Phoon, K.-K. & Kulhawy, F. H. (2005). *Géotechnique* **55**, No. 1, 45–54

Characterisation of model uncertainties for laterally loaded rigid drilled shafts

K.-K. PHOON* and F. H. KULHAWY†

This paper presents a critical evaluation of model factors for laterally loaded rigid drilled shafts (bored piles). Both the lateral or moment limit and hyperbolic capacity are considered to make explicit the dependence of model factors on the criterion for interpreting 'capacity' from load test data. Although the hyperbolic capacity may be closest to the theoretical ultimate state or upper bound, results indicate that it generally does not produce a mean model factor of 1. When the measured capacity is interpreted consistently from load test data, the coefficient of variation (COV) appears to remain relatively constant between 30% and 40%. The range of the mean bias for the lateral or moment limit is 0·67 to 1·49, whereas that of the hyperbolic capacity is 0·98 to 2·28. Based on available data, a log-normal probability model appears adequate for reliability analysis. Laboratory-scale load tests conducted in uniform soil deposits prepared under controlled laboratory conditions are ideal for establishing benchmarks on the probable magnitude of uncertainty arising from model idealisations alone. However, the limited range of geometric and geotechnical parameters in a laboratory load test database may not produce a representative mean model factor. A field load test database typically contains more diverse geometric and geotechnical parameters, but it entails an unknown degree of extraneous uncertainties. A comparative study indicates that model statistics are surprisingly robust and appear not to be seriously affected by the above concerns (possibly because of normalisation). Model factors from drained analysis seem to be more variable than those from undrained analysis. A more detailed examination indicates that the higher COV of about 40% for these drained model factors arises because they are not completely random. There are reasons to believe that applying a more complete force system for drained analysis could minimise some of the undesired correlations and reduce the COV to a level comparable to undrained analysis.

KEYWORDS: full-scale tests; model tests; limit state design/analysis; piles; statistical analysis

Cet exposé présente une évaluation critique de facteurs de modèles pour des puits percés rigides à charge latérale (piles forées). Nous prenons en compte la limite latérale et la limite de moment ainsi que la capacité hyperbolique afin de rendre explicite la dépendance des facteurs de modèles vis-à-vis du critère servant à interpréter la 'capacité' d'après les données d'essais de charge. Bien que la capacité hyperbolique puisse être la plus proche de l'état ultime théorique ou de la limite supérieure, les résultats indiquent qu'en général cela ne produit pas un facteur de modèle moyen de 1. Lorsque la capacité mesurée est interprétée de manière consistante d'après les résultats des essais de charge, le coefficient de variation (COV) paraît rester relativement constant, entre 30% et 40%. La gamme de l'inclinaison moyenne pour la limite latérale ou la limite de moment se situe entre 0,67 et 1,49, alors que celle de la capacité hyperbolique se situe entre 0,98 et 2,28. D'après les données disponibles, un modèle à probabilité log-normale semble adéquat pour les analyses de fiabilité. Des essais de charge en laboratoire menés sur des dépôts de sol uniforme préparés en conditions contrôlées sont parfaits pour établir des points de référence sur la magnitude probable de l'incertitude venant des idéalisations de modèles seules. Cependant, la gamme limitée de paramètres géométriques et géotechniques dans une base de données d'essais de charge en laboratoire peut ne pas reproduire un facteur de modèle moyen représentatif. Une base de données d'essais de charge sur le terrain contient de manière typique des paramètres géométriques et géotechniques plus divers mais elle possède un degré inconnu d'incertitudes étrangères. Une étude comparative indique que les statistiques du modèle sont étonnamment robustes et semblent ne pas être affectées de manière sérieuse par les problèmes susmentionnés (probablement en raison de la normalisation). Les facteurs de modèles des analyses drainées semblent être plus variables que ceux des analyses non drainées. Un examen plus détaillé indique que le COV plus élevé d'environ 40% pour ces facteurs de modèles drainés se produit car ceux-ci ne sont pas complètement aléatoires. Nous avons des raisons de penser que le fait d'appliquer un système de force plus complet pour les analyses drainées pourrait minimiser certaines des corrélations indésirables et réduire le COV à un niveau comparable à celui des analyses non drainées.

INTRODUCTION

Progress in the development of reliability-based codes for geotechnical engineering remains limited and slow. More often than not, attempts to bridge the gap between structural and geotechnical design focus primarily on producing the same 'look and feel' as simplified reliability-based structural code formats such as the partial factor approach or load and resistance factor design (LRFD). Numerical values for the resistance factors or soil parameter partial factors essentially are selected by splitting the existing global factors of safety judiciously to reproduce working stress designs. Kulhawy & Phoon (2002) advocated the need to re-focus on basic design issues, rather than the format of the design check, and the way in which the original global factor of safety is re-arranged. There is little dispute that the current geotechnical

Manuscript received 22 April 2004; revised manuscript accepted 23 September 2004.
Discussion on this paper closes on 1 August 2005, for further details see p. ii.
* Department of Civil Engineering, National University of Singapore.
† School of Civil and Environmental Engineering, Cornell University, Ithaca, USA.

design process could be improved significantly by integrating the various design components (loads, soil parameters, calculation models, and factors of safety) in a more consistent way within a framework that recognises uncertainties explicitly. Reliability-based design is the only methodology available to date that can ensure self-consistency from both physical and probabilistic requirements and is compatible with the theoretical basis underlying other disciplines such as structural design. In addition, geotechnical design will be subjected to increasing codification as a result of code harmonisation across material types and national boundaries. It also is clear that regulatory pressure will eventually bring geotechnical design within an umbrella framework established predominantly by structural engineers. For example, EN1990 : 2002 (BSI, 2002) describes the principles and requirements for safety, serviceability, and durability of structures, and the basis for their design and verification; it also gives guidelines for related aspects of structural reliability in the annexes. In the United States, this process is also under way for highway bridge design (e.g. Withiam, 2003). Therefore there are strong practical reasons to consider geotechnical LRFD as a simplified reliability-based design procedure, rather than an exercise in rearranging the original global factor of safety. Philosophically, this approach calls for a willingness to accept reliability analysis as a *necessary* basis for geotechnical LRFD calibrations (Phoon *et al.*, 2003).

One widely expressed concern is that it is somewhat difficult to embrace uncertainty directly because it is complex in geotechnical problems (e.g. Green & Becker, 2001). After all, even the uncertainty underlying the evaluation of a design soil parameter is a function of inherent soil variability, degree of equipment and procedural control maintained during site investigation, and precision of the correlation model used to relate a field measurement with the resulting design soil parameter. A total soil variability analysis that lumps all these components together produces only site-specific statistics that cannot be used in a design code. Significant effort is needed to compile statistics on each component so that they can be combined in a more general way. One such study has been conducted (Phoon & Kulhawy, 1996, 1999a, 1999b), and first-order estimates of inherent soil variability, measurement errors, and correlation uncertainties are available for more rigorous calibration of geotechnical reliability-based design equations.

A similar effort is under way to quantify uncertainties associated with geotechnical calculation models. Although many geotechnical calculation models are 'simple', reasonable predictions of fairly complex soil–structure interaction behaviour can still be achieved through empirical calibrations. Because of our geotechnical heritage, which is steeped in such empiricisms, model uncertainties can be significant. Even a simple estimate of the average model bias is crucial for reliability-based design. If the model is conservative, it is obvious that the probabilities of failure calculated subsequently will be biased, because those design situations that belong to the safe domain could be assigned incorrectly to the failure domain, as a result of the built-in conservatism.

Robust model statistics can only be evaluated using (a) realistically large-scale prototype tests, (b) a sufficiently large and representative database, and (c) reasonably high-quality testing where extraneous uncertainties are well controlled. With the possible exception of foundations, insufficient test data are available to perform robust statistical assessment of the model error in many geotechnical calculation models. The development of a fully rigorous reliability-based design code that can handle the entire range of geotechnical design problems is currently impeded by the scarcity of these important statistics. Sidi (1986) was among the first to report model statistics that were established firmly using a large load test database assembled by Olson & Dennis (1982). The focus of the study was on friction piles in clay subjected to axial loading. Briaud & Tucker (1988) conducted a similar study using a 98-pile load test database obtained from the Mississippi State Highway Department. Recent literature includes estimation of model statistics for the calibration of deep foundation resistance factors for AASHTO (American Association of State Highway and Transportation Officials) (Paikowsky, 2002). A substantial part of the study pertains to the evaluation of driven pile axial capacity using dynamic methods. None of these studies addresses the applicability of model statistics beyond the conditions implied in the database. This question mirrors the same concern expressed previously on the possible site-specific nature of soil variabilities.

This paper presents a critical evaluation of model factors using an extensive database collected as part of an EPRI (Electric Power Research Institute) research study on transmission line structure foundations (Chen & Kulhawy, 1994). Small-scale laboratory tests and full-scale field tests on rigid drilled shafts subjected to lateral-moment loading are analysed to illustrate important statistical issues that deserve rigorous examination, and to furnish model statistics for subsequent reliability calibrations. In particular, it is natural to question whether the computation of a theoretical capacity from geotechnical parameters that are potentially affected by spatial variabilities, measurement errors, and transformation uncertainties could lead to significant overestimation of model uncertainty. A comparison between small-scale laboratory tests and full-scale field tests would help to clarify the interaction between aleatoric and epistemic influences.

LATERAL AND MOMENT CAPACITY

A rigorous analysis of a drilled shaft under lateral loading is very complex because of the asymmetrical three-dimensional force system that acts on the shaft. These forces include the active and passive soil resistances acting on the front and back faces, vertical and horizontal shear forces acting along the lateral surface, and axial and shear tip resistances. A detailed analysis of this 3-D problem would require the use of numerical methods. However, most con-

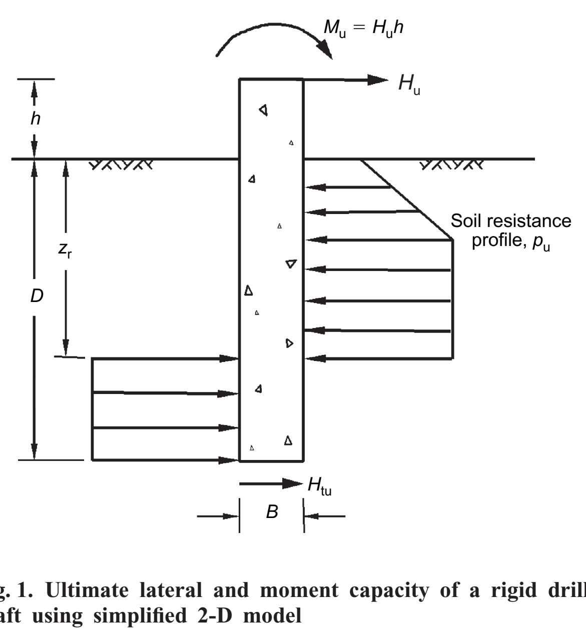

Fig. 1. Ultimate lateral and moment capacity of a rigid drilled shaft using simplified 2-D model

ventional analysis approaches are based on simplified 2-D models that assume the soil resistance develops only from lateral soil stress and perhaps tip shear, as shown in Fig. 1.

The two unknowns in this simplified problem are the ultimate lateral capacity (H_u) and the depth of rotation (z_r). These two unknowns can be determined by balancing forces and moments as follows:

$$H_u = H_{su} - H_{tu} = \int_0^{z_r} p_u B \mathrm{d}z - \int_{z_r}^{D} p_u B \mathrm{d}z - A_{tip}\tau_{tip} \quad (1a)$$

$$M_u = H_u h = -\int_0^{z_r} p_u B z \mathrm{d}z + \int_{z_r}^{D} p_u B z \mathrm{d}z + H_{tu} D \quad (1b)$$

in which M_u is the ultimate moment capacity, H_{su} and H_{tu} are the side and tip resistances, h is the eccentricity of the lateral load, p_u is the ultimate lateral soil stress, B is the shaft diameter, D is the shaft depth, A_{tip} is the tip area, τ_{tip} is the unit tip shear, and z is the depth. The ultimate lateral soil stress typically increases with depth to a limiting value that is produced by a deep failure mechanism in which soil flows around the shaft. Common lateral soil stress distributions include: (a) the undrained models proposed by Reese (1958), Hansen (1961), Broms (1964a), Stevens & Audibert (1979), and Randolph & Houlsby (1984); and (b) the drained models proposed by Hansen (1961), Broms (1964b), and Reese *et al.* (1974). Broms (1964b) further simplified his lateral soil stress distribution by replacing the passive stresses developed below the depth of rotation with a concentrated load at the shaft tip; this is called the 'simplified Broms' method herein.

This study focuses on free-head rigid drilled shafts because full mobilisation of soil strength, as illustrated in Fig. 1, is applicable only if plastic hinges do not form anywhere along the shaft. A summary of the databases reported by Chen & Kulhawy (1994) is shown in Table 1. In practice, it is not easy to ascertain whether a shaft is 'rigid' in the sense mentioned above, because drilled shafts usually are not exhumed after a load test, and the failure mechanism is therefore not documented. The drilled shafts in the databases were classified as 'rigid' based on recommendations proposed by Poulos & Davis (1980), Poulos & Hull (1989), and Carter & Kulhawy (1992). Undrained laboratory tests were conducted in Cornell clay, which is an inorganic silty clay of low plasticity index (about 10%) with undrained shear strength less than 10 kPa. Details of the tests are given by Mayne *et al.* (1992). Drained laboratory tests were conducted in filter sand ranging from loose to dense. The effective stress friction angles range from 38° to 49°, and the in-situ horizontal soil stress coefficients range from 0·3 to 0·9. Details of the tests are given by Agaiby *et al.* (1992). Undrained field load test data were collected from 14 sites, with predominantly clayey soils. The plasticity indices range from about 10% to 30%, and the undrained shear strengths range from about 20 to 250 kPa. Drained field test data were collected from 10 sites, with predominantly sandy soils. The effective stress friction angles range from 31° to 41°, and the in-situ horizontal soil stress coefficients range from 0·4 to 2·4. Details of the field tests are given by Chen & Kulhawy (1994). It should be noted that there are small variations in the numbers of tests cited herein in tables and figures because some tests in the database do not provide adequate data for full interpretation.

There are various methods to interpret the lateral 'capacity' from load tests, such as the displacement limit, rotation limit, lateral or moment limit, and hyperbolic capacity. The displacement and rotation limits are rather arbitrary and do not relate directly to soil–shaft behaviour. The lateral or moment limit (Hirany & Kulhawy, 1988, 1989) is based on the mode of soil–shaft failure and essentially represents a first yield or lower bound. The hyperbolic capacity represents an ultimate limit or upper bound because it is the asymptotic limit of the load–displacement curve. However, it requires extrapolation from measured data, and the asymptote is computed mathematically using a hyperbolic equation with no reference to actual shaft behaviour. One could argue that the lateral or moment limit is a better choice because it is evaluated from the mode of soil–shaft failure. Both criteria are discussed below to evaluate the impact of capacity interpretation method on model statistics.

STATISTICS OF MODEL FACTORS

A plausible and common method of correcting for model error is to assume the following multiplicative model (e.g. Ang & Tang, 1984, Sidi, 1986):

$$H_L \text{ or } H_h = M H_u \quad (2)$$

in which H_L is the interpreted lateral or moment limit, H_h is the interpreted hyperbolic capacity, H_u is the computed lateral capacity, and M is a model factor, typically assumed to be a log-normal random variable.

Lateral or moment limit of rigid drilled shafts

The empirical distributions of M for rigid drilled shafts subjected to undrained and drained loadings are summarised in Figs 2 and 3 respectively. The 'ultimate' lateral capacity is interpreted from load test data using the lateral or moment limit method. The laboratory-scale load tests were conducted in uniform kaolinite clay and filter sand deposits prepared under controlled laboratory conditions. Therefore uncertainties arising from evaluation of soil parameters are minimal. In addition, construction variabilities and measurement errors associated with load tests are minimal. Therefore model uncertainties computed from laboratory tests should be a relatively accurate indicator of errors arising from the use of simplified calculation models. The main concern is whether the model factors are applicable beyond the uniform profile and specific soil type used in the laboratory. Model factors from field tests are expected to be more general because they are computed from load tests conducted in more diverse

Table 1. Description of databases on laterally loaded rigid drilled shafts

Description	No. of tests	B	D/B	h/D
Undrained loading:				
Laboratory tests	48	89–175 mm	3·00–7·98	0·03–4·01
Field tests	27	0·08–1·98 m	2·25–10·49	0·03–6·83
Drained loading:				
Laboratory tests	55	76–152 mm	2·61–9·03	0·06–4·99
Field tests	22	0·05–1·62 m	2·49–7·03	0·00–5·37

B, shaft diameter; D, shaft depth; h, lateral load eccentricity.

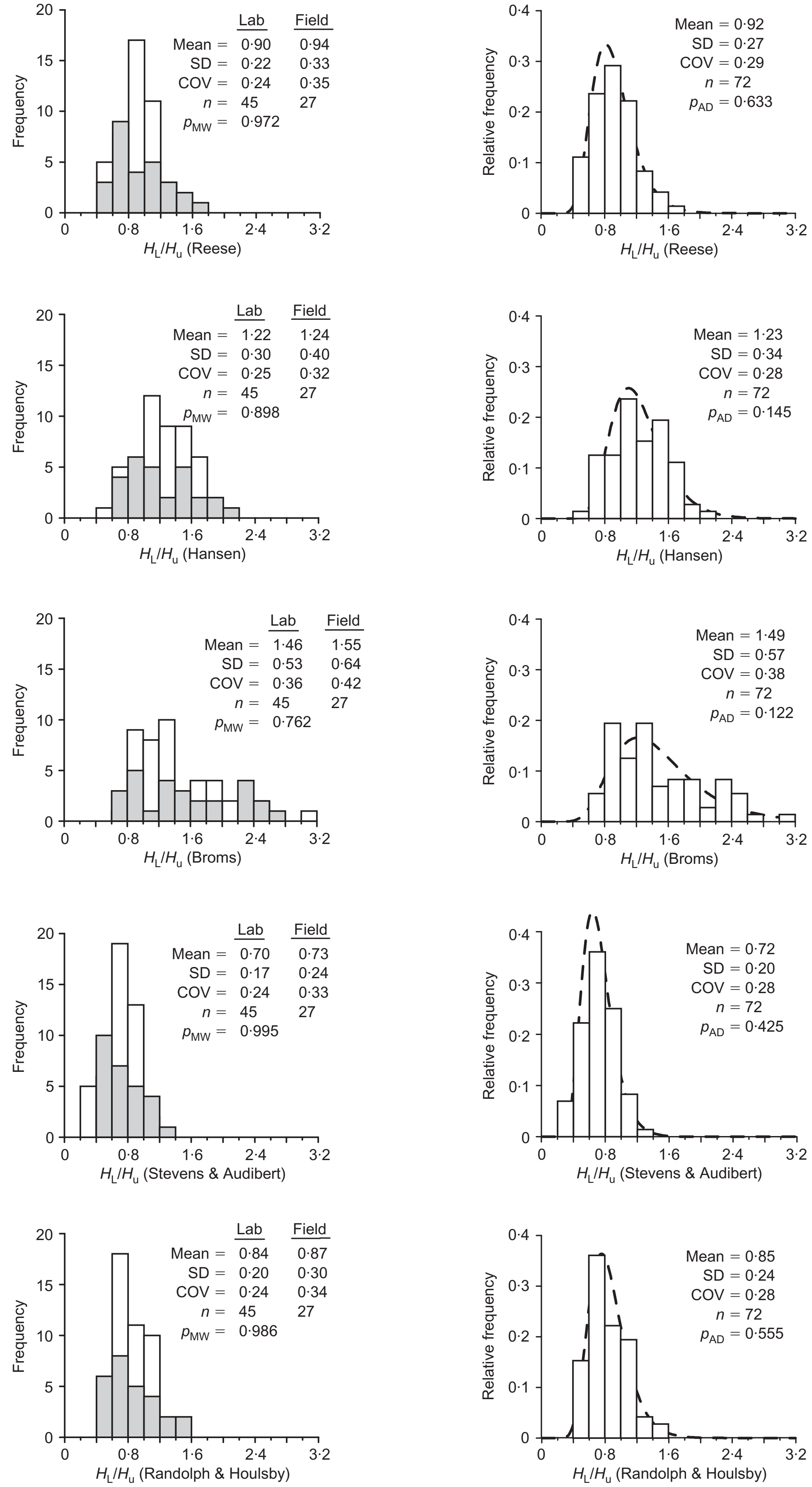

Fig. 2. Undrained model factors of rigid drilled shafts based on lateral or moment limit

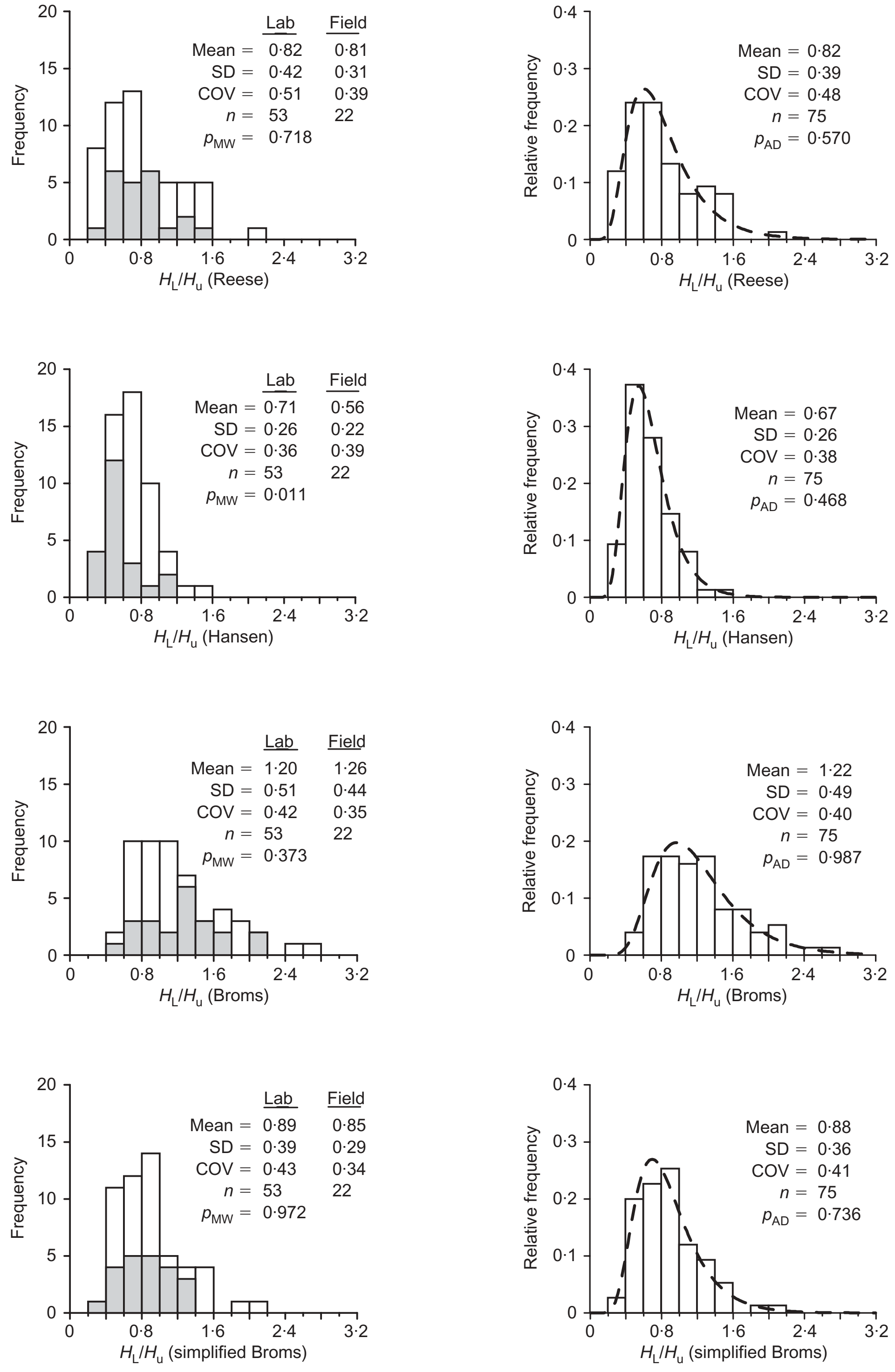

Fig. 3. Drained model factors of rigid drilled shafts based on lateral or moment limit

site environments. However, it is reasonable to query whether the statistics of such model factors are lumped statistics, in the sense that extraneous sources of uncertainties (e.g. construction variabilities, measurement errors incurred during load tests, and uncertainties in soil parameter evaluation) are inextricably included in the computation.

A comparison between laboratory and field data such as those shown in Figs 2 and 3 is illuminating. Laboratory and field results are plotted as white and shaded histograms in the left panel of each figure respectively. Visual inspection and simple statistics (mean, standard deviation (SD), coefficient of variation (COV)) show that the histograms are similar. The p-values from the Mann–Whitney test (p_{MW}) show formally that the null hypothesis of equal medians cannot be rejected at the customary 5% level of significance, with the exception of the drained factor for the Hansen model. Therefore it is reasonable to argue that the results presented in Figs 2 and 3 have wider applicability beyond the conditions implied by the underlying databases, and the model uncertainties are caused mainly by idealisations intrinsic to the respective analytical models. This observation is quite interesting because the geometric and geotechnical parameters in the laboratory and field databases are noticeably different. For example, the average diameter is about one order of magnitude smaller in the laboratory tests. Moreover, the undrained shear strengths and in-situ horizontal soil stress coefficients in the respective undrained and drained laboratory tests are in the lower ranges of those found in the field tests, whereas the effective stress friction angles in the drained laboratory tests are in the upper range of those found in the field tests.

A more robust estimate of the empirical distribution is obtained by combining the laboratory and field data, as shown in the right panel of each figure. Note that the vertical axis is relative frequency, which is defined as the ratio of frequency to sample size (n). For rigid drilled shafts subjected to lateral-moment loading, the COV of the model factor appears to fall within a narrow range of about 30–40%, with the exception of the drained Reese *et al.* (1974) model. However, the mean bias can vary from 0·67 to 1·49, depending on the drainage condition and the analytical model. Therefore it is more important to estimate a realistic mean bias for the analytical model used in reliability calibration. Fortunately, fewer data are required to estimate the mean bias than the COV, for a given level of statistical precision. The reason is that the variance of the sample mean (σ^2/n) is smaller than the variance of the sample variance [$2\sigma^4/(n-1)$], in which σ^2 is the variance of the underlying population. In the absence of sufficient data to estimate the sample variance reliably, it is reasonable to assume a COV between 30% and 40%, given its narrow range. If the above observations prove to be generic, reliability calibration incorporating model statistics would be more economically feasible, given that the cost per load test is significantly more expensive than the cost per soil test.

The log-normal probability density function is plotted over the combined laboratory and field data (as a dashed line) for visual comparison. It appears to be a reasonable probability model for M. However, for sample sizes commonly encountered in load test databases (say $n \approx 75$ or so), histograms are known to be notoriously misleading in identifying the underlying probability model. In fact, it can be proven that histograms derived from small sample sizes are not expected to look like the population probability model. Histograms that provide a good fit are suspicious. At present, it is safe to say that the log-normal probability model is more of a reasonable hypothesis, rather than a fact firmly established by empirical data, in contrast to the mean bias and COV discussed above. Using the Anderson–Darling test, one could formally state that there is no evidence to reject the null hypothesis of log-normality at the customary 5% level of significance, because the p-values (p_{AD}) are larger than 0·05. The Anderson–Darling test is a modification of the more well-known Kolmogorov–Smirnov (K-S) test that gives more weight to the probability tails than does the K-S test. It has the advantage of being a more sensitive test.

Note that a crude estimate of the mean and COV of the model factor is still better than making an unwarranted assumption that the mean = 1 and COV = 0. This point can be illustrated by considering a simple reliability problem involving one load (F) and one capacity (H) random variables, both of which are log-normally distributed. The exact solutions for the reliability index (β) and probability of failure (p_f) are:

$$\beta = \frac{\ln\left(\mu_{FS}\sqrt{\dfrac{1+\mathrm{COV}_F^2}{1+\mathrm{COV}_H^2}}\right)}{\sqrt{\ln[(1+\mathrm{COV}_F^2)(1+\mathrm{COV}_H^2)]}} \tag{3a}$$

$$p_f = \Phi(-\beta) \tag{3b}$$

in which μ_{FS} is the mean factor of safety, and COV_F and COV_H are the coefficients of variation of the load and capacity respectively. Assuming $\mu_{FS} = 3$ and $\mathrm{COV}_F = \mathrm{COV}_H = 0{\cdot}3$ with no model influence, the reliability index is 2·646 and the probability of failure is 0·0041. If the mean and COV of the model factor were 0·8 and 0·3 respectively, the mean factor of safety would decrease to $\mu_{FS} = 3 \times 0{\cdot}8 = 2{\cdot}4$ and the COV of the capacity would increase to $\mathrm{COV}_H = (0{\cdot}3^2 + 0{\cdot}3^2)^{0{\cdot}5} = 0{\cdot}42$. The reliability index and probability of failure in this case are 1·679 and 0·0465, which represents a one order of magnitude change in p_f.

Hyperbolic capacity of rigid drilled shafts

The model statistics evaluated based on the hyperbolic capacity of rigid drilled shafts are summarised in Table 2. Because the hyperbolic capacity is the asymptotic limit of the load–displacement curve, it is not surprising that the mean bias is consistently larger in this case (range of mean bias is between 0·98 and 2·28). However, it is surprising to find that the COVs remain relatively unchanged. This lends support to the previous claim that it is sufficient to estimate the mean bias from limited load test data and assume a log-normal probability model with a COV between 30% and 40% for reliability-based design. The similar COVs also indicate that the hyperbolic capacity (H_h) and lateral or moment limit (H_L) could be interpreted from load test data with similar degrees of consistency.

Despite the different theories and empirical data used to justify the analytical models considered in this study, no one particular model results in a model uncertainty that is significantly smaller than the rest. However, the Broms (1964a) method for undrained loading appears to produce more variable model factors, regardless of the capacity interpretation method used. It is interesting to note that the Anderson–Darling p-value is significantly higher in many cases (particularly for drained loading) when the lateral or moment limit is adopted. This result implies that the model factor defined using H_L in the numerator more closely resembles a log-normal random variable.

Undrained versus drained analysis

Another evident trend is that the COV for drained analysis (≈ 40%) is generally higher than the COV for undrained analysis (≈ 30%). One hesitates to conclude that analytical models for drained analysis are less precise, because some of the components in the complex 3-D force system induced by lateral-moment loading should not be neglected in drained analysis. One complicating factor is that the lateral soil stress distributions between undrained and drained analyses are not strictly comparable, because they were not derived from a unified model that can handle a general c–ϕ soil. The only exception is the Hansen (1961) model. For H_L, the COV for drained analysis is significantly higher, but less so when H_h is adopted. The lateral soil stress distributions proposed by Reese (1958) and Reese *et al.* (1974) could perhaps be considered as relatively unified because they are based on the same conceptual models for shallow and deep failures. In this case, the COV for drained analysis is significantly higher regardless of the capacity interpretation method.

Davidson *et al.* (1982) proposed a more complete equilibrium system for drilled shafts with D/B between 1 and 10 that includes lateral soil resistance, vertical side shear, tip shear, and tip moment. Lateral soil resistance is modelled using the Hansen (1961) lateral soil stress distribution. The ultimate lateral capacity is computed numerically using the p–y curve approach. The model factors computed using the

Table 2. Model factors for rigid drilled shafts based on hyperbolic capacity

Model Undrained:		Laboratory $n = 47$	Field $n = 27$		Combined $n = 74$
Reese (1958)	Range	0·75–2·41	0·82–2·72	Range	0·75–2·72
	Mean	1·43	1·40	Mean	1·42
	COV	0·26	0·33	COV	0·29
	p_{MW}	0·315		p_{AD}	0·186
Hansen (1961)	Range	0·86–3·56	1·13–3·61	Range	0·86–3·61
	Mean	1·95	1·85	Mean	1·92
	COV	0·28	0·31	COV	0·29
	p_{MW}	0·296		p_{AD}	0·175
Broms (1964a)	Range	1·08–4·31	1·09–4·49	Range	1·08–4·49
	Mean	2·28	2·29	Mean	2·28
	COV	0·35	0·41	COV	0·37
	p_{MW}	0·875		p_{AD}	0·149
Stevens & Audibert (1979)	Range	0·55–2·03	0·63–2·13	Range	0·55–2·13
	Mean	1·12	1·09	Mean	1·11
	COV	0·28	0·32	COV	0·29
	p_{MW}	0·435		p_{AD}	0·367
Randolph & Houlsby (1984)	Range	0·67–2·33	0·77–2·52	Range	0·67–2·52
	Mean	1·33	1·30	Mean	1·32
	COV	0·27	0·32	COV	0·29
	p_{MW}	0·351		p_{AD}	0·270
Drained:		$n = 55$	$n = 22$		$n = 77$
Reese *et al.* (1974)	Range	0·40–3·35	0·85–2·07	Range	0·40–3·35
	Mean	1·19	1·19	Mean	1·19
	COV	0·48	0·30	COV	0·43
	p_{MW}	0·440		p_{AD}	0·168
Hansen (1961)	Range	0·55–2·33	0·55–1·55	Range	0·55–2·33
	Mean	1·05	0·83	Mean	0·98
	COV	0·32	0·30	COV	0·33
	p_{MW}	0·002		p_{AD}	0·229
Broms (1964b)	Range	0·90–3·40	0·85–3·07	Range	0·85–3·40
	Mean	1·77	1·89	Mean	1·80
	COV	0·40	0·33	COV	0·38
	p_{MW}	0·308		p_{AD}	0·064
Simplified Broms (1964b)	Range	0·64–2·62	0·59–1·95	Range	0·59–2·62
	Mean	1·31	1·27	Mean	1·30
	COV	0·40	0·32	COV	0·38
	p_{MW}	0·852		p_{AD}	0·141

Davidson *et al.* (1982) model are compared with those computed using the Hansen (1961) model in Fig. 4. Only laboratory load test data were considered to avoid introducing extraneous uncertainties associated with field data. Model factors based on drained and undrained analyses are plotted as white and shaded histograms respectively. It is interesting to observe that the COV for drained analysis does decrease when a more complete force system is modelled. On the other hand, the COV for undrained analysis remains relatively constant.

A more direct means of verifying whether the absence of some force components contributes to the larger COV for drained analysis is to plot the model factors against the side and tip shears. The computation of the actual side and tip shears is complex, and probably requires 3-D finite element analysis. For the purpose of this study, only nominal measures that could provide a relative indicator of side and tip shears between the load tests are needed. Reasonable measures of the side and tip shears that include the relevant geometric and geotechnical parameters can be evaluated as follows:

Undrained:

$$\text{Side shear} = \pi B D s_u(\text{DSS}) \tag{4a}$$

$$\text{Tip shear} = \frac{\pi B^2}{4} s_u(\text{DSS}) \tag{4b}$$

Drained:

$$\text{Side shear} = \pi B D \overline{\sigma}_{va} K_0 \tan \overline{\phi} \tag{5a}$$

$$\text{Tip shear} = \frac{\pi B^2}{4} \overline{\sigma}_{tip} \tan \overline{\phi} \tag{5b}$$

in which s_u(DSS) is the undrained shear strength from the direct simple shear test; $\overline{\sigma}_{va}$ is the average effective vertical stress along the shaft length, which is assumed to be $(17)(0{\cdot}5D)$ kPa because the unit weight of dry filter sand $\approx$ 17 kN/m^3; K_0 is the in-situ horizontal soil stress coefficient; $\overline{\phi}$ is the effective stress friction angle; and $\overline{\sigma}_{tip}$is the vertical effective stress at the shaft tip, which is simply assumed to be caused by the self-weight of the shaft, i.e. $(24)(D)$ kPa, although the vertical side resistance is a contributing factor as well.

Model factors for undrained and drained analyses are plotted against the nominal side and tip shears in Fig. 5. For undrained analysis, it is clear that the model factors are not correlated to the nominal side or tip shear. This lack of correlation appears regardless of the capacity interpretation method. For drained analysis, large model factors are associated with small side or tip shear and vice-versa. The above observations apply to other analytical models of lateral soil stress distribution, with the exception of the

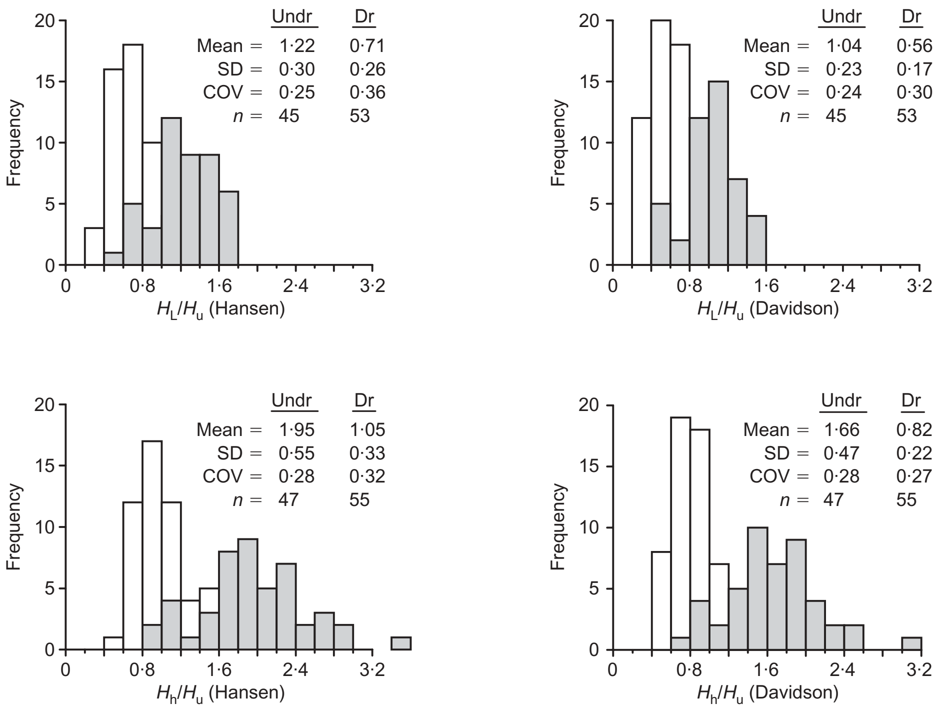

Fig. 4. Comparison between Hansen (1961) and Davidson *et al.* (1982) models using laboratory load test databases

Broms (1964a) model for undrained analysis. Model factors produced by this model also exhibit a similar correlation behaviour seen in drained analysis. This may explain the higher COV associated with this model (~40%) in comparison with the rest of the undrained models (~30%). The drained model factors produced by the Davidson *et al.* (1982) numerical method exhibit a somewhat weaker correlation, primarily because the large model factors at small side or tip shear are reduced so that they fall in line with the rest of the data.

CONCLUSIONS

Ideally, a theoretical model should capture the key features of the physical system, and the remaining difference between the model and reality should be random in nature because it is caused by numerous minor factors that were left out of the model. The statistics of model factors should capture these random differences resulting from model idealisations. In practice, model factors could only be evaluated by comparison with load test data. It is important to question whether model factors so derived would entail extraneous uncertainties arising from:

(*a*) loose definition of measured 'capacity'
(*b*) variabilities associated with construction and load test
(*c*) computation of theoretical capacity from geotechnical parameters that are potentially affected by spatial variabilities, measurement errors, and transformation uncertainties
(*d*) the limited range of geometric and geotechnical parameters inherent in any load test database.

Attempts are made to understand these fairly difficult and fundamental issues, using load test databases for laterally loaded rigid drilled shafts.

In this study, both lateral or moment limit and hyperbolic capacity are considered to make explicit the dependence of model factors on the criterion for interpreting 'capacity' from load test data. Contrary to popular belief, the relationship between measured and theoretical capacities is neither obvious nor well defined. For example, it is tempting to assume that the hyperbolic capacity is associated with a theoretical ultimate state involving full mobilisation of soil strength, because it is the asymptotic limit of the load–displacement curve. However, results in this study show that this tempting definition of measured capacity generally does not produce a mean model factor of 1. The practical solution is to interpret the measured capacity consistently, as adopted in this study. Although the mean model factor would change depending on the interpretation method, the COV appears to remain relatively constant between 30% and 40%, when a reasonably well-defined and consistent criterion is followed. The range of the mean bias for the lateral or moment limit is 0·67 to 1·49, and that of the hyperbolic capacity is 0·98 to 2·28. Based on available data, a simple log-normal probability model appears to be adequate for subsequent reliability analysis.

Laboratory-scale load tests conducted in uniform soil deposits prepared under controlled laboratory conditions are very useful for establishing benchmarks on the probable magnitude of uncertainty arising from model idealisations alone. However, the limited range of geometric and geotechnical parameters in a laboratory load test database may not produce a representative mean model factor. A field load test database typically contains more diverse geometric and geotechnical parameters, but it entails an unknown degree of extraneous uncertainties. A careful comparison study conducted herein indicates that model statistics are surprisingly robust and appear not to be seriously affected by the above concerns (possibly because of normalisation).

Model factors from drained analysis seem to be more variable than those from undrained analysis. A more detailed examination reveals that the higher COV of about 40% for these drained model factors arises because they are not completely random. There are reasons to believe that apply-

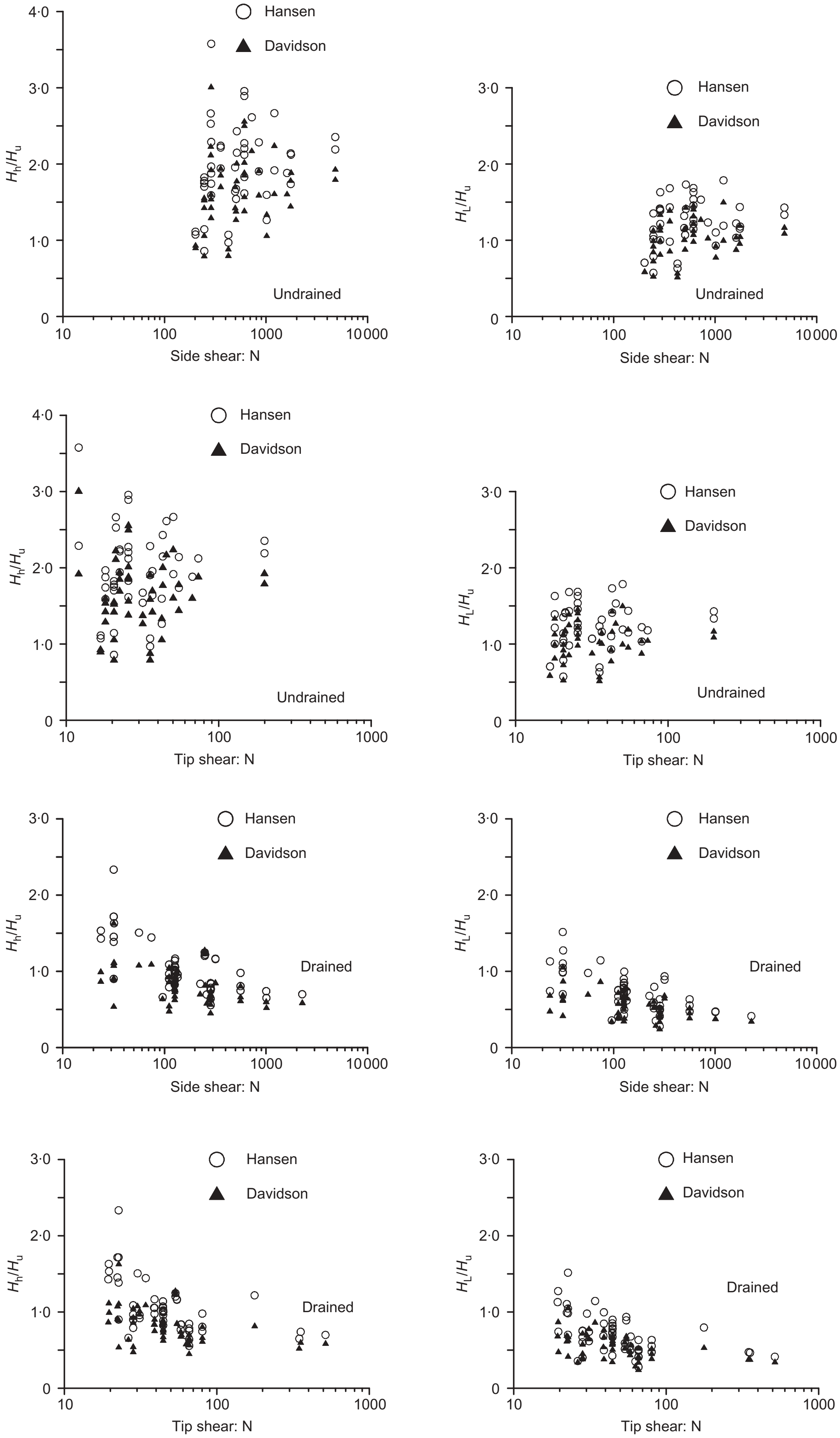

Fig. 5. Effect of nominal measures of side and tip shears on model factors (laboratory load test databases)

ing a more complete force system for drained analysis could minimise some of the undesired correlations and reduce the COV to a level comparable to undrained analysis.

REFERENCES

Agaiby, S. W., Kulhawy, F. H. and Trautmann, C. H. (1992). *Experimental study of drained lateral and moment behavior of drilled shafts during static and cyclic loading*, Report TR-100223. Palo Alto, CA: Electric Power Research Institute.

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

Briaud, J. L. & Tucker, L. M. (1988). Measured and predicted axial response of 98 piles. *J. Geotech. Engng, ASCE* **114**, No. 9, 984–1001.

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

Broms, B. B. (1964a). Lateral resistance of piles in cohesive soils. *J. Soil Mech. Found. Div., ASCE* **90**, No. SM2, 27–63.

Broms, B. B. (1964b). Lateral resistance of piles in cohesionless soils. *J. Soil Mech. Found. Div., ASCE* **90**, No. SM3, 123–156.

Carter, J. P. & Kulhawy, F. H. (1992). Analysis of laterally loaded shafts in rock. *J. Geotech. Engng, ASCE* **118**, No. 6, 839–855.

Chen, Y.-J. & Kulhawy, F. H. (1994). *Case history evaluation of behavior of drilled shafts under axial and lateral loading*, Report TR-104601. Palo Alto, CA: Electric Power Research Institute.

Davidson, H. L., Cass, P. G., Khilji, K. H. & McQuade, P. V. (1982). *Laterally loaded drilled pier research*, Report EL-2197 (1, 2). Palo Alto, CA: Electric Power Research Institute.

Green, R. & Becker, D. (2001). National report on limit state design in geotechnical engineering: Canada. *Geotech. News* **19**, No. 3, 47–55.

Hansen, J. B. (1961). *Ultimate resistance of rigid piles against transversal forces*, Bulletin 12, pp. 5–9. Copenhagen: Danish Geotechnical Institute.

Hirany, A. & Kulhawy, F. H. (1988). *Conduct and interpretation of load tests on drilled shaft foundation: detailed guidelines*, Report EL-5915 (1). Palo Alto, CA: Electric Power Research Institute.

Hirany, A. & Kulhawy, F. H. (1989). Interpretation of load tests on drilled shafts. Part 3: lateral and moment. In *Foundation. engineering: Current principles and practices* (GSP 22) (ed. F. H. Kulhawy), pp. 1160–1172. New York: ASCE.

Kulhawy, F. H. & Phoon, K. K. (2002). Observations on geotechnical reliability-based design development in North America. *Proceedings of the international workshop on foundation design codes and soil investigation in view of international harmonization and performance based design*, Tokyo, pp. 31–48.

Mayne, P. W., Kulhawy, F. H. & Trautmann, C. H. (1992). *Experimental study of undrained lateral and moment behavior of drilled shafts during static and cyclic loading*, Report TR-100221. Palo Alto, CA: Electric Power Research Institute.

Olson, R. E. & Dennis, N. D. (1982). *Review and compilation of pile test results, axial pile capacity*, Geotechnical Engineering Report CR83-4. Austin: Department of Civil Engineering, University of Texas.

Paikowsky, S. G. (2002). *Load and resistance factor design (LRFD) for deep foundations*, NCHRP Report 24–17. Washington, DC: Transportation Research Board.

Phoon, K. K. & Kulhawy, F. H. (1996). On quantifying inherent soil variability. In *Uncertainty in the geologic environmen*t (GSP 58) (eds C. D. Shackelford, P. P. Nelson and M. J. S. Roth), pp. 326–340. New York: ASCE.

Phoon, K. K. & Kulhawy, F. H. (1999a). Characterization of geotechnical variability. *Can. Geotech. J.* **36**, No. 4, 612–624.

Phoon, K. K. & Kulhawy, F. H. (1999b). Evaluation of geotechnical property variability. *Can. Geotech. J.* **36**, No. 4, 625–639.

Phoon, K. K., Becker, D. E., Kulhawy, F. H., Honjo, Y., Ovesen, N. K. & Lo, S. R. (2003). Why consider reliability analysis in geotechnical limit state design? *Proceedings of the international workshop on limit state design in geotechnical engineering practice* (LSD2003), Massachusetts Institute of Technology, Cambridge (CD-ROM).

Poulos, H. G. & Davis, E. H. (1980). *Pile foundation analysis and design*. New York: John Wiley & Sons.

Poulos, H. G. & Hull, T. S. (1989). Role of analytical geomechanics in foundation engineering. In *Foundation engineering: Current principles and practices* (GSP 22) (ed. F. H. Kulhawy), pp. 1578–1606. New York: ASCE.

Randolph, M. F. & Houlsby, G. T. (1984). Limiting pressure on a circular pile loaded laterally in cohesive soil. *Géotechnique* **34**, No. 4, 613–623.

Reese, L. C.(1958). Discussion of 'Soil modulus for laterally loaded piles. *Trans. ASCE* **123**, 1071–1074.

Reese, L. C., Cox, W. R. & Coop, F. D. (1974). Analysis of laterally loaded piles in sand. *Proc. 6th Offshore Technology Conf., Houston* **2**, 473–483.

Sidi, I. D. (1986). *Probabilistic prediction of friction pile capacities*. PhD thesis, University of Illinois, Urbana-Champaign, 1986.

Stevens, J. B. & Audibert, J. M. E. (1979). Re-examination of p-y curve formulations. *Proc. 11th Offshore Technology Conf., Houston* **1**, 397–403.

Withiam, J. L. (2003). Implementation of the AASHTO LRFD bridge design specifications for substructure design. *Proceedings of the international workshop on limit state design in geotechnical engineering practice* (LSD2003), Cambridge, MA (CD-ROM).