



# OPEN Performance evaluation of machine learning algorithms for predicting liquefaction-induced lateral displacement

Mahmood Ahmad<sup>1,2</sup>, Mohammad Al Zubi<sup>3</sup>, Shaikat Biswas<sup>4</sup>, Abdur Rahman Rakib<sup>4</sup>, Abdullah Alzlfawi<sup>5</sup>, Sabahat Hussan<sup>6</sup>, Shay Haq<sup>7</sup>, Rohayu Che Omar<sup>1</sup>, Zia Ullah<sup>8</sup> & Zsolt Tóth<sup>9</sup>✉

Accurate prediction of liquefaction-induced lateral displacement is essential for seismic risk assessment, resilient infrastructure design, and cost-effective mitigation. Such predictions are complex and cannot be reliably addressed using conventional analytical approaches. This research utilizes XGBoost, CatBoost and AdaBoost algorithms with 247 post-liquefaction in-situ free-face ground condition case studies to model and investigate liquefaction induced lateral displacements. The models are assessed according to the coefficient of determination ( $R^2$ ), coefficient of correlation ( $r$ ), mean absolute error (MAE), mean squared error (MSE), root means square error (RMSE), and Root Mean Squared Error-Observations standard deviation Ratio (RSR) and Nash Sutcliffe Efficiency (NSE) coefficient. The develop models are compared with each other, and also to the Gaussian process regression, artificial neural network, Evolutionary polynomial regression and Multiple linear regression models described in the literature. The XGBoost model had the best prediction performance, with  $R^2 = 0.9905$ ,  $r = 0.9952$ ,  $MAE = 0.1491$ ,  $MSE = 0.0485$ ,  $RMSE = 0.2203$ ,  $RSR = 0.0981$ , and  $NSE = 0.9904$  for training and  $R^2 = 0.9251$ ,  $r = 0.9618$ ,  $MAE = 0.3642$ ,  $MSE = 0.3723$ ,  $RMSE = 0.6101$ ,  $RSR = 0.278$ , and  $NSE = 0.9227$  for testing results respectively. The rank score analysis confirmed XGBoost to be superior, as it reached the highest total score of 40. Sensitivity analysis further revealed that  $T_{15}$  was the most sensitive parameter to output variable.

**Keywords** Lateral displacement, Soil liquefaction, Machine learning, Performance measure

Lateral displacement has been recognized as a particularly destructive type of ground failure in previous large-magnitude earthquake disasters<sup>1</sup>. Case studies show that liquefaction-induced lateral displacement is one of the principal causes of damage to pile foundations subjected to seismic stresses<sup>2</sup>. Liquefaction and its induced hazards such as lateral displacement can cause significant deformations and damage to existing structures, including ports, bridges, and pipelines<sup>3–5</sup>. Lateral displacement refers to the horizontal movement that occurs in gently inclined ground surfaces (typically with slopes less than 5%) when saturated soils experience excess pore water pressure or undergo liquefaction<sup>1</sup>. Lateral displacement can result in shear failure in pile foundations, causing surface structures to crack, strain, or even collapse<sup>6</sup>. Lateral displacement can cause numerous types of ground deformation, ranging in size from a few centimeters to several meters<sup>7</sup>. Lateral ground movement caused by liquefaction is a complicated geotechnical phenomenon that is governed by seismic loads, soil characteristics, and site conditions. As a result, conventional approaches for predicting such displacements frequently have intrinsic limits and cannot adequately reflect the complexity of the problem<sup>8</sup>. Several variables influence lateral

<sup>1</sup>Institute of Energy Infrastructure, Universiti Tenaga Nasional, Kajang 43000, Malaysia. <sup>2</sup>Department of Civil Engineering, University of Engineering and Technology Peshawar (Bannu Campus), Bannu 28100, Pakistan. <sup>3</sup>Department of Mechanical Engineering, Hijjawi Faculty for Engineering Technology, Yarmouk University, Irbid, Jordan. <sup>4</sup>Department of Computer Science, Troy University, 600 University Ave, Troy, AL 36082, USA. <sup>5</sup>Department of Civil and Environmental Engineering, College of Engineering, Majmaah University, Al Majmaah 11952, Saudi Arabia. <sup>6</sup>Department of Civil Engineering, HITEC University, Taxila 47040, Pakistan. <sup>7</sup>Department of Civil Engineering, National University of Sciences and Technology, Balochistan Campus, Quetta 87300, Pakistan. <sup>8</sup>Broadridge Financial Solutions Inc, 51 Mercedes Way, Brentwood, NY 11717, USA. <sup>9</sup>Faculty of Wood Engineering and Creative Industries, University of Sopron, Sopron, Hungary. ✉email: toth.zsolt@uni-sopron.hu

displacement, including groundwater depth, the thickness and particle size of liquefiable soils, and the magnitude of the earthquake<sup>7</sup>. Several analytical, empirical, and computational models have been developed to anticipate how liquefiable soils will behave and how much lateral displacement may be triggered during seismic activity<sup>3</sup>. It is still difficult to adequately predict a liquefiable soil, despite intensive laboratory experimentation. These attempts are complicated by the challenge of getting “undisturbed” testing samples from the in situ deposit<sup>7</sup>. Predicting lateral displacements is a significant difficulty in geotechnical engineering because of the highly nonlinear interactions involved, which make it impossible to accurately handle using traditional methods. Therefore, using cutting-edge and creative assessment methodologies becomes crucial to designing practical and affordable mitigation strategies against liquefaction-induced lateral displacement<sup>8</sup>.

Liquefaction-induced lateral displacement has been examined through deformation monitoring and excavation-induced ground response analyses, providing insight into subsurface movement mechanisms relevant to seismic spreading<sup>9</sup>. Data-driven prediction of underground settlement, landslide formation mechanisms, hybrid displacement modeling, and nanoscale pore evolution collectively improve understanding of progressive ground deformation under instability<sup>10–14</sup>. Seismic detection techniques, discrete element simulations, tidal-stiffness profiling, seepage–erosion experiments, and intelligent risk models further enhance assessment of dynamic soil weakening and infrastructure vulnerability<sup>15–19</sup>. Advanced seismic interaction analyses, jacking-force prediction, unified critical state modeling, hydraulic control strategies, and seabed wave-response studies strengthen the theoretical and computational framework for evaluating liquefaction-driven lateral displacement in complex soil–structure systems<sup>20–24</sup>.

To assess the liquefaction induced lateral displacement, a number of analytical and empirical methods have been put forth. A useful and successful empirical tool among them is the Multiple Linear Regression (MLR) model, which was developed by regression analysis of recorded case histories<sup>7</sup>. Liquefaction-induced lateral displacement in earth dams, embankments, and slopes is made simple by an analytical method that combines Newmark’s sliding-block method with recorded acceleration time histories<sup>25</sup>. To assess the lateral displacement caused by liquefaction, a number of models have been developed using the finite element method. Gu and Robertson<sup>26</sup> evaluated the Lower San Fernando Dam’s post-earthquake deformations using an incremental finite element method. Gu and Robertson<sup>27</sup> applied a plane-strain model to analyze liquefaction-induced deformations, successfully capturing the displacement pattern at a wildlife site in California, though the displacements were overestimated by approximately 30%. In contrast, a stochastic and non-parametric method called Gaussian Process Regression (GPR) has become a viable machine learning (ML) strategy for dealing with challenging and nonlinear issues<sup>28</sup>.

In recent years, ML have received growing attention in civil engineering for their ability to address complex tasks, where conventional deterministic methods remain limited<sup>29–38</sup>. The application of AI methods has increased considerably, primarily due to their ability to model and predict highly complex non-linear interactions<sup>29,30,39–41</sup>. ML is an emerging set of diverse techniques that tend to reconcile the low precision and uncertainty as well as the lack of information using some statistical, probabilistic, and optimization tools in order to analyze sets of data, categorize the information, discover novel patterns and predict forthcoming trends with the briefest delay<sup>42</sup>. By gradually integrating several weak learners, the gradient boosting technique creates a powerful prediction model. In order to minimize total loss, these learners are connected in phases during training. The model iteratively modifies the weights at each stage to lower the error between the predicted and observed values<sup>43</sup>. In contrast to models like random forest or support vector machine, the algorithms Extreme Gradient Boosting (XGBoost), Adaptive Boosting (AdaBoost), and Categorical Boosting (CatBoost) were chosen for their superior accuracy and computational efficiency, as well as their capacity to handle high-dimensional data and capture non-linear relationships<sup>44–46</sup>. The frequent usage of these algorithms in research across fields demonstrates their versatility and adaptability to a wide range of domains; nevertheless, literature reviews indicate that their applications in seismic liquefaction-induced hazards are limited.

This study’s primary contributions are: (1) development of novel predictive models for liquefaction-induced lateral displacement with high precision and accuracy; (2) rigorous validation of the proposed models through comparative performance assessment against established models in the literature; and (3) execution of a sensitivity analysis to identify the most influential input parameters.

## Data catalog and analytics

This study’s dataset includes 247 free-face lateral displacement recordings obtained from Youd et al.<sup>47</sup>, Cetin et al.<sup>48</sup>, and Chu et al.<sup>49</sup> (see Supplementary information file, Table S1). Regarding data quality, the dataset was thoroughly screened prior to analysis, with no missing values identified. Additionally, the data distribution was carefully examined to confirm the absence of significant outliers, ensuring the integrity and robustness of the model training process. These factors support the reliability of the reported evaluation metrics and overall model performance. The dataset was adopted from published case histories in a consistent form; therefore, no additional pre-processing or normalization was applied. This study used seven major input parameters to evaluate lateral displacement ( $D_H$ ): earthquake magnitude ( $M$ ), attenuation-based peak ground acceleration ( $a_{max}$ ) estimates derived from the ground-motion relationships of Sadigh et al.<sup>50</sup>, horizontal distance from the seismic energy source ( $R$ ), mean particle size within  $T_{15}$  ( $D_{50,15}$ ), cumulative thickness of saturated layers with corrected SPT values ( $N_{1,60} < 15$  ( $T_{15}$ ), and average fines content ( $< 0.075$  mm) in  $T_{15}$  ( $F_{15}$ ). These input parameters were selected based on key variables consistently reported in previous numerous studies that are widely accepted for predicting lateral displacement (e.g., Ahmad et al.<sup>28</sup>, Javadi et al.<sup>51</sup>, and Jafarian and Nasri<sup>52</sup>). This study incorporates peak ground acceleration (PGA,  $a_{max}$ ) as an additional intensity measure, estimated using fault-mechanism-based attenuation model by Sadigh et al.<sup>50</sup> to improve prediction accuracy.

The dataset of 247 case histories was partitioned into training and testing subsets. The training set, comprising 198 samples (80% of the data), was used to develop the model, while the testing set, consisting of 49 samples

| Dataset  | Parameter          |         |        |               |              |                 |              |             | Output    |
|----------|--------------------|---------|--------|---------------|--------------|-----------------|--------------|-------------|-----------|
|          | Statistical        | Seismic |        |               | Geotechnical |                 |              | Topographic |           |
|          |                    | M       | R (km) | $a_{max}$ (g) | $F_{15}$ (%) | $D50_{15}$ (mm) | $T_{15}$ (m) | W (%)       | $D_H$ (m) |
| Training | Minimum            | 6.4     | 0.5    | 0.15          | 1            | 0.04            | 0.2          | 1.64        | 0         |
|          | Average            | 7.26    | 15.34  | 0.4           | 17.25        | 0.39            | 8.22         | 11.54       | 2.49      |
|          | Maximum            | 9.2     | 100    | 0.68          | 70           | 7.7             | 16.7         | 57.7        | 10.16     |
|          | Standard Deviation | 0.48    | 11.03  | 0.14          | 13.06        | 0.67            | 5.1          | 10.34       | 2.25      |
| Testing  | Minimum            | 6.4     | 0.5    | 0.15          | 3            | 0.07            | 0.5          | 2.11        | 0         |
|          | Average            | 7.27    | 15.12  | 0.42          | 20.33        | 0.27            | 6.27         | 10.63       | 2.03      |
|          | Maximum            | 9.2     | 60     | 0.68          | 66           | 1.47            | 15.6         | 48.98       | 9.29      |
|          | Standard Deviation | 0.62    | 12.93  | 0.16          | 15.18        | 0.29            | 5.13         | 8.08        | 2.22      |

**Table 1.** Statistical Parameters for free face condition.

| Parameters      | M     | R (km) | $a_{max}$ (g) | $F_{15}$ (%) | $D50_{15}$ (mm) | $T_{15}$ (m) | W (%) | $D_H$ (m) |
|-----------------|-------|--------|---------------|--------------|-----------------|--------------|-------|-----------|
| M               | 1     |        |               |              |                 |              |       |           |
| R (km)          | 0.76  | 1      |               |              |                 |              |       |           |
| $a_{max}$ (g)   | −0.34 | −0.72  | 1             |              |                 |              |       |           |
| $F_{15}$ (%)    | −0.37 | −0.37  | 0.56          | 1            |                 |              |       |           |
| $D50_{15}$ (mm) | 0.03  | 0.01   | −0.11         | −0.23        | 1               |              |       |           |
| $T_{15}$ (m)    | 0.21  | 0.36   | −0.57         | −0.59        | 0.24            | 1            |       |           |
| W (%)           | 0.00  | −0.05  | 0.18          | 0.25         | 0.03            | −0.15        | 1     |           |
| $D_H$ (m)       | 0.18  | 0.23   | −0.25         | −0.35        | −0.08           | 0.52         | 0.15  | 1         |

**Table 2.** Input and output parameters correlation.

(20% of the data), was employed to assess its performance. The statistical distributions of both input and output parameters for the two subsets are presented in Table 1.

To assess the significance of associations between variables, correlation coefficients ( $\rho$ ) were calculated. The parameters showed moderate to weak correlations, and the largest absolute correlation coefficient ( $|\rho|$ ) was  $< 0.8$  (see Table 2). It was also tested whether multicollinearity could have played a role, since highly correlated predictor variables can inflate rather than decrease variance, and it may be difficult to detect true effects of individual predictors. The outcomes showed that some parameters were moderately associated (e.g.,  $M$ – $R$ ,  $R$ – $a_{max}$ ) but all the results did not surpass the generally accepted benchmark of multicollinearity ( $|\rho| > 0.8$ )<sup>53</sup>. Therefore, multicollinearity was unlikely and did not negatively affect the performance of the model.

The box plot distribution of liquefaction-induced parameters was plotted as presented in Fig. 1. The findings reveal that the magnitude of earthquake ( $M$ ) and peak ground acceleration ( $a_{max}$ ) has low dispersion, which helps in expressing the similarities in case histories of seismic input conditions. In comparison, horizontal distance to seismic energy source ( $R$ ) records highest variation and positive skew, implying that there is a wide variation of values in the dataset in connection to horizontal distance to seismic energy source. Average fines (particles  $< 0.075$  mm) in  $T_{15}$  ( $F_{15}$ ) and free-face ratio ( $W$ ) also have significant variability and high value outliers and hence indicate that there is strong variability in the soil composition and free-face conditions. Parameters, like average particle size in  $T_{15}$  ( $D50_{15}$ ), and accumulative thickness of saturated layers with adjusted SPT number ( $N_{1,60} < 15$  ( $T_{15}$ )), are considered to have relatively compact distributions, which suggest some more even site-specific geotechnical properties. Parameters i.e.,  $R$ ,  $F_{15}$  and  $W$  are highly variable, which shows their significant impact on spatial variability of liquefaction manifestation.

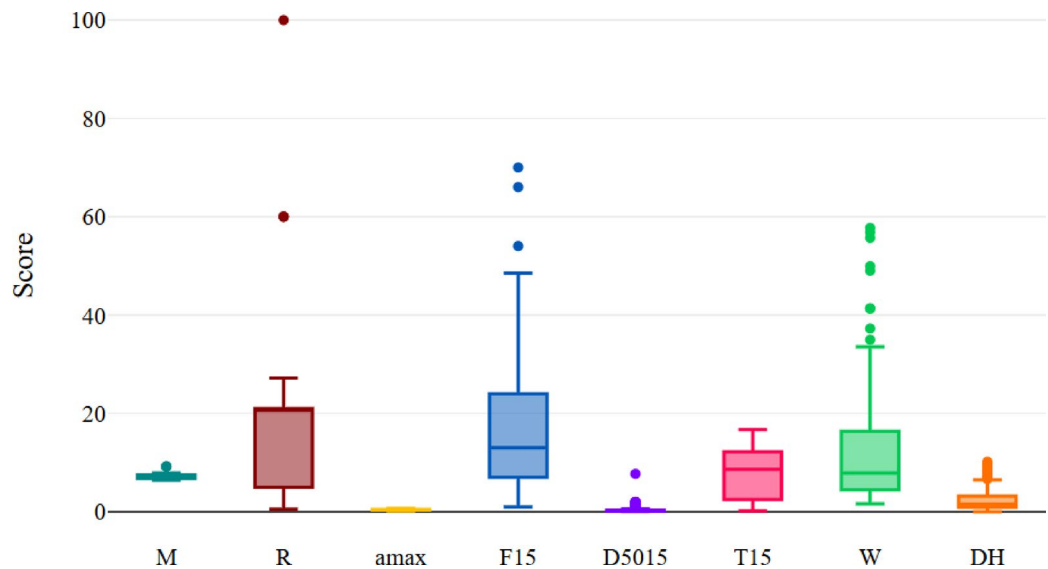
## Methodology

### Adaptive boosting (AdaBoost)

With a strong theoretical foundation and a proven track record of effectiveness in real-world applications, AdaBoost is among the excellent boosting algorithms as it introduces an innovative and new concept to learning algorithm creation by using a weak learning algorithm with accuracy only marginally better than random guessing to create an arbitrarily accurate strong learning algorithm<sup>45</sup>. The goal of the machine learning technique known as “boosting” is to combine several, incorrect, and weak rules to produce a single, highly accurate prediction rule<sup>54</sup>.

$$\text{Final Hypothesis, } F(x) = \sum a_t h_t(x) \quad (1)$$

where



**Fig. 1.** Box plot distribution of liquefaction-induced lateral displacement parameters.

weak hypothesis =  $h(x)$ .

$$\alpha_t = \frac{1}{2} \ln \left( \frac{1 - \epsilon_t}{\epsilon_t} \right) \quad (2)$$

The weak learner aims to identify a weak hypothesis that has a low weighted error,  $\epsilon_t$ , in relation to weak learner distribution. The combined or final hypothesis determines the value of a weighted set of weak hypotheses<sup>54</sup>.

### Extreme gradient boosting (XGBoost)

Sequential decision trees are produced by the enhanced gradient tree boosting technique known as XGBoost. In all computing environments, it can swiftly do correlated calculations<sup>55</sup>. XGBoost employs a gradient boosting framework, sequentially constructing trees based on the residuals of preceding iterations, while incorporating regularization terms into its objective function to mitigate overfitting and enhance both generalization and predictive performance<sup>56</sup>. Through hardware optimization, tree pruning, and parallelization, the XGBoost method enhances the system's performance<sup>55</sup>.

### Categorical boosting (CatBoost)

CatBoost is a ML ensemble technique that belongs to the GBDT family<sup>58</sup>. CatBoost is a gradient boosting method that makes use of binary decision trees as its key predictor<sup>59</sup>. Ordered target encoding is the basic concept of CB's preprocessing of categorical features. It works by randomly sorting the dataset and then only using the objects that came before the current object to compute the categorical feature's numerical conversion<sup>60</sup>. In the event where the  $k^{\text{th}}$  sample's  $i^{\text{th}}$  feature is a categorical feature, the conversion formula can be written as

$$x_k^i = \frac{\sum_{x_j \in D_K} \{x_k^i = x_j^i\} \times y_j + a \times p}{\sum_{x_j \in D_K} \{x_k^i = x_j^i\} + a} \quad (3)$$

where  $p$  is the added prior item, ' $a$ ' is usually a weighting coefficient greater than 0.  $D_k$  is the dataset before the  $k^{\text{th}}$  sample in the random ordering, and  $\{x_k^i = x_j^i\} = 1$  when  $x_k^i$  and  $x_j^i$  belong to the same category and  $\{x_k^i = x_j^i\} = 0$  when they belong to different categories<sup>60</sup>.

### Performance evaluation

The evaluation stage involves the computation of diverse assessment metrics, encompassing coefficient of determination ( $R^2$ ), root mean squared error (RMSE), mean absolute error (MAE), root mean square standard deviation ratio (RSR), and Nash-sutcliffe efficiency (NSE). These metrics serve to gauge the efficacy of the model's performance, shedding light on the extent to which the model's predictions correlate with the actual target values. The formulations used to calculate these performance metrics are expressed in Eqs. (4)–(10)<sup>61–67</sup>.

$$R^2 = 1 - \frac{\sum_{i=1}^n (d_i - y_i)^2}{\sum_{i=1}^n (d_i - d_{mean})^2} \quad (4)$$

$$r = \frac{\sum_{i=1}^n (d_i - d_{mean})(y_i - y_{mean})}{\sqrt{\sum_{i=1}^n (d_i - d_{mean})^2 \cdot \sum_{i=1}^n (y_i - y_{mean})^2}} \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_i - y_i)^2} \tag{6}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (d_i - y_i)^2 \tag{7}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |(y_i - d_i)| \tag{8}$$

$$RSR = \frac{\sqrt{\sum_{i=1}^n (d_i - y_i)^2}}{\sqrt{\sum_{i=1}^n (d_i - d_{mean})^2}} \tag{9}$$

$$NSE = 1 - \frac{\sum_{i=1}^n (d_i - y_i)^2}{\sum_{i=1}^n (d_i - d_{mean})^2} \tag{10}$$

where  $d_i$  is the  $i$ th observed value,  $y_i$  is the  $i$ th predicted value,  $d_{mean}$  is the mean value of the observed values,  $y_{mean}$  is the mean value of the predicted values and  $n$  is the number of data points. The optimal model exhibits  $R^2$ ,  $r$ , and NSE values close to 1, with minimal RMSE, MSE, MAE, and RSR approaching 0.

Model development and evaluation

The data set of 247 post-liquefaction in-situ free-face ground condition case studies were divided into training (80%) and testing (20%) sets to maintain the statistical consistency of all the input variables. AdaBoost, XGBoost and CatBoost algorithms were used to create three models based on the training dataset and tested on the testing dataset. The parameter range was predefined and used a random search strategy to tune hyperparameters, which was used only on the training subset to achieve the best results as indicated in Table 3. The early stopping criteria is used to stop training when the model’s performance i.e., root mean squared error (RMSE) value on a testing set no longer improves or starts to degrade. It helps prevent overfitting by halting the training process before the model starts to memorize the noise in the data. The performance of the model was tested based on seven statistical parameters:  $R^2$ ;  $r$ ; MAE; MSE; RMSE; RSR ratio and NSE coefficient.

| Algorithm | Hyperparameter                                     | Value         | Range    |
|-----------|----------------------------------------------------|---------------|----------|
| XGBoost   | Number of trees                                    | 70            | 50–200   |
|           | Learning rate                                      | 0.162         | 0.01–0.3 |
|           | Lambda (L2 regularization)                         | 1             | 0–5      |
|           | Maximum tree depth                                 | 7             | 3–10     |
|           | Subsample (fraction of training instances)         | 0.6           | 0.5–1.0  |
|           | Colsample_bytree (fraction of features per tree)   | 1             | 0.5–1.0  |
|           | Colsample_bylevel (fraction of features per level) | 1             | 0.5–1.0  |
|           | Colsample_bynode (fraction of features per split)  | 0.4           | 0.3–1.0  |
|           |                                                    |               |          |
| CatBoost  | Number of trees                                    | 69            | 50–200   |
|           | Learning rate                                      | 0.268         | 0.01–0.3 |
|           | Lambda (L2 regularization)                         | 3             | 0–5      |
|           | Maximum tree depth                                 | 6             | 3–10     |
|           | Colsample_bytree (fraction of features per tree)   | 1             | 0.5–1.0  |
| AdaBoost  | Base estimator                                     | Decision Tree | –        |
|           | Number of estimators                               | 181           | 50–250   |
|           | Learning rate                                      | 1             | 0.01–2.0 |
|           | Classification algorithm                           | SAMME         | –        |
|           | Regression loss function                           | Square        | –        |

Table 3. Optimal hyperparameters and the search range.

## Results and discussions

### Performance of AdaBoost model

Coefficient of determination ( $R^2$ ), correlation coefficient ( $r$ ), mean absolute error (MAE), root mean square error (RMSE), ratio of RMSE to standard deviation of observed value (RSR) and Nash-Sutcliffe coefficient (NSE) were employed to assess the accuracy and precision of AdaBoost model. The trend line for AdaBoost model, for both testing and training phase respectively is plotted by comparing the actual regression, depicted in Figs. 2 and 3. The AdaBoost model shows a comparatively larger error dispersion in both training and testing datasets, with clear deviations from the zero-error line, according to the residual error plots in Fig. 4. This implies that while the model captures broad patterns, it is inconsistent and may exhibit moderate prediction bias, especially for data that has for unseen data.

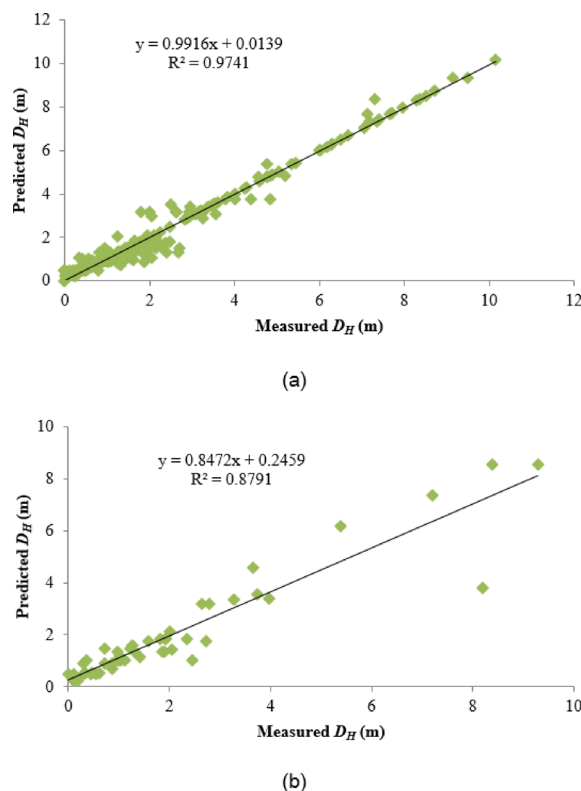
Table 4 shows obviously that  $R^2=0.9741$ ,  $r=0.987$ ,  $MAE=0.2037$ ,  $MSE=0.1324$ ,  $RMSE=0.3638$ ,  $RSR=0.162$  and  $NSE=0.9737$  for the training dataset whereas  $R^2=0.8791$ ,  $r=0.9376$ ,  $MAE=0.4501$ ,  $MSE=0.5918$ ,  $RMSE=0.7693$ ,  $RSR=0.3506$  and  $NSE=0.8771$  for the testing dataset.

### Performance of CatBoost model

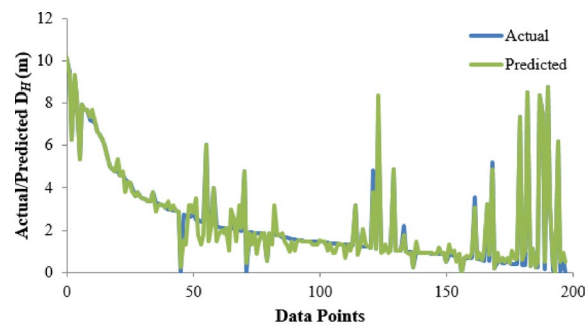
The graphs of actual and predicted data for testing and training datasets of CatBoost model have plotted in Figs. 5 and 6. For training model  $R^2=0.9618$ ,  $r=0.9807$ ,  $MAE=0.3272$ ,  $MSE=0.1988$ ,  $RMSE=0.4461$ ,  $RSR=0.1987$  and  $NSE=0.9605$  and for testing model  $R^2=0.9297$ ,  $r=0.9642$ ,  $MAE=0.3845$ ,  $MSE=0.0485$ ,  $RMSE=0.6145$ ,  $RSR=0.28$  and  $NSE=0.9216$ . In comparison to AdaBoost, the CatBoost model exhibits better residual clustering around zero, as seen in Fig. 7; nonetheless, considerable spread and outliers remain, particularly during the testing stage. Although there is still some variation in its generalization performance, this indicates improved learning capacity and decreased bias.

### Performance of XGBoost model

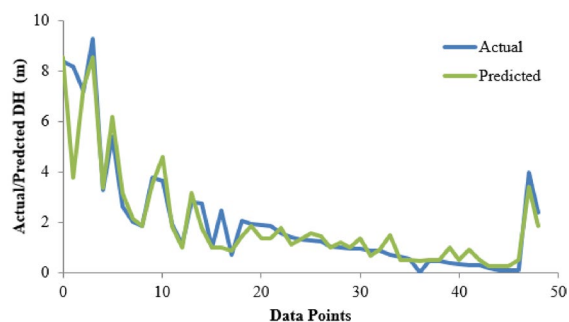
The trend line is provided for the observed regression,  $R^2=0.9905$  for training model and  $R^2=0.9251$  for testing XGBoost model (see Figs. 8 and 9). It could be noted from Table 4, for the training XGBoost model with  $r=0.9952$ ,  $MAE=0.1491$ ,  $MSE=0.0485$ ,  $RMSE=0.2203$ ,  $RSR=0.0981$  and  $NSE=0.9904$  and for the XGBoost model in testing dataset with  $r=0.9618$ ,  $MAE=0.3642$ ,  $MSE=0.3723$ ,  $RMSE=0.6101$ ,  $RSR=0.278$  and  $NSE=0.9227$ . Figure 10 shows that the XGBoost model's residuals around the zero line are more compact and uniformly distributed on both the training and testing datasets. The lower dispersion and fewer outliers suggest improved prediction accuracy, resilience, and excellent generalizability. According to the analysis of Figs. 4 and 7, and 10, the XGBoost model outperforms AdaBoost and CatBoost by having the most consistent and least dispersed residuals, indicating it as the most dependable and accurate model for predicting  $D_H$  in the study.



**Fig. 2.** Scatter plot of AdaBoost model (a) training and (b) testing datasets.

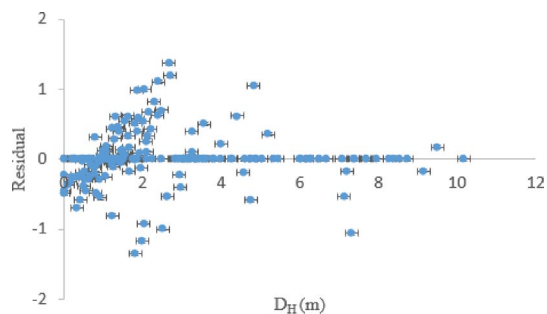


(a)

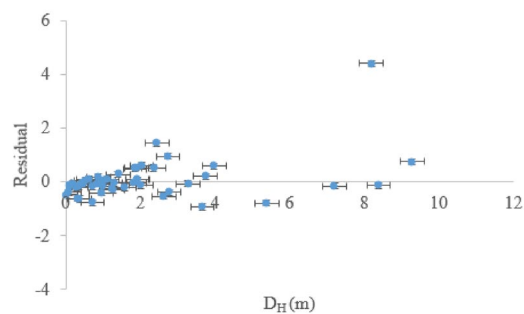


(b)

**Fig. 3.** Line graph of AdaBoost model showing the actual  $D_H$  values against predicted  $D_H$  (a) training and (b) testing datasets.



(a)



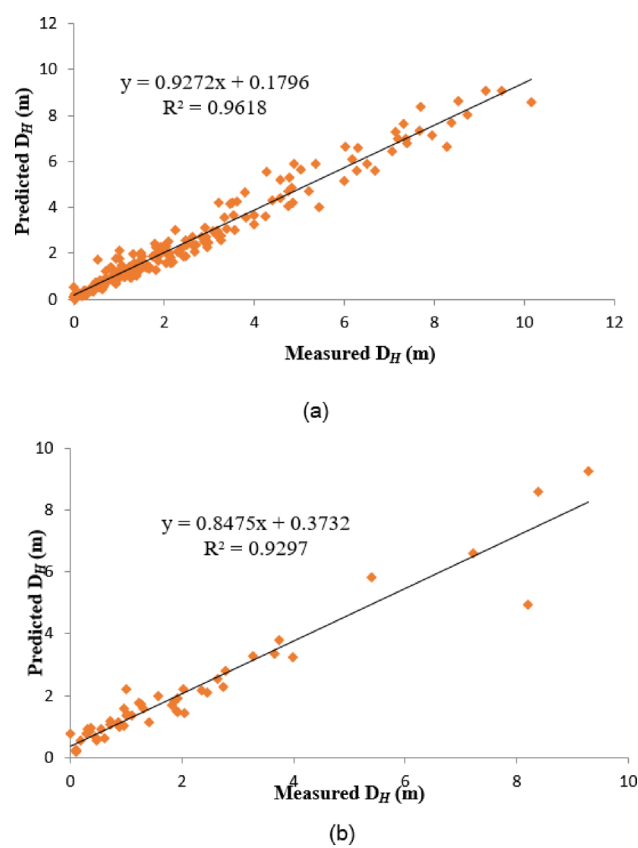
(b)

**Fig. 4.** Residual error plot of the AdaBoost model for (a) training and (b) testing datasets.



| Model                                    | Dataset  | R <sup>2</sup> | r      | MAE    | MSE    | RMSE   | RSR    | NSE    | Reference                    |
|------------------------------------------|----------|----------------|--------|--------|--------|--------|--------|--------|------------------------------|
| AdaBoost                                 | Training | 0.9741         | 0.987  | 0.2037 | 0.1324 | 0.3638 | 0.162  | 0.9737 | This study                   |
|                                          | Testing  | 0.8791         | 0.9376 | 0.4051 | 0.5918 | 0.7693 | 0.3506 | 0.8771 |                              |
| CatBoost                                 | Training | 0.9618         | 0.9807 | 0.3272 | 0.1988 | 0.4461 | 0.1987 | 0.9605 |                              |
|                                          | Testing  | 0.9297         | 0.9642 | 0.3845 | 0.3772 | 0.6145 | 0.28   | 0.9216 |                              |
| XGBoost                                  | Training | 0.9905         | 0.9952 | 0.1491 | 0.0485 | 0.2203 | 0.0981 | 0.9904 |                              |
|                                          | Testing  | 0.9251         | 0.9618 | 0.3642 | 0.3723 | 0.6101 | 0.278  | 0.9227 |                              |
| GPR                                      | Training | 0.9402         | 0.9697 | 0.3403 | –      | 0.5597 | 0.248  | 0.938  | Ahmad et al. <sup>28</sup>   |
|                                          | Testing  | 0.894          | 0.9455 | 0.544  | –      | 0.8438 | 0.387  | 0.851  |                              |
| Evolutionary Polynomial Regression (EPR) | Training | 0.913          | –      | 0.537  | –      | 1.003  | –      | –      | Rezania et al. <sup>68</sup> |
|                                          | Testing  | 0.883          | –      | 0.291  | –      | 1.158  | –      | –      |                              |
| Artificial Neural Networks (ANN)         | Training | 0.875          | –      | 0.702  | –      | 1.074  | –      | –      |                              |
|                                          | Testing  | 0.872          | –      | 0.82   | –      | 1.21   | –      | –      |                              |
| MLR                                      | Training | 0.868          | –      | 0.81   | –      | 1.24   | –      | –      |                              |
|                                          | Testing  | 0.875          | –      | 0.43   | –      | 1.196  | –      | –      |                              |

**Table 4.** Performance statistics of models in comparison with extra available models in literature. “–” not reported in the respective reference.

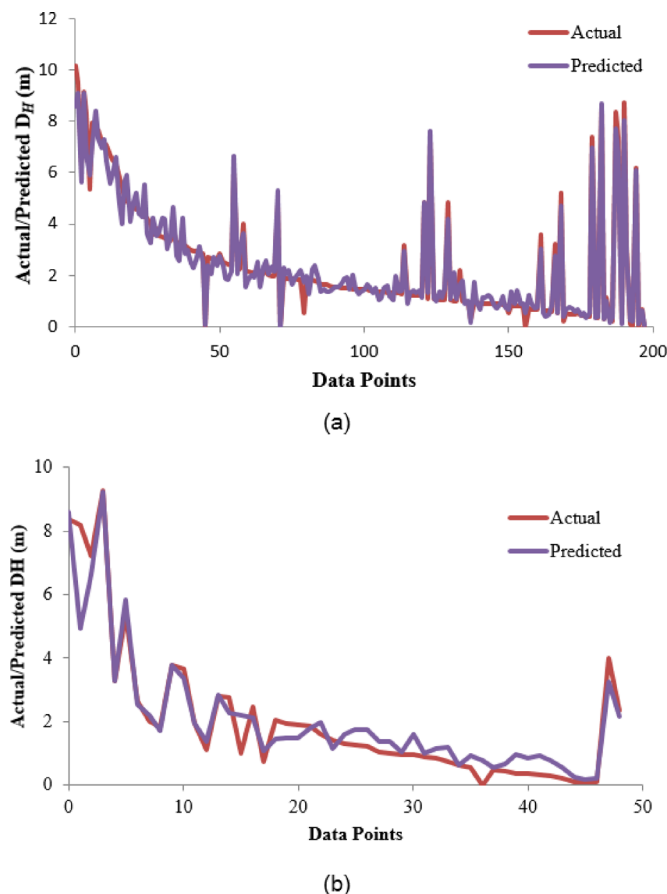


**Fig. 5.** Scatter plot of CatBoost model (a) training and (b) testing datasets.

**Performance evaluation with literature models**

In the testing phase, the performance of the developed ensemble learning models—AdaBoost, CatBoost, and XGBoost—was evaluated and compared with existing literature models using key statistical indicators, including  $R^2$ ,  $r$ , MAE, MSE, RMSE, RSR, and NSE. As in Table 4, the performance of XGBoost was most favorable as it scored an  $R^2$  of 0.9251,  $r$  of 0.9618, MAE of 0.3642, MSE of 0.3723, RMSE of 0.6101, RSR of 0.278, and NSE of 0.9227. These findings show great predictive reliability and show great concordance with observed values. CatBoost closely came next with highest  $R^2$  (0.9297) and  $r$  (0.9642), indicating slightly better fit and correlation with observed data, yet has slightly high prediction errors (MAE=0.3845, MSE=0.3772, RMSE=0.6145, RSR=0.280, NSE=0.9216) than the XGBoost. Conversely, AdaBoost produced the worst result on all the measures with





**Fig. 6.** Line graph of CatBoost model presenting the actual  $D_H$  values versus the predicted  $D_H$  (a) training and (b) testing datasets.

$R^2=0.8791$ ,  $r=0.9376$ ,  $MAE=0.4051$ ,  $MSE=0.5918$ ,  $RMSE=0.7693$ ,  $RSR=0.3506$  and  $NSE=0.8771$ , reflecting lower predictive capability. The results of testing dataset (see Table 4) demonstrated that XGBoost produced the greatest predictive accuracy with  $R^2=0.9905$  (training) and  $0.9251$  (testing), correlation coefficients ( $r$ )= $0.9952$  and  $0.9618$ , and the error metrics ( $MAE=0.1491$ ,  $MSE=0.0485$ ,  $RMSE=0.2203$  in training;  $MAE=0.3642$ ,  $MSE=0.3723$ ,  $RMSE=0.6101$  in testing) as compared to the Gaussian Process Regression (GPR) with ( $R^2=0.9402/0.894$ ,  $RMSE=0.5597/0.8438$  in training/testing) against EPR ( $R^2=0.913/0.883$ ,  $RMSE=1.003/1.157$ ), ANN ( $R^2=0.875/0.872$ ,  $RMSE=1.074/1.21$ ) and MLR ( $R^2=0.868/0.875$ ,  $RMSE=1.24/1.196$ ) in literature. For XGBoost, it was observed a slight gap between training and testing performance, indicating slight overfitting; this is likely due to its high model complexity and sensitivity to hyperparameter settings. Regarding AdaBoost, its comparatively weaker testing performance can be attributed to its sequential weighting mechanism, which makes it more sensitive to noisy data and outliers, reducing generalization compared to XGBoost and CatBoost. Overall, the ensemble models, particularly the XGBoost, showed superior predictive ability and generalization during the testing, compared with all the baseline models across all the evaluation metrics including,  $R^2$ ,  $r$ ,  $MAE$ ,  $MSE$ ,  $RMSE$ ,  $RSR$ , and  $NSE$ .

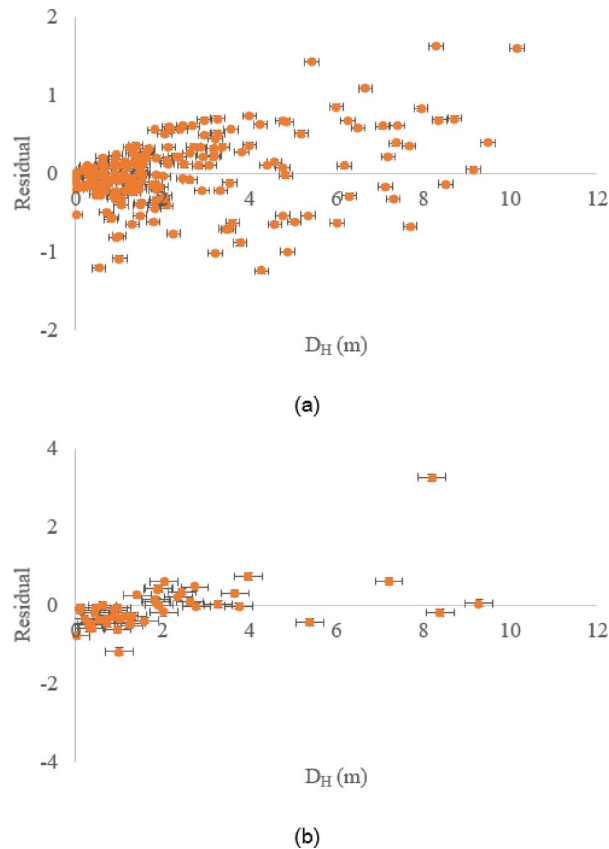
### Radar plot analysis

Figure 11 shows a radar chart that graphically demonstrates the relative strengths of each model, which supports the identified trends. The results combined make it clear that XGBoost is the most successful model in the given regression task when tested.

The XGBoost model here proposed is a superior than the current available models in literature since it is more accurate and robust in predictive power but low on computational overhead. It can capture high-dimensional, nonlinear interactions between features using little data preprocessing, leading to more efficient and better generalization of models.

### Rank analysis

To obtain a fair and comprehensive comparison among all the proposed machine learning models, rank analysis was carried out using several metrics (e.g.,  $RMSE$ ,  $MAE$ ,  $R^2$ ,  $NSE$ ). Models were ranked separately by each metric such that the model with the highest score acquires rank 1, then lower ranks sequentially. The numerical sum of ranks of all metrics was then calculated to obtain the overall position of each model. The statistical performance indicators were analyzed using rank analysis to determine the efficiency of the training and testing models.



**Fig. 7.** Residual error plot of the CatBoost model for (a) training and (b) testing datasets.

Table 5 indicates that XGBoost model scored the best in terms of accuracy on testing, with a rank value of 19, followed by CatBoost that had a rank value of 18. In the training phase, XGBoost was ranked the best one being ranked highest with a total rank value of 21 as compared to the other models therefore proving it to be robust and reliable irrespective of the dataset.

#### Wilcoxon signed-rank test

The Wilcoxon Signed-Rank Test was used to statistically compare the differences between the three developed models, AdaBoost, CatBoost, and XGBoost. The test was used to compare the predictive performance of the models. The null hypothesis ( $H_0$ ) states that the median difference between paired observations is zero, indicating no significant difference. A significance level of  $\alpha=0.05$  was used. The findings demonstrate that AdaBoost and CatBoost differ considerably ( $p=0.011$ ), as do CatBoost and XGBoost ( $p=0.010$ ), whereas AdaBoost and XGBoost do not differ significantly ( $p=0.712$ ). This implies that CatBoost's predictions diverge more than the others, although AdaBoost and XGBoost are more consistent. The statistical study shows that AdaBoost and XGBoost produce more consistent predictions in comparison to the observed values, as indicated by the non-significant differences. CatBoost, on the other hand, behaves statistically differently, which could reflect model bias or sensitivity to input features. The lack of a substantial difference between AdaBoost and XGBoost indicates that both models perform similarly in capturing the dataset's underlying patterns. As a result, these models may be considered reliable for prediction purposes in the specific application.

#### Sensitivity analysis

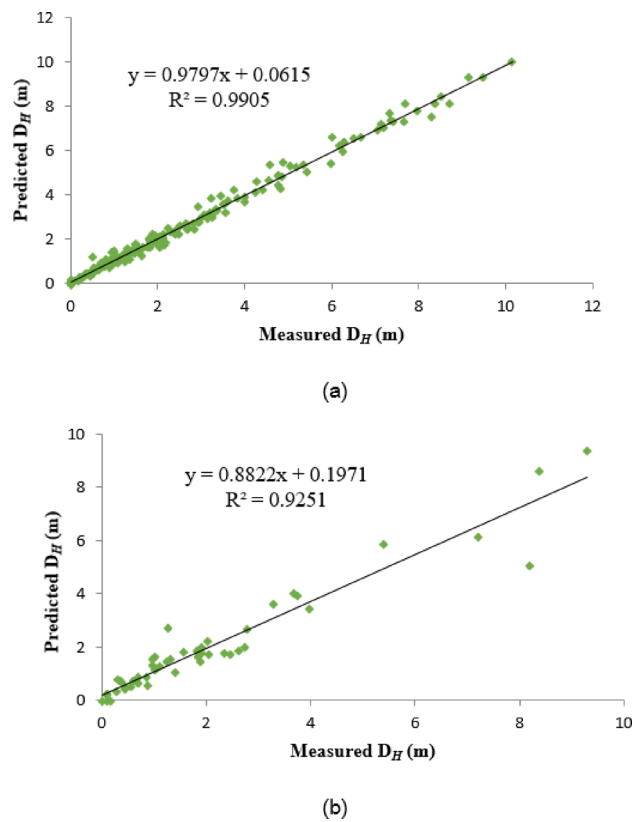
The top-ranked XGBoost model (rank value=21) was used in sensitivity analysis with the Yang and Zhang method as used by Ahmad et al. [69] to determine the individual effect of each input parameter on the prediction of liquefaction-induced lateral displacement and is presented in the Eq. (11):

$$r_{ij} = \frac{\sum_{k=1}^n (Y_{ik} \times Y_{ok})}{\sqrt{\sum_{k=1}^n Y_{ik}^2 \sum_{k=1}^n Y_{ok}^2}} \quad (11)$$

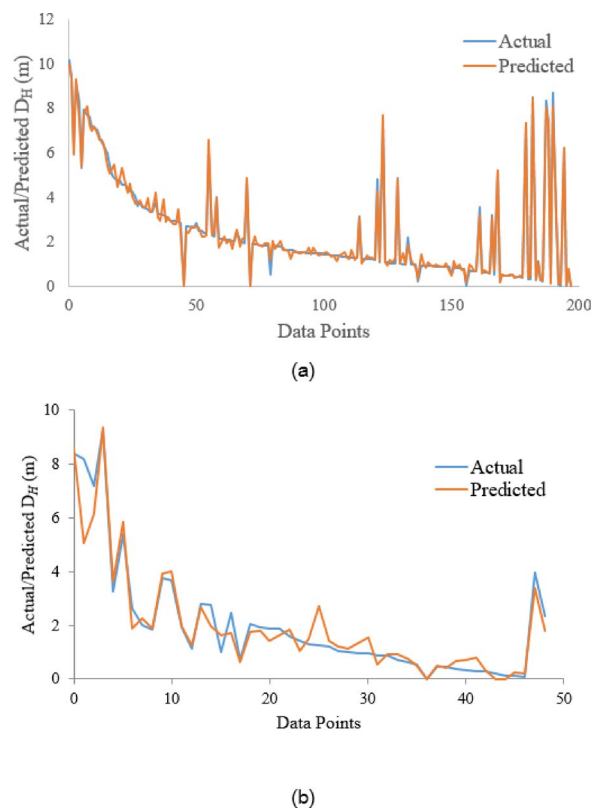
where:  $r_{ij}$  correlation coefficient representing the effect of the  $i$ -th input variable on the  $j$ -th output parameter. Its value is restricted between 0 and 1; a larger value indicates a stronger influence of the input on the output (here,  $D_H$ ).

$i$ : index of the input variable.

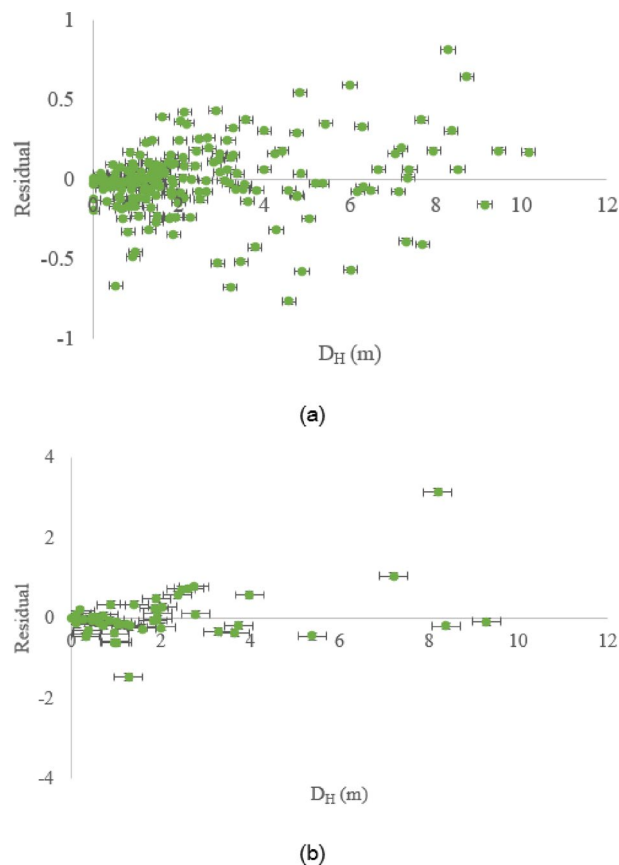
$j$ : index of the output parameter.



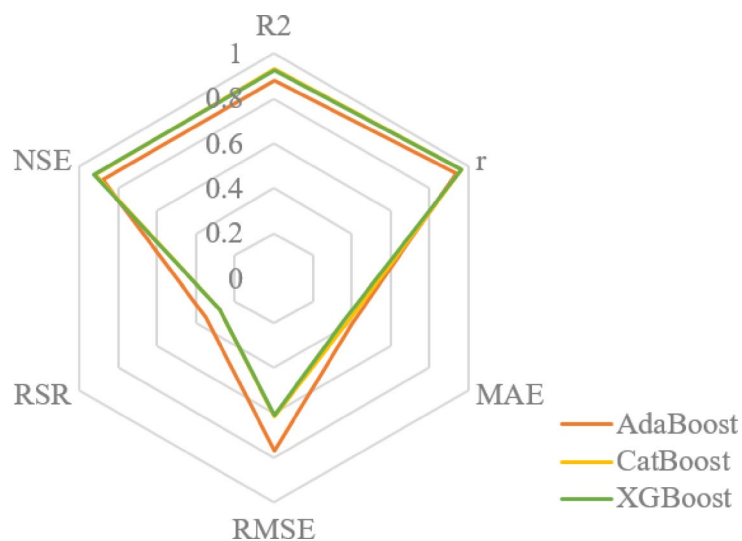
**Fig. 8.** Scatter plot of XGBoost model (a) training and (b) testing datasets.



**Fig. 9.** Line graph of the XGBoost model presenting the actual  $D_H$  values versus the predicted  $D_H$  (a) training and (b) testing datasets.



**Fig. 10.** Residual error plot of the XGBoost model for (a) training and (b) testing datasets.



**Fig. 11.** Radar plot shown on a comparison of AdaBoost performance, CatBoost performance and XGBoost performance on the testing phase of six evaluation metrics.

$o$ : refers to the observed output values.

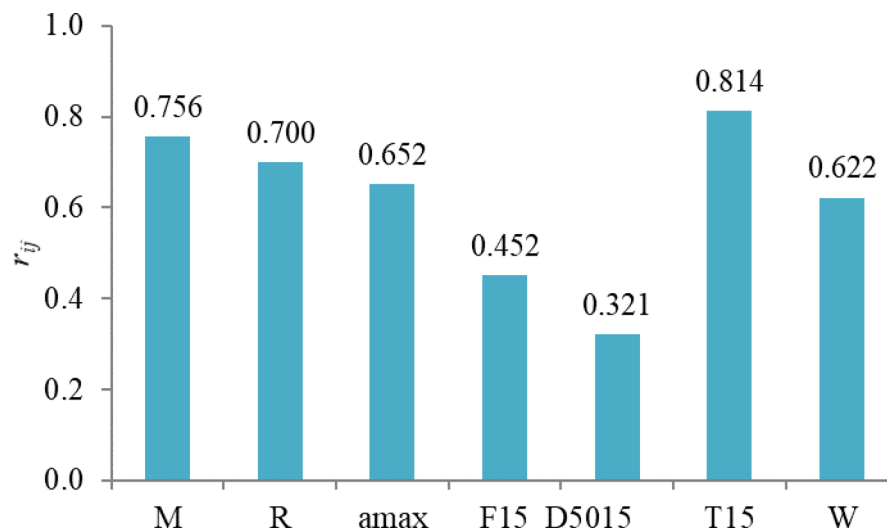
$k$ : index of the data point, with  $k = 1, 2, \dots, n$ .

In this study,  $n = 198$  (number of data points),  $Y_{ik}$  is the value of the  $i$ -th input for the  $k$ -th record, and  $Y_{ik}$  is the corresponding observed output.

The  $r_{ij}$  values reflect the specific input–output sensitivities observed for the XGBoost model used in this study and may differ under alternative modeling approaches. The findings from the Fig. 12 revealed that

| Dataset  | Model    | Performance Indicators |        |        |       |     |      | Score | References                   |
|----------|----------|------------------------|--------|--------|-------|-----|------|-------|------------------------------|
|          |          | $R^2$                  | MAE    | RMSE   | $R^2$ | MAE | RMSE |       |                              |
| Training | AdaBoost | 0.9741                 | 0.2037 | 0.3638 | 6     | 6   | 6    | 18    | Present study                |
|          | CatBoost | 0.9618                 | 0.3272 | 0.4461 | 5     | 5   | 5    | 15    |                              |
|          | XGBoost  | 0.9905                 | 0.1491 | 0.2203 | 7     | 7   | 7    | 21    |                              |
|          | GPR      | 0.9402                 | 0.3403 | 0.5597 | 4     | 4   | 4    | 12    | Ahmad et al. <sup>28</sup>   |
|          | EPR      | 0.913                  | 0.537  | 1.003  | 3     | 3   | 3    | 9     | Rezania et al. <sup>68</sup> |
|          | ANN      | 0.875                  | 0.702  | 1.074  | 2     | 2   | 2    | 6     |                              |
|          | MLR      | 0.868                  | 0.81   | 1.24   | 1     | 1   | 1    | 3     |                              |
| Testing  | AdaBoost | 0.8791                 | 0.4051 | 0.7693 | 3     | 4   | 5    | 12    | Present study                |
|          | CatBoost | 0.9297                 | 0.3845 | 0.6145 | 7     | 5   | 6    | 18    |                              |
|          | XGBoost  | 0.9251                 | 0.3642 | 0.6101 | 6     | 6   | 7    | 19    |                              |
|          | GPR      | 0.894                  | 0.544  | 0.8438 | 5     | 2   | 4    | 11    | Ahmad et al. <sup>28</sup>   |
|          | EPR      | 0.883                  | 0.291  | 1.158  | 4     | 7   | 3    | 14    | Rezania et al. <sup>68</sup> |
|          | ANN      | 0.872                  | 0.82   | 1.21   | 1     | 1   | 1    | 3     |                              |
|          | MLR      | 0.875                  | 0.43   | 1.196  | 2     | 3   | 2    | 7     |                              |

**Table 5.** Rank analysis of machine learning models based on statistical performance indicators for training and testing stages.

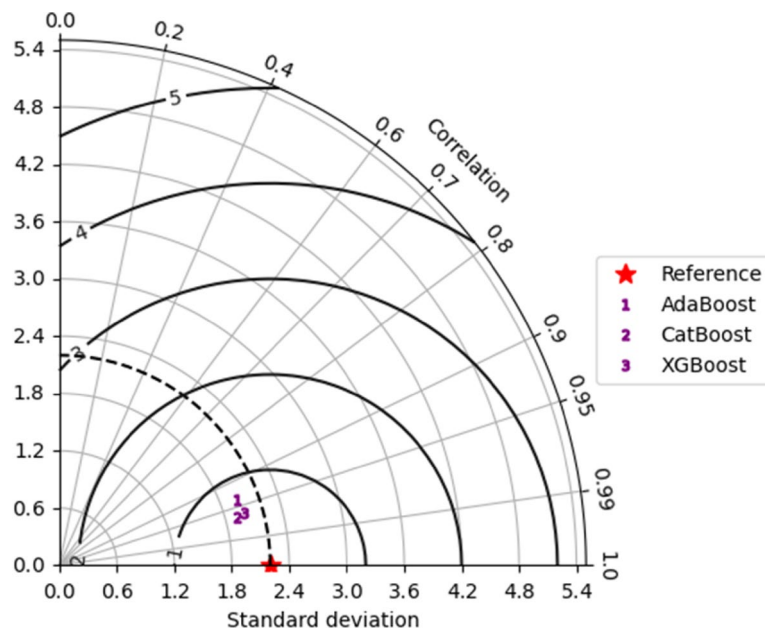


**Fig. 12.** Sensitivity analysis of input parameters.

accumulative thickness of saturated layers with adjusted SPT number SPT number  $(N_1)_{60} < 15$  ( $T_{15}$ ) was the most sensitive (0.814), followed by earthquake magnitude (M) at 0.756 and horizontal distance to the seismic energy source (R) at 0.700. Peak ground acceleration ( $a_{max}$ ) and free-face ratio (W) were also found to be highly influential with sensitivities of 0.652 and 0.622 respectively as compared to the average fines content in  $T_{15}$  ( $F_{15}$ ) and average particle size in  $T_{15}$  ( $D_{5015}$ ) with sensitivities of 0.452 and 0.321 respectively. From a geotechnical perspective, mitigation of the most influential factors may be realized by (i) decreasing  $T_{15}$  through densification or replacement of liquefiable layers, drainage or preloading to reduce saturation potential; (ii) mitigating the seismic demand level as a result of high M and  $a_{max}$  by site selection, setback, or methods of ground motion isolation techniques; and (iii) altering geometry or slope profile by decreasing the free-face ratio (W), and hence lowering driving forces of lateral displacement. These confined actions respond to the most sensitive parameters thus minimizing the lateral displacement facilitating the enhanced site resilience to the risks of liquefaction-induced hazards.

### Taylor diagram of model performance

The level of agreement can be characterized conveniently in terms of a Taylor diagram (see Fig. 13). The radial distance from the origin here is plotted as standard deviation. The RMS error scales with the difference in standard deviation units of actual and projected fields. The azimuthal angle refers to the correlation coefficient [70]. A Taylor diagram was used to assess the overall performance of the machine learning models in a simultaneous comparison of correlation coefficient, standard deviation and root mean square error to the observed data. This



**Fig. 13.** Taylor diagram for model scores by AdaBoost, CatBoost, and XGBoost on liquefaction-induced lateral displacement predictive performance.

graphical method gives a combined evaluation, with a model that is nearer to the reference point showing more precision and consistency.

The proposed AdaBoost, CatBoost and XGBoost models can predict the liquefaction induced lateral displacement with a high level of accuracy. In addition, the expected accuracy of the model is largely determined by values of parameters used in training and affects by distribution of variables.

## Conclusions

The three models were built for the prediction of liquefaction-induced lateral displacement using AdaBoost, CatBoost and XGBoost algorithms. The models were developed based on seven input parameters and one dependent parameter,  $D_H$ . Statistical indexes including  $R^2$ ,  $r$ , MAE, RMSE, RSR and NSE were applied to evaluate the models. The study dataset included 247 cases of which 198 were used for training and 49 were used to evaluate each model. The XGBoost has  $R^2=0.9905$ , MAE=0.1491, MSE=0.0485, RMSE=0.2203 in the training phase while for testing phase  $R^2=0.9251$ , model  $r=0.9618$ , MAE=0.3642, MSE=0.3723, RMSE=0.6101. For the training phase CatBoost has  $R^2=0.9618$ ,  $r=0.9807$ , MAE=0.3272, MSE=0.1988, RMSE=0.4461, whereas for the testing phase  $R^2=0.9297$ ,  $r=0.9642$ , MAE=0.3845, MSE=0.3772, RMSE=0.6145. AdaBoost has  $R^2=0.9741$ ,  $r=0.987$ , MAE=0.2037, MSE=0.1324, RMSE=0.3638 for the training phase and  $R^2=0.8791$ ,  $r=0.9376$ , MAE=0.4501, MSE=0.5918, RMSE=0.7693 for the testing phase. Based on the comparison, it can be concluded that XGBoost and CatBoost give better performances in comparison to AdaBoost, GPR, EPR ANN and MLR. The sensitivity analysis showed that the values of parameters  $T_{15}$  ( $r_{ij}=0.814$ ),  $M$  ( $r_{ij}=0.756$ ) and  $R$  ( $r_{ij}=0.700$ ) were important for the output parameter. The  $r_{ij}$  value of  $D50_{15}$  is the lowest (0.321) compared with other seven parameters, and the parameter  $T_{15}$  has the most influence on the output parameter,  $D_H$ . The study is confined to the free-face liquefaction scenario, training and validation of the models was carried out within the ranges of compiled geotechnical, seismic, and topographic parameters and is therefore not extrapolate to other site conditions or extreme loading. The data-driven model does not have constitutive behavior of soil embedded in it, which weakens physical interpretability. Future research should extend the data range to other sites and seismic conditions and implement physics-informed or hybrid ML solutions to combine empirical learning and the geomechanical framework to achieve better levels of robustness and generalization.

## Data availability

All data supporting the findings of this study are available within the paper and its Supplementary Information.

Received: 18 September 2025; Accepted: 22 April 2026

Published online: 30 April 2026

## References

1. A. F. Rauch, "EPOLLS: an empirical method for predicting surface displacements due to liquefaction-induced lateral spreading in earthquakes," 1997.
2. L. Su, L. Tang, X. Ling, C. Liu, X. J. S. D. Zhang, and E. Engineering, "Pile response to liquefaction-induced lateral spreading: a shake-table investigation," vol. 82, pp. 196–204, 2016.

3. W. Araujo and C. J. A. S. Ledezma, "Factors that affect liquefaction-induced lateral spreading in large subduction earthquakes," vol. 10, no. 18, p. 6503, 2020.
4. D. Ranjan Kumar and W. J. S. R. Wipulanusat, "Advancements in predicting soil liquefaction susceptibility: a comprehensive analysis of ensemble and deep learning approaches," vol. 15, no. 1, p. 26453, 2025.
5. Kumar, D. R., Samui, P., Burman, A., Biswas, R. & Vanapalli, S. Jjo. E. S. S. A novel approach for assessment of seismic induced liquefaction susceptibility of soil. *J. Earth Syst. Sci.* **133**(3), 128 (2024).
6. Y. Yang, W. Yang, S. Xie, C. Feng, and Z. Huang, "Experimental Research on Liquefied Sand by Compaction Piles with Equal Energy and Equal Deformation in High Intensity Seismic Region," in *International Conference on Civil Architecture and Structural Engineering*, 2024, pp. 338–347: Springer.
7. A. T. Goh and W. J. E. g. Zhang, "An improvement to MLR model for predicting liquefaction-induced lateral spread using multivariate adaptive regression splines," vol. 170, pp. 1–10, 2014.
8. Z. Chen, C. Tan, C. Wang, Y. Chen, J. Chen, and Y. J. G. J. Wang, "Evaluation of liquefaction-induced lateral spread displacement based on ensemble learning," vol. 59, no. 9, pp. 2534–2548, 2024.
9. Yin, Q., Xin, T., Zhenggang, H. & Minghua, H. Jai. C. E. Measurement and analysis of deformation of underlying tunnel induced by foundation pit excavation. *Adv. Civ. Eng.* **2023**(1), 8897139 (2023).
10. Y. Li, X. Weng, D. Hu, Z. Tan, J. J. o. P. S. E. Liu, and Practice, "Data-driven method for predicting long-term underground pipeline settlement induced by rectangular pipe jacking tunnel construction," vol. 16, no. 3, p. 04025046, 2025.
11. Feifei, W. Jjo. M. S. Formation mechanism of high-cold and high-altitude landslide disasters caused by complex goaf groups. *J. Min. Sci.* **61**(1), 71–79 (2025).
12. T. Ma, L. Wu, J. Zhou, H. Zhang, and H. J. N. H. Xiao, "An interpretable hybrid model for predicting step-like landslide displacement: a case study in the Three Gorges Reservoir," vol. 121, no. 18, pp. 21441–21458, 2025.
13. Y. Zheng and B. A. J. A. G. Baudet, "Evolution of nano-pores in illite-dominant clay during consolidation," vol. 19, no. 1, pp. 71–83, 2024.
14. X. Li, Y. Shen, W. Chen, S. Feng, Z. Zhang, and H. J. O. E. Hong, "Seismic response of ice-water-covered two-layer seabed under obliquely incident P/SV-wave considering imperfect interface condition," vol. 342, p. 122872, 2025.
15. Jia, C., Cheng, S., Li, L., Chen, Y. J. T., U. S. Technology. Seismic ahead-prospecting method based on delayed blasting excitation in the tunnel face: A case study. *Tunn. Undergr. Space Technol.* **161**, 106577 (2025).
16. H. Zhong, Z. Zhu, D. Zhang, B. Zhang, and X. J. n. H. S. Zhang, "Discrete element simulation of damage mechanisms in stone cultural relics under dynamic impact," vol. 13, no. 1, p. 249, 2025.
17. K. Fang, D. Y. Tan, C. C. Zhang, S. G. Shi, H. W. Sang, and B. J. G. R. L. Shi, "Nature-based profiling of subsurface soil stiffness driven by tidal forces," vol. 52, no. 21, p. e2025GL118702, 2025.
18. S. Peng *et al.*, "Laboratory investigation effects of control measures for leakage-induced erosion on seepage interactions in defective underground structures," vol. 161, p. 106593, 2025.
19. X. Qu, T. Zhao, J. J. R. E. Cao, and S. Safety, "Improved AHP and dual-hidden layer adaptive regularized neural network model for the quantitative risk assessment of oil and gas pipelines," p. 112028, 2025.
20. Z. Hu, J. Li, J. Cui, G. J. S. D. Lin, and E. Engineering, "Seismic evaluation of nuclear structures under structure-soil-structure interaction: Effects of ray-tracing wave propagation and damping calibration," vol. 199, p. 109649, 2025.
21. Yu, J. *et al.* Prediction of jacking force in circular press-in caisson in clay. *Can. Geotech. J.* **63**, 1–17 (2026).
22. Zhuang, P.-Z. *et al.* CASM-U: A unified critical state model for unsaturated clays and sands. *Acta Geotech.* **20**(1), 211–230 (2025).
23. Liu, X. *et al.* Cascade control method for hydraulic secondary regulation drive system based on adaptive robust control. *ISA Trans.* **156**, 479–489 (2025).
24. Z. Wang *et al.*, "Displacement Control Method of Hydraulic Servomotor Based on Swashplate Moment Disturbance Compensation," 2025.
25. Yegian, M., Marciano, E. & Ghahraman, V. G. Jjo. G. E. Earthquake-induced permanent deformations: Probabilistic approach. *Journal of Geotechnical Engineering* **117**(1), 35–50 (1991).
26. Gu, W., Morgenstern, N. & Robertson, P. J. Joge. Progressive failure of lower San Fernando dam. *Journal of Geotechnical Engineering* **119**(2), 333–349 (1993).
27. W. Gu, N. Morgenstern, and P. J. J. o. g. e. Robertson, "Postearthquake deformation analysis of Wildlife site," vol. 120, no. 2, pp. 274–289, 1994.
28. M. Ahmad *et al.*, "Prediction of liquefaction-induced lateral displacements using Gaussian process regression," vol. 12, no. 4, p. 1977, 2022.
29. I. H. J. S. c. s. Sarker, "AI-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems," vol. 3, no. 2, p. 158, 2022.
30. M. Madanchian and H. J. R. i. E. Taherdoost, "The impact of artificial intelligence on research efficiency," p. 104743, 2025.
31. Shehadeh, A., Alshboul, O., Al-Shboul, K. F. & Tatari, O. An expert system for highway construction: Multi-objective optimization using enhanced particle swarm for optimal equipment management. *Expert Syst. Appl.* **249**, 123621 (2024).
32. Alshboul, O., Al Mamlook, R. E., Shehadeh, A. & Munir, T. J. C. Empirical exploration of predictive maintenance in concrete manufacturing: Harnessing machine learning for enhanced equipment reliability in construction project management. *I. Engineering* **190**, 110046 (2024).
33. Alshboul, O., Almasabha, G., Shehadeh, A. & Al-Shboul, K. A comparative study of LightGBM, XGBoost, and GEP models in shear strength management of SFRC-SBWS. *Structures* **61**, 106009 (2024) (: Elsevier).
34. Al-Shboul, K. F., Almasabha, G., Shehadeh, A. & Alshboul, O. Exploring the efficacy of machine learning models for predicting soil radon exhalation rates. *Stoch. Environ. Res. Risk Assess.* **37**(11), 4307–4321 (2023).
35. Alshboul, O., Shehadeh, A., Almasabha, G., Mamlook, R. E. A. & Almuflih, A. S. Evaluating the impact of external support on green building construction cost: A hybrid mathematical and machine learning prediction approach. *Buildings* **12**(8), 1256 (2022).
36. Halalshah, N. *et al.* Breakthrough curves prediction of selenite adsorption on chemically modified zeolite using boosted decision tree algorithms for water treatment applications. *Water* **14**(16), 2519 (2022).
37. Almasabha, G., Al-Shboul, K. F., Shehadeh, A. & Alshboul, O. Machine learning-based models for predicting the shear strength of synthetic fiber reinforced concrete beams without stirrups. *Structures* **52**, 299–311 (2023) (: Elsevier).
38. Alshboul, O., Almasabha, G., Al-Shboul, K. F. & Shehadeh, A. A comparative study of shear strength prediction models for SFRC deep beams without stirrups using machine learning algorithms. *Structures* **55**, 97–111 (2023) (: Elsevier).
39. Kumar, M., Samui, P., Kumar, D. R. & Asteris, P. G. State-of-the-art XGBoost, RF and DNN based soft-computing models for PGPN piles. *Geomech. Geoengin.* **19**(6), 975–990 (2024).
40. M. Kumar, N. Z. Fathima, and D. R. Kumar, "A novel XGBoost and RF-based metaheuristic models for concrete compression strength," in *International Conference on Civil Engineering Innovative Development in Engineering Advances*, 2023, pp. 495–503: Springer.
41. Kumar, M., Kumar, D. R. & Wipulanusat, W. J. G. Reliability-based design for strip-footing subjected to inclined loading using hybrid LSSVM ML models. *G. Engineering* **42**(8), 7677–7697 (2024).
42. Zendeheboudi, S., Rezaei, N. & Lohi, A. Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review. *Appl. Energy* **228**, 2539–2566 (2018).
43. Karimai, K., Liu, W. & Maruyama, Y. Prediction and factor analysis of liquefaction ground subsidence based on machine-learning techniques. *Appl. Sci.* **14**(7), 2713 (2024).



44. A. V. Dorogush, V. Ershov, and A. J. a. p. a. Gulin, "CatBoost: gradient boosting with categorical features support," 2018.
45. Ying, C., Qi-Guang, M., Jia-Chen, L. & Lin, G. Advance and prospects of AdaBoost algorithm. *J. A. A. S.* **39**(6), 745–758 (2013).
46. T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
47. Youd, T. L., Hansen, C. M. & Bartlett, S. F. Revised multilinear regression equations for prediction of lateral spread displacement. *J. Geotech. Geoenviron. Eng.* **128**(12), 1007–1017 (2002).
48. K. O. Cetin *et al.*, "Liquefaction-induced lateral spreading at Izmit Bay during the Kocaeli (Izmit)-Turkey earthquake," vol. 130, no. 12, pp. 1300–1313, 2004.
49. Chu, D. B., Stewart, J. P., Youd, T. L. & Chu, B. Liquefaction-induced lateral spreading in near-fault regions during the 1999 Chi-Chi, Taiwan earthquake. *J. Geotech. Geoenviron. Eng.* **132**(12), 1549–1565 (2006).
50. Sadigh, K., Chang, C.-Y., Egan, J., Makdisi, F. & Youngs, R. R. Attenuation relationships for shallow crustal earthquakes based on California strong motion data. *J. S. r. l.* **68**(1), 180–189 (1997).
51. A. A. Javadi, M. Rezaia, M. M. J. C. Nezhad, and Geotechnics, "Evaluation of liquefaction induced lateral displacements using genetic programming," vol. 33, no. 4–5, pp. 222–233, 2006.
52. Y. Jafarian, E. J. A. J. o. C. Nasri, and E. Engineering, "Evaluation of uncertainties in the available empirical models and probabilistic prediction of liquefaction induced lateral spreading," vol. 48, no. 3, pp. 275–290, 2016.
53. C. F. Dormann *et al.*, "Collinearity: a review of methods to deal with it and a simulation study evaluating their performance," vol. 36, no. 1, pp. 27–46, 2013.
54. Schapire, R. E. "Explaining adaboost," in *Empirical inference: festschrift in honor of vladimir N 37–52* (Springer, 2013).
55. Z. A. Ali, Z. H. Abduljabbar, H. A. Tahir, A. B. Sallow, and S. M. J. A. J. o. N. U. Almufti, "eXtreme gradient boosting algorithm with machine learning: A review," vol. 12, no. 2, pp. 320–334, 2023.
56. Hsiao, C.-H., Kumar, K. & Rathje, E. M. JFi. B. E. Explainable AI models for predicting liquefaction-induced lateral spreading. *Front. Built Environ.* **10**, 1387953 (2024).
57. Zou, M., Jiang, W.-G., Qin, Q.-H., Liu, Y.-C. & Li, M.-L.J.M. Optimized XGBoost model with small dataset for predicting relative density of Ti-6Al-4V parts manufactured by selective laser melting. *Materials* **15**(15), 5298 (2022).
58. J. T. Hancock and T. M. J. J. o. b. d. Khoshgoftaar, "CatBoost for big data: an interdisciplinary review," vol. 7, no. 1, p. 94, 2020.
59. L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. J. A. i. n. i. p. s. Gulin, "CatBoost: unbiased boosting with categorical features," vol. 31, 2018.
60. X. Feng, J. He, and B. J. A. G. Lu, "Accurate and generalizable soil liquefaction prediction model based on the CatBoost algorithm," vol. 72, no. 5, pp. 3417–3426, 2024.
61. R. Daneshfaraz, E. Aminvash, S. Di Francesco, A. Najibi, and J. J. N. M. i. C. E. Abraham, "Three-dimensional study of the effect of block roughness geometry on inclined drop," vol. 6, no. 1, pp. 1–9, 2021.
62. F. Kalateh, E. Aminvash, R. J. A. W. I. Daneshfaraz, Ecosystems, and Society, "On the hydraulic performance of the inclined drops: the effect of downstream macro-roughness elements," vol. 73, no. 3, pp. 553–568, 2024.
63. F. Kalateh, E. Aminvash, J. J. E. J. o. E. Abraham, and C. Engineering, "On the effect of flow regime and slope of the channel bed on the hydraulic performance of the sharp-crested rectangular side weir: a numerical simulation," pp. 1–18, 2024.
64. R. Daneshfaraz, E. Aminvash, A. Ghaderi, J. Abraham, and M. J. A. s. Bagherzadeh, "SVM performance for predicting the effect of horizontal screen diameters on the hydraulic parameters of a vertical drop," vol. 11, no. 9, p. 4238, 2021.
65. Abdullah, G. M. S. *et al.* Boosting-based ensemble machine learning models for predicting unconfined compressive strength of geopolymer stabilized clayey soil. *Sci. Rep.* **14**(1), 2323 (2024).
66. Al-Zubi, M. A. *et al.* Long short term memory networks for predicting resilient Modulus of stabilized base material subject to wet-dry cycles. *Sci. Rep.* **14**(1), 27928 (2024).
67. Ahmad, M., Al-Mansob, R. A., Ramli, A. B. B., Ahmad, F. & Khan, B. J. Unconfined compressive strength prediction of stabilized expansive clay soil using machine learning techniques. *Multiscale Multidiscip. Model. Exp. Des.* <https://doi.org/10.1007/s41939-023-00203-7> (2023).
68. Rezaia, M., Faramarzi, A. & Javadi, A. A. An evolutionary based approach for assessment of earthquake-induced soil liquefaction and lateral displacement. *Eng. Appl. Artif. Intell.* **24**(1), 142–153 (2011).
69. Ahmad, M. *et al.* Predicting subgrade resistance value of hydrated lime-activated rice husk ash-treated expansive soil: A comparison between M5P, support vector machine, and Gaussian process regression algorithms. *Mathematics* **10**(19), 3432 (2022).
70. Taylor, K. E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **106**(D7), 7183–7192 (2001).

## Acknowledgements

This work was supported by Universiti Tenaga Nasional through BOLDREFRESH2025 fund under project code of J510050002-IC-6. Furthermore, the authors gratefully acknowledge Ms. Tehreem Javed for her assistance with preliminary documentation, initial model development, and validation. Her contribution during the early phase of this research is sincerely appreciated.

## Author contributions

Author Contribution: Mahmood Ahmad: Conceptualization, Formal Analysis, Investigation, Methodology, Writing—original draft, Writing—review and editing. Mohammad Al Zubi: Investigation, Validation, Resources, Writing—review and editing. Abdullah Alzlfawi: Investigation, Methodology, Resources, Supervision, Validation, Writing—review and editing. Sabahat Hussan: Formal Analysis, Investigation, Resources, Writing—review and editing. Shay Haq, Shaikat Biswas, Abdur Rahman Rakib, and Zia Ullah: Investigation, Formal analysis, Validation, Resources, Writing—review and editing. Rohayu Che Omar: Conceptualization, Investigation, Validation, Writing—review and editing. Zsolt Tóth: Supervision, Project administration, Formal Analysis, Investigation, Methodology, Writing—original draft, Writing—review and editing.

## Funding

Open access funding provided by University of Sopron. This research was funded by Universiti Tenaga Nasional, BOLDREFRESH2025 fund, Project Code J510050002-IC-6.

## Declarations

## Competing interests

The authors declare no competing interests.

## Informed Consent

Not applicable.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-026-50670-4>.

**Correspondence** and requests for materials should be addressed to Z.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026