Probabilistic Calibration and Prediction of Seismic Soil Liquefaction using quoFEM

Aakash Bangalore Satish[1*[0000-0003-4016-9903]], Sang-ri Yi[1[0000-0001-7196-127X]]
Adithya Salil Nair[2], Pedro Arduino[3]

[1] University of California, Berkeley, Berkeley CA 94804, USA
bsaakash@berkeley.edu
[2] The Ohio State University, Columbus OH 43210, USA
[3] University of Washington, Seattle WA 98195, USA

Abstract. Liquefaction under cyclic loads can be predicted through advanced (liquefaction-capable) material constitutive models. However, such constitutive models have several input parameters whose values are often unknown or imprecisely known, requiring calibration via lab/in-situ test data. This study proposes a Bayesian updating framework that integrates probabilistic calibration of the soil model and probabilistic prediction of lateral spreading due to seismic liquefaction. In particular, the framework consists of three main parts: (1) Parametric study based on global sensitivity analysis, (2) Bayesian calibration of the primary input parameters of the constitutive model, and (3) Forward uncertainty propagation through a computational model simulating the response of a soil column under earthquake loading. For demonstration, the PM4Sand model is adopted, and cyclic strength data of Ottawa F-65 sand from cyclic direct simple shear tests are utilized to calibrate the model. The three main uncertainty analyses are performed using quoFEM, a SimCenter open-source software application for uncertainty quantification and optimization in the field of natural hazard engineering. The results demonstrate the potential of the framework linked with quoFEM to perform calibration and uncertainty propagation using sophisticated simulation models that can be part of a performance-based design workflow.

Keywords: liquefaction, PM4Sand, Bayesian calibration, sensitivity analysis, uncertainty propagation, quoFEM

1 Introduction

Earthquake-induced soil liquefaction can lead to unexpected casualties and property loss, and therefore, it is essential to predict its risk in advance. Soil behavior under cyclic loads can be simulated using liquefaction-capable constitutive models [1]. For example, PM4Sand is a sand plasticity model capable of simulating liquefaction under a broad mix of conditions in the field, including a wide range of density, shear stress, confining stress, and drainage/loading conditions [2]. This model has been implemented in numerical analysis software, including the open-source computational framework OpenSees [3,4]. The flexibility of such a model is attained through a large number of parameters (although most of them are typically set to their recommended default values) as well as a wide search space of the possible parameter combinations [5]. While in-situ or lab test results provide information for calibrating unknown parameters [6], some challenges remain in using deterministic calibration methods. For example, the large search space of input parameters and potential multimodality of the calibration objective function make the optimizer susceptible to fall in local optima. Further, soil properties often can only be measured indirectly and reveal high spatial variability, thus limiting the credibility of parameter values estimated from limited test data. Consequently, after identifying a combination of optimal parameter values, a significant amount of uncertainty remains in the estimated parameter values as well as in the response predicted by the model [7]. Therefore, there is a need for adopting probabilistic calibration methods and performing forward uncertainty propagation. For example, by introducing the Bayesian calibration approach, correlations and interactions between parameters as well as multimodality can be captured in terms of multiple near-optimal parameter combinations, i.e., posterior distribution or samples, and such probabilistic representation allows prediction of the uncertainty propagating to the liquefaction-induced lateral spreading at a site of interest [8,9,10]. Additionally, probability-based sensitivity analysis of input parameters allows us to identify the importance of each parameter while taking inherent uncertainty into account [11,12].

This study proposes a systemized approach for probabilistic liquefaction prediction updating that consists of three steps: parametric study, parameter calibration based on experimental data, and response prediction. Further, it shows that each probabilistic analysis can be greatly accelerated by using the research tool quoFEM developed by the NHERI SimCenter at UC Berkeley [13]; this is an open-source software application developed to assist researchers and practitioners in the field of natural hazard engineering. quoFEM allows users to link different simulation engines, including OpenSees and FEAP [14] and advanced uncertainty quantification (UQ) and optimization methods with a user interface. Each step of the framework introduces variance-based global sensitivity analysis, transitional Markov chain Monte Carlo-based Bayesian updating, and forward resampling algorithms, respectively, among other alternatives supported in quoFEM. Cyclic Direct Simple Shear (CyDSS) test data of Ottawa F-65 sand from Morales *et al.* (2021) is utilized to calibrate the PM4Sand model to match the prediction of the cyclic strength curve (i.e., the number of cycles to the onset of liquefaction given the cyclic shear stress ratio) from the model to that obtained from the tests [3,15]. All the analyses are conducted on high-performance computers at DesignSafe-CI using the quoFEM user interface.

2 Probabilistic framework for calibration and prediction of lateral spreading induced by seismic soil liquefaction

The systematic framework for updating predictions of lateral spreading due to soil liquefaction given experimental data consists of three main steps: (1) Variance-based global sensitivity analysis to identify the influence of the primary input parameters on the onset of liquefaction under cyclic loading, (2) Bayesian calibration of the primary input parameters of a soil constitutive model (PM4Sand in this case), and (3) forward uncertainty propagation to update the probability distribution of the lateral spreading under earthquake loading with the updated distribution of parameters.

Global Sensitivity Analysis. Variance-based sensitivity analysis is first introduced for a preliminary parametric study of the soil constitutive model. By identifying the proportion of response variance attributed to each parameter, the sensitivity measures, also known as Sobol indices, are assessed. The main Sobol index accounts for the contribution of a variable to a response, while the total Sobol index additionally accounts for the interaction effects [12,16]. The main and total indices for i-th parameter, X_i, to a response $\boldsymbol{F}(\boldsymbol{X})$ are defined respectively as $S_i = \mathrm{Var}_{X_i}\left[\mathrm{E}_{\boldsymbol{X}_{\sim i}}[\boldsymbol{F}(\boldsymbol{X})|X_i]\right]/Var(Y)$ and $S_i^{\mathrm{T}} = \mathrm{E}_{\boldsymbol{X}_{\sim i}}\left[\mathrm{Var}_{X_i}[\boldsymbol{F}(\boldsymbol{X})|\boldsymbol{X}_{\sim i}]\right]/Var(Y)$ where $\mathrm{E}_x[\cdot]$ and $\mathrm{Var}_x[\cdot]$ respectively represent mean and variance operator over a preset range of a variable or joint variables of x. The symbol $\boldsymbol{X}_{\sim i}$ represents a set of all input parameters in the model except X_i, i.e. $\boldsymbol{X} = \{X_i, \boldsymbol{X}_{\sim i}\}$. For efficient estimation of the Sobol indices, a quasi-Monte Carlo-based estimation method is selected through the quoFEM Dakota UQ engine developed by Sandia National laboratory [11,16].

Bayesian Calibration. Through Bayes' Theorem [8], the joint posterior distribution of all input parameters $\boldsymbol{X}$ as well as measurement noise levels $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ are identified from the dataset $\boldsymbol{D}$ as $p(\boldsymbol{X}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}|\boldsymbol{D}) \propto p(\boldsymbol{D}|\boldsymbol{X}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})p(\boldsymbol{X})p(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})$, where $p(\boldsymbol{X})$ and $p(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})$ are respectively prior distributions of parameter values and measurement variance levels, which are set to be uniform, and $p(\boldsymbol{D}|\boldsymbol{X}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})$ represents the likelihood of input parameters given the data $\boldsymbol{D}$ and measurement noise $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$. The likelihood function is often defined by assuming that observation noises follow independent and identically distributed Gaussian distributions without loss of generality, i.e., if $\boldsymbol{D}$ are measurements corresponding to system outputs $\boldsymbol{F}(\boldsymbol{X})$ with additive noise, $\boldsymbol{D} = \boldsymbol{F}(\boldsymbol{X}) + \boldsymbol{\varepsilon}$ and the noise term $\boldsymbol{\varepsilon}$ follows a Gaussian distribution, $\boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})$, with $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ being a diagonal matrix. The likelihood function of $\boldsymbol{X}$ is defined as $p(\boldsymbol{D}|\boldsymbol{X}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) = \mathrm{N}(\boldsymbol{F}(\boldsymbol{X}), \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})$. The samples of the posterior distribution are obtained by transitional Markov chain Monte Carlo (MCMC) sampling through quoFEM's UCSD (University of California, San Diego) UQ engine [17,18]. quoFEM also supports other probabilistic/deterministic calibration methods through the Dakota UQ engine.

Forward Propagation. Forward propagation allows researchers to see how the remaining uncertainty of the input parameters affects other estimations; in this case liquefaction-induced lateral spreading under an earthquake load. For this, the posterior samples obtained in the *Bayesian Calibration* process are directly imported back into quoFEM. Nonparametric naive resampling is conducted to generate samples from the posterior sample set, which are used as inputs to free-field analysis.

3 Problem description

The framework described in Section 2 was applied to a problem of uncertainty analysis during seismic soil liquefaction simulations using the data and the models described in this section.

Experimental Data. Cyclic Direct Simple Shear (CyDSS) tests were conducted on Ottawa F-65 sand by Morales *et al.* (2021) for different Cyclic Stress Ratio (CSR) conditions and the data was shared on the DesignSafe-CI Data Depot [15]. From this data, the number of cycles to the onset of liquefaction was computed using a threshold of 3% on the amplitude of shear strain. The results of these computations are shown in **Table 1**, and this data is used to calibrate the parameters of the PM4Sand material model.

Table 1. Cyclic strength data from experiments

Cyclic shear stress ratio (CSR)*	Number of cycles to initial liquefaction
0.105	26
0.105	21
0.130	13
0.151	5
0.172	4
0.200	3

* CSR=τ_{cyc}/σ'_{vo}, where τ_{cyc} is horizontal cyclic shear stress, σ'_{vo} is vertical consolidation stress

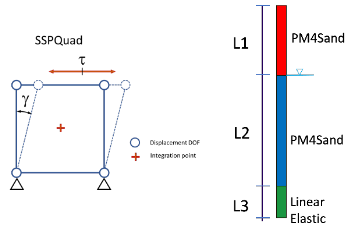


Fig. 1. [Left] – single element FE model used in sensitivity analysis and Bayesian calibration; [Right] – schematic of soil column used in free-field analysis

Material model. PM4Sand is a constitutive model capable of simulating liquefaction response of sandy soils. In OpenSees, this model has 24 parameters [4], but we considered the three primary input parameters - apparent relative density D_r, shear modulus coefficient G_o, and contraction rate parameter h_{po}, for sensitivity analysis and calibration while the other parameter values were set to the default recommendations in Boulanger and Ziotopoulou (2017) [2].

Finite Element (FE) models. During sensitivity analysis and calibration, a single element, shown on the left-hand side of **Fig. 1** was utilized to simulate the material response during a stress-controlled cyclic direct simple shear test. On the right-hand side of **Fig. 1**, a schematic representation of the synthetic layered soil profile used for 1D free-field analysis is shown. The soil column has a grade of 3%. The top layer, L1, represents a crust of thickness 2 m, L2 represents the liquefiable layer of 3 m thickness, and L3, the bottom 1 m of the soil, is linear elastic. The width of the soil column was 0.25 m, and the domain was discretized using undrained quadrilateral elements (SSPQuadUP) of size 0.25 m x 0.25 m. The response of this soil column was simulated by applying the Loma Prieta ground motion recorded at Gilroy Array #2.

4 Results and discussion

Global Sensitivity Analysis. Values of the input parameters were randomly sampled from the probability distributions in **Table 2** using the Latin Hypercube sampling method and the corresponding number of cycles to initial liquefaction was obtained from simulation with the single element FE model at five CSR values shown in **Table 1**. As results of variance-based global sensitivity analysis, quoFEM returns the main and total Sobol indices, sample values of the input parameters that were used to estimate the values of the Sobol indices, and corresponding model outputs.

Table 2. Probability distribution of the input parameters

Parameter	Distribution	Range
D_r	Uniform	0.1 – 0.9
h_{po}	Uniform	0.01 – 5
G_o	Uniform	200 – 2000

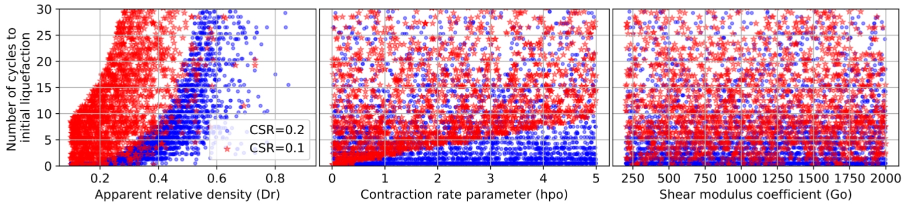


Fig. 2. Dependance of output on the primary input parameters of PM4Sand

Table 3. Main and total Sobol indices for the primary input parameters of PM4Sand to the number of cycles to initial liquefaction at different cyclic stress ratios

	CSR = 0.10		CSR = 0.13		CSR = 0.15		CSR = 0.17		CSR = 0.20	
	Main	Total	Main	Total	Main	Total	Main	Total	Main	Total
$\boldsymbol{D_r}$	0.849	0.915	0.826	0.896	0.827	0.895	0.817	0.900	0.805	0.885
$\boldsymbol{h_{po}}$	0.054	0.134	0.024	0.115	0.024	0.101	0.021	0.101	0.025	0.100
$\boldsymbol{G_o}$	0.000	0.002	0.000	0.002	0.001	0.002	0.003	0.003	0.004	0.006

Fig. 2 shows example scatter plots of the number of cycles to liquefaction, N_{cyc}, against each primary input parameter, X_i, given CSR of values 0.1 and 0.2. At any given value of one of the parameters X_i, there is a scatter in the output values which is due to the randomness in the sampled values of the other input parameters $\boldsymbol{X}_{\sim i}$. At low apparent relative densities D_r, the sand is in a loose state and N_{cyc} is low, as expected. It can be observed from the figure that N_{cyc} exhibits a clear dependance on the value of D_r, while there is minor dependance on the value of the contraction rate parameter h_{po} at low CSR values and no clear dependance on the value of the shear modulus coefficient G_o. These qualitative observations of the influence of the input parameters on the outputs are quantified by the Sobol indices enumerated in **Table 3**. The estimated Sobol

indices in the table correlate well with the observations from **Fig. 2**, with D_r having the highest main and total Sobol indices, followed by h_{po}, with G_o having near-zero Sobol index values at all CSR values studied. h_{po} has higher influence at lower CSR values than at higher CSR values. All the parameters, especially h_{po} have higher total Sobol indices compared to the main indices, which indicates that the inputs are interacting with each other when influencing N_{cyc} predicted by the model.

Bayesian calibration. During Bayesian calibration, the data of cyclic strength obtained from lab tests at 5 different CSR values, shown in **Table 1**, were used to estimate the values of the three primary input parameters of the PM4Sand model.

The prior probability density chosen for the three primary input parameters was the same as that used for the sensitivity analysis, shown in **Table 2**. The prior probability density must be set in a way that reflects all the available information/knowledge about the parameter values before any data are available. In this example, uninformative (uniform) priors were used to reflect lack of information about the parameter values. With this uniform prior, the posterior probability distribution was determined by the information contained in the data; the prior only defining the possible range of values that the parameter estimates can take.

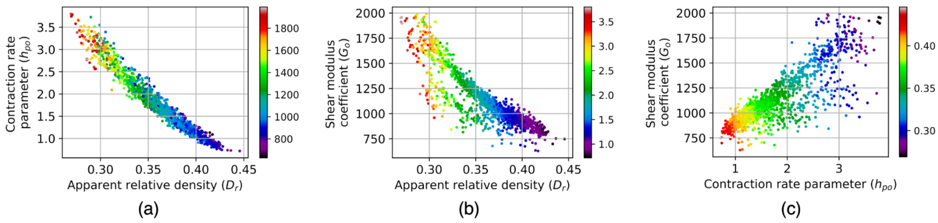
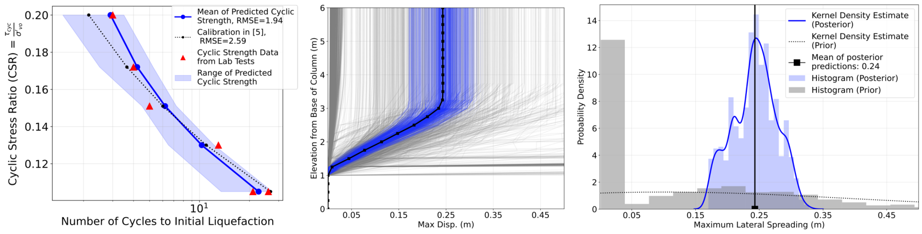


Fig. 3. Samples from posterior probability distribution, colored by (a) shear modulus coefficient, (b) contraction rate parameter, (c) apparent relative density

Fig. 3 shows 2000 samples drawn using quoFEM from the joint posterior probability distribution of the three primary input parameters, which characterizes the parameter estimation uncertainty. Since the data used consisted of only the number of cycles to reach a threshold shear strain, there are multiple settings of the primary input parameters that could predict the same or very similar cyclic strength curve but with different dynamic response history. Hence, a unique set of parameter values could not be identified from this data. Samples shown in **Fig. 3** capture the interdependence between the estimated parameter values that lead to predictions of similar cyclic strength curves.

quoFEM also returns predictions of the cyclic strength curve corresponding to the sample parameter values. The range of predicted cyclic strength shown in **Fig. 4** (a) is due to the parameter estimation uncertainty. The mean of the predicted cyclic strength is also shown in **Fig. 4** (a), along with the lab test data used for calibration, and a previous deterministic calibration to the same dataset by Ziotopoulou et. al. (2018) [5]. From the figure, it is evident that the cyclic strength data from lab tests is of limited amount and noisy, introducing uncertainty about the true response of the soil underlying the measured values. Additionally, the computational model might not represent the true response exactly, which introduces another layer of uncertainty. Bayesian calibration accounts for these sources of uncertainty in parameter estimation and can lead to more robust predictions. For example, as seen in **Fig. 4** (a), the range of predictions of cyclic strength from Bayesian estimation covers most of the data. The values of the root mean square error (RMSE) in the predictions of the number of cycles to initial liquefaction, shown in the legend in **Fig. 4** (a), indicate that the mean of the predicted cyclic strength obtained by Bayesian calibration matches the data more closely on average than the prediction achieved by deterministic calibration.



(a) (b) (c)

Fig. 4. Predictions from posterior sample values of the calibrated material model parameters

Uncertainty propagation. After Bayesian calibration, the uncertainty in the estimated values of the primary input parameters was represented by the 2000 posterior samples shown in **Fig. 3**. These values were used as inputs to the PM4Sand model representing the behavior of the top two layers of soil, shown in **Fig. 1**, during simulation of the response of the soil column to a single earthquake ground motion record (Loma Prieta Gilroy Array #2) corresponding to a PGA of 0.37g. The results of the uncertainty propagation analysis are shown in **Fig. 4** (b) and (c). **Fig. 4** (b) shows the updated uncertainty in the predicted maximum lateral displacement profiles using samples from the posterior probability density of the parameters (blue lines) compared with prediction by samples from the prior (gray lines). The mean of the predictions from the posterior is shown with the black line. **Fig. 4** (c) shows the distribution of the maximum lateral spreading at the top of the soil column (mean: 0.24 m, standard deviation: 0.03 m).

5 Conclusions

This paper presented a framework for characterizing the uncertainty in the site response due to seismic soil liquefaction using the open-source software application quoFEM. Further directions can be to quantify the probability of damage to structures and economic loss because of seismic soil liquefaction, and regional scale analysis of risk to infrastructure induced by seismic liquefaction.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. (1612843 & 2131111). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. Carey TJ, Kutter BL (2017) Comparison of liquefaction constitutive models for a hypothetical sand. *In Geotechnical Frontiers 2017*, pp. 389-398
2. Boulanger RW, Ziotopoulou K (2017) *PM4Sand (Version 3.1): A sand plasticity model for earthquake engineering applications.* Dept. of Civil and Env. Eng., Univ. of California, Davis, Davis, CA, Rep. UCD/CGM-17/01
3. McKenna F (2011) OpenSees: a framework for earthquake engineering simulation. *Computing in Science & Engineering*, vol 13(4), pp. 58-66.
4. Chen L, Arduino P (2021) *Implementation, verification, and validation of the PM4Sand model in OpenSees*, Pacific Earthquake Engineering Research (PEER) Center, University of California, Berkeley, Berkeley, USA, Rep. 2021/02
5. Ziotopoulou J, Montgomery J, Bastidas AMP, Morales B (2018) Cyclic Strength of Ottawa F-65 sand: laboratory testing and constitutive model calibration. *Geotechnical Earthquake Engineering and Soil Dynamics*, vol 293, pp. 180-189
6. Ziotopoulou K, Boulanger RW (2013) Calibration and implementation of a sand plasticity plane-strain model for earthquake engineering applications. *Soil Dynamics and Earthquake Engineering*, vol 53, pp. 268-280
7. National Academics of Sciences, Engineering, and Medicine (2016) *State of the Art and Practice in the Assessment of Earthquake-Induced Soil Liquefaction and Its Consequences.* The National Academies Press, Washington, DC, USA
8. Kruschke J (2014) *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan, 2nd ed.* Bloomington, IN, USA: Elsevier
9. Mercado V, Ochoa-Cornejo F, Astroza R, El-Sekelly W, Abdoun T, Pastén C, Hernández F (2019) Uncertainty quantification and propagation in the modeling of liquefiable sands. *Soil Dynamics and Earthquake Engineering*, vol 123, pp. 217-229
10. Chen L (2020) *Implementation, Verification, Validation, and Application of Two Constitutive Models for Earthquake Engineering Applications*. Diss University of Washington, 2020.
11. Adams BM et al. (2020) *Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.13 User's Manual* (No. SAND2020-12495). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).
12. Hu Z, Mahadevan S (2019) Probability models for data-driven global sensitivity analysis. *Reliability Engineering & System Safety*, vol 187, pp. 40-57
13. McKenna F, Zsarnoczay A, Gardner M, Elhaddad W, Yi S, Aakash BS (2021) NHERI-SimCenter/quoFEM: Version 2.4.0 (v2.4.0). Zenodo. https://doi.org/10.5281/zenodo.5558000
14. Taylor RL (2014) *FEAP-A finite element analysis program*.
15. Morales B, Humire F, Ziotopoulou K (accessed June 28th, 2021) *Data from: Cyclic Direct Simple Shear Test-ing of Ottawa F50 and F65 Sands (Feb. 1st, 2021).* Distributed by DesignSafe-CI Data Depot. doi: 10.17603/ds2-eahz-9466
16. Weirs VG., Kamm JR, Swiler LP, Tarantola S, Ratto M, Adams BM, Rider WJ, Eldred MS, (2012) Sensitivity analysis techniques applied to a system of hyperbolic conservation laws. *Reliability Engineering & System Safety*, vol 107, pp.157-170
17. Ching J, Chen YC (2007) Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging. *Journal of engineering mechanics*, vol 133(7), pp. 816-832
18. Minson SE, Simons M, Beck JL (2013) Bayesian inversion for finite fault earthquake source models I—Theory and algorithm. *Geophysical Journal International,* vol 194.3, pp. 1701-1726