Feasibility of Stochastic Gradient Boosting Approach for Evaluating Seismic Liquefaction Potential Based on SPT and CPT Case Histories

Jian Zhou[1]; Enming Li[2]; Mingzheng Wang[3]; Xin Chen[4]; Xiuzhi Shi[5]; and Lishuai Jiang[6]

Abstract: Earthquakes have always attracted civil and geotechnical engineers' attention, especially when it comes to the liquefaction potential of soil. This paper investigates the feasibility of classifier based on stochastic gradient boosting (SGB) to explore the liquefaction potential from actual cone penetration test (CPT) and standard penetration test (SPT) field data. SGB is composed of many classification and regression trees which meet the mechanism of ensemble learning and show strong predictive power compared with conventional statistical learning models in several engineering applications. The binary classifier was built by the database gathered from CPT and SPT filed data for predicting the non-liquefaction or liquefaction of soil, the SGB hyperparameters are optimized by grid search method with tenfolds cross validation methods. Three performance metric, namely Cohen's Kappa coefficient, classification accuracy rate and receiver operating characteristic curve, are used to evaluate the predictive performance of SGB approaches. With CPT and SPT test sets, highest classification accuracy rate of 88.62% and 95.45%, respectively, are achieved with SGB. It is confirmed that the SGB can be applied to characterize the complex relationship between the liquefaction potential and different soil and seismic parameters with great efficiency. Further, relative importance of influencing variables for each model are investigated and demonstrated that the SGB predictor is more sensitive to the indicators of initial soil friction angle for SPT data whereas cone tip resistance for CPT data. **DOI: 10.1061/(ASCE)CF.1943-5509.0001292.**

Author keywords: Liquefaction potential; Stochastic gradient boosting; Cone penetration test (CPT); Standard penetration test (SPT).

Introduction

The phenomenon of soil liquefaction during earthquakes has extremely negative impacts on lifelines, bridges, and buildings (Seed and Idriss 1971; Lai et al. 2004; Pal 2006; Goh and Goh 2007; Hanna et al. 2007; Samui 2007; Oommen et al. 2010; Kohestani et al. 2015; Li et al. 2017; Wang et al. 2017; Zhou et al. 2019). Therefore, in order to guarantee the safety of people and property, it is imperative to estimate the probability of soil liquefaction and its effects. Also, the Cone Penetration Test (CPT) and Standard Penetration Test (SPT) have been widely and increasingly applied in many geotechnical and civil inverstigations with aspect to earthquake (Zhou et al. 2019). This benefits from the CPT and SPT which have better accuracy, operability, and reliability in comparison to other field tests. Moreover, many liquefaction triggering curves also measured from CPT and SPT have been put into practice (e.g., Moss et al. 2006; Hanna et al. 2007; Idriss and Boulanger 2006, 2014). Consequently, this study concentrates on the estimation of the seismic liquefaction potential using the CPT and SPT data.

Over the past few decades, many approaches have been investigated to predict seismic liquefaction potential of soil deposits, including field methods (e.g., Shear wave velocity based, CPT based and SPT based), analytical (empirical or semi-empirical) methods, geographic information system methods, numerical methods, laboratory methods (e.g., cyclic triaxial test and centrifuge modeling), artificial intelligence and soft computing based methods (Seed and Idriss 1971; Goh 1996; Lai et al. 2004, 2006; Idriss and Boulanger 2006; Moss et al. 2006; Pal 2006; Samui 2007; Oommen et al. 2010; Goh and Goh 2007; Hanna et al. 2007; Xue and Yang 2013; Kohestani et al. 2015; Kaveh et al. 2018; Zhou et al. 2019). However, empirical or semi-empirical approaches are limited by the local monitoring data and result in the analysis results are not representative. As for the numerical methods, it tends to be complicated to determine the reliable values of model input parameters (Jiang et al. 2019a, b). In addition, at the aspect of predicting and evaluating the buildings, it is worthy more thorough research.

Utilizing artificial intelligence techniques in this field has been recently highlighted and extensively used by several researchers to predict soil liquefaction triggered by earthquakes in literature, in particular, artificial neural network (ANN), adaptive neuro fuzzy inference system (ANFIS), relevance vector machine, and support

[1]Assistant Professor, School of Resources and Safety Engineering, Central South Univ., #932 Lushan South Rd., Changsha 410083, China; Postdoctoral Researcher, Postdoctoral Scientific Research Workstation, Shenzhen Zhongjin Lingnan Nonfemet Co., Ltd., Shenzhen 518042, China. Email: csujzhou@hotmail.com

[2]Master Student, School of Resources and Safety Engineering, Central South Univ., #932 Lushan South Rd., Changsha 410083, China. Email: lem123456@csu.edu.cn

[3]Ph.D. Candidate, Mining Innovation Rehabilitation and Applied Research Corporation—Mining Innovation, Laurentian Univ., Sudbury, Canada P3E 2C6. Email: mz_wang_rock@outlook.com

[4]Ph.D. Candidate, School of Resources and Safety Engineering, Central South Univ., #932 Lushan South Rd., Changsha 410083, China. Email: chenxin@csu.edu.cn

[5]Professor, School of Resources and Safety Engineering, Central South Univ., #932 Lushan South Rd., Changsha 410083, China. Email: shixiuzhi@263.net

[6]Assistant Professor, State Key Laboratory of Mining Disaster Prevention and Control, Shandong Univ. of Science and Technology, Qingdao 266590, China (corresponding author). Email: lsjiang@sdust.edu.cn

Note. This manuscript was submitted on March 30, 2018; approved on October 17, 2018; published online on February 23, 2019. Discussion period open until July 23, 2019; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Performance of Constructed Facilities***, © ASCE, ISSN 0887-3828.

vector machine (SVM) have been successfully employed to assess liquefaction potential with more precise accuracy in comparison to some traditional statistical machine learning methods in the past few decades (Zhou et al. 2019). For example, Pal (2006) makes use of the SVM-based classifier to convey the potential of soil liquefaction from actual SPT and CPT data. Yazdi et al. (2012) combine ANFIS and a support vector data description to correct the class imbalance problem based on in-situ liquefaction data. In another work, Xue and Yang (2013) also applied the ANFIS to evaluate soil liquefaction based on CPT data. Kohestani et al. (2015) utilizes a random forest algorithm to detect the liquefaction potential of soil from CPT data. Kaveh et al. (2018) employed the patient ruled-induction method to unearth the potential of liquefaction occurrence. Recently, Hoang and Bui (2018) developed a novel soft computing model with a hierarchical structure for evaluating soil liquefaction triggered by earthquake combining a hybridization of kernel Fisher discriminant analysis and a least squares SVM. The aforementioned methods have achieved relative satisfactory results, but the major disadvantage of these methods is obvious. For example, both ANN and ANFIS model are seriously time-consuming during the process of building the membership functions and the rule base, especially the determination of ANN architecture is a difficult task because neither direction nor analytical method is available. SVM is difficult to determine the insensitivity parameter and the penalty weight parameters. Moreover, cross-validation (CV) process was not carried out and thus the overall predictive performance of these models are not well interpreted but far from proffering convincing results (e.g., Goh 1996; Pal 2006; Samui 2007; Kaveh et al. 2018; Hanna et al. 2007; Hoang and Bui 2018; Zhou et al. 2019). Therefore, it is necessary to advance the idea that making more systematic and in-depth research on predicting the liquefaction potential triggered by earthquakes.

In contrast to the aforementioned methods, supervised machine learning techniques such as the stochastic gradient boosting (SGB) (Friedman 2001, 2002; Kuhn and Johnson 2013) has received growing attention in different domains, dispensing with the assumptions of statistics, the SGB can rapidly recognize the nonlinear relationship between inputs and outputs and has provided valuable results for prediction performance. The SGB algorithm has promoted the development of many civil and mining subjects, such as the hardrock pillar stability (Zhou et al. 2015), rockbursts in hardrock (Zhou et al. 2016a, b), blast-induced residential structure damage prediction (Zhou et al. 2016c), hangingwall stability prediction (Qi et al. 2017); and energy consumption of commercial buildings (Touzani et al. 2018). But yet no study has been recorded or reported which employ the SGB algorithm to investigate the soil liquefaction triggered by earthquakes. To verify the feasibility of using the classification of the SGB model to predict liquefaction triggered by earthquakes in soil deposits, the CPT and SPT databases are used for developing the SGB model after the algorithm and procedure of SGB is generalized reviewed. Finally, the predictive performance of the SGB model is investigated.

Materials and Methods

Data Set and Predictor Variables

To achieve this objective, two datasets collected from the previous work are applied to evaluate the feasibility of the SGB model in the present study.

1. SPT database. Hanna et al. (2007) summarized a number of 620 real cases of soil liquefaction observations with SPT test. The available parameters are extracted from the outcomes of field tests from the earthquakes occurred in China and Turkey. Twelve soil and seismic parameters as available have been presented for liquefaction/non-liquefaction prediction, which includes the values of the following independent variables: M_{w}: earthquake magnitude (Richter); PHA or amax: peak horizontal acceleration at ground surface (g); Z: soil specimen depth (m); $N1$: corrected PST number (%); F_{75}: percent fines less than 75 μm (%); d_w: ground water table depth (m); V_s: shear wave velocity (m/s); φ': initial soil friction angle (degrees); $\mathrm{S_{vo}}$: total vertical stress (kPa); $\mathrm{S_{1vo}}$: effective vertical stress (kPa); $\tau_{\mathrm{av}}/\mathrm{S_{1vo}}$: cyclic stress ratio (CSR), which illustrates the seismic demand on soil; a_t: threshold acceleration, which contributes to the strain-based procedure and along with strain-based safety factor indicates the exceedance of the threshold strain (g). Note that the numbers of occurrence or non-occurrence of liquefaction cases are 256 and 364 in SPT database, respectively. Readers interested in more details of these cases can find worthwhile information in Hanna et al. (2007).
2. CPT database. Goh and Goh (2007) compiled a database from published literature which consisted of CPT-based 226 case studies collected from over 52 sites and field performance observations of six different earthquakes all over the world (Zhou et al. 2019), such as Chi-Chi earthquake in 1999, Haicheng earthquake in 1975, Imperial Valley earthquake in 1979 and 1981, Loma Prieta earthquake in 1989 and San Fernando Valley earthquake in 1971. The six input variables used for liquefaction potential assessment are identified namely M_w (Richter), $a_{\max}$ (g), the sleeve friction ratio ($\mathrm{R_f}$, %), the effective stress ($\mathrm{S_1}$, kPa), the total stress ($\mathrm{S_2}$, kPa), and the cone tip resistance (q_{c}, MPa). The numbers of non-liquefied and liquefied records are 133 and 93 in CPT database, respectively. More detailed information about these case histories can refer to ref. Juang et al. (2003) and Goh and Goh (2007).

Details of Stochastic Gradient Boosting Approach

SGB proposed by Friedman (2001, 2002), as a supervised machine learning approach, is considered promising and can achieve acceptable results evidenced from the reference Lu et al. (2016), Zhou et al. (2016a, c), Touzani et al. (2018), and Zhou et al. (2019). The two problem types for which SGB is used are regression and classification problems (Friedman 2001). In this work, the binary dependent variable of SPT and CPT database is liquefaction (L), and No for nonliquefied soils and Yes for liquefied soils, thus prediction of seismic liquefaction potential belongs to the binary classification problem. In this regard, the main idea behind SGB using binary classification is briefly introduced, in which a scalar score function is formed to distinguish the two labels. Suppose $X = \{\boldsymbol{x}_i\}_{i=1}^N$ with $\boldsymbol{x}_i \in R^D$ is a training set, and their labels $Y = \{y_i\}_{i=1}^N$ with $y_i \in \{0, 1\}$, which represents nonliquefied soils and liquefied soils in this work, the objective is to choose a classification function $\zeta(\boldsymbol{x})$ to minimize the aggregation of some specified loss function $\mathcal{L}(y_i, \zeta(\boldsymbol{x}_i))$ (Hastie et al. 2001):

$$\zeta^* = \arg\min_{\zeta} \sum_{i=1}^{N} \mathcal{L}(y_i, \zeta(\boldsymbol{x}_i)), i = 1, 2, \ldots, N \tag{1}$$

Gradient boosting considers the function estimation F in an additive form:

$$\zeta(\boldsymbol{x}) = \sum_{m=1}^{T} \xi_m(\boldsymbol{x}), \qquad m = 1, 2, \ldots, T \tag{2}$$

where the $\{\xi_m(\boldsymbol{x})\}$ are designed in an incremental fashion; T = the number of iterations; at the m-th stage, the newly added function;

ξ_m is chosen to optimize the aggregated loss while keeping the $\{\xi_j\}_{j=1}^{m-1}$ fixed.

Each function ξ_m belongs to a set of parametrized base-learners (BLS) (Friedman 2001). Let θ denotes the vector of parameters of the BLS. SGB uses decision trees to be the BLS. For this choice, θ consists of parameters that represent the tree structure (i.e., the threshold for splitting each node, the feature to split in each internal node).

An approximate function of the loss can be formed by Eq. (3) at stage m:

$$\mathcal{L}(y_i, \xi_m(\boldsymbol{x}_i) + \zeta_{m-1}(\boldsymbol{x}_i)) \approx 0.5\xi_m(\boldsymbol{x}_i)^2 + g_i\xi_m(\boldsymbol{x}_i) + \mathcal{L}(y_i, \zeta_{m-1}(\boldsymbol{x}_i)) \tag{3}$$

where $\zeta_{m-1}(\boldsymbol{x}_i) = \sum_{j=1}^{m-1} \xi_j(\boldsymbol{x}_i)$ and $g_i = \frac{\partial\mathcal{L}(y_i, \zeta(\boldsymbol{x}_i))}{\partial\zeta(\boldsymbol{x}_i)}\Big|_{\zeta(\boldsymbol{x}_i)=\zeta_{m-1}(\boldsymbol{x}_i)}$.

Get the estimate ξ_m by minimizing the right part of Eq. (3), it can be represented as the following minimization problem:

$$\arg\min_{\xi_m} \sum_{i=1}^{N} 0.5(\xi_m(\boldsymbol{x}_i) - g_i)^2 \tag{4}$$

Usually, a suitable step size (shrinkage parameter) is applied to ξ_m before it is added to ζ_{m-1}, then update ξ_m, and output $\zeta(\boldsymbol{x})$. To establish the model between soil liquefaction triggered by earthquakes and influential indicators, the SGB model was conducted by the gbm R-package (Ridgeway 2007) within R environment (R Core Team 2017). More detailed description about SGB can be found from the bibliography. (Friedman 2001; Hastie et al. 2001; Kuhn and Johnson 2013) and in the help and vignettes of the cited R packages.

Optimizing the SGB Hyperparameters by Grid Search Method with Cross Validation Methods

In the SGB there are mainly four hyper-parameters as described below. By tuning these parameters, the complexity and operability of the models can be controlled (Friedman 2001; Ridgeway 2007; Kuhn and Johnson 2013):

1. Maximum depth of variable interactions (interaction.depth), plays an important role controlling the maximum interaction order of the model and the complexity of the boosted ensemble;
2. The number of trees (or iterations) (n.trees), determines the numbers of decision trees, SGB can overfit if it is too large;
3. Minimum observations in terminal node (n.minobsinnode), represents the minimum number of observations in trees' terminal nodes. The default for the R *gbm* package (Ridgeway 2007) is 10, higher values mean smaller trees so make the SGB algorithm run faster and use less memory.
4. The learning rate (shrinkage), which is used for reducing (or shrinking) the size of incremental steps and thus penalizes the importance of each consecutive iteration. The value of shrinkage lies generally lies between 0 and 1 (typical values are 0.01, 0.05 or 0.001). Smaller values of shrinkage mean the convergency of larger numbers of iterations (n.trees).

In this work, the Kappa index is computed for each combination of the hyperparameters tuning with grid search strategy (Kuhn and Johnson 2013) and cross validation (CV) method. To avoid the bias in data selection, one of the most popular validation methods such as k-fold (CV) (Kohavi 1995; Zhou et al. 2016a, b, c) was applied in this work during the process of SGB hyperparameters tuning. Although, there is no definite rule of determining the value of k, according to the practice $k = 5$ (or 10) is used. But here k was set to be 10 as recommended by Kohavi (1995) associated with the trade-off between computation time and bias. It's, therefore, that ten rounds of training and validating were conducted using different partitions and their results were averaged to represent the performance of the SGB model on the training set.

Performance Metric

To benchmark model performances, the SGB classification model on seismic-liquefaction potential data was numerically and graphically evaluated by the classification accuracy rate (CAR), Cohen's Kappa coefficient and receiver operating characteristic (ROC) curve (Hastie et al. 2001; Kuhn and Johnson 2013). The classification accuracy rate (CAR), which can be taken as the percentage of samples that is correctly classified by the SGB model relative to the total number of samples by the classification models. The Cohen's Kappa coefficient or called Kappa (Kuhn and Johnson 2013; Zhou et al. 2016a, 2019) is a statistical measure which indicated the proportion of correctly classified units after the probability of chance agreement has been removed, and the Kappa index is more robust than CAR because it can consider the probability that a pixel is classified accidentally. As descripted by Kuhn and Johnson (2013), the CAR and the Kappa can be obtained using the following expressions:

$$\mathrm{CAR} = (\mathrm{TP} + \mathrm{TN})/\mathrm{N} \tag{5}$$

$$\mathrm{Kappa} = \frac{\mathrm{CAR} - Pe}{1 - Pe} \tag{6}$$

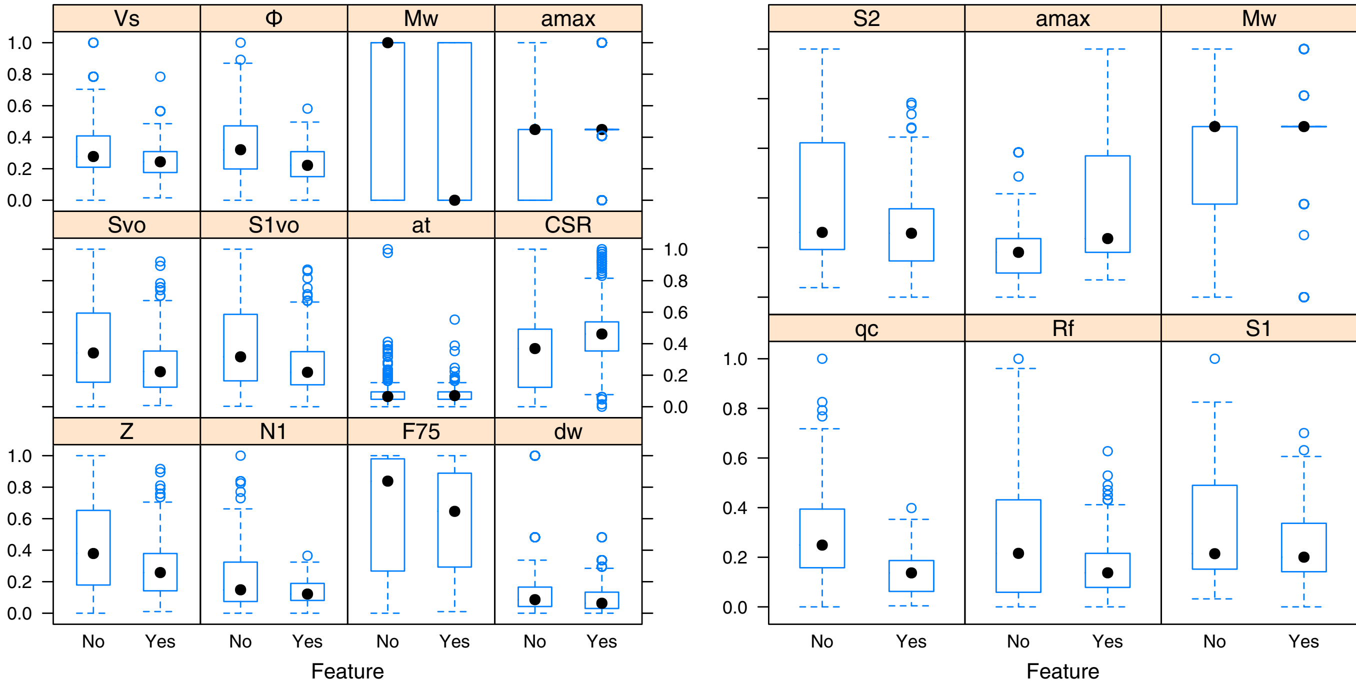
where $Pe = [(\mathrm{TN}+\mathrm{FN})\cdot(\mathrm{TN}+\mathrm{FP}) + (\mathrm{TP}+\mathrm{FP})\cdot(\mathrm{TP}+\mathrm{FN})]/\mathrm{N}^2$ can be seen as an expected accuracy based on confusion matrix; TN, FN, TP, FP and N denotes the numbers of true negatives, false negatives, true positives, false positives and the number of samples, respectively.

In addition to prediction accuracy and Kappa index, ROC was plotted and AUC (the area under the curve) value was calculated for each model to evaluate the prediction performance. The ROC (Bradley 1997) is a 2D plot of true positive rate (sensitivity) versus false positive rate (1-specificity) in vertical and horizontal axes, respectively. The ROC curve is also a binary tool with five degrees of rating (Bradley 1997): not discriminating (0.5–0.6), poor (0.6–0.7), fair (0.7–0.8), good (0.8–0.9) and excellent (0.9–1).

Results and Discussions

Descriptive Analysis

Fig. 1 demonstrates the all input parameters with respect to SPT and CPT data used to develop soil liquefaction triggered by earthquakes prediction models paralleling with their range, mean and outlets. The distributions of these variables and the relationship between seismic-liquefaction potential and each input variables in SGB models are depicted in the correlation matrix plot in Fig. 2, it can be seen that the pairwise relationship between soil and seismic parameters with corresponding correlation coefficients (in the upper panel) and the marginal frequency distribution (on the diagonal) for each parameter. As for the CPT data, it can be observed that all parameters have no relatively acceptable/valuable correlation with each other for CPT database but several parameters have appropriate correlation with each other for SPT database. In addition, as for the SPT data, it can be indicated that the parameter φ (CC or Correlation Coefficient $= -0.327$) and CSR (CC $= 0.267$) are slightly correlated with L, whereas for the CPT data,



(a) (b)

Fig. 1. Boxplot of indicators for original data sets: (a) SPT data; and (b) CPT data.

it can be observed that the parameter qc (CC = −0.457) and a_{max} (CC = 0.374) is highly correlated with L.

Results of Hyper-Parameters Tuning

This section applied the aforementioned methodology for modelling the SGB prediction of seismic liquefaction potential. The predictive models are established using input/output variables with training set and applied to testing sets as demonstrated in Fig. 3. Obviously, the SGB model with respect to input parameters are the indicators influencing an evaluation target (seismic-liquefaction potential). Thus SGB model with SPT data: input variables = M_w, a_{max}, Z, $N1$, F_{75}, d_w, V_s, φ, S_{vo}, S_{1vo}, CSR, and a_t; and SGB model with CPT data: input variables = M_w, a_{max}, S_1, S_2, R_f and qc. To carry out the experiments, the SPT/CPT dataset is randomly split into two subsets with stratified sampling method which divides the datasets into training and test datasets and creates the training set by selecting from these subsets with the same outcome distribution of the whole dataset. In the next step, by conducting a 10-fold CV procedure (Kuhn and Johnson 2013; Zhou et al. 2015, 2016a, b), the original training data are regarded as an integral database which randomly divided into ten portions. And then each fold of cross validation selects one portion of the data as testing data. The remaining nine portions constitute the training dataset, by this way, the model is established, as demonstrated as Fig. 3. Also, ten predictive accuracies were obtained. By this process, the hyper-parameters of SGB were more feasible and reasonable in contrast to manual search and random search. To investigate the behavior of the SGB hyper-parameters (n.trees, shrinkage, interaction.depth and n.minobsinnode), the model was tuned using a relatively fine grid search with ten folds CV methods, and tuning parameters are set as tabulated in Table 1. Furthermore, parameters of the SGB classification models which achieved the highest Kappa are defined as optimized during the tenfold CV procedure. The relationship between number of trees, the depth of the decision trees, the learning rate (the numbers in the top red boxes) and minimum number of observations in trees' terminal nodes (the numbers in the top green boxes) is shown in Fig. 4, where the vertical axis in each subplot illustrates the Kappa index by the 10-fold CV.

In the application of SPT dataset, the dataset compiled of 620 case histories is randomly split 80: 20 into two subsets: training (497) and test (123) datasets. The optimal parameters of the SGB model with SPT database were n · trees = 400, interaction · depth = 7, n · minobsinnode = 5 and shrinkage = 0.05. The CAR and Kappa of the SGB model with SPT data are obtained to be 87.12% and 0.7335 for 497 sets of training data using 10-folds CV procedure [Fig. 4(a)].

Similarly, to conduct the experiment for CPT dataset, the dataset consisting of 226 cases is also randomly split into two subsets (Zhou et al. 2019): training (182) and test (44) datasets. The optimal values leading to the highest Kappa based on tenfold CV were used in the final models. Thus the final values used for the CPT database [Fig. 4(b)] were n · trees = 800, interaction · depth = 2, n · minobsinnode = 3 and shrinkage = 0.1. The CAR and Kappa of the SGB model with CPT data are found to be 93.96% and 0.8732, respectively.

Additionally, the boxplot in Fig. 5 reports the performance of SGB classifiers using SPT data and CPT data on tenfold CV phrase with training set average accuracy/Kappa and its variance. The difference between lower and upper quartiles in SGB with SPT data is comparatively smaller than the SGB with CPT data that show relatively low variance of accuracies in different iterations.

Validation of the SGB

An ROC curve was implemented to evaluate the performance of SGB classification models in seismic liquefaction potential. The results of the prediction rate curves are represented in Fig. 6. According to the results, the AUC of ROC was calculated as 0.9417, and 0.9915 for SPT test data, and CPT test data, respectively.

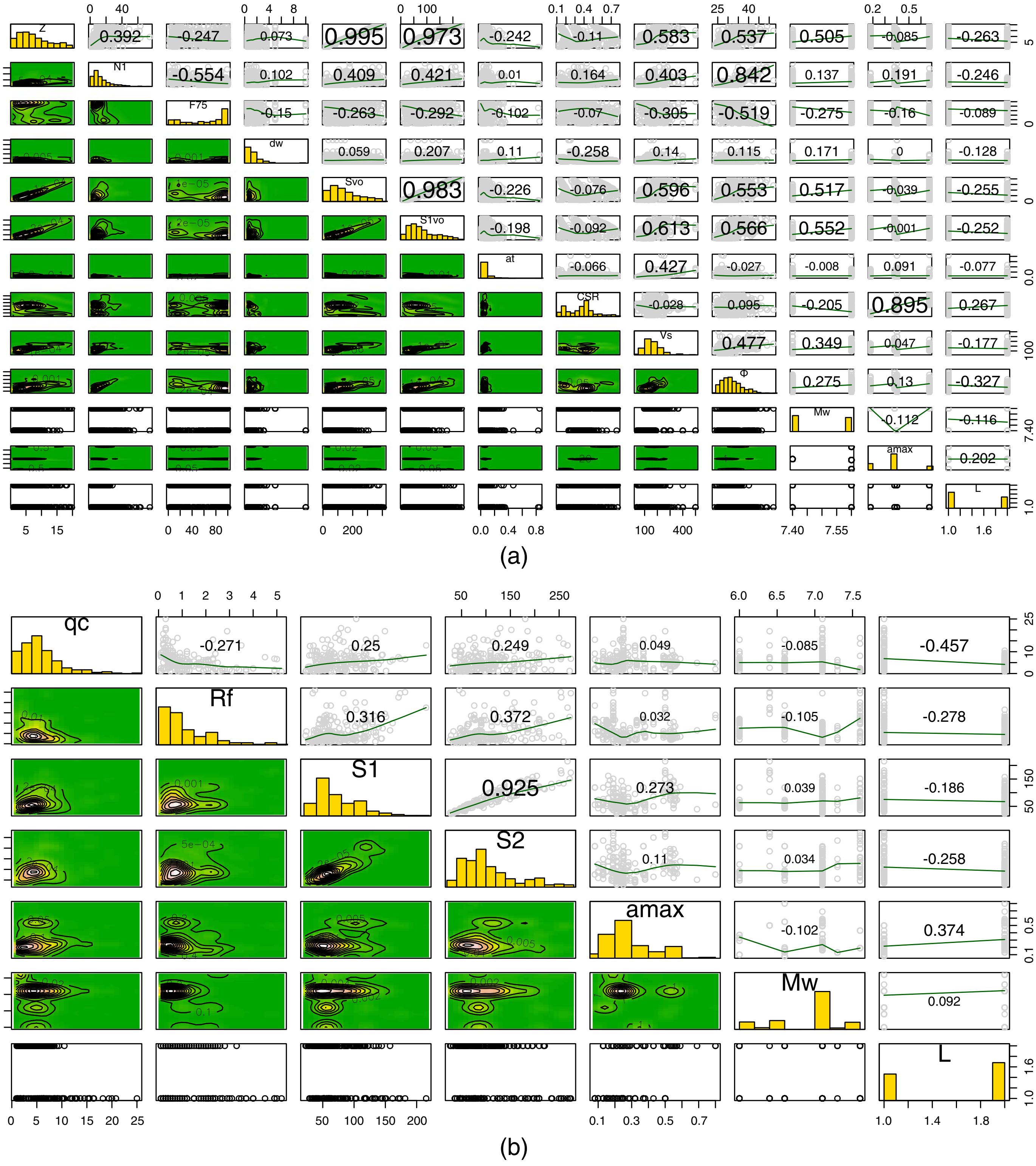


Fig. 2. Scatterplot matrix for all data sets with bivariate kernel density estimation: (a) SPT data; and (b) CPT data.

The area under the curve (AUC) determines the goodness of the performance of the SGB classifier in large extent where the value of 1.0 represents the ideal performance. Thus, it can be seen that the capability of the SGB was excellent for seismic liquefaction potential prediction.

To validate the generalization ability of the proposed model between the measured (real) and predicted values, 123 SPT testing cases and 44 CPT testing cases were validated by each optimized SGB model. The results are tabulated in Table 2 based on the confusion matrices (Zhou et al. 2015). Producer's accuracy (PA) and user's accuracy (UA) (Zhou et al. 2015) for each class using the SGB model is also shown in Table 2. The CAR and Kappa of the SGB model with SPT test data is obtained by 88.62% and 0.7682 for 123 cases of testing data, respectively. The CAR and Kappa of the SGB model with CPT data is achieved by 95.45% and 0.9043 for 44 cases of testing data, respectively. Obviously,

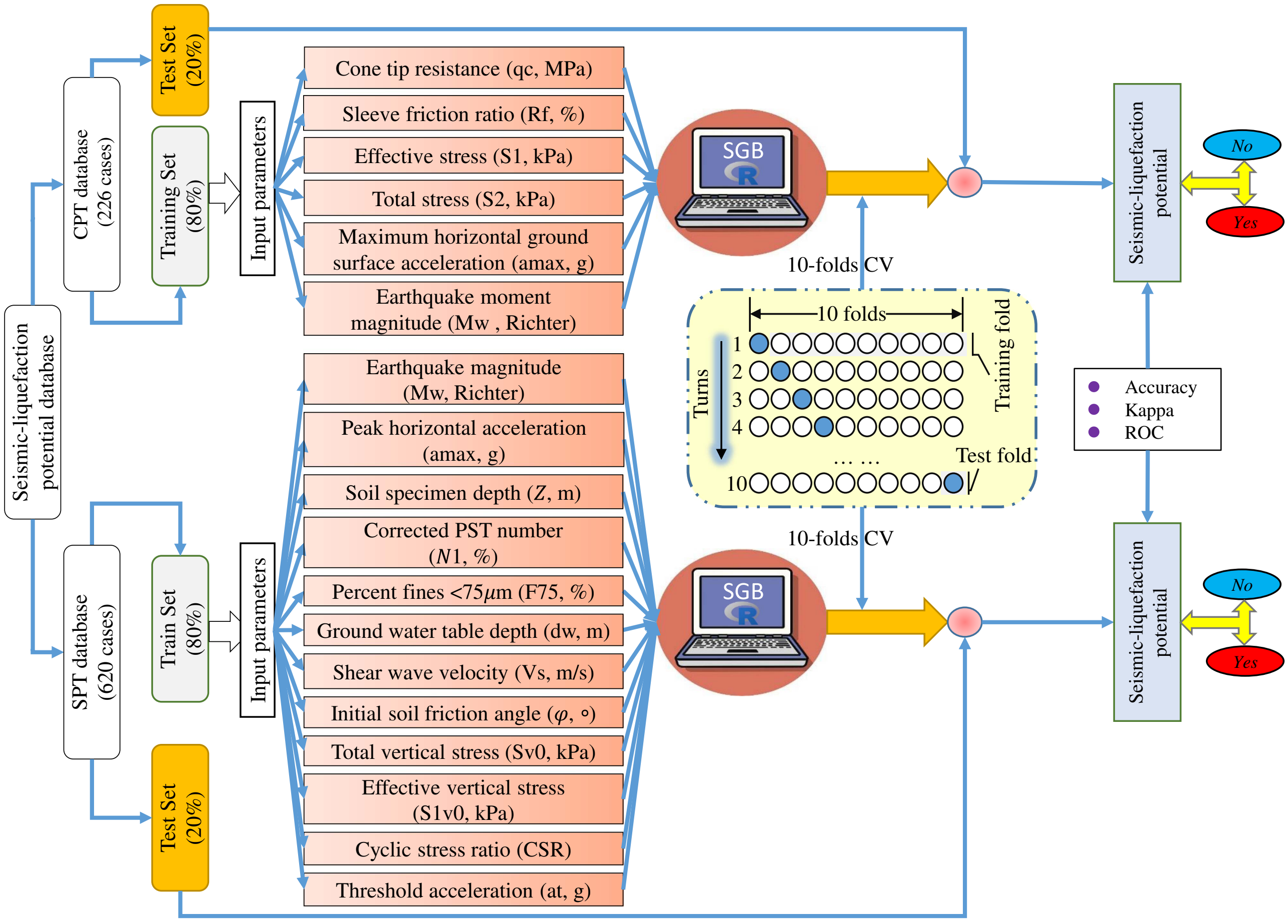


Fig. 3. Research architectures of overall procedure for performance evaluation for seismic-liquefaction potential using SGB methods.

Table 1. Tested values for the hyperparameters of the SGB prediction models and the optimal parameters revealed during parameter tuning

Hyperparameter	Tested values	Optimal value SPT	Optimal value CPT
n.trees	100 to 1,000 with increment 100	400	800
Interaction.depth	1 to NPV with increment 1	7	2
Shrinkage	{0.01,0.05,0.1}	0.05	0.1
n.minobsinnode	{3, 5, 10}	5	3

Note: NPV is the number of predictor variables.

these results are in accordance with actual observations and the accuracy of the SGB-based classification model is acceptable for estimation of seismic liquefaction potential. Additionally, the PA and UA indicate that with the available SPT/CPT data no liquefaction label react with the SGB classifier better than liquefaction one on SPT dataset whereas they yield the same accuracy on CPT dataset.

Subsequently, several additional SPT/CPT-based liquefaction trigged by earthquake case histories as reported in Idriss and Boulanger (2014), such as seven CPT liquefaction case histories were obtained from Tohoku earthquake in 2011 with Mw = 9.0 (Cox et al. 2013; Idriss and Boulanger 2014) and nine SPT-based case histories data were collected from Kocaeli earthquake in 1999 with Mw = 7.5 (Bray et al. 2004; Idriss and Boulanger 2014), have been validated by the developed SGB models, we achieved the satisfactory results using the optimized SGB models. Also, the civil and geotechnical engineers face challenges to present the potential of seismic liquefaction in new project for geotechnical design and site investigation, the aforementioned SGB-based model can provide a relatively reliable level of seismic liquefaction risk and thus reduces the level of losses, especially providing a reference for similar engineering surveys.

Variable Importance and Interpretation

The importance of influencing variables can be evaluated individually by means of the feature importance (Kuhn 2008). The variable importance of each SGB based model can be achieved accessible with the use of $varImp()$ function in *caret* package (Kuhn 2008). Fig. 7 provides the result for the SGB model and displays the relative variable importance for each models. The sensitivity analysis on these indicators is implemented meanwhile aiming to illustrate the contribution of each variant indicator for the estimation of these cases via *ROC* statistic. The SGB model in SPT database used the 12 parameters of M_w, a_{max}, Z, $N1$, F_{75}, d_w, V_s, φ, S_{vo},

J. Perform. Constr. Facil., 2019, 33(3): 04019024

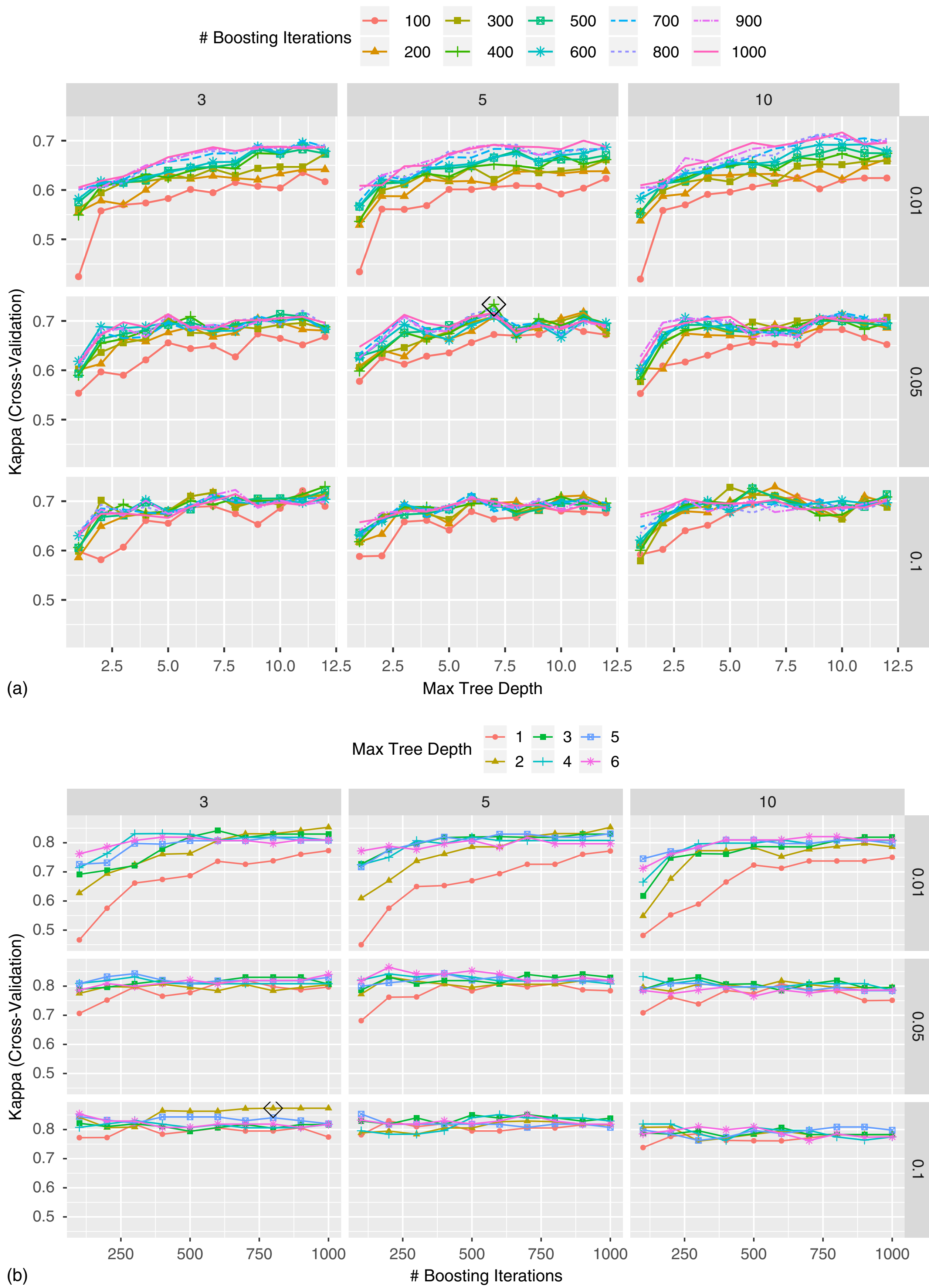


Fig. 4. Plots of resampling profiles to examine the relationship between estimates of performance and tuning parameters for SGB models with Kappa: (a) SPT data; and (b) CPT data.

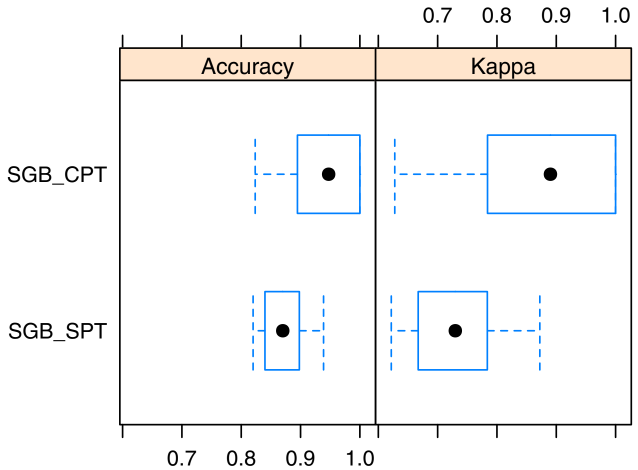


Fig. 5. Boxplot distributions of training set in terms of "accuracy" and "kappa" for SGB methods—resulting from repeated 10-fold CV procedure by the algorithm.

S_{1vo}, CSR, and a_t. Among all the variables, it is noteworthy that φ (internal friction angle of soil) is the most influential parameter among aforementioned indicators for seismic liquefaction potential prediction, which is consistent with the work of Hoang and Bui (2018). The indicators F_{75}, CSR and $N1$, are a bit sensitive, followed by the indicators S_{vo}, Z, S_{1vo}, V_s, d_w, a_t, a_{max}, and M_w. The factors of a_{max} and M_w are not as sensitive as the former factors, as depicted in Fig. 7(a).

Similarly, the SGB model in CPT database used the 6 parameters of Mw, a_{max}, S1, S2, qc, and R_f. It is highlighted that qc (the cone tip resistance) and a_{max} are the two most influential parameters among all these indicators for prediction seismic liquefaction potential, for the remaining parameters, the contribution of influence coefficients are sorted by R_f, $S2$, $S1$ and Mw in sequence as depicted in Fig. 7(b). The results are in line with Hoang and Bui (2018) reports and are also in keeping with the correlation matrix of the variables that showed the highest coefficients for these variables.

Findings demonstrated that the SGB predictor is more sensitive to the indicators of φ for SPT data whereas qc for CPT data. Interestingly, the factor of a_{max} is sensitive to SGB with SPT data but not sensitive to SGB with CPT data, which provides meaningful guiding significance for the subsequent prediction of seismic liquefaction potential. This results may be influenced by class imbalance and sampling bias as investigated by Hu et al. (2016) on the performances of proposed seismic liquefaction potential models, and thus it is significant to select a suitable class distribution of training data for reducing the class imbalance and sampling bias problems. However, it should be noted that the importance score of SGB models was relied on the SPT and/or CPT dataset used in this work. After that a larger dataset that contains more nonliquefaction and/or liquefaction events is desirable to achieve more reliable importance scores for influencing indicators.

Limitations

The SGB approach for seismic liquefaction potential, however, has some limitations that need to be addressed in future research. First, the limitation of SGB approach is that the datasets are relatively small, only a few dozen of cases involved in SGB modeling. It should be analyzed a larger dataset to improve the model's precision and reliability. Second, a further research about more

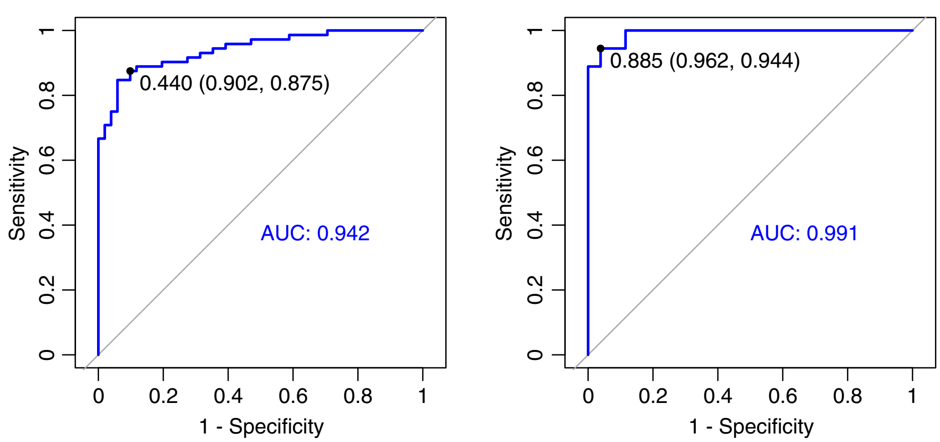


Fig. 6. ROC curve of the SGB model with the optimal hyper-parameters on SPT and CPT test set: (a) SGB_CPT ROC; and (b) SGB_SPT ROC.

Table 2. Confusion matrices and associated SGB classifier performance for prediction seismic liquefaction potential with test data

		Reference data			
Data set	Classified data	No	Yes	UA/%	PA/%
SPT database[a]	No	**63**	5	92.65	87.50
	Yes	9	**46**	83.64	90.20
CPT database[b]	No	**16**	0	100	88.89
	Yes	2	**26**	92.86	100

Note: The diagonal elements (correct decisions) are marked in bold. PA is the Producer's accuracy and UA is the user's accuracy.
[a]Kappa = 0.7682, CAR = 88.62% [95% CI: (0.8164, 0.9364)].
[b]Kappa = 0.9043, CAR = 95.45% [95% CI: (0.8453, 0.9944)].

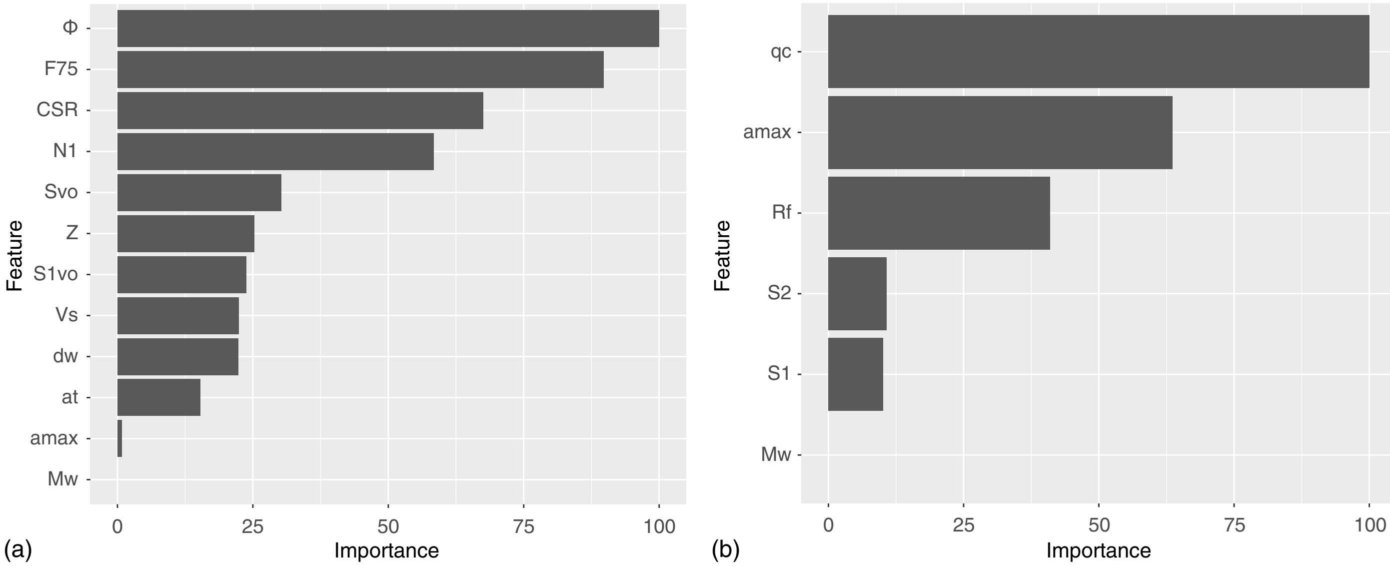


Fig. 7. Variable importance plot generated by the SGB algorithm: (a) SPT data; and (b) CPT data.

influencing variables of seismic liquefaction potential need to be explored and it remains to tune other hyper-parameters to guarantee the SGB algorithm more reliable and feasible and another optimization strategies such as Bayesian optimization (Snoek et al. 2012) and heuristic algorithms (Zhou et al. 2012) would be obtained desired results. Third, even though binary SGB classification models for seismic liquefaction potential showed high accuracy, multi-output classification models for seismic liquefaction potential analysis have not been investigated in this work. In addition, like other machine learning methods, it is sensitive to select the dataset in respect of the resulting tree structure of the SGB model and hence the data tend to be overfitted. Meanwhile, it is still difficult to interpret the complex relationships between the response and predictor variables because of the SGB algorithm's "black box" nature. Last, other advanced supervised machine learning approaches that exhibit satisfactory predictive accuracy on modelling complicated nonlinear systems, such as extreme gradient boosting (Chen and Guestrin 2016), random forest (Zhou et al. 2017) have not been investigated and compared on seismic liquefaction potential prediction.

Conclusions

Evaluation of seismic liquefaction potential of soil deposits has become a hot topic in earthquake science. In this work, a novel SGB-based approach is present to predict the seismic liquefaction potential of soil depoisits. Two database from CPT and SPT are validated the feasibility of the SGB model. Typical influential factors that characterize earthquake are generalized as the input parameters to discriminate that if the soil liquefaction phenomena occurs. To examine the fitness of the SGB model, tuning hyperparameters obtained the largest Kappa are considered ideal during the tenfolds CV process. To benchmark model performances, classification accuracy rate, Kappa index and area under the curve are employed. It is generally accepted that the SGB has been considered successful to evaluate the potential of soil liquefaction because of the strikes of the earthquake. Additionally, we also obtained the order of importance of input parameters which influence the prediction of soil liquefaction potential. It is suggested that the SGB method can be considered for modelling the sophisticated relationship between the liquefaction potential and relevant soil and seismic parameters, and also provide a novel approach for other geotechnical and civil engineering stability/failure issues depending on data availability.

Acknowledgments

The authors appreciate the support of the National Natural Science Foundation Project of China (Grant No. 41807259), the State Key Laboratory Breeding Base for Mining Disaster Prevention and Control (Grant No. MDPC201608), the Natural Science Foundation of Hunan Province (Grant No. 2018JJ3693), the China Postdoctoral Science Foundation funded project (Grant No. 2017M622610) and the Sheng Hua Lie Ying Program of Central South University (Principle Investigator: Dr. Jian Zhou).

References

Bradley, A. P. 1997. "The use of the area under the ROC curve in the evaluation of machine learning algorithms." *Pattern Recognit.* 30 (7): 1145–1159. https://doi.org/10.1016/S0031-3203(96)00142-2.

Bray, J. D., et al. 2004. "Subsurface characterization of ground failure sites in Adapazari Turkey." *J. Geotech. Geoenviron. Eng.* 130 (7): 673–685. https://doi.org/10.1061/(ASCE)1090-0241(2004)130:7(673).

Chen, T., and C. Guestrin. 2016. "Xgboost: A scalable tree boosting system." Accessed March 15, 2018. http://arxiv.org/abs/1603.02754.

Cox, B. R., et al. 2013. "Liquefaction at strong motion stations and in Urayasu City during the 2011 Tohoku-Oki earthquake." Supplement, *Earthquake Spectra EERI* 29 (S1): S55–S80. https://doi.org/10.1193/1.4000110.

Friedman, J. H. 2001. "Greedy function approximation: A gradient boosting machine." *Ann. Stat.* 29 (5): 1189–1232. https://doi.org/10.1214/aos/1013203451.

Friedman, J. H. 2002. "Stochastic gradient boosting." *Comput. Stat. Data Anal.* 38 (4): 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2.

Goh, A. T. 1996. "Neural-network modeling of CPT seismic liquefaction data." *J. Geotech. Eng.* 122 (1): 70–73. https://doi.org/10.1061/(ASCE)0733-9410(1996)122:1(70).

Goh, A. T., and S. Goh. 2007. "Support vector machines: Their use in geotechnical engineering as illustrated using seismic liquefaction data."

Comput. Geotech. 34 (5): 410–421. https://doi.org/10.1016/j.compgeo.2007.06.001.

Hanna, A. M., D. Ural, and G. Saygili. 2007. "Neural network model for liquefaction potential in soil deposits using Turkey and Taiwan earthquake data." *Soil Dyn. Earthquake Eng.* 27 (6): 521–540. https://doi.org/10.1016/j.soildyn.2006.11.001.

Hastie, T., R. Tibshirani, and J. H. Friedman. 2001. Vol. 1 of *The elements of statistical learning*, 339. New York: Springer.

Hoang, N. D., and D. T. Bui. 2018. "Predicting earthquake-induced soil liquefaction based on a hybridization of kernel Fisher discriminant analysis and a least squares support vector machine: A multi-dataset study." *Bull. Eng. Geol. Environ.* 77 (1): 191–204. https://doi.org/10.1007/s10064-016-0924-0.

Hu, J. L., X. W. Tang, and J. N. Qiu. 2016. "Analysis of the influences of sampling bias and class imbalance on performances of probabilistic liquefaction models." *Int. J. Geomech.* 17 (6): 04016134. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000808.

Idriss, I. M., and R. W. Boulanger. 2006. "Semi-empirical procedures for evaluating liquefaction potential during earthquakes." *Soil Dyn. Earthquake Eng.* 26 (2): 115–130. https://doi.org/10.1016/j.soildyn.2004.11.023.

Idriss, I. M., and R. W. Boulanger. 2014. *CPT and SPT based liquefaction triggering procedures*. Rep. No. UCD/CGM-14/01. London: Dept. of Civil and Environmental Engineering.

Jiang, L., P. Kong, J. Shu, and K. Fan. 2019a. "Numerical analysis of support designs based on a case study of a longwall entry." *Rock Mech. Rock Eng.* https://doi.org/10.1007/s00603-018-1728-2.

Jiang, L., Q. S. Wu, Q. L. Wu, P. Wang, Y. Xue, P. Kong, and B. Gong. 2019b. "Fracture failure analysis of hard and thick key layer and its dynamic response characteristics." *Eng. Fail. Anal.* 98: 118–130. https://doi.org/10.1016/j.engfailanal.2019.01.008.

Juang, C. H., H. Yuan, D. H. Lee, and P. S. Lin. 2003. "Simplified cone penetration test-based method for evaluating liquefaction resistance of soils." *J. Geotech. Geoenviron. Eng.* 129 (1): 66–80. https://doi.org/10.1061/(ASCE)1090-0241(2003)129:1(66).

Kaveh, A., S. M. Hamze-Ziabari, and T. Bakhshpoori. 2018. "Patient rule-induction method for liquefaction potential assessment based on CPT data." *Bull. Eng. Geol. Environ.* 77 (2): 849–865. https://doi.org/10.1007/s10064-016-0990-3.

Kohavi, R. 1995. "A study of cross-validation and bootstrap for accuracy estimation and model selection." In Vol. 2 of *Proc., 14th Int. Joint Conf. on Artificial Intelligence IJCAI' 95*, 1137–1143. San Francisco: Morgan Kaufmann.

Kohestani, V. R., M. Hassanlourad, and A. Ardakani. 2015. "Evaluation of liquefaction potential based on CPT data using random forest." *Nat. Hazards* 79 (2): 1079–1089. https://doi.org/10.1007/s11069-015-1893-5.

Kuhn, M. 2008. "Building predictive models in R using the caret package." *J. Stat. Software* 28 (5): 1–26. https://doi.org/10.18637/jss.v028.i05.

Kuhn, M., and K. Johnson. 2013. *Applied predictive modeling*. New York: Springer.

Lai, S. Y., W. J. Chang, and P. S. Lin. 2006. "Logistic regression model for evaluating soil liquefaction probability using CPT data." *J. Geotech. Geoenviron. Eng.* 132 (6): 694–704. https://doi.org/10.1061/(ASCE)1090-0241(2006)132:6(694).

Lai, S. Y., S. C. Hsu, and M. J. Hsieh. 2004. "Discriminant model for evaluating soil liquefaction potential using cone penetration test data." *J. Geotech. Geoenviron. Eng.* 130 (12): 1271–1282. https://doi.org/10.1061/(ASCE)1090-0241(2004)130:12(1271).

Li, X. B., J. Zhou, S. F. Wang, and B. Liu. 2017. "Review and practice of deep mining for solid mineral resources." *Chin. J. Nonferrous Metals* 27 (6): 1236–1262. https://doi.org/10.19476/j.ysxb.1004.0609.2017.06.021.

Lu, J., D. Lu, X. Zhang, Y. Bi, K. Cheng, M. Zheng, and X. Luo. 2016. "Estimation of elimination half-lives of organic chemicals in humans using gradient boosting machine." *Biochimica et Biophysica Acta (BBA)-General Subjects* 1860 (11): 2664–2671. https://doi.org/10.1016/j.bbagen.2016.05.019.

Moss, R. E., R. B. Seed, R. E. Kayen, J. P. Stewart, A. Der Kiureghian, and K. O. Cetin. 2006. "CPT-based probabilistic and deterministic assessment of in situ seismic soil liquefaction potential." *J. Geotech. Geoenviron. Eng.* 132 (8): 1032–1051. https://doi.org/10.1061/(ASCE)1090-0241(2006)132:8(1032).

Oommen, T., L. G. Baise, and R. Vogel. 2010. "Validation and application of empirical liquefaction models." *J. Geotech. Geoenviron. Eng.* 136 (12): 1618–1633. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000395.

Pal, M. 2006. "Support vector machines-based modelling of seismic liquefaction potential." *Int. J. Numer. Anal. Methods Geomech.* 30 (10): 983–996. https://doi.org/10.1002/nag.509.

Qi, C., A. Fourie, G. Ma, X. Tang, and X. Du. 2017. "Comparative study of hybrid artificial intelligence approaches for predicting hangingwall stability." *J. Comput. Civ. Eng.* 32 (2): 04017086. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000737.

R Core Team. 2017. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ridgeway, G. 2007. "Generalized boosted models: A guide to the gbm package." Accessed June 10, 2017. http://cran.r-project.org/web/packages/gbm/index.html.

Samui, P. 2007. "Seismic liquefaction potential assessment by using relevance vector machine." *Earthquake. Eng. Eng. Vibr.* 6 (4): 331–336. https://doi.org/10.1007/s11803-007-0766-7.

Seed, H. B., and I. M. Idriss. 1971. "Simplified procedure for evaluating soil liquefaction potential." *J. Soil Mech. Found. Div.* 97 (9): 1249–1273.

Snoek, J., H. Larochelle, and R. P. Adams. 2012. "Practical Bayesian optimization of machine learning algorithms." In *Proc., 26th Int. Conf. on Advances in Neural Information Processing Systems (NIPS'12)*. Edited by P. Bartlett, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, 2960–2968. Cambridge, MA: MIT Press.

Touzani, S., J. Granderson, and S. Fernandes. 2018. "Gradient boosting machine for modeling the energy consumption of commercial buildings." *Energy Build.* 158: 1533–1543. https://doi.org/10.1016/j.enbuild.2017.11.039.

Wang, Z. W., D. P. Zhao, X. Liu, C. Chen, and X. B. Li. 2017. "P and S wave attenuation tomography of the Japan subduction zone." *Geochem. Geophys. Geosyst.* 18 (4): 1688–1710. https://doi.org/10.1002/2017GC006800.

Xue, X., and X. Yang. 2013. "Application of the adaptive neuro-fuzzy inference system for prediction of soil liquefaction." *Nat. Hazards* 67 (2): 901–917. https://doi.org/10.1007/s11069-013-0615-0.

Yazdi, J. S., F. Kalantary, and H. S. Yazdi. 2012. "Investigation on the effect of data imbalance on prediction of liquefaction." *Int. J. Geomech.* 13 (4): 463–466. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000217.

Zhou, J., X. B. Li, and X. Z. Shi. 2012. "Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines." *Saf. Sci.* 50 (4): 629–644. https://doi.org/10.1016/j.ssci.2011.08.065.

Zhou, J., X. B. Li, and H. S. Mitri. 2015. "Comparative performance of six supervised learning methods for the development of models of pillar stability." *Nat. Hazards* 79 (1): 291–316. https://doi.org/10.1007/s11069-015-1842-3.

Zhou, J., X. B. Li, and H. S. Mitri. 2016a. "Classification of rockburst in underground projects: comparison of ten supervised learning methods." *J. Comput. Civ. Eng.* 50 (4): 629–644. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000553.

Zhou, J., X. Chen, M. Z. Wang, E. M. Li, H. Chen, and X. Z. Shi. 2019. "Classification of seismic-liquefaction potential using Friedman's stochastic gradient boosting based on the cone penetration test data." In *Transportation and geotechniques: Materials, sustainability and climate*, edited by M. Barman, M. Zaman, and J. R. Chang, 67–78. Cham, Switzerland: Springer.

Zhou, J., X. Z. Shi, K. Du, X. Y. Qiu, X. B. Li, and H. S. Mitri. 2017. "Feasibility of random-forest approach for prediction of ground settlements induced by the construction of a shield-driven tunnel." *Int. J. Geomech.* 79 (1): 291–316. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000817.

Zhou, J., X. Z. Shi, R. D. Huang, X. Y. Qiu, and C. Chen. 2016b. "Feasibility of stochastic gradient boosting approach for predicting rockburst damage in burst-prone mines." *Trans. Nonferrous Met. Soc. China* 26 (7): 1938–1945. https://doi.org/10.1016/S1003-6326(16)64312-1.

Zhou, J., X. Z. Shi, and X. B. Li. 2016c. "Utilizing gradient boosted machine for the prediction of damage to residential structures owing to blasting vibrations of open pit mining." *J. Vibr. Control* 22 (19): 3986–3997. https://doi.org/10.1177/1077546314568172.