Probabilistic identification of soil stratification

J. LI*, M. J. CASSIDY†, J. HUANG‡, L. ZHANG§ and R. KELLY‡

Identification of soil stratification is vital to geotechnical structural design and construction where the soil layer, soil type and properties are necessary inputs. Although methods are available for classifying the soil profile using measured cone penetration test (CPT) data, the identification of soil stratification at unsampled locations is still difficult due to significant variability of natural soil. The identification is further complicated by the considerable uncertainties in the CPT measurements and soil classification methods. This study aims to develop a probabilistic method to predict soil stratification at unsampled locations by explicitly filtering the uncertainties in soil classification systems. An established Kriging interpolation technique is used to estimate the CPT parameters which are further interpreted to identify the soil stratification. Equations are derived to quantify the degree of uncertainties reduced by this method. The approaches are illustrated using a database of 26 CPT tests recently sourced from a dike near Ballina, Australia. Results show that the majority of the uncertainties in the soil parameters are screened by a soil classification index. The remaining uncertainties are further filtered by the soil classification systems. A clear stratification with a high degree of confidence is obtained in both horizontal plane and vertical unsampled locations, which shows excellent agreement with the existing CPT tests. This study provides a methodology to clearly identify the soil strata and reduce the uncertainties in prediction of design properties, paving the way for a more cost-effective geotechnical design.

KEYWORDS: geology; site investigation; soil classification; statistical analysis

INTRODUCTION

Soil stratification is essential in geotechnical site characterisation and structural design (Houlsby & Houlsby, 2013; Wang et al., 2014). Recent studies have reported the significant effect of soil stratification on the design of foundations, tunneling and pipelines (Burd & Frydman, 1997; Padrón et al., 2008; Huang & Griffiths, 2010; Zhang et al., 2012; Lee et al., 2013; White et al., 2014). The identification of soil stratification includes determining soil types, the number of soil layers, the thickness of each layer and soil properties.

The standardised cone penetration test (CPT) and piezocene (CPTU) have been widely used to infer the soil type by soil layers, the thickness of each layer and soil properties. Although methods are available for classifying the soil profiling using measured cone penetration test (CPT) data, the identification of soil stratification at unsampled locations is still difficult due to significant variability of natural soil. The identification is further complicated by the considerable uncertainties in the CPT measurements and soil classification methods. This study aims to develop a probabilistic method to predict soil stratification at unsampled locations by explicitly filtering the uncertainties in soil classification systems. An established Kriging interpolation technique is used to estimate the CPT parameters which are further interpreted to identify the soil stratification. Equations are derived to quantify the degree of uncertainties reduced by this method. The approaches are illustrated using a database of 26 CPT tests recently sourced from a dike near Ballina, Australia. Results show that the majority of the uncertainties in the soil parameters are screened by a soil classification index. The remaining uncertainties are further filtered by the soil classification systems. A clear stratification with a high degree of confidence is obtained in both horizontal plane and vertical unsampled locations, which shows excellent agreement with the existing CPT tests. This study provides a methodology to clearly identify the soil strata and reduce the uncertainties in prediction of design properties, paving the way for a more cost-effective geotechnical design.

KEYWORDS: geology; site investigation; soil classification; statistical analysis
estimation to the soil stratification are then significantly reduced through a soil classification system. Ultimately, a clear soil stratification in three dimensions is identified with a limited number of CPTU tests.

**KRIGING ESTIMATION AT UNSAMPLED LOCATIONS**

Kriging is an advanced interpolation procedure that uses the correlation among existing data (Matheron, 1971). The correlation is often described by a semivariogram, the value of which for separation distance \( h \) is calculated by the average squared difference in property values between pairs of input sample points separated by \( h \)

\[
\gamma_h = \frac{1}{2} E \left[ (p_i - p_{i+h})^2 \right]
\]

(1)

where \( \gamma_h \) is the semivariogram value for the data pairs with property values of \( p_i \) and \( p_{i+h} \). In geotechnical practice, CPT locations are often irregularly spaced, requiring the distance and directions to be grouped into subsets (Olea, 1999; Dasaka & Zhang, 2012) and for the semivariogram to be evaluated based on the subsets. To evaluate the spatial variability using Kriging techniques it is essential that the data are stationary; that is, the mean and covariance of the data depend only upon separation, not on absolute location.

If the semivariogram does not level off for long separation distances, this indicates that the data set is non-stationary (Kulatilake & Ghosh, 1988). If the data are non-stationary, treatment must be given to transform the data to a stationary set by removing the deterministic component called the trend, and the stationary residual random component is then analysed.

The properties at an unsampled location can be estimated using a linear weighting of all measured data within the effective domain around the location

\[
\hat{p}_0 = \sum_{i=1}^{N} \lambda_i p_i
\]

(2)

where \( \lambda_i \) is the weight for the measured property \( p_i \), and \( \hat{p}_0 \) is the estimated property at the unsampled location. Kriging uses the semivariogram to assign a weight to each measured data point. To ensure the estimated value is unbiased, the sum of the weights \( \lambda_i \) must equal to one. The weights are selected to minimise the expected mean squared error between the estimation and the true value. The solution to the minimisation, constrained by the unbiased estimation, gives the Kriging equations (Murakami et al., 2006)

\[
\begin{bmatrix}
\gamma_{11} & \cdots & \gamma_{1N} & \gamma_{1N} \gamma_{2N} & \cdots & \gamma_{1N} & \gamma_{1N} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\gamma_{N1} & \cdots & \gamma_{NN} & \gamma_{N1} & \cdots & \gamma_{N1} & \gamma_{NN} \\
1 & \cdots & 1 & 0 & m & 1 & 1
\end{bmatrix}
\begin{bmatrix}
\lambda_1 \\
\vdots \\
\lambda_N
\end{bmatrix}
= \begin{bmatrix}
\gamma_0 \\
\vdots \\
\gamma_{0N} \\
1
\end{bmatrix}
\]

(3)

where \( \gamma_0 \) denotes the modelled semivariogram values based on the distance between the two samples pertaining to the \( i \)th and \( j \)th locations. The unknown \( m \) is a Lagrange multiplier to ensure the unbiased constraint.

The semivariogram can be evaluated using nugget (i.e. a small offset of the semivariogram when the separation is zero), range (i.e. the distance where the model first flattens out) and sill (i.e. the value that the semivariogram model attains at the range). If a high-quality semivariogram with a small nugget/sill ratio and a large range is obtained, Kriging can be performed with only a few data (Jernigan, 1986). An important parameter to quantify the accuracy of the estimation is the Kriging variance \( \sigma^2 \)

\[
\sigma^2 = E(\hat{p}_0 - p_0)^2
\]

(4)

where \( \sigma \) is the Kriging standard error. This Kriging variance depends only on the locations of the data but not on their values. To measure the relative variation of the estimation, a Kriging variation is defined as

\[
e = \frac{\sigma}{\hat{p}_0}
\]

(5)

The quality of semivariogram affects the Kriging variation, an indicator measuring the estimation accuracy. Both the quality of semivariogram and the Kriging variation are dependent on the following aspects.

(a) CPT plan and data quality. The distances between the CPTs and the number of CPTs are essential to the estimation. If the distances between the CPTs are larger than the correlation length, the correlation structure cannot be obtained properly, which directly influences the accuracy of the estimation. High-quality data that lead to a small nugget and a large range in the semivariogram maintain small measurement errors and give better estimation.

(b) Correlation of existing data. Good correlation of the existing data results in a good estimation of the semivariogram, which, in turn, leads to a small Kriging variation.

(c) Distance between the location of interest and existing CPTs. As the estimation relies more heavily on neighbouring CPTs, the accuracy is better in the vicinity of existing tests.

(d) Estimated value of the property. Kriging variance is dependent on location and correlation structure, which are not related to the estimated value of the property. However, the Kriging variation is a value relative to the estimated value, which depends on the estimated value. If the estimated value is relatively small, its Kriging variation tends to be large.

The range within which the true value of the estimated property lies with a prescribed probability is often of practical interest. This range (with lower and upper limits) can be defined by assuming a normally distributed parameter with a mean value of \( \hat{p}_0 \) and a variance of \( \sigma^2 \) (Ang & Tang, 2007)

\[
(p_0)^{1-\alpha} = (\hat{p}_0 + k_{a/2}\sigma; \hat{p}_0 + k_{(1-a)/2}\sigma)
\]

(6)

in which \( (p_0)^{1-\alpha} \) is the \( (1-\alpha) \) confidence interval for the estimation, \( k_{a/2} = -\Phi^{-1}(1-\alpha/2) \) and \( k_{(1-a)/2} = \Phi^{-1}(1-\alpha/2) \). \( \Phi \) is the cumulative distribution function of the standard normal variate. \( \alpha \) is a parameter to specify the probability. Particularly, the 95% confidence interval (with \( \alpha \) equal to 0.05) of the estimation is \( (\hat{p}_0 - 1.96\sigma; \hat{p}_0 + 1.96\sigma) \).

**UNCERTAINTY IN SOIL CLASSIFICATION BASED ON CPTU TESTS**

The soil classification method proposed by Robertson (1990) is one of the most widely used methods and adopts three parameters based on CPTU tests

(a) normalised cone resistance

\[
Q_1 = \frac{q_d - \sigma_{qD}}{\sigma_{qD}}
\]

(7)
where \( q \) is total overburden stress, \( \sigma_{o} \) is effective overburden stress, \( f_s \) is sleeve friction, \( u \) is pore pressure measured between the cone tip and the friction sleeve, \( u_0 \) is equilibrium pore pressure, \( q_t \) is corrected cone penetration resistance to account for unequal end area effects

\[
q_t = q_c + (1 - a)u
\]

where \( q_c \) is measured cone penetration resistance, \( a \) is the net area ratio between load cell support diameter, \( d \), and cone diameter, \( D \)

\[
a = \frac{d^2}{D^2}
\]

A soil classification chart using these three parameters was proposed by Robertson (1990). This method was further improved by Jefferies & Davies (1993) to make the chart more amendable to spreadsheet analysis by introducing a soil classification index, \( I_c \), to define the boundary of soil type

\[
I_c = \sqrt{\left[ 3 - \log \left( \frac{Q_t}{1 - B_q} \right) \right]^2 + \left[ 1.5 + 1.3 \log (F_R) \right]^2}
\]

Within this classification, soil types are attributed from \( I_c \) as summarised in Table 1. This CPT classification method may not provide accurate predictions of soil type based on grain size distribution but provides a guide to soil behaviour type (Lunne et al., 1997). This classification inevitably contains uncertainties, such as measurement errors and transformation uncertainty. Experience has shown that measurement errors are mainly from the measurement of cone penetration resistance \( (q) \), sleeve friction \( (f_s) \) and pore pressure \( (u) \).

In the following sections, the cone penetration resistance, sleeve friction and pore pressure are interpolated separately at locations of interest, which quantify the variation in each parameter. The soil types are examined using the soil classification method proposed by Jefferies & Davies (1993) based on inferred parameters. The transformation uncertainty can then be expressed in an explicit manner. In the next section, equations will be derived to quantify the uncertainties that are filtered by soil classification.

**UNCERTAINTY FILTERING THROUGH SOIL CLASSIFICATION**

The change in soil classification index, \( \Delta I_c \), can be expressed in terms of changes in normalised cone resistance \( (\Delta Q_t) \), normalised friction ratio \( (\Delta F_R) \) and pore pressure ratio \( (\Delta B_q) \)

\[
\Delta I_c = \frac{\partial I_c}{\partial Q_t} \Delta Q_t + \frac{\partial I_c}{\partial F_R} \Delta F_R + \frac{\partial I_c}{\partial B_q} \Delta B_q
\]

in which

\[
\frac{\partial I_c}{\partial Q_t} = \left( \frac{3 - \log \left( \frac{Q_t}{1 - B_q} \right)}{1 - \frac{Q_t}{\ln 10}} \right)^{1/2} \times \left( \frac{3 - \log \left( \frac{Q_t}{1 - B_q} \right)}{1 - \frac{Q_t}{\ln 10}} \right)^{1/2} \times \left[ 1.5 + 1.3 \log (F_R) \right]
\]

\[
\frac{\partial I_c}{\partial F_R} = \left( \frac{3 - \log \left( \frac{Q_t}{1 - B_q} \right)}{1 - \frac{Q_t}{\ln 10}} \right)^{1/2} \times \left[ 1.5 + 1.3 \log (F_R) \right] \frac{1.3}{F_R \ln 10}
\]

\[
\frac{\partial I_c}{\partial B_q} = \left( \frac{3 - \log \left( \frac{Q_t}{1 - B_q} \right)}{1 - \frac{Q_t}{\ln 10}} \right)^{1/2} \times \left( \frac{3 - \log \left( \frac{Q_t}{1 - B_q} \right)}{1 - \frac{Q_t}{\ln 10}} \right)^{1/2} \times \left[ 1.5 + 1.3 \log (F_R) \right] \frac{1}{(1 - B_q)}
\]

The change in soil classification index with the only change in normalised cone resistance (assuming the changes in \( F_R \) and \( B_q \) are zero) can then be expressed as

\[
\frac{\Delta I_c}{I_c} = \frac{\log \left( \frac{Q_t}{1 - B_q} \right) - 3}{I_c^2 \ln 10}
\]

Similarly, the change in soil classification index with the only change in normalised friction ratio or pore pressure ratio can be expressed by equations (18) and (19)

\[
\frac{\Delta I_c}{I_c} = \frac{1.3 - 1.5 \log (F_R)}{I_c^2 \ln 10}
\]

\[
\frac{\Delta I_c}{I_c} = \frac{B_q \left( 3 - \log \left( \frac{Q_t}{1 - B_q} \right) \right)}{(1 - B_q) I_c^2 \ln 10}
\]

Equations (17)–(19) quantify the relative change of soil classification index with a relative change in cone penetration resistance, sleeve friction or pore pressure. A ratio smaller than 1 indicates that the magnitude of change in the soil classification system is smaller than the magnitude of change in the measured \( q_c, f_s \) and \( u \), which means that uncertainties in the predictions of these parameters are filtered by the soil classification index as described by equation (12). Furthermore, the uncertainty in the soil classification index can be filtered by the soil classification system, as indicated by Table 1. For example, if a variation of 5% in \( \Delta I_c/I_c \) occurs for \( I_c \) of 3.02, \( I_c \) can then vary from 2.87 to 3.17. However, any value in this range is classified as clay (see Table 1). This soil classification system further filters the uncertainty in the soil classification index.

**CASE STUDY**

In total, 26 CPTU tests (as shown in Fig. 1) have been performed on a soft clay site near the town of Ballina, New South Wales, Australia. The tests were conducted over a total area of 100 x 100 m. In each test, the cone penetration resistance, sleeve friction and pore pressure were measured.
every 0.02 m in depth. The profiles of the measured \(q_c\), \(f_s\) and \(u\) of a typical CPTU test are shown in Fig. 2. These parameters at a designated depth can be used to estimate the same parameters at unsampled locations at the same depth, which is the first step towards the identification of soil stratification at unsampled locations. The estimation of these parameters and their accuracy for a plane at 2 m depth are first detailed. The uncertainty filtering process is then performed and the soil types at the same plane are predicted. Finally, the soil profile and properties at two unsampled locations, B1 and B2 (also shown in Fig. 1), are examined. B1 is located within the sampling cluster where enough information can be used to evaluate the rationality of the prediction. B2 is far away from the existing CPTU tests and represents a typical prediction where only limited information is available in the field. The accuracy of the predictions from the two locations will be examined separately.

Prediction of CPTU parameters at unsampled locations and their accuracy

The cone penetration resistance, sleeve friction and pore pressure measured at 2 m depth of the 26 CPTU tests are demonstrated in Fig. 3. The vertical axis of the dot points gives the property values, and the vertical lines attached to the points indicate the locations of the tests. The soil at 1.82 m depth at location (14.95, 0.75) is classified as silt mixtures according to the Jefferies & Davies method (1993). Laboratory sieve analysis of the soil taken at the same location shows that the soil has 43.4% clay, 53.6% silt and 3.0% sand. Hence the soil is classified as a silt mixture, which is consistent with the classification using the Jefferies & Davies method (1993). As this method is not the focus of the present study, more discussion on its reliability can be found in, among other sources, the book by Lunne et al. (1997).

Previous researchers have reported that the range of semivariograms in the horizontal direction is larger than that in the vertical direction. The ranges in different directions in a horizontal plane are often assumed the same with an isotropic correlation structure, especially for small data sets (Jernigan, 1986). In this study no direction-dependency is observed among the data in the horizontal plane at a depth of 2 m (see Fig. 3). Hence a lateral isotropic correlation structure is assumed. Trends which are equal to the averages of the cone penetration resistance, sleeve friction and pore pressure are first removed, respectively. Then the semivariograms of the three parameters are calculated using a self-developed code according to equation (1) and plotted as shown in Fig. 4. Results show that the semivariance increases first with the separation distance and then levels off. A theoretical exponential model is used here to fit the sample semivariogram. For the cone penetration resistance and sleeve friction the model starts from zero, which indicates that the measurement error is negligible. In contrast, the nugget effect in pore pressure is much higher. This means that the measurement error in pore pressure is large, which is consistent with the field observation. The semivariograms for the cone penetration resistance and sleeve friction have small nugget/sill ratios and relatively large ranges, which indicates fairly good semivariograms. However, with the same set of
data the quality of the semivariogram for the pore pressure is unsatisfactory. The influence of the quality of the semivariogram on the estimation will be further discussed in the following sections.

The Kriging interpolation was performed, which results in the estimated $q_c$, $f_s$, and $u$ values at 2 m depth, as shown in Fig. 5. The accuracy of this estimation can be evaluated by the Kriging variation, as defined by equation (5). The Kriging variation values of $q_c$, $f_s$, and $u$ for unsampled locations at 2 m depth are shown in Fig. 6. The variations in cone penetration resistance are smaller than 0.27. In particular, the variations are extremely low at locations near the existing CPTs (see Fig. 6(a)). The variations in sleeve friction are relatively large (a maximum variation of 0.74), especially in the areas with sparse CPT tests (Fig. 6(b)). The variations in pore pressure are extremely large (a maximum variation of 11) in the areas with small pore pressures (Fig. 6(c)). Such large variation is partly due to the poor quality in the

![Fig. 3. Measured (a) cone penetration resistance ($q_c$), (b) sleeve friction ($f_s$) and (c) pore pressure ($u$) at 2 m depth](image)

![Fig. 4. Semivariogram for: (a) cone penetration resistance ($q_c$); (b) sleeve friction ($f_s$); (c) pore pressure ($u$) at 2 m depth](image)

![Fig. 5. Estimated (a) penetration resistance ($q_c$), (b) sleeve friction ($f_s$) and (c) pore pressure ($u$) at 2 m depth](image)

![Fig. 6. Kriging variation of (a) penetration resistance; (b) sleeve friction; and (c) pore pressure at 2 m depth](image)
The variations in the uncertainties in the predicted intervals of pore pressure, the quantification of the variations due to the large variations in pore pressure and the poor standard error values for pore pressure are extremely large for locations B1 and B2 are shown in Table 2. The Kriging and pore pressure. The estimators, standard errors and 95% standard error at location B2 results in a wider confidence interval of 2·86 MPa at location B1 and 0·038 MPa at location B2. A larger standard error occurs at location B2 because the CPT data around B2 are sparse. The 95% confidence intervals are (0·165, 0·238) at location B1 and (0·096, 0·246) at location B2. The larger standard error at location B2 results in a wider confidence interval at this location.

A similar procedure can be performed for sleeve friction and pore pressure. The estimators, standard errors and 95% confidence intervals (shown as CI on column headings) for locations B1 and B2 are shown in Table 2. The Kriging standard error values for pore pressure are extremely large due to the large variations in pore pressure and the poor semivariogram. Although such large variations lead to wide intervals of pore pressure, the quantification of the variations is one of the strengths of the proposed method.

### Soil type

If the soil parameters (i.e. $q_c$, $f_c$ and $u$) at a certain location are estimated, the soil type at this location can be classified using equation (12). As the soil parameters may be distributed in an interval, the soil type may also vary. An alternative soil type can be obtained using the lower bound values of the 95% confidence interval for the soil parameters at each location. Similarly, a soil type can be obtained using the upper bound of the confidence interval.

The soil classification index, $I_c$, for location B1 calculated from the estimator is 3·02. The soil classification index is 2·86 using the lower bound value of cone penetration resistance (0·165 MPa), sleeve friction (2·41 kPa) and pore pressure (−26·72 kPa). The soil classification index is 3·09 using the upper bound value of cone penetration resistance (0·238 MPa), sleeve friction (7·29 kPa) and pore pressure (64·00 kPa). Hence the uncertainties in the measured CPTU data lead to a range of soil classification index values from 2·86 to 3·09. A poor quality of data with large uncertainties will result in a wide range of the soil classification index. The variations that propagate from the estimated soil parameters to the soil classification index can be calculated using equations (17) and (18). The variations in the estimated $q_c$, $f_c$ and $u$ are 0·08, 0·26 and 1·24, respectively. The variations in $I_c$ caused by those in $q_c$, $f_c$ and $u$ are 0·009, 0·034 and 0·062, respectively. The variations in $I_c$ are only approximately 10% of the variations in $q_c$, $f_c$ and $u$, which indicates that the majority of the uncertainties in the predicted $q_c$, $f_c$ and $u$ are filtered.

The soil type can be classified by the soil classification index according to Table 1. When the soil classification index is larger than 2·82 and smaller than 3·22, the soil is classified as a clay. The soil classification index at location B1 is in the range of 2·86–3·09, which indicates that the soil is clay. In other words, with 95% probability, the soil at location B1 at 2 m depth is clay. This process shows that the relatively small variations in the soil classification index are further filtered by the soil classification systems, which leads to a consistent conclusion of clay at unsampled location B1. Although the variations in the estimated $q_c$, $f_c$ and $u$ are large, the classification of soil type consistently indicates that the soil is clay. Such uncertainty filtering gives a clear indicator of the soil type with high confidence.

At location B2, the soil classification index at 2 m depth is 3·22, which is on the boundary of clay and organic clay. The soil classification index calculated from the lower bound values of the parameters is 3·27, which indicates that the soil behaves like organic clay. The soil classification index calculated from the upper bound values is 3·17, which is classified as clay. The results indicate that the soil at location B2 at 2 m depth may situate in a transition zone from clay to organic clay. The transition may be due to

\[(a)\] soil behaviour, which is influenced by the soil ahead and behind the cone tip (Ahmadi & Robertson, 2005)

\[(b)\] the variation in soil classification (e.g. the variations in the soil classification index, $I_c$, caused by variations in $q_c$, $f_c$ and $u$ are 0·02, 0·04 and 0·02, respectively; only limited influence is caused by such small variations)

\[(c)\] soil located in a transition zone that consists of mixtures of clay and organic clay. The presence of a transition zone can be further examined by the neighbouring CPTs around location B2. The nearest neighbouring CPT tests are NS12, NS28 and NS30. The soils at 2 m depth at NS12, NS28 and NS30 are clay, organic clay and organic clay, respectively, which confirms that B2 is located in a transition zone from clay to organic clay from south to north.

### Soil stratification

#### Soil stratification in the horizontal direction.

Soil stratification in the horizontal direction can be examined by exploring the soil types at a horizontal plane. The soil types of the whole site at 2 m depth are calculated for each small zone (1·5 x 1·5 m) using the aforementioned method. Fig. 7(a) shows a soil type map that is classified based on the estimated cone penetration resistance, sleeve friction and pore pressure. Organic clay behaviour is present at the west of this area, which changes to clay towards the east and further to silt mixtures at the southeast corner of the site. The soil at location (14·95, 0·75) is classified as clay in this predicted soil type map. The particle size distribution at this location at 2·27 m depth shows that there is 54·7% of clay, 41·3% of silt and 4% of sand. Therefore, the soil particle size distribution also indicates that this is fine-grained soil (i.e. silty clay).

Different soil types may be present due to variations in the estimated soil parameters. For example, the soil types classified using the lower bound and upper bound values of the 95% confidence interval are not exactly the same as those in Fig. 7(a). Therefore, three soil types at each location are predicted using the estimator, the lower bound and the upper bound of the soil parameters, respectively. By randomly

---

**Table 2. Predicted value, standard error and 95% confidence interval of $q_c$, $f_c$ and $u$ at locations B1 and B2**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Location B1</th>
<th>Location B2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimator</td>
<td>Standard error</td>
</tr>
<tr>
<td>Cone penetration resistance, $q_c$</td>
<td>0·202</td>
<td>0·019</td>
</tr>
<tr>
<td>Sleeve friction, $f_c$, kPa</td>
<td>4·85</td>
<td>1·24</td>
</tr>
<tr>
<td>Pore pressure, $u$, kPa</td>
<td>18·64</td>
<td>23·15</td>
</tr>
</tbody>
</table>

Downloaded by [University Of Western Australia] on [14/11/16]. Copyright © ICE Publishing, all rights reserved.
Sampling the soil type from the three classifications that reflect the variation in the soil parameters, different soil type maps can be obtained. Figs 7(b)–7(f) show five random samples, which demonstrate the possible presence of soft, sensitive clay, silt mixture and sand mixture in the northeast of the site. The soil type maps show a consistent transition trend, despite the large variations in the estimated $q_c$, $f_s$ and $u$. The boundary between different soil types is vague when considering the variation in estimation.

Soil profile in the vertical direction. The soil profile in the vertical direction is examined at locations B1 and B2 in this section. Values of the cone penetration resistance, sleeve friction and pore pressure at 2, 5, 8 and 11 m depths are first estimated for location B1. The estimated values and 95% confidence intervals of $q_c$, $f_s$ and $u$ are shown in Fig. 8. Considerable variations are present in these parameters. Soil classification at location B1 is performed and shown in Fig. 9. At 2 m depth, clay is consistently identified, despite
considerable variation in the estimated $q_c$, $f_s$ and $u$. At 5 m depth, the prediction from the estimator and upper bound values indicate that the soil in this area is clay. However, the classification from the lower bound values of the interval indicates organic clay behaviour. Hence, the soil at 5 m depth may be in a transition zone from clay to organic clay. At 8 m depth, the estimator and the lower bound values lead to the conclusion of organic clay. The upper bound values identify the soil as clay, which further confirms the presence of a transition zone from clay to organic clay. At 11 m depth, all of the estimators, lower bound values and upper bound values of the 95% confidence interval indicate that the soil behaves as organic clay.

These results indicate clear soil stratification at location B1, which is clay at 2 m depth, transitioning to organic clay at 11 m depth. This estimation can be examined by comparing the stratification at location B1 with neighbouring tests (see Fig. 10). CPT test NS14 is located approximately 2 m south of B1. NS15 is located 3 m north of B1. NS19 is located 5 m east of B1. The stratification of the three existing CPT tests and the soil types at 2, 5, 8 and 11 m are clearly demonstrated in the figure. The soils at 2 m depth at the three locations are all clay, which supports the conclusion that the soil at B1 is clay. At 5 m depth, the soil at NS15 is in a transition zone, and the soils at NS14 and NS19 are clay, which also indicate that a transition zone at B1 is possible. At 8 m depth, the soil in NS19 is clay. The soils in NS15 and NS14 are located in the transition zone. Hence, with a high probability, location B1 is situated in a transition zone. At 11 m depth, the soils in the three CPT tests are all organic clay. Organic clay at 11 m depth at location B1 is logical. The boundary of the stratification can be accurately identified if predictions are obtained at more depths.

Figure 11 shows the estimators and 95% confidence intervals of $q_c$, $f_s$ and $u$ at location B2. The variations in the three parameters at location B2 are larger than those at location B1. The soil types classified from the estimator, the lower bound and upper bound values of the confidence interval are shown in Fig. 12. In the upper 2 m, the soil varies from clay to organic clay. From approximately 5 m, the prediction from the estimator and the 95% confidence interval consistently indicate that the soil behaves like organic clay. Hence, a layer of organic clay from 5 to 11 m can be concluded with high confidence.

**Soil properties**

With a clear definition of soil stratification, the property of each layer of soil can be established. Here, the layer of organic clay at location B2 is used as an example. The undrained shear strength, $s_u$, of clay can be calculated based on CPTU tests as
where $N_c$ is a cone factor. The cone factor values typically range from 7 to 15 (Robertson & Campanella, 1983), depending on the sensitivity and the overconsolidation ratio of the clay. Although affecting the absolute value of shear strength, the cone factor does not affect the trend of the strength. Kelly et al. (2014) have calibrated the cone factor for the Ballina clay relative to laboratory triaxial compression strength, which results in $N_c = 11.2$.

Parameters $q_c$, $f_s$ and $u$ are first estimated using the Kriging technique at various depths, which are further used to derive the corrected penetration resistance, $q_t$, according to equation (10). The profile of corrected cone penetration resistance for the layer of organic clay (from 5 to 11 m) at location B2 is shown in Fig. 13. The cone penetration resistance increases with depth. The range of cone penetration resistance with 95% probability is also given in Fig. 13. The shear strength of this organic clay layer can then be obtained using equation (20). The estimated shear strength and its 95% confidence interval along the depth are shown in Fig. 14. The undrained shear strength increases almost linearly with depth. The strength profile in this layer is estimated by $s_u = 8.0 + 1.5z$ kPa, with a strength of 15.5 kPa at the layer surface (i.e. 5 m) and a gradient of 1.5 kPa/m.

**CONCLUSIONS**

A method has been developed for the prediction of soil stratification at unsampled locations. The strength of this method is in its ability to greatly filter the uncertainties in
prediction in an explicit manner, which leads to clear stratification with a high degree of confidence. The resulting soil types and stratification were validated against existing CPTU tests. Excellent agreement was obtained between the predicted stratification and that based on sampling and sieve analysis.

Equations were derived to quantify the degree of uncertainties reduced by the proposed method. The majority of the uncertainties in the soil parameters (i.e., $q_c$, $f_s$, and $u$) are screened by the soil classification index. The remaining uncertainties are further filtered by the use of a soil classification system. This method leads to a consistent conclusion on soil type, despite considerable variations in soil parameters. Better prediction accuracy can be obtained for CPTU data of higher quality, better correlation of existing data and locations closer to existing tests.

Another advantage of this method is that the soil properties at locations without CPTU tests can be estimated at various depths with high confidence (and with confidence intervals calculated). The undrained soil strength profile of an organic clay layer at an unsampled location was found to increase linearly with soil depth, with a strength of 15.5 kPa at the layer surface and a gradient of 1.5 kPa/m with depth. Although this paper illustrates the procedure using a clay layer, the method is equally applicable to sand layers.

ACKNOWLEDGEMENTS

This research was supported by the Australian Research Council Centre of Excellence for Geotechnical Science and Engineering and by the Lloyd’s Register Foundation. This support is gratefully acknowledged.

NOTATION

- $a$: net area ratio
- $B_n$: pore pressure ratio
- $F_n$: normalised friction ratio
- $f_s$: sleeve friction
- $h$: separation distance
- $I_c$: soil classification index
- $m$: Lagrange multiplier
- $N_c$: cone factor
- $p_i$: property value at location $i$
- $p_{i+h}$: property value at location $i+h$
- $p_0$: estimated property at the unsampled location
- $\rho_{1-a}$: (1 – $a$) confidence interval
- $Q_i$: normalised cone resistance
- $q_c$: measured cone penetration resistance
- $q_t$: corrected cone penetration resistance
- $s_u$: undrained shear strength
- $u$: pore pressure
- $v_0$: equilibrium pore pressure
- $y_n$: semivariogram value for the data pairs with property values of $p_i$ and $p_{i+h}$
- $\varepsilon$: Kriging variation
- $\lambda_i$: weight for measured property $p_i$
- $\sigma$: Kriging standard error
- $\sigma_o$: total overburden stress
- $\sigma_o^\prime$: effective overburden stress
- $\Phi$: cumulative distribution function of standard normal variate

REFERENCES


![Fig. 14. Shear strength and 95% confidence interval of the organic clay layer at location B2](image-url)

Prediction $s_u = 8.0 + 1.5z$
$R^2 = 0.93$

Fitted shear strength model

PROBABILISTIC IDENTIFICATION OF SOIL STRATIFICATION 25


