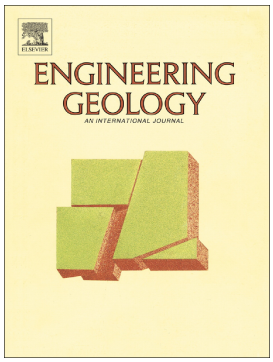


Model selection in geological and geotechnical engineering in the face of uncertainty - Does a complex model always outperform a simple model?

C. Hsein Juang, Wenping Gong, James R. Martin, Qiushi Chen

PII: S0013-7952(17)30184-9
DOI: doi:10.1016/j.enggeo.2018.05.022
Reference: ENGEO 4854

To appear in: *Engineering Geology*

Received date: 3 February 2017
Revised date: 18 May 2018
Accepted date: 23 May 2018

Please cite this article as: C. Hsein Juang, Wenping Gong, James R. Martin, Qiushi Chen , Model selection in geological and geotechnical engineering in the face of uncertainty - Does a complex model always outperform a simple model?. Engineering Geology (2017), doi:10.1016/j.enggeo.2018.05.022

Model Selection in Geological and Geotechnical Engineering in the Face of Uncertainty – Does A Complex Model Always Outperform A Simple Model?

C. Hsein Juang[1,4], Wenping Gong[2,3,*], James R. Martin, II[3], Qiushi Chen[3]

[1]MOE Key Laboratory of Road and Traffic Engineering, College of Transportation Engineering, Tongji University, Shanghai 201804, China

[2]Faculty of Engineering, China University of Geosciences (Wuhan), Wuhan, Hubei 430074, China.

[3]Glenn Department of Civil Engineering, Clemson University, Clemson, SC 29634, USA.

[4]Department of Civil Engineering, National Central University, Jhongli 32001, Taiwan.

[*]Corresponding author: Wenping Gong (wenpinggong@cug.edu.cn)

Abstract

In order to achieve higher fidelity (i.e., higher accuracy) in the model predictions, the solution models employed in geological and/or geotechnical engineering are becoming complex and sophisticated. However, simple and robust models are preferred by engineers in practice. Thus, a dilemma exists between the choice of a complex model for fidelity and that of a simple model for high robustness (i.e., lower variation in the discrepancy between model prediction and observation). This issue becomes more profound when model parameters exhibit uncertainty, which is quite common in geological and geotechnical problems. In this paper, we examine the issue of model selection in the face of uncertainty with three problems: the selection of the order of polynomial fit (i.e., lower order vs. higher order) in the development of data-driven empirical model, the selection of the level of sophistication (i.e., random variable vs. random field) in the probabilistic characterization of the detrended soil property, and the selection of the level of complexity (i.e., simple vs. complex) of the soil constitutive model in numerical modelling. The results illustrate that although the complex and sophisticated models could yield predictions that are more accurate, the simple models might yield predictions that are more robust. This paper provides an insight regarding the question, "Does a complex model always outperform a simple model?"

Keywords: model selection; uncertainty; polynomial fitting; site characterization; drilled shaft; braced excavation.

1. Introduction

In that most models adopted in geological and geotechnical engineering are abstractions of the real world, discrepancies between model predictions and field observations are generally a norm, rather than an exception (Ang and Tang 1984; Cheung and Tang 2005). With advances in mathematics and computational power, researchers could be afforded to derive complex models, which offer an opportunity to improve the fidelity in the model prediction; as such, the models in geological and geotechnical engineering are becoming more complex and sophisticated. On the other hand, simple and robust models are preferred by engineers in practice, as the complex and sophisticated models are more difficult to apply and the model predictions might be less robust. Thus, a dilemma exists between the choice of a complex model for fidelity and that of a simple model for usability and robustness.

The issue of model fidelity and robustness is examined herein at a deep level. Generally, the residual (i.e., the discrepancy between model prediction and field observation) of a complex model is much lower compared to that of a simple model because of its higher fidelity. However, the parameters of the complex model are more difficult to characterize accurately and precisely, as a larger number of model parameters are involved in this complex model. This issue would be especially profound in the situation where only limited test data are available. Further, test errors (i.e., discrepancies between measurements and in situ performances) always exist. The limited availability of test data and the randomness of test errors can lead to difficulty in characterizing model parameters, and as such, the characterized model parameters are often uncertain. Thus, it is more appropriate to characterize the model parameters in a probabilistic manner, and the uncertainty in model parameters must be explicitly considered in the selection of solution models (Chiu et al. 2012; Zhang et al. 2014; Ching and Wang 2016). Note that at a given (or same) level

of variation (or uncertainty) in the model parameters, the variability in the model prediction obtained from the complex model tends to be larger than that obtained from the simple model. For example, an overfitted model could be more sensitive to the measurement and modeling error (Yuen and Mu 2015). In other words, the complex model is generally less robust against (or more sensitive to) the uncertainty in the model parameters. Moreover, given the same amount of data, the characterized uncertainty in the model parameters of the complex model might be larger than that of the simple model, as the increase in the number of model parameters could lead to increased difficulty in the model parameters characterization. Thus, the predictive ability of the complex model is strongly degraded by the uncertainty in the model parameters.

It is worth noting that model selection in geological and geotechnical engineering is often a significant challenge, which is especially profound in the face of uncertainty. For example, in a probabilistic characterization of a soil property at a site, the results could be strongly influenced by the selection of the trend functions and autocorrelation functions (e.g., Cao and Wang 2014; Ching and Wang 2016; Ching et al. 2017; Ching and Phoon 2017). Nevertheless, the issue of model selection has not been adequately examined in most studies reported in the literature. This paper presents a study on the issue of model selection from the perspective of model fidelity and model robustness. To provide a basis for discussion of this issue, three representative problems are studied. The first problem considers the polynomial fitting of a set of test data, this type of problem is usually encountered in the development of data-driven empirical models. The second problem considers the probabilistic characterization of the detrended soil property at an assumed site, in which the residual between the spatial trend of the soil property and the in situ property is modelled either by a random variable or a random field (Phoon and Kulhawy 1999; Ching et al. 2016a). The last problem considers the selection of soil constitutive models in the numerical

modelling of a braced excavation, as many soil constitutive models have been developed with different levels of complexity and capability (Stallebrass and Taylor 1997; Oliveira and Lemos 2014; Goldstein et al. 2016). Through analysis of these three problems, a qualitative insight can be gained regarding the question whether a complex model always outperforms a simple model from the perspective of model fidelity and model robustness.

2. Problem 1: Choice of Low versus High Order Polynomial Fit

Despite advances in numerical methods and testing techniques during past decades, data-driven empirical models are widely adopted in engineering practice (Chiu et al. 2012; Zhang and Goh 2013; Yuen and Mu 2015), especially in geological and geotechnical engineering (Juang et al. 2003; Rezania et al. 2010; Chiu et al. 2012; Kuok et al. 2015; Khoshnevisan et al. 2015; Yuen et al. 2016) and landslide problems (Ren et al. 2015; Huang et al. 2017; Ma et al. 2017; Miao et al. 2018). These data-driven empirical models can be either parametric or non-parametric models. Regardless, the procedures to develop these models are fundamentally the same: 1) a database of case histories is established; 2) a solution model, either parametric or non-parametric function, based upon the relationship between the inputs and the output is derived; 3) the derived model is validated; and 4) the model is further updated when additional case history data are available.

Although the derivation and the selection of the data-driven empirical models have been extensively studied in the literature (Chiu et al. 2012; Zhang et al. 2014; Yuen and Mu 2015; Ching and Wang 2016), a direct comparison between the simple and complex models from the perspective of model predictive ability, based upon discrepancies between model predictions and true model values, has been rarely reported. In the first problem examined, the influence of the level of model complexity on the model predictive ability is studied through a direct comparison

between the simple and complex models; in which, the model parameters are determined with the Bayesian inference method (Beck and Au 2002; Cao and Wang 2012; Gong et al. 2017), and the model predictive ability is evaluated based upon the discrepancies between model predictions and true model values.

A set of synthetic test data are obtained with the assumptions of the true model and test error. With these synthetic data, a set of data-driven empirical models could readily be derived through polynomial curve fitting. Then, the performance of the derived models (i.e., polynomial fits) is evaluated through comparing their predictions with the true model values. Of particular interest is the influence of the order of the polynomial fit, which indicates the complexity of the model, on the model predictive ability.

Here, a synthetic test data, denoted as $d(x)$, is generated with the following formulation.

$$d(x) = g(x) + \varepsilon, \qquad x \in [0, 10.0] \qquad (1)$$

where x represents the position, $x \in [0, 10.0]$; ε represents the test error, which is assumed to be a normal variable with mean of $\mu_\varepsilon = 0.0$ and standard deviation (Std. Dev.) of $\sigma_\varepsilon = 0.3$; and, $g(x)$ represents the true model value at position x, which is assumed to follow the equation below.

$$g(x) = \sqrt[2]{x} + \sqrt[3]{x}, \qquad x \in [0, 10.0] \qquad (2)$$

Meanwhile, it is assumed that the test positions are equally spaced in the range of [0, 10.0]. For example, given a specified number of tests, denoted as n_d, the test positions are determined: 0, $\frac{10}{n_d - 1}$, $\frac{20}{n_d - 1}$, ···,10.0. Then, a corresponding set of synthetic test data, denoted as $\mathbf{d} = [d_1, d_2, \cdots, d_{nd}]$, could be obtained with Eq. (1). It is noted that the correlation among the test errors, denoted as $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{nd}]$, is not considered in this study for simplicity. Shown in Figure 1 are a set of synthetic test data obtained with $n_d = 10$, denoted as $\mathbf{d}_1$.

With the obtained synthetic test data $\mathbf{d} = [d_1, d_2, \cdots, d_{nd}]$, the polynomial fit, denoted as $f(x)$, is then adopted for deriving the data-driven empirical models.

$$f(x) = \sum_{i=0}^{i=n_p} p_i x^i, \qquad x \in [0, 10.0] \qquad (3)$$

where $\mathbf{p} = [p_1, p_2, \cdots, p_{np}]$ represent the model parameters, in which n_p represents the order of the polynomial fit.

In the absence of the knowledge of the true model $g(x)$ and the underlying physics, the selection of the order of the polynomial fit n_p can be a significant challenge (Ting 1987; Ooi and Ramsey 2003). In this problem, the following orders of the polynomial fit are studied, n_p = 1, 2, 3, 4, 5, and 6; and, the corresponding models are denoted as Models P1, P2, P3, P4, P5, and P6, respectively. Here, the model parameters $\mathbf{p} = [p_1, p_2, \cdots, p_{np}]$ are estimated with the synthetic test data **d** utilizing the least square regression analysis. For example, Models P1, P2, P3, P4, P5, and P6 that are derived with test data $\mathbf{d}_1$ are also plotted in Figure 1, and the formulation of these empirical models is tabulated in Table 1.

As can be seen in Figure 1, discrepancies exist between the predictions obtained from the empirical models and test data, and the discrepancy between model prediction and test data tends to decrease with the order of the polynomial fit. For example, a larger discrepancy is observed between prediction from Model P1 (or Model P2) and test data $\mathbf{d}_1$; whereas, the prediction from Model P6 (or Model P5) matches test data $\mathbf{d}_1$ well. As can be seen in Table 1, the fidelity of the polynomial model (in matching the test data $\mathbf{d}_1$) increases with the order of the polynomial fit, indicated by the decrease of Std. Dev. of $[f(x) - d(x)]$. Thus, a more reliable empirical model may be achieved by adopting a higher-order polynomial fit. However, this viewpoint is not always correct, as will be evidenced by the results presented in the following paragraphs.

Now, let's consider the randomness of the test errors and the discrepancy between model prediction $f(x)$ and test data $\mathbf{d}$ in the model parameters characterization and the subsequent model prediction. For illustration purposes, model parameters $\mathbf{p} = [p_1, p_2, \cdots, p_{np}]$ are assumed to follow a multivariate normal distribution; and, the statistics of which, in terms of mean and covariance, denoted as $\boldsymbol{\mu}_\mathbf{p}$ and $\mathbf{C}_\mathbf{p}$, respectively, are determined with the Markov Chain Monte Carlo (MCMC) simulation-based Bayesian inference method (Beck and Au 2002; Zhang et al. 2009; Cao and Wang 2012; Juang et al. 2013; Gong et al. 2017), in which the Metropolis-Hastings algorithm (Metropolis et al. 1953; Hastings 1970) is employed. The procedure for the MCMC simulation-based Bayesian inference method using Metropolis-Hastings algorithm is given in Appendix A.

The likelihood function of the synthetic test data $\mathbf{d} = [d_1, d_2, \cdots, d_{nd}]$, denoted as $L(\mathbf{d}|\boldsymbol{\theta})$, is an essential element within the MCMC simulation-based Bayesian inference method, which can be constructed as follows.

$$L(\mathbf{d}\,|\,\boldsymbol{\theta}) = \prod l(d_i\,|\,\boldsymbol{\theta}) \qquad (4)$$

where $l(d_i|\boldsymbol{\theta})$ represents the likelihood of observing test data d_i at position x_i given the unknown parameters $\boldsymbol{\theta}$. In this problem, the unknown parameters $\boldsymbol{\theta}$ are referred to model parameters $\mathbf{p} = [p_1, p_2, \cdots, p_{np}]$, expressed as $\boldsymbol{\theta} = \mathbf{p}$. In that test error ε follows a normal distribution with mean of μ_ε and Std. Dev. of σ_ε, the likelihood function $l(d_i|\boldsymbol{\theta})$ is readily formulated as:

$$l(d_i\,|\,\boldsymbol{\theta}) = \phi\left(\frac{d_i - \left[f(x_i) + \mu_\varepsilon\right]}{\sigma_\varepsilon}\right) \qquad (5)$$

where $\phi(\cdot)$ represents the probability density function (PDF) of the standard normal variable.

In this problem, the prior distribution of the unknown parameters $\boldsymbol{\theta}$ is assumed to follow a multivariate normal distribution: the mean values are taken as the model parameters obtained

with the least square regression analysis (see Table 1), the coefficients of variation (COVs) are taken as 0.10, and the correlation coefficients among the unknown parameters are taken as 0. The lengths of the burn-in samples and the Markov chain, in the MCMC simulation-based Bayesian inference method, are taken as 5,000 and 100,000, respectively. Through which, the converged statistics of the unknown parameters $\boldsymbol{\theta}$, in terms of mean $\boldsymbol{\mu}_{\boldsymbol{\theta}}$ (i.e., $\boldsymbol{\mu}_{\mathbf{p}}$) and covariance $\mathbf{C}_{\boldsymbol{\theta}}$ (i.e., $\mathbf{C}_{\mathbf{p}}$), can be derived. With this parameters setting, the statistics of the model parameters $\mathbf{p} = [p_1, p_2, \cdots, p_{np}]$, in this problem, are readily obtained, as shown in Table 2.

With the obtained statistics of the model parameters $\mathbf{p} = [p_1, p_2, \cdots, p_{np}]$ (see Table 2), the model prediction at a given position can be predicted probabilistically. To evaluate the predictive ability of the data-driven empirical models (i.e., Models P1 to P6), the predictions obtained from the empirical models (i.e., Eq. 3) and the true model values obtained from the true model (i.e., Eq. 2) are compared. Here, the position is assumed to be uniformly distributed on [0, 10]. With the aid of 100,000 Monte Carlo simulations (MCSs), the distribution of the discrepancy between the model prediction and the true model value, in terms of $[g(x) - f(x)]$, for each of these empirical models can be obtained; and, the results are plotted in Figure 2.

It can be seen from Figure 2 that the predictive ability of the empirical model is strongly affected by the order of the polynomial fit. The variation in the model prediction, indicated by the width of the distribution of the discrepancy $[g(x) - f(x)]$, decreases first and then increases with the order of the polynomial fit. Thus, the robustness of the model prediction, which might be measured by the variation in the discrepancy $[g(x) - f(x)]$ and a lower variation signals higher robustness, increases first and then decreases with the order of the polynomial fit. The mean and Std. Dev. of the discrepancy $[g(x) - f(x)]$ of the six empirical models (i.e., Models P1 to P6) are calculated and illustrated in Figure 3(b) and Figure 3(c), respectively. In Figure 3(b), the mean of

the discrepancy is slightly influenced by the order of the polynomial fit. In Figure 3(c), the Std. Dev. of the discrepancy decreases first and then increases with the order of the polynomial fit.

Next, two other sets of synthetic test data, denoted as $\mathbf{d}_2$ and $\mathbf{d}_3$, are obtained with Eq. (1) and shown in Figure 3(a): $\mathbf{d}_2$ and $\mathbf{d}_3$ are generated with the assumptions of $\sigma_\varepsilon = 0.4$ (and $\mu_\varepsilon = 0.0$) and $\sigma_\varepsilon = 0.5$ (and $\mu_\varepsilon = 0.0$), respectively. With the test data $\mathbf{d}_2$ and $\mathbf{d}_3$ shown in Figure 3(a), the statistics of the model parameters of the six empirical models (i.e., Models P1 to P6) can be derived with the aforementioned MCMC simulation-based Bayesian inference method. Then, the statistics of the discrepancy between the model prediction and the true model value are obtained; the results are plotted in Figure 3(b) and Figure 3(c). In Figure 3(b), the mean of the discrepancy decreases slightly with the order of the polynomial fit with respect to test data $\mathbf{d}_2$ and $\mathbf{d}_3$; whereas, the influence of the order of the polynomial fit on the mean of the discrepancy is not significant with respect to test data $\mathbf{d}_1$. In Figure 3(c), the variation in the discrepancy decreases first and then increases with the order of the polynomial fit. Thus, the results are slightly affected by the synthetic test data or the statistics of the test error: the robustness of the model prediction tends to increase first and then decreases with the order of the polynomial fit.

It should be noted that the selection of the most suitable order of the polynomial fit in the development of the data-driven empirical model is most likely problem-specific. In the problem studied, the following assumptions have been adopted: 1) the true model follows the form of Eq. (2); 2) the test errors are normally distributed and uncorrelated; and 3) the statistics of the test errors are known in the calibration of the model parameters. The results obtained in this problem provide an insight into the choice of the order of the polynomial fit in the development of a data-driven empirical model. Although a higher fidelity (i.e., in matching test data) could be achieved with a higher-order polynomial fit, the robustness of the model prediction might degrade in the

face of uncertainty in the model parameters; and, a higher-order polynomial fit tends to yield a larger variation in the discrepancy between the model prediction and the true model value. In other words, a complex and sophisticated model is not necessarily a "better" model than a simple model, and the latter may outperform the former in the face of uncertainty (from the perspective of model robustness). Indeed, model fidelity and model robustness, the two objectives in a model selection, usually conflict with each other (Gong et al. 2016), and the optimal model could be the one that balances the objective of higher robustness (in the model prediction) and that of higher fidelity (in matching the test data).

3. Problem 2: Choice of Random Variable versus Random Field

Geomaterials are a product of natural deposits, not artificial materials. The property of the *in situ* geomaterials is dependent upon the depositional history and the stress history. In the absence of the knowledge of depositional history and stress history, the geotechnical property at a site cannot be known prior to the site investigation. Furthermore, only a limited amount of site-specific test data can be obtained from the site investigation especially due to budget constraints. Thus, the geotechnical property is only known at limited locations at a given site; whereas, the geotechnical property at other positions cannot be known and has to be estimated from that obtained from the site investigation. In light of the inherent spatial variability of the geotechnical property (e.g., the soil property at a site tends to vary spatially) and the limited availability of the site investigation data, the geotechnical property at a site oftentimes has to be characterized in a probabilistic manner; and, the topic of probabilistic site characterization has gained a significant interest during the last three decades (Degroot and Baecher 1993; Christian et al. 1994; Fenton 1999a&b; Phoon and Kulhawy 1999; Phoon et al. 2003; Uzielli et al. 2005; Cao and Wang 2012;

Cassidy et al. 2013; Gong et al. 2014; Ching et al. 2015; Wang and Zhao 2016; Bong and Stuedlein 2017&2018).

The studies on the probabilistic site characterization can be categorized into the following groups: 1) the identification of soil layers in a profile at a given location (Phoon et al. 2003; Cao and Wang 2012; Huang et al. 2014; Ching et al. 2015); 2) the characterization of stratigraphic uncertainty from a limited number of boreholes (Li et al. 2016; Wang et al. 2017; Xiao et al. 2017; Chen et al. 2018); 3) the random field or random variable characterization of soil property in a homogeneous layer (Ching and Wang 2016; Ching and Phoon 2017; Ching et al. 2017; Gong et al. 2017); and 4) the influence of the statistical uncertainty in the site characterization on the analysis and design of the interested geotechnical system (Ching et al. 2016a&2016b; Ching et al. 2017; Gong et al. 2014&2018). Although the topic of the probabilistic site characterization has been frequently reported, the focus is usually on the development and the application of the characterization methods; whereas, the direct comparison among the characterization methods has been rarely reported. For example, the performance of the characterization method could be influenced by the level of model complexity. Thus, a comparison study between the random field and random variable characterizations of the soil property in a homogeneous layer is carried out in the second problem. The Bayesian inference method is adopted to calibrate the statistical information of the random variable and that of the random field; as an example, the performance of these two methods is evaluated through the task of the bearing capacity prediction in a drilled shaft.

Here, a set of synthetic test data of the undrained shear strength are obtained from the assumptions of an in situ soil profile, as shown in Figure 4, and the test errors. Then, the spatial trend of the synthetic test data is fitted with a polynomial fit, and the residual between the spatial

trend and the in situ property is characterized by a random variable or a random field. With the probabilistic characterization of the soil property, the performance of the geotechnical system of concern could be evaluated. For example, the bearing capacity of a drilled shaft in this soil, as shown in Figure 5, is evaluated in this problem; then, the predicted bearing capacity is compared with the benchmark bearing capacity of the drilled shaft that is evaluated with the true in situ soil property. The influence of the probabilistic characterization of the residual (i.e., the discrepancy between the spatial trend and the in situ soil property), in terms of the model selection between a random variable model and a random field model, on the prediction of the bearing capacity of the drilled shaft is studied.

In that the undrained shear strength often increases with the vertical effective stress (or simply depth), the focus of this problem is on the characterization of the normalized undrained shear strength, denoted as $c_u^{\mathrm{n}}(z) = c_u(z)/z$, where $c_u(z)$ is the undrained shear strength at depth z. It is noted that the influence of the pore water pressure is not included in this problem, the normalization of the shear strength with the depth is considered equivalent to that with the vertical effective stress, and the normalization of the shear strength with the depth is employed for illustration purposes; alternatively, the un-normalized shear strength could be characterized directly. In reference to the in situ property of the normalized undrained shear strength shown in Figure 4, a synthetic test data of the normalized undrained shear strength at depth z, denoted as $c_u^{\mathrm{n}}(z)$, can be obtained.

$$c_u^{\mathrm{n}}(z) = g(z) + \varepsilon, \qquad z \in [0, 15.0\,\mathrm{m}] \quad (6)$$

where $g(z)$ represents the in situ property of the normalized undrained shear strength at depth z (see soil profile $\mathbf{S}_1$ in Figure 4); and, ε represents the test error, which is simulated by a normal

variable with mean of μ_ε = 0.0 and Std. Dev. of σ_ε = 1.00 kPa/m, and the correlation among the test errors $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{nd}]$ is not considered, in which n_d represents the number of test data.

In an assumed site investigation program, the test depths are equally spaced in the range of [0, 15.0 m]. For a specified number of tests n_d, the test depths are determined: $\frac{1.5}{n_d}$ m, $\frac{3.0}{n_d}$ m, ⋯, and 15.0 m. Then, a corresponding set of synthetic test data of the normalized undrained shear strength, denoted as $\mathbf{d} = [\, c_u^{\mathrm{n}}(z_1),\ c_u^{\mathrm{n}}(z_2),\ \cdots,\ c_u^{\mathrm{n}}(z_{nd})\,]$, can be obtained with Eq. (6). For example, Figure 4 shows a set of synthetic test data of the normalized undrained shear strength obtained with n_d = 10, denoted as $\mathbf{d}_1$.

With the obtained synthetic test data of the normalized undrained shear strength $\mathbf{d} = [\, c_u^{\mathrm{n}}(z_1),\ c_u^{\mathrm{n}}(z_2),\ \cdots,\ c_u^{\mathrm{n}}(z_{nd})\,]$, the in situ property of the normalized undrained shear strength can be characterized as follows.

$$f(z) = \sum_{i=0}^{i=n_p} p_i z^i + \omega(z), \qquad z \in [0, 15.0\,\mathrm{m}] \qquad (7)$$

where $\mathbf{p} = [p_1, p_2, \cdots, p_{np}]$ represent the model parameters, in which n_p represents the order of the polynomial fit of the spatial trend of the normalized undrained shear strength $c_u^{\mathrm{n}}(z)$, in terms of $\sum_{i=0}^{i=n_p} p_i z^i$; and, $\omega(z)$ represents the residual between the spatial trend and the in situ property.

In this problem, the spatial trend of the normalized undrained shear strength is captured by a 1st order polynomial fit (n_p = 1; denoted as P1) and a 2nd order polynomial fit (n_p = 2; denoted as P2), and the residual between the spatial trend and the in situ property is characterized by a normal variable and a stationary Gaussian random field, denoted as V and F, respectively. Thus, four soil models are studied in this problem, namely (P1+V), (P2+V), (P1+F), and (P2+F).

For simplicity, the model parameters $\mathbf{p} = [p_1, p_2, \cdots, p_{np}]$, in this problem, are treated as constant values and determined with the least square regression analysis, the resulting model P1 and P2 (i.e., spatial trend) are illustrated in Figure 4; and, the statistics of the residual ω, in terms of the statistics of the random variable (i.e., mean μ_ω and Std. Dev. σ_ω) and those of the random field (i.e., mean μ_ω, Std. Dev. σ_ω and scale of fluctuation λ_ω) are determined using the aforementioned MCMC simulation-based Bayesian inference method. In a strict manner, the trend functions and associated model parameters should be treated as uncertain and characterized probabilistically (Ching and Wang 2016; Ching and Phoon 2017; Ching et al. 2017). Nonetheless, the purpose of this problem is to compare these two modeling approaches (i.e., random field versus random variable) in the site characterization. Thus, the spatial trends, in this problem, may be adequately captured by these two polynomial functions (i.e., P1 and P2), and the related model parameters could be treated as constant values.

In that the likelihood function of the synthetic test data $\mathbf{d} = [\, c_u^{\mathrm{n}}(z_1),\, c_u^{\mathrm{n}}(z_2),\, \cdots,\, c_u^{\mathrm{n}}(z_{nd})\,]$ can be an essential element within the MCMC simulation-based Bayesian inference method, the detailed formulation of this likelihood function is presented below. Note that when the residual ω between the spatial trend and the in situ property is characterized by a normal variable (V), the unknown parameters $\boldsymbol{\theta}$ are referred to the statistics of this random variable, expressed as $\boldsymbol{\theta} = [\mu_\omega, \sigma_\omega]$; and then, the likelihood function of the synthetic test data $\mathbf{d} = [\, c_u^{\mathrm{n}}(z_1),\, c_u^{\mathrm{n}}(z_2),\, \cdots,\, c_u^{\mathrm{n}}(z_{nd})\,]$, denoted as $L(\mathbf{d}|\boldsymbol{\theta})$, can be formulated as follows.

$$L(\mathbf{d}\,|\,\boldsymbol{\theta}) = \prod l(d_i\,|\,\boldsymbol{\theta}) \qquad (8)$$

where $l(d_i|\boldsymbol{\theta})$ represents the likelihood of observing test data $d_i = c_u^{\mathrm{n}}(z_i)$ at depth z_i given the unknown parameters $\boldsymbol{\theta}$, which may be calculated as:

$$l(d_i \mid \boldsymbol{\theta}) = l[c_u^{\mathrm{n}}(z_i) \mid \boldsymbol{\theta}] = \phi\left(\frac{c_u^{\mathrm{n}}(z_i) - \left[f(z_i) + \mu_\omega + \mu_\varepsilon\right]}{\sqrt{\sigma_\omega{}^2 + \sigma_\varepsilon{}^2}}\right) \qquad (9)$$

where $\phi(\cdot)$ represents the PDF of the standard normal variable. When the residual ω between the spatial trend and the in situ property is characterized by a stationary Gaussian random field, the unknown parameters $\boldsymbol{\theta}$ are referred to the statistics of this random field, expressed as $\boldsymbol{\theta} = [\mu_\omega, \sigma_\omega, \lambda_\omega]$; and, the likelihood function of the synthetic test data $\mathbf{d} = [c_u^{\mathrm{n}}(z_1), c_u^{\mathrm{n}}(z_2), \cdots, c_u^{\mathrm{n}}(z_{nd})]$ must be formulated in a different manner (Gong et al. 2014&2017):

$$L(\mathbf{d} \mid \boldsymbol{\theta}) = \frac{1}{(2\pi)^{n_d/2}} \frac{1}{\left[\det(\mathbf{C})\right]^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{d}-\boldsymbol{\mu})\mathbf{C}^{-1}(\mathbf{d}-\boldsymbol{\mu})^{\mathrm{T}}\right] \qquad (10)$$

where $\boldsymbol{\mu} = [f(z_1) + \mu_\omega + \mu_\varepsilon, f(z_2) + \mu_\omega + \mu_\varepsilon, \cdots, f(z_{nd}) + \mu_\omega + \mu_\varepsilon]$ represent the mean of the synthetic test data $\mathbf{d} = [c_u^{\mathrm{n}}(z_1), c_u^{\mathrm{n}}(z_2), \cdots, c_u^{\mathrm{n}}(z_{nd})]$, and $\mathbf{C}$ represents the covariance among the synthetic test data $\mathbf{d}$ with the element $\boldsymbol{C}_{i,j}$ ($i, j = 1, 2, \cdots, n_d$):

$$C_{i,j} = \mathrm{Cov}\left[d_i, d_j\right] = \mathrm{Cov}[c_u^{\mathrm{n}}(z_i), c_u^{\mathrm{n}}(z_j)] = \begin{cases} \sigma_\omega{}^2 + \sigma_\varepsilon{}^2, & i = j \\ \rho_{\omega i,\omega j} \cdot \sigma_\omega{}^2, & i \neq j \end{cases} \qquad (11)$$

where $\rho_{\omega i,\omega j}$ represents correlation coefficient between the residual at depth z_i and that at depth z_j. The exponential autocorrelation structure, which is widely adopted in the literature (Cho 2007; Jiang et al. 2014), is employed in this study and $\rho_{\omega i,\omega j}$ is computed as follows.

$$\rho_{\omega i,\omega j} = \exp\left(-\frac{2\left|z_i - z_j\right|}{\lambda_\omega}\right) \qquad (12)$$

where $|z_i - z_j|$ represents the absolute distance between the depth z_i and the depth z_j.

With the aforementioned likelihood functions of the synthetic test data $\mathbf{d}$, the statistics of the residual (between the spatial trend and the in situ property) can readily be obtained using the

aforementioned MCMC simulation-based Bayesian inference method. In this problem, the prior distribution of the unknown parameters $\boldsymbol{\theta}$ is assumed to follow a multivariate normal distribution: 1) the mean of μ_ω is taken as the mean of the residuals between the spatial trend (obtained with the least square regression analysis; see Table 3) and the synthetic test data **d**, and the COV of μ_ω is taken as 0.10; 2) the mean of σ_ω is taken as the Std. Dev. of the residuals between the spatial trend (obtained with the least square regression analysis; see Table 3) and the synthetic test data **d**, and the COV of σ_ω is taken as 0.20; and 3) the mean and COV of λ_ω are taken as 3.0 m and 0.30, respectively. The lengths of the burn-in samples and the Markov chain are taken as 5,000 and 100,000, respectively. With this parameters setting, the posterior distribution of the statistical parameters $\boldsymbol{\theta}$ could be obtained, the resulting soil models, including (P1+V), (P2+V), (P1+F) and (P2+F), are tabulated in Table 3. The resulting soil models are readily adopted for evaluating the ultimate bearing capacity of the drilled shaft. Note that the term of "ultimate bearing capacity" is referred to herein as the bearing capacity based on the ultimate limit state (ULS), as opposed to the bearing capacity based on the serviceability limit state (SLS). The ultimate bearing capacity, denoted as Q_{ULS}, of the drilled shaft in this soil, as shown in Figure 5, is evaluated as:

$$Q_{\mathrm{ULS}} = \pi B\alpha \int_0^D c_{u1} dz + \frac{\pi}{4} B^2 N_c c_{u2} \qquad (13)$$

where D and B represent the depth and the diameter of the drilled shaft, respectively, and which are taken as 15.0 m and 0.45 m, respectively; c_{u1} and c_{u2} represent the undrained shear strength along the shaft length and that at the shaft base (kPa), respectively; α represents the adhesion coefficient, which is taken as 0.5 (Meyerhof 1976); and, N_c represents the base baring capacity factor, which is taken as 9.0 (Meyerhof 1976). In order to estimate the side resistance, in terms of $\pi B\alpha \int_0^D c_{u1} dz$, the length of this drilled shaft is sub-divided into 300 equal elements or segments.

For illustration purposes, the bearing capacity model shown in Eq. (13) is assumed to be perfect (i.e., the model error is negligible).

With the aid of 100,000 MCSs, the distribution of the predicted bearing capacity of this drilled shaft, denoted as Q_{ULS}^{P}, can be obtained with the characterized soil models shown in Table 3. The benchmark bearing capacity of this drilled shaft, denoted as Q_{ULS}^{T}, is evaluated with the *in situ* undrained shear strength shown in Figure 4, assuming that this drilled shaft is strictly located at the position where the soil profile is taken from. In this scenario, the discrepancy between the predicted bearing capacity and the benchmark capacity could be used to evaluate the predictive ability of the soil model, and a lower discrepancy signals a better soil model. The distribution of the discrepancy between the predicted bearing capacity and the benchmark bearing capacity, in terms of $(Q_{ULS}^{T} - Q_{ULS}^{P})/Q_{ULS}^{T}$, could then be obtained, and the results are shown in Figure 6. The plots in Figure 6 illustrate that the variation in the discrepancy obtained from model (P1+V) or (P2+V) is smaller than that obtained from model (P1+F) or (P2+F). Thus, the random variable (V) model could yield a more robust prediction of the bearing capacity of this drilled shaft, compared to the random field (F) model in this problem. In Table 3, the mean of the discrepancy obtained from model (P1+V) or (P1+F) is closer to 0 than that obtained from model (P2+V) or (P2+F); and, the difference in the mean of the discrepancy is negligible between model (P1+V) and (P1+F), or between model (P2+V) and (P2+F). Among these four soil models studied, model (P2+V) yields the smallest variation in the discrepancy while model (P1+F) yields the largest variation; and, the mean of the discrepancy obtained from model (P1+V) is the closest to 0 while that obtained from model (P2+V) is the farthest from 0. As such, in this site characterization, the spatial trend can be better characterized with the 1st order polynomial (P1), indicated by a smaller distance between the mean of the discrepancy and 0; and, the residual can be better characterized

with the random variable (V), indicated by a smaller variation in the discrepancy. It is noted that a smaller variation in the discrepancy signals higher model robustness. Thus, the complex model (i.e., the spatial trend is captured by a higher order polynomial and the residual is characterized with a random field model) does not necessarily represent a better model considering the model robustness.

Note that the predictive ability of the soil model can be affected by both the mean and Std. Dev. of the discrepancy between the predicted bearing capacity and the benchmark capacity. For ease of comparison, the discrepancy, in terms of $(Q_{ULS}^{T} - Q_{ULS}^{P})/Q_{ULS}^{T}$, is fitted herein by a normal distribution and the statistics of which are calculated with 100,000 MCSs. In a utopia situation (i.e., a perfect site characterization), the discrepancy between the predicted bearing capacity and the benchmark capacity is approaching to zero. Thus, the likelihood of the discrepancy equating to zero, denoted as $\Pr[(Q_{ULS}^{T} - Q_{ULS}^{P})/Q_{ULS}^{T} = 0]$, is estimated for the above soil models, including (P1+V), (P2+V), (P1+F), and (P2+F), and the model yielding the highest likelihood is identified as the "best" model. Given the in situ soil profile $\mathbf{S}_1$ and synthetic test data $\mathbf{d}_1$ shown in Figure 4, model (P1+V) is identified as the "best" model in this problem (see Table 3).

The identified "best" model potentially varies with the *in situ* soil property, the number of test data n_d, the standard deviation of the test error σ_ε, and different simulations of the test errors $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{nd}]$. To address this concern, an extensive series of analyses are undertaken in this section, involving four different *in situ* soil profiles (i.e., soil profile $\mathbf{S}_1$ shown in Figure 4 and soil profiles $\mathbf{S}_2$, $\mathbf{S}_3$, $\mathbf{S}_4$ shown in Figure 7), four different numbers of test data (i.e., n_d = 10, 20, 30, and 40), five different degrees of variation in the test errors (i.e., σ_ε = 0.25 kPa/m, 0.50 kPa/m, 0.75 kPa/m, 1.00 kPa/m, and 1.50 kPa/m), and three different simulations of the test errors (i.e., three MCSs are adopted for generating the test errors $\boldsymbol{\varepsilon}$). Here, a total of 240 sets of synthetic test

data **d** are generated, and then 240 "best" models are identified. The distribution of the resulting "best" model among these four soil models is depicted in Figure 8. The plots in Figure 8 depict that model (P1+V) and (P2+V) yield high probabilities of being identified as the "best" model with respect to in situ soil profiles $\mathbf{S}_1$ and $\mathbf{S}_2$, and model (P2+V) yields high probability of being identified as the "best" model with respect to in situ soil profiles $\mathbf{S}_3$ and $\mathbf{S}_4$; and, the probabilities of model (P1+F) and (P2+F) being identified as the "best" model are relatively low. The spatial trend of the soil property is best characterized with the 1st order polynomial (P1) with respect to in situ soil profile $\mathbf{S}_1$, and the spatial trend is best characterized with the 2nd order polynomial (P2) with respect to in situ soil profiles $\mathbf{S}_3$ and $\mathbf{S}_4$. Thus, the residual between the spatial trend and the in situ property could be better characterized by the simple random variable model; on the other hand, the characterization of the spatial trend could be influenced by the in situ soil profile. The mean and Std. Dev. of the discrepancy $(Q_{ULS}^{T} - Q_{ULS}^{P})/Q_{ULS}^{T}$ of these four soil models are depicted in Figure 9. In Figure 9(a), the difference in the mean of the discrepancy is negligible between model (P1+V) and (P1+F), or between model (P2+V) and (P2+F); and, in Figure 9(b), model (P2+V) yields the smallest variation in the discrepancy, followed by model (P1+V) and (P2+F), and model (P1+F) yields the largest variation in the discrepancy.

The choice of modeling a soil property with random variable versus random field can be a challenging issue. The random variable is simple to use and easy to characterize; whereas, the random field model is theoretically sound but complex and hard to characterize. In this problem, the following assumptions were adopted: 1) the true *in situ* soil profile at the position of concern is known; 2) the test errors are normally distributed and uncorrelated; 3) the statistics of the test errors are known in the site characterization; 4) the drilled shaft is strictly located at the position where the soil profile is taken from so that the benchmark bearing capacity could be evaluated; 5)

the adopted bearing capacity model is assumed to be perfect (i.e., the model error is negligible); and 6) the spatial trends are captured by the two polynomial functions and the associated model parameters are treated as constant values. Though several assumptions were made, the results presented provide an insight into the choice of the random variable model versus the random field model in the probabilistic site characterization. Although the residual between the spatial trend and the in situ soil property can be characterized either by a random variable or a random field, the simple random variable model can lead to a better prediction of the bearing capacity of the drilled shaft in the face of uncertainty (in the model parameters), based upon a comparison with the benchmark bearing capacity. Thus, the complex random field model is not necessarily a "better" model than the simple random variable model; and, the latter may outperform the former in the site characterization from the perspective of the predicted performance of the geotechnical system.

4. Problem 3: Choice of Simple versus Complex Soil Constitutive Model

The stress-strain relationship, or constitutive model, of a soil plays a vital role in the numerical simulation of the performance of a geotechnical structure. To capture the influences of geomaterial compositions, particle-size distributions, water contents and stress-strain histories, various soil constitutive models have been proposed, including for examples, Mohr-Coulomb model (Smith 2005; Robert 2017), Cam-Clay model (Suebsuk et al. 2010; Goldstein et al. 2016), cap-yield model (Hinchberger and Qu 2009), small-strain model (Stallebrass and Taylor 1997; Kung et al. 2009), and anisotropic models (Whittle 1993; Hashash and Whittle 1996; Jung et al. 2004; Oliveira and Lemos 2014). As can be seen, in order to achieve higher fidelity (i.e., higher accuracy) in capturing the stress-strain behaviors, soil constitutive models are becoming complex

and sophisticated. Consequently, a larger number of model parameters are involved in these complex soil constitutive models; and, the calibration of these model parameters is a significant challenge, especially in the situation where the test data are limited. Although the subject of soil constitutive models has been extensively studied, few studies have been carried out and reported on the calibration of model parameters and the optimal selection of soil constitutive models. To this end, in the third problem we investigated the optimal selection of the soil constitutive model in the setting of a braced excavation.

In this problem, a braced excavation case published in Plaxis Tutorial Manual (2002) and FLAC version 7.0 User's Manual (2011), as shown in Figure 10, is studied. The selection of soil constitutive models is a significant challenge in the numerical modelling of a braced excavation problem (Lim et al. 2010). For example, ideally the small-strain behavior (Stallebrass and Taylor 1997; Kung et al. 2009), anisotropic behavior (Whittle 1993; Jung et al. 2004) and confining pressure-dependent behavior should all be included; however, such a comprehensive soil model is not available in a commercial code. Note that although no soil constitutive model is perfect, each and every model has its own merits and shortcomings, which poses a significant challenge to the optimal selection of the constitutive model. Further, the limited availability of site-specific data, randomness of test errors and inherent variability in natural deposits can lead to uncertainty in the characterized soil model parameters. In the face of uncertainty in the model parameters, the challenge in the selection of soil constitutive models becomes more profound. The influence of the selection of soil constitutive models on the evaluated maximum wall deformation (induced by the braced excavation) is studied herein. The Mohr-Coulomb plastic model and the cap-yield model, known as MC model and CY model, respectively, are adopted for the intended study, as these two models have been well calibrated in the FLAC (2011). However, the calibration in the

numerical method-based software is only for the mean of the model parameters. The levels of uncertainty in these model parameters are unknown and are likely different for these two models, such uncertainties require to be characterized from a site investigation program. In this problem, however, the levels of uncertainties in the soil model parameters are assumed and taken to be the same for both soil models. Such simplified assumptions are adopted throughout the analysis in this problem. It should be noted that although only two models (MC model and CY model) are studied and compared in this example, the analysis could be easily extended to cover other soil constitutive models.

In reference to Figure 10, this excavation is 20 m wide and the final excavation depth is 10 m. The diaphragm walls are extended to a 30 m depth and are braced by horizontal struts at 2 m interval. The ground consists of two soil layers: a 20 m thick soft clay underlain by a stiff sand layer that extends to a greater depth. The initial water table is at the ground surface. This braced excavation involves several stages of dewatering, excavation, adding of support, and excavation. In this study, only three main construction stages are considered: 1) dewatering to a 20 m depth in the region to be excavated; 2) excavation to a 2 m depth; and 3) installation of the horizontal strut and excavation to a 10 m depth. Note that both stability and serviceability of the diaphragm wall system can be signaled by the wall deformation, the maximum wall deformation at the final stage is focused in this problem.

The finite difference method (FDM)-based program FLAC version 7.0 (2011) is adopted herein to analyze the wall responses during the excavation. It should be noted that any other suitable code can be used. In the numerical modelling of this braced excavation, the plane-strain condition is assumed for the FDM analysis and only half-width of the excavation is analyzed due to symmetry. To minimize the boundary effect on the wall responses during the excavation, the

right-side boundary is set at 35 m away from the diaphragm wall and the bottom boundary is set at 30 m below the final excavation depth. The bottom boundary is restrained vertically and the left- and right-side boundaries are restrained horizontally. The uniform mesh density with a mesh size of 1.0 m×1.0m is adopted in the FDM mesh. In the built FLAC model, the diaphragm wall and horizontal strut are modelled with beam elements, and the soil-structure interfaces are modeled with interface elements. The material properties of the diaphragm wall, horizontal strut, and soil-structure interfaces are listed in Table 4.

The hydraulic conductivity and porosity of the soils are assumed to be 10^{-6} m/sec and 0.3, respectively. The stress-strain behavior of the sand is modelled with the Mohr-Coulomb plastic model (MC model), the model parameters are tabulated in Table 5. The stress-strain behavior of the clay, in this study, is modelled with the MC model, and then the CY model for a comparison study. The model parameters of the clay are tabulated in Table 6. It is worth noting that the soil constitutive models, model parameters, and other parameter settings listed in the FLAC User's Manual (2011) are strictly followed in the built FLAC model. Figure 11 depicts that similar wall deformations are derived with MC model and CY model. Thus, the model parameters of the MC model and those of the CY model could be considered "equivalent" in the perspective of the wall deformation evaluation in this braced excavation problem.

As can be seen in Table 6, many more model parameters are involved in the complex CY model, while only a few model parameters are required in the simple MC model. Moreover, the engineers are generally more familiar (or comfortable) with both the physical meaning of the MC model parameters and use of the MC model in numerical modelling. Thus, the model parameters in the MC model could be determined accurately and with ease. In order to analyze the selection of soil constitutive models (i.e., MC model versus CY model) on the maximum wall deformation,

all the non-zero model parameters of the clay tabulated in Table 6 are assumed to be normally distributed. For simplicity, the spatial variability of the model parameters is not included in this problem. The mean of these non-zero model parameters are assumed to be identical to those shown in Table 6, the COVs of these non-zero model parameters are assumed to be 0.10, and the correlations among these non-zero model parameters are ignored.

Based upon the assumed statistical information of the model parameters, the distribution of the predicted maximum wall deformation is obtained with 5,000 MCSs, as shown in Figure 12. In the context of the MCS of these non-zero model parameters, the generated model parameters that are meaningless or invalid are omitted (i.e., not analyzed by the adopted FLAC model). For the CY model, only 4,924 sets of CY model parameters (among the 5,000 sets screened) can lead to converged wall deformation; whereas, in the case of MC model, converged wall deformation can always be obtained. Thus, the built FLAC model is more sensitive to (or less robust against) the CY model parameters. The histograms in Figure 12 depict that the CY model could result in a larger variation in the computed maximum wall deformation, indicated by a wider distribution of the histogram. The COVs of the maximum wall deformation evaluated with the MC model and CY model are 0.11 and 0.14, respectively.

Note that even though the stress-strain behavior of the clay at small strain can be better represented by the complex CY model (Oettl et al. 1998; Surarak et al. 2012), the CY model could lead to a larger variation in the model prediction considering the uncertainty in the model parameters. The implication is that although the complex CY model could be more accurate, it is less robust compared to the simple constitutive models such as the MC model. It should be noted that in this problem, no benchmark maximum wall deformation is available; as such, the direct comparison cannot be made through an examination of the discrepancy between the predicted

performance and the benchmark performance (as shown previously in Figure 2 and Figure 6). For a quantitative comparison of the model robustness, an amplifying factor, denoted as AF, is defined as follows.

$$\mathrm{AF} = \frac{\delta_{MP}}{\delta_{mp}} \qquad (14)$$

where δ_{MP} represents the COV of the computed maximum wall deformation, and δ_{mp} represents the COV of the model parameters of the soil constitutive model. Note that the amplifying factor (AF) defined in Eq. (14) is not a bias factor that indicates the ratio of the predicted performance over the benchmark performance; rather, it may be interpreted as a sensitivity index of the soil constitutive model, which signals the degree of variation in the system response given a specified degree of variation in the input model parameters. A smaller AF value indicates that the system performance is less sensitive to (or more robust against) the uncertainties in the input model parameters of the soil constitutive model, and thus implies higher robustness. This is consistent with the robustness concept originated by Taguchi (1986). With the MCS results, the COV of the predicted maximum wall deformation (δ_{MP}) can be derived, and then the amplifying factors (AFs) of the soil constitutive models could be calculated. The AFs are 1.43 and 1.10 with respect to the CY model and MC model, respectively; and thus, the simple MC model in this problem yields a more robust prediction than the CY model does.

Next, different degrees of variation in the model parameters are studied and the resulting amplifying factors are illustrated in Figure 13. It can be seen from Figure 13 that the MC model consistently yields a lower AF value and thus offers a more robust prediction of the maximum wall deformation. In other words, even though the stress-strain behavior of a soil could be more accurately simulated by the complex soil constitutive model, a larger variation in the computed

system performance can be resulted utilizing this complex model. Thus, for numerical modeling and analysis, use of a complex soil constitutive model (i.e., CY model in this problem) does not necessarily outperform a simple soil constitutive model (i.e., MC model in this problem) from the perspective of model robustness.

It is noted that the selection of the most suitable soil constitutive model in the numerical modelling and analysis is a significant challenge that has been rarely analyzed, especially in the situation where the model parameters cannot be characterized with certainty. In this problem, the following assumptions were made: 1) the levels of uncertainties in the model parameters of the MC model and those of the CY model are assumed to be identical; 2) in both MC model and CY model, the uncertainties in all those non-zero model parameters are assumed to be identical; 3) in both MC model and CY model, the correlations among these model parameters are ignored; and 4) the depth-dependent feature of the soil properties (e.g., the strength and deformation modulus), in both MC model and CY model, is ignored. Though several assumptions were made, the results presented provide an insight into the choice between the simple soil constitutive model and the complex soil constitutive model in numerical modelling and analysis. Although the stress-strain behavior of a soil could be more accurately simulated by a complex constitutive model, a larger variation in the computed system performance might be resulted with this complex model, thus a simple constitutive model could be preferred in such scenarios. For a more comprehensive study of this problem, it is desirable to consider the following aspects: 1) given a specified amount of site-specific data, the characterized uncertainties of different soil constitutive models could be different; 2) within a specified soil constitutive model, the characterized uncertainties in different model parameters can be different, and these model parameters can be correlated; and 3) within a specified constitutive model, the spatial variability of the model parameters cannot be negligible.

5. Concluding Remarks

This paper presented a preliminary study on the selection of solution models in the face of uncertainty (in the model parameters) by means of three representative geotechnical problems. The first problem considers the selection of the order of the polynomial fit in the development of data-driven empirical models, the second problem considers the selection between the random variable and the random field in the probabilistic characterization of a soil property at a site, and the third problem considers the selection of soil constitutive models in the numerical modelling of a braced excavation. The following conclusions are drawn based upon the results presented:

(1) Due to the limited availability of site-specific data, and the existence of test error and inherent variability in natural deposits, the model parameters of the geotechnical model could not be characterized with certainty, which tends to result in a significant uncertainty in the predicted system performance. Because of the uncertainty in the model parameters, a complex model may be more accurate but less robust. Further, the uncertainty in the model parameters of the complex model could be more difficult to be characterized, as a larger number of model parameters are involved. Thus, a complex model does not necessarily indicate a better model; instead, a simple model might outperform a complex model in the face of uncertainty in the model parameters.

(2) In the development of data-driven empirical models in geological and/or geotechnical engineering, the fidelity could be guaranteed by adopting a higher-order polynomial fit. Even so, discrepancies between model predictions and field observations always exist; and as such, model parameters should be characterized as uncertain variables and the intended system performance should be evaluated in a probabilistic manner. The robustness of a model, which can be indicated by the variation of the model prediction, increases first and then decreases with the order of the

polynomial fit. Indeed, fidelity and robustness are two conflicting objectives in the development of data-driven empirical models. In consideration of both model fidelity and model robustness, a lower-order polynomial fit might outperform a higher-order polynomial fit.

(3) In the site characterization of a soil property, the residual between the spatial trend, which can be captured by a polynomial fit, and the in situ property can be characterized either by a random variable or a random field. Although the random field model is more sophisticated and theoretical sound, the characterization of the random field is much more difficult than that of the random variable, and this issue becomes more profound when the site-specific data are limited. The larger uncertainty in the characterization of the random field could lead to larger uncertainty in the predicted system performance, compared to the random variable. Thus, in consideration of the uncertainty in the model parameters, the simple random variable model might outperform the complex random field model in predicting the geotechnical system performance.

(4) In the selection of soil constitutive models in the numerical modelling and analysis, the stress-strain behavior of the soil could be more precisely simulated by a complex constitutive model (e.g., CY model). Consequently, more model parameters are involved in the complex soil constitutive model, in comparison to the simple soil constitutive model (e.g., MC model). Due to the limited availability of site-specific data, the uncertainty in the model parameters of the complex soil constitutive model can be more challenging to be characterized. Given a specified level of uncertainty in the model parameters, the complex soil constitutive model may lead to a larger variation in the computed system performance. Thus, the complex soil constitutive model can be precise but not robust, and a complex soil constitutive model does not necessarily indicate a better soil constitutive model in the numerical modelling of a geotechnical system.

It should be noted that significant simplifications are made in the three problems studied, and the findings reported herein on the choice of solution models in the face of uncertainty are deemed preliminary. Model selection in geological and/or geotechnical engineering is indeed a great challenge, it deserves further research. For example, the measures of both model fidelity and model robustness have not been formally defined, and the tradeoff between model fidelity and model robustness in the optimal selection of the solution model has not been examined. Nevertheless, the observations and insight reported in this paper can add to the general literature of the model selection in, and beyond, the field of geological and/or geotechnical engineering.

Acknowledgments

The first author wishes to acknowledge the support of Shanghai Summit Discipline Development Program. The second author wishes to acknowledge the financial support provided by the Natural Science Foundation of China (No. 41702294), the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan), of China, and the Glenn Department of Civil Engineering, Clemson University, South Carolina, USA.

References

Ang, A.H.S., and Tang, W.H. (1984). Probability Concepts in Engineering Planning and Design: Decision, Risk and Reliability Vol. 2, New York, USA: Wiley.

Beck, J.L., and Au, S.K. (2002). "Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation." Journal of Engineering Mechanics, 128(4), 380-391.

Bong, T., and Stuedlein, A.W. (2017). "Spatial variability of CPT parameters and silty fines in liquefiable beach sands." Journal of Geotechnical and Geoenvironmental Engineering, 143(12), 04017093.

Bong, T., and Stuedlein, A.W. (2018). "Effect of cone penetration conditioning on random field model parameters and impact of spatial variability on liquefaction-induced differential settlements." Journal of Geotechnical & Geoenvironmental Engineering, 144(5), 04018018.

Cao, Z., and Wang, Y. (2012). "Bayesian approach for probabilistic site characterization using cone penetration tests." Journal of Geotechnical and Geoenvironmental Engineering, 139(2), 267-276.

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

Cassidy, M.J., Uzielli, M., and Tian, Y. (2013). "Probabilistic combined loading failure envelopes of a strip footing on spatially variable soil." Computers and Geotechnics, 49, 191-205.

Chen, G., Zhu, J., Qiang, M., and Gong, W. (2018). "Three-dimensional site characterization with borehole data – A case study of Suzhou area." Engineering Geology, 234, 65-82.

Cheung, R.W.M., and Tang, W.H. (2005). "Realistic assessment of slope reliability for effective landslide hazard management." Géotechnique, 55(1), 85-94.

Ching, J., Wang, J.S., Juang, C.H., and Ku, C.S. (2015). "Cone penetration test (CPT)-based stratigraphic profiling using the wavelet transform modulus maxima method." Canadian Geotechnical Journal, 52(12), 1993-2007.

Ching, J., and Wang, J.S. (2016). "Application of the transitional Markov chain Monte Carlo algorithm to probabilistic site characterization." Engineering Geology, 203, 151-167.

Ching, J., Phoon, K.K., and Wu, S.H. (2016a). "Impact of statistical uncertainty on geotechnical reliability estimation." Journal of Engineering Mechanics, 142(6), 04016027.

Ching, J., Hu, Y.G., and Phoon, K.K. (2016b). "On characterizing spatially variable soil shear strength using spatial average." Probabilistic Engineering Mechanics, 45, 31-43.

Ching, J., and Phoon, K.K. (2017). "Characterizing uncertain site-specific trend function by sparse Bayesian learning." Journal of Engineering Mechanics, 143(7), 04017028.

Ching, J.Y., Phoon, K.K., Beck, J.L., and Huang, Y. (2017). "Identifiability of geotechnical site-specific trend functions." ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 3(4), 04017021.

Chiu, C.F., Yan, W.M, and Yuen, K.V. (2012). "Estimation of water retention curve of granular soils from particle-size distribution – A Bayesian probabilistic approach." Canadian Geotechnical Journal, 49(9), 1024-1035.

Cho, S.E. (2007). "Effects of spatial variability of soil properties on slope stability." Engineering Geology, 92(3), 97-109.

Christian, J.T., Ladd, C.C., and Baecher, G.B. (1994). "Reliability applied to slope stability analysis." Journal of Geotechnical Engineering, 120(12), 2180-2207.

Degroot, D.J., and Baecher, G.B. (1993). "Estimating autocovariance of in - situ soil properties." Journal of Geotechnical Engineering, 119(1), 147-166.

Fenton, G.A. (1999a). "Estimation for stochastic soil models." Journal of Geotechnical and Geoenvironmental Engineering, 125(6), 470-485.

Fenton, G.A. (1999b). "Random field modeling of CPT data." Journal of Geotechnical and Geoenvironmental Engineering, 125(6), 486-498.

FLAC version 7.0. (2011). Fast Lagrangian Analysis of Continua. Minneapolis, USA: Itasca Consulting Group Inc.

Goldstein, R.V., Dudchenko, A.V., and Kuznetsov, S.V. (2016). "The modified Cam-Clay (MCC) model: Cyclic kinematic deviatoric loading." Archive of Applied Mechanics, 86(12), 2021-2031.

Gong, W., Luo, Z., Juang, C.H., Huang, H., Zhang, J., and Wang, L. (2014). "Optimization of site exploration program for improved prediction of tunneling-induced ground settlement in clays." Computers and Geotechnics, 56, 69-79.

Gong, W., Tien, Y.M., Juang, C.H., Martin, J.R., and Zhang, J. (2016). "Calibration of empirical models considering model fidelity and model robustness – Focusing on predictions of liquefaction-induced settlements." Engineering Geology, 203, 168-177.

Gong, W., Tien, Y.M., Juang, C.H., Martin, J.R., and Luo, Z. (2017). "Optimization of site investigation program for improved statistical characterization of geotechnical property based on random field theory." Bulletin of Engineering Geology and the Environment, 76(3), 1021-1035.

Gong, W., Juang, C.H., Martin, J.R., Tang, H., Wang, Q., and Huang, H. (2018). "Probabilistic analysis of tunnel longitudinal performance based upon conditional random field simulation of soil properties." Tunnelling and Underground Space Technology, 73, 1-14.

Hashash, Y.M.A. and Whittle, A.J. (1996). "Ground movement prediction for deep excavations in soft clay." Journal of Geotechnical Engineering, 122, 474-486.

Hastings, W.K. (1970). "Monte Carlo sampling methods using Markov chains and their applications." Biometrika, 57(1), 97-109.

Hinchberger, S.D., and Qu, G.F. (2009). "Viscoplastic constitutive approach for rate-sensitive structured clays." Canadian Geotechnical Journal, 46(6), 609-626.

Huang, F.M., Huang, J.S., Jiang, S.H., and Zhou, C.B. (2017). "Landslide displacement prediction based on multivariate chaotic model and extreme learning machine." Engineering Geology, 218, 173-186.

Huang, K., Wang, Y., and Cao, Z. (2014). "Bayesian identification of soil strata in London clay." Géotechnique, 64(3), 239-246.

Jiang, S.H., Li, D.Q., Zhang, L.M., and Zhou, C.B. (2014). "Slope reliability analysis considering spatially variable shear strength parameters using a non-intrusive stochastic finite element method." Engineering geology, 168, 120-128.

Juang, C.H., Yuan, H., Lee, D.H., and Lin, P.S. (2003). "Simplified cone penetration test-based method for evaluating liquefaction resistance of soils." Journal of Geotechnical and Geoenvironmental Engineering, 129(1), 66-80.

Juang, C.H., Luo, Z., Atamturktur, S., and Huang, H. (2013). "Bayesian updating of soil parameters for braced excavations using field observations." Journal of Geotechnical and Geoenvironmental Engineering, 139(3), 395-406.

Jung, Y.H., Chung, C.K., and Finno, R.J. (2004). "Development of nonlinear cross-anisotropic model for the pre-failure deformation of geomaterials." Computers and Geotechnics, 31(2), 89-102.

Khoshnevisan, S., Juang, H., Zhou, Y.G., and Gong, W. (2015). "Probabilistic assessment of liquefaction-induced lateral spreads using CPT – Focusing on the 2010-2011 Canterbury earthquake sequence." Engineering Geology, 192, 113-128.

Kung, T.C., Ou, C.Y., and Juang, C. H. (2009). "Modeling small-strain behavior of Taipei clays for finite element analysis of braced excavations." Computers and Geotechnics, 36(1-2), 304-319.

Kuok, S.C., Ng, I.T., and Yuen, K.V. (2015). "Study of the attenuation relationship for the Wenchuan Ms 8.0 earthquake." Earthquake Engineering and Engineering Vibration, 14(1), 1-11.

Li, Z., Wang, X., Wang, H., and Liang, R.Y. (2016). "Quantifying stratigraphic uncertainties by stochastic simulation techniques based on Markov random field." Engineering Geology, 201(2016), 106-122.

Lim, A., Ou, C.Y., and Hsieh, P.G. (2010). "Evaluation of clay constitutive models for analysis of deep excavation under undrained conditions." Journal of GeoEngineering, 5(1), 9-20.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. (1953). "Equation of state calculations by fast computing machines." The Journal of Chemical Physics, 21(6), 1087-1092.

Meyerhof, G.G. (1976). "Bearing capacity and settlement of pile foundations (11th Terzaghi 612 Lecture)." Journal of the Geotechnical Engineering Division 102(3), 195-228.

Ma, J., Tang, H., Liu, X., Hu, X., Sun, M., and Song, Y. (2017). "Establishment of a deformation forecasting model for a step-like landslide based on decision tree C5.0 and two-step cluster algorithms: a case study in the Three Gorges Reservoir area, China." Landslides, 14(3), 1275-1281.

Miao, F.S., Wu, Y.P., Xie, Y.H., and Li, Y.N. (2018). "Prediction of landslide displacement with step-like behavior based on multialgorithm optimization and a support vector regression model." Landslides, 15(3), 475-488.

Oettl, G., Stark, R.F., and Hofstetter, G. (1998). "A comparison of elastic-plastic soil models for 2D FE analyses of tunnelling." Computers and Geotechnics, 23(1), 19-38.

Oliveira, P.J.V., and Lemos, L.J.L. (2014). "Experimental study of isotropic and anisotropic constitutive models." Journal of Geotechnical and Geoenvironmental Engineering, 140(8), 06014008.

Ooi, P.S.K., and Ramsey, T.L. (2003). "Curvature and bending moments from inclinometer data." International Journal of Geomechanics, 3(1), 64-74.

Phoon, K.K., and Kulhawy, F.H. (1999). "Characterization of geotechnical variability." Canadian Geotechnical Journal, 36(4), 612-624.

Phoon, K.K., Quek, S.T., and An, P. (2003). "Identification of statistically homogeneous soil layers using modified Bartlett statistics." Journal of Geotechnical and Geoenvironmental Engineering, 129(7), 649-659.

Plaxis, B.V. (2002). PLAXIS Version 8, Tutorial Manual. R.B.J. Brinkgreve, ed. Delft: Plaxis.

Ren, F., Wu, X.L, Zhang, K.X., and Niu, R.Q. (2015). "Application of wavelet analysis and a particle swarm-optimized support vector machine to predict the displacement of the

Shuping landslide in the Three Gorges, China." Environmental Earth Sciences, 73(8), 4791-4804.

Rezania, M., Javadi, A.A., and Giustolisi, O. (2010). "Evaluation of liquefaction potential based on CPT results using evolutionary polynomial regression." Computers and Geotechnics, 37(1), 82-92.

Robert, D.J. (2017). "A modified Mohr-Coulomb model to simulate the behavior of pipelines in unsaturated soils." Computers and Geotechnics, 91, 146-160.

Smith, C.C. (2005). "Complete limiting stress solutions for the bearing capacity of strip footings on a Mohr-Coulomb soil." Géotechnique, 55(8), 607-612.

Stallebrass, S.E., and Taylor, R.N. (1997). "The development and evaluation of a constitutive model for the prediction of ground movements in overconsolidated clay." Géotechnique, 47(2), 235-253.

Suebsuk, J., Horpibulsuk, S., and Liu, M.D. (2010). "Modified Structured Cam Clay: a generalised critical state model for destructured, naturally structured and artificially structured clays." Computers and Geotechnics, 37(7-8), 956-968.

Surarak, C., Likitlersuang, S., Wanatowski, D., Balasubramaniam, A., Oh, E., and Guan, H. (2012). "Stiffness and strength parameters for hardening soil model of soft and stiff Bangkok clays." Soils and Foundations, 52(4), 682-697.

Taguchi, G. 1986. Introduction to quality engineering: designing quality into products and processes. Quality Resources, White Plains, N.Y.

Ting, J.M. (1987). "Full-scale cyclic dynamic lateral pile responses." Journal of Geotechnical Engineering, 113(1), 30-45.

Uzielli, M., Vannucchi, G., and Phoon, K.K. (2005). "Random field characterisation of stress-nomalised cone penetration testing parameters." Géotechnique, 55(1), 3-20.

Wang, Y., and Zhao, T. (2016). "Interpretation of soil property profile from limited measurement data: A compressive sampling perspective." Canadian Geotechnical Journal, 53, 1547-1559.

Wang, X., Wang, H., and Liang, R.Y. (2017). "A method for slope stability analysis considering subsurface stratigraphic uncertainty." Landslides(8), 1-12.

Whittle, A.J. (1993). "Evaluation of a constitutive model for overconsolidated clays." Géotechnique, 43(2), 289-313.

Xiao, T., Zhang, L.M., Li, X.Y., and Li, D.Q. (2017). "Probabilistic stratification modeling in geotechnical site characterization." ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 3(4), 04017019

Yuen, K.V., and Mu, H.Q. (2015). "Real-time system identification: An algorithm for simultaneous model class selection and parametric identification." Computer-Aided Civil and Infrastructure Engineering, 30(10), 785-801.

Yuen, K.V., Ortiz, G.A., and Huang, K. (2016). "Novel nonparametric modeling of seismic attenuation and directivity relationship." Computer Methods in Applied Mechanics and Engineering, 311, 537-555.

Zhang, J., Zhang, L.M., and Tang, W.H. (2009). "Bayesian framework for characterizing geotechnical model uncertainty." Journal of Geotechnical and Geoenvironmental Engineering, 135(7), 932-940.

Zhang, J., Huang, H.W., Juang, C.H., and Su, W.W. (2014). "Geotechnical reliability analysis with limited data: consideration of model selection uncertainty." Engineering Geology, 181, 27-37.

Zhang, W.G., and Goh, A.T.C. (2013). "Multivariate adaptive regression splines for analysis of geotechnical engineering systems." Computers and Geotechnics, 48, 82-95.

Appendix A. Procedures for the MCMC Simulation-Based Bayesian Inference Method

The procedures for the MCMC simulation-based Bayesian inference method for deriving the statistics of the unknown parameters $\boldsymbol{\theta}$ using Metropolis-Hastings algorithm are summarized as follows (Gong et al. 2017):

Step 1: At stage $m = 1$, sample a starting point, denoted as $\boldsymbol{\theta}_1$, in the Markov chain. This starting point might be randomly generated from the prior distribution $f'(\boldsymbol{\theta})$ or assigned the mean of the prior distribution $\boldsymbol{\mu}_{\boldsymbol{\theta}}$.

Step 2: At next stage (m starts from 2), sample a candidate point $\boldsymbol{\theta}^*$ from a proposal distribution of $J(\boldsymbol{\theta}^*|\boldsymbol{\theta}_{m-1})$, which is assumed to be a multivariate normal distribution, in which the mean is set to the current sample point $\boldsymbol{\theta}_{m-1}$ and the covariance is set to $s\mathbf{C}_{\boldsymbol{\theta}}$, where s represents a scaling factor and $\mathbf{C}_{\boldsymbol{\theta}}$ represents the covariance of the prior distribution of $\boldsymbol{\theta}$. The scaling factor s should be selected so that the acceptance rate of the MCMC samples is between 20% and 40%.

Step 3: Calculate the ratio of the un-normalized posterior probability density function (PDF) of $\boldsymbol{\theta}^*$ over that of $\boldsymbol{\theta}_{m-1}$, denoted as ζ, as follows.

$$\zeta = \min\left[\frac{L(\mathbf{d}\,|\,\boldsymbol{\theta}^*)\,f'(\boldsymbol{\theta}^*)}{L(\mathbf{d}\,|\,\boldsymbol{\theta}_{m-1})\,f'(\boldsymbol{\theta}_{m-1})}, 1\right] \qquad \text{(A1)}$$

where $L(\mathbf{d}|\boldsymbol{\theta})$ represents the likelihood of observing test data $\mathbf{d}$ given the unknown parameters of $\boldsymbol{\theta}$.

Step 4: Sample a random variable, denoted as ξ, from a uniform distribution of $U(0, 1)$.

Step 5: Determine whether the candidate point $\boldsymbol{\theta}^*$ is acceptable with the following rule: if $\xi \le \zeta$, then the candidate point $\boldsymbol{\theta}^*$ is acceptable and set $\boldsymbol{\theta}_m = \boldsymbol{\theta}^*$; otherwise, the candidate point $\boldsymbol{\theta}^*$ is not acceptable and set $\boldsymbol{\theta}_m = \boldsymbol{\theta}_{m-1}$. Note that when the number of test data is large, the likelihood

function $L(\mathbf{d}|\boldsymbol{\theta}^*)$ becomes significantly small and the mathematical operation of the density ratio ζ becomes a challenge. To avoid such occurrences, the logarithm of the density ratio ζ, denoted as $\ln\zeta$, might be computed for evaluating the acceptability of the candidate point $\boldsymbol{\theta}^*$. Then, Eq. (A1) is modified into:

$$\ln\zeta = \min\left\{\ln\left[L\left(\mathbf{d}\,|\,\boldsymbol{\theta}^*\right)\right] + \ln\left[f\left(\boldsymbol{\theta}^*\right)\right] - \ln\left[L\left(\mathbf{d}\,|\,\boldsymbol{\theta}_{m-1}\right)\right] - \ln\left[f\left(\boldsymbol{\theta}_{m-1}\right)\right], 0\right\} \quad \text{(A2)}$$

Accordingly, the acceptance rule is modified as: if $\ln\xi \le \ln\zeta$, then accept the candidate point $\boldsymbol{\theta}^*$ and set $\boldsymbol{\theta}_m = \boldsymbol{\theta}^*$; otherwise, reject the candidate point $\boldsymbol{\theta}^*$ and set $\boldsymbol{\theta}_m = \boldsymbol{\theta}_{m-1}$.

Step 6: Repeat Steps 2-5 until a target number of MCMC samples, in terms of Markov chain length, is reached. Based upon the obtained MCMC samples, the statistics of the unknown parameters $\boldsymbol{\theta}$ can be derived straightforwardly.

Table 1. Deterministic models obtained with the synthetic test data $\mathbf{d}_1$ shown in Figure 1 using the least square regression analysis (Problem 1)

Model	Model coefficients **p**	Fidelity in matching the test data **d**	
		Mean of $[f(x) - d(x)]$	Std. Dev. of $[f(x) - d(x)]$
P1	1.4534 0.4341	-8.92×10^{-6}	0.4875
P2	0.9065 0.8033 -0.0369	9.47×10^{-6}	0.3403
P3	0.6116 1.2888 -0.1649 0.0085	7.86×10^{-7}	0.2622
P4	0.7536 0.7616 0.1009 -0.0341 0.0021	-8.44×10^{-5}	0.2216
P5	0.8959 -0.3204 0.9675 -0.2737 0.0292 -0.0011	-7.70×10^{-4}	0.1392
P6	0.8305 0.5215 -0.0247 0.1468 -0.0515 0.0061 -0.0002	1.57×10^{-2}	0.1122

Table 2. Probabilistic models obtained with the synthetic test data $\mathbf{d}_1$ shown in Figure 1 using the MCMC-based Bayesian inference method (Problem 1)

Model	Mean $\boldsymbol{\mu}_p^T$	Covariance $\mathbf{C}_p$
P1	1.4525 0.4341	1.11E-02 -1.38E-03 -1.38E-03 3.93E-04
P2	0.9049 0.8034 -0.0369	6.57E-03 -1.03E-03 1.97E-05 -1.03E-03 9.81E-04 -7.96E-05 1.97E-05 -7.96E-05 1.01E-05
P3	0.6122 1.2890 -0.1648 0.0085	3.27E-03 -8.73E-04 4.47E-05 9.54E-07 -8.73E-04 3.11E-03 -4.24E-04 8.59E-06 4.47E-05 -4.24E-04 9.64E-05 -5.06E-06 9.54E-07 8.59E-06 -5.06E-06 4.66E-07
P4	0.7473 0.7624 0.0997 -0.0337 0.0021	5.15E-03 -1.06E-03 -3.26E-05 1.53E-05 -6.49E-07 -1.06E-03 2.59E-03 -2.42E-04 -3.10E-05 2.91E-06 -3.26E-05 -2.42E-04 8.66E-05 -6.82E-06 3.68E-08 1.53E-05 -3.10E-05 -6.82E-06 3.41E-06 -2.45E-07 -6.49E-07 2.91E-06 3.68E-08 -2.45E-07 2.24E-08
P5	0.8944 -0.3103 0.9596 -0.2730 0.0293 -0.0011	7.95E-03 -7.21E-04 -1.20E-03 9.31E-05 2.82E-05 -2.59E-06 -7.21E-04 4.34E-04 1.25E-04 -3.03E-05 -1.40E-06 2.90E-07 -1.20E-03 1.25E-04 1.24E-03 -3.02E-04 1.69E-05 1.08E-07 9.31E-05 -3.03E-05 -3.02E-04 9.82E-05 -9.54E-06 2.71E-07 2.82E-05 -1.40E-06 1.69E-05 -9.54E-06 1.51E-06 -7.23E-08 -2.59E-06 2.90E-07 1.08E-07 2.71E-07 -7.23E-08 4.45E-09
P6	0.8537 0.4780 -0.0237 0.1476 -0.0508 0.0059 -0.0002	4.81E-03 -4.25E-04 5.47E-06 2.34E-05 -3.20E-05 6.06E-06 -3.11E-07 -4.25E-04 4.51E-04 3.09E-06 -8.51E-05 2.24E-05 -2.12E-06 6.91E-08 5.47E-06 3.09E-06 1.29E-06 -3.58E-06 9.47E-07 -8.85E-08 2.80E-09 2.34E-05 -8.51E-05 -3.58E-06 3.88E-05 -9.81E-06 7.98E-07 -1.92E-08 -3.20E-05 2.24E-05 9.47E-07 -9.81E-06 2.83E-06 -2.72E-07 8.46E-09 6.06E-06 -2.12E-06 -8.85E-08 7.98E-07 -2.72E-07 3.19E-08 -1.24E-09 -3.11E-07 6.91E-08 2.80E-09 -1.92E-08 8.46E-09 -1.24E-09 5.84E-11

Table 3. Probabilistic characterization of the normalized undrained shear strength using the synthetic test data $\mathbf{d}_1$ shown in Figure 4 (Problem 2)

Spatial trend, $\sum_{i=0}^{i=n_p} p_i z^i$		Residual, ω		Fidelity in matching the true bearing capacity, $(Q_{ULS}^T - Q_{ULS}^P)/Q_{ULS}^T$		
Model	Model coefficients, $\mathbf{p}^T$	Model	Statistics of the residual, θ	Mean	Std. Dev.	Pr[$(Q_{ULS}^T - Q_{ULS}^P)/Q_{ULS}^T = 0$]
P1	15.5377 -0.3350	V	1.40E-07 1.6862	-2.79×10^{-3}	3.15×10^{-2}	12.5963 [a]
		F	1.40E-07 1.7019 2.982	-3.14×10^{-3}	7.16×10^{-2}	5.5631
P2	18.4582 -1.3085 0.059	V	-3.70E-07 1.3485	-1.97×10^{-2}	2.53×10^{-2}	11.6548
		F	-3.70E-07 1.3662 2.930	-1.96×10^{-2}	5.72×10^{-2}	6.5707

[a] The model that yields the highest likelihood is identified as the "best" model.

Table 4. Material properties of diaphragm wall, horizontal strut, and soil-structure interfaces (Problem 3)

Category	Parameter	Value
Diaphragm wall	Equivalent thickness, t (m)	1.26
	Density, ρ_w (kg/m^3)	2,000
	Young's modulus, E_w (GPa)	5.712
	Poisson ratio, ν	0.2
	Moment of inertia, I_w (m^4)	0.167
Horizontal strut	Cross-sectional area, A_s (m^2)	1.0
	Spacing, H_s (m)	2.0
	Density, ρ_s (kg/m^3)	3,000
	Young's modulus, E_s (GPa)	4.0
	Moment of inertia, I_s (m4)	0.083
Soil-structure interface	Friction angle, φ_i (°)	12.5
	Cohesion, c_i (kPa)	2.5
	Normal-stiffness, K_{ni} (MPa)	550
	Shear-stiffness, K_{si} (MPa)	550

Table 5. Mohr-Coulomb drained properties of the sand (Problem 3)

Model parameter	Value
Dry density, ρ_{ss} (kg/m^3)	1,700
Young's modulus, E_{ss} (MPa)	40.0
Poisson ratio, ν_{ss}	0.3
Cohesion, c_{ss} (kPa)	0
Friction angle, φ_{ss} (°)	32.0
Dilation angle, ψ_{ss} (°)	2.0

Table 6. Drained model parameters of the clay (Problem 3)

Mohr-Coulomb model (MC model)		Cap-yield model (CY model)	
Model parameter	Value	Model parameter	Value
Dry density, ρ_{sc} (kg/m^3)	1,600	Dry density, ρ_{sc} (kg/m^3)	1,600
Young's modulus, E_{sc} (MPa)	10.0	Cap-yield surface parameter, α	1.0
Poisson ratio, ν_{sc}	0.35	Ultimate friction angle, φ_f (°)	25.0
Cohesion, c_{sc} (kPa)	0	Ultimate dilation angle, ψ_f (°)	0
Friction angle, φ_{sc} (°)	25.0	Multiplier, R	3.333
Dilation angle, ψ_{sc} (°)	0	G_{ref}^{e} (MPa)	10.0
		K_{ref}^{iso} (MPa)	4.0
		Reference pressure, p_{ref} (MPa)	0.1
		Poisson ratio, ν_{ur}	0.2
		Cohesion, c_{sc} (kPa)	0
		Power, m	1.0
		K_0^{nc}	0.5
		Initial mobilized friction angle, φ_m (°)	19.47
		Failure ratio, R_f	0.9

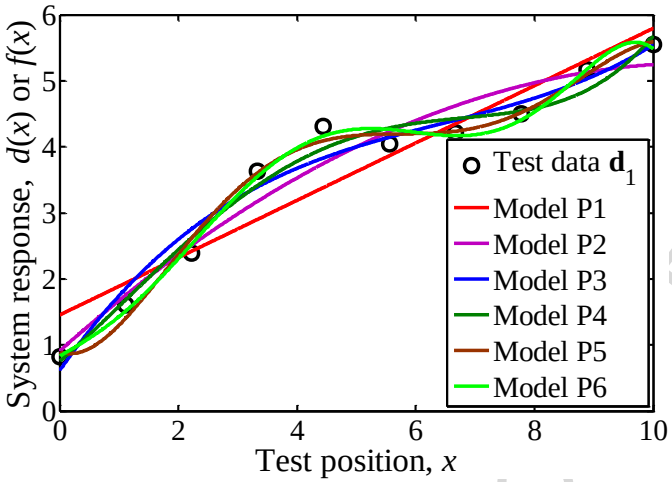


Figure 1. Polynomial fitting of the synthetic test data $\mathbf{d}_1$

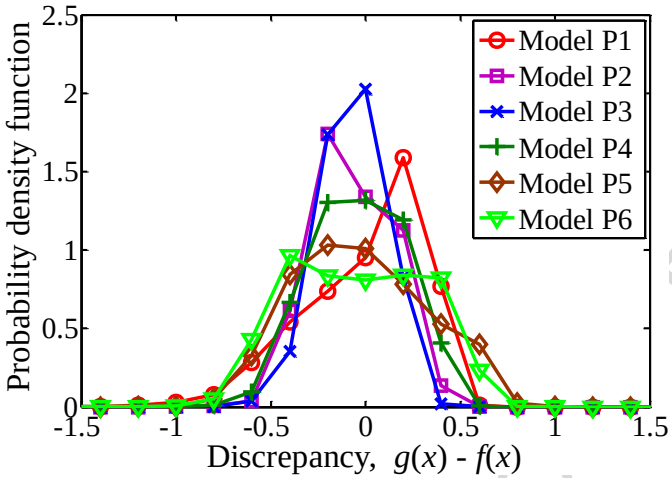
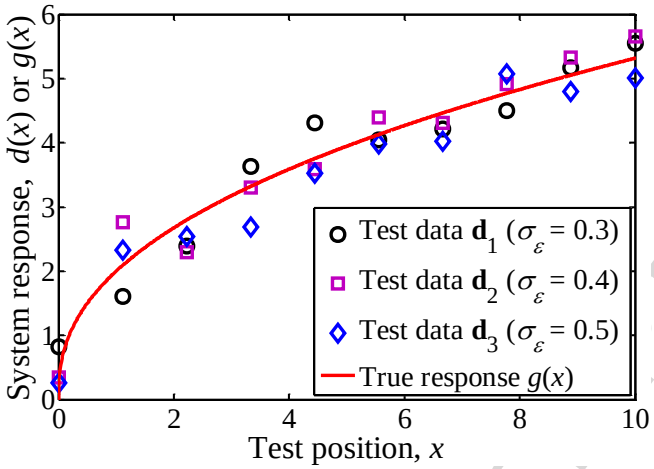
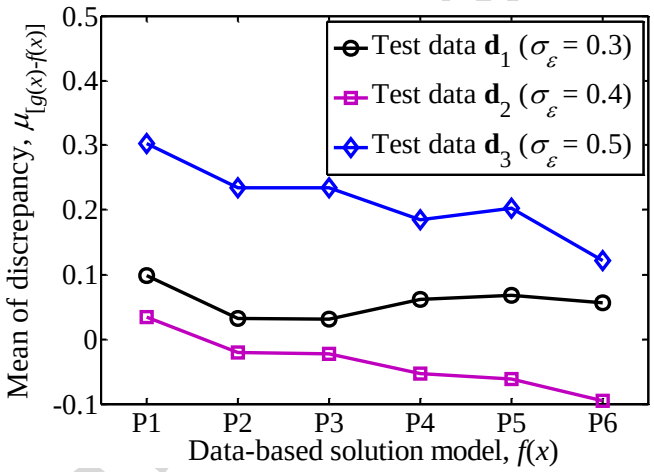


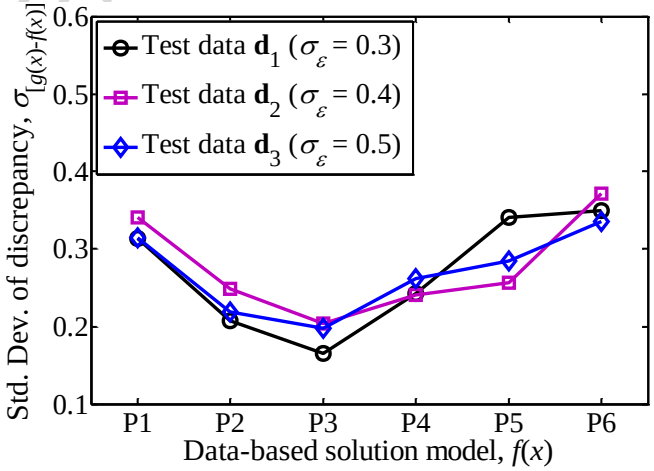
Figure 2. Distribution of the discrepancy between the model prediction and the true model value



(a) 3 sets of synthetic test data studied



(b) Mean of the discrepancy



(c) Std. Dev. of the discrepancy

Figure 3. Statistics of the discrepancy between the model prediction and the true model value

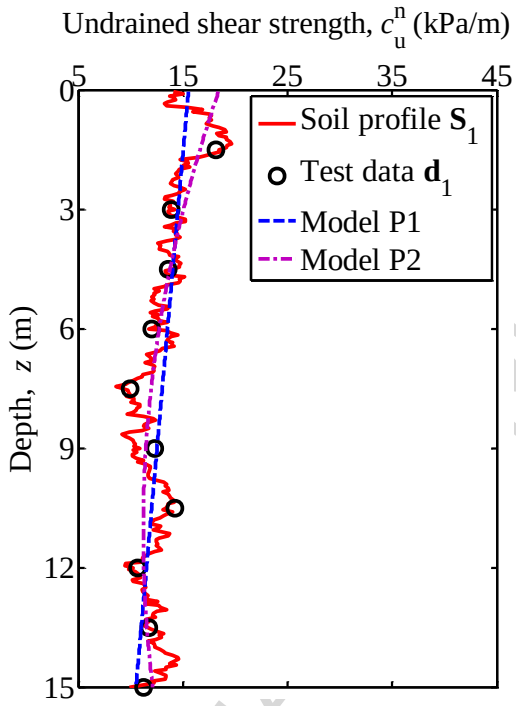


Figure 4. Polynomial fitting of the spatial trend of the normalized undrained shear strength based upon synthetic test data $\mathbf{d}_1$

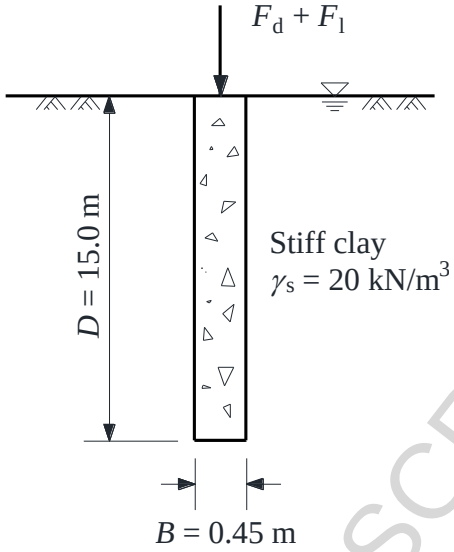


Figure 5. Schematic diagram of a drilled shaft in clay

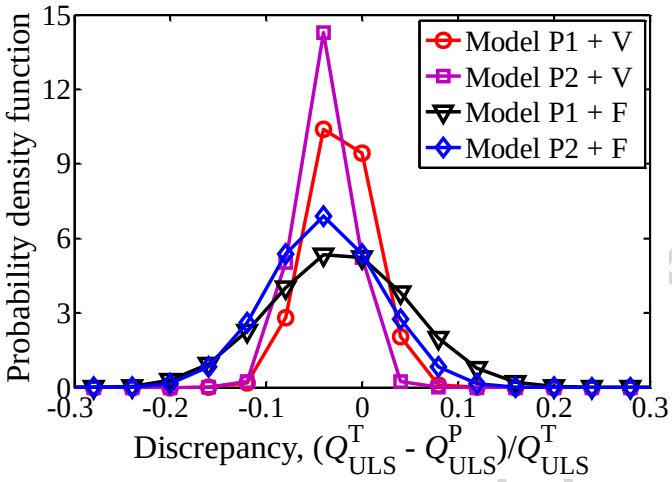


Figure 6. Distribution of the discrepancy between the predicted bearing capacity and the benchmark bearing capacity

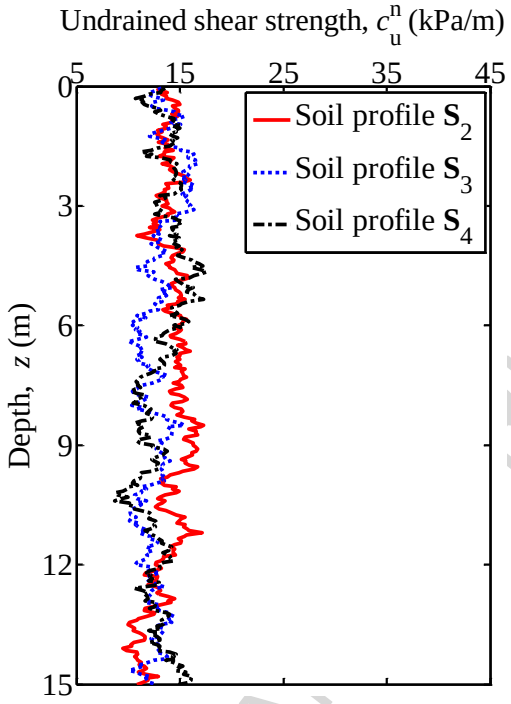


Figure 7. Three other in situ soil profiles studied in Problem 2

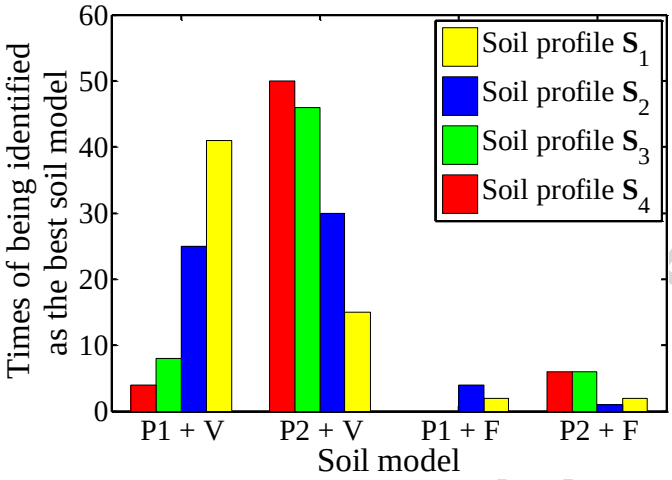
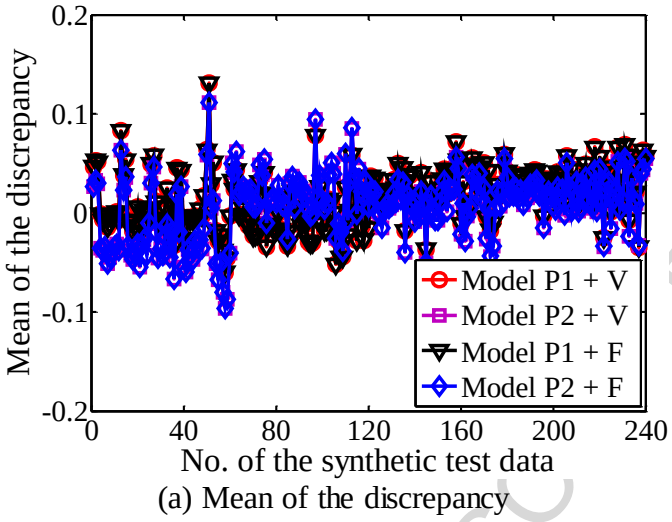
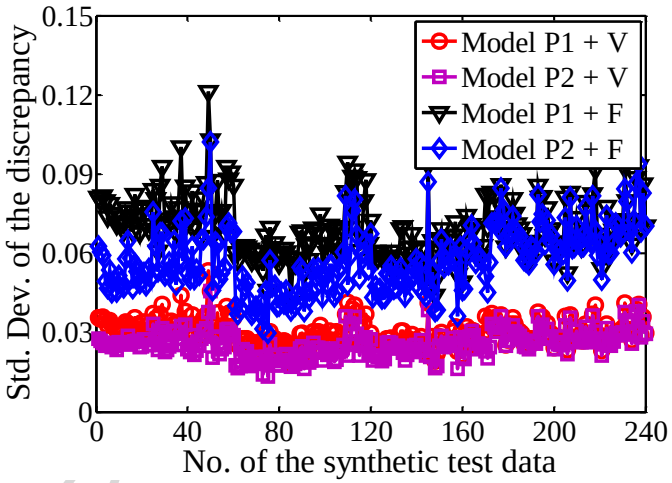


Figure 8. Distribution of the best model among the studied soil models



(a) Mean of the discrepancy



(b) Std. Dev. of the discrepancy

Figure 9. Statistics of the discrepancy $(Q_{ULS}^{T} - Q_{ULS}^{P})/Q_{ULS}^{T}$ obtained with the generated 240 sets of synthetic test data of the normalized undrained shear strength

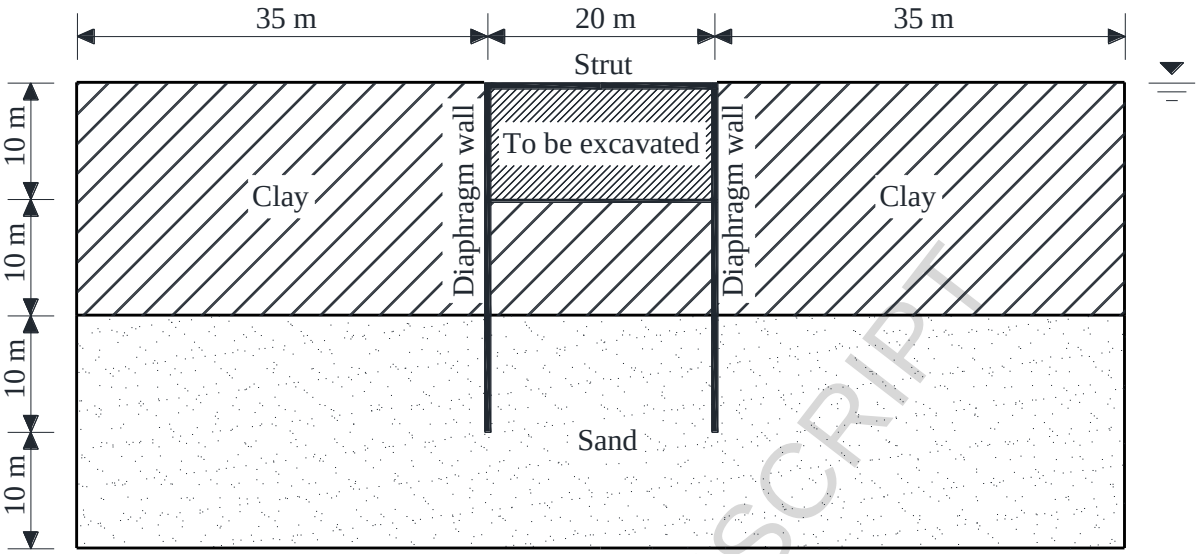


Figure 10. Geometry for the braced excavation in Problem 3

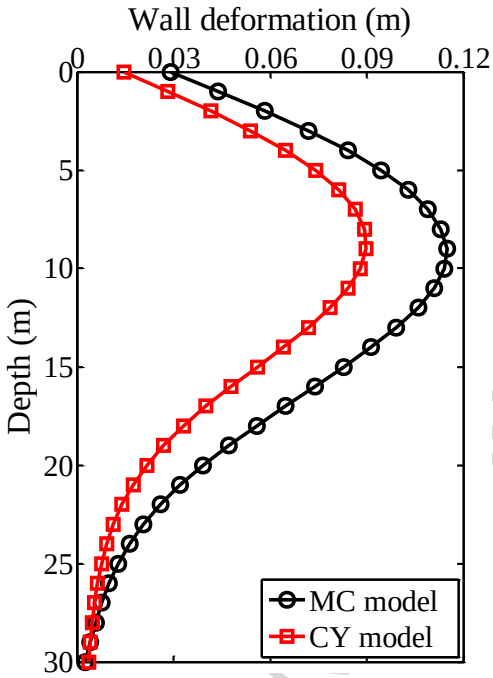


Figure 11. Deterministic analysis results of the MC model and the CY model

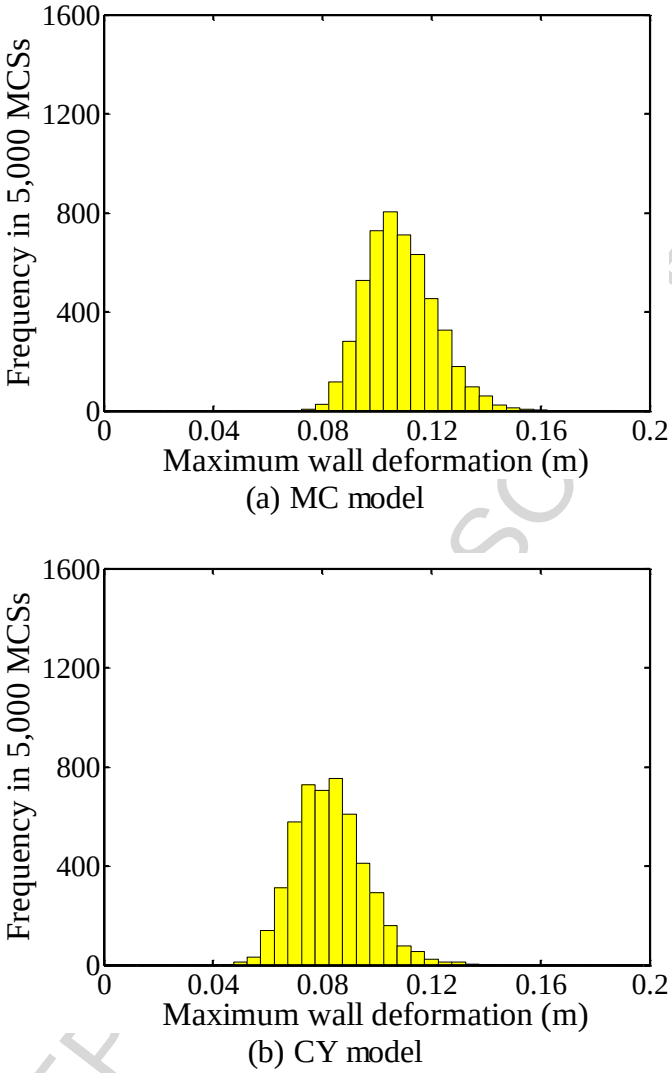


Figure 12. Distribution of the predicted maximum wall deformation

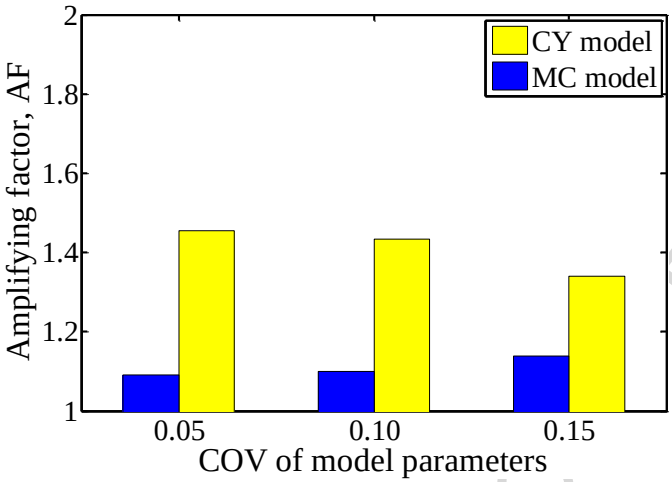


Figure 13. Amplifying factors of the MC model and the CY model

Highlights

- This paper provides an insight on the question, “does a complex model always outperform a simple model in the face of uncertainty?”
- he question is examined in three problem settings: low versus high order of polynomial fits, random variable versus random field, and simple versus complex constitutive model.
- Analysis shows that although a complex model may yield a more accurate prediction, a simple model may yield a more robust prediction.