Geotechnical and Geophysical Site Characterisation 5 – Lehane, Acosta-Martínez & Kelly (Eds)
© 2016 Australian Geomechanics Society, Sydney, Australia, ISBN 978-0-9946261-1-0

Session Report: Application of Statistical Techniques

M.B. Jaksa
School of Civil, Environmental and Mining Engineering, University of Adelaide, Australia

ABSTRACT: This report presents an overview of the aims, key aspects and conclusions associated with the 10 papers submitted to the conference relating to the application of statistical techniques, as well as some areas that the authors may wish to consider in future work. The papers are divided into 3 sections dealing with statistics and probability, spatial variability and artificial intelligence. The papers demonstrate the broad application of the statistics and probability methods to geotechnical site characterization and underline the important role these techniques play in, and value they deliver to, geotechnical engineering.

1 INTRODUCTION

It is well appreciated that, because they are natural, geotechnical materials are inherently variable and their behavior is complex. As such, stochastic methods have been shown, since the mid-1960s, to be of great benefit in geotechnical engineering. In particular, the application of statistical and probabilistic methods to geotechnical engineering can provide valuable insights into the nature of uncertainties in site characterization and geotechnical design and the influence of spatial variability on the behavior of geotechnical systems. In addition, these methods are essential for providing a framework for understanding and quantifying risk and reliability and also in facilitating accurate geotechnical predictions.

This report presents an overview of the aims, key aspects and conclusions associated with the 10 papers assigned to the *Application of Statistical Techniques* session, as well as some suggestions that the authors may wish to consider in future work. The papers are divided into 3 categories: *Statistics and Reliability*, with 4 papers (Baziw & Verbeek 2016, Foti & Passeri 2016, Huang et al. 2016, Styler & Weemees 2016); *Spatial Variability* incorporating 5 papers (Lehane et al. 2016, Krage et al. 2016, Parida et al. 2016, Wierzbicki et al. 2016, Zho et al. 2016); and *Artificial Intelligence* which includes a single paper dealing with an artificial neural networks model (Sastre et al. 2016). The papers are arranged in alphabetical order of the first author.

2 STATISTICS AND RELIABILITY

2.1 *Gaussian distribution fitting for reliability of shear wave velocity*

Baziw & Verbeek (2016) present an extension to their interval velocity classification (IVC) technique, which they originally proposed in 2015. The IVC method utilizes linearity estimates from polarization analysis, in conjunction with cross correlation coefficient calculations of the full waveforms obtained from the seismic cone penetration test (SCPT). The paper outlines details of the mathematics and implementation of a new parameter, termed spectrum rank (*SR*), which is introduced into the IVC technique. The spectrum rank quantifies the deviation of the source wave frequency spectrum from a desirable Gaussian-shaped curve. The seismic traces from the SCPT, with high signal-to-noise ratios, were observed to exhibit a similar shape to that of a Gaussian (i.e. normal) distribution.

The authors show that, when applied to actual SCPT data, the *SR* value is strongly correlated to the signal-to-noise ratio (SNR) of the acquired trace. In addition, for seismic traces with a low SNR, the authors recommend that batch or automated processing should be avoided and arrival times should be obtained visually from vertical seismic profiles in order to identify first breaks or dominant peaks or troughs.

2.2 *Reliability of soil porosity estimation*

Foti & Passeri (2016) present a research study investigating the reliability of porosity estimation from shear wave velocities using error propagation theory. They make use of the relationship in Equation (1) that was developed by the authors in an earlier study where the porosity, n, is expressed as a function of: ρ^s and ρ^w, which are, respectively, the mass densities of the soil particles and pore water; K^w is the bulk modulus of the pore water; V_p and V_s are the velocities of propagation of the dilatational and shear waves, respectively; and ν_{sk} is the Poisson's ratio of the soil skeleton.

$$n = \frac{\rho^s - \sqrt{(\rho^s)^2 - \dfrac{4\,(\rho^s - \rho^w)\,K^w}{V_p^{\,2} - 2\left(\dfrac{1-\nu_{sk}}{1-2\,\nu_{sk}}\right)V_s^{\,2}}}}{2\,(\rho^s - \rho^w)} \qquad (1)$$

The authors consider n as a function of ρ^s, ρ^w, K^w, $V_p = d/t_p$, $V_s = d/t_s$ and ν_{sk}, where d is the travel distance and t_i the travel times for each seismic wave (with $i = p, s$).

Error propagation theory is used to characterize the influence of each parameter that appears in the porosity formulation in (1), in this example specialized for a cross-hole test configuration. In their study, the authors assume that each variable is normally and randomly distributed and independent. In addition, the analyses assume $\nu_{sk} = 0.25 \pm 0.1$, $\rho^s = 2.7$ g/cm^3 and $\rho^w = 1$ g/cm^3.

The authors present data from two case studies, one from the site of the Zelazny Most tailings dam in Poland and the second from a site in the Italian town of Mirandola. From their analyses, the authors conclude that, for cross-hole tests, the distance between boreholes has the most significant influence on soil porosity estimation, whereas travel times have a modest effect. The influence of the velocity of compressional waves in water and the Poisson's ratio of the soil skeleton also affect the estimate of n. The effect of P-waves is also found to be significant.

2.3 *Bayesian updating using in situ test data*

The study by Huang et al. (2016) seeks to apply Bayesian updating to include laboratory testing and associated uncertainties to enhance the accuracy of seismic measurements. Their work involves the measurement of shear wave velocity, V_s, using the seismic dilatometer test (SDMT), supplemented with laboratory determined preconsolidation pressure data from constant rate of strain oedometer tests.

The authors adopt the Markov Chain Monte Carlo sampling method to sample the posterior distribution. Empirical relationships are then derived between the preconsolidation pressure, σ', and V_s. A prior distribution of preconsolidation pressures is obtained assuming a log-normal distribution, and using a linear trend and the geostatistical technique of kriging.

The posterior mean preconsolidation pressure is then used to update the SDMT data such that its trend and magnitude more closely approximates the laboratory values. The authors observe that the posterior prediction is a better fit to the laboratory test data and the use of a scale of fluctuation with the kriged prior also provides greater accuracy when interpolating between data points.

The authors conclude that the uncertainties associated with preconsolidation pressure can be significantly reduced by incorporating shear wave velocity measurements. In addition, whilst the paper presents a 1D example, the authors suggest that the technique is equally relevant to 2D and 3D situations.

2.4 *Quantifying and reducing uncertainty in down-hole shear wave velocities using signal stacking*

The research study presented by Styler & Weemees (2016) seeks to quantify the improvement in the interpreted shear wave propagation time, in down-hole seismic testing using a seismic piezocone (SCPTU), that can be realized through signal stacking of multiple traces. The extensive paper demonstrates how to: calculate the noise in a set of down-hole seismic traces; quantify the signal-to-noise ratio (SNR) for a trace and for stacked signals; and evaluate the error in the propagation time when comparing two seismic traces.

The authors found that the SNR increases with signal stacking, the error in the propagation time decreases with higher SNR, and the decrease in SNR with depth can be overcome by signal stacking.

In addition, the authors presented interpretation methods that can be used to: quantify signal noise through signal-subtraction; quantify SNR in down-hole seismic signals; and calculate the error in the shear wave propagation time using the cross-correlation function. The authors state that, while the techniques outlined in the paper are applicable to the SCPTU, they can also be applied to any SCPTU that recorded more than one trace per seismic test.

3 SPATIAL VARIABILITY

3.1 *Probabilistic assessment of laterally loaded pile performance in sand*

Lehane et al. (2016) present a probabilistic assessment of the lateral response of a pile in a uniformly graded dune sand using cone penetration test (CPT) and load test data. The paper examines the effects of the spatial variability of the ground parameters using a large series of CPTs performed at the site of the lateral pile tests. The authors use a direct CPT approach developed by Suryasentana & Lehane [S&L] (2014, 2016) to predict the sand's lateral load-displacement (*p-y*) curves and these are compared with the measurements of the test piles. A Monte Carlo simulation is performed which involves the generation of a series of random q_c profiles consistent with the assessed site variability and their combination with the *LAP* program to generate probability density functions for the load that would cause a 1% rotation gradient (=0.57°) at the pile head.

The paper examines the sensitivity of variations in q_c to the performance of two laterally loaded test piles conducted at the Shenton Park site in Perth, Western Australia. The stratigraphy at the Shenton Park site consists of a 5–7 m thick deposit of siliceous dune sand overlying weakly cemented limestone. A total of 12 CPTs were performed on the site. Two 225 mm diameter x 3.5 m long grout piles, constructed using the continuous flight auger technique, were adopted in the study and pushed laterally apart.

The lateral load-displacement relationship is predicted using the *LAP* program incorporating the S&L Method. Good agreement is observed between the predicted and measured behaviours.

A total of 50 CPTs were randomly generated, using *Microsoft Excel* and a vertical scale of fluctuation of 0.5 m, which was estimated from the CPTs performed on site.

The authors observed that the lateral load range likely to induce a given level of head rotation for a pile in sand is significantly lower than the range anticipated from the q_c variability and the CPT-based S&L Method provides good predictions for the lateral response of a test pile in a medium dense sand site. The authors also observed that the predictions using the S&L and the API sand methods have a low sensitivity to randomly generated q_c or ϕ' profiles that are normally distributed at any given depth.

3.2 *Identification of geological depositional variations using CPT-based conditional probability mapping*

Krage et al. (2016) present research aimed at improving the characterization and understanding of subsurface stratigraphy at a project site using transition probability geostatistics conditioned to CPT soundings and combined with geological information.

The authors use geostatistics in two ways to augment site investigations to: (1) identify or estimate the soil type at unknown locations; and (2) estimate the engineering properties at these unknown spatial locations. Their integrated site characterization framework is summarized in Figure 1.

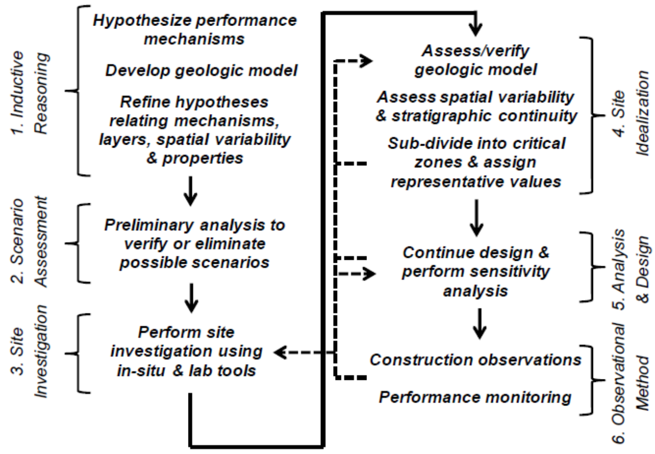

Figure 1. Flowchart of the integrated site characterization framework.

The authors incorporate the transition probability in their analyses, which describes the likelihood of transitioning from one category (where categories are user defined, such as soil type or engineering property based) to another over some separation distance. The main advantage of this approach is the ability to model ordered systems, such as geological facies environments.

The authors present an example case of a site for a new 11 m high embankment dam with respect to liquefaction assessment to which they applied their framework. Geostatistical simulation was performed conditioned to the CPTs taken along the dam wall alignment.

The authors concluded that the simulations indicate liquefaction is expected to be most prevalent at shallow depths: 1–2 m deep on the west side and to 4–6 m deep on the east side. Liquefaction is also expected at greater depths.

3.3 *Stochastic waveform inversion for probabilistic geotechnical site characterization*

Parida et al. (2016) propose a stochastic inverse analysis methodology to estimate probabilistically the Young's modulus from geophysical test measurements by accounting for uncertainties from spatial variability, measurement errors and limited data. They focus on the spectral analysis of surface waves (SASW) seismic geophysical test and adopt Monte Carlo simulation.

The authors simulate a 3D virtual site of soil moduli using an anisotropic random field. The methodology employs the finite element method with the stochastic collocation approach to solve probabilistically the forward problem for the SASW test. The virtual site is then excited using a chirp signal and the ground is assumed to be linear-elastic.

From the simulations performed the authors observe that the amount of information gained decreases with depth, implying that the sensors towards the bottom contribute modestly to the inverse estimation process. The authors suggest that the developed methodology is mathematically rigorous and computationally efficient, and is general enough to be extended widely, including inverse estimation of additional elastic, as well as elastoplastic, soil parameters.

3.4 *3D mapping of organic layers by means of CPTU and statistical data analysis*

Wierzbicki et al. (2016) present a research study which seeks to examine selected methods of statistical data analysis to determine the spatial extent of organic soil layers using piezocone (CPTU) data. Geotechnical characterization is undertaken for a site located 50 km from Poznań, Poland. The ground at the study area consists of glacial clay, layered sands and gravels, silts and organic soils.

CPTU data are subjected to clustering analysis using the *k*-means method and the inverse distance weighting (IDW) method is used to develop 2D and 3D models. The authors conclude that only simultaneous use of all available data result in accurate identification of the organic soil layer. The complete 3D IDW model, however, yields unsatisfactory results.

3.5 *Assessment of ground improvement on silt based on spatial variability analysis of CPTU data*

The research study undertaken by Zho et al. (2016) seeks to examine the difference in spatial variability characteristics of silt before and after compaction. A silt site located in the Jiangsu province, China, was improved using a new deep resonance compaction technique to increase the liquefaction resistance of the silt. A total of 17 CPTUs were performed prior to compaction and 26 after ground improvement (9 after 14 days, 7 after 54 days and 10 after 60 days) to assess the efficacy of the technique.

The study undertakes spatial variability analyses on the CPTU data before and after compaction. In particular, random field theory is used to examine cone tip resistance, q_t. The mean, coefficient of variation, *COV*, and the vertical scale of fluctuation, δ_v, of q_t are examined prior to and after compaction.

The results show that both the mean and δ_v decreased immediately after compaction, but gradually increased with the strength and density recovery; and the *COV* consistently decreased after compaction. The authors also suggest that the variation in the spatial variability of q_t could possibly be explained by the actions of the CPTU test and resonance compaction, their effect on the structure of the silt and the increase in soil strength with time.

4 ARTIFICIAL INTELLIGENCE

In the final paper in this session, Sastre et al. (2016), using data from the Panda2 variable energy dynamic cone penetrometer, develop an artificial intelligence-based model to predict the grain size class of the subsurface profile. The Panda is a lightweight dynamic cone penetrometer that is used to investigate ground profiles to depths of up to 5 m. It uses variable energy which is delivered manually by repeated blows of a standardized hammer. After each blow, the dynamic cone resistance, q_d, is calculated at the current depth using the Dutch formula (Cassan 1988):

$$q_d = \frac{\frac{1}{2}MV^2}{A(1+P/M)x_{90°}} \quad (2)$$

where: M is the weight of the striking mass; V is the speed of impact of the hammer; A is the area of the cone; P is the weight of the struck mass; and $x_{90°}$ is the penetration due to a single blow of the hammer (for a 90° cone).

The authors develop an artificial neural networks (ANNs) model using a database consisting of 218 Panda2 penetrograms (soundings), incorporating 149 tests performed in a laboratory-based calibration chamber (370 mm in diameter and 800 mm in depth) and 69 in situ tests from various locations in France. The tests were undertaken on relatively homogeneous soils containing particles with a grain size smaller than 50 mm, and the nature and geotechnical properties of the soils were characterized using standard laboratory classification tests.

In order to determine the appropriate set of ANN input parameters, the authors used a technique inspired by speech recognition, termed the '4 signal analysis' approach, which the authors state incorporates a statistical, nonlinear, morphological and spectral, and pattern vector of 26 parameters for each penetrogram. Using a 'one-at-a-time approach' the authors examined the sensitivity of each of the 26 parameters and reduced the number of inputs to the following 17: (1) q_d mean; (2) q_d median; (3) q_d standard deviation; (4) q_d coefficient of variation; (5) q_d variance; (6) q_d range; (7) q_d interquartile range; (8) q_d skewness; (9) q_d kurtosis; (10) q_d Shannon entropy; (11) q_d logarithm entropy range; (12) q_d skewness; (13) q_d slope changes; (14) q_d waveform; (15) linear coefficient of linear trend; (16) independent coefficient of linear trend; and (17) maximum spectral power. The sole

ANN output is termed by the authors *GTR classification*, which is essentially a soil type classification into one of the following 4 classes: Class 1: fine-grained soils; Class 2: fine sands; Class 3: sands/gravels; and Class 4: gravels.

Between 2 and 25 hidden nodes are examined, as well as one and two hidden layers. The final, optimal ANN model consists of 17 input nodes, 12 nodes in a single hidden layer, and a single output node. The authors observe that the optimal ANN model assigned the correct class to of 97% of the samples (212 out of 218). The authors also report an accuracy of 94% and 97% for the testing and validation and sets, respectively.

In the interests of progressing and improving the work in the future, that authors may wish to consider the following comments. Closer examination of the 17 input nodes shows some degree of redundancy and potential inefficiency. For example, inputs (8) and (12) appear to be identical and it is difficult, from the description provided in the paper, to appreciate parameters (15) and (16). Furthermore, parameter (4), the coefficient of variation, is simply the ratio of the standard deviation [parameter (3)] to the mean [parameter (1)]. Hence, the inclusion of parameter (4) adds no new information to the model. Furthermore, parameter (5), the variance, is simply the square of the standard deviation [parameter (3)]. In addition, it is likely that other statistical parameters, such as (2), (6)–(9), (12), may also provide marginal value. If so, one would need to question the validity of the 'one-at-a-time' approach, which is used to assess the sensitivity of the individual input parameters.

In order to develop a parsimonious model, and one that perhaps yields an equation which is suitable for simple hand calculation [see Shahin et al. (2002); Shahin & Jaksa (2005), for example], the authors are encouraged to explore redeveloping the ANN with the following input parameters: (1) q_d mean; (3) q_d standard deviation; (8) q_d skewness; (9) q_d kurtosis; (10) q_d Shannon entropy; (11) q_d logarithm entropy range; (13) q_d slope changes; (14) q_d waveform; (15) linear coefficient of linear trend; (16) independent coefficient of linear trend; and (17) maximum spectral power.

Finally, it is unclear from the paper how the authors plan to disseminate the model that it can be used in practice. The development of a parsimonious model and an equation would certainly assist in this endeavor.

5 CONCLUSION

The 10 papers presented in this session demonstrate the wide range of geotechnical site characterization problems that can be successfully addressed using statistical and probabilistic methods.

6 REFERENCES

Baziw, E. & Verbeek, G. 2016. Frequency spectrum "bell-curve" fitting as a component of SCPT interval velocity accuracy assessment. In Editor (ed.) *5th International Conference on Geotechnical and Geophysical Site Characterisation*; *Proc. intern. conf., Gold Coast, 3-4 September 2016*. Rotterdam: Balkema.

Cassan, M. 1988. Les essais in situ en mechanique des sols. Volume I realisation et interpretation, Eyrolles, 146-151.

Foti, S. & Passeri, F. 2016. Reliability of soil porosity estimation from seismic wave velocities. In Editor (ed.) *5th International Conference on Geotechnical and Geophysical Site Characterisation*; *Proc. intern. conf., Gold Coast, 3-4 September 2016*. Rotterdam: Balkema.

Huang, J., Kelly, R. & Sloan, S.W. 2016. Enhanced data interpretation: combining in-situ test data by Bayesian updating. In Editor (ed.) *5th International Conference on Geotechnical and Geophysical Site Characterisation*; *Proc. intern. conf., Gold Coast, 3-4 September 2016*. Rotterdam: Balkema.

Krage, C., DeJong, J.T. & Boulanger, R.W. 2016. Identification of geological depositional variations using CPT-based conditional probability mapping. In Editor (ed.) *5th International Conference on Geotechnical and Geophysical Site Characterisation*; *Proc. intern. conf., Gold Coast, 3-4 September 2016*. Rotterdam: Balkema.

Lehane, B.M., Glisic, F. & Doherty J.P. 2016. Probabilistic assessment of laterally loaded pile performance in sand. In Editor (ed.) *5th International Conference on Geotechnical and Geophysical Site Characterisation*; *Proc. intern. conf., Gold Coast, 3-4 September 2016*. Rotterdam: Balkema.

Parida, S.S., Sett, K. & Singla, P. 2016. Stochastic waveform inversion for probabilistic geotechnical site characterization. In Editor (ed.) *5th International Conference on Geotechnical and Geophysical Site Characterisation*; *Proc. intern. conf., Gold Coast, 3-4 September 2016*. Rotterdam: Balkema.

Sastre, C., Benz, M., Gourvès, R., Breul, P. & Bacconnet, C. 2016. Automatic methodology to predict grain size class from dynamic penetration test using neural networks. In Editor (ed.) *5th International Conference on Geotechnical and Geophysical Site Characterisation*; *Proc. intern. conf., Gold Coast, 3-4 September 2016*. Rotterdam: Balkema.

Shahin, M.A., Jaksa, M.B. & Maier, H.R. 2002. Artificial neural network–based settlement prediction formula for shallow foundations on granular soils. *Australian Geomechanics* 37(4): 45-52.

Shahin, M.A. & Jaksa, M.B. 2005. Neural network prediction of pullout capacity of marquee ground anchors. *Computers and Geotechnics* 32(3): 153-163.

Styler, M.A. & Weemees, I. 2016. Quantifying and reducing uncertainty in down-hole shear wave velocities using signal stacking. In Editor (ed.) *5th International Conference on Geotechnical and Geophysical Site Characterisation*; *Proc. intern. conf., Gold Coast, 3-4 September 2016*. Rotterdam: Balkema.

Suryasentana, S.K. & Lehane, B.M. 2014. Numerical derivation of CPT-based p–y curves for piles in sand. *Géotechnique* 64(3): 186-194.

Suryasentana, S.K. & Lehane, B.M. 2016. Updated CPT-based p-y formulation for laterally loaded piles in cohesionless soil under static loading. *Géotechnique*, dx.doi.org/10.1680/ jgeot.14.P.156.

Wierzbicki, J., Smaga, A., Stefaniak, K., Wołyński, W. 2016. 3D mapping of organic layers by means of CPTU and statistical data analysis. In Editor (ed.) *5th International Conference on Geotechnical and Geophysical Site Characterisation*; *Proc. intern. conf., Gold Coast, 3-4 September 2016*. Rotterdam: Balkema.

Zou, H.F., Cai, G.J., Liu, S.Y., Lin, J., Bheemasetti, T.V. & Puppala, A.J. 2016. Assessment of ground improvement on silt based on spatial variability analysis of CPTU data. In Editor (ed.) *5th International Conference on Geotechnical and Geophysical Site Characterisation*; *Proc. intern. conf., Gold Coast, 3-4 September 2016*. Rotterdam: Balkema.