trials – if the failure probability is too high, the design factors are suitably adjusted, and so on. The detailed steps are as follows;

1. for a particular geotechnical system (e.g., shallow foundation) and limit state (e.g., bearing capacity), choose a geotechnical resistance factor to be used in the design,
2. simulate a random field of ground properties, having a specified variance and correlation structure,
3. *virtually* sample the ground at some location to obtain 'observations' of the ground properties. The distance between the sample and the geotechnical system acts as a proxy for site and model understanding – the farther the sample is from the geotechnical system, the more the uncertainty about the system performance (decreased site and model understanding),
4. design the geotechnical system using the characteristic geotechnical parameters determined from the sample taken in step 3. The definition of 'characteristic' depends on the design code being used. For example, in Europe, the characteristic values would be a lower 5-percentile. In North America, a 'cautious estimate of the mean' is probably a more common definition, as discussed previously. In most of the calibration exercises undertaken for the CHBDC, the characteristic values were taken as the geometric average of the sampled 'observations'. The geometric average is always at least slightly lower (more so for higher variability) than the arithmetic average, and so can be viewed as a 'cautious estimate of the mean',
5. *virtually construct* the geotechnical system according to the design in the previous step and place it on (or in) the random field generated in step 2,
6. employ a sophisticated numerical model (e.g., the finite element method) to determine if the geotechnical system exceeds the limit state being designed against (this is a *failure*),
7. repeat from step 2 a large number of times, recording the number of failures.

8. the probability of failure is then estimated as the number of failures divided by the number of trials. If this probability is too high, the geotechnical resistance factor needs to be decreased, if too low, the geotechnical resistance factor should be increased. After adjusting the geotechnical resistance factor appropriately, the entire procedure can be repeated from step 1 using the new geotechnical resistance factor.

For example, Figure 11 presents the theoretically determined geotechnical resistance factors for the bearing capacity design of a shallow foundation where the target maximum lifetime failure probability is $p_m = 0.001$, corresponding to a reliability index of about $\beta = 3.1$. If the worst case resistance factor for a reasonable coefficient of variation of the ground shear strength, $v_c = 0.3$, is examined, it can be seen from Figure 11 that the typical `understanding' (assumed to be $r = 4.5$ m) geotechnical resistance factor is about 0.45. For 'high' understanding ($r = 0$ m), a similar plot (not shown) suggests a geotechnical resistance factor of about 0.65 when $v_c = 0.3$. At the other extreme, for `low' understanding ($r = 9$ m), the theory suggests a geotechnical resistance factor of about 0.4. These theoretical results seem to be in reasonable agreement with the range suggested by other codes.

Calibration of the Geotechnical Consequence Factor

The basic idea of the consequence factor is to adjust the target maximum lifetime failure probability, p_m, to a value which is appropriate for the failure consequences. For example, if the geotechnical system supports a storage warehouse which is rarely visited, the failure consequences are slight and its failure probability should be higher than that for typically supported structures. If, on the other hand, the geotechnical system supports a hospital or lifeline bridge, then the failure probability should probably be lower than that for typically supported structures.

Using the random finite element method (RFEM), which includes modeling of the ground's spatial variability, Fenton et al. (2011) produced Figure 13, which illustrates how the probability of bearing capacity failure changes with the consequence factor for the typical site understanding case ($r = 4.5$ m), correlation length $\theta = 6$ m, and using a design resistance factor of $\varphi_g = 0.5$. It can be seen that fairly small changes in the consequence factor, Ψ, can make large differences in the failure probability, p_f. As expected, the soil variability (v_c), also has a very significant effect on p_f. The two horizontal lines in Figure 13 bound failure consequence acceptable probabilities, $p_m = 1/1000$ to $p_m = 1/10,000$.

To illustrate how Figure 13 works, one additional curve was produced for $v_c = 0.23$. When $\Psi = 1.0$ (typical consequence), the $v_c = 0.23$ case has failure probability $p_f \approx 2\times10^{-4} = 1/5000$, which is the maximum acceptable failure probability for typical consequences ($\beta = 3.5$). To adjust this case to have failure probability $p_f = 1\times10^{-4} = 1/10,000$ (high consequence), a consequence factor of about $\Psi = 0.93$ should be used – the required Ψ value occurs where the $v_c = 0.23$ curve intersects the horizontal $p_m = 1/10,000$ line. The recommended consequence

factor for this case has been rounded down to 0.90, as discussed shortly. Similarly, to adjust the $v_c = 0.23$ case for a low consequence design ($p_m = 1/1000$), the consequence factor is obtained at the intersection of the $v_c = 0.23$ curve and the upper horizontal line. This occurs at about $\Psi = 1.13$ (which will be rounded to $\Psi = 1.15$ shortly).

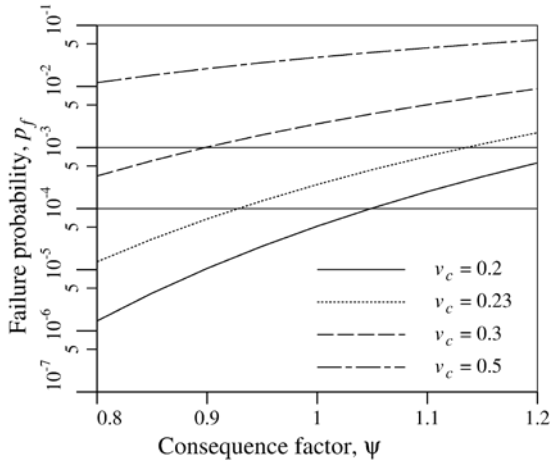

Fig. 13. Failure probability versus consequence factor for $\theta = 6$ m, $r = 4.5$ m, and $\varphi_g = 0.5$.

For the high consequence level case, Fenton et al. (2011) found that the range in Ψ values (for *r* ranging from 0 to 9 m, θ ranging from 0 to 50 m, and v_c ranging from 0.1 to 0.5) is from 0.91 to 0.976. For the low consequence level case, Fenton et al. (2011) found the range in Ψ to be from 1.06 to 1.28. However, when the resistance factor is held fixed with respect to θ and v_c, the range in the consequence factor is increased considerably. For example, when the resistance factor is fixed at $\varphi_g = 0.40$ for low site and model understanding ($r = 9$ m), the consequence factor for the high consequence case varies from 0.38 to 2.47. The equivalent range for the low consequence case is 0.52 to 2.69. More details can be found in Fenton et al. (2011).

Figure 14 shows how the consequence factor varies with correlation length, θ, for $r = 4.5$ m (typical understanding) and $\varphi_g = 0.5$. The upper Figure 14a is for the high failure consequence case ($p_m = 1/10,000$) and the lower Figure 14b is for the low failure consequence case ($p_m = 1/1000$). Considering Figure 14a, the task is to choose a factor for the high consequence case which is sufficiently conservative and yet not excessively so. Reducing the consequence factor results in more conservative designs (lower failure probability). A solid horizontal line has been drawn across the plot at $\Psi = 0.9$ and it can be seen that this value is conservative for all $v_c \le 0.25$ (approximately), in that the curves for $v_c = 0.1$ and 0.2 lie entirely above $\Psi = 0.9$. What this means is that if v_c is known to be 0.1, for example, then using $\Psi = 0.9$ in the design would result in a failure probability well below the target of $p_m = 1/10,000$. On the other hand, if v_c is not clearly known, then $\Psi = 0.9$ is reasonably conservative for all but sites with large soil variability (e.g. $v_c \ge 0.3$). If

site investigation is sufficient to keep the residual variability below this level, then $\Psi = 0.9$ is a reasonable design value for the high failure consequence case which will almost always lead to a failure probability well below $p_m = 1/10,000$ ($\beta = 3.7$).

A similar argument can be applied to Figure 14b for the low consequence case, where a solid line at $\Psi = 1.15$ has been drawn across the plot. It can be seen that this value is not quite as conservative as the high consequence factor (selected above) in that the $v_c = 0.2$ curve comes somewhat closer to $\Psi = 1.15$. The authors feel, however, that conservatism is not quite as important for the low failure consequence case, and so selected the somewhat less conservative value of 1.15.

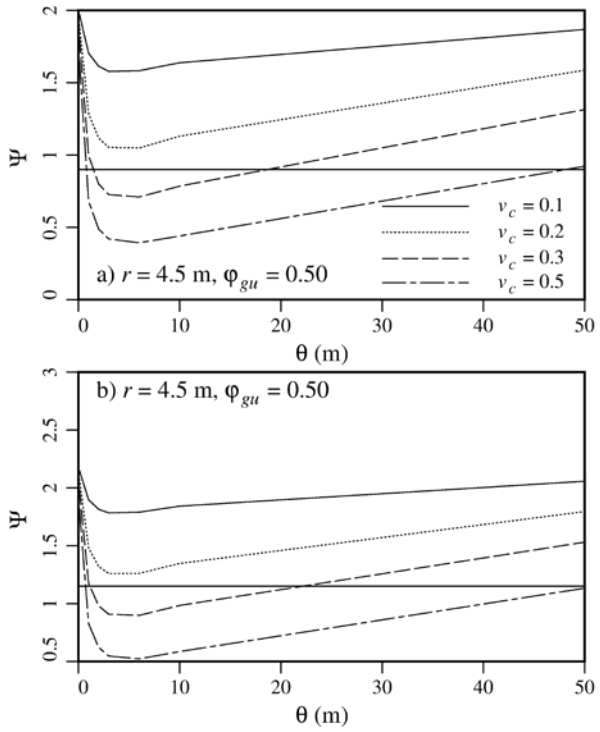

Fig. 14. Consequence factor versus correlation length for $r = 4.5$ m and $\varphi_{gu} = 0.5$ at high consequence level ($p_m = 1/10,000$) in (a), where $\Psi = 0.9$ is proposed, and at low consequence level ($p_m = 1/1000$) in (b), where $\Psi = 1.15$ is proposed.

Research into the consequence values for deep foundation design (Naghibi et al., 2013) yields similar consequence factors for both ULS and SLS design. Thus, it appears that the consequence factors selected for the 2014 CHBDC are reasonably appropriate for other limit states of geotechnical system.

FUTURE DIRECTIONS

Previous sections illustrated how spatial variability models can be used both to more realistically represent the ground and its failure mechanism as well as to serve as a mathematical proxy for site understanding. Although computers are becoming

fast enough to bring random field models into the design office, there are still a number of impediments to their widespread use;

1. While the mean values of the ground properties may be reasonably well known, estimating their variances requires significantly more samples of the ground throughout the site. In many cases, such intensive sampling will not have been done, and so variance estimates commonly come from the literature. The entire issue of specifying the variance of the random ground model is complicated by a number of factors;
   a. it is really the uncertainty between sample locations (e.g., between CPT soundings) whose variance needs to be estimated, and this variance will be a function of distance from the sampling points. This level of variance characterization is best accomplished by using conditioned random fields which take on the sample values at the sample locations (possibly with an added measurement error component) and which are increasingly random with distance from the sample locations.
   b. the specified ground property variance is often scale dependent. For example, if one is trying to specify the variance of the ground's permeability field, it is necessary to consider the scale at which the model is to be applied. At the micro-scale in, say, a sandy soil, one point could be within a particle of granite, having virtually zero permeability, and in a void at another point, having virtually infinite permeability. At that scale, the permeability variance will be extremely large. In practice, however, interest is usually in the permeability of some representative volume representing the property averaged over that representative volume. This variance will be much reduced from the "point" variance – how much reduced depends on the size of the representative volume. Similar scale, or *local averaging*, effects are present in most geotechnical problems. For example, the stability failure of a slope involves the average shear strength along the weakest failure path through the soil. The ultimate capacity of a pile involves the average of the shear strength over the pile surface along with the bearing capacity at the pile tip, which itself involves an average of ground shear strength below and around the pile.
2. At most sites, it is unlikely that sufficient investigation will have been done in all three spatial directions to clearly estimate the directional correlation lengths. If the correlation lengths are unknown, the design can conservatively be carried out at the "worst case" correlation length, which is typically approximately equal to some dimension of the geotechnical system being designed (e.g., distance to nearest sample, width of the footing, etc). However, using the worst case correlation lengths can incur a significant design penalty. For example, Figure 11 suggests that if $v_c = 0.3$ and the worst case correlation length is used, the geotechnical resistance factor is $\varphi_g \approx 0.45$. On the other hand, if the correlation is actually known to be about 20 m, the resistance factor required in design rises to about 0.60 – a fairly significant increase.

As mentioned previously, the current edition of the Canadian Highway Bridge Design Code includes a geotechnical consequence factor to account for the severity of failure consequences. This idea is not new. The so-called importance factor, which is applied to site specific and highly variable loads (e.g., snow, wind, and earthquake) does

exactly the same thing – modifies the target system reliability depending on failure consequences.

Adjusting the target reliability, depending on failure consequence severity, is one way of accomplishing a risk-based design. The perhaps more precise way is to actually perform a risk assessment of the design, where risk is defined as the product of failure probability times the cost of failure and choose the design having the lowest risk. Such an approach is not commonly taken in civil engineering for the simple fact that failure often involves loss of life and assigning a cost to loss of human life (or any lives, for that matter) is a difficult and sensitive issue.

Nevertheless, there are many geotechnical design issues that do not involve loss of lives which would definitely benefit from a risk assessment (cost-benefit) approach to design. Serviceability limit states, for example, have fairly well defined limits (e.g., excessive settlement) and entering such a state will have a cost which can be estimated (e.g., the cost of improving/stiffening a foundation).

For example, Figures 8 and 10 can be used to perform a risk assessed design of a shallow foundation against entering a serviceability limit state. For a fixed failure probability, $p_f = 0.05,$ the various sampling schemes shown in Figure 8 result in different resistance factors, which in turn directly influence the cost of constructing the foundation. This construction cost can then be balanced against the cost of sampling and the optimum sampling scheme determined. For example, assuming an unsophisticated sample average is used to estimate the soil properties, it appears from Figure 10 that the best sampling scheme (highest resistance factor, lowest construction cost) is #5, where a single sample is taken directly under the footing. It must be remembered, however, that in order to achieve maximum construction savings, a sample would have to be obtained under every footing, which is often not practical. The system level cost-benefit analysis is a relatively straightforward extension of a single footing cost-benefit analysis.

CONCLUSIONS

Geotechnical design has advanced considerably in the last decade or so. Reliability-based concepts are common in modern geotechnical design codes – most codes now have acknowledged target reliabilities and have implemented load and resistance factors to achieve those targets.

There is still much to be improved. For example, the direct consideration of spatial variability in geotechnical response predictions and design methodologies is still relatively rare. Modern computers are fast enough that random field representations of the ground, combined with advanced finite element models, in a Monte Carlo simulation framework can be done relatively easily. Such analyses would augment more traditional geotechnical engineering, allowing for improved design decisions to be made in the face of uncertainty.

Design codes which allow the resistance factors to change depending on the level of uncertainty (e.g., the *Canadian Highway Bridge Design Code* and the Australian Standard 5100 *Bridge Design*) are indirectly using a risk assessment to carry out the design. Increased site and model understanding generally involves additional cost, but these codes allow this cost to be offset by construction savings through an increased

resistance factor. The next logical step is to routinely perform such risk assessments for individual designs in order to optimize overall savings while maintaining system reliability.

Design codes of the future will increasingly allow, and promote, flexibility in the design process if rigorous probabilistic and risk (cost-benefit) assessments have been performed.

REFERENCES

British Standard BS EN 1990:2002 (2002). *Eurocode – Basis of Structural Design*, CEN (European Committee for Standardization), Brussels.

Canadian Standards Association (2014). *Canadian Highway Bridge Design Code*, CAN/CSA-S6-14, Mississauga, Ontario.

Fenton, G.A. and Griffiths, D.V. (2008). *Risk Assessment in Geotechnical Engineering*, John Wiley & Sons, Hoboken, New Jersey.

Fenton, G.A., Zhang, X.Y. and Griffiths, D.V. (2007). Reliability of shallow foundations designed against bearing failure using LRFD, *Georisk*, **1**(4), 202–215.

Fenton, G.A. and Griffiths, D.V. (2002). Probabilistic foundation settlement on spatially random soil, ASCE *J. Geotech. Geoenv. Engrg.*, **128**(5), 381–390.

Fenton, G.A., Griffiths, D.V. and Ojomo, O.O. (2011). Consequence factors in the ultimate limit state design of shallow foundations, *Can. Geotech. J.*, **48**(2), 265–279.

Fenton, G.A. and Vanmarcke, E.H. (1990). Simulation of random fields via Local Average Subdivision, ASCE *J. Engrg. Mech.*, **116**(8), 1733–1749.

Fenton, G.A., Griffiths, D.V. and Cavers, W. (2005). Resistance factors for settlement design, *Can. Geotech. J.*, **42**(5), 1422–1436.

Fenton, G.A., Griffiths, D.V. and Zhang, X.Y. (2008). Load and resistance factor design of shallow foundations against bearing failure, *Can. Geotech. J.*, **45**(11), 1556–1571.

Fenton, G.A., Naghibi, F., Dundas, D., Bathurst, R.J. and Griffiths, D.V. (2016). Reliability-based geotechnical design in 2014 Canadian Highway Bridge Design Code, *Can. Geotech. J.*, **53**(2), 236–251.

Griffiths, D.V. and Fenton, G.A. (2009). Probabilistic settlement analysis by stochastic and random finite element methods, *ASCE J. Geotech. Geoenv Engrg.*, **135**(11), 1629–1637.

Li, C.-C. and Kiureghian, A. (1993). "Optimal discretization of random fields," ASCE J. Engrg. Mech., **119**(6), 1136–1154.

Littlejohn, G.S., Arber, N.R., Craig, C. and Forde, M.C. (1991). *Inadequate Site Investigation*, Institution of Civil Engineers, Thomas Telford, London, UK.

Matthies, H.G., Brenner, C.E., Bucher, C.G. and Soares, C.G. (1997). Uncertainties in probabilistic numerical analysis of structures and solids -- stochastic finite elements, *Structural Safety*, **19**(3), 283–336.

Meyerhof, G. G. (1963). Some recent research on the bearing capacity of foundations, *Can. Geotech. J.*, **1**(1), 16–26.

Meyerhof, G. G. (1951). The ultimate bearing capacity of foundations, *Géotechnique*, **2**(4), 301–332.

Naghibi, F., Fenton, G.A. and Griffiths, D.V. (2013). Resistance and consequence factor calibration for deep foundations, *Proceedings of the 66th Canadian Geotechnical Conference*, Canadian Geotechnical Society, Sep 29 – Oct 2, Montreal, Paper #224.

National Research Council (2010). *National Building Code of Canada*, 13th Ed., National Research Council of Canada, Ottawa.

Prandtl, L. (1921). Uber die Eindringungsfestigkeit (Harte) plastischer Baustoffe und die Festigkeit von Schneiden, *Zeitschrift fur angewandte Mathematik und Mechanik*, **1**(1), 15–20.

Smith, I.M., and Griffiths, D.V. (2004). *Programming the Finite Element Method*, J. Wiley & Sons, 4th ed., Hoboken, NJ.

Sokolovski, V.V. (1965). *Statics of Granular Media*, 270 pages, Pergamon Press, London, UK.

Standards Australia (2004). *Bridge Design, Part 3: Foundations and Soil-Supporting Structures*, Australian Standard AS 5100.3--2004, Sydney, Australia.

Terzaghi, K. (1943). *Theoretical Soil Mechanics*, John Wiley & Sons, New York, NY.

NOTATION

a, b	lower and upper bounds of tanh distribution
B	footing width
c	cohesion or threshold
$\hat{c}$	characteristic cohesion = geometric average of observations
c_i^o	*i*'th observed soil cohesion measurement
E	elastic modulus field
$\hat{E}$	estimate of effective elastic modulus, derived from soil samples
E_{eff}	effective uniform elastic modulus giving same settlement as actual settlement
E_g	local geometric average of elastic modulus field
f_δ	probability density function of foundation settlement
F	actual (random) load effect
$\hat{F}$	characteristic (nominal) load effect
$\hat{F}_D$	characteristic dead load
$\hat{F}_L$	characteristic live load
F_s	effective total factor of safety, accounting for resistance, load, and bias factors
G	standard normal (Gaussian) random field
G_e	the random field G averaged over each finite element
H	depth to bedrock
L_x	random field dimension, similarly for L_y and L_z
m	location parameter of tanh distribution or sample mean of subscripted variable or number of samples
n	number of simulations or points
N_c	bearing capacity factor
$\hat{N}_c$	characteristic bearing capacity factor based on characteristic soil properties
p_f	failure probability
p_m	maximum tolerable failure probability
q_c	CPT tip resistance
q_u	ultimate bearing resistance stress
$\hat{q}_u$	characteristic ultimate bearing resistance stress
r	distance from foundation to sample location
$R_{D/L}$	ratio of dead to live load

$\hat{R}$ characteristic geotechnical resistance (based on characteristic soil properties)

$\hat{R}_u$ ultimate characteristic geotechnical resistance

s scale parameter of tanh distribution or sample standard deviation of subscripted variable

u_1 influence factor for settlement prediction

v coefficient of variation $(=\sigma/\mu)$ of subscripted variable

x variable or horizontal component of spatial position

$\underset{\sim}{x}$ spatial coordinate or position

X random variable or random field

X_c random field conditioned on a set of observations

X_e random field formed by averaging X over each finite element

X_k best linear unbiased estimate field based on observations

X_s best linear unbiased estimate field based on X_u at observation locations

X_u unconditional random field

y horizontal component of spatial position

z vertical component of spatial position

α load factor (α_i = i'th load factor)

α_D dead load factor

α_L live load factor

α_T effective total load factor

β reliability index

γ variance reduction function

δ foundation settlement

δ_{det} deterministic foundation settlement using mean elastic modulus

δ_{max} maximum acceptable foundation settlement

θ correlation length of subscripted random field (if present), having directional components $\theta_x, \theta_y,$ and θ_z

μ mean, or mean of the subscripted variable

ϕ angle of internal friction

$\hat{\phi}$ characteristic angle of internal friction = arithmetic average of friction angle observations

ϕ_i^o i'th observed friction angle measurement

φ resistance factor (subscript g for geotechnical, u for ULS, s for SLS)

φ_c resistance factor for concrete

φ_s resistance factor for steel

ρ correlation coefficient

σ standard deviation, or standard deviation of the subscripted variable

$\underset{\sim}{\tau}$ spatial lag vector, having directional components $\tau_x, \tau_y,$ and τ_z

Ψ consequence factor