Geotech Geol Eng (2023) 41:4415–4429
https://doi.org/10.1007/s10706-023-02523-4

ORIGINAL PAPER

Estimation of Uncertainties in Soil Using MCMC Simulation and Effect of Model Uncertainty

Himanshu Rana · Bhardwaj Pandit · G. L. Sivakumar Babu

Received: 27 January 2023 / Accepted: 5 June 2023 / Published online: 14 June 2023

Abstract The simulation of field conditions for seismically induced slope failures incorporates model uncertainties, which account for the difference between simulated and observed slope behaviour. The quantification of this uncertainty is mandatory to understand the field response of the geotechnical system and make decisions for geotechnical systems. Previous studies have partially studied uncertainty for slope systems under seismic loading. To this aim, this study proposes a methodology based on probabilistic back analysis to estimate uncertainties in soil parameters considering the observed slope response under seismic loading. The proposed method involves support vector regression (SVR) model to map the relationship between soil parameters and seismically induced slope displacement. The SVR model is generated using the data from the numerical simulation of slope system under seismic loading using FLAC 2D. Further, the developed SVR model is used for probabilistic back analysis using Markov Chain Monte Carlo (MCMC) simulation. The Noto Hanto earthquake in 2007 and the subsequent slope failure along Noto Yuryo Road, Japan, are considered as a case study to validate the proposed methodology. The results of the case study show that the updated or inferred soil parameters have less variability than the prior distribution. Further, the uncertainties in the slope system influence the inferred soil parameters. Hence, a parametric study is conducted to investigate the effect of model uncertainty on the posterior statistics of soil parameters. The study results facilitate a better understanding of the slope deformation mechanism and the effect of model uncertainty on the updated statistics of soil parameters.

Keywords Probabilistic back analysis · MCMC simulation · Model uncertainty · Seismically induced slope displacement

H. Rana (✉)
Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede, The Netherlands
e-mail: h.rana@utwente.nl; himanshurana@iisc.ac.in; himanshur93@gmail.com

B. Pandit
Department of Mining Engineering, Indian Institute of Technology (BHU), Varanasi, India
e-mail: bhardwajp.min@iitbhu.ac.in

G. L. Sivakumar Babu
Civil Engineering Department, Indian Institute of Science, Bengaluru 560012, India
e-mail: gls@iisc.ac.in

1 Introduction

Slope failures due to seismically induced instabilities pose risk to infrastructure, environment, and lives in earthquake-prone mountainous regions (Gratchev and Towhata 2011; Huang et al. 2016; Wang and Li 2021). The seismic slope displacement is an important measure to determine the severity of landslide hazard and risk. To this aim, numerical models are commonly used to simulate the field conditions of slope under seismic loading using non-linear dynamic analysis (Bray et al. 2018; Arab Khaburi and Mortazavi 2019; Wang and Li 2021). However, the input parameters for numerical models, such as soil parameters, incorporate inherent uncertainties due to non-homogeneous and variable soil (Abbaszadeh et al. 2011; Metya and Bhattacharya 2016; Cheng and He 2020; Rana and Sivakumar Babu 2023). These uncertainties lead to divergence between simulated behaviour using numerical models and observed behaviour of slope in the field.

The inherent uncertainties in soil parameters involve measurement, testing, and observational uncertainties (Abbaszadeh et al. 2011; Tietje et al. 2014; Kong et al. 2022). These uncertain parameters form the basis of further decisions for geotechnical systems and planning. Hence, the accurate quantification of these parameters with inherent uncertainties is necessary for risk management (Alitabar et al. 2021). The accurate estimation of these parameters is not feasible through testing (Abdulai and Sharifzadeh 2019), which led to rise of back or inverse analysis in study of geotechnical systems. The back analysis is conducted using deterministic and probabilistic approach. The deterministic approach utilises one set of parameters with respect to one observation (Feng et al. 2004). On the other hand, the probabilistic approach facilitates the estimation of uncertainties in the input parameters in terms of variability for a given observation of geotechnical system (Gilbert et al. 1998). Therefore, probabilistic back analysis is widely utilised to assess accurate soil parameters with uncertainties. The probabilistic back analysis uses the observed response of the geotechnical system (i.e., the factor of safety or slope displacement) to infer multiple sets of input parameters with different likelihoods responsible for the observed response (Zhang et al. 2010; Wang et al. 2013; Ng et al. 2014; Ering and Sivakumar Babu 2016; Contreras and Brown 2019; Jiang et al. 2020). Several methods, such as Bayesian, least squares, and maximum likelihood, are utilised for probabilistic back analysis of geotechnical systems. The Bayesian method is used to infer the input parameters and the uncertainty of input parameters with the available observations (Gilbert et al. 1998; Miranda et al. 2009; Zhang et al. 2010; Li et al. 2016; Ering and Sivakumar Babu 2016; Rana and Sivakumar Babu 2022a, b).

Probabilistic back analysis of geotechnical systems is also conducted using Markov Chain Monte Carlo (MCMC) simulation (Zhang et al. 2010; Juang et al. 2013; Wang et al. 2013; Wu et al. 2017; Contreras and Brown 2019). Juang et al. (2013) proposed a method for probabilistic back analysis to update the statistics of uncertain soil parameters using MCMC simulation in conjunction with the Bayesian theorem for multi-staged braced excavation. The study results suggest that the proposed method efficiently reduces the uncertainty involved in soil parameters. Zhang et al. (2010) proposed a probabilistic approach to back analyse soil parameters for slope using MCMC simulation and response surface method (RSM). The field performance of slope, i.e., safety factor, is utilised to update soil parameters. They concluded that the correlation between input soil parameters does not influence the posterior distribution. However, the type of prior distribution significantly affects the posterior distribution. Wang et al. (2013) utilised MCMC simulation and the maximum likelihood method to back analyse the anchor force and strength parameters for slope failure in Taiwan. They concluded that improved knowledge of these parameters better explains the slope failure mechanism and forms the basis for remedial measures.

Large-scale projects require extensive computational resources to conduct non-linear dynamic analysis using numerical models for probabilistic back analysis. Therefore, several researchers have adopted surrogate models to replicate the numerical models for back analysis (Feng et al. 2004; Xu et al. 2013; Dilip and Sivakumar Babu 2013; Li et al. 2016; Contreras and Brown 2019). Feng et al. (2004) and Zhao and Yin (2009) used the support vector machine (SVM) model as a surrogate model to replace the numerical model for back analysis of the soil parameter values for given observations. Li et al. (2016) proposed a method for probabilistic back analysis of high-cut rock slope. This method involves a multi-output SVM model as a surrogate model for numerical simulation and Bayesian analysis for back analysis of slope displacement. The study results suggested that back-calculated values of geomechanical parameters closely match the practical values of parameters with associated uncertainty. Wu et al. (2017) presented a polynomial chaos expansion-based MCMC simulation for probabilistic back analysis of rainfall-induced slope failure. The polynomial chaos expansion acts as a surrogate model to replicate the coupled hydromechanical model for unsaturated slope material under rainfall and makes the method faster and more efficient. However, there is limited research on the probabilistic back analysis of seismically induced slope failures. The seismically induced slope displacement requires investigation to elucidate the slope failure mechanism and to estimate variability in soil parameters. The accurate statistics of these parameters form the basis of the provision of remedial measures and reliability-based analysis.

The numerical models used for probabilistic back analysis replicate the field conditions, incorporating model uncertainties (Cheung and Tang 2005). The negligence of model uncertainties may lead to biased geotechnical decisions as model uncertainties affect the predicted values of geotechnical parameters. These geotechnical parameters are further used for geotechnical decisions. Hence, the quantification of the model uncertainties is essential for understanding the soil parameters and their effect on the slope systems. The geotechnical predictions made without considering model uncertainties do not accurately estimate failure probabilities for slope systems (Zhang et al. 2009, 2015).

Huang et al. (2012) proposed a Bayesian network to determine the model uncertainty of a geotechnical system considering uncertainties associated with input parameters. They concluded that ignoring model uncertainty may lead to underestimating or overestimating the reliability index. Zhang et al. (2012) presented a methodology based on Bayesian analysis to characterise the model uncertainty for the geotechnical model. This method considers the uncertainty in input parameters and error in the observed data. Model uncertainty parameters are considered as random variables and estimated using the observed data and input parameters using Bayesian analysis. The updated posterior statistics of model uncertainty determine the characteristics of model uncertainty. Li et al. (2021) proposed a methodology for probabilistic back analysis based on MCMC simulation, considering model bias, parameter uncertainty, and observation error. They concluded that the proposed method efficiently reduces the prediction uncertainty and accurately predicts the geosystem's performance. However, the effect of model uncertainty on the updated posterior statistics of input parameters is partially studied.

The present study aims to propose a methodology for probabilistic back analysis of seismically induced slope displacement to infer the posterior statistics of input parameters for observed seismically induced slope displacement. The proposed method adopts the SVR model as a surrogate model to map the relationship between input parameters, i.e., soil parameters, and response variable, i.e., seismically induced slope displacement. The SVR model is trained using the obtained data from slope simulation under seismic loading using the finite difference method (FDM). The trained model is further used to estimate the posterior distribution of input parameters utilising the MCMC simulation based on the Bayes theorem. A slope failure in 2007 due to the Noto Hanto earthquake along the Noto Yuryo Road, Japan is used as a case study to validate the proposed methodology. Results of the case study suggest that posterior statistics of the input parameter show less variability than prior statistics. The numerical model utilised in this study also incorporates model uncertainty, defined as the difference between the simulated and field responses of the slope system. The effect of the model uncertainty on the estimated posterior distribution is investigated using a parametric study, which facilitate the understanding of model uncertainty. The results of the probabilistic back analysis can be used for further design of remedial measures for slope and reliability-based slope design.

2 Methods

2.1 Support Vector Regression (SVR) Algorithm

The SVR algorithm is a commonly used model to map a non-linear relationship between input parameters and the response variable of geotechnical systems (Feng et al. 2004; Zhao and Yin 2009; Li et al. 2016). The SVR algorithm projects input data into higher dimensional space. The set of sample data is $\{x_i, y_i\}$, where $x_i \in R^n$, $y_i \in R$ and $i = 1,2,\ldots\ldots,m$ for a non-linear problem $f(x)$. The support vectors $(f(x))$ are calculated by solving the following optimisation problem.

$$\text{Maximise } w(\alpha, \alpha^*) = \frac{-1}{2}\sum_{i,j=1}^{m}\left(\alpha_i - \alpha_i^*\right)\left(\alpha_j - \alpha_j^*\right)K\left(x_i.x_j\right) + \sum_{i=1}^{m} y_i(\alpha_i - \alpha_i^*) - \sum_{i=1}^{m}(\alpha_i + \alpha_i^*) \tag{1}$$

$$\text{Subjected to}\begin{cases} \sum_{i=1}^{m}\left(\alpha_i-\alpha_i^*\right)=0 \quad 0\le\alpha_i \\ \alpha_i^*\le C, \quad i=1,2,\ldots.m \end{cases} \tag{2}$$

where C is a regularisation parameter that decides the trade-off between overfitting and excess error ε. α_i and α_i^* are lagrangian multipliers. $K(x_i.x_j)$ is a kernel function defined as

$$K\left(x_i.x_j\right)=\Phi(x_i)\Phi(x_j) \tag{3}$$

where Φ is the formulation of non-linear mapping. The kernel function $K(x_i.x_j)$ *is* generally used in three forms (a) Polynomial kernel, (b) Sigmoid kernel, and (c) Radial basis function kernel.

The regression function can be obtained as follows.

$$f(x)=\sum_{i=1}^{m}\left(\alpha_i-\alpha_i^*\right)K\left(x_i.x_j\right)+b \tag{4}$$

where b is an offset.

2.2 Methodology for Probabilistic Back Analysis

Probabilistic displacement back analysis is performed using the SVR algorithm and MCMC simulation. The SVR model is used as a surrogate model to efficiently replicate the numerical (simulation) model, which substantially reduces effort and computational time. The SVR algorithm mapped the relationship between input soil parameters and slope displacement. Bayes theorem is exploited to back-calculate the posterior distribution of input parameters and the uncertainties associated with these parameters for a given value of displacement. The MCMC simulation is used to obtain the samples from the updated posterior input parameters.

2.2.1 Mapping of Soil Parameters and Displacement

The SVR model is used to map the relationship between the observed slope displacement and soil parameters, i.e., cohesion, internal friction angle, and maximum shear modulus. The relationship is defined as

$$SVR(X) : R^n \rightarrow R \tag{5}$$

$$y=SVR(X) \tag{6}$$

where $X=(x_1, x_2, \ldots..x_n)$ is a vector which incorporates values of soil parameters, i.e., cohesion, internal friction angle, and maximum shear modulus. y is a vector of observed displacements. This study uses the polynomial kernel to develop the SVR model.

The SVR model is data-driven. Hence, a known dataset was required to train the SVR model. To generate a known dataset, soil parameters were considered as random variables with specified mean and standard deviation. Several realisations of these random variables were generated using the Latin hypercube method. For each realisation of soil parameters, numerical model using the finite difference method was developed to calculate displacement at a point on the slope surface. The calculated slope displacement and the corresponding realisation of soil parameters formed the training dataset. The SVR model was trained using this training dataset to further use the trained SVR model to predict slope displacement for an unknown set of soil parameters.

2.2.2 Back analysis Based on Bayes Theorem and the MCMC Simulation

Bayes theorem utilises the observed information along with the prior knowledge of input parameters to update the posterior distribution of input parameters. This study uses the Bayes theorem for back analysis of seismically induced slope displacement. The estimation of slope displacement involves uncertainties, defined as the difference between observed displacement and predicted displacement by the model.

$$y=h(\theta)+\varepsilon \tag{7}$$

where y represents the observed slope displacement. $h(\theta)$ denotes the predicted value of slope displacement by the SVR model. θ and ε represent vectors of uncertain input parameters and model correction factor, respectively. The model correction factor (ε) characterises the model and measurement uncertainty involved in the slope stability model. ε is assumed to follow normal distribution and have mean μ_ε and standard deviation σ_ε. The likelihood function (L) is defined as the conditional probability density function (pdf) of θ for known μ_ε and σ_ε; and the observed value of slope displacement (Y).

$$L(\theta/(y=Y)) = \Phi\left(\frac{h(\theta)+\mu_\varepsilon - Y}{\sigma_\varepsilon}\right) \quad (8)$$

where Φ is the pdf of standard normal variable. The likelihood function represents the chance of the predicted slope displacement being equal to the observed slope displacement.

According to Bayes theorem, the posterior distribution of θ is defined as

$$f(\theta/y) = k\Phi\left(\frac{h(\theta)+\mu_\varepsilon - Y}{\sigma_\varepsilon}\right)f(\theta) \quad (9)$$

where k is a normalisation constant. $f(\theta)$ represents the distribution of prior knowledge of uncertain input parameters. Hence, the updated distribution of uncertain input parameters can be obtained by introducing the likelihood function and solving Eq. (9). The updated distribution (or posterior distribution) reflects the change in the input parameters, which led to the observed slope displacement.

The present study used the MCMC simulation to determine the posterior distribution using Eq. (9). The MCMC simulation draws samples from a random distribution. The drawn sample in the MCMC simulation depends on the previously drawn sample, which leads to the formation of the Markov chain. These samples finally converge to a given target distribution. Several algorithms, i.e., Gibbs sampler, Metropolis–Hastings algorithm, Slice sampler etc., are used to build Markov chains. The Metropolis–Hastings algorithm is the most used algorithm as it does not require knowledge of the shape of the target distribution.

The Metropolis–Hastings algorithm can be described as the following steps.

1. Initialise θ_0 as the starting point of the Markov chain. It can be randomly chosen from the starting prior distribution.
2. Sample a candidate state θ^* from the proposal distribution or jumping distribution $g(\theta^*/\theta^{i-1})$. θ^* is dependent on the previously drawn sample θ^{i-1}.
3. Draw a sample u from the uniform distribution $U(0,1)$.
4. Estimate the probability of acceptance of Θ^*

$$A\left(\theta^*/\theta^{i-1}\right) = \min\left(1, \frac{P(\theta^*)g(\theta^*/\theta^{i-1})}{P(\theta^{i-1})g(\theta^{i-1}/\theta^*)}\right) \quad (10)$$

where $P(\Theta)$ is the target density function and is defined as

$$P(\theta) = \Phi\left(\frac{h(\theta)+\mu_\varepsilon - Y}{\sigma_\varepsilon}\right)f(\theta) \quad (11)$$

5. If $u \le A(\theta^*/\theta^{i-1})$, accept the candidate state θ^* and assign $\theta^i = \theta^*$, otherwise reject the candidate state and assign $\theta^i = \theta^{i-1}$.
6. Repeat all the previously described steps to obtain the required number of samples.

The abovementioned Metropolis–Hastings algorithm is governed by two factors: the selected proposal distribution and the burn-in period. The burn-in period is the required number of runs for Markov chain to reach a stationary density, making the candidate state follow the target distribution. In the present study, the proposal distribution is assumed to follow the normal distribution, and the probability of move is estimated using Eq. (11).

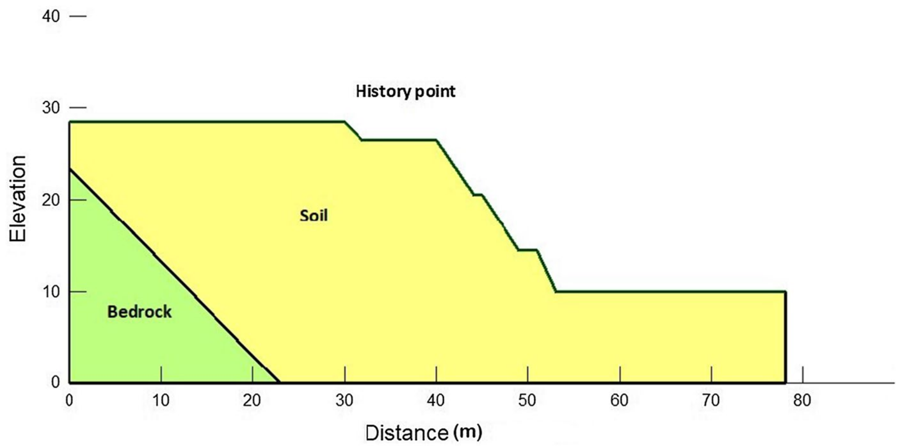
3 Case Study

The Noto Hanto earthquake (magnitude 6.9 on the Richter scale) in 2007 led to many slope failures along the Noto Yuryo road, Japan (Hamada et al. 2007). One of these slope failures is considered as a case study for the proposed methodology. The input parameters for slope simulation, such as slope geometry and soil properties, are taken from the previous study by Lu et al. (2015) (Table 1). The slope geometry is shown in Fig. 1. The slope is formed of two materials, i.e., overlain soil and bedrock. The overlain material is sandy loamy soil. This benched slope is 28.5 m in height and 78 m in length. The acceleration time history for the Noto Hanto earthquake, 2007 was obtained from the CESMD database (Center

Table 1 Properties of the slope soil mass

Sr. No	Properties	Soil	Bedrock
1	Unit weight (kN/m^3)	17.5	20
2	Cohesion(kPa)	6.8	8
3	Internal friction angle	28.5	30
4	Shear Modulus (MPa)	30.5	174.5
5	Poisson's ratio	0.3	0.35

Fig. 1 The outline of slope geometry considered for case study



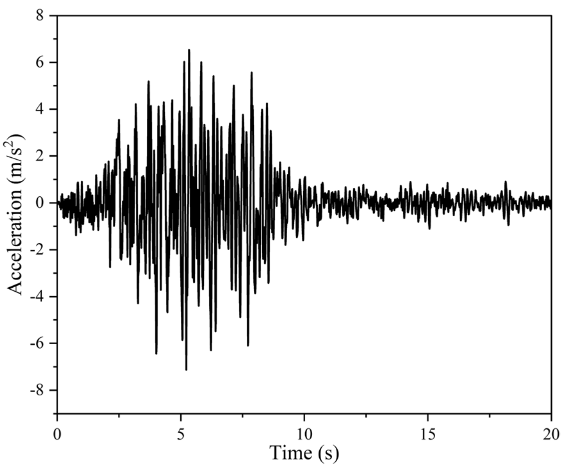


Fig. 2 Acceleration time history of applied at the base of the slope

for strong motion data by USGS, CGS, and ANSS) https://strongmotioncenter.org (Fig. 2). This earthquake caused permanent deformation of soil slope.

3.1 Numerical Simulation of Slope Displacement

The slope was simulated using 2D non-linear dynamic analysis in finite difference code FLAC 2D. The permanent horizontal slope displacement was observed at a point on the surface of the slope (Fig. 1). The acceleration time history was baseline corrected and applied at the base of the slope. The frequency content of the acceleration time history governs the mesh size of the model, as the mesh should facilitate accurate wave transmission. The mesh size was chosen for the overlain material having element sizes of 1–2.08 to 2.85–2 m. A different element size of 2.85–2 m was adopted for bedrock material. The free field boundary condition was applied along the left and right boundary of the model. The quiet boundary condition was applied at the base of the model to decrease the effect of reflected waves.

The Mohr–Coulomb elastoplastic model was utilised in dynamic analysis to model the slope material. The non-associative flow rule was used for shear failure, and the associative flow rule was utilised for tension failure. Rayleigh damping was provided to take into account the energy dissipation. The shear modulus reduction and damping ratio variation with shear strain was considered from Seed and Idriss (1970) (Fig. 3). Static analysis was conducted prior to dynamic analysis to establish the initial stress condition in the model. The safety factor was estimated for the slope prior to earthquake loading. The calculated safety factor was 1.03, consistent with the safety factor of 1.06 calculated by Lu et al. (2015).

The acceleration time history was applied at the base of the slope for dynamic analysis. The slope displacement was noted at a point on the slope surface as the output or response of the dynamic analysis. The noted slope displacement using dynamic analysis was 0.0429 m, comparable to the slope displacement of 0.045 m calculated by Lu et al. (2015). The state of stress after dynamic loading is shown in Fig. 4. Hence, the developed dynamic model accurately simulated the seismically induced slope displacement and can be used for further analysis.

For soil slopes under seismic loading, the slope material does not behave as a rigid block and deforms continuously with seismic loading. It may act as

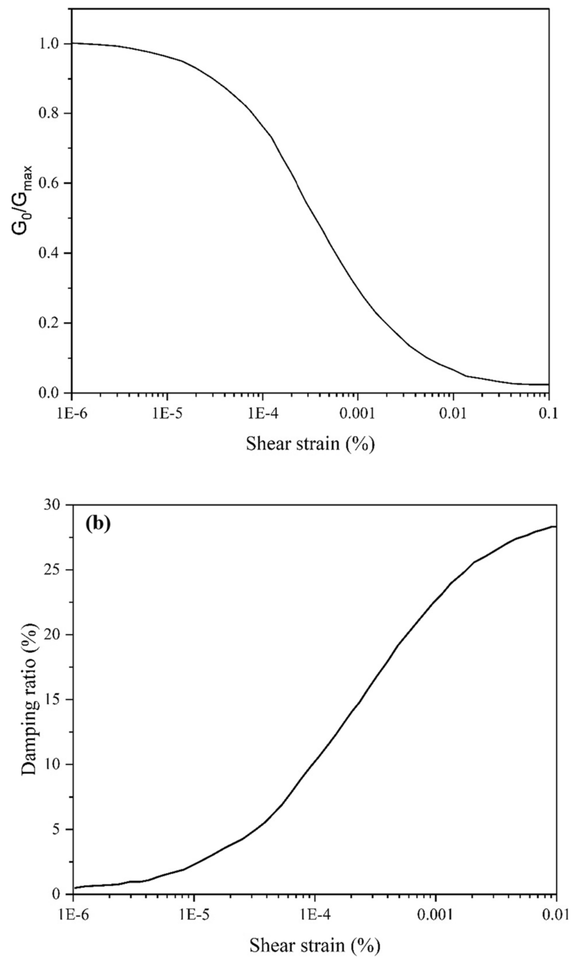


Fig. 3 **a** Shear modulus reduction (G_0/ G_{max}) with shear strain and **b** damping ratio variation with shear strain for sandy loamy soils (Seed and Idriss 1970). G_{max} and G_0 represent maximum shear modulus and small strain shear modulus, respectively

debris flow due to the change in soil shear strength parameters under seismic loading (Kokusho and Ishizawa 2007). Hence, soil shear strength and stiffness parameters (i.e., cohesion, internal friction angle, and maximum shear modulus) are considered for probabilistic back analysis of seismically induced slope displacement.

3.2 Mapping of Input and Output Parameters

The displacement back analysis is an effective approach for accurately estimating soil parameters and associated uncertainties as the soil is a non-homogeneous and anisotropic material. In the present study, Bayes theorem-based MCMC simulation was utilised to conduct displacement back analysis, which required mapping uncertain soil parameters to permanent slope displacement. The relationship between uncertain soil parameters and permanent slope displacement was established using the SVR model, which involves training the SVR model. The overlain soil material parameters were considered as input parameters for displacement back analysis as the deformation occurred in the overlain soil material. Li et al. (2016) generated 50 input parameters to train the machine learning model for displacement back analysis. In the present study, 40 sets of input parameters (cohesion, internal friction angle, and maximum shear modulus) were generated using the Latin hypercube method, considering the values of the coefficient of variation (COV) summarised in Table 2. The COV of cohesion and friction were taken from Sivakumar Babu and Murthy (2005). The COV of maximum shear modulus was taken from Tran et al. (2020). The slope displacement for each set of input parameters was estimated using the developed numerical dynamic model.

The dataset was assembled, incorporating input parameters and the corresponding permanent horizontal slope displacement, calculated by non-linear dynamic analysis. This dataset was utilised for training the SVR model using Eq. (6). To validate the trained SVR model, additional 20 sets of input parameters were also generated, and the corresponding permanent horizontal slope displacement was estimated using the developed dynamic model. This additional dataset (called the testing dataset) was used to determine the accuracy of the trained SVR model. The accuracy of the trained SVR model was determined using the regression error characteristic (REC) curve (proposed by Bi and Bennett (2003)). The REC curves were plotted for training and testing datasets (Fig. 5). The area between the REC curve and the left Y-axis indicates the mean absolute error, which was very small (1.15e-4). The prediction loss for the trained SVR model was also estimated and found as 2.99e-08, which was also very small.

The predicted slope displacement by the trained SVR model was also compared with the calculated slope displacement by the dynamic model (Fig. 6). It was observed that the testing samples shows more variability as compared to training samples. However,

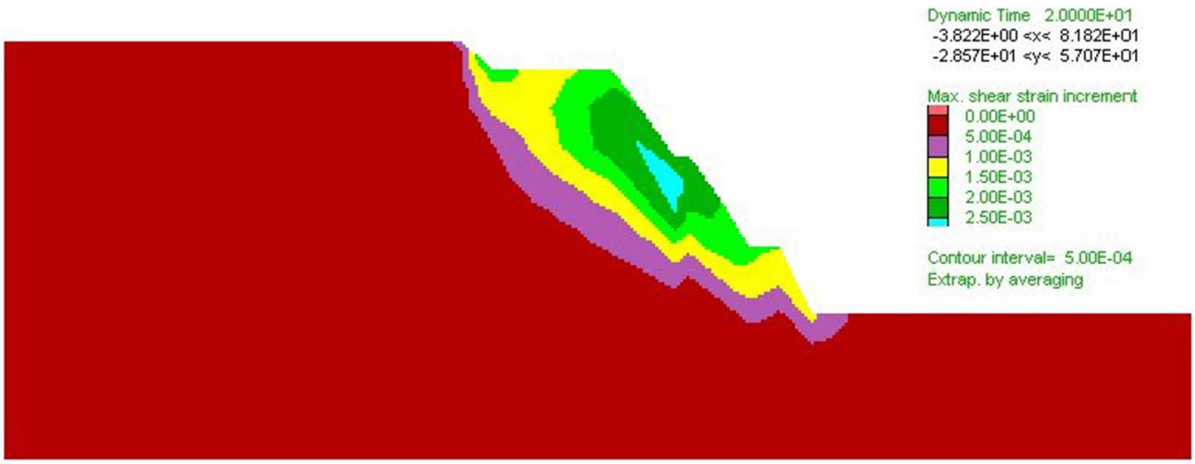


Fig. 4 Shear strain response of slope after the earthquake

Table 2 Prior statistics of soil parameters

Sr no	Parameter	Mean value	COV (%)
1	Cohesion	6.8 kPa	10
2	Internal friction angle	28.5	10
3	Shear Modulus	30.5 MPa	52

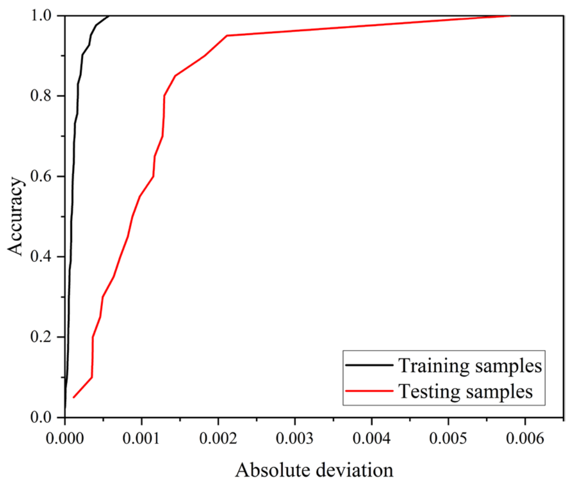


Fig. 5 Regression error characteristics (REC) curve for the trained SVM model

the mean absolute error and prediction loss were very small for both training as well as testing datasets, which suggests that the trained SVM model efficiently predicted slope displacement values. Hence, it

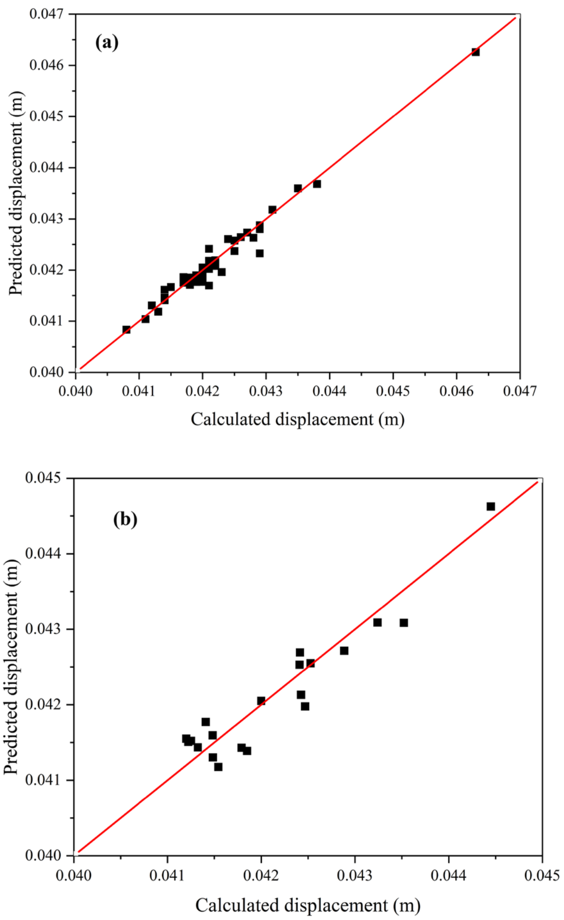


Fig. 6 Comparison of predicted displacement by SVM model and calculated displacement by slope simulation for, **a** training samples and **b** testing samples

can be used to further estimate updated statistics of soil parameters using the MCMC simulation.

4 Results

4.1 Results of MCMC Simulation

In the present study, the MCMC simulation based on the Bayes theorem was conducted to determine posterior statistics of soil parameters. The selection of the proposal or jumping distribution was one of the crucial factors in the MCMC simulation (as mentioned in step 2 of the Metropolis–Hastings algorithm). A multivariate lognormal distribution was chosen as the jumping distribution in this study. The current point in the Markov chain acted as the mean point of the proposal distribution. The covariance of the proposal distribution was equal to ζC_{Θ}. ζ and C_{Θ} represent the scaling factor and covariance of uncertain input parameters (cohesion, internal friction angle, and maximum shear modulus), respectively.

The scaling factor (ζ) governs the movement speed of the Markov chain from one side to the other side of the posterior space. The scaling function should be selected so that the Markov chain does not spend too much time standing still and should not take exceptionally long to move from one side to the other side of the posterior space. The scaling function also controls the acceptance ratio, defined as the ratio of accepted samples and total samples. Gelman et al. (1995) suggested that an acceptance ratio of 20 to 40% is sufficient for the Markov chain to move actively in the posterior space. In the present study, the scaling factor was chosen as 2, which provided acceptance ratio of 29%.

The model correction factor (ε) accounts for uncertainties involved in the developed model. The mean of ε was assumed to be zero with an unknown standard deviation (σ_{ε}). Since σ_{ε} was not known, it was treated as an additional unknown in the Bayesian analysis with weakly informative prior distribution. It was inferred jointly with the parameters of the slope. In this study, this weakly informative distribution was selected as uniform distribution with a lower value equal to zero and a higher value equal to the observed displacement of the slope. In other words, the COV of the model uncertainty ranged from a minimum of zero to a maximum of 100%. This range of COV incorporated most of the model uncertainties estimated using field data. In this study, thirty Markov chains with different initial states were run to determine the posterior statistics of input parameters. With the first half of samples rejected as it was assumed to belong to the burn-in phase, it was observed that the estimated posterior statistics were similar for all chains for large number of drawn samples. The posterior statistics for different numbers of drawn samples are summarised in Table 3. It was noted that the variability of posterior statistics for different Markov chains was minimal. Hence, 100,000 samples were enough to determine robust posterior statistics. The prior and posterior distribution of input parameters were plotted in Fig. 7. The standard deviation of

Table 3 Posterior statistics of input parameters for different number of drawn samples

Sr. no	Parameter	No. of samples	Mean value	Standard deviation
1	Cohesion (kPa)	10,000	6.83	0.69
		20,000	6.81	0.68
		50,000	6.71	0.66
		100,000	6.70	0.66
2	Internal friction angle (degree)	10,000	30.35	2.80
		20,000	28.52	2.78
		50,000	27.74	2.76
		100,000	27.51	2.75
3	Shear Modulus (MPa)	10,000	29.87	14.86
		20,000	28.51	14.68
		50,000	24.91	14.19
		100,000	24.56	14.15

cohesion and internal friction angle is slightly reduced by 3% and 1.78%, respectively. The mean of cohesion and internal friction angle is slightly reduced. At the same time, the mean of maximum shear modulus is decreased considerably (Table 4), and the standard deviation of maximum shear modulus is reduced by 10.62%. Hence, maximum shear modulus governs the mechanism of seismic slope displacement as the statistics of maximum shear modulus showed the maximum variation after back analysis.

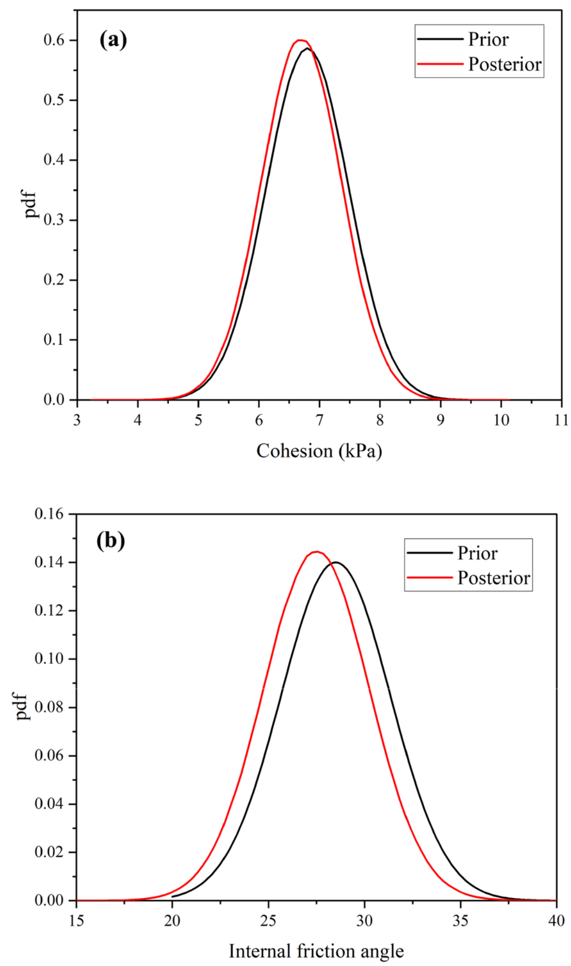


Fig. 7 Prior and Posterior distribution of soil parameters, **a** cohesion, **b** internal friction angle, and **c** maximum shear modulus

Garevski et al. (2013) proposed an advanced methodology to estimate seismic slope instability and evaluated the effect of different uncertain parameters on slope stability using sensitivity analysis. They concluded that the shear modulus influences the slope deformation significantly. Kumar et al. (2017) studied the dynamic behaviour of soil at high strains using experimental study. They concluded that strain level or soil deformation significantly affects the shear modulus. The results of this study are consistent with the findings of these previous studies as statistics of maximum shear modulus are changed considerably after updating using MCMC simulation in this study, which suggests that the permanent slope displacement is more influenced by shear modulus than cohesion and internal friction angle.

The updated values of shear strength parameters are further validated by simulating the slope under seismic loading using updated values of these parameters. The obtained value of slope displacement is 0.045 m, which is very close to the obtained slope displacement by Lu et al. (2015). Hence, the obtained values of shear strength parameters are successfully validated and can be used to plan remedial measures. These values can also be further utilised for the reliability-based design of slopes.

Several studies are conducted to determine the posterior statistics of input parameters using MCMC simulation based on the Bayes theorem. These studies consider model uncertainties as the difference between actual performance and model prediction. The model uncertainties, considered using the model correction factor (ε), affect the estimation of posterior statistics using MCMC simulation. The model

Table 4 Prior and posterior statistics of soil parameters

	Parameters	Cohesion	Internal friction angle	Maximum shear modulus
Prior	Mean	6.8 kPa	28.5 degree	30.5 MPa
	Standard deviation	0.68	2.85	15.86
Posterior	Mean	6.7 kPa	27.5 degree	24.5 MPa
	Standard deviation	0.65	2.75	14.15

uncertainties are considered normally distributed in the present study. According to previous studies, there is limited research on the effect of model uncertainties on posterior statistics. Hence, in the present study, the standard deviation of the model correction factor (ε) is considered a variable to investigate its influence on posterior statistics. This knowledge will facilitate a better understanding of the back analysed model correction factor and its impact on the posterior distribution of parameters.

4.2 Effect of Model Uncertainties on Posterior Statistics

To observe the influence of model uncertainty on the posterior characteristics of soil parameters, the back analysis was performed for permanent slope displacement considering different variability of model uncertainty. For simplicity, the model uncertainty was considered to have zero mean, and the investigation was conducted for different values of fixed and known values of standard deviation. In this parametric study, the standard deviation of the model uncertainty was varied in multiples of the observed slope displacement (i.e., displacement/5, displacement/4, displacement/3, displacement/2, displacement, displacement × 2, displacement × 3, displacement × 4, displacement × 5), i.e., COV ranging from 20 to 500%.

Figure 8 shows the posterior mean of the cohesion with increasing COV of the model uncertainty. It can be seen that the mean value decreases exponentially with an increase in COV of model uncertainty till a value of 200% is reached, after which it remains constant. However, the posterior standard deviation of cohesion first decreases, attaining a minimum at COV of 33%, after which it increases to COV of 200% and remains approximately constant, as evident from Fig. 8. The posterior mean of internal friction angle of soil decreases exponentially with increasing COV of model uncertainty (Fig. 9), showing similar behaviour to that of cohesion. However, the standard deviation of internal friction angle shows an exponential increase and subsequently constant value with increasing COV of model uncertainty (Fig. 9). Similar behaviour of shear modulus is observed, i.e., exponential decrease and subsequently constant value at COV of 200% (Fig. 10). The posterior standard deviation of shear modulus decreases exponentially to

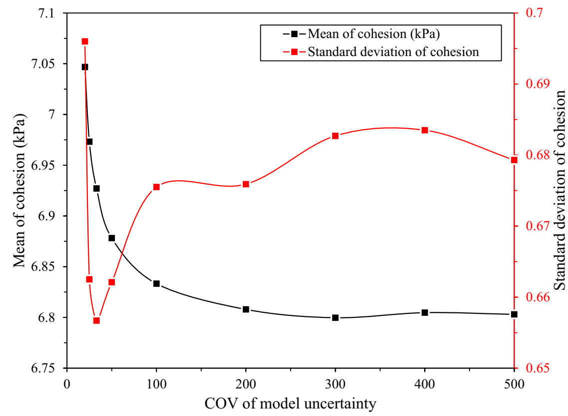


Fig. 8 Variation of mean and standard deviation of cohesion with COV of model uncertainty

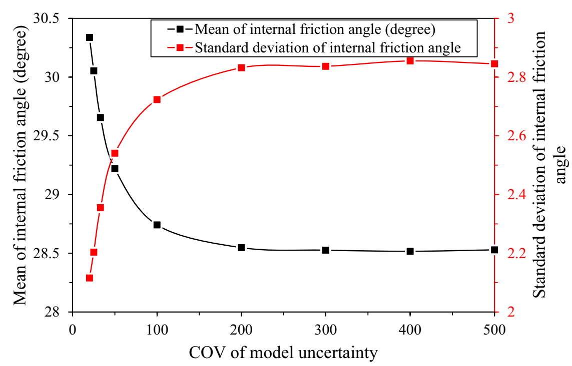


Fig. 9 Variation of mean and standard deviation of internal friction angle with COV of model uncertainty

reach a minimum at model uncertainty COV of 33% and marginally increased to achieve a constant value at COV of 200% (Fig. 10). The variation of calculated displacement using the updated values of soil parameters considering different values of COV of model uncertainty is also plotted in Fig. 11. It was noted from Fig. 11 that the noted displacement did not show a specific pattern with increasing standard COV of model uncertainty model uncertainty is a dynamic parameter and controlled by multiple factors. The variation shown in Fig. 11 is specific to the case study considered in the present study.

It should be noted that the posterior COV of soil parameters also shows a similar variation as the standard deviation of the posteriors. The posterior standard deviation of cohesion and internal friction angle becomes approximately equal to the prior

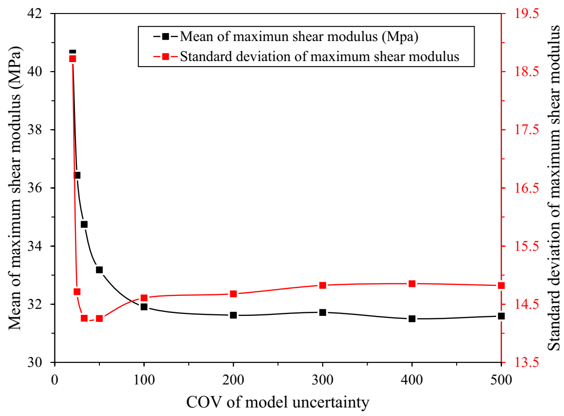


Fig. 10 Variation of mean and standard deviation of maximum shear modulus with COV of model uncertainty

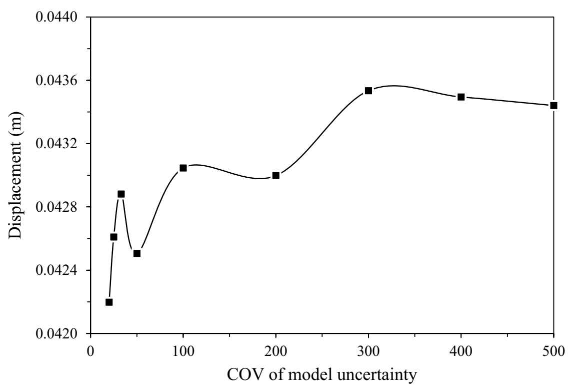


Fig. 11 The variation of calculated displacement using the updated values of soil parameters considering different values of standard deviation of model uncertainty

standard deviation after 300% COV of model uncertainty. However, the posterior standard deviation of the shear modulus is less than the prior standard deviation for higher COV.

The parametric study results suggest that the mean value of all soil parameters decreases with increased variability of model uncertainty. The variability of soil parameters has shown different patterns with the increase in the variability of model uncertainty. As the variability or standard deviation of model uncertainty increases, the observed information regarding the output becomes less informative, resulting in less efficacy in updating the variability of soil parameters (Zhang et al. 2009). This might be the reason for the different behaviour of standard deviation of soil parameters with model uncertainty.

5 Concluding Remarks

The estimation of soil parameters in field conditions is a challenging task due to inherent uncertainties and the natural variability of the soil. Further, this task becomes more tedious for slopes under seismic conditions due to dynamic loading. To this aim, a framework for probabilistic back analysis is proposed to determine the soil parameters with inherent uncertainties for slope systems under seismic loading. The proposed methodology adopts Markov Chain Monte Carlo (MCMC) simulation based on the Bayes theorem. This method utilises the SVR algorithm to map the relationship between uncertain soil parameters and slope displacement. The SVR algorithm was trained using the data obtained by simulation of slope under seismic loading using the FDM model, considering input parameters as random variables. The efficiency of the trained SVR model was determined using the REC curve. The obtained SVR model was utilised in probabilistic back analysis using the Bayes theorem to update the statistics of uncertain input parameters. The MCMC simulation was adopted to draw samples from updated input parameters.

The results of probabilistic back analysis provide updated statistics of soil parameters. The variability of cohesion and internal friction angle is slightly reduced after updating using the Bayes theorem. The variability of maximum shear modulus is considerably reduced. The same pattern is observed for the mean of these parameters, i.e., the slight reduction in mean of cohesion and internal friction angle; and considerable reduction in mean of maximum shear modulus. Hence, the results of the study suggest that maximum shear modulus is greatly influenced by deformation or induced strain under seismic loading, which is consistent with the results of the previous studies conducted by Garevski et al. (2013) and Kumar et al. (2017). However, the numerical simulations utilised in the present study incorporate model uncertainty. The estimation of the model uncertainty is a challenging task as it depends on the values of other uncertainties, such as testing and measurement uncertainty. However, the decisions made for geotechnical systems without considering model uncertainties could be biased. Hence, understanding the model uncertainty for numerical simulations is necessary to make appropriate decisions for geotechnical systems.

In the present study, the effect of model uncertainties on the updated statistics of input parameters is determined by conducting a parametric study. The parametric study involves probabilistic back analysis considering a constant value of zero for the mean of model error and the variable standard deviation of model uncertainty. The standard deviation of model error ranges from 20 to 500% of slope displacement in the field. The mean of all the soil parameters (cohesion, internal friction angle, and maximum shear modulus) showed similar behaviour, i.e., exponential decrease till a certain value of standard deviation and then constant value. On the other hand, the standard deviation of cohesion and maximum shear modulus first decreased until a certain standard deviation of model uncertainty and then increased. The variability of internal friction angle was increased exponentially. The results of the parametric study facilitate the understanding of model uncertainty for seismically induced slope displacement.

Author contribution Dr. Himanshu Rana: Conceptualization, methodology, code writing, data processing, writing- original draft. Dr. Bhardwaj Pandit: Conceptualization, code writing, data processing, writing- original draft. Prof. G L Sivakumar Babu: supervision, conceptualization, writing- review.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data Availability statement The data, models and code used or generated for this study are available with the corresponding author on request.

Declarations

Competing interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abbaszadeh M, Shahriar K, Sharifzadeh M, Heydari M (2011) Uncertainty and reliability analysis applied to slope stability: a case study from Sungun copper mine. Geotech Geol Eng 29:581–596. https://doi.org/10.1007/s10706-011-9405-1

Abdulai M, Sharifzadeh M (2019) Uncertainty and reliability analysis of open pit rock slopes: a critical review of methods of analysis. Geotech Geol Eng 37:1223–1247. https://doi.org/10.1007/s10706-018-0680-y

Alitabar A, Noorzad R, Qolinia A (2021) Evaluation of the instability risk of the dam slopes simulated with monte Carlo method (case study: Alborz Dam). Geotech Geol Eng 39:4237–4251. https://doi.org/10.1007/s10706-021-01755-6

Arab Khaburi M, Mortazavi A (2019) Slope stability analysis of sarcheshmeh copper mine west wall under seismic loads. Geotech Geol Eng 37:3141–3155. https://doi.org/10.1007/s10706-019-00830-3

Bi J, Bennett KP (2003) Regression error characteristic curves. In: International conference on machine learning. Washington

Bray JD, Macedo J, Travasarou T (2018) Simplified procedure for estimating seismic slope displacements for subduction zone earthquakes. J Geotech Geoenviron Eng. https://doi.org/10.1061/(ASCE)GT.1943

Cheng Y, He D (2020) Slope reliability analysis considering variability of shear strength parameters. Geotech Geol Eng 38:4361–4368. https://doi.org/10.1007/s10706-020-01266-w

Cheung RWM, Tang WH (2005) Realistic assessment of slope reliability for effective landslide hazard management. Geotechnique 55:85–94

Contreras LF, Brown ET (2019) Slope reliability and back analysis of failure with geotechnical parameters estimated using Bayesian inference. J Rock Mech Geotech Eng 11:628–643. https://doi.org/10.1016/j.jrmge.2018.11.008

Dilip DM, Sivakumar Babu GL (2013) Methodology for pavement design reliability and back analysis using Markov chain monte Carlo simulation. J Transp Eng 139:65–74. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000455

Ering P, Sivakumar Babu GL (2016) Probabilistic back analysis of rainfall induced landslide—a case study of Malin landslide, India. Eng Geol 208:154–164. https://doi.org/10.1016/j.enggeo.2016.05.002

Feng XT, Zhao H, Li S (2004) A new displacement back analysis to identify mechanical geo-material parameters based on hybrid intelligent methodology. Int J Numer Anal Methods Geomech 28:1141–1165. https://doi.org/10.1002/nag.381

Garevski M, Zugic Z, Sesov V (2013) Advanced seismic slope stability analysis. Landslides 10:729–736. https://doi.org/10.1007/s10346-012-0360-6

Gelman A, Carlin JB, Stern HS, et al (1995) Bayesian data analysis, 3rd edn

Gilbert RB, Wright SG, Liedtke E (1998) Uncertainty in back analysis of slopes: Kettleman hills case history. J Geotech Geoenviron Eng 124:1167–1176

Gratchev IB, Towhata I (2011) Analysis of the mechanisms of slope failures triggered by the 2007 Chuetsu Oki earthquake. Geotech Geol Eng 29:695–708. https://doi.org/10.1007/s10706-011-9411-3

Hamada M, Aydan O, Sakomoto A (2007) A quick report on Noto Peninsula earthquake on March 25, 2017. Tokyo

Huang HW, Zhang J, Zhang L (2012) Bayesian network for characterizing model uncertainty of liquefaction potential evaluation models. KSCE J Civ Eng 16:714–722. https://doi.org/10.1007/s12205-012-1367-1

Huang S, Lv Y, Peng Y (2016) Dynamic response of sandy slope under coupling of earthquake and groundwater. Geotech Geol Eng 34:889–899. https://doi.org/10.1007/s10706-016-0014-x

Jiang SH, Huang J, Qi XH, Zhou CB (2020) Efficient probabilistic back analysis of spatially varying soil parameters for slope reliability assessment. Eng Geol. https://doi.org/10.1016/j.enggeo.2020.105597

Juang CH, Luo Z, Atamturktur S, Huang H (2013) Bayesian updating of soil parameters for braced excavations using field observations. J Geotech Geoenviron Eng 139:395–406. https://doi.org/10.1061/(asce)gt.1943-5606.0000782

Kokusho T, Ishizawa T (2007) Energy approach to earthquake-induced slope failures and its implications. J Geotech Geoenviron Eng 133:828–840. https://doi.org/10.1061/ASCE1090-02412007133:7828

Kong D, Luo Q, Zhang W et al (2022) Reliability analysis approach for railway embankment slopes using response surface method based Monte Carlo simulation. Geotech Geol Eng 40:4529–4538. https://doi.org/10.1007/s10706-022-02168-9

Kumar SS, Krishna AM, Dey A (2017) Evaluation of dynamic properties of sandy soil at high cyclic strains. Soil Dyn Earth Eng 99:157–167

Li S, Zhao H, Ru Z, Sun Q (2016) Probabilistic back analysis based on Bayesian and multi-output support vector machine for a high cut rock slope. Eng Geol 203:178–190. https://doi.org/10.1016/j.enggeo.2015.11.004

Li Z, Gong W, Li T et al (2021) Probabilistic back analysis for improved reliability of geotechnical predictions considering parameters uncertainty, model bias, and observation error. Tun under Space Tech Trenchless Tech Res. https://doi.org/10.1016/j.tust.2021.104051

Lu L, Wang ZJ, Song ML, Arai K (2015) Stability analysis of slopes with ground water during earthquakes. Eng Geol 193:288–296. https://doi.org/10.1016/j.enggeo.2015.05.001

Metya S, Bhattacharya G (2016) Reliability analysis of earth slopes considering spatial variability. Geotech Geol Eng 34:103–123. https://doi.org/10.1007/s10706-015-9932-2

Miranda T, Gomes Correia A, Ribeiro e Sousa L (2009) Bayesian methodology for updating geomechanical parameters and uncertainty quantification. Int J Rock Mech Min Sci 46:1144–1153. https://doi.org/10.1016/j.ijrmms.2009.03.008

Ng SM, Ashraf M, Ismail M, Abustan I (2014) Back analysis of slope failure using finite element with point estimate method (FEM-PEM). J Civ Eng Res 2014:31–35. https://doi.org/10.5923/c.jce.201402.04

Rana H, Sivakumar Babu GL (2022b) Evaluating soil parameters considering probabilistic back analysis for slope failures. In: Geo-Congress 2022b. American Society of Civil Engineers, pp 360–369

Rana H, Sivakumar Babu GL (2023) Reliability analysis of spatially variable soil slope using deep learning algorithm. In: Geo-congress 2023. American Society of Civil Engineers, Reston, VA, pp 553–562

Rana H, Sivakumar Babu GL (2022a) Probabilistic back analysis for rainfall-induced slope failure using MLS-SVR and Bayesian analysis. Georisk. https://doi.org/10.1080/17499518.2022.2084555

Seed B, Idriss IM (1970) Soil moduli and damping factors for dynamic response analysis. Berkeley

Sivakumar Babu GL, Murthy DSN (2005) Reliability analysis of unsaturated soil slopes. J Geotech Geoenviron Eng 131:1423–1428. https://doi.org/10.1061/ASCE1090-02412005131:111423

Tietje O, Fitze P, Schneider HR (2014) Slope stability analysis based in autocorrelated shear strength parameters. Geotech Geol Eng 32:1477–1483. https://doi.org/10.1007/s10706-013-9693-8

Tran T-T, Salman K, Han S-R, Kim D (2020) Probabilistic models for uncertainty quantification of soil properties on site response analysis. ASCE ASME J Risk Uncertain Eng Syst A Civ Eng. https://doi.org/10.1061/ajrua6.0001079

Wang Z, Li B (2021) Comparative study on the stability of hydropower station slope under different earthquake conditions. Geotech Geol Eng 39:4511–4521. https://doi.org/10.1007/s10706-021-01780-5

Wang L, Hwang JH, Luo Z et al (2013) Probabilistic back analysis of slope failure—a case study in Taiwan. Comput Geotech 51:12–23. https://doi.org/10.1016/j.compgeo.2013.01.008

Wu F, Zhang LL, Li HW (2017) Probabilistic back analysis based on polynomial chaos expansion for rainfall-induced soil slope failure. In: Geo-Risk 2017. American Society of Civil Engineers, pp 173–182

Xu F, Huang BQ, Wang K (2013) Study and application of slope displacement back analysis based on SVM-CTS. In: Applied mechanics and materials. pp 163–166

Zhang J, Zhang LM, Tang WH (2009) Bayesian framework for characterizing geotechnical model uncertainty. J Geotech Geoenviron Eng 135:932–940. https://doi.org/10.1061/ASCEGT.1943-5606.0000018

Zhang J, Tang WH, Asce HM et al (2010) Efficient probabilistic back-analysis of slope stability model parameters. J Geotech Geoenviron Eng 136:99–109. https://doi.org/10.1061/ASCEGT.1943-5606.0000205

Zhang J, Tang WH, Zhang LM, Huang HW (2012) Characterising geotechnical model uncertainty by hybrid Markov Chain Monte Carlo simulation. Comput Geotech 43:26–36. https://doi.org/10.1016/j.compgeo.2012.02.002

Zhang DM, Phoon KK, Huang HW, Hu QF (2015) Characterization of model uncertainty for cantilever deflections in

undrained clay. J Geotech Geoenviron Eng. https://doi.org/10.1061/(ASCE)GT.1943
Zhao H, bo, Yin S, (2009) Geomechanical parameters identification by particle swarm optimization and support vector machine. Appl Math Model 33:3997–4012. https://doi.org/10.1016/j.apm.2009.01.011