Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly

doi: 10.3850/978-981-18-5182-7_07-004-cd

ISGSR 2022

Bayesian Regression Models for Predicting Undrained Shear Strength from Piezocone Penetration Tests

Mats Kahlström[1] and Nezam Bozorgzadeh[1]

[1]Norwegian Geotechnical Institute, Oslo, Norway.
mats.kahlstrom@ngi.no
nezam.bozorgzadeh@ngi.no

Abstract: This paper presents Bayesian regression models for predicting undrained shear strength from piezocone penetration test (CPTU) measurements. A database of laboratory tests from high-quality block samples in clays and the corresponding CPTU measurements are re-analyzed. Attention is paid to reporting and communicating the results of regression analyses, particularly statistical parameter uncertainty, a task that is often performed poorly in geotechnical engineering literature. Four regression models are fitted to the data, and some preliminary goodness-of-fit checks are discussed. Furthermore, the predictive accuracy of the models is compared. The models are then used in an example of predicting undrained shear strength from CPTU data at a new location. It is discussed how such probabilistic predictions relate to the notion of characteristic values prescribed in geotechnical design standards. Finally, updating regression models based on site-specific data is briefly discussed.

Keywords: Undrained shear strength; CPTU; Bayesian data analysis; uncertainty quantification.

1 Introduction

Bayesian methods have been used for predicting undrained shear strength of clays from effective stress and laboratory test data (e.g. Cao and Wang 2014). More recently, and with increasing availability of the piezocone penetration test (CPTU), CPTU-based empirical relationships are gaining popularity in practice. However, more often than not, the results of statistical analyses of such data is communicated inadequately in the literature; best-fit values for regression coefficients and residual standard deviations are usually reported but statistical parameter uncertainty is overlooked (e.g., Karlsrud et al. 2005; Mayne and Peuchen 2018; Paniagua et al. 2019). Quantification and propagation of these uncertainties is crucial for probabilistic engineering design and risk assessment in general, and for deriving characteristic/representative profiles that comply with modern design standards such as EN 1997 (CEN, 2004) in particular. This is also essential for combining site-specific data and generic correlations in a Bayesian framework (e.g. Bozorgzadeh et al., 2019; Ching et al., 2021).

In this paper, using Bayesian regression models, we re-analyze a database of laboratory tests from high-quality block samples in clay and the corresponding CPTU measurements. The database includes information from multiple sites with comparable ground conditions, primarily from Norway. Multiple regression models are fitted to the database relating the undrained shear strength S_u^C, obtained from anisotropically-consolidated undrained triaxial compression tests (CAUC), to CPTU measurement data such as the net cone resistance q_{net}, and the effective pore pressure response Δu. Water content w obtained from index lab testing is also used as a predictor in one of the regression models. The paper is organized as follows. Section 2 introduces the data and the empirical correlations. Section 3 reports the results of the regression analyses. Section 4 compares the fitted models and presents an example of probabilistic predictions of undrained shear strength for a new CPTU profile, and then discusses model updating based on site-specific data.

2 Database and empirical equations

The database contains laboratory test results from high-quality block samples prepared using the Sherbrooke block sampler (Lefebvre and Poulin, 1979) and the corresponding CPTU measurements on soft to medium stiff clays. The database has evolved throughout the years (Karlsrud et al. 2005; Karlsrud and Hernandez-Martinez 2013) and was most recently described by Paniagua et al. (2019). In its present form, the Paniagua et al. (2019) database contains 61 samples from 19 sites, primarily from Norway.

Except for updating the values of Δu for two samples, the database analyzed in this paper is identical to Paniagua et al. (2019). All samples in the current database have a sample quality of 1 or 2 based on the criteria described by Lunne et al. (2006). The data are shown in Figure 1 along with one of the fitted regression models (Section 3). For more details about the data see Paniagua et al. (2019).

Paniagua et al. (2019) report different empirical equations for predicting S_u^C. Here, we first re-examine a model that includes q_{net}, w and Δu as predictors

$$S_u^C = \beta_0 \times (\frac{q_{net}}{w})^{\beta_{qnet/w}} \times \Delta u^{\beta_{\Delta u}} \quad (1)$$

Three simpler variations of the above equation are also considered:

$$S_u^C = \beta_0 \times q_{net}^{\beta_{qnet}} \times \Delta u^{\beta_{\Delta u}} \quad (2)$$

$$S_u^C = \beta_0 \times \Delta u^{\beta_{\Delta u}} \quad (3)$$

$$S_u^C = \beta_0 \times q_{net}^{\beta_{qnet}} \quad (4)$$

Eqs. (2) - (4) only include CPT-based predictors. Such models are gaining more attraction in practice because they do not rely on additional laboratory test results such as water content or plasticity index.

3 Bayesian statistical modelling

3.1 Regression models

The regression model is fitted to log-transformed data. This transformation is useful, first because S_u^C is constrained to be positive, and second, because larger values of S_u^C exhibit larger variance (e.g. Fig 1b). Therefore, for i = 1, 2, ..., 61 data points the regression model is:

$$\ln\left(s_{u[i]}^C\right) = \mu_{[i]} + \varepsilon_{[i]} = \ln(\beta_0) + \beta_{qnet/w} \times \left(\ln\left(q_{net[i]}\right) - \ln\left(w_{[i]}\right)\right) + \beta_{\Delta u} \times \ln\left(\Delta u_{[i]}\right) + \varepsilon_{[i]} \quad (5a)$$

where μ is mean log-transformed S_u^C and defined as a linear function of the predictors: $\ln(q_{net})$ - $\ln(w)$ and $\ln(\Delta u)$. When fitting the linear regression, the mean of each predictor is subtracted from them, i.e. the predictors are *mean centered.* So, the regression model becomes:

$$\ln\left(s_{u[i]}^C\right) = \ln(\beta_0) + \beta_{qnet/w} \times \left(\ln\left(q_{net[i]}\right) - 6.09 - \ln\left(w_{[i]}\right) - 0.89\right) + \beta_{\Delta u} \times \left(\ln\left(\Delta u_{[i]}\right) - 5.82\right) + \varepsilon_{[i]} \quad (5b)$$

The constant term $\ln(\beta_0)$ in the centered model has a physically appealing interpretation: it is the average $\ln(S_u^C)$ when the predictors are set to their averages. The mean-centering also removes the correlation between the intercept $\ln(\beta_0)$ and the coefficients of the predictors (i.e. the slopes) of the linear regression model. For comprehensive discussions about centering and scaling predictors and response variables in regression analysis see e.g. Gelman et al. (2020).

The εs are assumed to be normally distributed errors with mean zero and unknown standard deviation $\sigma_{\ln(suc)}$ (hereafter σ for brevity):

$$\varepsilon_i \sim \text{Normal}\left(0, \sigma_{ln(S_u^C)}\right) \quad (5c)$$

The model parameters (i.e. $\ln(\beta_0)$, $\beta_{qnet/w}$, $\beta_{\Delta u}$ and σ) are assigned vague prior distributions that convey little information (relative to the data) so the posteriors are essentially not affected by them. The regression models based on Eqs. (2) - (4) are formulated similar to the regression model of Eq. (5) but are not stated here due to limited space. All models are fitted using the probabilistic programming software Stan (Stan development team, 2019) which uses Markov chain Monte Carlo to simulate from the posterior distributions of the parameters of the model.

3.2 Results

A common visualization of linear regression with two or more continuous predictors is plotting the fitted model against each predictor while *the other predictors are held at their average values*. This allows a visual examination of the average effect of the predictors on the response variable one at a time. Such visualizations can be shown on either the log-log scale or the original scale of the data.

Figure 1 shows two examples of this for the regression model of Eq. 5. Figure 1a shows mean $\ln(S_u^C)$ vs. $\ln(\Delta u)$, and can be interpreted as follows. For a clay with q_{net} of about 444 kPa (corresponding to the mean $\ln(q_{net})$ of 6.09 in Eq. 5b) and w of about 0.41 (corresponding to mean $\ln(w)$ of −0.89 in Eq. 5b), a unit change in $\ln(\Delta u)$ corresponds to an average increase of 0.67 in $\ln(S_u^C)$ (see $\beta_{\Delta u}$ in Table 1). Figure 1b shows mean S_u^C vs. $q_{net/w}$; plotting on the original scale of the data might be more familiar to engineers, but the non-linearity of the regression curve does not allow the straightforward interpretation permitted on the log-scale. Figure 1 also shows the pointwise 95% credible interval for posterior mean and the 95% posterior predictive interval for the response variable. Additional visual summaries of the regression model of Eq. 5, and visual summaries for the other regression models are not provided here due to limited space.

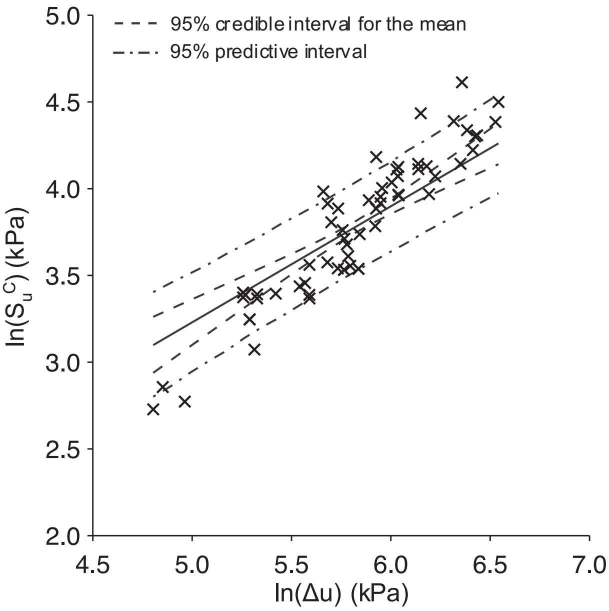
Table 1 gives posterior summary statistics for all regression models. The posterior means can be thought of as best estimates with posterior standard deviations quantifying the associated uncertainty. As mentioned earlier, uncertainty in estimates is usually not reported in the geotechnical literature. Aside from the fact that these are part of the fitted model and should not be overlooked, they are particularly crucial for predicting the "mean" response (here, mean S_u^C) in design situations where the *average* geotechnical parameter is of interest. Furthermore, it should be noted that estimated regression coefficients can be correlated, and this should also be reported. As mentioned earlier, mean-centering of predictors removes the correlation between the intercept and the slopes. For the regression model of Eq. (5), the correlation coefficient between $\beta_{qnet/w}$ and $\beta_{\Delta u}$ is -0.86, and for the model based on Eq. (2), the correlation coefficient between β_{qnet} and $\beta_{\Delta u}$ is -0.81. Another instance where complete reporting of regression results is important is when one wants to formulate prior distributions based on published models but does not have access to the original data. This is not possible if only the point estimates are reported.

Table 1. Posterior summary statistics.

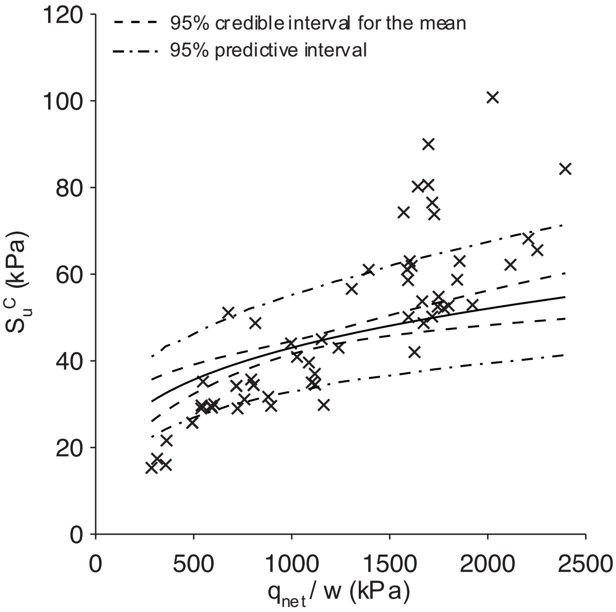
Eq.	Parameter							
	$\ln(\beta_0)$		$^*\beta_{qnet}$		$\beta_{\Delta u}$		σ	
	mean$_{(sd)}$	95% CI	mean$_{(sd)}$	95% CI	mean$_{(sd)}$	95% CI	mean$_{(sd)}$	95% CI
1	3.78$_{(0.02)}$	(3.74, 3.81)	0.27$_{(0.06)}$	(0.16, 0.39)	0.67$_{(0.08)}$	(0.51, 0.83)	0.13$_{(0.01)}$	(0.11, 0.16)
2	3.78$_{(0.02)}$	(3.74, 3.81)	0.28$_{(0.09)}$	(0.11, 0.46)	0.79$_{(0.08)}$	(0.63, 0.93)	0.14$_{(0.01)}$	(0.12, 0.17)
3	3.78$_{(0.02)}$	(3.74, 3.81)	-	-	0.99(0.05)	(0.89,1.08)	0.15$_{(0.01)}$	(0.13, 0.18)
4	3.78$_{(0.03)}$	(3.72,3.84)	1.03$_{(0.09)}$	(0.85, 1.20)	-	-	0.23$_{(0.02)}$	(0.20, 0.28)

* For model in Eq. (1), this is $\beta_{qnet/w}$

Figure 2 shows the standardized residuals of $\ln(S_u^C)$ for all regression models. For each model. it is expected that 95% of the residuals (about 58 out of 61) lie within (-1.96, 1.96) which is the 95% probability range for a standard normal distribution. This is the case for all four models. The 9th observation is particularly interesting because it seems to be an outlier (i.e. has a low probability of occurrence) according to all models. No obvious problems could be associated with the measured values for this data point from a preliminary investigation of the lab and CPTU data. The regression model could be improved to account for this observation, e.g. by replacing the normal error (Eq. 5c) with a heavy-tail student *t*-distribution. Such improvements are not pursued in this paper; an example could be found in Bozorgzadeh and Bathurst (2019).



a) with $\ln(q_{net})$ and $\ln(w)$ kept at their averages



b) with $\ln(\Delta u)$ kept at its average

Figure 1. Example visualizations of the fitted regression model of Eq. (5).

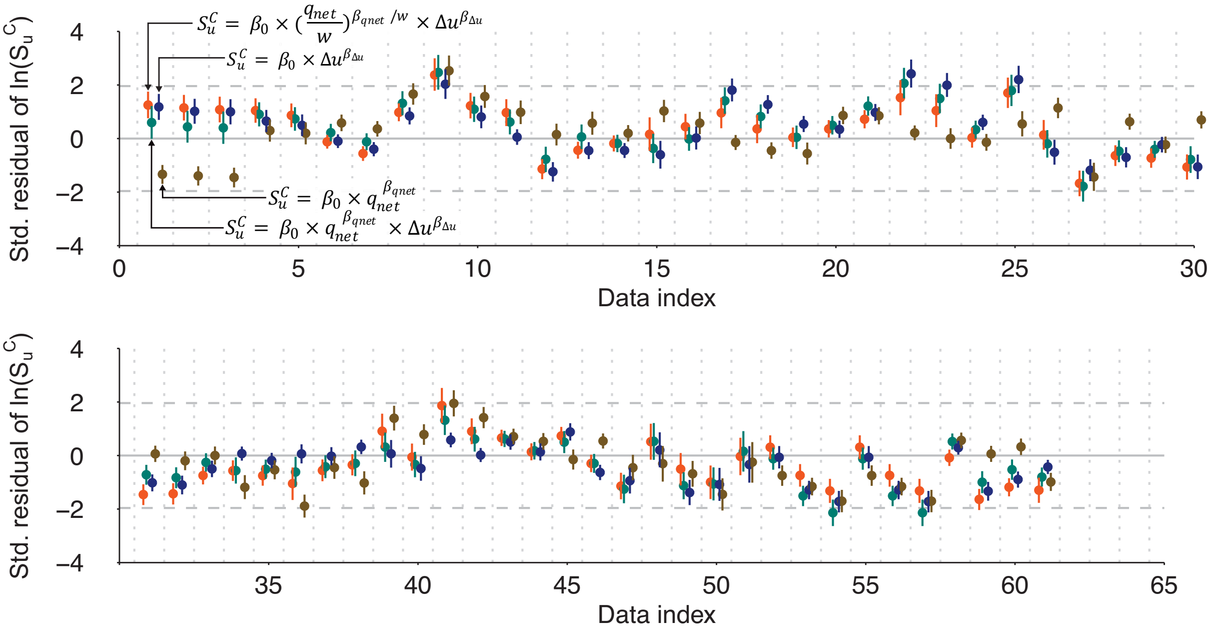


Figure 2. Standardized residuals of $\ln(S_u^C)$ from the four regression models.

4 Model comparison and example predictions of S_u^C from new CPTU data

In the Bayesian framework, predictive accuracy of models can be compared using different methods as discussed by Gelman et al. (2013) and summarized by Bozorgzadeh and Bathurst (2019) in the context of geotechnical applications. Here, we use the leave-one-out information criterion (LOOIC) which is a log-likelihood based measure of goodness-of-fit. Calculating the exact LOOIC requires excluding data points one at a time from the analysis, fitting the model to the remaining n-1 data points, and using the resulting posteriors to evaluate the log-density of the left-out data point; this is computationally expensive. Vehtari et al. (2017) introduced an efficient computation of LOOIC using Pareto-smoothed importance sampling where it is possible to estimate LOOIC based on simulations from the posteriors obtained using all the data.

Table 2 reports the values of LOOIC and approximate standard errors estimated using this method. Smaller LOOIC indicates higher predictive accuracy, so the first model seems to be the most accurate. However, the difference in LOOIC for the second model is less than two standard errors away from zero, and the third one about two standard errors away from zero, indicating that these models should not be necessarily ruled out. The model with qnet as the only predictor has considerably less predictive accuracy. This model also has a larger residual standard deviation compared to the other three models (see Table 1). This is in agreement with general findings of Karlsrud et al. (2005). This model is not suggested for predictions.

Table 2. Model comparison.

Empirical equation	$LOOIC_{(SE)}$	Difference* in $LOOIC_{(SE)}$
$S_u^C = \beta_0 \times (\frac{q_{net}}{w})^{\beta_{qnet/w}} \times \Delta u^{\beta_{\Delta u}}$	$-73.93_{(8.78)}$	$0.0_{(0.0)}$
$S_u^C = \beta_0 \times q_{net}^{\beta_{qnet}} \times \Delta u^{\beta_{\Delta u}}$	$-64.20_{(11.19)}$	$9.7_{(7.6)}$
$S_u^C = \beta_0 \times \Delta u^{\beta_{\Delta u}}$	$-56.29_{(10.53)}$	$17.6_{(8.6)}$
$S_u^C = \beta_0 \times q_{net}^{\beta_{qnet}}$	$-2.72_{(9.99)}$	$71.2_{(8.4)}$

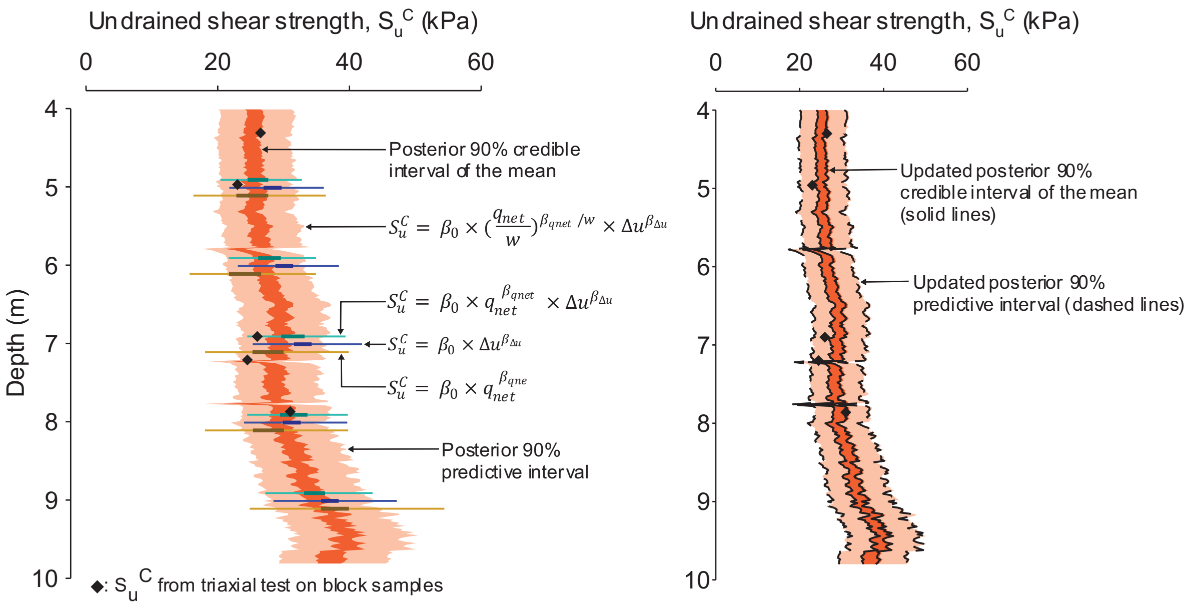
* Difference from LOOIC of the first model

The fitted regression models can be used to predict S_u^C of a new location, i.e., to transform a CPTU profile (and the water content data) to distributions of S_u^C. For demonstration, a clay site outside Stockholm in Sweden (Hov and Garcia de Herreros, 2020) is selected where CPTU measurements are available. The parameter ranges for the predictors at this new site are within the envelope of parameters in the database used to fit the regression models (q_{net}:143-437 kPa, Δu: 162-347 kPa, w: 0.45-0.65). Other index parameters for the clay such as plasticity index, although not used in the regression models, are also within the envelopes found in the database. Thus, it is reasonable to assume that the new site falls within the domain of applicability of the regression models.

Starting with regression model of Eq. (5), the dark-shaded range in Figure 4a is the 90% probability interval for *mean* S_u^C; this uncertainty stems from uncertainty in the regression parameters. The light-shaded range is the 90% predictive interval, i.e. predictions of yet-to-be-observed values of S_u^C. The boundaries of the shaded ranges are the 5th and 95th percentiles of the distributions of mean S_u^C and S_u^C, and are of interest because they correspond to definitions of characteristic values in modern geotechnical design standards such as EN 1997 (CEN, 2004). Therefore, the lower bounds of the dark- and light-shaded areas are the *characteristic* S_u^C *profiles* for design situations (limit states) where large and small ground volumes are involved, respectively.

To prevent visual clutter, yet allow for some form of visual comparison, predictions from the other models are shown at few selected depths as error bars. For each error bar, the central range marked by a thicker line is the 90% interval for the mean (comparable to dark-shaded) and the range represented by the thinner line is the 90% predictive interval (comparable to light-shaded). The model with both q_{net} and Δu gives predictions that are approximately the same as the first model. The model with only Δu gives slightly higher predictions, and the model with only q_{net} predicts lower S_u^C and is also the most uncertain. If the difference between predictions from different models is judged to be important from an engineering perspective, rather than relying on a single (best) model, model averaging can be used. Exploring this is beyond the scope of this paper; Bozorgzadeh and Bathurst (2019) discuss Bayesian model averaging for geotechnical applications.

For the selected new location, in addition to CPTU, undisturbed sampling using the mini block-sampler (Emdal et al. 2016) was performed. Figure 3 shows S_u^C from five CAUC tests that were performed on specimens between 4 to 8 m depth. All measurements fall within the 90% predictive distributions of the four regression models.



(a) From regression models (b) Updated prediction with site-specific data (model of Eq. 5)

Figure 3. Predictions of undrained shear strength profiles for a new CPTU measurement.

The current predictions can be updated in light of the new triaxial test results. The updated profiles for regression model of Eq. (5) are shown in Figure 3b. The updated model results in essentially the same profiles as before. This is because the regression model assumes that given the predictors, the average undrained shear strength for all sites is the same. The mean SuC for all sites is estimated from 61 data points; adding five data points does not change the posterior distributions of the parameters. This assumption of identical mean SuC for all sites can be relaxed by fitting hierarchical regression models that allow for model parameters to be site-dependent (e.g. Bozorgzadeh et al., 2019; Bozorgzadeh and Bathurst, 2020; Ching et al., 2021). Extending the regression models presented in this paper to hierarchical models is the topic of ongoing research by the authors.

5 Summary and conclusions

Using a previously published empirical equation for predicting undrained shear strength from CPTU data this paper discussed better communication of regression results, particularly when it comes to uncertainties and also making it possible for others to use a reported regression model for making probabilistic predictions or formulating prior distributions. The paper then discussed preliminary goodness-of-fit checks of the regression models using residuals. Predictive accuracy of four regression models was compared. The best model was used in a demonstrative example to predict undrained shear strength at a new site where CPTU data are available. The notion of characteristic values as described in modern design standards was discussed in the context of probabilistic predictions from the regression models. Updating the regression model with site-specific data was briefly explored.

The (usually not explicitly stated) restrictive assumption of identical regression parameters was discussed, and hierarchical Bayesian modeling identified as the way forward for more efficient utilization of site-specific data.

Acknowledgments

The authors are grateful to the Norwegian Research Council for supporting this research through its annual base funding to NGI.

References

Bozorgzadeh, N., and Bathurst, R. J. (2019). Bayesian model checking, comparison and selection with emphasis on outlier detection for geotechnical reliability-based design. *Computers and Geotechnics*, 116, 103181.

Bozorgzadeh, N., Harrison, J. P., & Escobar, M. D. (2019). Hierarchical Bayesian modelling of geotechnical data: application to rock strength. *Géotechnique*, 69(12), 1056-1070.

Bozorgzadeh, N., & Bathurst, R. J. (2020). Hierarchical Bayesian approaches to statistical modelling of geotechnical data. *Georisk*. doi: https://doi.org/10.1080/17499518.2020.1864411.

Cao, Z. and Wang, Y. (2014). Bayesian Model Comparison and Characterization of Undrained Shear Strength. *Journal of Geotechnical and Geoenvironmental Engineering*, 140(6).

CEN (European Committee for Standardisation) (2004). EN-1997-1: Geotechnical design: Part 1, general rules. *Brussels, Belgium: CEN*.

Ching, J., Wu, S., & Phoon, K. K. (2021). Constructing quasi-site-specific multivariate probability distribution using hierarchical Bayesian model. *Journal of Engineering Mechanics*, 147(10), 04021069.

Emdal, A., Gylland, A., Amundsen H. A., Kåsin, K., and Long, M. (2016). Mini-block sampler. *Canadian Geotechnical Journal*, 53(8): 1235-1245.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. Bayesian Data Analysis. Boca Raton, FL, USA: *CRC Press*.

Gelman, A., Hill, J., & Vehtari, A. (2020). Regression and other stories. *Cambridge University Press*.

Hov, S., Garcia de Herreros, C. (2020). Jämförelse av miniblock- och standardkolvprovtagaren i en varvig östsvensk lera. *Branschsamverkan i grunden (BIG)*, projekt A2018-26. Trafikverket.

Karlsrud, K., Lunne, T., Kort, D. A., and Strandvik, S. (2005). CPTU correlations for clays. *Proceedings of the international conference on soil mechanics and geotechnical engineering, Rotterdam: Balkema*, 693–702.

Lefebvre, G., Poulin, C. (1979) A new method of sampling in sensitive clay. *Canadian Geotechnical Journal*, 16(1), 226–233.

Lunne, T., Berre, T., Andersen, K.H., Strandvik, S., Sjursen, M. (2006) Effects of sample disturbance and consolidation procedures on measured shear strength of soft marine Norwegian clays. *Canadian Geotechnical Journal*, 43(7).

Mayne P. W. and Peuchen, J. (2018) Evaluation of CPTU Nkt cone factor for undrained strength of clays. *Proceedings of the 4th International Symposium on Cone Penetration Testing*.

Paniagua, P., D'Ignazio, M., L'Heureux J. S., Lunne, T., and Karlsrud, K. (2019). CPTU correlations for Norwegian clays: an update. *Aims Geosciences*, 5(2).

Stan Development Team. 2019 Stan Modeling Language Users Guide and Reference Manual, 2.21.0. https://mc-stan.org

Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing*, 27(5), 1413-1432.