Proceedings of ISFOG 2025
5TH INTERNATIONAL SYMPOSIUM ON
FRONTIERS IN OFFSHORE GEOTECHNICS
Nantes, France | June 9-13 2025

ISBN 978-2-85782-758-0

Bayesian inference for constitutive model calibration

C. Edmonds, G. Saturnino, U. Cilingir*
Ørsted, Gentofte, Denmark

**ULACI@orsted.com*

ABSTRACT: This paper presents an automated method for calibrating several parameters in constitutive soil models using Bayesian inference. The probabilistic machine learning method combines triaxial test data with prior geotechnical knowledge to estimate constitutive model parameters and quantify their uncertainties using a Markov Chain Monte Carlo algorithm. The method has been tested and validated on two different constitutive models: the phenomenological non-linear Mohr-Coulomb model for sands and the critical state Modified Cam-Clay model for clays with Hvorslev surface. The known parameter values (from artificially generated data or previous projects) consistently reside within the 95% credible intervals, with the estimated means frequently approximating the known values. The approach can easily be extended to other constitutive models and experimental data types as it does not require knowing gradients with respect to model parameters.

Keywords: Bayesian analysis; constitutive modelling; model calibration

1 INTRODUCTION

Constitutive models are used to model the complex behaviours of soils and simulate ground conditions for various applications, including in the geotechnical modelling of offshore site conditions. The calibration of constitutive models can be a repetitive and manual process which requires inputs from in-situ measurements and laboratory data.

Bayesian inference is a probabilistic method used in machine learning which can provide automatic estimates and credible intervals for the parameters in such constitutive models.

Various optimization methods (Machaček et al, 2022) and machine learning methods (Zhang et al, 2023) have previously been applied to the task of constitutive model calibration. Our paper suggests a gradient-free method which provides reliable results on sparse data and includes uncertainty quantification.

2 METHODS

In this paper we will consider two constitutive models: a phenomenological model for sands, and a critical state model for clays.

The non-linear isotropic hypoelastic model coupled by a strain hardening/softening Mohr-Coulomb failure criterion for sands is inspired by Potts and Zdravkovic (1999) and referred to as the Non-linear Ørsted Mohr-Coulomb Model (NØMM) in the remainder of this paper. The NØMM model combines Potts and Zdravkovic (1999) strain dependent hardening-softening Mohr-Coulomb model with a Kondner and Zelasko (1963) type hyperbolic model generalized small strain stiffness model. The hardening and softening of the yield and potential surfaces are controlled by accumulated total deviatoric strain as shown in Figure 1.

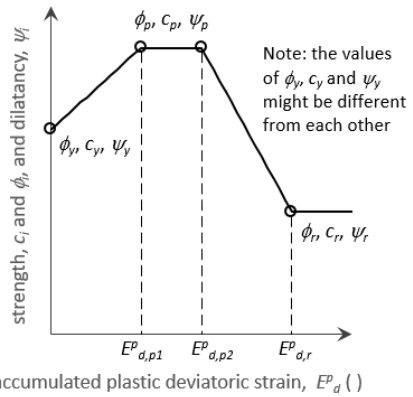

Figure 1. Tri-linear isotropic hardening/softening rule for strength parameters, c_i, ϕ_i and dilatancy ψ_i with accumulated deviatoric total strain $E^p_{d,p1}$, $E^p_{d,p2}$ and $E^p_{d,r}$ at first yield, peak, and residual respectively. Modified from Potts and Zdravkovic (1999).

For clays, we consider an elastoplastic strain hardening/softening model with a Hvorslev surface named Modified Cam-Clay (MCC) (Tsiampousi et al, 2013), based on critical state theory.

For the NØMM model, we focus on calibrating the plastic parameters, such as the friction angle at the first yielding, peak, and residual state, and the plastic deviatoric strain at the start, end of the peak and the start of the residual state. For the MCC model, we seek to calibrate empirical Hvorslev surface parameters ($m, n, \alpha,$

and β) and the Yield Stress Ratio (R) for each test to prove the concept and for brevity of the presentation.

2.1 Data preparation

In the case of triaxial tests, the recorded laboratory data y and corresponding simulation data $\hat{y}(\theta)$ can be either deviatoric stresses, volumetric strains (in drained tests), excess pore water pressure (in undrained tests), or a linear combination of these. The results reported below are based on using deviatoric stresses only, but similar results have also been obtained with volumetric strains and excess porewater pressure. To ensure comparable results across pressure conditions, all data in a given test were normalized by the maximum value in that test.

2.2 Theory of Bayesian Inference

Bayesian inference is based on Bayes' theorem. The posterior distribution $p(\theta|y)$ for parameter θ given data y is updated using the prior probability $p(\theta)$ of observing parameter θ and the likelihood $p(y|\theta)$ of observing data y given the parameter θ

$$p(\theta|y) \propto p(y|\theta)p(\theta) \quad (1)$$

In geotechnical applications, where many constitutive model parameters θ need to be calibrated simultaneously, computing the posterior distribution $p(\theta|y)$ directly using Equation 1 becomes intractable. Instead, a Markov Chain Monte Carlo (MCMC) algorithm is used to representatively sample the posterior distribution $p(\theta|y)$.

2.3 Markov Chain Monte Carlo algorithm

A MCMC algorithm with a Metropolis-Hastings step was used to sample the posterior distribution $p(\theta|y)$ and obtain probability distributions for the constitutive model parameters θ. The algorithm is initiated with random values for the parameters θ within a set of user-defined boundaries. In each step of the algorithm, new parameter values for θ are proposed and either accepted or rejected in a Metropolis-Hastings step.

In detail, a normal jumping distribution was used to propose the parameters θ to be assessed in each step. The jumping distribution has a mean equal to the previous parameter value, and a variance which is either user-defined or automatically defined. If the proposed parameter values lie outside the user-defined boundaries, these values are discarded and new values generated.

The prior distributions $p(\theta)$ for the various parameters are assumed to be independent and uniform for each parameter θ being estimated. The bounds of these prior distributions are set by the user, thereby offering the opportunity to encode prior knowledge about typical parameter ranges into the algorithm. When data is scarce, the prior distribution has greater impact on inference, emphasising the importance of selecting reasonable bounds.

The likelihood distribution measures how likely observing y in the laboratory is, given the model parameters θ. To calculate the likelihood, we first evaluate the simulated response of the constitutive model $\hat{y}(\theta)$ assuming parameter values θ. Then, the difference between the data recorded in the laboratory y and the simulated response $\hat{y}(\theta)$ is quantified using the root mean square error (RMSE)

$$L_i(\theta) = \sqrt{\frac{1}{M_i}\sum_{j=1}^{M_i}\left(y_{i,j} - \hat{y}_{i,j}(\theta)\right)^2} \quad (2)$$

where $y_{i,j}$ is the laboratory data for step j and laboratory test i, $\hat{y}_{i,j}(\theta)$ is the simulated soil response with model parameters θ and same conditions as test i, and M_i is the number of data points for the test i.

The loss term $L_i(\theta)$ defined in Equation 2 is then employed in the likelihood distribution, which is a half normal distribution with mean 0 and variance v_i^2

$$p(y|\theta) = \prod_{i=1}^{N} HN\left(L_i(\theta) \mid \mu = 0, v_i^2\right). \quad (3)$$

This half normal distribution assigns higher probabilities to low RMSE values $L_i(\theta)$, to encourage learning parameter values θ that produce a simulated response $\hat{y}(\theta)$ that is close to true laboratory data y.

In the Metropolis-Hastings step, the posterior distribution is evaluated using Equations 1 to 3 and either accepted or rejected with a certain probability.

The MCMC algorithm is updated in this manner for 4000 steps, with the first half discarded as the burn-in period. Parameter values for the second half are aggregated into histograms and thus represent posterior distributions for the parameters θ. The accepted parameter values in the second half of the chain are saved and used to compute percentiles, forming two-sided Bayesian credible intervals for $\hat{y}(\theta)$ based on the marginal posterior distributions. As the Metropolis-Hastings step contains an inherent element of randomness, four Markov chains were run and averaged to obtain more accurate results.

Convergence of the MCMC algorithm was assessed using the R-hat measure derived by Gelman et al (2013). The R-hat measure compares the variance of the estimated parameters between and within the four Markov chains, checking whether the chains have

mixed well and are sampling from the same posterior distribution. No convergence issues were observed.

3 RESULTS

We first present results for calibrating parameters of the NØMM model, followed by results for the MCC model. For each model, we include results from simulated data, as well as laboratory data from offshore sites in Europe and the United States (US).

3.1 NØMM model

We have used the Bayesian inference algorithm described in Section 2 to obtain estimates for two plastic parameters of the phenomenological NØMM model: peak friction angle ϕ_p and the cumulative plastic deviatoric strain at the start of the peak state $E^p_{d,p1}$.

As a preliminary test of the algorithm before applying it to true triaxial data, artificial data was generated by simulating drained triaxial compression of a soil sample at two confining pressures ($p_0 = 100$ kPa and $p_0 = 200$ kPa) with set parameter values of $\phi_p = 38°$ and $E^p_{d,p1} = 0.05$. White Gaussian noise was added to the computed deviatoric stress to imitate variations seen in laboratory data, as seen in Figure 2. We adopted uniform prior distributions in the ranges $[30°,\ 45°]$ for ϕ_p and $[0.001, 0.20]$ for $E^p_{d,p1}$. Other NØMM model parameters are assumed to be known exactly.

The posterior distributions obtained for ϕ_p and $E^p_{d,p1}$ are shown in Figure 3 and Table 1. Inferred and reference values for two parameters of the NØMM model. As only two parameters from the constitutive model are being inferred, it is also possible to obtain estimates of the posterior distributions using a grid search. The grid search solutions are shown with the red lines in Figure 3 and resemble the posterior distributions obtained using the MCMC algorithm. Narrow Bayesian 95% credible intervals are observed in Figure 2.

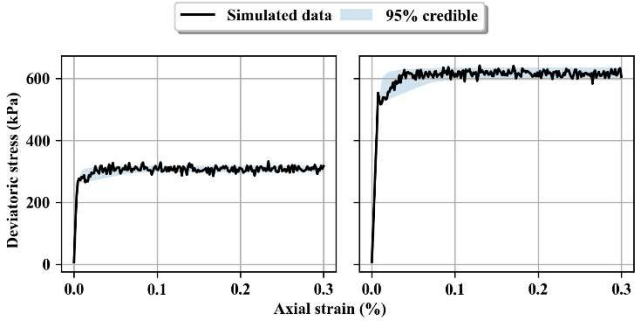

Figure 2. Simulated triaxial data used as input, with 95% credible intervals as determined by the algorithm.

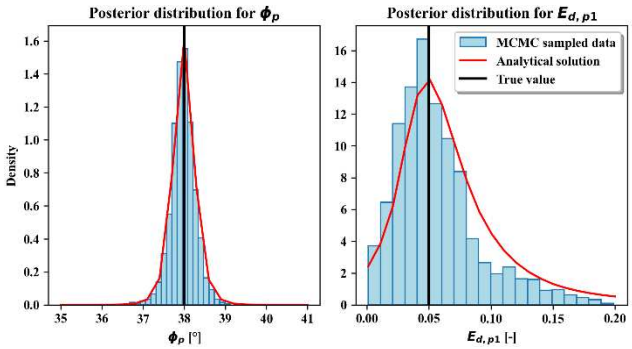

Figure 3. Posterior distributions estimated by the Bayesian inference algorithm for the ϕ_p and $E^p_{d,p1}$ parameters of the NØMM model.

Table 1. Inferred and reference values for two parameters of the NØMM model.

Parameter	Reference value	Mean	95% credible interval
ϕ_p	38.00	37.97	(37.38, 38.55)
$E^p_{d,p1}$	0.050	0.057	(0.009, 0.149)

3.1.1 Application to US offshore project

Next, we apply the NØMM model to three representative triaxial tests for a sand type from an offshore project on the East coast of the United States. The material is a dense, fine to medium sand occasionally with gravel. The parameters to be calibrated are extended to include six plastic parameters: the friction angle at the first yielding state ϕ_y, peak state ϕ_p, and residual state ϕ_r; and the plastic deviatoric strain at the start of the peak state $E^p_{d,p1}$, end of the peak state $E^p_{d,p2}$, and start of the residual state $E^p_{d,pr}$. The bounds selected for the uniform prior distributions of these six parameters are detailed in Table 2. Uniform prior distribution bounds for six plastic parameters of the NØMM model.

Table 2. Uniform prior distribution bounds for six plastic parameters of the NØMM model.

Parameter	Uniform prior distribution bounds
ϕ_y	$[20°,\ 60°]$
ϕ_p	$[20°,\ 60°]$
ϕ_r	$[20°,\ 60°]$
$E^p_{d,p1}$	[0.0001, 0.02]
$E^p_{d,p2}$	[0.02, 0.12]
$E^p_{d,pr}$	[0.12, 0.25]

The posterior distributions obtained using the Bayesian inference algorithm for these six plastic parameters are shown in Figure 4 and summarized in Table 3. The engineer-chosen values, which were picked independently of these results, all lie within the 95% credible intervals. For the ϕ_p, ϕ_r, and $E^p_{d,p2}$ parameters, the means of the posterior distributions lie

close to the engineer-chosen values. Wider posterior distributions are observed for the ϕ_y and $E^p_{d,p1}$ parameters, which describe the start of the peak state. This suggests that precise values for these two parameters contribute less to determining the deviatoric stress response to triaxial loading than parameters with sharply peaked distributions, such as ϕ_p and ϕ_r, for the case where we focus on comparing the response for the whole range of strain values.

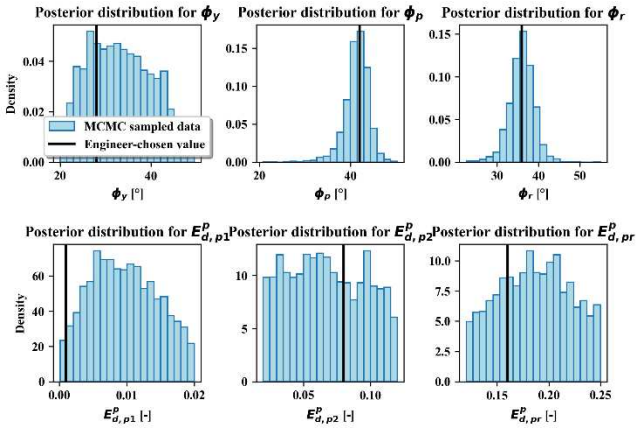

Figure 4. Posterior distributions and engineer-chosen values for six plastic parameters of the NØMM model, based on two triaxial tests from a sand unit in an offshore US project.

Table 3. Inferred and engineer-chosen values and credible intervals for six plastic parameters of the NØMM model, based on triaxial tests from a sand unit in an offshore US project.

Parameter	Engineer-chosen value	Mean (Inferred)	95% credible interval
ϕ_y	28	33.2	(21.4, 45.6)
ϕ_p	42	41.3	(34.5, 45.9)
ϕ_r	36	36.0	(29.7, 41.7)
$E^p_{d,p1}$	0.001	0.010	(0.001, 0.018)
$E^p_{d,p2}$	0.08	0.07	(0.02, 0.12)
$E^p_{d,pr}$	0.16	0.19	(0.13, 0.25)

The three triaxial tests used for the Bayesian inference algorithm are displayed in Figure 5. The simulation of mean parameters from the Bayesian inference algorithm, as well as the simulated response with the engineer-chosen values are also shown and resemble each other closely.

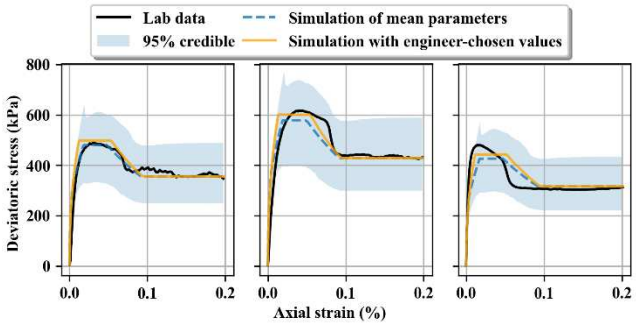

Figure 5. Laboratory data, simulated response with the mean inferred parameters, and simulated response with engineer-chosen parameters, and 95% credible intervals for sand in offshore US project.

3.2 MCC model

The Bayesian inference algorithm described in Section 2 was applied to infer empirical parameters of the critical state MCC model for clays.

The soil response to undrained triaxial compression with a confining pressure of 100 kPa and two different initial states (yield stress ratio) $R_1 = 40$ and $R_2 = 30$ and Hvorslev surface parameters ($\beta = 0.4$ and $m = 1.0$) was simulated and used as input for the Bayesian inference algorithm. Uniform prior distributions were adopted with bounds shown in Table 4.

Table 4. Uniform prior distribution bounds for four parameters of the MCC model on simulated response.

Parameter	Uniform prior distribution bounds
β	(0.15, 0.70)
m	(0.30, 1.50)
R_1	(30, 50)
R_2	(20, 40)

The posterior distributions determined by the Bayesian inference algorithm are compared to the true values in Figure 6 and Table 5. Note that the yield stress ratio R is calibrated individually for each of the two triaxial tests (R_1 for Test 1 and R_2 for Test 2), while the α and n parameters are calibrated using both triaxial tests.

The distributions for β, R_1, and R_2 are peaked around the reference values, whereas the distribution for m is much wider. This suggests that the algorithm is not able to access the true m value with precision, as this parameter has less importance in determining the soil deviatoric stress response to loading under triaxial conditions.

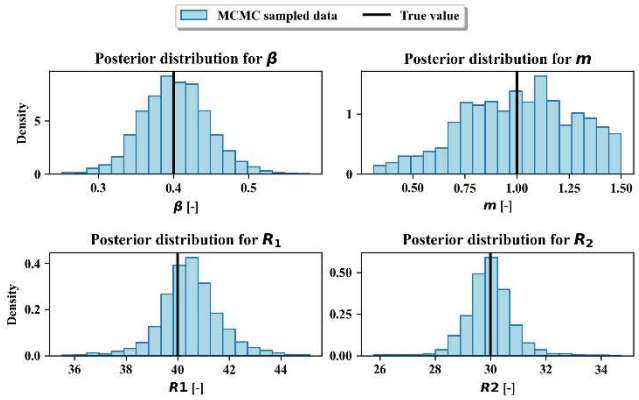

Figure 6. Posterior distributions and reference values for $\beta, m, R_1,$ and R_2 MCC model parameters.

Table 5. Inferred and reference values and credible intervals for Hvorslev surface and initial state parameters for MCC model on simulated data.

Parameter	Reference value	Mean (Inferred)	95% credible interval
$\boldsymbol{\beta}$	0.40	0.40	(0.31, 0.49)
$\boldsymbol{m}$	1.00	1.01	(0.45, 1.46)
$\boldsymbol{R_1}$	40.0	40.5	(38.1, 43.0)
$\boldsymbol{R_2}$	30.0	30.0	(28.4, 31.8)

The two triaxial tests used for the Bayesian inference algorithm are shown in Figure 7 with 95% credible intervals. The credible intervals are narrow, reflecting narrow posterior distributions and low levels of noise in the data.

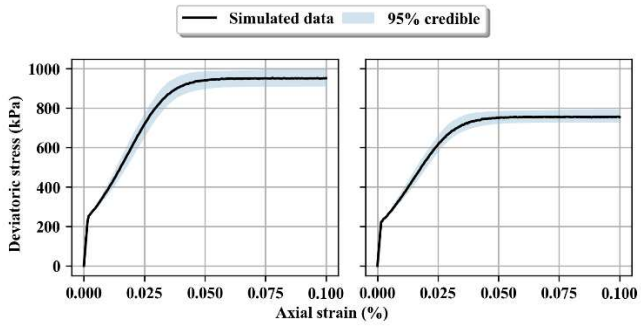

Figure 7. Reference data and simulated response with the mean inferred parameters, reference parameters, and 95% credible interval for soil response for the MCC model.

3.2.1 Application to North Sea offshore project

Two triaxial tests from a clay unit located in the North Sea were chosen as representative samples from an offshore project. The clay is overconsolidated, low to medium plasticity clay. The parameters α and n, which describe the shape of the Hvorslev surface, and the yield stress ratio R were inferred using the Bayesian Inference algorithm. The prior distributions were assumed to be uniform within the bounds detailed in Table 6.

Table 6. Uniform prior distribution bounds for four parameters of the MCC model on triaxial response from North Sea clay.

Parameter	Uniform prior distribution bounds
$\boldsymbol{\alpha}$	(0.05, 0.50)
$\boldsymbol{\beta}$	(0.05, 0.70)
$\boldsymbol{R_1}$	(1.1, 5.0)
$\boldsymbol{R_2}$	(1.1, 5.0)

The posterior distributions shown in Figure 8 and summarized in Table 7 reflect that not all parameters are equally easy to estimate, as the distributions for α and n are much wider than the distributions for yield stress ratios R_1 and R_2. The engineer-chosen values for all four parameters lie within the 95% Bayesian credible intervals of the posterior distributions, but the means of the posterior distributions are not as close to the engineer-chosen values compared to results presented in the above section.

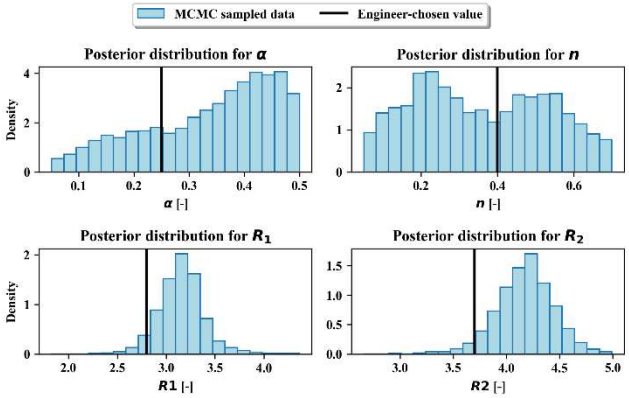

Figure 8. Posterior distributions and engineer-chosen values for $\alpha, n, R_1,$ and R_2 MCC model parameters, for two triaxial tests from North Sea project.

Table 7. Inferred and engineer-chosen values for Hvorslev surface and initial state parameters for MCC model on triaxial data from North Sea project.

Parameter	Engineer-chosen value	Mean	95% credible interval
$\boldsymbol{\alpha}$	0.25	0.34	(0.09, 0.49)
$\boldsymbol{n}$	0.40	0.36	(0.08, 0.67)
$\boldsymbol{R_1}$	2.8	3.2	(2.7, 3.7)
$\boldsymbol{R_2}$	3.7	4.2	(3.6, 4.7)

The 95% credible intervals and triaxial test data used as input are shown in Figure 9. In this case, the simulated response assuming the mean parameters from the Bayesian inference algorithm lie closer to the laboratory data than the engineer-chosen values. It should be noted that the engineer-chosen values reflect the results from a calibration strategy that is based on the ground model characteristics and not merely as an

exercise to find the best matches to the triaxial test results.

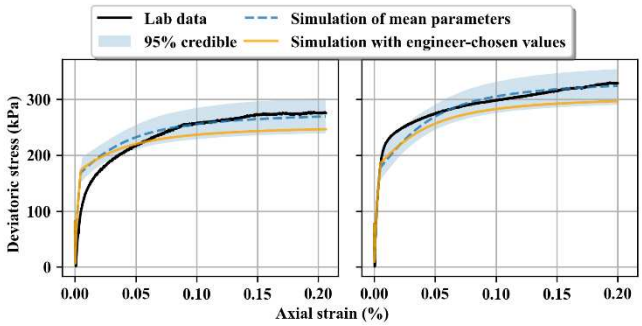

Figure 9. Laboratory data, simulated response with the mean inferred parameters, and simulated response with engineer-chosen parameters, and 95% credible intervals for clay in North Sea project.

4 CONCLUSION

A Bayesian inference algorithm which can reliably estimate constitutive model parameters on sparse laboratory data has been developed. The algorithm uses Markov Chain Monte Carlo simulations to infer posterior distributions and quantify uncertainties in parameter estimates.

The Bayesian inference algorithm has been tested and validated on two constitutive models: the phenomenological non-linear Mohr-Coulomb model for sands and the critical state Modified Cam-Clay model for clays with Hvorslev surface.

The work presented in this paper opens opportunities for further research in the automatic calibration of additional model parameters, including elastic parameters, and various constitutive models. Future work could explore the impact of modifying the boundaries of assumed uniform prior distributions or testing alternative prior distributions on the results. Different likelihood functions could also be investigated to tailor results to specific requirements.

The implementation of this algorithm is recommended in cases where many parameters need to be calibrated at once, and the algorithm can be used to suggest initial estimates, including uncertainties, to guide engineers in their calibration of constitutive models. The algorithm is easy to generalize for new constitutive models as it does not require knowing gradients with respect to model parameters.

AUTHOR CONTRIBUTION STATEMENT

Clara Edmonds: Investigation, Software, Formal Analysis, Writing - Original draft and Editing. **Guilherme B. Saturnino**: Conceptualization, Methodology, Software. **Ulas Cilingir**: Conceptualization, Supervision, Writing - Reviewing and Editing.

ACKNOWLEDGEMENTS

The authors would like to thank their colleagues at Ørsted for their support: V. Avgerinos, Y. Zhou, S. Kularathna, and R. Baginski.

REFERENCES

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., & Rubin, D.B. (2013). *Bayesian Data Analysis*, 3rd ed., Chapman and Hall/CRC. https://doi.org/10.1201/b16018

Kondner, R.L. and Zelasko, J.S. (1963). A hyperbolic stress-strain formulation for sands, in: *Proceedings of the 2nd Pan American Conference on Soil Mechanics and Foundations Engineering*, São Paulo, Brazil, vol. I, 289-324.

Machaček, J., Staubach, P., Tavera, C.E.G., et al (2022). On the automatic parameter calibration of a hypoplastic soil model, *Acta Geotech.* 17, pp. 5253–5273. https://doi.org/10.1007/s11440-022-01669-4

Potts, D.M. and Zdravkovic, L. (1999). *Finite element analysis in geotechnical engineering*, Thomas Telford Publishing, London, U.K..

Tsiampousi, A., Zdravkovic, L. and Potts, D.M. (2013). A new Hvorslev surface for critical state type unsaturated and saturated constitutive models, *Computers and Geotechnics,* 48, pp. 156–166.

Zhang, P., Yin, Z.Y., Sheil, B. (2023). Interpretable data-driven constitutive modelling of soils with sparse data, *Comput. Geotech.* 160, pp. 105511. https://doi.org/10.1016/j. compgeo.2023.105511