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
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_13-007-cd

ISGSR 2022

Using Bayesian Updating to Improve Forecasts of Embankment Settlements on Soft Soils

Xiao Wan[1] and James Doherty[2]

[1]Department of Civil, Environmental and Mining Engineering, The University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Perth, Australia.
E-mail: xiao.wan@research.uwa.edu.au
[2]Department of Civil, Environmental and Mining Engineering, The University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Perth, Australia.
E-mail: james.doherty@uwa.edu.au

Abstract: In this paper, Bayesian updating is combined with a data-driven approach known as CRACA to provide settlement forecasts for embankments on soft soils. Oedometer test data are ingested directly in the CRACA method and processed automatically to ensure minimal subjective intervention. The uncertainties involved in the laboratory measurements and the automated interpretation process are represented using scaling factors with an initial value of unity (returning the prior). These scaling factors are treated as random variables in Bayesian updating using early-stage settlement monitoring data to improve future forecasts. This method is applied to an embankment constructed at the Bloemendalerpolder site, Netherlands, and is shown to provide a significant improvement in the accuracy of the four-year surface settlement forecast using settlement monitoring data over the first year, reducing the different between model and measured settlement from 8 % (prior) to 2%. Using 600 days of data, the forecast settlement matches the 4-year measured settlement almost precisely.

Keywords: Data-driven; Bayesian updating; Embankment; Settlement; Consolidation; Creep.

1 Introduction

Geotechnical models that rely on parameters that are subjectively evaluated/interpreted from laboratory and in situ test data have been shown to be unreliable. This has been demonstrated by a number of prediction exercises (Kelly et al., 2018; Doherty et al., 2018a; Lehane et al., 2008) which have shown engineers using the same data and the same model often end up with very different model parameters and therefore different model forecasts. Bayesian methods are becoming popular in geotechnical engineering to help improve model forecasts by using monitoring data as it becomes available. This type of Bayesian updating is particularly well suited to forecasting the settlement of embankments on soft soils. This is because embankments settle over a long period of time (several years) and there is potential to modify future construction activities (for example, when to remove surcharge) to take advantage of improved model forecasts (Kelly and Huang, 2015).

When Bayesian updating is combined with conventional parameter-based models the subjective parameters are treated as random variables. Therefore, the Bayesian analysis starts from a highly subjective prior, and it is unclear how this may influence the overall results of the analysis. This paper aims to overcome this subjective starting point by combining Bayesian updating with a data-driven method for predicting the settlement of embankments on soft soil. The data driven method, known as CRACA (i.e. CReep And Consolidation Analysis), was originally developed by Doherty and Bransby (2021) and was extended to account for multi-stage loading by Wan and Doherty (2022a) to improve the settlement forecast during and shortly after embankment construction. Importantly, the CRACA method directly ingests measured oedometer data and applies an automated and strictly defined procedure to generate a forecast of the time-settlement response of embankments on soft soils. This ensures that there is minimal user subjectivity in the application of this method. Therefore, the method may be applied in a consistent manner from one site to another with confidence that the user applying the method has minimal influence over the results. The forecasting capability of the process can therefore be fairly and objectively assessed. This is not the case with parameter-based models. Because CRACA does not use parameters, scaling factors are introduced that account uncertainty associated with the laboratory measurements and the automated interpretation process. These factors have an initial value of unity (returning the prior) and are updated in a Bayesian framework as settlement monitoring data is revealed over time to improve future forecasts. This method was developed by Wan and Doherty (2020b) and applied using an embankment case history at the Australia National Field-Testing Facility, in Ballina. This paper demonstrates a further application of this method for another case history at Bloemendalerpolder, Netherlands (Hoefsloot, 2015). The following sections of the paper describe the field test site and the available test data. The approach for incorporating Bayesian updating within the CRACA framework is then briefly described before performance of the model is demonstrated. All data used in this paper are freely available in digital format (Doherty et al., 2018b).

2 Application of CRACA at The Bloemendalerpolder Site

The Bloemendalerpolder site is located 10 km south-east of Amsterdam, the Netherlands. As part of the Geo-Impuls project (Cools, 2011; Van Staveren et al., 2013), two identical sand fill embankments were constructed in 2010. Figure 1 shows the plan view of one of the embankments and the cross-section is illustrated in Figure 2. This embankment is 3 m high, 36 m long and 26 m wide with a side-slope of 2H:1V. Three oedometer tests were conducted on samples obtained beneath this embankment. The test depths are indicated by red hollow circles in Figure 2 and listed in Table 1 together with other soil properties. The embankment was constructed with an initial 1 m lift, followed by 0.5 m lifts approximately every 3 weeks and the construction timeline is listed in Table 2. Prefabricated vertical drains (PVDs) were installed beneath the footprint of the embankments in a triangular grid with a spacing of 1.0 m. The groundwater table is around 0.2 m below the ground surface (Hoefsloot, 2015).

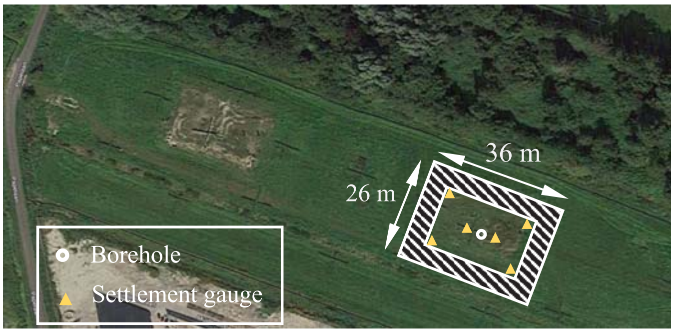

Figure 1. Plan view of the test embankment at Papelaan, Weesp, Netherlands

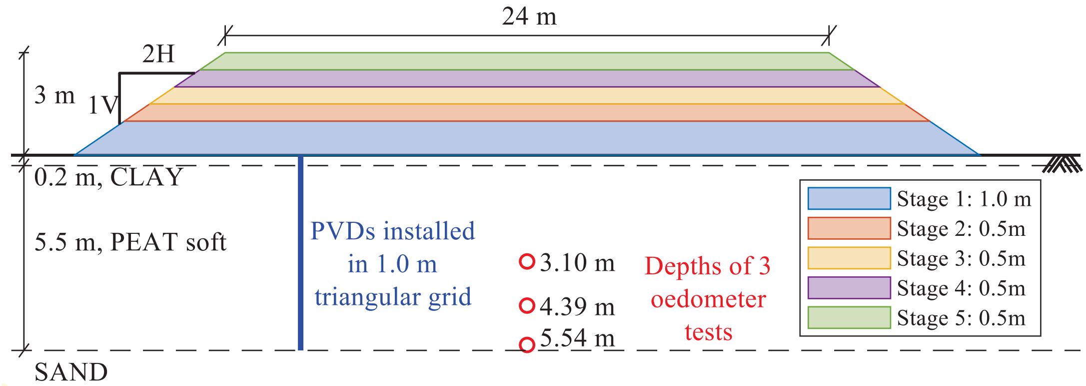

Figure 2. Illustration of embankment cross-section, ground conditions and depth of oedometer test samples

Table 1. Sample depths for each oedometer test

Oedometer test #	Sample depth (m)	Initial void ratio	Saturated unit weight (kN/m^3)
1	3.1	9.95	10.00
2	4.39	8.44	9.82
3	5.54	2.38	13.45

Table 2. The Blemendalerpolder site embankment construction timeline

Stage #	Day	Height increment (m)	Total Height (m)
1	0	1.0	1.0
2	25	0.5	1.5
3	45	0.5	2.0
4	90	0.5	2.5
5	111	0.5	3.0

In the CRACA method, the initial in-situ effective stress prior to the construction of the embankment is calculated using the unit weights measured from the oedometer tests and the ground water level. The final effective stress, assuming full consolidation, is calculated by adding the incremental stress due to the embankment. The change in total stress is divided into stages corresponding to the construction timeline using an elastic solution for a flexible strip load (Poulos and Davis 1974), where the load is based on the embankment height, width and the fill unit weight, which is taken as 17 kN/m^3 for this embankment.

With the known initial and final effective stresses profiles for each construction stage, a primary strain profile for each construction stage can be determined. This is achieved using the measured stress-strain response from each oedometer test and an automated interpolation procedure that projects the in-situ effective stress at the start and the final effective stress of a construction stage after full consolidation to determine corresponding strains. The difference in strain is taken as the primary strain increment ($\varepsilon^p{}_i$) for the i^{th} load stage. This process is repeated for all oedometer test depths to develop a primary strain profile for each construction stage ($\varepsilon^p{}_i(z)$), as shown in Figure 3a.

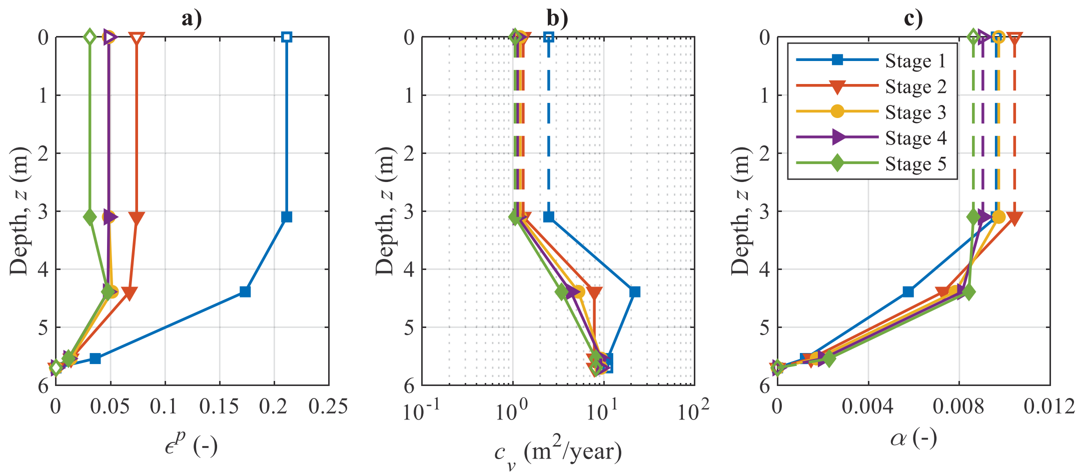

Figure 3. The profiles of a) increment primary strain, b) coefficient of consolidation and c) creep compression index

As described by Doherty and Bransby (2021), the primary strain at the ground surface is assumed to be identical to the value at the upper oedometer test and the strain value at the lower oedometer test is linearly extrapolated to a value of zero at the base of the consolidating material, which corresponds to the top of the sand layer. In this case, the closest sample depth to the ground is 3.1 m and the depth of the top of the sand layer is 5.7 m. These extrapolated points are indicated hollow markers in Figure 3 to distinguish them from measured values.

A consolation analysis is also performed using either a one-dimensional consolidation model, or an axisymmetric unit cell for embankments with PVDs. The coefficient of consolidation (c_v) profile for the consolidation model is determine automatically by evaluating c_v for each oedometer load increment using a numerical optimisation algorithm described in detail by Doherty and Bransby (2021). The optimised c_v values are paired with the average of the stress at the start and end of the oedometer load increment. By interpolation, the average stress values of each construction stage are used to determine the c_v value at the depth of the oedometer test sample. This is repeated at all oedometer test depths to form c_v profiles for each construction stage, which are plotted in Figure 3b. A similar approach is applied to determine the creep compression index, α, from the strain-time data, which are plotted in Figure 3c. As described by Doherty and Bransby (2021), the coefficient of consolidation at the base of the model and the ground surface and the creep compression index at the ground surface are assumed to be identical to the closest sample depth; the creep compression is linearly extrapolated from the lowest test depth to the value of zero at the base of the model, which are indicated as dash lines and hollow markers.

Given the coefficient of consolidation profile and the boundary conditions (i.e. the permeable top, bottom sand layer and the installed PVDs in triangular pattern with 1.0 m spacing), the time-dependent degree of consolidation for each construction stage can be determined using an axisymmetric unit cell model with free draining boundaries at the top, bottom and the axis of symmetry. The radius of the model was 0.53 m, which is based on the spacing of the PVDs. Independent consolidation models are established for each load stage.
The diffusion equations are solved using the c_v profile, the initial conditions and boundary conditions to model the decay in the excess pore pressure over time, resulting in the degree of consolidation ranging from 0 to 1. The degree of consolidation profiles are used to scale the primary strain profile (see Figure 3a) to evaluate the primary settlement over time. The primary settlement for i^{th} construction stage $s^p{}_i$ can be estimated by calculating the area of the strain profiles scaled by the degree of consolidation. The overall time-dependent primary settlement is then calculated as the superposition of primary settlement for each construction stage.

The creep profile, $\alpha_i(z)$ (see Figure 3c) and simulated time-dependent degree of consolidation profile can be used to estimate a creep strain profile for load stage i at sample depth z. The time dependent creep strain profile can be integrated between the top and bottom of the profile to estimate the creep surface settlement over time. To prevent double counting, the secondary settlement is taken as the maximum secondary settlement among all load stages.

The total surface settlement is the sum of primary settlement and creep/secondary settlement and is plotted with time in Figure 4, along with the measured embankment settlement.

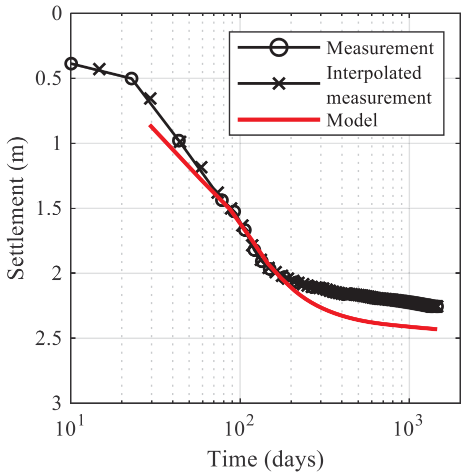

Figure 4. Time-dependent surface settlement comparison between the measurement and model

3 Combining the CRACA method with Bayesian updating

As monitoring data becomes available it can be utilized to update the long-term performance of the embankment and improve the reliability and accuracy of the forecast and therefore provide valuable information for decision making for the broader project. In Bayesian updating, both measured data and model parameters are assumed as random variables with a given mean, standard deviation, and probability distribution. As described above, in CRACA method, the embankment settlement depends on primary strain (ε_p), coefficient of consolidation (c_v) and creep index (α) profiles. As these values are determined using oedometer tests data in an automated way, there remains uncertainty with the measurements and the automated interpretation process. To capture this uncertainty and to incorporate a Bayesian updating, each of the profile values are scaled as follow

$$\varepsilon^i_p = \theta^i_\varepsilon \times \varepsilon^i_p;\ c^i_v = \theta^i_{cv} \times c^i_v;\ \alpha^i = \theta^i_\varepsilon \times \alpha^i \tag{1}$$

where i ranges from 1 to the number of points in the profile (n) (i.e. the number of oedometer tests). These factors are assembled into an overall scaling vector, $\boldsymbol{\theta}$. A scaling factors of unity will return an initial forecast prior to model updating with embankment monitoring data (i.e. Figure 4). According to Bayes' theorem, in-situ monitored data can be used to update the factors to improve the model forecast and the posterior probability is determined based on the likelihood given the monitored data and the prior probability as

$$P(\boldsymbol{\theta}|\boldsymbol{M}) = cL(\boldsymbol{\theta}|\boldsymbol{M})\, P(\boldsymbol{\theta}) \tag{2}$$

where c is a normalization constant and $\boldsymbol{M}$ contains the measured settlement values at particular observation times (i.e. $\boldsymbol{M} = \{m_1, m_2, \dots, m_N\}^T$, where N is the number of measurements). $P(\boldsymbol{\theta})$ is the prior probability before obtaining monitoring data, which can be determined as the product of the prior probability density value of each scaling factor. In this case, to prevent negative values, each scaling factor is assumed to follow a log-normal distribution with an assumed mean of unity and a standard deviation of 0.3. $L(\boldsymbol{\theta}|\boldsymbol{M})$ is the likelihood of $\boldsymbol{\theta}$ given $\boldsymbol{M}$ and determined as the product of individual likelihood of scaling vector given each measurement. For example, for the i^{th} measurement, $L_i(\boldsymbol{\theta}|m_i)$ is determined as the probability density of the difference between the calculated and the measured value by assuming it follows a normal distribution with zero mean and 0.1 standard deviation and can be calculated as

$$L_i\left(\boldsymbol{\theta}|m_i\right) = \frac{1}{\sqrt{2\pi} \times 0.1} exp\left(-\frac{\left(f_i - m_i\right)^2}{2 \times 0.1^2}\right) \tag{3}$$

where f_i is the i^{th} calculated value and i ranges from 1 to the number of observation times (N).

The likelihood value (Eq. (3)) and posterior probability (Eq. (2)) are both function of the scaling vector and with a better match between the calculated and measured settlement, the more likely these scaling factors are the optimal values and the larger the value of posterior probability.

Markov Chain Monte Carlo (MCMC) algorithms are often used in Bayesian analysis to generate samples for variables (in this case the scaling factors). Slice sampling is a type of MCMC algorithm that was first introduced by Neal (2003) to generate samples from complex multivariant distributions, such as the posterior distribution in Eq. (2). Slice sampling works through an iteration process and generates samples that have higher probabilities.

4 Application to Bloemendalerpolder Embankment Data

At the Bloemendalerpolder site, embankment surface settlement was measured at 25 irregularly spaced times over four years. To ensure equal weighting for each time period over the four years, these settlement measurements were converted into 100 uniformly spaced observations using linear interpolation (see the markers in Figure 4)

In order to study how the settlement forecast accuracy changes as more settlement observations becomes available after the embankment is constructed, 6 cases were considered. As summarised in Table 3, Case 1 used the first 180 days of monitoring data in Bayesian updating and the following cases revealed more data with Case 6 using 3 years of monitoring data.

Table 3. Cases summary

Data type	Time period (starting from the beginning of the construction up to)					
Measured surface settlement data	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
	180 days	270 days	1 year	1.5 years	2 years	3 years

For each of the cases, the iteration process starts with an initial scaling vector of unity and the first 100 samples are discarded to ensure stationarity of the Markov chain. A total 1000 scaling vector samples were generated using the slice sampler MCMC algorithm. The optimal scaling vector was taken as the mean of the 1000 samples for each case.

5 Results

Figure 5a shows the time-dependent surface settlement forecasts using the optimised factors of all cases. It can be seen that all of the updated settlement forecasts after Bayesian updating are improved compared to the prior forecast, especially for settlements using 200 or more days of monitoring data (Case 2 and above). According to Kelly et al. (2018), a settlement accuracy within ±20% is sufficient for practical purposes. The updated 4-year surface settlements of all cases are plotted in Figure 5b, where ethe horizontal axis represents the number of days of monitoring data used to update ethe model. The prior prediction is represented by the point at the time of 0 days. It shows that the updated forecasts for all cases are more accurate than the prior prediction and well within 20% of the measured settlement. As more monitoring data is revealed, the updated 4-year settlement forecast improves. Assuming these four-year settlement values follow a normal distribution, the 95% CI can be calculated (see Figure 5b). It shows that for all cases, the measured four-year surface settlement is always within the 95% CI.

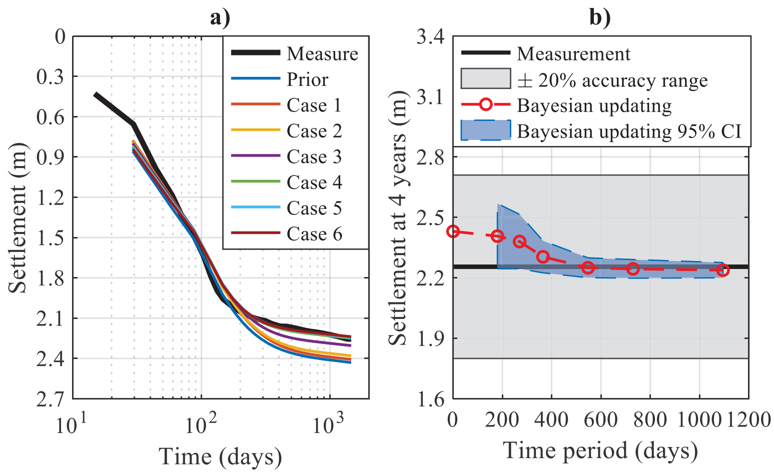

Figure 5. a) time-dependent surface settlement comparison between the measurement and the calculation after Bayesian updating; b) the 4-year surface settlement comparison between the measurement and the calculation after Bayesian updating and the 95% CI

6 Conclusion

In this paper, Bayesian updating is combined with the CRACA method to improve forecasts of embankment settlements over time using early-stage monitoring data. A significant advantage of this is that the prior forecast is generated using the automated and strictly defined CRACA method that ingests oedometer data directly. This ensures that there is minimal user subjectivity in the application of this method. Therefore, the method may be applied in a consistent manner from one site to another with confidence that the user applying the method has minimal influence over the results. The forecasting capability of the process can therefore be fairly and objectively assessed. This is clearly not the case with parameter-based models.

Data from an embankment constructed in the Bloemendalerpolder site was used to test the method and it was shown that the prior forecast was within around 8% of the measured settlement after 4 years. By feeding in monitored settlement data over the first year, this improved to around 2%. Using 600 days of data, the forecast settlement matched the 4-year measured settlement almost precisely.

Because the method enforces a strict procedure for processing laboratory data that minimizes subjective judgement, the range of possible forecasts for a given set of input data is likely to be small among different users. Therefore, the model can be tested in a meaningful and falsifiable way. The application of the method to the Bloemendalerpolder site is therefore another demonstration of this methods significant potential for improving embankment settlement forecasts, which may lead to improved decision making for important infrastructure projects.

Acknowledgment

The first author would like to acknowledge the University of Western Australia's support through a 'UWA International Fee Scholarship' and 'University Postgraduate Award'.

References

Cools, P. (2011). The Geo-Impuls Programme reducing geotechnical failure in the Netherlands. *Geotechnical Safety and Risk. ISGSR 2011*, 191-198.

Doherty, J. P., & Bransby, M. F. (2021). A data-driven approach for predicting the time-dependent settlement of embankments on soft soils. *Géotechnique*, 71(11), 1014-1027, https://doi.org/10.1680/jgeot.19.SiP.015.

Doherty, J. P., Gourvenec, S., & Gaone, F. M. (2018a). Insights from a shallow foundation load-settlement prediction exercise. *Comput Geotech*., 93, 269-279, https://doi.org/10.1016/j.compgeo.2017.05.009.

Doherty, J. P., Gourvenec, S., Gaone, F. M., Pineda, J. A., Kelly, R., O'Loughlin, C. D., Cassidy, M. J. & Sloan, S. W. (2018b). A novel web based application for storing, managing and sharing geotechnical data, illustrated using the national soft soil field testing facility in Ballina, Australia. *Comput Geotech*., 93, 3-8, https://doi.org/10.1016/j.compgeo.2017.05.007.

Hoefsloot, F. J. M. (2015). Long Term Monitoring Test Embankments Bloemendalerpolder–Geo-Impuls Program. In *Geotechnical Safety and Risk V (pp. 621-627). IOS Press*.

Kelly, R., & Huang, J. (2015). Bayesian updating for one-dimensional consolidation measurements. *Can. Geotech. J.*, *52*(9), 1318-1330, https://doi.org/10.1139/cgj-2014-0338@cgj-ec.2015.01.issue-4.

Kelly, R. B., Sloan, S. W., Pineda, J. A., Kouretzis, G., & Huang, J. (2018). Outcomes of the Newcastle symposium for the prediction of embankment behaviour on soft soil. *Comput Geotech*., *93*, 9-41, https://doi.org/10.1016/j.compgeo.2017.08.005.

Lehane, B., Doherty, J., & Schneider, J. A. (2008). Settlement prediction for footings on sand. *In Settlement prediction for footings on sand (pp. 133-150). IOS Press*.

Neal, R. M. (2003). Slice sampling. *The annals of statistics*, *31*(3), 705-767, https://doi.org/10.1214/aos/1056562461.

Poulos, H. G., & Davis, E. H. (1974). *Elastic solutions for soil and rock mechanics (Vol. 582). New York: Wiley*.

Van Staveren, M. T., Litjens, P. P. T., & Cools, P. M. C. B. M. (2013, September). Embedding Geo Risk Management. The Geo-Impuls Approach. *In Proceedings of the 18th International Conference on Soil Mechanics and Geotechnical Engineering (pp. 1847-1850)*.

Wan, X., & Doherty, J. P. (2022a). A data-driven approach for forecasting embankment settlement accounting for multi-stage construction. *Comput Geotech*., 152, https://doi.org/10.1016/j.compgeo.2022.105001.

Wan, X., & Doherty, J. P. (2022b). From data to decision: combining Bayesian updating with a data-driven prior to forecast the settlement of embankments on soft soils. *J. Geotech. Geoenviron. Eng*. Submitted under review.