Integrating laboratory and geophysical data considering measurement errors

Jiawei Xie[1], *Jinsong* Huang[1#], and *Shan* Huang[1]

[1] *The University of Newcastle, Discipline of Civil, Surveying & Environmental Engineering, Callaghan, Australia*
[#]*Corresponding author: Jinsong.Huang@newcastle.edu.au*

ABSTRACT

Laboratory and geophysical tests are commonly used in site characterization. Combining these data sets based on empirical relationships can essentially enhance data interpretation. While in traditional approaches, the uncertainties in the relationship between these data sets are ignored. The Bayesian updating method is used to consider these uncertainties. Besides, the uncertainties due to measurement errors in the laboratory tests, particularly for preconsolidation pressure, are considered based on the kriging fitting method. The outcomes of kriging fitting are utilized to establish the prior distribution, and these outcomes are then compared against the baseline established by the trend fitting method. The Markov chain Monte Carlo (MCMC) algorithm is applied to incorporate the shear wave velocity measurements from a seismic dilatometer test to derive the posterior distribution. Bayesian updating of parameters considering measurement errors is able to get a more convincing design profile.

Keywords: Laboratory data; geophysical data; Bayesian updating; measurement errors.

1. Introduction

Multi-source data are commonly included in the process of site characterizations, such as geophysical, in-situ testing and laboratory data. While these data are usually considered as separated information or combined based on determining methods in site investigation programs. Geophysical tests provide an economical and quick way to identify subsoil stratigraphy. However, geophysical data can only provide indirect information to assess material parameters. To enhance the data interpretation, in-situ tests and laboratory data can be combined with geophysical data to interpret stratigraphic parameters.

Integration of different data sets can provide more reliable site characterization results, reducing the uncertainties associated with a single measurement (e.g., Xie et al. 2022). Commonly this integration is done manually and largely based on engineering judgment and experience, which will result in additional uncertainties. Probabilistic numerical methods can be used to combine multi-source data and extract more comprehensive information. Varied types of data may have different units and resolutions. Probabilistic numerical methods such as the Bayesian updating method can be used to increase the resolution of stratigraphic assessments by combining geophysics and high-quality laboratory measurements (Foti 2012). Bayesian updating method is well suited for geotechnical data fusion, particularly when limited information is available. Kriging interpolation method is widely used to estimate the values in unsampled locations (Huang et al. 2016, 2018) based on limited data. However, kriging interpolation technique assumes no measurement error is incurred, which is not the case in geotechnical engineering. To consider the uncertainties due to measurement errors in the test data, kriging fitting method is applied to overcome this deficiency.

In this paper, Bayesian updating method is used to combine in-situ shear wave velocity measurements and laboratory tests at the same location based on a real case study. The preconsolidation pressure values were obtained from constant rate of strain consolidation tests and the shear wave velocity data were obtained from a seismic dilatometer test. Related information can be found in the previous work in Huang et al. (2016). To consider the uncertainties due to measurement errors in the laboratory tests for the preconsolidation pressure, kriging fitting method is applied. The fitting results are used to derive the prior distribution of the preconsolidation pressures. Trend fitting method is also used as a baseline. The theory can be extended to two and three dimensions (Huang et al. 2018). The procedure is general and can be used for more realistic applications.

2. Integration framework

Bayesian updating method provides a useful framework for combining data sets by quantitatively analyzing uncertainties and providing safer predictions in risk-based design (Kelly and Huang 2015). Bayes' formula can be written as Eq. (1):

$$P(\theta \mid y) \propto P(y \mid \theta)P(\theta) \tag{1}$$

where $P(\theta)$ is the prior probability distribution of the material parameters, $P(y \mid \theta)$ is the probability of measurements y conditional on the material parameters θ, and $P(\theta \mid y)$ is the posterior distribution of the material parameters updated by the measurements.

2.1. Considering measurement errors in the prior distribution

Kriging method can be applied to get the prior probability distribution of the material parameters along with the depth (Huang et al. 2016). Kriging is often used to predict the unsampled locations based on the known values. The Kriging method averages over the errors to make the predictions spatial smoother (Griffiths and Fenton 2008). Kriging for data with measurement errors is called Kriging fitting.

For separating the random measurement errors associated with a particular test from the inherent spatial variability of the soil properties, Baecher (1982) proposed the following measurement error model for observation at x_s:

$$\boldsymbol{X}_\varepsilon(x_s) = \boldsymbol{X}(x_s) + \varepsilon(x_s) \tag{2}$$

where $\boldsymbol{X}(x_s)$ are the real soil properties at sampled locations, $\varepsilon(x_s)$ is a site-specific zero-mean measurement error with known variance of σ_m^2 for each measurement.

In the kriging fitting method, the weighting coefficients vector can be represented as:

$$\boldsymbol{\beta} = (\boldsymbol{C} + \boldsymbol{C}_m)^{-1}\boldsymbol{b} \tag{3}$$

where $\boldsymbol{C}$ is the matrix of covariances between the known values at the sampled locations. $\boldsymbol{b}$ is the vector of the covariances between the known values and the prediction locations. And $\boldsymbol{C}_m$ is the matrix of covariances between measurement errors at the sampled locations. $\boldsymbol{C}_m$ is a diagonal matrix which diagonal terms are the variances of measurement error σ_m^2 represents the variance of the known sample values. The kriging variance at each location can be expressed in a matrix form:

$$\sigma_K^2 = \sigma^2 - \boldsymbol{\beta}'\boldsymbol{b} \tag{4}$$

2.2. Likelihood function

Let $\boldsymbol{y}_i$ denotes a type of observation, the likelihood function includes the mechanical model $f(\boldsymbol{\theta})$ to convert the measurement to the property of concern. The transformed results are unavoidably different from the actual value due to the measurement error and model transfer errors, which can be known as mean error $\boldsymbol{\mu}_\varepsilon$. If this error is assumed to be normally distributed, the likelihood function of measurement $\boldsymbol{y}$ can be written as:

$$P(\boldsymbol{y}_i \mid \boldsymbol{\theta}) = \phi\left(\frac{y_i - f(\boldsymbol{\theta}) - \boldsymbol{\mu}_\varepsilon}{\sigma_\varepsilon}\right) \tag{5}$$

where σ_ε is the standard deviation of the mean error, and ϕ is the probability density function of the standard normal distribution.

2.3. Posterior distribution

Posterior distribution indicates the degree to which the real values are known when the new measurement is taken into account. The posterior distribution is often difficult to get the analytical representation. Therefore, the posterior distribution is sampled numerically by sampling method. MCMC is applied to sample the posterior distribution. In MCMC, a Markov chain is used to generate several steps in which a test realization is proposed for the posterior distribution parameters. Each sample is drawn from the probability distribution which is dependent upon the last samples. The chain will settle on the desired quantity which is independent of the initial startup implementation. The accepted realizations of the Markov Chain are sampled with their distribution corresponding to that of the posterior distribution (Beck and Au 2002).

3. Case study

A case study of combining preconsolidation pressures obtained from the laboratory test with the seismic data is provided. The real engineering data come from a soft soil test site in Ballina, New South Wales, Australia, which is operated by the Australian Research Council Centre of Excellence for Geotechnical Science and Engineering. Note that the data start from reduced level -1 mAHD to avoid the distortion of near-surface results. The semi-empirical relationship is used to convert the seismic data into preconsolidation pressure data. Then the Bayesian updating method is employed to combine multi-source data and get the more convincing posterior data. The kriging fitting method is used to estimate the prior distribution of preconsolidation pressures with measurement errors.

3.1. Prior interpretation of preconsolidation pressures

The laboratory test data is treated as prior information. Assuming the prior data has a log-normal distribution cause the pressure cannot be a negative value. The preconsolidation pressure is assumed to increase with a linear trend as the baseline, which can be achieved by fitting with the least square method. The standard deviation of the detrended data is 6.39kPa. Kriging fitting method is applied to consider the laboratory test measurement errors to get a better estimation of the random field between known data. The measurement errors are assumed to be 10% ($\sigma_m =$ 0.639 kPa) of the standard deviation of the detrended data (Yang et al., 2022). The prior standard deviation in each location can be calculated based on Eq. (4). To get the matrix of covariances used in Eq. (4), Markov correlation function is used in this paper due to its simplicity. The Markov correlation function is shown in Eq. (6):

$$\rho(\tau_{ij}) = e^{(-2\tau_{ij})/(\theta_Z)} \tag{6}$$

where $\rho(\tau_{ij})$ is correlation coefficient between properties assigned to two points i and j in the random field separated by an absolute distance τ_{ij}.

For one-dimensional problem, the scale of fluctuation is only applied to the depth direction, which is assumed to be 5m in this paper. Scale of fluctuation can be determined by fitting the autocorrelation function model if more data are available (Huang et al. 2018).

Linear trend fitting is compared with the kriging fitting method. From Fig. 1, it can be found that trend fitting cannot consider the spatial correlation between the known points and may lead to a large deviation at partial points. The kriging fitting method can take the measurement errors into consideration, which can reflect the prior distribution more reasonably.

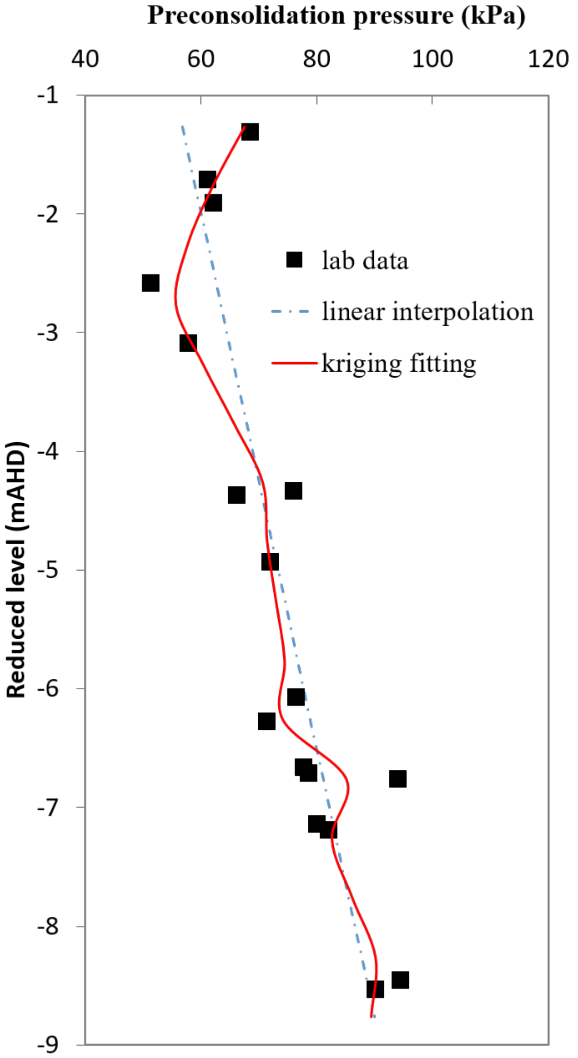

Figure 1. Prior distribution of preconsolidation pressures.

3.2. Transformation between preconsolidation pressures and shear wave velocity

Empirical relationship (Atkinson 2017) is used to correlate the preconsolidation pressure with the shear wave velocity (see Eq. (7)).

$$\frac{G_0}{p_a} = A\left(\frac{\sigma_v'}{p_a}\right)^n OCR^m \quad (7)$$

where G_0 is the small strain stiffness, p_a is normalizing pressure taken to be 1kpa, A, n and m are constants, σ_v' is the effective vertical stress. OCR is related to the preconsolidation pressure. The shear wave velocity is related to the G_0.

Eq. (7) was developed based on tests on reconstituted samples of several fine-grained soils (Viggiani and Atkinson, 1995). The relationship was shown to hold for both reconstituted and undisturbed clay under isotropic and anisotropic stress states not close to failure. This model function can be used to convert the geophysical data into a form that can be compared with more accurate laboratory test data. A relationship between preconsolidation pressure and shear wave velocity is obtained by adopting the values n = 0.9, m = 0.35 and A = 170 estimated from Atkinson (2017). Preconsolidation pressures assessed from a constant rate of compression tests and the ones interpreted from a seismic dilatometer test are compared in Fig. 2. The laboratory test data are considered as the prior set of material parameters and the seismic dilatometer data are the new observations.

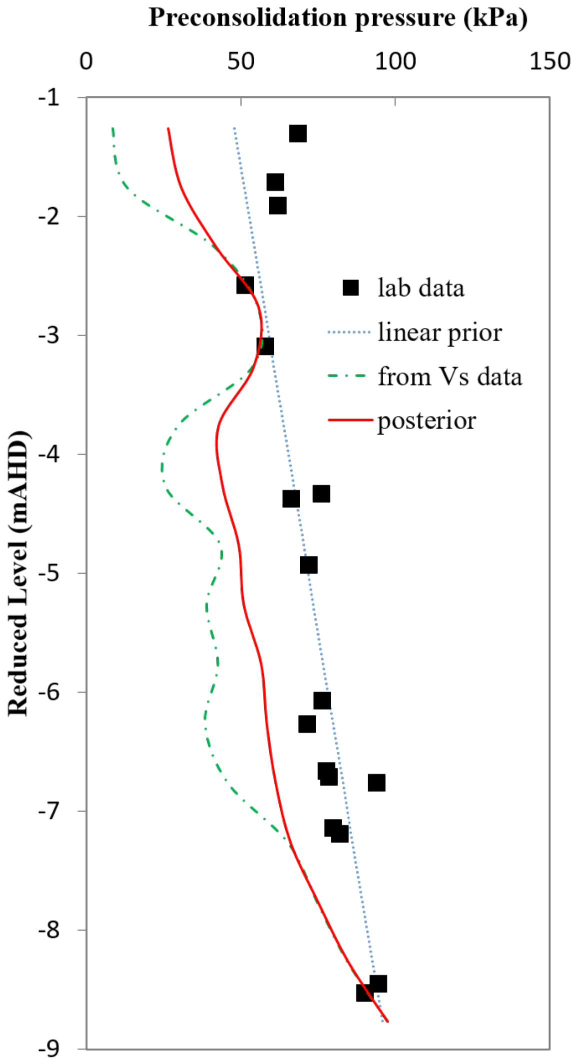

Figure 2. Posterior mean preconsolidation pressures using trend analysis as prior distribution.

3.3. Posterior preconsolidation pressure results

The posterior preconsolidation pressure results are highly related to the mean and standard deviation of the uncertainties in the seismic dilatometer data and the empirical model function. However, these uncertainties are hard to obtain for calculation. To illustrate the proposed framework, small standard deviation of measurement errors (1kPa) is allocated to the shear wave velocities (Foti 2012). The posterior solution will be dominated by the one with smaller errors in prior information and measurements.

The posterior preconsolidation pressure results based on the kriging fitting and trend analysis are shown in Figs. 2 and 3.

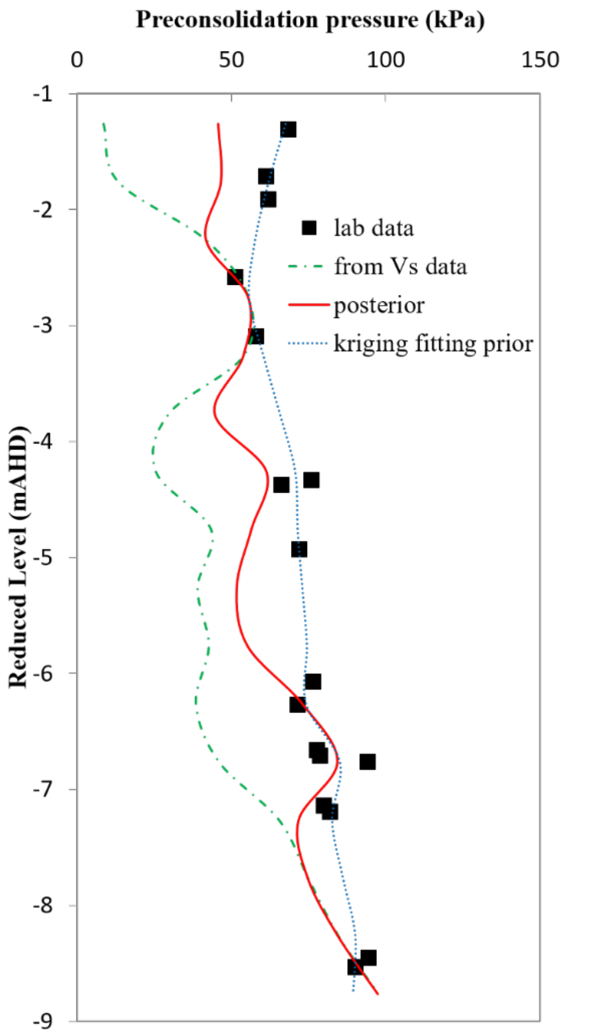

Figure 3. Posterior mean preconsolidation pressures using kriging fitting as prior distribution.

Figs. 2 and 3 reveal that the mean values of the posterior distributions updated from both trend fitting and kriging fitting methods are positioned between the laboratory and geophysical data. Notably, the data results from kriging fitting demonstrate more significant fluctuations in the depth direction, particularly in areas with laboratory test data. This observation can be attributed to the kriging fitting method's enhanced ability to leverage data from test points where available, effectively capturing the variability in those regions.

4. Discussion

The above sections demonstrated the details of combining multi-source data based on the Bayesian updating method. This framework is an enhancement of conventional methods based on engineering judgment and experience to combine multi-source data sets.

The Bayesian updating approach can be extended to include data both in the horizontal and vertical directions. Then the material parameters can be updated based on the laboratory and geophysical data in two or three dimensions so that the spatial relationship of the material grid points can be considered instead of assigning uniform distribution to a specific stratum. The biggest challenge in this framework is to quantify the model errors. Further research for determining the model errors can be found in Ching et al. (2016).

5. Conclusion

This paper developed a framework that is well-suited to enhance data interpretation based on multi-source geotechnical data. How to consider the errors for the prior and measurement data are discussed. Geotechnical data are used consistently and based on rigorous statistical principles; the uncertainties of pre-consolidation pressure can be reduced. However, more work needs to be done in determining the value of measurement errors.

Acknowledgements

This study was funded by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (Project DP220103381), the National Natural Science Foundation of China (No. 52222905, 52179103).

References

Atkinson, J., 2017. "The Mechanics of Soils and Foundations", 2nd ed., CRC Press, London, UK. https://doi.org/10.1201/9781315273549

Baecher, G.B., 1982. "Statistical methods in site characterization", Updating Subsurface Samplings of Soils and Rocks and Their In-situ Testing, pp. 463-492.

Beck, J.L., Au, S.K., 2002. "Bayesian Updating of Structural Models and Reliability using Markov Chain Monte Carlo Simulation", J. Eng. Mech., 128(4), pp. 380-391. http://doi.org/10.1061/(ASCE)0733-9399(2002)128:4(380).

Ching, J., Phoon, K.K., Wu, T.J., 2016. "Spatial correlation for transformation uncertainty and its applications", Georisk: Assess. Manag. Risk Eng. Syst. Geohaz., 10, pp. 294-311. http://doi.org/10.1080/17499518.2016.1205749.

Foti, S., 2012. "Combined use of geophysical methods in site characterization", In: Proceedings of the Fifth International Conference on Geotechnical and Geophysical Site Characterization (ISC-4), Brazil, pp. 43-61.

Griffiths, D., Fenton, G., 2008. "Risk Assessment in Geotechnical Engineering", Hoboken, New Jersey: John Wiley & Sons, Inc.

Huang, J., Kelly, R., Sloan, S.W., 2016. "Enhanced data interpretation: Combining in-situ test data by Bayesian updating", In: Proceedings of the Fifth International Conference on Geotechnical and Geophysical Site Characterization (ISC'5), Gold Coast, pp. 1437-1441.

Huang, J., Zheng, D., Li, D.Q., Kelly, R., Sloan, S.W., 2018. "Probabilistic characterization of two-dimensional soil profile by integrating cone penetration test (CPT) with multi-channel analysis of surface wave (MASW) data", Can. Geotech. J., 55, pp. 1168-1181. http://doi.org/10.1139/cgj-2017-0429

Kelly, R., Huang, J., 2015. "Bayesian updating for one-dimensional consolidation measurements", Can. Geotech. J., 52, pp. 1318-1330. http://doi.org/10.1139/cgj-2014-0338

Viggiani, G., Atkinson, J., 1995. "Stiffness of fine-grained soil at very small strains", Géotechnique, 45, pp. 249–265.

Xie, J., Huang, J., Lu, J., Burton, G.J., Zeng, C., Wang, Y., 2022. "Development of two-dimensional ground models by combining geotechnical and geophysical data", Eng. Geol., 300, pp. 106579. https://doi.org/10.1016/j.enggeo.2022.106579

Yang, R., Huang, J., Griffiths, D.V., 2022. "Optimal geotechnical site investigations for slope reliability assessment considering measurement errors", Eng. Geol., 297, pp. 106497. https://doi.org/10.1016/j.enggeo.2021.106497