Bayesian model selection for sand with generalization ability evaluation

Yin-Fu JIN[1], Zhen-Yu YIN[1,*], Wan-Huan ZHOU[2] and Jian-Fu SHAO[3]

Affiliation:

1 Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China

2 State Key Laboratory of Internet of Things for Smart City and Department of Civil and Environmental Engineering, University of Macau, Macau S.A.R., China

3 University of Lille, CNRS, Centrale Lille, FRE 2016 – LaMcube – Laboratoire de mécanique multiphysique multiéchelle, F-59000, Lille, France

* Corresponding author: Dr Zhen-Yu YIN, Tel: +85 3400-8470;

E-mail: zhenyu.yin@polyu.edu.hk; zhenyu.yin@gmail.com

Abstract: Current studies have focused on selecting constitutive models using optimization methods, or selecting simple formulas or models using Bayesian methods. In contrast, this paper deals with the challenge to propose an effective Bayesian-based selection method for advanced soil models accounting for the soil uncertainty. Four representative critical state based advanced sand models are chosen as database of constitutive model. Triaxial tests on Hostun sand are selected as training and testing data. The Bayesian method is enhanced based on transitional Markov chain Monte Carlo method, whereby the generalization ability for each model is simultaneously evaluated, for the model selection. The most plausible/suitable model in terms of predictive ability, generalization ability, and model complexity is selected using training data. The performance of the method is then validated by testing data. Finally, a series of drained triaxial tests on Karlsruhe sand is used for further evaluating the performance.

Key words: Bayesian theory; constitutive relation; sand; transitional Markov Chain Monte Carlo; generalization ability; critical state

1 Introduction

Reliable predictions depend heavily on plausible constitutive models with reasonable parameters in geotechnical engineering [1-7]. To achieve this purpose, many constitutive models have been proposed for soils [8-11]. However, different constitutive models render dissimilar results in the process of numerical simulations, prompting variations in engineering decisions that affect the levels of safety, economy, and risk in construction [12-17]. Therefore, selecting the appropriate model is a critical issue in the application of constitutive models to practical use. Lack of attention to the problem of model selection has become a primary source of risk for accident [18].

Model selection is the task of choosing a model having the correct inductive bias from a set of candidate models. In practice, selection of a constitutive model often depends on the user's preferences and experiences, both of which are subjective. Accordingly, an efficient approach plays an important role in conducting model selection. Among current studies, model selection primarily relies on optimization methods and the Bayesian approach. However, model selection based on optimization methods is less attractive due to the lack of considering soil uncertainty [8, 19]. The Bayesian approach appears to be primarily responsible for this problem, in which the ability to select the most plausible/suitable model while simultaneously obtaining the posterior uncertainty of parameters would be of considerable use to engineers [20-34]. Unfortunately, Bayesian methods are only applied to selecting some simple formula or simple soil models in geotechnical engineering [20, 22, 25, 26, 35]. Accordingly, an investigation of Bayesian model selection for advanced soil models is desirable.

Furthermore, a reliable soil model should offer a reasonable trade-off between predictive capability and robustness. The predictive capability is usually evaluated using the magnitude of difference between predictions and experiments for optimization-based model selections [7, 8, 11, 19, 36-41] or by the maximum likelihood value for Bayesian-based model selections [21, 22, 25, 26, 28, 42, 43]. Robustness can be measured by generalization ability, which is a measure of how accurately a model is able to predict outcome values for previously unseen data. For a given soil, simple models with insufficient features are likely to miss some of the characteristics of soil behaviour. In contrast, advanced models with a large number of identified parameters will be better able to capture different soil behaviours, but they are therefore likely to lead to over-fitting the data, thereby reducing generalization ability. Accordingly, it is desirable to select a soil model that offers an outstanding predictive ability and generalization ability so that both efficient and robust application performance can be expected in practice [29].

This paper thus conducts soil model selection using the Bayesian method while evaluating the model's generalization ability. For this purpose, a set of model classes is selected that includes four representative sand models: (1) the critical-state-based sand model (SIMSAND), (2) the simple anisotropic sand model (SANISAND), (3) the critical-state-based hypoplastic sand model (HYPOSAND), and (4) the Cam-Clay-based sand model (MCCSAND). Three drained triaxial tests on Hostun sand are selected to make five combinations having different number of tests, making five groups of training data. For each combination, Bayesian model selection is performed using the enhanced transitional Markov chain Monte Carlo (TMCMC) method, in which the generalization ability for each selected sand model is simultaneously evaluated by testing data on the same sand. The most appropriate model in terms of predictive ability, generalization ability and model complexity is selected and validated. Finally, a series of drained triaxial tests on Karlsruhe sand is used for further evaluating the performance.

2 Representative advanced sand models

An impressive variety of sand models that can be classified as (1) elastic perfectly plastic models (such as the Drucker-Prager model, the Mohr-Coulomb model), (2) nonlinear simple models (such as the nonlinear Mohr-Coulomb (NLMC) [8], hardening soil model [44, 45]), (3) critical-state-based advanced models (such as Nor-Sand model [46]; CSAM model [47]; Severn–Trent sand model [48]; UH models [49-52]; SANISAND model [53]; the critical-state-based simple sand model (SIMSAND) [8, 11, 19, 54]; and (4) hypoplasticity models [55-58] have been developed. Not surprisingly, such models display different performances when modelling sand behaviours. The MC and NLMC models are fundamentally limited from physics perspective, which is well known for most geotechnical engineers. Thus, both two models are not considered in model selection. To compare the performance of comparable models (same level of numerical sophistication) and an engineer cannot easily distinguish which model to be adopted in practice, a set of model classes that included four representative sand models was chosen for their popularity to perform model selection: (1) the SIMSAND model, the critical state and interlocking effect are incorporated so that the stress dilatancy and contraction can be described, (2) the SANISAND model, which incorporates the concept of bounding surface compared to SIMSAND, (3) the critical-state-based hypoplastic sand model (HYPOSAND) by Wang et al. [59], which belongs to the framework of hypoplasticity and (4) the MCCSAND, which belongs to the framework of modified Cam-Clay (Yao et al. [50, 51]). The basic constitutive equations of all selected sand models are summarized in Appendix I.

Although the equations are basic, some specific points should be addressed. For HYPOSAND, Young's modulus is constant. For SIMSAND and SANISAND, Young's modulus is expressed as follows, according to Richart et al. [60],

$$E = E_0 \cdot p_{at} \frac{(2.97-e)^2}{(1+e)} \left(\frac{p'}{p_{at}} \right)^{\zeta} \tag{1}$$

For the MCCSAND, Young's modulus is expressed as:

$$E = \frac{3(1-2\upsilon)(1+e_0)}{\kappa} (p' + p_s) \text{ with } p_s = \exp\left(\frac{N-Z}{\lambda} \right) - 1 \tag{2}$$

where E_0 is the reference value of Young's modulus; e is the void ratio and e_0 is the initial void ratio; p' is the mean effective stress; p_{at} is the atmospheric pressure used as reference pressure (p_{at} = 101.3 kPa); ζ is a constant; υ is Poisson's ratio; κ is the swelling index; λ is the compression index; N and Z are two constants of the MCCSAND.

For SIMSAND, SANISAND and HYPOSAND, the nonlinear formulation of critical state line (CSL) [61] was adopted.

$$e_c = e_{ref} \exp\left[-\lambda \left(\frac{p'}{p_{at}} \right)^{\xi} \right] \tag{3}$$

where e_c is the critical void ratio; e_{ref} is the initial critical void ratio at $p' = 0$; λ and ξ are two parameters controlling the shape of CSL in the e-logp' plane.

The parameters of each selected model can be divided into: (1) elastic parameters, (2) plastic shear-hardening related parameters, (3) stress-dilatancy-related parameters and (4) critical-state related parameters for critical-state-based models.

3 Model selection approach and generalization ability evaluation

3.1 Bayesian model class selection

In this section, the Bayesian approach for parametric identification and model class selection is briefly outlined. Further details of basic Bayesian model class selection can be found in Appendix II and Yuen [28].

Following a Bayesian formulation [29, 62, 63] and assuming that the observation data and the model predictions satisfy the prediction error equation:

$$U_{\text{obs}} = U_{\text{num}}(\mathbf{b}) + \varepsilon \tag{4}$$

where $\mathbf{b}$ is the vector containing model parameters, such as friction angle and critical state related parameters; ε is a zero-mean Gaussian random variable with variance σ_ε^2 representing the prediction error variance and σ_ε^2 is another unknown parameter besides the soil model parameters $\mathbf{b}$. Thus, the uncertain parameter vector $\boldsymbol{\theta}$ includes the model parameters $\mathbf{b}$ and the prediction-error variance σ_ε^2, i.e., $\boldsymbol{\theta}=\left[\mathbf{b},\ \sigma_\varepsilon^2\right]$. Table 1 shows the uncertain parameters and the number of parameters for each sand model class.

Note that the two elastic parameters, E_0 and ζ in model classes SIMSAND and SANISAND, can easily be obtained from isotropic compression tests. Accordingly, the method for determining the values of E_0 and ζ, as presented in [8, 19, 37], is adopted in this study. The cohesion c for HYPOSAND is set to zero because the tests are performed on dry sand in this study. A typical value of Poisson's ratio υ=0.25 is assumed for all model classes.

Uncertainties of parameters can be evaluated using the posterior PDFs, with the expression of the posterior PDF for data D written as follows:

$$p(\boldsymbol{\theta}|D) = \frac{p(\boldsymbol{\theta})\,p(D|\boldsymbol{\theta})}{p(D)} \tag{5}$$

where $\boldsymbol{\theta}=[\mathbf{b},\sigma_\varepsilon]$ is the uncertain parameters; $p(D)$ is the evidence; $p(\boldsymbol{\theta})$ is the prior PDF of the uncertain parameters $\boldsymbol{\theta}$, which is based on the previous knowledge or user's judgment; and $p(D|\boldsymbol{\theta})$ is the likelihood function expressing the level of data fitting.

Generally, deformation and stress are two extremely important indicators for soil behaviours. The measurement produced by a laboratory test usually contains two curves, such as the curves ε_a–q and ε_a–e for the drained triaxial test or the curves ε_a–q and ε_a–u for the undrained triaxial test (where ε_a is axial strain, q is deviatoric stress, e is void ratio, and u is excess pore water pressure). Accordingly, a goodness-of-fit function involving these two important indicators is reasonable. Note that the measured q has no correlation with the measured e or u for a given sand. Actually, the measured q, e and u are mainly influenced by the confining pressure, the relative density of sand and the initial void ratio. According to [8, 19], a normalized goodness-of-fit function is adopted due to the error independent of the magnitude of different variables (e.g., q and e or u), which is expressed as:

$$J_g(\mathbf{b};D) = \frac{1}{N_0 N}\sum_{j=1}^{N_0}\left[\sum_{i=1}^{N}\left(\frac{U^i_{\text{obs}} - U^i_{\text{num}}}{U^i_{\text{obs}}}\right)^2\right]_j \tag{6}$$

where N is the number of measured values, N_0 is the number of curves for one test, U^i_{obs} is the value of measurement point i, and U^i_{num} is the value of calculation at point i.

With multiple observations and types of observations, likelihood values for each observation must be combined into an overall value for each candidate parameter set [64]. For laboratory tests of sand, the multiple tests can be a series of triaxial tests with different relative densities (from loose to dense) under different confining pressures (from low to high) and different drainage conditions (drained and undrained). All the tests are assumed independent each other. When the measured data D involve M tests during Bayesian parameter identification, the likelihood function is expressed as:

$$\ln p\left(D|\boldsymbol{\theta}\right)=\sum_{i=1}^{M}\mathrm{w}_i \ln p\left(D_i|\boldsymbol{\theta}\right) \tag{7}$$

where M is the number of involved tests, w_i is weight of $p\left(D_i|\boldsymbol{\theta}\right)$, and $p\left(D_i|\boldsymbol{\theta}\right)$ is the likelihood corresponding to the test i. In this study, the weight of each likelihood for all involved tests is considered the same and thus equal to 1.

The posterior PDF $p\left(\boldsymbol{\theta}|C,D\right)$ represents the updated belief about the parameter vector $\boldsymbol{\theta}$ after obtaining the evidence D. An accurate estimator of the parameters $\boldsymbol{\theta}$ for the adopted soil model is the Maximum a Posteriori (MAP) estimation. The MAP parameter vector $\boldsymbol{\theta}^*$ can be obtained by maximizing the posterior $p\left(\boldsymbol{\theta}|C,D\right)$, or equivalently, $p\left(\boldsymbol{\theta}|C\right)p\left(D|\boldsymbol{\theta},C\right)$. Considering that the model classes involve high-dimensional nonlinear functions, the evidence integral must be evaluated numerically. As the TMCMC method has been proven more efficient for high-dimensional problems and can also evaluate the evidence for each model class [65-68], it was used to quantify the uncertainty of model parameters and conduct the model class selection.

3.2 Enhancement of TMCMC method

The TMCMC method was originally developed by Ching and Chen [65] as a combination of the sequential particle filter method [69] and MCMC. The method begins with the prior distribution $p\left(\boldsymbol{\theta}\right)$ and makes a gradual transition to the posterior by optimization at each round of samplings. The key idea of TMCMC is that of proposal density, which corresponds to the jth round of sampling $p\left(\boldsymbol{\theta}\right)_j$ determined as,

$$p\left(\boldsymbol{\theta}\right)_j \propto p\left(\boldsymbol{\theta}\right)\cdot L\left(\boldsymbol{\theta}|D\right)^{q_j} \tag{8}$$

where $q_j \in [0, 1]$ is chosen following $q_0=0< q_1<...< q_m=1$ with j=0, 1, ….., m denoting the stage level. Consequently, $p(\boldsymbol{\theta})_0$ equals the prior distribution $p(\boldsymbol{\theta})$ for j=0, and $p(\boldsymbol{\theta})_m$ is the posterior distribution $p(\boldsymbol{\theta}|D)$ for j=m.

The details of the original TMCMC method, with its MATLAB code, can be found in Ching and Wang [67]. In the original TMCMC, the new samples are generated from a normal distribution with the mean and standard error calculated from the samples of last iteration. However, some observations have indicated that the inappropriate mean value and standard deviation error can result in the estimated posteriors tending to fall into local convergence [68]. Therefore, ~~T~~to improve the performance of original TMCMC, a differential evolution–Markov chain algorithm proposed by Vrugt [70] was adopted in this study to replace the process to generate a new sample from the normal distribution in original TMCMC, which can be generated as:

$$\boldsymbol{\theta}_{(j,l)}^{\text{new}} = \boldsymbol{\theta}_{(j,l)}^{c} + d\boldsymbol{\theta}_{(j,l)} \tag{9}$$

with

$$d\boldsymbol{\theta}_{(j,l)} = (1+\lambda)\cdot\gamma\cdot\left[\left(\boldsymbol{\theta}_j^{best} - \boldsymbol{\theta}_{(j,l)}^{c}\right) + \left(\boldsymbol{\theta}_{(j,a)} - \boldsymbol{\theta}_{(j,b)}\right)\right] + \zeta \tag{10}$$

where $\boldsymbol{\theta}_{(j,l)}^{\text{new}}$ is the new sample; $\boldsymbol{\theta}_{(j,l)}^{c}$ is the current sample; $\boldsymbol{\theta}_j^{best}$ is the sample corresponding to the maximum weight in the current iteration; d is the dimension of $\boldsymbol{\theta}$; $\boldsymbol{\theta}_{(j,a)}$ and $\boldsymbol{\theta}_{(j,b)}$ are two vectors consisting of d variables, where the indices a and b are two integers drawn from [1,…, N_s]; $\gamma = 2.38/\sqrt{2\delta d}$ is the jump rate; δ denotes the number of chain pairs used to generate the jump with a default value of δ=3 according to Vrugt [70]. The values of λ and ζ are sampled independently from the uniform distribution [–c, c] and the normal distribution $N(0, c^*)$, respectively. In this study, the c=0.1 and $c^*=10^{-12}$ were employed, as recommended by Vrugt [70].

After differential evolution, a binomial crossover operation forms the final sample,

$$\boldsymbol{\theta}_{(j,l)}^{\text{new}} = \begin{cases} \boldsymbol{\theta}_{(j,l)}^{\text{new}}, & \text{if } \text{rand}(0,1) \le CR \text{ or } l = l_{rand} \\ \boldsymbol{\theta}_{(j,l)}^{c}, & \text{otherwise} \end{cases} \tag{11}$$

where rand(0, 1) is a uniform random number within [0, 1]; l_{rand}=randint (1, d) is an integer randomly chosen from 1 to d and is newly generated for each l; the crossover probability $CR \in [0, 1]$ corresponds roughly to the average fraction of the vector components that are inherited from the mutation vector, with CR=0.9 taken in this study.

After the enhanced TMCMC algorithm is executed, the importance weights produced during the algorithm can be used to estimate the model evidence $p\left(D \middle| C_j\right)$, which can be estimated by the enhanced TMCMC algorithm as a by-product:

$$p\left(D \middle| C_j\right) \approx S = \prod_{j=0}^{m-1}\left(\frac{1}{N_s}\sum_{k=1}^{N_s} w_{j,k}\right) \tag{12}$$

where S is asymptotically unbiased estimation of the model evidence, $w_{j,k}$ are the importance weights, and m is the total number of transitional stages.

The local convergence problem of drawing posteriors can be solved by the enhanced DE-TMCMC. Thus, the model selection process conducted by enhanced DE-TMCMC is more robust than the same works done by original TMCMC. Finally, the most plausible/suitable model can be selected for a given problem.

3.3 Evaluation of generalization ability

Since sand models are always evaluated on the basis of finite samples/tests, the evaluation of a sand model is sensitive to sampling error. As a result, measurements of prediction error for the current data may not provide much information about predictive ability for new data. To track this problem, generalization ability, a measure of how accurately a model is able to predict outcome values for previously unseen data, was adopted for this case.

Due to the difficulties in computing the unknown joint probability distribution for generalization error, generalization ability is usually measured by empirical error, a function of the difference between the actual and predicted results for out-of-sample data. To make the empirical error independent of the type of test and the number of measurement points, a normalized empirical error function was adopted, expressed as:

$$\text{Error}(\boldsymbol{\theta}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{U_{\text{obs}}^{i} - U_{\text{num}}^{i}}{U_{\text{obs}}^{i}}\right)^{2}} \tag{13}$$

where U_{obs}^{i} is the ith observed data point; U_{num}^{i} is the ith numerical data point; N is the number of data points; $\boldsymbol{\theta}$ is a set of model parameters; $\text{Error}(\boldsymbol{\theta})$ is the error between observed and numerical data for the set of model parameters $\boldsymbol{\theta}$.

With multiple observations, the combined empirical error is expressed as:

$$\text{Error}(\boldsymbol{\theta}) = \sum_{j=1}^{X} l_i \left[\text{Error}(\boldsymbol{\theta})\right]_i \tag{14}$$

where X is the number of observations contributing to the computation of generalization ability, and l_i is weight. In this study, $l_i=1/X$ for a uniform weight.

3.4 General procedure of model selection

Fig. 1 shows the procedure for model class selection and hence evaluating generalization ability. In this procedure, a set of candidate models was first selected. Then, the measured data serving as the training data for model selection were selected. Next, the likelihood function for multiple observations was determined according to Eq.(7). Subsequently, the Bayesian model class selection using the enhanced TMCMC method was successfully carried out, incidentally giving the posteriors of all identified parameters for each selected model class. Based on preliminary results, the most plausible model with posteriors of its parameters was gained.

Continuing the procedure, after obtaining the posteriors of all identified parameters for each selected model class, simulations of the new tests that were different from the training data used in the model class selection, were performed using each model class with N optimum sets of parameters drawn from the obtained posterior distributions. The empirical error was simultaneously calculated. The model class showing minimum empirical error was identified as the suitable model having strong generalization ability.

The ideal model class should survive during Bayesian model class selection as well as the evaluation of generalization ability.

3.5 Illustration case of Bayesian model selection

To show the proposed procedure of model selection, an illustration case on selecting a polynomial equation is presented. The measured data was generated by the equation $y=2x+5x^2$ with $x\in[0, 3]$. The general model class can be defined as follows:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \varepsilon \tag{15}$$

where a_0~a_4 are uncertain model parameters; x is input and y is output.

Then, a total of 15 ($\mathrm{C}_4^1 + \mathrm{C}_4^2 + \mathrm{C}_4^3 + \mathrm{C}_4^4$) models were generated. Based on the measured data, an appropriate model will be finally selected from 15 models. The prior PDF of all uncertain parameters for each model class was assumed in uniform distribution. All uncertain parameters were independently and uniformly distributed within [0, 10]. The results of model selection are summarized in Table 2. According to the values of $\ln p\left(D\middle|C_j\right)$, the model 8 is most appropriate and the identified parameters are a_1=2.000009, a_2=4.9999961 with the uncertainty of ε=3.63e-06. The obtained model approximately equals to the predefined equation

4 Selection of sand models

4.1 Model selection based on different test sets

A series of laboratory triaxial tests performed on Hostun sand by Liu et al. [71] and Li et al. [72] were chosen for the model selection. The initial void ratio and confining pressure for each test are listed in Table 3. Fig. 2 shows all results of selected triaxial tests. To form a comprehensive group of experiments that can effectively reflect common sand behaviours, a total of three drained triaxial tests including one dilative test (test 1) and two contractive tests (tests 4 and 5) were chosen as training data for model class selection and parameter identification. The exact information of used training data is data points on deviatoric stress-axial strain and void ratio or volumetric strain-axial strain curves (i.e. q-ε_a and e-ε_a). The other tests were considered as testing data used to calculate the empirical error for evaluating the generalization ability of each sand model. Note that the experimental data have strain localization, such as the tests on very dense sand, are not considered since the numerical results are obtained by an element integrated with one single Gauss point in this study. The model selection when the used data have strain localization can be conducted using finite element method for boundary value problems.

To investigate the effect of a number of tests on model class selection and the evaluation of generalization ability, the model class selection was successively conducted based on different number of tests. Based on three selected drained triaxial tests (test 1, 4 and 5), a total of seven combinations (selecting one, two or three tests from tests of 1, 4 and 5, respectively) were obtained as summarized in Table 4. However, combinations 2 and 3 were overlapped because of containing a similar contractive test. Thus, only combination 2 was selected. A similar situation was found for combinations 4 and 5. Thus, only combination 4 was selected. As such, combinations 1, 2, 4, 6 and 7—five cases in total—were finally adopted in the model class selection.

The prior PDF of all uncertain parameters for each model class was assumed to show uniform distribution. All uncertain parameters were independently uniformly distributed [*lower_bound*, *upper_bound*], e.g., the friction angle was uniformly distributed [20, 50]. The bounds of each prior PDF for all model classes are summarized in Table 5. Note that when the bounds for the uncertain parameters are very narrow, a very small step size is needed in the model selection.

For all cases, the enhanced TMCMC with 2000 was implemented, meaning 2000 samples per stage. Therefore, $2000N_s$ total samples were generated in the enhanced TMCMC simulation. According to previous results [26, 65-68], the total number of stages for different model class selections was not always the same.

4.2 Results and discussion

Table 6 summarizes the results of evidence obtained by enhanced TMCMC and the results of plausibility using Bayes' theorem based on various test combinations. To intuitively present the results of model class selection while choosing the most plausible model class, all plausibilities are exhibited in Fig. 3. The MAP parameters that correspond to the maximum values of their posterior PDFs and the uncertainties for all selected model classes are summarized in Tables 6~11. It can be seen that some parameters with physical meanings (e.g., friction angle, critical state parameters) identified from three tests for SIMSAND, SANISAND, HYPOSAND and MCCSAND agree well with experimental measurements by Liu et al. [71] and Li et al. [72].

For model class selection based on test 1 (dense sand) (see Fig. 3(a)), SANISAND was the most plausible model class while the HYPOSAND was the worst one. The results indicate that HYPOSAND had a limitation in capturing the dilative behaviour of dense sand. For selection based on test 4 (loose sand) (see Fig. 3(b)), the most plausible model class was still SANISAND, while the worst was the HYPOSAND. However, little difference in plausibilities for other model classes was found, demonstrating that such model classes show outstanding ability in the simulation of contractive behaviours for sands.

When the number of involved tests in the model selection increased to 2, a different result for model class selection was found. The most plausible model class was SIMSAND for selection based on tests 1 and 4 (see Fig. 3(c)), while it was SANISAND for selection based on tests 4 and 5 (see Fig. 3(d)). However, the difference of plausibility between SIMSAND and SANISAND was small for both cases, and thus it can sometimes be ignored. The worst model was HYPOSAND for case involving tests 1 and 4 and MCCSAND for case involving tests 4 and 5.

Based on preliminary results, it is possible to infer that the most plausible model class is either SIMSAND or SANISAND, which can perform well in capturing the sand behaviours of dilatancy and contractiveness. Furthermore, HYPOSAND and MCCSAND displayed similar performance but were inferior to SIMSAND and SANISAND in simulation.

Unsurprisingly, this surmise is confirmed by the results of model selection based on three tests, as depicted in Fig. 3(e). Although the performance of HYPOSAND and MCCSAND was inferior to that of SIMSAND and SANISAND, it is also acceptable in practice. However, more effective parameter identification methods for them should be considered to find more accurate parameters, such as optimization-based methods [7, 8, 19, 38-40].

As mentioned, a reliable soil model should have a reasonable trade-off between predictive capability and generalization ability. However, for a given soil model, the predictive performance for unseen test data indicating generalization ability relies heavily on the precision of the parameters used. Accordingly, the evaluation of generalization ability can serve as a remedy for Bayesian model class selection. To evaluate the generalization ability of each model using its optimal parameters, the testing tests were simulated and the empirical errors in terms of q and e for drained tests or u for undrained tests were simultaneously calculated according to Eq.(14), as shown in Fig. 4. The magnitude of empirical error reflects predictive ability along with the accuracy and reliability of optimal parameters obtained by Bayesian model selection from limited training tests for a given sand model. The large value of empirical error indicates a poor performance in terms of prediction.

Unlike the results of Bayesian model class selection, some different findings from the standpoint of generalization ability were noted. When only one training test on a dense sample was involved in the Bayesian model class selection, the best model was SANISAND and the worst was HYPOSAND in terms of the generalization ability (see Fig. 4(a)), consistent with the result obtained by Bayesian model class selection. When the involved test in model class selection was performed on a loose sample, the best mode was still SANISADN while the worst was MCCSAND (see Fig. 4(b)). It can be shown that the parameters identified from tests on dense sand are more reliable that those identified from tests on loose sand for the MCCSAND, which indicate that the parameters of the MCCSAND model are sensitive to the types of tests involved. Therefore, the MCCSAND model is not recommended to be applied to geotechnical practice when only one test on either dense or loose sample is available. Furthermore, HYPOSAND exhibit similar but poor performance for both cases using one test, and thus more attentions should be paid when using them in practice.

For cases with two tests, the most suitable models were SIMSAND and SANISNAD (see Fig. 4(c) and (d)). Due to a lack of the ability to balance the predictions between dilatancy and contractive behaviours, the inaccurate optimum parameters can lead to a weak generalization ability of HYPOSAND and MCCSAND, and thus neither is recommended when only one test on a dense sample or only one test on a loose sample is available for parameter identification in practice. For the case with three tests, the most appropriate models are SIMSAND and SANISAND, while the worst is still MCCSAND (see Fig. 4(e)).

In summary, the poor generalization ability of HYPOSAND and MCCSAND when the involved training tests are performed on both dense and loose sand samples is mainly attributable to poor predictive ability of the models ~~and inaccurate optimum parameters~~. The performance of HYPOSAND is not quite satisfactory compared to SIMSAND and SANISAND regardless of the number and types of involved tests. The poor generalization ability of the MCCSAND is found when the involved tests on a loose sample are dominant, which can be attributed to the ~~inaccurate identified parameters because its~~ predictive ability of the model ~~is acceptable based on the results of Bayesian model selection~~. The results also suggest that the tests performed on dense sand are beneficial to identify reliable parameters of the MCCSAND.

As stated by Wood [73], simple yet adequate models are favoured on the basis of practicality. In general, how simple of a model can be expressed using the number of parameters and the complexity of model equations. Compared to Bayesian model class selection, the Akaike information criterion (AIC) [74] and Bayesian information criterion (BIC) [75] are two commonly used criteria for model selection accounting for the effect of the number of parameters. For AIC, the penalty term is taken to be the number of adjustable parameters of model class. When the number of data points N is large, the penalty term will disappear which means that the contribution of the number of model parameters is little. Accordingly, the AIC cannot replace the Bayesian model class selection method if N is sufficiently large. For BIC, the penalty term increases with the number of data points N. According to Yuen [28], for large N, the BIC is equivalent to the Bayesian approach using equal prior plausibilities for all model candidates. However, in practice, the Bayesian model class selection is especially useful when N is not large so the selection of model class is difficult by the user's judgement. Therefore, the BIC cannot replace the Bayesian model class selection method if N is not sufficiently large since the residual term has an important contribution. Overall, the Bayesian model class selection is superior to AIC and BIC.

Numerical convergence is easier to obtain when simple formulas are used to deal with complex geotechnical problems. Due to the implementation of the bounding surface concept-based hardening law with a small elastic domain for SANISAND, increasing the model complexity, SIMSAND was finally selected as the most appropriate sand model in terms of the predictive ability and generalization ability for simulating monotonic behaviours, consistent with results obtained by the optimization method [8]. Therefore, SIMSAND is selected as an outstanding sand model regardless of the methods (deterministic and probabilistic) used.

To show the performance of the selected model class SIMSAND and its MAP parameters, a series of simulations on testing data that have been used in the evaluation of generalization ability were performed, as shown in Fig. 5. Satisfactory agreement is found between the numerical simulations and experiments that contractive and dilatancy behaviours, even static liquefaction, can be easily reproduced, demonstrating an excellent predictive ability for SIMSAND as well the rationality of identified parameters.

5 Validation by other sand

5.1 Model selection based on tests of Karlsruhe sand

A series of drained triaxial tests performed on Karlsruhe sand by Wichtmann and Triantafyllidis [76] was selected for this case. According to Wichtmann and Triantafyllidis [76], the test Karlsruhe sand has a mean grain size d_{50}=0.14 mm and a uniformity coefficient C_u=d_{60}/d_{10}=1.5. The minimum and maximum void ratios were e_{min}=0.677 and e_{max}=1.054. Most of the grains have a subangular shape.

The initial void and confining pressure for each test are summarized in Table 11. The experimental results for all triaxial tests are shown in Fig. 6. Three drained triaxial tests (tests 3, 4 and 9) with approximate density and different confining pressures following the industrial standard were selected as training data. The remaining tests were used to evaluate generalization ability. K_0=44 and n=0.68 were determined from the 1D compression tests following those studies [8, 19, 37].

5.2 Results and discussion

Fig. 7 shows the model selection results based on three drained tests of Karlsruhe sand. It was found that the most plausible model class was SIMSAND and the worst was still HYPOSAND, which is consistent with preliminary results obtained on tests of Hostun sand. SANISAND also had an acceptable performance and was inferior only to SIMSAND. Table 12 summarizes the MAP parameters of all sand models for Karlsruhe sand. It can be seen that the parameters identified for SIMSNAD, SANISAND, HYPOSAND and MCCSAND are in appropriate ranges compared to the experimental investigations (ϕ=33.1° and e_{ref}=1.067) by Wichtmann and Triantafyllidis [76].

Using the obtained optimal parameters, the evaluation of generalization ability for all model classes was conducted as shown in Fig. 8. SIMSAND and SANISAND were found to be the most suitable models, while the HYPOSAND model was worst in terms of generalization ability evaluated using finite testing data. Furthermore, the MCCSAND model also showed good generalization ability due to the parameters identified from tests on dense sand samples, confirming the validity of previous suggestions.

Overall results demonstrate that the SIMSAND is the most plausible/suitable sand model in terms of predictive ability and generalization ability for Karlsruhe sand. Fig. 9 shows the comparison of testing tests on Karlsruhe sand between experiments and simulations. Acceptable agreement is found between the numerical simulations and experiments that the contractive and dilatancy behaviours can be adequately reproduced, demonstrating an excellent predictive ability for SIMSAND as well the rationality of identified parameters.

6 Discussions

In the above investigation, four advanced constitutive models were adopted to describe the typical mechanical behaviours of sand under monotonic loading. However, the behaviours of granular materials are complex and an accurate modelling remains an open challenge. One of the main difficulties lies in a clear and efficient definition of the internal state of such materials. Usually two kinds of internal parameters have to be considered: (1) one or several scalar parameters characterizing the density state with respect to a critical density, such as the evolution of e/e_c in considered models of this study, which is a usual way to define the isotropic internal state; (2) one or several tensorial parameters characterizing the anisotropic internal state, such as the critical state considering fabric anisotropy [77-81]. In complex loading paths, experimental studies have indicated that the behaviour of a granular soil under shear is predominantly anisotropic. Such anisotropic behaviour of sand can be effectively modelled by incorporating the fabric tensor and its evolution.

Another effective way to simulate such an anisotropic behaviour is the use of micromechanics-based models [82-85]. Note that the four adopted sand models in this study only consider the internal isotropic state variables and the anisotropic behaviour of sand can't be described, which is a limitation of this study. Nevertheless, the approach can also be applied to models considering anisotropy with objective tests of complex loading paths.

The considered tests (training and testing tests) are only monotonous triaxial tests (drained and undrained) in this study. The proposed analysis can give an estimation of the ability of the considered models to give good simulation performance of monotonic loading defined from an initial isotropic internal state. However, the loading paths in practice are more complex, such as: (1) a loading with the stress principal directions different from the ones of an anisotropic initial internal state; (2) a cyclic loading; (3) a loading with significant evolutions of the principal directions. Again, the approach can also be applied to these cases if adopted constitutive models further considering anisotropy with effect of stress reversal.

Furthermore, the test results sometimes are not consistent themselves, which is probably caused by: (a) an inherent spatial variability of soil properties, (b) experimental uncertainty (measurement scatter) due to limitations of the experimental techniques and (c) sampling uncertainty (statistical uncertainty) due to the limited number of soil samples used in the investigation. Such inconsistency would lead to variability of parameters for model of interest, which can be quantified by the proposed Bayesian approach with DE-TMCMC from available experimental data. In practice, such inconsistency (experimental and sampling uncertainties) can be incorporated into probabilistic analyses using random field methods, such as by the approach of David Mašín [86].

The model selection using Bayesian approach from triaxial tests can be considered as a basic work before extending the approach to the practice. The presented work using triaxial tests to select "best model" can give a comprehensive understanding of different sand models on modelling the mechanical behaviours. Although the appropriateness of various constitutive models in FEM (not at the meso level of triaxial tests, but at the level of responses of real structures such as settlements, deflections) is well studied, choosing a suitable model for the engineering of interest solved by FEM is still a challenge. Therefore, it is important to engage the proposed approach to shed a different light on past conclusions from these studies. The proposed approach is more applicable to comparable models (same level of numerical sophistication) and an engineer cannot easily distinguish which models to be adopted in FEM.

7 Conclusions

This study has presented a selection of sand models along with parameter identification using Bayes' theorem, thus evaluating generalization ability based on test data. First, the principles of Bayesian model selection and the enhanced TMCMC method were briefly introduced. Then, the procedure for model class selection and evaluation of generalization ability was presented. To conduct the model class selection, four representative advanced sand models (SIMSAND, SANISAND, HYPOSAND and MCCSAND) were chosen. Three drained triaxial tests on Hostun sand were selected to make five combinations as five groups of training data. Then, for each combination, Bayesian model selection was performed using the enhanced TMCMC method. The optimum parameters corresponding to the maximum of posterior PDF with their uncertainties for each sand model were summarized. The plausibilities were compared for ~~six~~ four selected sand model classes. The generalization ability of each selected sand model was then evaluated by empirical error computed on new test data. Finally, a validation case based on Karlsruhe sand was carried out. Based on the obtained results, some conclusions were drawn, as follows:

(1) The SIMSAND model was demonstrated to be the most plausible/suitable sand model in terms of predictive ability, generalization ability and model complexity regardless of the number and types of training tests involved.

(2) Although the performance of HYPOSAND was not quite satisfactory compared to SIMSAND and SANISAND regardless of the number of tests involved, it was better than the MCCSAND– because of its insensitivity to parameters.

(3) The poor generalization ability of the MCCSAND was found when the involved training tests on a loose sample were dominant, which can be attributed to inaccurately identified parameters because its predictive ability was acceptable based on the results of Bayesian

model class selection. The results also suggest that the tests performed on dense sand are beneficial to identify reliable parameters of the MCCSAND.

Note that the anisotropic behaviour of granular materials was not considered and tests with more complex loading conditions were not included in this study. In fact, this study of model class selection on laboratory tests can be extended for any kind of soils and advanced constitutive models with considering material anisotropy and complex loading conditions, and will be a base for the future studies on boundary value problems using finite element method.

Acknowledgments

The financial support for this research came from the National Natural Science Foundation of China (51579179), and the Macau Science and Technology Development Fund (FDCT/125/2014/A3).

Appendix I

(1) Typical constitutive relations of SIMSAND, SANISAND and MCCSAND

Model class	SIMSAND	SANISAND	MCCSAND
Elasticity	$d\varepsilon_{ij}^{e}=\frac{1+\upsilon}{E}\sigma'_{ij}-\frac{\upsilon}{E}\sigma'_{kk}\delta_{ij}$		
Yield function	$f=\frac{q}{p'}-H=0$	$f=\left(q-p\alpha\right)^{2}-m^{2}p^{2}=0$	$f=\ln\left[\left(1+\frac{q^{2}}{M^{2}p^{2}-\chi q^{2}}\right)p+p_{s}\right]-\ln\left(p_{x0}-p_{s}\right)-\frac{1}{c_{p}}H=0$
Potential function	$\frac{\partial g}{\partial p'}=A_{d}\left(M_{pt}-\frac{q}{p'}\right)$ $\frac{\partial g}{\partial q}=1$ $M_{pt}=\frac{6\sin\phi_{pt}}{3-\sin\phi_{pt}}$	$D=A_{d}\left(M_{pt}-\alpha\right)$ $M_{pt}=\frac{6\sin\phi_{pt}}{3-\sin\phi_{pt}}$	$g=\ln\frac{p}{p_{x}}+\ln\left(1+\frac{q^{2}}{M_{c}^{2}p^{2}}\right)$ $M_{c}=M\exp\left(-m\xi\right)$ $\xi=e_{\eta}-e$
Hardening law	$H=\frac{M_{p}\varepsilon_{d}^{p}}{k_{p}+\varepsilon_{d}^{p}}$ with $M_{p}=\frac{6\sin\phi_{p}}{3-\sin\phi_{p}}$	$h=k_{p}\frac{\lvert\mathbf{b}:\mathbf{n}\rvert}{b_{ref}-\lvert\mathbf{b}:\mathbf{n}\rvert}$	$H=\int\frac{M_{f}^{4}-\eta^{4}}{M_{c}^{4}-\eta^{4}}d\varepsilon_{v}^{p}$ $M_{f}=6\left[\sqrt{\frac{k}{R}\left(1+\frac{k}{R}\right)}-\frac{k}{R}\right]$ $k=\frac{M^{2}}{12\left(3-M\right)},R=\exp\left(-\frac{e_{\eta}-e}{\lambda-\kappa}\right)$
Critical state	$e_{c}=e_{ref}\exp\left[-\lambda\left(\frac{p'}{p_{at}}\right)^{\xi}\right]$	$e_{c}=e_{ref}\exp\left[-\lambda\left(\frac{p'}{p_{at}}\right)^{\xi}\right]$	$e_{\eta}=Z-\lambda\ln\left(\frac{p+p_{s}}{1+p_{s}}\right)-\left(\lambda-\kappa\right)\ln\left[\frac{\left(1+\eta^{2}/\left(M^{2}-\chi\eta^{2}\right)\right)p+p_{s}}{p+p_{s}}\right]$
Inter-locking	$\tan\phi_{p}=\left(\frac{e_{c}}{e}\right)^{n_{p}}\tan\phi_{\mu}$; $\tan\phi_{pt}=\left(\frac{e_{c}}{e}\right)^{-n_{d}}\tan\phi_{\mu}$	$\tan\phi_{p}=\left(\frac{e_{c}}{e}\right)^{n_{p}}\tan\phi_{\mu}$; $\tan\phi_{pt}=\left(\frac{e_{c}}{e}\right)^{-n_{d}}\tan\phi_{\mu}$	$m=-\frac{1}{\xi_{c}}\ln\left(\frac{M_{c}^{\xi_{c}}}{M}\right)$ $\xi_{c}=\left(e_{\eta}-e\right)_{c}$ corresponding to $M_{c}^{\xi_{c}}$
Number of parameters	10	10	10

(2) Typical constitutive relations of HYPOSAND

Components	Constitutive equations
Constitutive equation	$\dot{\boldsymbol{\sigma}} = C_1(\mathrm{tr}\boldsymbol{\sigma})\dot{\boldsymbol{\varepsilon}} + C_2(\mathrm{tr}\dot{\boldsymbol{\varepsilon}})\boldsymbol{\sigma} + C_3\frac{\mathrm{tr}(\boldsymbol{\sigma}\dot{\boldsymbol{\varepsilon}})}{\mathrm{tr}\boldsymbol{\sigma}} + C_4(\boldsymbol{\sigma}+\boldsymbol{\sigma}^*)\lVert\dot{\boldsymbol{\varepsilon}}\rVert I_e$
Critical state line	$I_e = (e/e_c)^{\alpha}$ $e_c = e_{ref}\exp\left[-\lambda\left(p'/p_{at}\right)^{\xi}\right]$
Translated tensor for cohesion	$\boldsymbol{\sigma}_c = \boldsymbol{\sigma} - p_t\delta_{ij}$, and $p_t = c/\tan\varphi$
Number of parameters	8

in which C_i (i = 1; 2; 3; 4) are dimensionless parameters. The deviatoric stress tensor $\boldsymbol{\sigma}^*$ is defined by $\boldsymbol{\sigma}^* = \boldsymbol{\sigma} - 1/3(\mathrm{tr}\boldsymbol{\sigma})\boldsymbol{\delta}_{ij}$ with $\boldsymbol{\delta}_{ij}$ being the Kronecker delta. $\lVert\dot{\varepsilon}\rVert$ stands for the Euclidean norm of the stretching tensor defined as $\lVert\dot{\varepsilon}\rVert = \sqrt{tr(\dot{\varepsilon}^2)}$. I_e is the critical state function that describes the effects of void ratio and stress level. $\boldsymbol{\sigma}_c$ is a translated stress tensor. By replacing the stress tensor with the translated stress tensor, the practical hypoplastic model can describe the effects of cohesion. Note that four material parameters C_1, C_2, C_3, and C_4 were replaced by E, υ, ψ, and φ via a relationship according to Wu et al. [57] in this study.

Appendix II-basic information of Bayesian model class selection

Let D denote the input–output or output-only data from a physical system or phenomenon. Note that the data D can be stress-strain curve from laboratory tests and monitoring data from field tests in geotechnical engineering. The goal is to use D to select the most plausible/suitable class of models representing the system out of N_C given classes of models C_1, C_2, . . . , C_{NC}, such as the advanced sand models. Since probability may be interpreted as a measure of plausibility based on specified information, the probability of a class of models conditional on the set of dynamic data D is required. This can be obtained by using Bayes' theorem as follows:

$$p\left(C_j\middle|D\right) = \frac{p\left(D\middle|C_j\right)P\left(C_j\right)}{p(D)}, j = 1,\ 2,\ ...,\ N_c \tag{16}$$

where $p\left(C_j \middle| D\right)$ is the plausibility of a predictive model class C_j given the data D, $p\left(D \middle| C_j\right)$ is called the evidence of model class C_j provided by the data D, $P\left(C_j\right)$ is the prior plausibility of a model class C_j, $p(D)$ is the denominator, given by the law of total probability:

$$p\left(D\right)=\sum_{j=1}^{N_c} p\left(D \middle| C_j\right) p\left(C_j\right) \tag{17}$$

The prior plausibilities are normalized in the same way as probabilities:

$$\sum_{j=1}^{N_c} p\left(C_j\right)=1 \tag{18}$$

$$p\left(D \middle| C_j\right)=\int_{\Theta} p\left(D \middle| \boldsymbol{\theta}, C_j\right) p\left(\boldsymbol{\theta} \middle| C_j\right) d\boldsymbol{\theta},\ j=1,\ 2,\ ...,\ N_c \tag{19}$$

Note that $p\left(C_j \middle| D\right)$, $p\left(D \middle| C_j\right)$, $p(D)$, $p\left(C_j\right)$, $p\left(D \middle| \boldsymbol{\theta}, C_j\right)$ and $p\left(\boldsymbol{\theta} \middle| C_j\right)$ are probabilities, not probability density functions.

References

[1] S.L. Shen, Y.S. Xu, Numerical evaluation of land subsidence induced by groundwater pumping in Shanghai, Canadian Geotechnical Journal, 48 (2011) 1378-1392.

[2] S.-L. Shen, H.-N. Wu, Y.-J. Cui, Z.-Y. Yin, Long-term settlement behaviour of metro tunnels in the soft deposits of Shanghai, Tunnelling and Underground Space Technology, 40 (2014) 309-323.

[3] M. Karstunen, Z.Y. Yin, Modelling time-dependent behaviour of Murro test embankment, Geotechnique, 60 (2010) 735-749.

[4] M. Jiang, D. Harris, H. Zhu, Future continuum models for granular materials in penetration analyses, Granular Matter 9(2007) 97-108.

[5] J.-C. Chai, J.S.-L. Shen, M.D. Liu, D.-J. Yuan, Predicting the performance of embankments on PVD-improved subsoils, Computers and Geotechnics, 93 (2018) 222-231.

[6] Y.-F. Jin, Z.-Y. Yin, Z.-X. Wu, A. Daouadji, Numerical modeling of pile penetration in silica sands considering the effect of grain breakage, Finite Elem. Anal. Des. , 144 (2018) 15-29.

[7] Y.-F. Jin, Z.-Y. Yin, W.-H. Zhou, H.-W. Huang, Multi-objective optimization-based updating of predictions during excavation, Eng. Appl. Artif. Intell. , 78 (2019) 102-123.

[8] Y.-F. Jin, Z.-Y. Yin, S.-L. Shen, P.-Y. Hicher, Selection of sand models and identification of parameters using an enhanced genetic algorithm, Int. J. Numer. Anal. Methods Geomech. , 40 (2016) 1219-1240.

[9] M. Jiang, D. Harris, H. Yu, Kinematic models for non - coaxial granular materials. Part I: theory, Int. J. Numer. Anal. Methods Geomech. , 29 (2005) 643-661.

[10] M. Jiang, F. Zhang, Y. Sun, An evaluation on the degradation evolutions in three constitutive models for bonded geomaterials by DEM analyses, Computers and Geotechnics, 57 (2014) 1-16.

[11] Y.-F. Jin, Z.-Y. Yin, Z.-X. Wu, W.-H. Zhou, Identifying parameters of easily crushable sand and application to offshore pile driving, Ocean Eng. , 154 (2018) 416-429.

[12] M. Jiang, T. Li, B. Chareyre, Fabric rates applied to kinematic models: evaluating elliptical granular materials under simple shear tests via discrete element method, Granular Matter 18 (2016) 1-15.

[13] H.-M. Lyu, J.S. Shen, A. Arulrajah, Assessment of geohazards and preventative countermeasures using AHP incorporated with GIS in Lanzhou, China, Sustainability, 10 (2018) 304.

[14] D.-J. Ren, S.-L. Shen, A. Arulrajah, H.-N. Wu, Evaluation of ground loss ratio with moving trajectories induced in DOT tunnelling, Canadian Geotechnical Journal, 55 (2018) 894-902.

[15] H.-M. Lyu, W.-J. Sun, S.-L. Shen, A. Arulrajah, Flood risk assessment in metro systems of mega-cities using a GIS-based modeling approach, Sci. Total Environ. , 626 (2018) 1012-1025.

[16] H.-M. Lyu, S.-L. Shen, A. Zhou, J. Yang, Perspectives for flood risk assessment and management for mega-city metro system, Tunnelling and Underground Space Technology, 84 (2019) 31-44.

[17] Y.-X. Wu, H.-M. Lyu, J. Han, S.-L. Shen, Dewatering–Induced Building Settlement around a Deep Excavation in Soft Deposit in Tianjin, China, Journal of Geotechnical and Geoenvironmental Engineering, 145 (2019) 05019003.

[18] A. Whittle, R. Davies, Nicoll Highway collapse: evaluation of geotechnical factors affecting design of excavation support system, in: International conference on deep excavations, 2006, pp. 30.

[19] Y.-F. Jin, Z.-X. Wu, Z.-Y. Yin, J.S. Shen, Estimation of critical state-related formula in advanced constitutive modeling of granular material, Acta Geotech. , 12 (2017) 1329-1351.

[20] O.V. Akeju, K. Senetakis, Y. Wang, Bayesian parameter identification and model selection for normalized modulus reduction curves of soils, J. Earthquake Eng. , (2017) 1-29.

[21] J.L. Beck, K.-V. Yuen, Model selection using response measurements: Bayesian probabilistic approach, Journal of engineering mechanics, 130 (2004) 192-203.

[22] Z. Cao, Y. Wang, Bayesian model comparison and selection of spatial correlation functions for soil parameters, Structural Safety, 49 (2014) 10-17.

[23] S.H. Cheung, J.L. Beck, Bayesian model updating using hybrid Monte Carlo simulation with application to structural dynamic models with many uncertain parameters, Journal of engineering mechanics, 135 (2009) 243-255.

[24] M. Muto, J.L. Beck, Bayesian updating and model class selection for hysteretic structural models using stochastic simulation, J. Vib. Control 14 (2008) 7-34.

[25] T. Ritto, L. Nunes, Bayesian model selection of hyperelastic models for simple and pure shear at large deformations, Computers & Structures, 156 (2015) 101-109.

[26] F. Tan, W.-H. Zhou, K.-V. Yuen, Modeling the soil water retention properties of same-textured soils with different initial void ratios, J. Hydrol. , 542 (2016) 731-743.

[27] K. Worden, J. Hensman, Parameter estimation and model selection for a class of hysteretic systems using Bayesian inference, Mechanical Systems and Signal Processing, 32 (2012) 153-169.

[28] K.-V. Yuen, Recent developments of Bayesian model class selection and applications in civil engineering, Structural Safety, 32 (2010) 338-346.

[29] K.-V. Yuen, Bayesian methods for structural dynamics and civil engineering, John Wiley & Sons, 2010.

[30] L. Zhang, D.-Q. Li, X.-S. Tang, Z.-J. Cao, K.-K. Phoon, Bayesian model comparison and characterization of bivariate distribution for shear strength parameters of soil, Computers and Geotechnics, 95 (2017) 110-118.

[31] J. Zhang, H. Wang, H. Huang, L. Chen, System reliability analysis of soil slopes stabilized with piles, Eng. Geol. , 229 (2017) 45-52.

[32] W.-H. Zhou, K.-V. Yuen, F. Tan, Estimation of maximum pullout shear stress of grouted soil nails using Bayesian probabilistic approach, Int. J. Geomech. , 13 (2012) 659-664.

[33] F. Tan, W.H. Zhou, K.V. Yuen, Effect of loading duration on uncertainty in creep analysis of clay, Int. J. Numer. Anal. Methods Geomech. , 42 (2018) 1235-1254.

[34] W.-H. Zhou, F. Tan, K.-V. Yuen, Model updating and uncertainty analysis for creep behavior of soft soil, Computers and Geotechnics, 100 (2018) 135-143.

[35] C.F. Chiu, W.M. Yan, K.-V. Yuen, Estimation of water retention curve of granular soils from particle-size distribution—a Bayesian probabilistic approach, Canadian Geotechnical Journal, 49 (2012) 1024-1035.

[36] Y.-F. Jin, Z.-Y. Yin, Y. Riou, P.-Y. Hicher, Identifying creep and destructuration related soil parameters by optimization methods, KSCE Journal of Civil Engineering, 21 (2017) 1123-1134.

[37] Y.-F. Jin, Z.-Y. Yin, S.-L. Shen, P.-Y. Hicher, Investigation into MOGA for identifying parameters of a critical-state-based sand model and parameters correlation by factor analysis, Acta Geotech. , 11 (2016) 1131-1145.

[38] Y.-F. Jin, Z.-Y. Yin, S.-L. Shen, D.-M. Zhang, A new hybrid real-coded genetic algorithm and its application to parameters identification of soils, Inverse Problems in Science and Engineering, 25 (2017) 1343-1366.

[39] Z.-Y. Yin, Y.-F. Jin, J.S. Shen, P.-Y. Hicher, Optimization techniques for identifying soil parameters in geotechnical engineering: Comparative study and enhancement, Int. J. Numer. Anal. Methods Geomech. , 42 (2018) 70-94.

[40] Z.-Y. Yin, Y.-F. Jin, S.-L. Shen, H.-W. Huang, An efficient optimization method for identifying parameters of soft structured clay by an enhanced genetic algorithm and elastic–viscoplastic model, Acta Geotech. , 12 (2017) 849-867.

[41] Y.-F. Jin, Z.-Y. Yin, W.-H. Zhou, J.-H. Yin, J.-F. Shao, A single-objective EPR based model for creep index of soft clays considering L2 regularization, Eng. Geol. , 248 (2019) 242-255.

[42] T. Toni, D. Welch, N. Strelkowa, A. Ipsen, M.P. Stumpf, Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems, Journal of the Royal Society Interface, 6 (2009) 187-202.

[43] K.V. Yuen, H.Q. Mu, Real‐time system identification: an algorithm for simultaneous model class selection and parametric identification, Computer‐Aided Civil and Infrastructure Engineering, 30 (2015) 785-801.

[44] T. Schanz, P. Vermeer, P. Bonnier, The hardening soil model: formulation and verification, Beyond 2000 in computational geotechnics, (1999) 281-296.

[45] P. Vermeer, A double hardening model for sand, Geotechnique, 28 (1978) 413-433.

[46] M. Jefferies, Nor-Sand: a simle critical state model for sand, Geotechnique, 43 (1993) 91-103.

[47] H. Yu, CASM: A unified state parameter model for clay and sand, Int. J. Numer. Anal. Methods Geomech. , 22 (1998) 621-653.

[48] A. Gajo, M. Wood, Severn–Trent sand: a kinematic-hardening constitutive model: the q–p formulation, Geotechnique, 49 (1999) 595-614.

[49] Y. Yao, W. Hou, A. Zhou, UH model: three-dimensional unified hardening model for overconsolidated clays, Geotechnique, 59 (2009) 451-469.

[50] Y. Yao, D. Sun, T. Luo, A critical state model for sands dependent on stress and density, Int. J. Numer. Anal. Methods Geomech. , 28 (2004) 323-337.

[51] Y. Yao, D. Sun, H. Matsuoka, A unified constitutive model for both clay and sand with hardening parameter independent on stress path, Computers and Geotechnics, 35 (2008) 210-222.

[52] Y.-P. Yao, L.-M. Kong, A.-N. Zhou, J.-H. Yin, Time-dependent unified hardening model: Three-dimensional elastoviscoplastic constitutive model for clays, Journal of engineering mechanics, 141 (2014) 04014162.

[53] M. Taiebat, Y.F. Dafalias, SANISAND: Simple anisotropic sand plasticity model, Int. J. Numer. Anal. Methods Geomech. , 32 (2008) 915-948.

[54] Z.-X. Wu, Z.-Y. Yin, Y.-F. Jin, X.-Y. Geng, A straightforward procedure of parameters determination for sand: a bridge from critical state based constitutive modelling to finite element analysis, European Journal of Environmental and Civil Engineering, (2017) 1-23.

[55] D. Kolymbas, An outline of hypoplasticity, Archive of applied mechanics, 61 (1991) 143-151.

[56] D. Mašín, A hypoplastic constitutive model for clays, Int. J. Numer. Anal. Methods Geomech. , 29 (2005) 311-336.

[57] W. Wu, E. Bauer, D. Kolymbas, Hypoplastic constitutive model with critical state for granular materials, Mech. Mater. , 23 (1996) 45-69.

[58] W. Wu, D. Kolymbas, Hypoplasticity then and now, Constitutive modelling of granular materials, (2000) 57-105.

[59] S. Wang, W. Wu, Z.-Y. Yin, C. Peng, X.-Z. He, Modelling time-dependent behaviour of granular material with hypoplasticity, Int. J. Numer. Anal. Methods Geomech. , 42 (2018) 1331-1345.

[60] F. Richart, J. Hall, R. Woods, Vibrations of soils and foundations. International Series in Theoretical and Applied Mechanics, in, Englewood Cliffs, NJ: Prentice-Hall, 1970.

[61] Z.-Y. Yin, Z.-X. Wu, P.-Y. Hicher, Modeling Monotonic and Cyclic Behavior of Granular Materials by Exponential Constitutive Function, Journal of engineering mechanics, 144 (2018) 04018014.

[62] J.L. Beck, Bayesian system identification based on probability logic, Structural Control and Health Monitoring, 17 (2010) 825-847.

[63] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, Journal of engineering mechanics, 124 (1998) 455-461.

[64] J. He, J.W. Jones, W.D. Graham, M.D. Dukes, Influence of likelihood function choice for estimating crop model parameters using the generalized likelihood uncertainty estimation method, Agricultural Systems, 103 (2010) 256-264.

[65] J. Ching, Y.-C. Chen, Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging, Journal of engineering mechanics, 133 (2007) 816-832.

[66] W. Betz, I. Papaioannou, D. Straub, Transitional markov chain monte carlo: observations and improvements, Journal of engineering mechanics, 142 (2016) 04016016.

[67] J. Ching, J.-S. Wang, Application of the transitional Markov chain Monte Carlo algorithm to probabilistic site characterization, Eng. Geol. , 203 (2016) 151-167.

[68] S.-H. Lee, J. Song, System Identification of Spatial Distribution of Structural Parameters Using Modified Transitional Markov Chain Monte Carlo Method, Journal of engineering mechanics, 143 (2017) 04017099.

[69] N. Chopin, A sequential particle filter method for static models, Biometrika, 89 (2002) 539-552.

[70] J.A. Vrugt, Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation, Environmental Modelling & Software, 75 (2016) 273-316.

[71] Y.-J. Liu, G. Li, Z.-Y. Yin, C. Dano, P.-Y. Hicher, X.-H. Xia, J.-H. Wang, Influence of grading on the undrained behavior of granular materials, C.R. Mec. , 342 (2014) 85-95.

[72] G. Li, Y.-J. Liu, C. Dano, P.-Y. Hicher, Grading-dependent behavior of granular materials: from discrete to continuous modeling, Journal of engineering mechanics, 141 (2014) 04014172.

[73] D.M. Wood, Geotechnical modelling, CRC Press, 2003.

[74] H. Akaike, On entropy maximization principle, Application of statistics, (1977) 27-41.

[75] G. Schwarz, Estimating the dimension of a model, The annals of statistics, 6 (1978) 461-464.

[76] T. Wichtmann, T. Triantafyllidis, An experimental database for the development, calibration and verification of constitutive models for sand with focus to cyclic loading: part I—tests with monotonic loading and stress cycles, Acta Geotech. , 11 (2016) 739-761.

[77] Z. Gao, J. Zhao, A non-coaxial critical-state model for sand accounting for fabric anisotropy and fabric evolution, Int. J. Solids Struct. , 106 (2017) 200-212.

[78] Z. Gao, J. Zhao, X.S. Li, Y.F. Dafalias, A critical state sand plasticity model accounting for fabric evolution, Int. J. Numer. Anal. Methods Geomech. , 38 (2014) 370-390.

[79] Y.F. Dafalias, A.G. Papadimitriou, X.S. Li, Sand plasticity model accounting for inherent fabric anisotropy, Journal of engineering mechanics, 130 (2004) 1319-1333.

[80] X. Li, Y. Dafalias, A constitutive framework for anisotropic sand including non-proportional loading, Géotechnique, 54 (2004) 41-55.

[81] J. Zhao, N. Guo, Unique critical state characteristics in granular media considering fabric anisotropy, Géotechnique, 63 (2013) 695.

[82] C.S. Chang, Z.Y. Yin, Micromechanical Modeling for Inherent Anisotropy in Granular Materials, Journal of Engineering Mechanics-Asce, 136 (2010) 830-839.

[83] Z.Y. Yin, C.S. Chang, P.Y. Hicher, Micromechanical modelling for effect of inherent anisotropy on cyclic behaviour of sand, Int. J. Solids Struct. , 47 (2010) 1933-1951.

[84] H. Xiong, F. Nicot, Z. Yin, A three - dimensional micromechanically based model, Int. J. Numer. Anal. Methods Geomech. , 41 (2017) 1669-1686.

[85] H. Xiong, F. Nicot, Z. Yin, From micro scale to boundary value problem: using a micromechanically based model, Acta Geotech. , (2018) 1-17.

[86] D. Mašín, The influence of experimental and sampling uncertainties on the probability of unsatisfactory performance in geotechnical applications, Géotechnique, 65 (2015) 897-910.

Tables

Table 1 Six model class candidates and their uncertain parameters

Model class	Uncertain parameter vector $\boldsymbol{\theta}$	Number of uncertain parameters
SIMSAND	$\left[e_{ref},\ \lambda,\ \xi,\ \phi,\ k_{\mathrm{p}},\ A_d,\ n_p,\ n_d,\ \sigma_\varepsilon^2\right]$	9
SANISAND	$\left[e_{ref},\ \lambda,\ \xi,\ \phi,\ h_0,\ A_d,\ n_p,\ n_d,\ \sigma_\varepsilon^2\right]$	9
HYPOSAND	$\left[e_{ref},\ \lambda,\ \xi,\ E,\ \phi,\ A_d,\ \sigma_\varepsilon^2\right]$	7
MCCSAND	$\left[\phi,\ \kappa,\ \lambda,\ N,\ Z,\ \chi,\ m,\ \sigma_\varepsilon^2\right]$	8

Table 2 Model class selection results for illustration case

Model number	Model	$\ln p\left(D\middle\|C_j\right)$	Model number	Model	$\ln p\left(D\middle\|C_j\right)$
1	$y=a_0+\varepsilon$	-4879.1	9	$y=a_1x+a_3x^3+\varepsilon$	-45.83
2	$y=a_1x+\varepsilon$	-1009.5	10	$y=a_2x^2+a_3x^3+\varepsilon$	-48.78
3	$y=a_2x^2+\varepsilon$	-44.17	11	$y=a_0+a_1x+a_2x^2+\varepsilon$	36.62
4	$y=a_3x^3+\varepsilon$	-394.3	12	$y=a_0+a_1x+a_3x^3+\varepsilon$	-49.27
5	$y=a_0+a_1x+\varepsilon$	-850.2	13	$y=a_0+a_2x^2+a_3x^3+\varepsilon$	-33.81
6	$y=a_0+a_2x^2+\varepsilon$	-27.21	14	$y=a_1x+a_2x^2+a_3x^3+\varepsilon$	14.59
7	$y=a_0+a_3x^3+\varepsilon$	-181.3	15	$y=a_0+a_1x+a_2x^2+a_3x^3+\varepsilon$	-19.60
8	$y=a_1x+a_2x^2+e$	**165.1**			

Table 3 Summary of triaxial tests on Hostun sand

Triaxial tests	Drained tests					Undrained tests				
Test number	**1**	2	3	**4**	**5**	6	7	8	9	10
e_0	**0.66**	0.75	0.85	**0.83**	**0.82**	0.69	0.72	0.70	0.72	0.78
p' / kPa	**100**	100	100	**200**	**400**	100	100	400	400	400

Table 4 Test combinations for model class selection

Test combination	**1**	**2**	3	**4**	5	**6**	**7**
Test number	**1**	**4**	5	**1 and 4**	1 and 5	**4 and 5**	**1, 4 and 5**

Table 5 Bounds of all uncertain parameters in the prior PDF used in TMCMC

Soil models	Parameters								σ_{ε}^{2}
SIMSAND	e_{ref}	λ	ξ	ϕ /°	k_{p}	A_{d}	n_{p}	n_{d}	[0, 1]
	[0.5, 1.5]	[10^{-3}, 10^{-1}]	[0.1, 1.0]	[20, 50]	[10^{-3},10^{-1}]	[0.1, 3]	[0, 10]	[0, 10]	
SANISAND	e_{ref}	λ	ξ	ϕ /°	h_0	A_{d}	n_{p}	n_{d}	
	[0.5, 1.5]	[10^{-3}, 10^{-1}]	[0.1, 1.0]	[20, 50]	[1, 100]	[0.1, 3]	[0, 10]	[0, 10]	
HYPOSAND	e_{ref}	λ	ξ	E / kPa	ϕ /°	A_{d}			
	[0.5, 1.5]	[10^{-3}, 10^{-1}]	[0.1, 1.0]	[10^3, 5×10^4]	[20, 50]	[0.1, 3]			
MCCSAND	ϕ /°	κ	λ	N	Z	χ	m		
	[20, 50]	[10^{-3}, 10^{-1}]	[0.01, 1.0]	[0.5, 3.0]	[0.5, 2.0]	[0, 1]	[0, 10]		

Table 6 Results of model class selection based on limited number of tests

Tests	Test 1		Test 4		Tests 1 and 4		Tests 4 and 5		Tests 1, 4 and 5	
Model class	Ln(S)	$p(C\|D)$	Ln(S)	$p(C\|D)$	Ln(S)	$p(C\|D)$	Ln(S)	$p(C\|D)$	Ln(S)	$p(C\|D)$
SIMSAND	70.29	0.114	328.02	0.209	385.25	0.401	678.79	0.218	697.85	0.351
SANISAND	361.91	0.589	340.57	0.217	341.22	0.355	819.48	0.263	676.03	0.340
HYPOSAND	47.40	0.077	253.25	0.161	178.79	0.186	557.30	0.179	384.78	0.193
MCCSAND	54.96	0.089	273.10	0.174	279.40	0.291	466.10	0.149	522.10	0.262

Table 7 Identified MAP parameters of SIMSAND and their uncertainties

SIMSAND	e_{ref}	λ	ξ	ϕ /°	k_{p}	A_{d}	n_{p}	n_{d}
Based on test 1	0.765 (3.59E-3)	0.0122 (2.05E-3)	0.835 (3.78E-2)	27.8 (1.78E-1)	1.47E-03 (7.84E-5)	0.46 (1.96E-2)	1.8 (7.28E-2)	3.9 (2.16E-1)
Based on test 4	0.718 (8.07E-4)	0.0076 (9.56E-4)	0.313 (2.75E-2)	29.3 (4.54E-2)	7.75E-03 (4.04E-4)	0.73 (2.24E-2)	1.4 (1.06E-1)	4.2 (1.51E-1)
Based on test 1 and 4	0.750 (7.29E-4)	0.0254 (6.21E-4)	0.684 (1.54E-2)	29.3 (3.61E-2)	3.72E-03 (1.17E-4)	0.55 (6.54E-3)	2.8 (3.28E-2)	6.8 (7.45E-2)
Based on test 4 and 5	0.775 (2.47E-3)	0.0478 (2.63E-3)	0.464 (1.53E-2)	29.1 (3.92E-2)	9.36E-03 (3.59E-4)	0.64 (1.84E-2)	1.5 (1.04E-1)	6.0 (2.73E-1)
Based on test 1, 4 and 5	0.751 (1.03E-3)	0.0265 (9.47E-4)	0.628 (1.17E-2)	29.0 (4.11E-2)	4.54E-03 (1.35E-4)	0.56 (2.35E-2)	3.0 (3.66E-2)	7.0 (4.06E-1)

Table 8 Identified MAP parameters of SANISAND and their uncertainties

SANISAND	e_{ref}	λ	ξ	ϕ /°	h_0	A_{d}	n_{p}	n_{d}
Based on test 1	0.735 (1.02E-3)	0.0158 (5.80E-4)	0.832 (7.94E-3)	28.6 (0.0358)	19.3 (0.37)	0.54 (7.0E-3)	1.1 (7.5E-2)	6.7 (8.59E-2)
Based on test 4	0.710 (1.19E-3)	0.0052 (8.70E-4)	0.501 (3.13E-2)	28.6 (0.0593)	19.9 (0.38)	0.80 (2.25E-2)	1.1 (8.65E-2)	3.4 (1.48E-1)
Based on test 1 and 4	0.740 (1.23E-3)	0.0169 (8.81E-4)	0.803 (2.57E-2)	28.6 (0.0429)	29.0 (0.85)	0.61 (1.85E-2)	2.2 ()5.95E-2	6.1 (2.07E-1)
Based on test 4 and 5	0.738 (8.13E-4)	0.0225 (6.23E-4)	0.615 (8.66E-3)	28.1 (0.0259)	18.6 (0.12)	0.52 (9.7E-3)	0.6 (4.15E-2)	7.7 (1.92E-1)
Based on test 1, 4	0.739	0.0156	0.795	28.7	25.5	0.76	2.3	4.2

and 5	(5.87E-4)	(4.71E-4)	(1.15E-2)	(0.0275)	(0.46)	(3.96E-2)	(3.96E-2)	(1.89E-1)

Table 9 Identified MAP parameters of HYPOSAND and their uncertainties

HYPOSAND	e_{ref}	λ	ξ	E / kPa	ϕ /°	A_d
Based on test 1	0.803 (3.72E-3)	0.056 (3.40E-3)	0.759 (0.031)	14109 (192)	27.1 (0.092)	1.44 (0.051)
Based on test 4	0.669 (3.74E-3)	0.003 (5.58E-4)	0.919 (0.101)	4132 (84)	31.4 (0.079)	0.91 (0.032)
Based on test 1 and 4	0.734 (1.85E-3)	0.015 (1.23E-3)	0.733 (0.023)	7883 (83)	28.7 (0.093)	2.47 (0.032)
Based on test 4 and 5	0.746 (3.80E-3)	0.035 (3.06E-3)	0.574 (0.030)	4509 (78)	30.7 (0.057)	1.64 (0.048)
Based on test 1, 4 and 5	0.740 (1.65E-3)	0.021 (1.50E-3)	0.682 (0.026)	7860 (71)	28.8 (0.073)	2.53 (0.032)

Table 10 Identified MAP parameters of MCCSAND and their uncertainties

SANISAND	ϕ /°	κ	λ	N	Z	χ	m
Based on test 1	33.2 (0.26)	0.044 (0.004)	0.18 (0.009)	2.22 (0.069)	0.94 (0.008)	0.61 (0.024)	1.45 (0.133)
Based on test 4	30.3 (0.3)	0.045 (0.007)	0.12 (0.003)	1.47 (0.016)	1.45 (0.032)	0.52 (0.018)	7.26 (0.469)
Based on test 1 and 4	29.8 (0.3)	0.021 (0.003)	0.25 (0.006)	2.51 (0.049)	0.88 (0.004)	0.48 (0.008)	2.41 (0.031)
Based on test 4 and 5	29.6 (0.5)	0.035 (0.002)	0.22 (0.003)	2.59 (0.021)	0.78 (0.004)	0.96 (0.013)	1.30 (0.112)
Based on test 1, 4 and 5	28.4 (0.3)	0.066 (0.006)	0.17 (0.003)	2.15 (0.031)	0.72 (0.002)	0.22 (0.020)	0.95 (0.043)

Table 11 Summary of drained triaxial tests on Karlsruhe fine sand

Test number	1	2	3	4	5	6	7	8	9
e_0	0.996	0.840	0.734	0.735	0.706	0.697	0.960	0.840	0.718
p' / kPa	50	50	50	100	200	300	400	400	400

Table 12 MAP parameters of all sand models for Karlsruhe sand

SIMSAND	e_{ref}	λ	ξ	ϕ /°	k_p	A_d	n_p	n_d
	1.053	0.0238	0.743	33.6	3.84×10^{-3}	0.75	1.6	1.8
SANISAND	e_{ref}	λ	ξ	ϕ /°	h_0	A_d	n_p	n_d
	1.046	0.0137	0.881	33.2	35.8	0.60	1.2	2.3
HYPOSAND	e_{ref}	λ	ξ	E / kPa	ϕ /°	A_d		
	1.183	0.020	0.832	18940	29.0	0.65		
MCCSAND	ϕ /°	κ	λ	N	Z	χ	m	
	34.3	0.044	0.18	2.33	0.96	0.05	3.19	

Figure captions

Fig. 1 Procedure of Bayesian model class selection with generalization ability evaluation

Fig. 2 Results of triaxial tests of Hostun sand

Fig. 3 Model class selection results based on: (a) test 1; (b) test 4; (c) tests 1 and 4; (d) tests 4 and 5; (e) tests 1, 4 and 5

Fig. 4 Results of empirical errors based on: (a) test 1; (b) test 4; (c) tests 1 and 4; (d) tests 4 and 5; (e) tests 1, 4 and 5

Fig. 5 Comparison of triaxial tests between simulations and experiments on Hostun sand

Fig. 6 Results of drained triaxial tests on Karlsruhe fine sand (test 1: e_0=0.996, p'= 50 kPa; test 2: e_0=0.840, p'= 50 kPa; test 3: e_0=0.734, p'= 50 kPa; test 4: e_0=0.735, p'= 100 kPa; test 5: e_0=0.706, p'= 200 kPa; test 6: e_0=0.697, p'= 300 kPa; test 7: e_0=0.960, p'= 400 kPa; test 8: e_0=0.840, p'= 400 kPa; test 9: e_0=0.718, p'= 400 kPa)

Fig. 7 Model class selection results based on training data on Karlsruhe sand

Fig. 8 Results of empirical errors based on testing data for Karlsruhe sand

Fig. 9 Comparison of triaxial tests between simulations and experiments for Karlsruhe sand

Figure 1

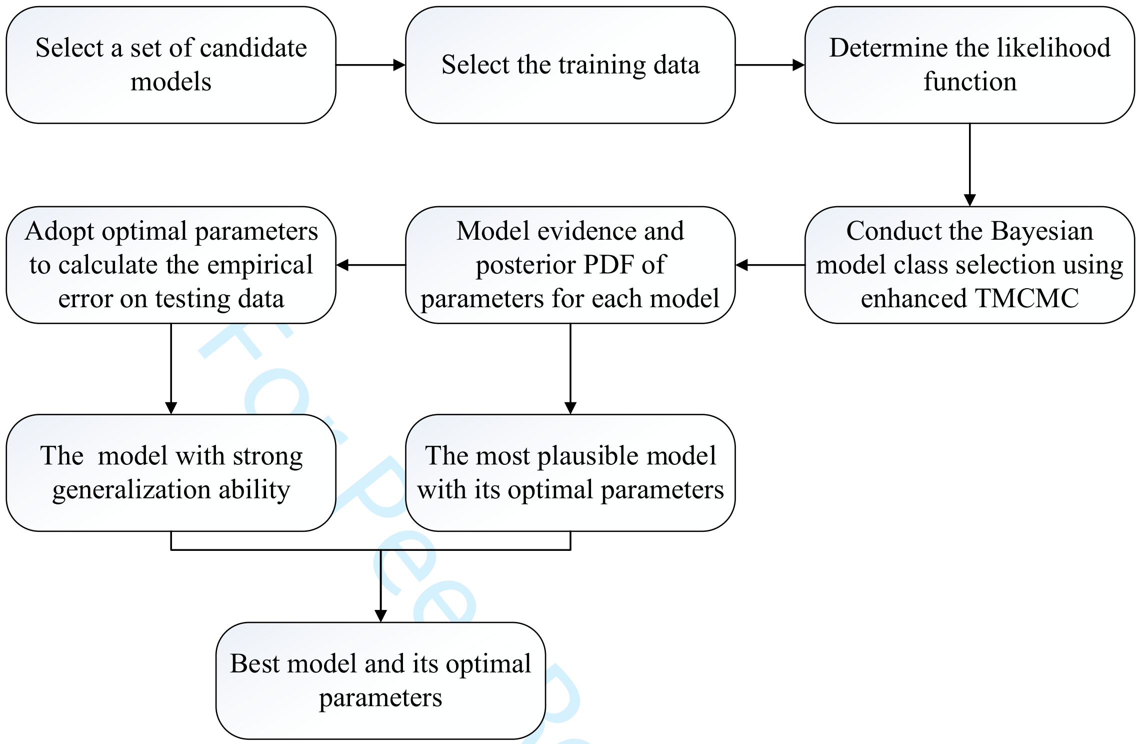

Select a set of candidate models
Select the training data
Determine the likelihood function
Conduct the Bayesian model class selection using enhanced TMCMC
Model evidence and posterior PDF of parameters for each model
Adopt optimal parameters to calculate the empirical error on testing data
The model with strong generalization ability
The most plausible model with its optimal parameters
Best model and its optimal parameters

Figure 2

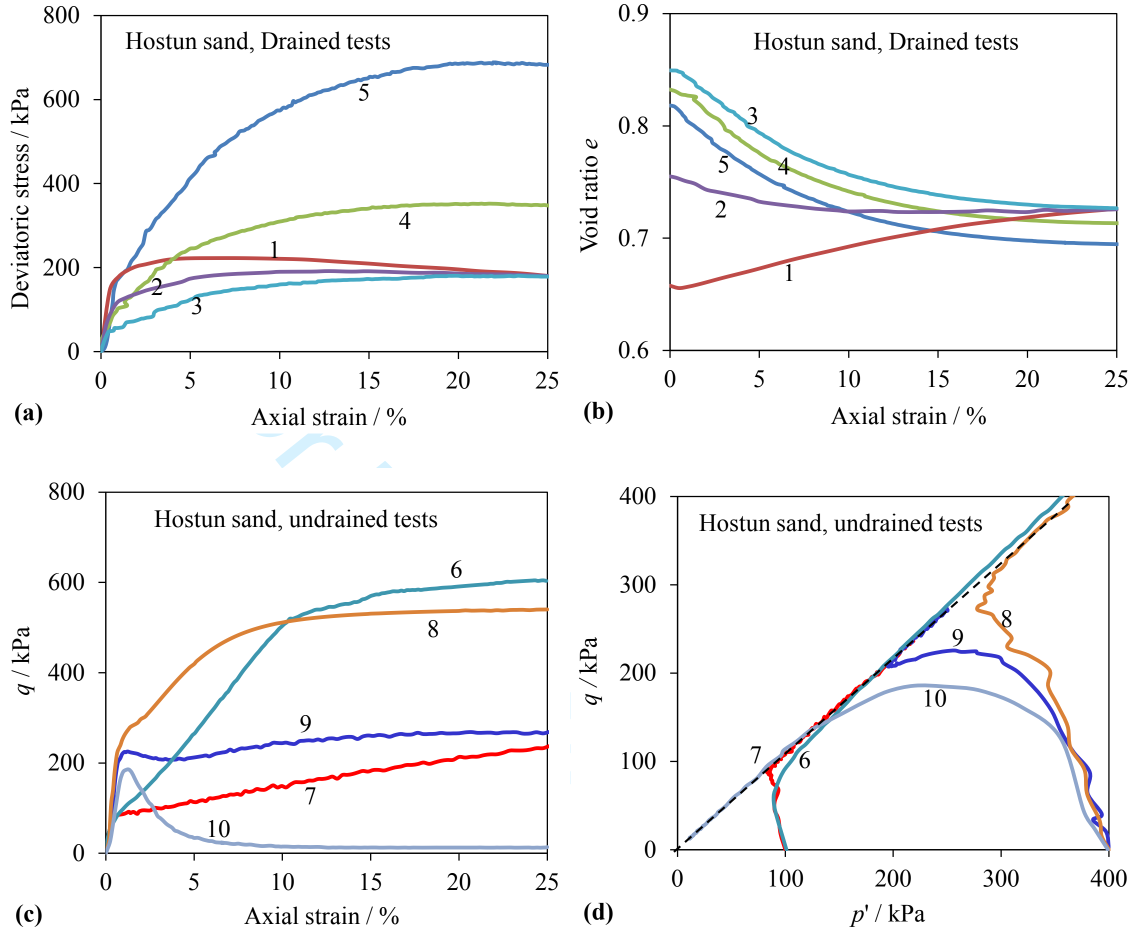
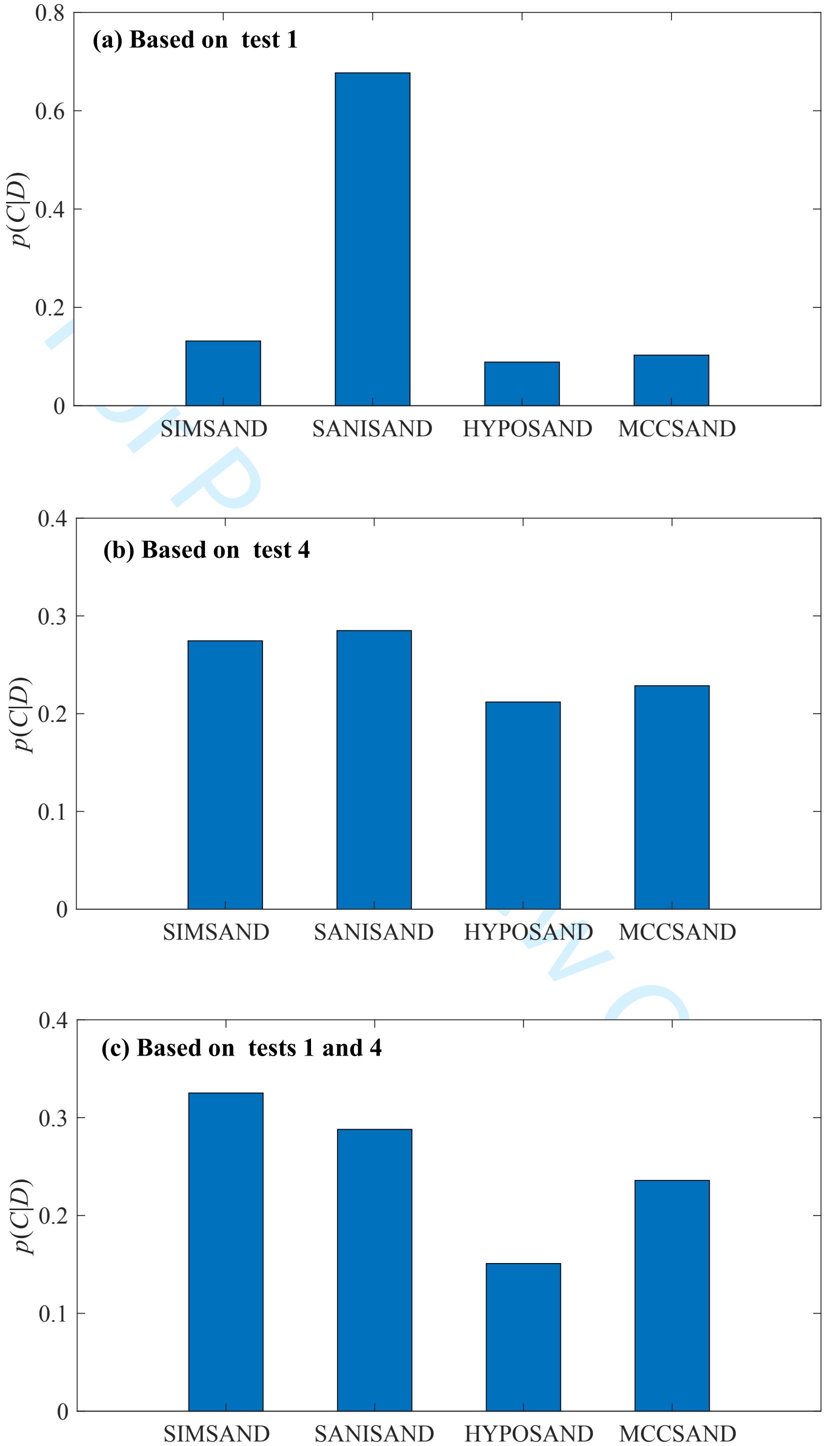
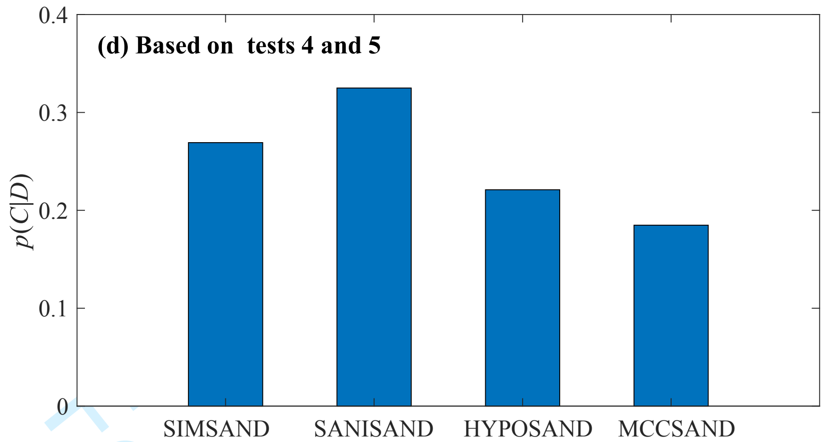
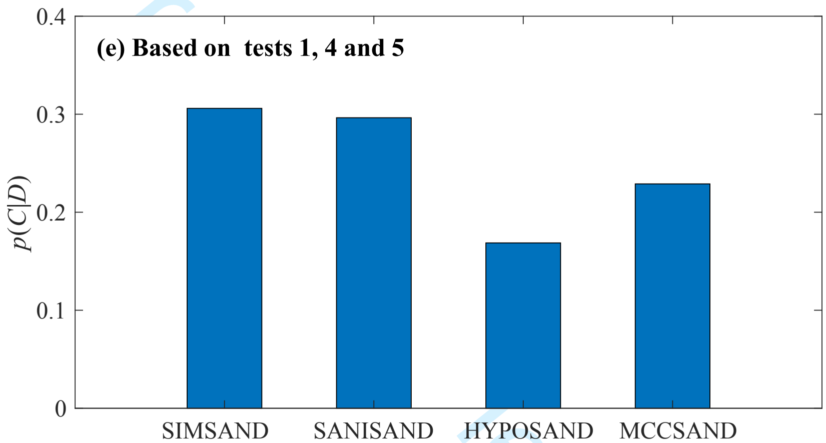

Hostun sand, Drained tests
Deviatoric stress / kPa
Axial strain / %
(a)
Hostun sand, Drained tests
Void ratio e
Axial strain / %
(b)
Hostun sand, undrained tests
q / kPa
Axial strain / %
(c)
Hostun sand, undrained tests
q / kPa
p' / kPa
(d)

Figure 3


(a) Based on test 1
p(C|D)
0.8
0.6
0.4
0.2
0
SIMSAND
SANISAND
HYPOSAND
MCCSAND
(b) Based on test 4
p(C|D)
0.4
0.3
0.2
0.1
0
SIMSAND
SANISAND
HYPOSAND
MCCSAND
(c) Based on tests 1 and 4
p(C|D)
0.4
0.3
0.2
0.1
0
SIMSAND
SANISAND
HYPOSAND
MCCSAND

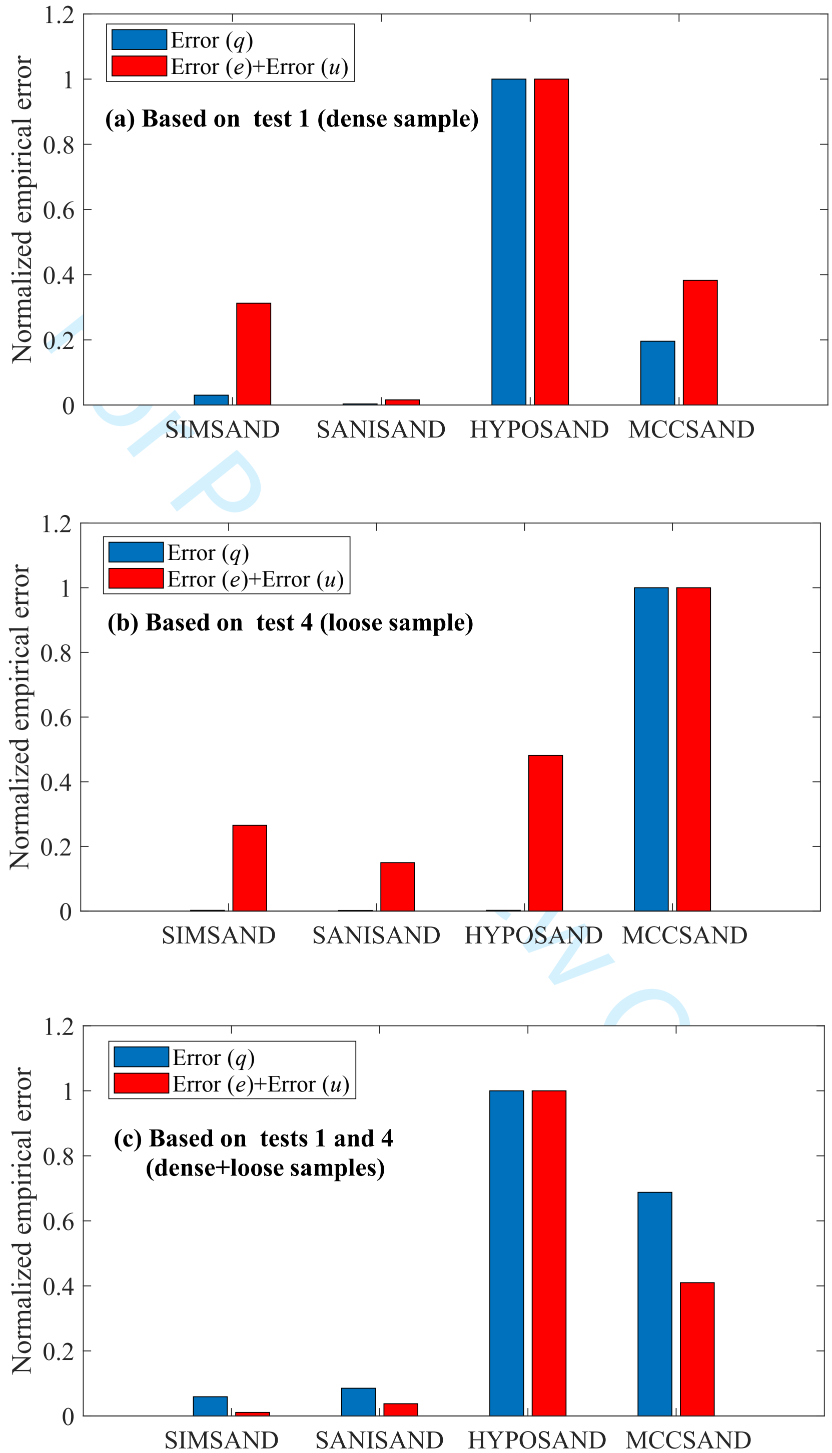


(d) Based on tests 4 and 5
p(C|D)
0
0.1
0.2
0.3
0.4
SIMSAND
SANISAND
HYPOSAND
MCCSAND

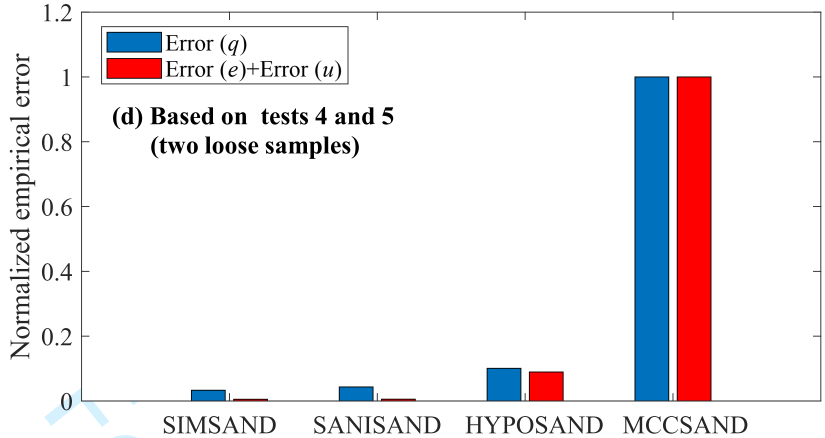


(e) Based on tests 1, 4 and 5
p(C|D)
0
0.1
0.2
0.3
0.4
SIMSAND
SANISAND
HYPOSAND
MCCSAND

Figure 4



1.2
1
0.8
0.6
0.4
0.2
0
Normalized empirical error
Error (q)
Error (e)+Error (u)
(a) Based on test 1 (dense sample)
SIMSAND
SANISAND
HYPOSAND
MCCSAND
(b) Based on test 4 (loose sample)
(c) Based on tests 1 and 4
(dense+loose samples)


Error (q)
Error (e)+Error (u)
(d) Based on tests 4 and 5
(two loose samples)
Normalized empirical error
1.2
1
0.8
0.6
0.4
0.2
0
SIMSAND
SANISAND
HYPOSAND
MCCSAND

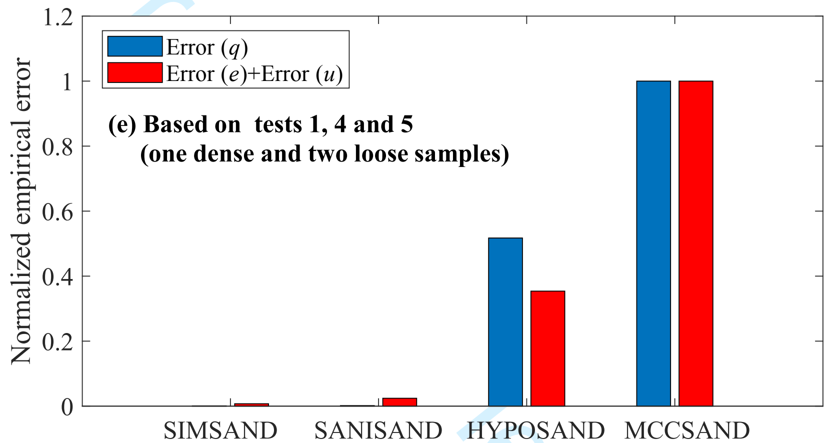

Error (q)
Error (e)+Error (u)
(e) Based on tests 1, 4 and 5
(one dense and two loose samples)
Normalized empirical error
1.2
1
0.8
0.6
0.4
0.2
0
SIMSAND
SANISAND
HYPOSAND
MCCSAND

Figure 5

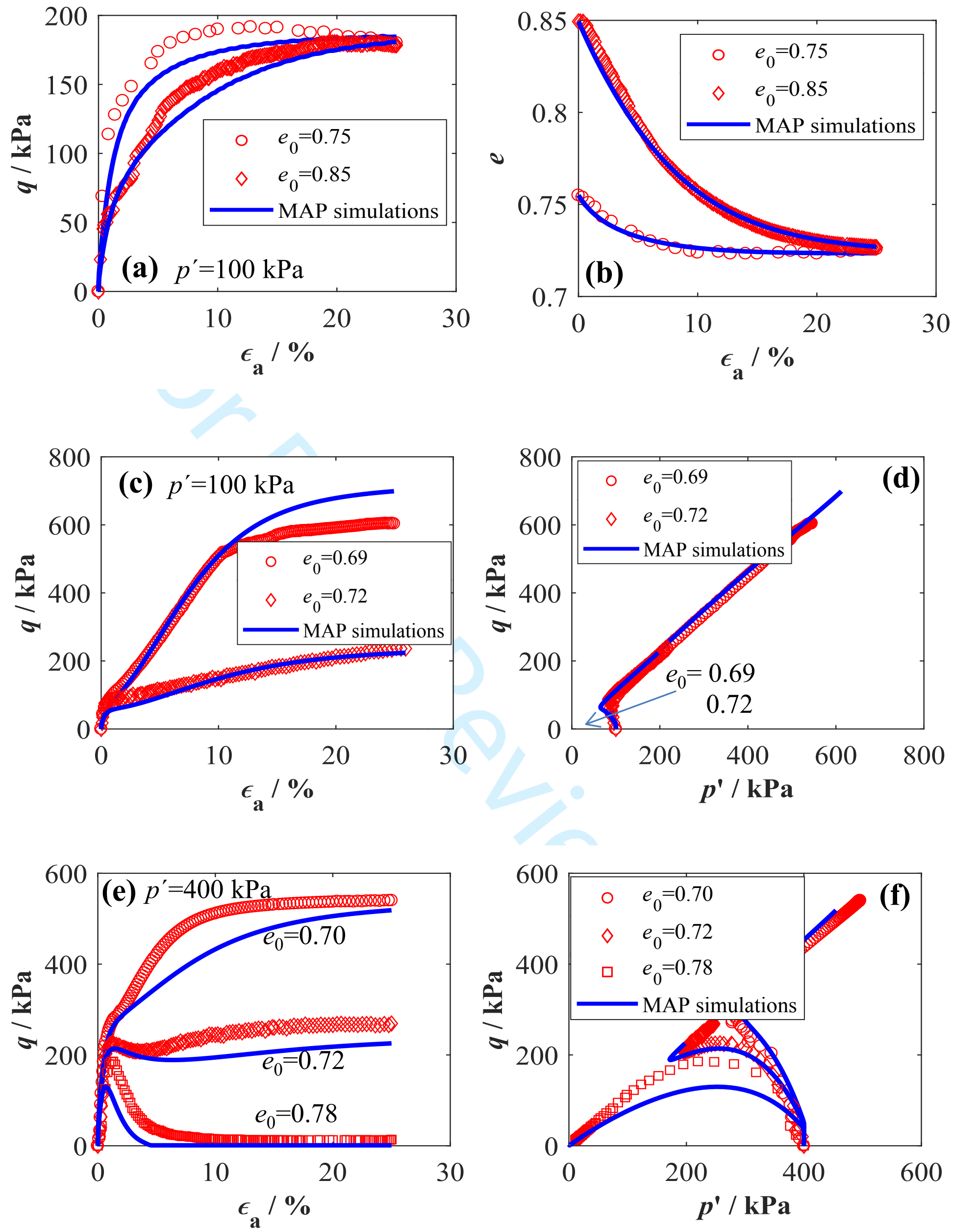

(a) p′=100 kPa
e0=0.75
e0=0.85
MAP simulations
q / kPa
ϵa / %
(b)
e
(c) p′=100 kPa
e0=0.69
e0=0.72
(d)
e0= 0.69
0.72
p' / kPa
(e) p′=400 kPa
e0=0.70
e0=0.72
e0=0.78
(f)

Figure 6

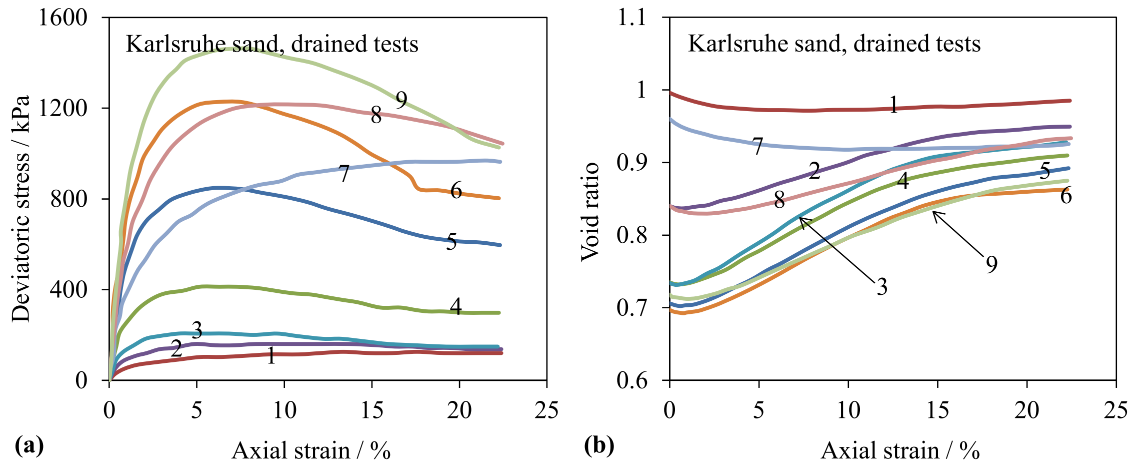

Karlsruhe sand, drained tests
Deviatoric stress / kPa
Axial strain / %
(a)
Karlsruhe sand, drained tests
Void ratio
Axial strain / %
(b)

Figure 7

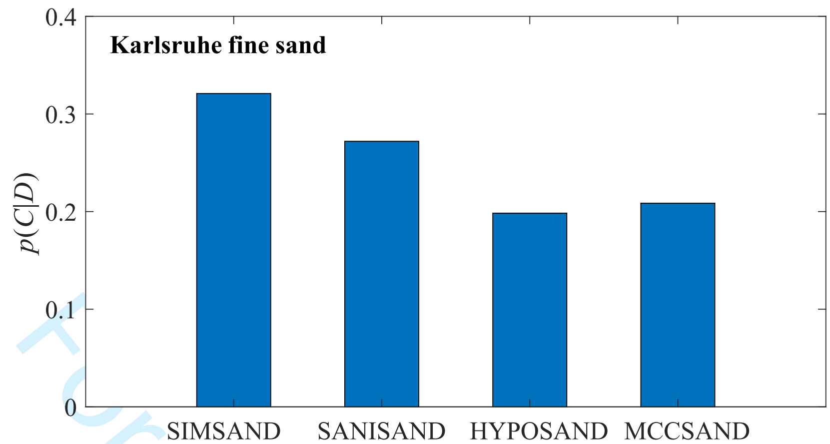
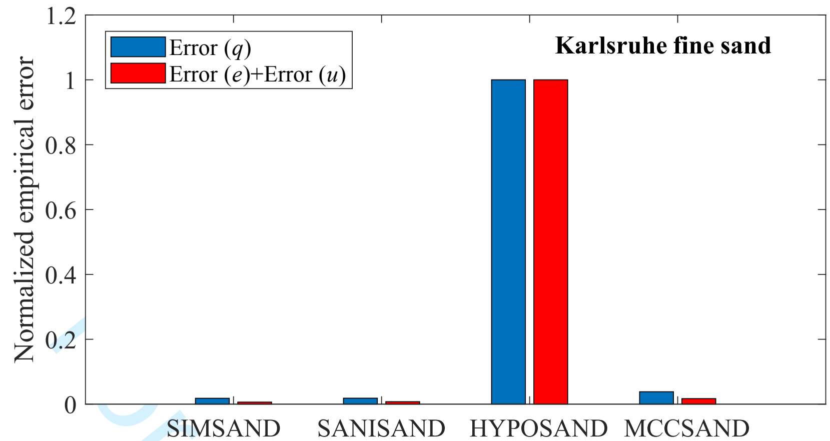
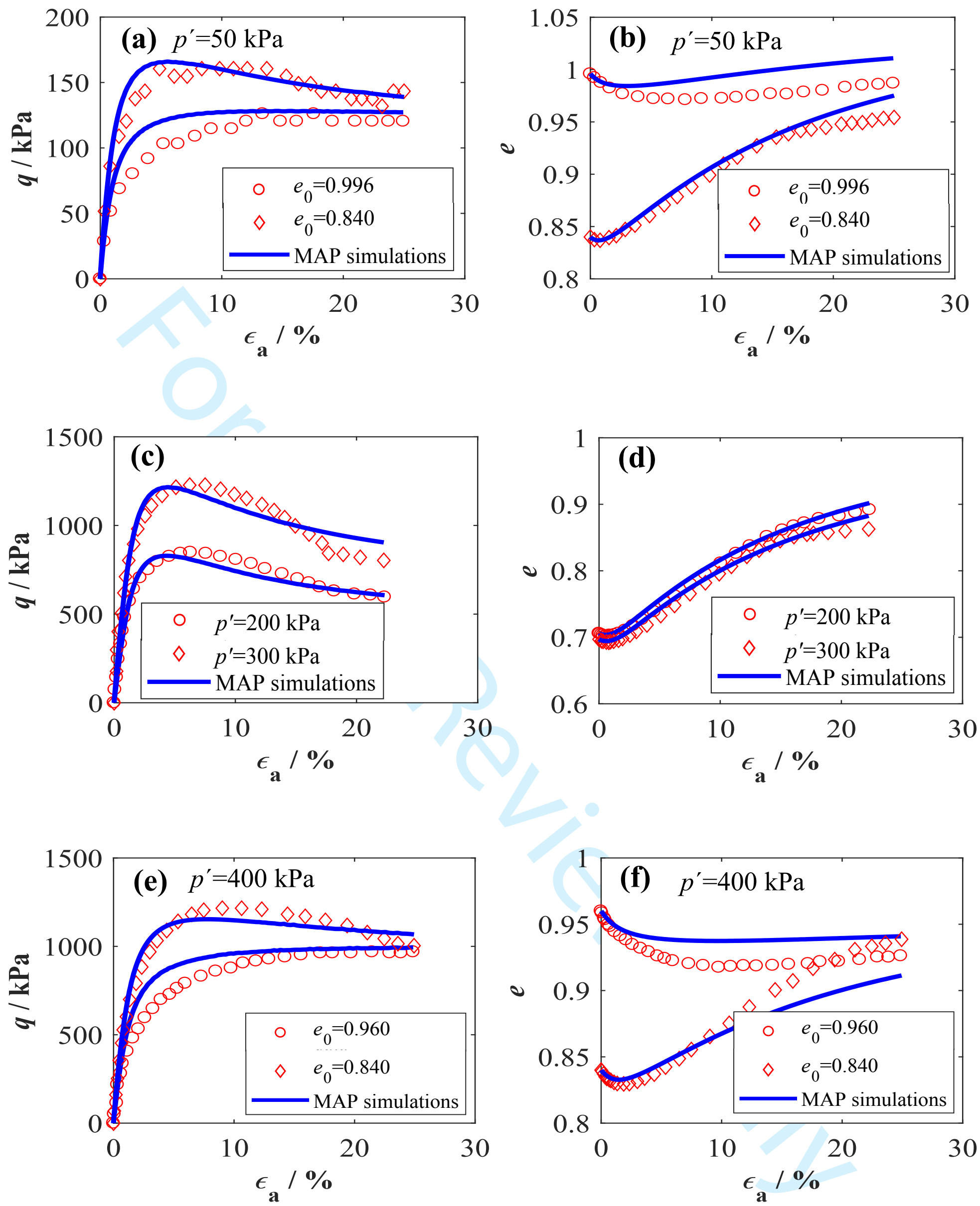

Karlsruhe fine sand
$p(C|D)$
0.4
0.3
0.2
0.1
0
SIMSAND
SANISAND
HYPOSAND
MCCSAND

Figure 8



1.2
1
0.8
0.6
0.4
0.2
0
Normalized empirical error
Error (q)
Error (e)+Error (u)
Karlsruhe fine sand
SIMSAND
SANISAND
HYPOSAND
MCCSAND

Figure 9


(a) p′=50 kPa
q / kPa
ϵa / %
e0=0.996
e0=0.840
MAP simulations
(b) p′=50 kPa
e
(c)
p′=200 kPa
p′=300 kPa
(d)
(e) p′=400 kPa
e0=0.960
(f) p′=400 kPa