INTERNATIONAL JOURNAL FOR NUMERICAL AND ANALYTICAL METHODS IN GEOMECHANICS, VOL. 6, 409–424 (1982)

# FACTORS OF SAFETY AND PILE LOAD TESTS

GREGORY B. BAECHER*
*Massachusetts Institute of Technology, U.S.A.*
AND
RÜDIGER RACKWITZ‡
*Technical University Munich, West Germany*

## SUMMARY

The factor of safety used in designing pile foundations for vertical load should depend on three things, prior information on load capacity summarized by empirical correlations with load capacity models, site specific information derived from load tests, and an objective function reflecting economic and safety considerations. A statistical approach to factor of safety selection was developed in order to suggest improvements of current standards for driven pile design. This approach recognizes a distinction between the variability of pile load capacity within individual sites, and the global variability upon which model correlations are based. Charts have been prepared for determining the FS required to achieve specified reliability indices, as a function of the number of load tests at a particular site and their outcomes.

## INTRODUCTION

Few problems in geotechnical engineering have received more attention than the vertical load capacity of individual piles, yet the analytical prediction of load capacity (yield load or load for given deformation) remains uncertain. In practice, the design of pile supported foundations is based primarily on the results of limited numbers of pile load tests carried out at the construction site.

Despite this situation, many formulae for predicting pile capacity have been proposed, and to one extent or another have been empirically correlated to observed pile capacities. Many of these formulae are based on an energy balance of the dynamics of pile driving (e.g., Reference 9), but soil statics models and methods based on cone penetration resistance have also been developed (e.g., Reference 12). The precision with which any of these formulae or methods can predict the load capacity of an actual pile is limited, but they serve to correlate the driving records or other qualities of placed piles with the observed performance of tested piles. They serve as indices against which the empirical record of pile load testing can be organized.

To predict pile capacity at an individual site prior information, usually that contained in empirical correlations, must be combined with the results of a limited number of pile load tests to draw conclusions. This is so whether statistical procedures are used or not, because the amount of information contained in the results of one, two, or perhaps three on site tests is very small. The purpose of the present work has been to develop a formal method for

* Associate Professor of Civil Engineering.
‡ Research Associate.

0363–9061/82/040409–16$01.60

*Received 23 December 1980*
*Revised 4 May 1981*

combining these two types of information, giving each its proper weight, and leading to a single prediction with an associated statement of uncertainty.

Attention in the present study has been restricted to individual driven friction piles in cohesionless soils. The reason for restricting attention to individual piles is that group capacity is seldom tested directly, given the large reactions required. Predictions of group capacity are made by analogy with model tests, by analytical modelling, or by heuristic rules. Thus, the statistical clarity of direct observation is lost. However, because present standards specify factors of safety on individual piles, the usefulness of the study is not seriously limited. The reasons for excluding cohesive soils were that less prior information is available (e.g. Reference 7), and that test results are sensitive to installation procedures, time after driving, drainage conditions and loading rates.

The study has considered site information only to the extent of the results of pile load tests. The quality and amount of information coming from borings, penetrations, laboratory tests and other measurement of soil properties have not been considered here. Thus, the work focuses on dynamic pile capacity formulae. However, the procedure developed is equally applicable to other methods for predicting load capacity (e.g., those based on penetrometer profiles), and four are compared.

It must be firmly stressed that the present work is no apology for dynamic formulae, and takes no position on whether they should or should not be a basis for practice. The reader is free to interpret the results as he wishes to judge the relative usefulness of dynamic *vs.* other formulae.

## BACKGROUND

The question of estimating pile capacity from dynamic formulae combined with load tests is, from a statistical point of view, more straight forward than many other problems in geotechnical engineering. Specifically, these estimates involve direct measurements rather than estimates of soil parameters propagated through a soil mechanics model. Thus, the problem is similar to a number of traditional sampling problems in statistics.

Obviously, any inference of pile capacity from load tests requires first, a clear definition of ultimate or working capacity as manifest in the load—settlement curve; and second, control of loading rate and therefore drainage conditions such that the results can be interpreted. No matter how sophisticatedly test results are analysed, predicted capacities are of little value if the basic data are of questionable reliability.

In any investigation, the variation of results from one test to another is, in principle, explainable by causal variables. That is, pile capacities vary from test to test for specific physical reasons. Were these causal variables known, more accurate predictions of pile capacity could be made. However, the fact that such causal mechanisms exist is of little help if they are unknown at the time of prediction. Statistical methods are introduced to deal with these residual uncertainties by maintaining logical consistency among prior information, observations and inferences.

Recently, Kay[7,8] has proposed a Bayesian statistical approach to combining the prior information of pile capacity formulae with load test results. This approach has been taken as a starting point for the present work, and certain limitations of that earlier work have been eliminated. Specifically, the present work develops a consistent procedure for dealing with variations of pile capacity within individual construction sites and among them, incorporates statistical uncertainty on within site variability, and refines the use of reliability indices. This is viewed as an extension of Kay's important contribution.



## PRIOR INFORMATION

Prior information for inferring pile load capacity comes from two sources: global correlations of observed capacity to capacity formulae, and results of multiple pile load tests at individual sites from which within site variation is estimated.

Several authors have presented data showing correlations between observed test results and the prediction of pile capacity formulae (e.g., References 4, 14, 17). These correlations are typically presented as regression of observed capacity to predicted or as distributions on the ratio of observed capacity to predicted. Figure 1 shows a typical result from Olson and Flaate,[14] in which the logarithm of observed to predicted capacity by Janbu's formula

$$R = \log_{10}\{\text{observed capacity}/\text{predicted capacity}\} \tag{1}$$

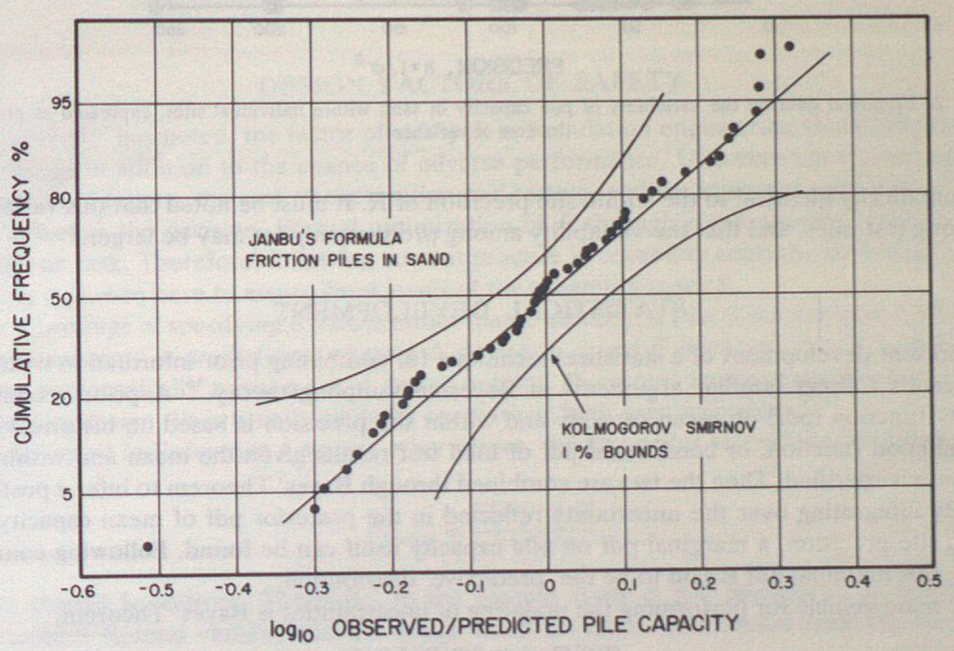


Figure 1. Empirical distribution of observed to predicted pile capacity based on Janbu's formula for steel piles in sand showing Kolmogorov–Smirnov 1 per cent goodness-of-fit bounds

is plotted against cumulative frequency on normal probability grid. It should be noted that summarizing the empirical record in the ratio $R$ is computationally convenient, but presumes a regression line of observed to predicted passing through the origin.

Correlations of the sort shown in Figure 1 have been widely commented on in the literature, and require little further discussion. An important aspect to note, however, is that empirically the ratio of observed to predicted capacity seems well modelled by a lognormal distribution, and passes common goodness-of-fit tests.

The second source of prior information concerns variation within individual sites. To the three cases reported by Kay, thirteen have been added,[3,15] to estimate the distribution of within site variation shown in Figure 2. This variation is measured by the precision, $h$, of the logarithm of pile capacity, defined as the reciprocal of within site variance $h = 1/\sigma^2$. This is

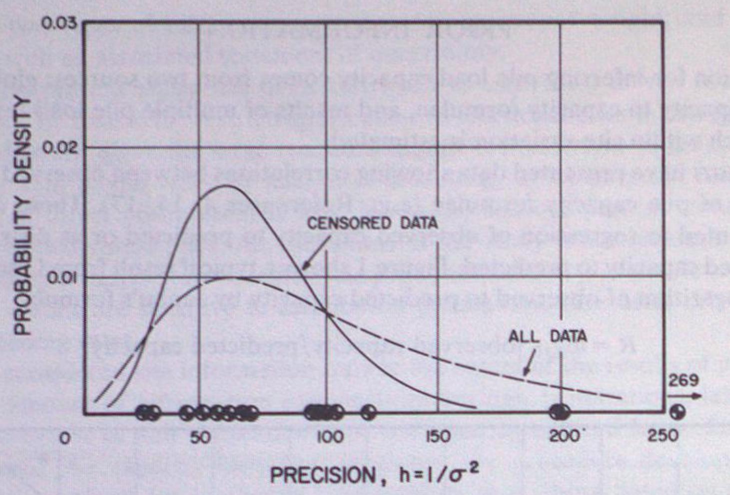


Figure 2. Empirical data on the variability of pile capacity in sand within individual sites, expressed as precision (inverse of variance)

mathematically identical to the within site precision of $R$. It must be noted that this variability is among test piles, and that the variability among production piles may be larger.*

## STATISTICAL DEVELOPMENT

The present development of a statistical technique for combining prior information with load test results follows familiar arguments in Bayesian sampling theory.[16] A prior probability density function (pdf) on mean capacity and within site precision is based on historical data. A likelihood function, or conditional pdf of load test results given the mean and within site precision is specified. Then the two are combined through Bayes' Theorem to infer a posterior pdf. By integrating over the uncertainty reflected in the posterior pdf of mean capacity and within site precision, a marginal pdf on pile capacity itself can be found. Following common usage, this marginal pdf is said to be the 'predictive' distribution.

The main vehicle for performing the updating of uncertainties is Bayes' Theorem,

$$f''(\boldsymbol{\theta}|D, \mathbf{z}) \propto f'(\boldsymbol{\theta}|D) L(\mathbf{z}|\boldsymbol{\theta}) \tag{2}$$

where $f'(\boldsymbol{\theta}|D)$ is a prior pdf over a set of distribution parameters $\boldsymbol{\theta}$ given previous data $D$, $L(\mathbf{z}|\boldsymbol{\theta})$ is the likelihood or conditional pdf of the test results $\mathbf{z}$ given values of $\boldsymbol{\theta}$, and $f''(\boldsymbol{\theta}|D, \mathbf{z})$ is the posterior pdf of $\boldsymbol{\theta}$ given $D$ and $\mathbf{z}$. The predictive pdf of further observations $\mathbf{y}$, for example the capacity of the designed piles, is then

$$f(\mathbf{y}|D, \mathbf{z}) = \int_{\boldsymbol{\theta}} f(\mathbf{y}|\boldsymbol{\theta}) f''(\boldsymbol{\theta}|D, \mathbf{z})\, d\boldsymbol{\theta} \tag{3}$$

The main statistical model in equation (3) is the likelihood function $L(\mathbf{z}|\boldsymbol{\theta})$, where the prior data $D$ are assumed implicit, describing the probability of observing the test results $\mathbf{z}$ given various values of the statistical parameters $\boldsymbol{\theta}$. From Figure 1 the empirical log-normality of the ratio of observed to predicted pile capacity for the Janbu formula is apparent. Therefore

* The authors acknowledge the helpful comments of a reviewer on this point.



the likelihood for the logarithm of the ratio of observed to predicted capacity $R$ is taken as normally distributed with mean $\mu$ and standard deviation $\sigma$,

$$L(R|\boldsymbol{\theta}) \sim N(R|\mu, \sigma) \tag{4}$$

In common practice, the prior distribution over $\mu$ and $\sigma^2$ in Normal sampling is taken to be Normal-Inverted Gamma ($N\Gamma^{-1}$), as this distribution is closed under multiplication by (4). That is, if the prior is $N\Gamma^{-1}$, so is the posterior.[16] Any distribution with this property is said to be a natural conjugate to the likelihood function, and its use simplifies calculations. However, it should be noted that the natural conjugate distribution also arises if prior to having observed $D$, little or no information was available on $\boldsymbol{\theta}$; that is, if the distribution on $\boldsymbol{\theta}$ were taken as non-informative or diffuse.[16]

In traditional Bayesian sampling for Normal processes, uncertainty on $\mu$ and $\sigma$ are both assumed to arise from the same sample of $n$ observations. Therefore, these two uncertainties are uniquely linked through the sample size $n$. In the present application, however, this is not the case. Prior information on $\mu$, after appropriate modification, comes from global correlations with the Janbu formula. Prior information on $\sigma^2$, the within site variance, comes from groups of load tests carried out within individual sites. Therefore, in the prior $f^0(\mu, \sigma)$ these uncertainties must be decoupled. Defining, as before, the precision of within site variability to be the inverse variance, $h = (1/\sigma^2)$, the prior on $\mu$ and $h$ is taken to be Normal-Gamma

$$f'(\mu, h) \propto h^{1/2} \exp\left(-\tfrac{1}{2} h n'(\mu - \mu')^2\right) \cdot \exp\left(-\tfrac{1}{2} h \nu' v'\right) h^{1/2(\nu'-1)} \tag{5}$$

where

$n'$ = prior equivalent sample size for the mean of $R$,
$\mu'$ = prior expected value of the mean of $R$,
$\nu'$ = prior degrees of freedom for the precision, $h$,
$v'$ = prior location parameter for the distribution of $h$.
Note that $\nu'$ does not necessarily equal $(n'-1)$.

Given a new sample of observations $r = (r_1, \ldots, r_n)$, for example the results of on site load tests, with the statistics

$$\bar{r} = \frac{1}{n} \sum_{i=1}^{n} r_i \tag{6}$$

$$s^2 = \frac{1}{\nu} \sum_{i=1}^{n} (r_i - \bar{r})^2 \tag{7}$$

where

$$\nu = n - 1 \tag{8}$$

the posterior density function over $(\mu, h)$ calculated through equation (2) becomes

$$f''(\mu, h|\mathbf{r}) \propto h^{1/2} \exp\left[-\tfrac{1}{2} h n''(\mu - \mu'')^2\right] \cdot h^{1/2(\nu''-1)} \exp\left[-\tfrac{1}{2} h \nu'' v''\right] \tag{9}$$

in which

$$n'' = n + n'$$

$$\mu'' = (n\bar{r} + n'\bar{r}')/n''$$

$$\nu'' = n + \nu'$$

$$\nu'' v'' = \nu' v' + (n-1)s^2 + n'\mu'^2 + n\bar{r}^2 - n''\mu''^2$$

The predictive distribution of $r$ is then calculated by equation (3) to be (Appendix I)

$$f(r|\mathbf{r}) \propto \left[1 + \frac{H}{\nu''}(r-\mu'')\right]^{-1/2(\nu''-1)} \tag{10}$$

This is a Student $t$ distribution of $\nu''$ degrees of freedom having parameters $\mu''$ and

$$H = \frac{n''}{(n''+1)v''} \tag{11}$$

The mean and variance of the predicted $r$ are

$$E[r|\mathbf{r}] = \mu''$$
$$V[r|\mathbf{r}] = \frac{H^{-1}\nu''}{\nu''-2} \tag{12}$$

## DESIGN FACTORS OF SAFETY

As Meyerhof[13] has noted, the factor of safety $F$ in foundation engineering should depend on many things in addition to the chance of adverse performance. Uncertainties in loading, the seriousness of failure, the cost of exploration and testing, and the marginal cost of increasing $F$ all influence the decision. Formal optimization of $F$ through comprehensive risk analysis is a difficult task. Therefore, following current practice in reliability analysis, the design factor of safety is chosen here to assure fixed levels of the reliability index $\beta$.

The advantage of specifying $\beta$ values rather than $F$ directly, is that $\beta$ incorporates a measure of the uncertainty in predictions, whereas $F$ does not. $\beta$ and $F$ need not be related uniquely or even monotonically across a set of design options. In current applications the selection of appropriate design $\beta$'s is usually made by comparison with existing codes.

As discussed in Reference 1 among other places, the first-order second-moment $\beta$ defined as

$$\beta_{\mathrm{FOSM}} = \frac{E[F]-1\cdot0}{\sqrt{(V[F])}} \tag{13}$$

suffers several limitations. Therefore, in the present study $\beta$ was defined as an equivalent standardized Normal variate, and the design factor of safety determined from the inverse $t$ distribution. From equation (10)

$$\beta_{\mathrm{st}} = \frac{r_0-\mu''}{\sqrt{H}} \tag{14}$$

is a standardized $t$ variate with $\nu''$ degrees of freedom. The design value of $r$, $r_0$, was chosen to equate the probability content of the tail of the $t$ distribution beyond $r_0$ with that of the standardized Normal distribution beyond $\beta$ (Figure 3).

Thus

$$r_0 = T_{\nu''}^{-1}[\Phi(-\beta)\sqrt{H}+\mu''] \tag{15}$$

where $T_{\nu''}^{-1}$ is the inverse $t$ distribution with $\nu''$ degrees of freedom, and $\Phi$ is the cumulative Normal. Equivalent $t$ variables for various $\beta$'s and degrees of freedom are shown in Figure 4.

Once the design value $r_0$ is obtained, the required factor safety is

$$F = 10^{-r_0}$$



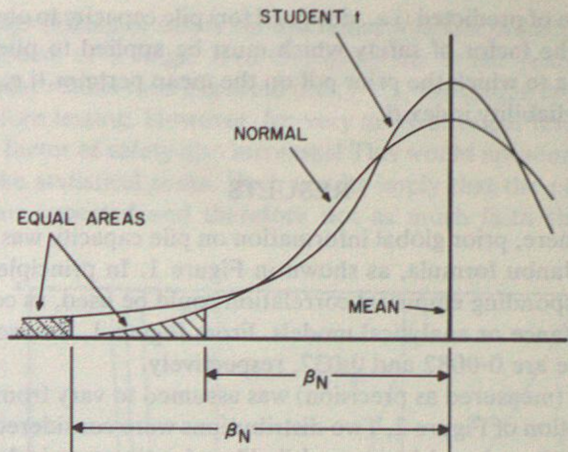


Figure 3. Reliability index for Student $t$ distribution defined as standardized deviate with same tail area as corresponding standardized Normal deviate $\beta$

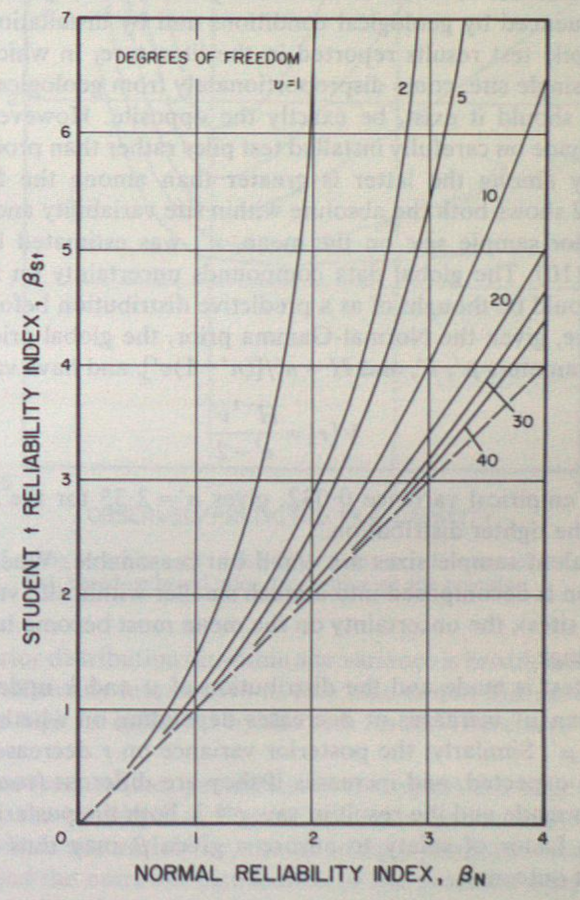


Figure 4. Relation between standardized Student $t$ and Normal variates with same tail area, shown as a function of the degrees of freedom of the Student $t$ variate

which is simply the ratio of predicted (i.e., designed for) pile capacity to observed (i.e., expected) pile capacity. This is the factor of safety which must be applied to pile capacity predictions made using the formula to which the prior pdf on the mean pertains (i.e., correlation) in order to insure the chosen reliability index $\beta$.

## RESULTS

For the case analysed here, prior global information on pile capacity was taken from empirical correlations with the Janbu formula, as shown in Figure 1. In principle, of course, any such formula and the corresponding empirical correlation could be used, as could correlations with cone penetration resistance or analytical models. From Figure 1, the mean and variance of $R$ for the global data base are 0·0082 and 0·032, respectively.

Within site variance (measured as precision) was assumed to vary from site to site according to the Gamma distribution of Figure 2. Two distributions were considered, a broad distribution ($\nu' = 4{\cdot}42$, $v' = 0{\cdot}0122$) based on Maximum-Likelihood estimators including all sixteen data, and a narrower distribution ($\nu' = 9{\cdot}28$, $v' = 0{\cdot}0152$) based on Maximum-Likelihood estimators excluding the three high precision data. The distinction is discussed below.

The bias in these estimates of within site variability is unclear at present.* Certainly, within site variability is influenced by geological conditions and by installation procedures. It may or may not be that pile test results reported in the literature, in which multiple tests have been conducted at a single site, come disproportionately from geologically uniform sites. It is possible that a bias, should it exist, be exactly the opposite. However, load tests that are reported have been made on carefully installed test piles rather than production piles. It seems likely that variability among the latter is greater than among the former. Nevertheless, inspection of Figure 2 shows both the absolute within site variability and its range to be large.

The equivalent prior sample size on the mean, $n'$, was estimated by inverting the logic leading to equation (10). The global data compounds uncertainty on the mean with within site variability, and could be thought of as a predictive distribution before any load tests have been made. Therefore, given the Normal-Gamma prior, the global prior distribution should be Student $t$ with parameters $\mu'$, $\nu'$, and $H = n'/[(n'+1)v']$, and have variance

$$V[r] = \frac{H^{-1}\nu'}{\nu' - 2} \tag{16}$$

Equating this to the empirical variance 0·032, gives $n' = 2{\cdot}35$ for the broad distribution on $h$, and $n' = 1{\cdot}55$ for the tighter distribution.

These prior equivalent sample sizes are small but reasonable. When the variance of the prior global prediction is decomposed into a much smaller within site variance and a variance on the mean (among sites), the uncertainty on the mean must become large if the prior global variance is not to change.

When a pile load test is made and the distribution of $\mu$ and $h$ updated through equation (3), the posterior mean $\mu''$ increases or decreases depending on whether $\bar{r}$ is greater or less than the prior mean $\mu'$. Similarly, the posterior variance on $r$ decreases if the results of the tests are about those expected, and increases if they are different from those expected. For example, if one test is made and the result is, say, $r = 3$, both the posterior mean and variance increase. The design factor of safety to ensure a given $\beta$ may thus increase or decrease, depending on the test outcomes.

* The authors wish to acknowledge the helpful comments of the referees on this point.



Figure 5 shows design factors of safety on the Janbu formula prediction for various $\beta$'s, as a function of the outcome of a single load test using the broad prior distribution on $h$ of Figure 2. For weaker test results than expected (i.e., $r < 1$) the design factor of safety increases from that required before testing. However, for very much stronger test results than expected (e.g., $r = 3$) the design factor of safety also increases! This would not seem to make engineering sense, but it does make statistical sense. Such results imply that the pile capacity formula is much less precise than expected and therefore not as much faith should be placed in its

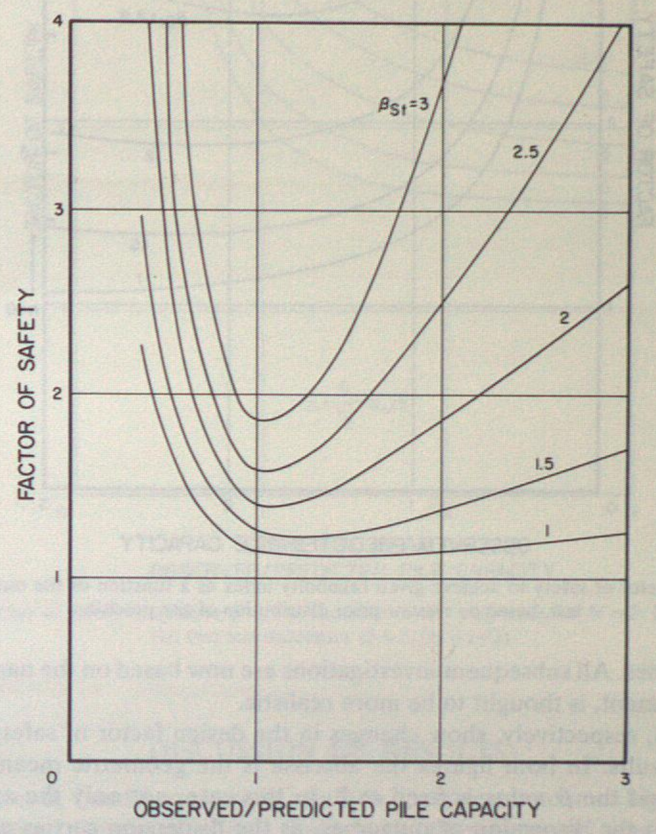


Figure 5. Required factor of safety to achieve given reliability index as a function of the outcome of one pile load test, based on broad prior distribution on site precision

predictions. Since the prior distribution on within site variance is broad, however, the individual load test result is not sufficiently informative to conclude that at this particular site the mean pile capacity will be as high as indicated by the test result. Therefore, the design factor of safety must increase.

The present data, however, are not yet sufficient to support this large range of within site variability, indicated by the broad pdf in Figure 2. In order to investigate the effect of a less broad distribution of within site variability, the three high precision results of Figure 2 were taken to be 'outliers', and the narrower distribution of site precision adopted. Figure 6 shows the parallel results using this distribution. The within site variance in this case is sufficiently narrowly distributed that the result of each load test does in fact carry substantial information

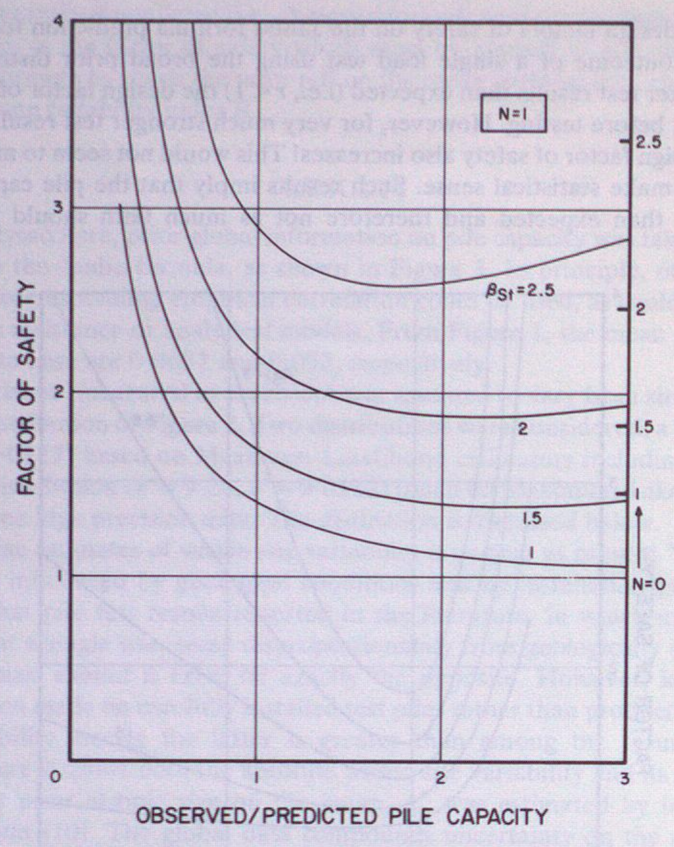


Figure 6. Required factor of safety to achieve given reliability index as a function of the outcome of one pile load test, based on narrow prior distribution of site precision

on the site properties. All subsequent investigations are now based on the narrower distribution which, for the moment, is thought to be more realistic.

Figures 7 and 8, respectively, show changes in the design factor of safety for 2 and 3 load tests and their results. In both figures the abscissa is the geometric mean of the arithmetic outcome ratios, and the $\beta$ value is fixed at 2. In this case, not only the average outcome is important but also the dispersion of outcomes, as the dispersion carries evidence about the within site variance. For convenience, this dispersion is expressed as the ratio of test outcomes in Figure 7, $\Delta = \sqrt{(2)}s$, and as the sum of squared deviations of the log ratios in Figure 8, $s^2 = \Sigma(r_i - \bar{r})^2$. Qualitatively the results are similar to those of Figure 6, with the modification that as the discrepancy among test results increases, so must $F$ in order to maintain $\beta = 2$.

Figure 9 shows a comparison of the factors of safety required to achieve given $\beta$ values for four pile capacity formulae. The *a priori* calibrations (i.e., the equivalent of Figure 1) for the Engineering News formula has been taken as $\mu' = -0{\cdot}152$, $n' = 0{\cdot}145$ (14); for the Meyerhof method as $\mu' = 0{\cdot}0004$, $n' = 0{\cdot}354$ (17); and for Rollberg's method as $\mu' = -0{\cdot}0016$, $n' = 2{\cdot}04$ (17). These curves do not correct for calibration bias. For example, although Rollberg's method appears to be a more precise (lower variance) predictor of pile capacity than Janbu's formula, because it *systematically* predicts higher capacities than Janbu's formula does, the design factors of safety must by comparison be larger. Thus, for $r = 1$ the two $F$'s would be approximately the same even though the uncertainty with Rollberg's method is smaller.



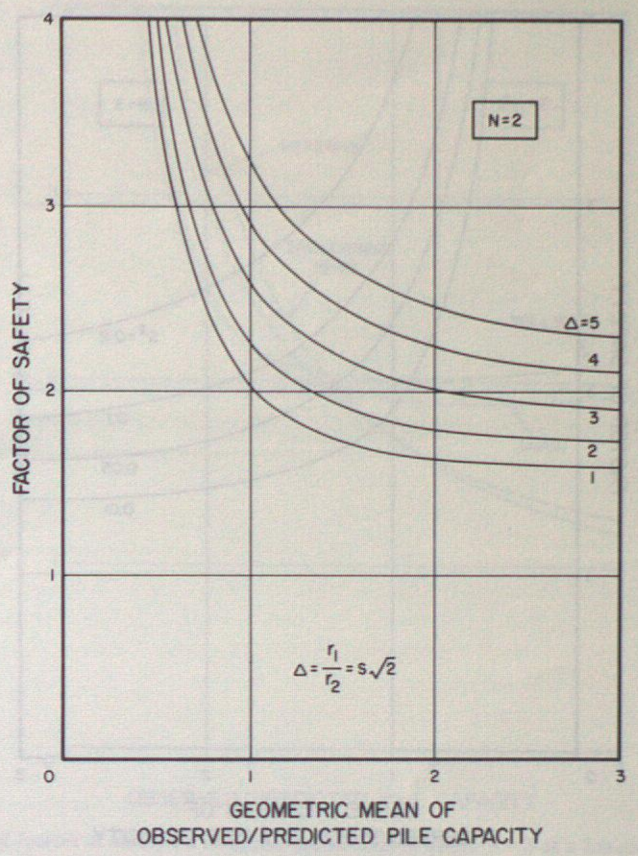


Figure 7. Required factor of safety to achieve $\beta = 2$ as function of test outcomes for $N = 2$. Factor $\Delta$ is the ratio of the two test outcomes ($\Delta = r_1/r_2 = s\sqrt{2}$)

## DISCUSSION OF RESULTS

The results of Figures 6, 7 and 8 conform to intuition. Load test outcomes higher than expected tend to lower required factors of safety; those lower than expected tend to raise required factors of safety. Greater dispersion of test results also raises required factors of safety, because it increases uncertainty on the mean and also increases the estimate of within site variance. However, until more data become available on within site variability which confirm the narrow distribution in Figure 2, the results of Figure 5 cannot be overlooked. Specifically, these results imply that if the test outcome is very much greater than expected one may have to increase the factor of safety placed on formulae predictions to achieve fixed levels of safety as reflected in $\beta$. This can even be seen at the higher values of $\beta$ and $\bar{r}$ in Figure 6.

For comparison, required factors of safety without load test results are shown along the ordinate of Figure 6. Again, typically, any test outcome of the magnitude that might have been expected reduces the required $F$ to be placed on the Janbu formula prediction.

An interesting comparison is the design factors of safety suggested by Figures 6, 7 and 8 with those now specified by DIN 1054 and 4026, the German standards for driven pile design. DIN 1054 and 4026 rely heavily on load tests, and allow dynamic formulae only when verified

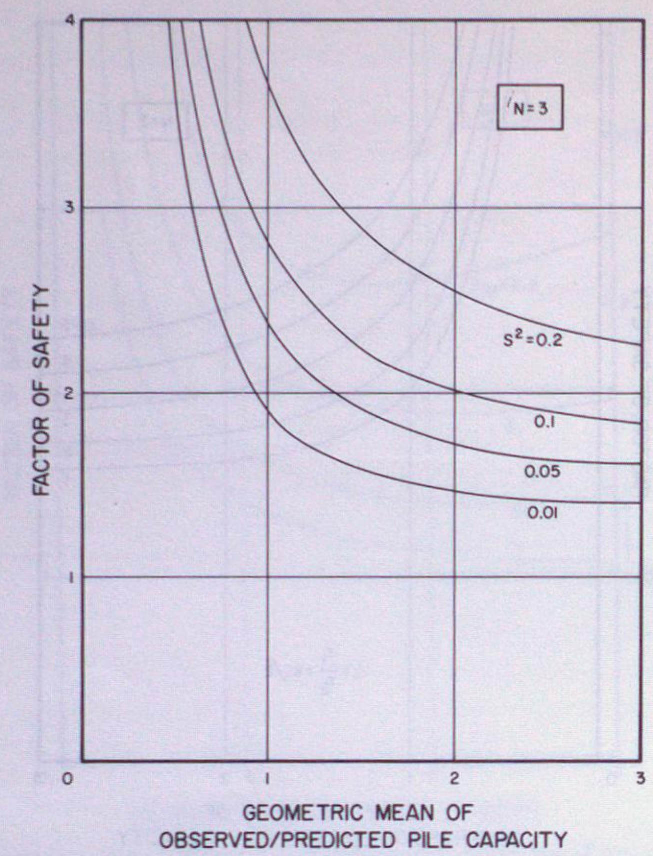


Figure 8. Required factor of safety to achieve $\beta = 2$ as function of test outcomes for $N = 3$

by on site tests. These factors of safety are shown in Table I. Also shown are the required $F$'s from Figures 6, 7 and 8 for all $r = 1$, and for various $\beta$'s.

Table I. Design factors of safety specified by DIN 4026 compared with the statistical analysis

| Number of load tests | DIN | Calculated $\beta = 2$ | 1·75 | 1·5 | 1·0 | Kay 2 |
|---|---|---|---|---|---|---|
| 1 | 2·0 | 2·05 | 1·84 | | 1·4 | 2·08 (2·18) |
| 2 | 1·75 | 1·95 | 1·75 | 1·67 | | 1·99 (2·06) |

While the internal consistency of the recommended factors of safety is reassuring, the low $\beta$ values that would seem implied are not. $\beta$ values common in the structures area have been shown to be in the order of 3 to 4. The probability of a deterministic design load exceeding pile capacity at $\beta = 2$ is about 2 per cent. Of course, group effects and conservatism in estimating the design load reduce this probability substantially.



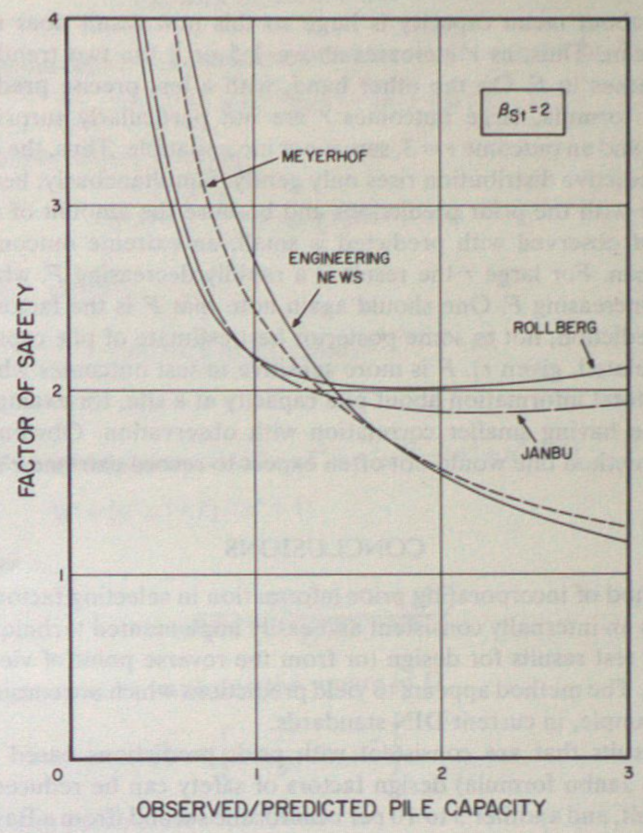


Figure 9. Comparison of factors of safety on formulae predictions to attain $\beta = 2$ as a function of the outcome of a single load test (observed/predicted)

In comparison with Kay's[7] results, the present analysis leads to lower required factors of safety for $r > 1$, and somewhat higher $F_s$ for $r = 1$. The main reasons for this agreement would seem to be that Kay's point estimate of site variability ($h = 69$) is consistent with the narrower distribution of Figure 5, and that the increase in $F$ due to Kay's larger prior uncertainty on $\mu$ is balanced by the decrease caused by his use of $\beta$ as defined in equation (13) rather than that of equation (14). Therefore, the agreement is somewhat illusory.

Again, the major limitation of the present analysis is lack of data on within site variability. The 13 or 16 data of Figure 2 yield an inaccurate estimate of the prior distribution of $h$. Not only from the present point of view, but also from that of analysing differential loads and displacements across a piled foundation, future collection of data on within site variability is important. Because the present data are collected from 'test' rather than 'production' piles, actual within site variability may be greater.

The comparisons of Figure 9 are at first surprising, and more easily understood by considering changes in the mean and standard deviation of predicted pile capacity as a function of a test outcome $r$. With the more precise prediction of, say, the Rollberg or Janbu formulae large test outcomes $r$ are surprising because they are incompatible with the rather precise *a priori* predictions. Thus, the standard deviation of the predictive pdf of pile capacity increases sharply when $r$ diverges significantly from that predicted. Simultaneously, however, the amount of prior information about mean capacity is large so this new result does not lead to strong updating of the mean. Thus, as $r$ increases above 1·5 or 2 the two trends balance or even lead to slight increases in $F$. On the other hand, with a less precise prediction of, say, the Engineering News formula, large outcomes $r$ are not particularly surprising. The *a priori* prediction is broad and an outcome $r = 3$, say, is not incompatible. Thus, the standard deviation of the posterior predictive distribution rises only gently. Simultaneously, because the outcome is not incompatible with the prior predictions and because the amount of information in the prior correlation of observed with predicted is small, an extreme outcome leads to strong updating of the mean. For large $r$ the result is a rapidly decreasing $F$, while for small $r$ the result is a rapidly increasing $F$. One should again note that $F$ is the factor of safety applied to the formulae prediction, not to some posterior best estimate of pile capacity (although the two are uniquely related, given $r$). $F$ is more sensitive to test outcomes when they constitute a large part of the total information about pile capacity at a site, for example when using the prediction formulae having smaller correlation with observation. Obviously, when using a precise prediction method one would not often expect to record extreme $r$'s.

## CONCLUSIONS

The proposed method of incorporating prior information in selecting factors of safety for pile load capacity offers an internally consistent and easily implemented technique for interpreting the implications of test results for design (or from the reverse point of view, for judging the worth of load tests). The method appears to yield predictions which are consistent with practice, as reflected, for example, in current DIN standards.

For load test results that are consistent with prior predictions based on good dynamic formulae (here the Janbu formula) design factors of safety can be reduced 5 to 10 per cent for the first test result, and another 5 to 10 per cent for the second (from a Bayesian perspective, the expected reduction from the first test is 4·8 per cent). However, the numerical results reported here depend on an assessment of within site variability based on a small number of test sites. Further rationalization of the uncertainties in predictions of pile performance can only follow the development of better empirical data on variability. This is true whether or not statistical techniques are used.

The procedure illustrates that for certain classes of geotechnical problems, particularly those for which prototype testing is possible, statistical methods provide a logical framework within which to evaluate field experimental data. Furthermore, these methods do not require arbitrary assumptions on distributional forms, inferences dependent on highly sensitive tail probabilities, subjective prior probabilities, or esoteric stochastic models. They are, rather, statistical procedures of precisely the same sort as widely and successfully used in a number of other disciplines.

## APPENDIX I: DERIVATION OF PREDICTIVE DISTRIBUTION OF PILE CAPACITY (EQUATION 9)

The predictive (or marginal) probability distribution of the logarithm of the ratio of observed to predicted pile capacity ($R$) is found by integrating over the joint distribution of $R$ and the parameters $(\mu, h)$.

From equation (9)

$$f''(\mu, h|\mathbf{r}) \propto \exp\{-\tfrac{1}{2}hn''(\mu-\mu'')^2\}h^{1/2}\exp\{-\tfrac{1}{2}h\nu''v''\}h^{1/2(v''-1)} \qquad \text{(A1)}$$



which is Normal-Gamma. Since the conditional distribution of $r$ given $(\mu, h)$ is Normal

$$f(r|\mu, h) \propto h^{1/2}\exp\{-\tfrac{1}{2}h(r-\mu)^2\} \qquad \text{(A2)}$$

the predictive distribution over $r$ becomes

$$f(r) \propto \int_\mu \int_h f(r|\mu, h) f(\mu, h|\mathbf{r})\, \mathrm{d}\mu\, \mathrm{d}h \qquad \text{(A3)}$$

Integrating first over $\mu$ yields

$$f(r) \propto \int_h \exp\{-\tfrac{1}{2}hQ\}h^{1/2(\nu''-1)}\, \mathrm{d}h \qquad \text{(A4)}$$

where

$$Q = [v'' + (n-1)S^2 + n\bar{r}^2 + n''\mu''^2 - (n''+n)W^2] \qquad \text{(A5)}$$

$$W = (n''\mu'' + \bar{r})/(n''+1) \qquad \text{(A6)}$$

The integrating over $h$,

$$f(R) \propto \frac{\Gamma[(1/2\nu'' - 1/2)]}{Q^{(1/2\nu''-1/2)}} \qquad \text{(A7)}$$

Inserting the constants and completing the square in $D$

$$f(r) \propto \left[1 + \frac{H}{\nu''}(r-\mu'')\right]^{-1/2(\nu''-1)} \qquad \text{(A8)}$$

which is Student-$t$ with parameters $\nu''$, $\mu''$, and

$$H = \left(\frac{n''}{n''+1}\right)\frac{1}{v''} \qquad \text{(A9)}$$

## APPENDIX II: LIST OF SYMBOLS

| | |
|---|---|
| $D$ | data |
| $E(.)$ | expectation |
| $f'(.)$ | prior probability density function (pdf) |
| $f''(.)$ | posterior pdf |
| $h$ | precision $(1/\sigma^2)$ |
| $H$ | parameter of Student $t$ pdf |
| $L(.)$ | likelihood function |
| $N(.)$ | Normal distribution |
| $n$ | number of pile load tests |
| $n'$ | equivalent prior number of tests for distribution on mean pile capacity |
| $n''$ | equivalent posterior number of tests |
| $r$ | test result, $\log_{10}$ (observed/predicted pile capacity) |
| $\mathbf{r}$ | vector of test results |
| $\bar{r}$ | average of test results, $\mathbf{r}$ |
| $r_0$ | design value of $r$ |
| $R$ | generic variable $\log_{10}$ (observed/predicted pile capacity) |
| $s^2$ | sum of squared deviations from mean of test results |

| | |
|---|---|
| $T_{\nu}^{-1}$ | inverse Student $t$ distribution |
| $v', v''$ | prior and posterior parameter of the distribution of within site precision |
| $V(.)$ | variance |
| $\mathbf{y}$ | predicted variable |
| $\beta_{FOSM}$ | first-order second-moment reliability index |
| $\beta_N$ | standardized Normal variate and reliability index |
| $\beta_{St}$ | standardized Student $t$ variate and reliability index |
| $\Gamma(.)$ | Gamma function |
| $\theta$ | parameters of statistical model |
| $\mu, \mu', \mu''$ | mean pile capacity, prior mean, posterior mean |
| $\nu', \nu''$ | parameter of prior and posterior pdf on $h$ |
| $\sigma^2$ | within site variance |
| $\Phi(.)$ | cumulative Normal distribution |

## REFERENCES


1. Comité Europeen du Béton, *Bulletin D'Information No. 112*, Joint Committee on Structural Safety, 1976.
2. DIN 4026: Rammpfähle, Aug. 1975; DIN 1054: Zulässige Belastungen des Baugrunds, Nov. 1976.
3. A. Evangelista, A. Pelligrino and C. Viggiani, 'Variability among piles of the same foundation', *Proc. 9th International Conference on Soil Mechanics and Foundation Eng.* Tokyo (1977).
4. G. G. Goble, F. K. Walker and F. Rausche, 'Pile bearing capacity prediction versus performance', *ASCE Speciality Conference on Performance of Earth and Earth Supported Structure* (1972).
5. W. S. Housel, 'Pile load capacity: estimates and test results', *J. Soil Mech. Found. Div., Proc. Am. Soc. Civ. Eng.* **92**, 1–30 (1966).
6. H. Jeffreys, *Theory of Probability*, The Clarendon Press, Oxford, 1961.
7. J. N. Kay, 'Safety factor evaluation for single piles in sand, *J. Geotech. Engng Div., Proc. Am. Soc. Civ. Eng.* **102**, 1093–1108 (1976).
8. J. N. Kay, 'Factor of safety for piles in cohesive soils', *Proc. 9th International Conference on Soil Mechanics and Foundation Engineering*, Tokyo (1977).
9. G. König, 'Die Tragfähigkeit von Pfählen', *Bauingenieur*, **45**, 23–29 (1970).
10. P. K. K. Leung, 'Modeling uncertainty in geotechnical predictions', *Thesis* presented to the Massachusetts Institute of Technology in partial fulfillment of the requirements for the degree Master of Science (1977).
11. M. R. Madhav and A. Arumugam, 'Pile capacity—a reliability approach', *Proc. 3rd International Conference on Applications of Statistics and Probability to Soil and Structural Engineering*, Sydney (1979).
12. G. G. Meyerhof, 'Bearing capacity and settlement of pile foundation', *ASCE J. Geotech. Engng Div., Proc. Am. Soc. Civ. Eng.* **102**, 195–228 (1976).
13. G. G. Meyerhof, 'Discussion to safety factor evaluation for single piles in sand, by J. N. Kay', *J. Geotech. Engng Div., Proc. Am. Soc. Civ. Eng.* **103**, 362 (1977).
14. R. E. Olsen and K. S. Flaate, 'Pile driving formulas for friction piles in sand', *J. Soil Mech. Found. Div., Proc. Am. Soc. Civ. Eng.* **93**, 279–296 (1967).
15. H. Petermann, E. Lackner and W. Schenek, 'Tragfähigkeit von Pfählen, Groβversuche und ihre Auswertung', *Berichte aus der Bauforschung*, H. 49, W. Ernst & Sohn, Berlin, 1958/59.
16. H. Raiffa and R. L. Schlaifer, *Applied Statistical Decision Theory*, Harvard University, School of Business Administration Research Division, Boston, 1961.
17. D. Rollberg, 'Bestimmung der Tragfähigkeit und des Rammwiderstandes von Pfählen und Sondierungen', *Veröffentlichungen des Institutes für Grundbau, Bodenmechanik, Felsmechanik und Verkehrswasserbau der RWTH Aachen*, Heft 3, 1977.
18. F. Tavernas and R. Andy, 'Limitations of the driving formulae for predicting the bearing capacities of piles in sand, *Can. Geotech. J.* **9**, 47–62 (1972).
19. F. A. Tavernas, 'Load test results on friction piles in sand', *Can. Geotech. J.* **8**, 7–22 (1971).
20. S. P. Wagner and R. J. Krizele, 'Statistical Optimization of friction pile foundations', *Proc. 2nd International Conference on Applications of Statistics and Probability to Soil and Structural Engineering*, Aachen (1975).
21. 'Grundlagen für die Festlegung von Sicherheitsanforderungen für bauliche Anlagen', *Institut für Bautechnik*, Entwurf, Berlin, 1977.
22. *General Principles on Reliability for Structural Design*, Joint Committee for Structural Safety, Lind, 1979.