RESEARCH ARTICLE

Ensemble unit and AI techniques for prediction of rock strain

Pradeep T[a], **Pijush SAMUI**[a], **Navid KARDANI**[b*], **Panagiotis G ASTERIS**[c]

[a] *Civil Engineering Department, National Institute of Technology, Patna 800005, India*

[b] *Civil and Infrastructure Discipline, School of Engineering, Royal Melbourne Institute of Technology (RMIT), Melbourne, Victoria, Australia*

[c] *Computational Mechanics Laboratory, School of Pedagogical and Technological Education, Heraklion, GR 14121, Greece*

[*]*Corresponding author. E-mail: navid.kardani@rmit.edu.au*

ABSTRACT The behavior of rock masses is influenced by a variety of forces, with measurement of stress and strain playing the most critical roles in assessing deformation. The laboratory test for determining strain at each location within rock samples is expensive and difficult but rock strain data are important for predicting failure of rock material. Many researchers employ AI technology in order to solve these difficulties. AI algorithms such as gradient boosting machine (GBM), support vector regression (SVR), random forest (RF), and group method of data handling (GMDH) are used to efficiently estimate the strain at every point within a rock sample. Additionally, the ensemble unit (EnU) may be utilized to evaluate rock strain. In this study, 3000 experimental data are used for the purpose of prediction. The obtained strain values are then evaluated using various statistical parameters and compared to each other using EnU. Ranking analysis, stress-strain curve, Young's modulus, Poisson's ratio, actual vs. predicted curve, error matrix and the Akaike's information criterion (AIC) values are used for comparing models. The GBM model achieved 98.16% and 99.98% prediction accuracy (in terms of values of R^2) in the longitudinal and lateral dimensions, respectively, during the testing phase. The GBM model, based on the experimental data, has the potential to be a new option for engineers to use when assessing rock strain.

KEYWORDS prediction, strain, ensemble unit, rank analysis, error matrix

1 Introduction

Rock deformation research is important for estimating crack pattern in rock masses. Due to the rising number of underground excavations under high mountains or at greater depths where the rock mass is subjected to severe pressures, recent study on the mechanical behavior of hard rocks has become more significant [1]. Stress-strain relationships are used to describe how rock deforms. Deformations occur as strains; the three types of deformation that can occur as a result of pressure in rock are elastic deformation, ductile strain, and fracture. Many researchers have studied the stress-strain relationship of rock materials through experiments. The effect of cyclic thermal shock on the physico-mechanical characteristics

of granite has been investigated by Yu et al. [2]. He evaluated the shape of the standard deviation curve of surface principal strain. Zhang et al. [3] used a combination of acoustic emission and ultrasonic transmission methods in a quasi-static uniaxial compression tests to investigate the damage characterizations of granites. Zhao et al. [4] evaluated the energy changes in the brittle granite under uniaxial compression at loading rates of 0.001, 0.005, 0.01, and 0.05 mm/s. Duan [5] investigated the unloading-induced collapse of brittle rock by conducting experimental and numerical research on granite. Soft computing approaches have recently been used to solve science and engineering challenges in a further variety of research projects [6,7, 8–18].

Direct examination is difficult to use in order to assess the strain in a rock mass [19,20]. To overcome these

concerns, many researchers have concentrated on numerical and AI-based approaches for determining rock properties [21–25]. Because it does not require any prior knowledge of the type of relationship that exists between the predictors and the forecasted variables, AI is a viable technique for establishing and simulating such a relationship [26]. A great benefit of AI over traditional empirical and statistical methods is that it can make predictions that were previously impossible [27]. Random forests (RF), also known as neural nets, provide estimates for variable relevance. They also provide way of dealing with data that is missing. Missing values are filled in by the value that appears the most in a specific node. RF outperforms all other classification methods in terms of accuracy. The RF algorithm can also handle large datasets with thousands of variables. When a class is more infrequent than other classes in the data, it can automatically balance data sets. The approach also works quickly with variables, making it suited for more complex tasks [28–30]. Gradient Boosting Machines (GBM) create a series of shallow and weak trees that learn from and improve on each other; it frequently gives exceptional forecasting accuracy, lots of flexibility, without need to pre-process the data, and handles missing data [31–34]. Support vector regression (SVR) has been shown to be a useful method for estimating real-value functions. It is robust to outliers. It is simple to update the decision model. It has high prediction accuracy and great generalization capabilities [35,36]. Group method of data handling (GMDH) can identify the influential variables and generate an explicit model formulation. It estimates the quantity of network layers and neurons in each layer automatically, successfully reducing artificiality in the simulation process. In contrast to other neural networks using black box models, polynomials are used to relate the selected parameters to the output [37,38].

The aim of this paper is to develop and apply soft computing approaches for estimating strain in rock. This was performed using a large experimental dataset generated from uniaxial compression testing on rocks. A well-equipped test setup was used to measure the rock sample's deformation at different locations. The cylindrical rock material was gradually loaded along its longitudinal axis. At the same time, the longitudinal and lateral deformations of the rock material were recorded. The strain gauge-based transducer was fitted on the perimeter of the cylindrical rock material at a different angle and height. After the experiment on the rock sample, the data from the acquisition system was obtained. The proposed framework requires input and output in order to predict model data. The stress and position of stain gauge (height and angle) in the rock material were employed as input parameters. The longitudinal and lateral strain of the rock sample was employed as output parameters. In this work,

predicting rock strain is performed using developed algorithms such as GBM, RF, SVR, and GMDH. In order to compare the developed model, the stress-strain curve, elastic modulus, and Poisons ratio were used to identify a robust model.

2 Details of data

During experimental testing, data was collected from the laboratory to forecast strain in a granite rock material. The work by Isah et al. [39] contained descriptions of strain measurements for uniaxial compression testing on rocks. Uniaxial compression tests were used to assess the load and deformation of a cylindrical granite rock (diameter = 40 mm and height = 108 mm). A load cell and the number of strain gauge-based transducers were used to measure the load on the longitudinal axis and deformation on the perimeter of the cylindrical rock. The details of the location of the strain gauge are shown in Fig. 1. In total, 48 electronic strain gauges were used on the perimeter (lateral and longitudinal direction). These deformation and load data were accumulated in data acquisition system. According to the dimension of the rock, the obtained data was used to determine stress and strain. The uniaxial strength of the granite rock was 336.110 MPa, with maximum strains of 2.9947×10^{-3} and -0.0037×10^{-3} in the lateral and longitudinal direction, respectively. This granite rock material has a Poisson's ratio of 0.27. The strain gauge's height, angle, and stress in the rock sample are considered input parameters, whereas the strains in the lateral and longitudinal direction of the rock material are considered output parameters. Using these input and output data, soft computing frameworks were employed to forecast the strain in a rock sample. Finally, 3000 measurements from
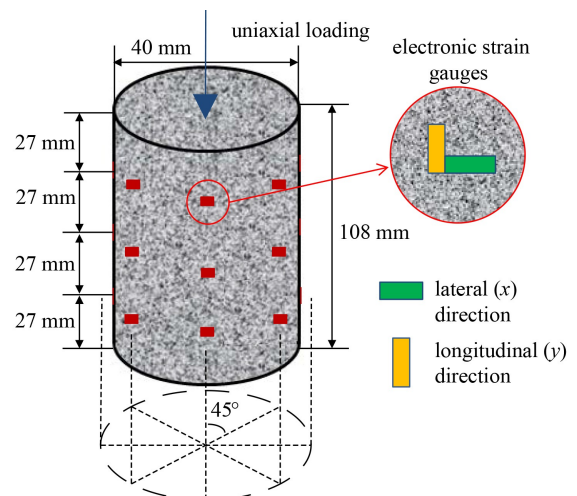

Fig. 1 Details of the strain gauge when a rock sample is subjected to uniaxial load.

the granite rock sample were gathered for the sake of this study [8,23].

These values were standardized between 0 and 1 in order to use a common scale for numeric column values in the dataset, using Eq. (1). 70% of the main data (2100 measurements) were used to train the models from the training data (2100). Similarly, the trained models were tested using 30% of the main data (900 measurements).

$$Normalized\ value = \left(\frac{Observed\ value - Minimum\ value}{Maximum\ value - Minimum\ value} \right). \tag{1}$$

3 Soft computing techniques: design and details

3.1 Gradient boosting machine

One of the most well-known supervised machine learning algorithms introduced by Friedman [40] is the GBM. The standard regression problem can be expressed as follows:

$$D = \{(x_1, y_1), \ldots, (x_N, y_N)\}, \tag{2}$$

where x_i is a member of the set $\chi \subset \mathbb{R}^m$ and represents a feature vector with m elements, and $y_i \in \mathbb{R}^m$ represents the actual outcome value such that $y_i = f(x_i) + \varepsilon$. Here ε represents random noise and has an expectation of 0 and an unknown finite variance; f is an unknown continuous function that must be computed. The aim of machine learning is to create a regression model or an approximation g of a function f that minimizes the predictable risk or loss function.

$$L(g) = \mathbb{E}_{(x,y) \sim p} L(y, g(x)) = \int_{\chi \times \mathbb{R}} L(y, g(x)) \, dP(x, y), \tag{3}$$

where a joint probability distribution of x and y is denoted by $P(x, y)$, $L(\cdot, \cdot)$ and this is the loss function. GBMs build an additive ensemble model of size M by iteratively improving the predictions of y from x with respect to L by adding new weak or base learners that improve on the prior ones.

$$g_0(x) = c, \tag{4}$$

$$g_i(x) = g_{i-1}(x) + \gamma_i h_i(x); \ i = 1, \ldots, M, \tag{5}$$

where i is the number of iteration level; the ith base model is h; γ_i is the ith base model weight or coefficient.

The aim of the algorithm is to reduce the loss function L as much as possible. If decision trees are considered to be the base models, each iteration results in the construction of a single decision tree to fit the negative gradients. The parameters θ_i can be used to define the function h, i.e., $h_i(x) = h(x, \theta_i)$. The depths of trees, the

learning rate, and the number of iterations are all GBM parameters.

3.2 Random forest algorithm

The aim of RF is to use a number of weak models, such as decision trees, to make a robust model. It is an ensemble learning technique for regression, classification and other tasks that works by creating a large number of decision trees during training [41,42]. The mean or average prediction of the individual trees is returned for regression tasks. The RF algorithm works primarily as follows.

1) The bagging sampling method is used to generate K training sets from the initial training set M, with each training set having a total of N samples.

2) To produce K CART decision tree models, the model is trained using K training sets.

3) Each time, the ideal division attributes of the current node is selected according to the GINI index to develop branch nodes, and finally a single decision tree is generated for the characteristic attributes of a single decision tree model.

4) The created K decision trees are developed into a random forest.

Improving the combined classification model's extrapolation prediction performance is critical to the final result. The final classification choice is made after k rounds of training, $\{h_1(x), h_2(x), \ldots, h_k(x)\}$, and a simple majority vote:

$$H(x) = \arg \max_{|Y} \sum_{i=1}^{k} I(h_i(x) = Y), \tag{6}$$

where $H(x)$ denotes the combined classification model, h_i denote a single decision tree result, Y denote output of model, $I(\cdot)$ denotes an indicative function. The following is the formula for determining the importance of variables:

$$V = \frac{A - A_1}{OOB}, \tag{7}$$

where A denotes the number of samples properly categorized prior to re-placement. A_1 denotes the number of samples that were correctly classified after being replaced, OOB is the observed score of RF.

3.3 Support vector regression

The support vector machine (SVM) is a set of learning methods for handling real issues with a small sample, non-linearity, and high dimensionality techniques that were developed in the late twentieth century [36,43]. Consider the following sample set, $D = (x_i, y_i), i = 1, 2, \ldots, n$; x_i is the input vector; y_i is the goal output; and n is the

sample set's number of samples. The issue is an autoregressive problem, with x_i denoting daily influence variables. Many non-linear problems in real life are solved by mapping the sample point into a high-dimensional space ($\phi \to \phi(x)$) using a mapping function ϕ. Only the dot product of the feature space, $\phi(x_i) \cdot \phi(x_j)$, is employed to generate the ideal hyper-plane. Consequently, if a function $K(\cdot)$ satisfies $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$, it is considered a kernel function. The Gaussian function, given by Eq. (8), is the most commonly utilized type of kernel function.

$$K(x_i, x_j) = \exp\left(-0.5x_i - \frac{x_j^2}{\sigma^2}\right). \tag{8}$$

The SVM was developed to resolve classification problems, but it's also useful for regression analysis, which comes in two types: linear and non-linear. The linear regression function is defined as follows:

$$y = f(x) = \omega \cdot x + b. \tag{9}$$

The original data can be written as $\{(x_i, y_i)\}_{i=1}^{l}$; the regression function in Eq. (9) must be as small as possible to that it is flat. Consequently, its Euclidean space norm should be minimized. In Eq. (9), ω and b are the regression function's normal vectors and offsets, respectively. A linear function with an error of ϵ is considered to fit all of the training data. The following optimization problem can then be solved.

$$\min \Phi(\omega) = \frac{1}{2}\omega^2, \tag{10}$$

with the limitations,

$$\begin{cases} y_i - \omega x_i - b \leqslant \epsilon, \\ \omega x_i + b - y_i \leqslant \epsilon, \end{cases} i = 1, 2, \ldots, n. \tag{11}$$

When the above limitations cannot be totally satisfied, the relaxation variables ξ_i and ξ_i^* are introduced, and the optimization problem is converted into the following issue,

$$\min \Phi(\omega) = \frac{1}{2}\omega^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*), \tag{12}$$

with the limitations,

$$\begin{cases} y_i - \omega x_i - b \leqslant \epsilon + \xi_i, \\ \omega x_i + b - y_i \leqslant \epsilon + \xi_i^*, \end{cases} \xi_i, \xi_i^* \geqslant 0; i = 1, 2, \ldots, IC > 0. \tag{13}$$

The constraints are linear since function $\Phi(\omega)$ is quadratic, hence the problem is a standard quadratic programming problem that can be resolved using Lagrange multipliers.

3.4 Group method of data handling

GMDH is a collection of methods for estimating system behavior and solving various challenges. The GMDH algorithm is used to develop an adaptable, supervised learning network that has been used in a wide range of applications, including medical, automobile systems, and so on. It is predicted that this method can find the best answer by minimizing the value of an external criterion [44]. The GMDH algorithms allow for the automatic selection of mathematical expressions such as polynomials, non-linear models, and probabilistic functions in order to discover the best modelling structure or network depending on the input data.

This method consists of a primary layer of neurons connected by a polynomial that generates new neurons in succeeding layers. The following procedure was employed for a given set of 'n' observations of the m independent variables $(x_1, x_2, \ldots, x_m)$ and their corresponding matrix of dependent values $(y_1, y_2, \ldots, y_n)$. The total number of combinations for a pair of variables is $m \times (m-1)/2$, because there are m input variables. Equation (14) can be used to assess the output y using a polynomial expression.

$$y = a + bx_i + cx_j + dx_i^2 + ex_j^2 + fx_ix_j. \tag{14}$$

Equation (14) can then be used to evaluate the polynomial for all n observations, resulting in matrix Z. The Z matrix can be thought of as a set of new, improved variables that are more predictable than the original generation $x_1, x_2, \ldots, x_m$ variables.

$$z_{ij} = a + bx_i + cx_j + dx_i^2 + ex_j^2 + fx_ix_j. \tag{15}$$

Therefore, the algorithm computes the *RMS* value in order to exclude the variables that are the least effective. The test data set for each column of the Z matrix ($j = 1$ to $m \times (m-1)/2$) can be used to assess this. The regularity criterion is given by Eq. (16).

$$r_j^2 = \frac{\left(\sum_{j=1}^{n}(y_i - z_{ij})^2\right)}{y_j^2}. \tag{16}$$

The columns of Z should be sorted in increasing order of r_j, with the original columns of the input matrix X being replaced by the columns of Z satisfying $r_j < R$ (R stands for a user-specified value). The process is repeated until the method termination criterion is met, and additional generations are obtained.

3.5 Ensemble unit

Combining multiple model outputs has shown that ensembling techniques as post-process procedures can improve

model prediction [45]. When numerous approaches are utilized, the results do not appear to be extremely sensitive to the particular choice of methods, which is an advantage of combining predictions. Using a combination of predictors is thus safer and less risky than depending on a single technique. Experimental and theoretical studies suggest that combining the outputs of several models can be a great way to improve the overall efficiency of time series prediction. The linear weighted averaging method was evaluated in this paper for combining the outputs of the employed models to improve prediction performance. In this method, four techniques were applied for combining the outputs of the GBM, SVR, RF, and GMDH models. The weighted averaging model can be written in Eq. (17).

$$f(x) = \sum_{i=1}^{n} w_i f(x), \tag{17}$$

where w_i is the applied weight on the ith model which can be determined based on the model performance as:

$$w_i = \frac{R_i^2}{\sum_{i=1}^{n} R_i^2}, \tag{18}$$

where R_i^2 is the coefficient of determination on the ith single model.

3.6   Evaluation of models

The model accuracy is analyzed by using some statistical parameter including determination coefficient (R^2), Weighted Mean Absolute Percentage Error ($WMAPE$), Root Mean Square Error ($RMSE$), Variance Account Factor (VAF), Performance Index (PI) [46], Root mean square error to observation's standard deviation ratio (RSR) [47], Willmott's Index of agreement (WI) [48], Mean absolute error (MAE) [49], Mean absolute percentage error ($MAPE$), and Mean Bias Error (MBE).

$$R^2 = \frac{\sum_{i=1}^{N} (A_i - P_{mean})^2 - \sum_{i=1}^{N} (A_i - P_i)^2}{\sum_{i=1}^{N} (A_i - A_{mean})^2}, \tag{19}$$

$$WMAPE = \frac{\sum_{i=1}^{n} \left| \frac{A_i - P_i}{A_i} \right| \times A_i}{\sum_{i=1}^{n} A_i}, \tag{20}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (A_i - P_i)^2}, \tag{21}$$

$$VAF = \left(1 - \frac{var(A_i - P_i)}{var(A_i)} \right) \times 100, \tag{22}$$

$$PI = adj.R^2 + (0.01 \times VAF) - RMSE, \tag{23}$$

$$RSR = \frac{RMSE}{\sqrt{\frac{1}{N} \sum_{i=1}^{n} (A_i - A_{mean})^2}}, \tag{24}$$

$$WI = 1 - \left[\frac{\sum_{i=1}^{N} (A_i - P_i)^2}{\sum_{i=1}^{N} \{|P_i - A_{mean}| + |A_i - A_{mean}|\}^2} \right], \tag{25}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |(P_i - A_i)|, \tag{26}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{A_i - P_i}{A_i} \right| \times 100, \tag{27}$$

$$MBE = \frac{1}{N} \sum_{i=1}^{N} (P_i - A_i), \tag{28}$$

where A_i is the observed ith value; P_i is the predicted ith value; A_{mean} is the average of actual value; N is the number of samples.

4   Result and discussion

In this section, the models' performances are compared. The models were developed with the MATLAB 2015a version of software. Tuning factors such as the number of trees, learning rate, Sigma, and others were important during the model training process. The tuning parameters of all models are displayed in Table 1.

4.1   Stress−strain curve

Rock is primarily a brittle substance with high compression strength but low tensile strength. Tensile deformation in the lateral (x) dimension occurred as a reaction to the tensile force in the rock material. The maximum strain indicates the start of a failure form in the rock. In this

study, data from the rock was used to examine this maximum strain. We investigated the strain values and found that the highest values were achieved at sample heights of 27 and 54 mm and angles of 270° and 90° in the lateral and longitudinal dimensions, respectively. Figures 2, 3, 4, and 5 show the strain in the longitudinal and lateral direction of the rock, related to the height and angle of the strain gauge. Different strain gauges mounted in lateral and longitudinal dimensions were used to evaluate the behavior of each rock sample. Figures 6 and 7 show the stress-strain curves obtained from models such as GBM, SVR, RF, and GMDH. These stress−strain curves were used for comparison of the model outputs with actual curve. Some models (SVR, RF, and GMDH) failed to perform effectively, despite the fact that the actual curve began at the origin. However, in both

Table 1   Tuning parameters of the developed models

model	tuning parameters
GBM	i. No. of trees (train) = 100,200,…,1000 ii. Max. tree leaves = 3,5,7,9,11,13,15 iii. Max. tree depth = 4,5,6 iv. learning rate = 0.001,0.005,0.01,0.05,0.1 v. Min. No. of data in a leaf = 5,10,15,20
SVR	i. Penalty of the error = 1000 (max.) ii. Sigma = 1 (max.)
RF	i. No. of trees = 1000 (max.) ii. No. of input (best split) = 3 iii. Min. No. of samples (internal node) = 20 iv. Min. No. of samples (leaf node) = 15
GMDH	i. N-layer = 4 ii. α = 0.6 iii. Max-Neurons = 20
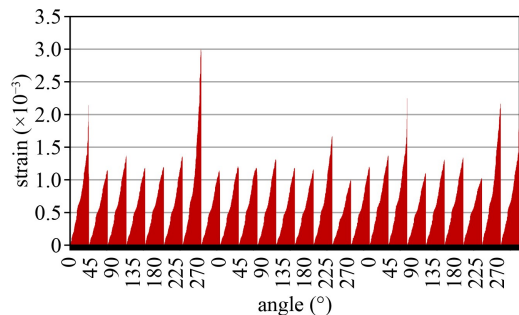
Fig. 4   Strain−height variation in the *y* direction.

Fig. 5   Strain−angle variation in the *y* direction.
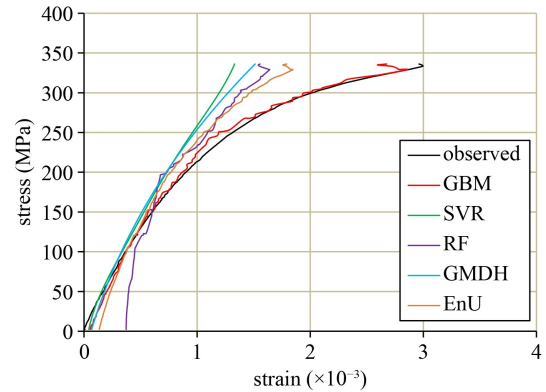
Fig. 6   Rock sample stress−strain behavior (*x* direction).

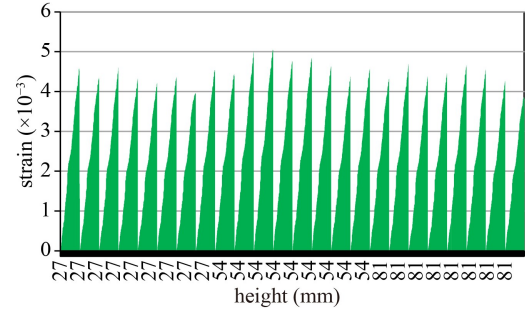**Fig. 2**   Strain−height variation in the *x* direction.

Fig. 3   Strain−angle variation in the *x* direction.
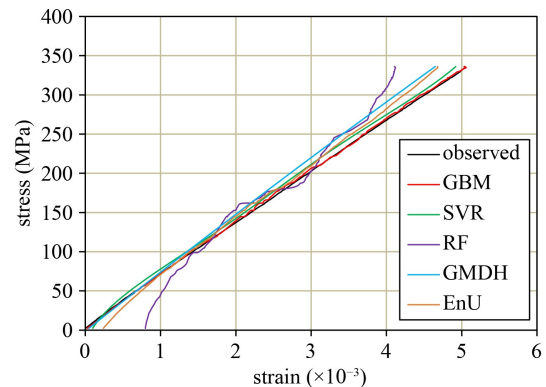

Fig. 7   Rock sample stress−strain behavior (*y* direction).

longitudinal and lateral dimensions, it is clear that GBM is perfectly fitted to the actual curve.

4.2 Young's modulus (E)

The ratio of the applied stresses in longitudinal axis to the corresponding strain in the same axis under tension or compression of rock sample is known as Young's modulus (E). Values of E were used to compare actual and model results in this section. Saturated cores rock has an E value of up to 90 GPa [50]. The actual value is 83.39 GPa in this case. When compared to other models, the GBM ($E = 83.11$ GPa) proved to be the most accurate model. These results are visualized in form of bubble chart (Fig. 8). EnU and RF, only, have values less than 75 GPa.

4.3 Poisson's ratio (μ)

The ratio of lateral strain to longitudinal strain in rock samples under compression or tension stress is known as Poisson's ratio (μ). Here, actual and predicted μ values were determined by using lateral and longitudinal strain values. With a mean value of 0.28, the values ranged from 0.19 to 0.35 [51]. All values were within the range and nearly equal to mean values. Therefore, from this point of view, all models were considered to be equally valid. The values are shown in the form of radar diagram (Fig. 9).

4.4 Statistical parameters with rank

Tables 2 and 3 demonstrate the statistical assessment and score of the developed model for lateral and longitudinal strains, respectively. The values of performance parameters are displayed on the basis of normalized outputs. To consider better model efficiency, *WMAPE*, *MAE*, *RMSE*, *MAPE*, *MBE*, and *RSR* should have values equal to 0, *VAF* should be close to 100, PI should be close to 2 and R^2 should be close to 1. *WI* is a scale that is in the range 0 to 1 and represents the of error level in model predictions. All models achieved good values compared to the limits and range of parameters. In this study, the performances of the predictive models and EnU was compared using a most likely rank system [52]. The Rank was calculated using the training and testing parameter values from each of the four models and EnU. The range of Rank from 1 to 5 was defined by the number of models. A Rank system's causality value is the ideal value, whereas the comparative best model has a first rank. The most likely rank for training and testing data is used to calculate the model's overall performance. As a result of this, the GBM (1) achieved first rank in both lateral and longitudinal dimension. SVR (3 and 2), EnU (2 and 3), GMDH (4), and RF (5) appeared in next place,

respectively, in both lateral and longitudinal dimensions. Therefore, GBM is regarded as the most accurate performer in lateral and longitudinal dimensions.

4.5 Actual vs. predicted

Figure 10 illustrates the strain predicted by various proposed models compared to actual value for the training and testing datasets in lateral and longitudinal dimension. As the points approached the regression line, the developed model's performance improved. The GBM model outperformed other models, as indicated by the R^2 values. In both the dimensions, the GBM equation was very close to ($x = y$) in training and testing. The SVR and EnU also outperformed the RF and GMDH. As a result, GBM is thought to be the most effective model.

4.6 Error matrix

Figures 11 and 12 in this section illustrate the degree of inaccuracy associated with the models depending on a variety of statistical parameters. This is a heat map matrix created from comparing ideal values of the statistical parameter. More details of this matrix are discussed in
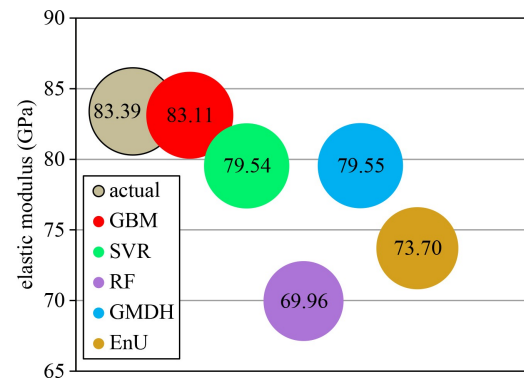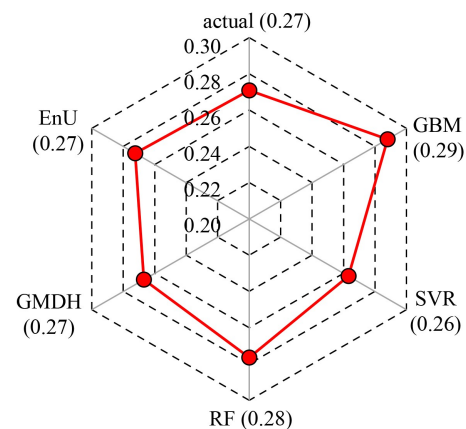
Fig. 8 Comparison of predicted elastic modulus.

Fig. 9 Predicted Poisson's ratio.

Table 2 Statistical parameter in lateral dimension

parameter		model									
		GBM		SVR		RF		GMDH		EnU	
		train	test	train	test	train	test	train	test	train	test
R^2	value	0.9931	0.9816	0.8915	0.9052	0.9391	0.9347	0.8784	0.8951	0.9503	0.9484
	rank	1	1	4	4	3	3	5	5	2	2
WMAPE	value	0.0391	0.0483	0.1065	0.1004	0.1734	0.1672	0.1316	0.1226	0.0928	0.0889
	rank	1	1	3	3	5	5	4	4	2	2
RMSE	value	0.0125	0.0194	0.0501	0.0443	0.0544	0.0522	0.0521	0.0464	0.0371	0.0348
	rank	1	1	3	3	5	5	4	4	2	2
VAF	value	99.3044	98.1576	88.8898	90.4512	86.4918	86.6854	87.8359	89.4863	93.8477	94.0761
	rank	1	1	3	3	5	5	4	4	2	2
PI	value	1.9737	1.9437	1.7301	1.7651	1.7496	1.7492	1.7045	1.7433	1.8516	1.8542
	rank	1	1	4	3	3	4	5	5	2	2
RSR	value	0.0831	0.1334	0.3348	0.3046	0.3670	0.3585	0.3476	0.3187	0.2474	0.2391
	rank	1	1	3	3	5	5	4	4	2	2
MAPE	value	8.6163	9.4273	15.0512	17.0682	53.0258	50.0705	22.5055	21.8960	29.5867	31.2156
	rank	1	1	2	2	5	5	3	3	4	4
WI	value	0.9982	0.9955	0.9678	0.9743	0.9526	0.9558	0.9666	0.9730	0.9822	0.9838
	rank	1	1	3	3	5	5	4	4	2	2
MAE	value	0.0083	0.0101	0.0226	0.0210	0.0359	0.0349	0.0280	0.0256	0.0197	0.0186
	rank	1	1	3	3	5	5	4	4	2	2
MBE	value	2.8E–05	3.5E–04	–6.2E–03	–3.6E–03	–1.8E–03	1.1E–03	–5.1E–04	1.5E–03	–1.6E–03	–1.5E–04
	rank	1	2	5	5	4	3	2	4	3	1
most likely rank		1		3		5		4		2	

Table 3 Statistical parameter in longitudinal dimension

| parameter		model									
		GBM		SVR		RF		GMDH		EnU	
		train	test	train	test	train	test	train	test	train	test
R^2	value	0.9998	0.9884	0.9965	0.9860	0.9804	0.9693	0.9920	0.9819	0.9974	0.9869
	rank	1	1	3	3	5	5	4	4	2	2
WMAPE	value	0.0067	0.0114	0.0321	0.0345	0.1298	0.1332	0.0390	0.0406	0.0368	0.0397
	rank	1	1	2	3	5	2	4	5	3	4
RMSE	value	0.0038	0.0287	0.0167	0.0319	0.0755	0.0778	0.0241	0.0359	0.0218	0.0335
	rank	1	1	2	2	5	5	4	4	3	3
VAF	value	99.9802	98.8321	99.6486	98.5662	92.1973	91.3857	99.2047	98.1734	99.3504	98.4038
	rank	1	1	2	2	5	5	4	4	3	3
PI	value	1.9958	1.9480	1.9763	1.9397	1.8268	1.8052	1.9600	1.9277	1.9691	1.9373
	rank	1	1	2	2	5	5	4	4	3	3
RSR	value	0.0139	0.1081	0.0614	0.1201	0.2769	0.2933	0.0884	0.1352	0.0801	0.1263
	rank	1	1	2	2	5	5	4	4	3	3
MAPE	value	4.6528	6.8070	29.7569	46.4693	63.0917	71.4464	13.4474	19.0890	32.8977	35.2632
	rank	1	1	3	4	5	5	2	2	4	3
WI	value	1.0000	0.9971	0.9991	0.9964	0.9749	0.9717	0.9980	0.9954	0.9983	0.9958
	rank	1	1	2	2	5	5	4	4	3	3
MAE	value	0.0030	0.0051	0.0146	0.0155	0.0590	0.0599	0.0177	0.0183	0.0167	0.0178
	rank	1	1	2	2	5	5	4	4	3	3
MBE	value	2E–05	1E–03	–5E–03	–3E–03	–6E–04	2E–04	–7E–05	1E–03	–1E–03	–5E–06
	rank	1	3	5	5	3	2	2	4	4	1
most likely rank		1		2		5		4		3	
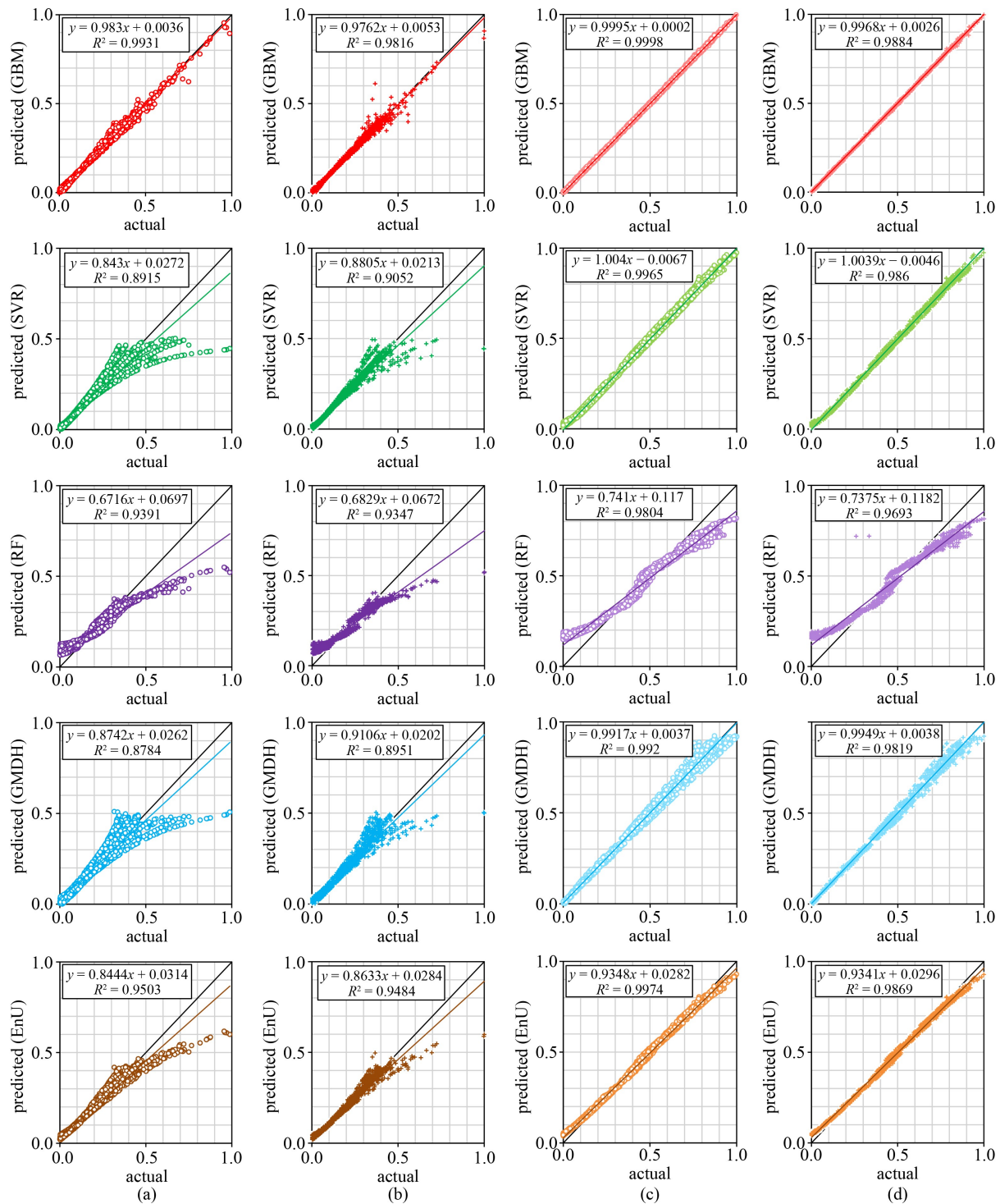
Fig. 10 Actual vs. predicted. (a) Training in lateral dimension; (b) testing in lateral dimension; (c) training in longitudinal dimension; (d) testing in longitudinal dimension.

[25]. This matrix clearly indicated that the GBM had a lower error rate in both training and testing of both dimensions. When compared to other models, the reddish tint indicates a higher level of RF inaccuracy. As a consequence, the GBM model beats the others, while the RF model performed the worst in both training and testing of both dimensions.

4.7 Akaike's information criterion

One of the most sensitive challenges in the construction of data models is evaluating the model's generalization potential, which is defined as the model's greatest performance in the study dataset that is being used. The Akaike's information criterion (AIC) criterion was
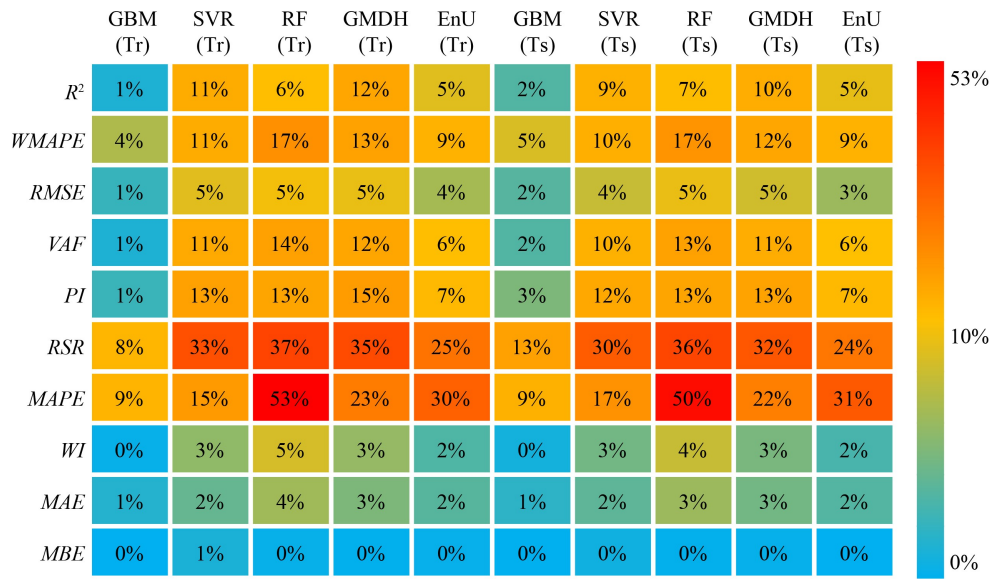
	GBM (Tr)	SVR (Tr)	RF (Tr)	GMDH (Tr)	EnU (Tr)	GBM (Ts)	SVR (Ts)	RF (Ts)	GMDH (Ts)	EnU (Ts)
R^2	1%	11%	6%	12%	5%	2%	9%	7%	10%	5%
WMAPE	4%	11%	17%	13%	9%	5%	10%	17%	12%	9%
RMSE	1%	5%	5%	5%	4%	2%	4%	5%	5%	3%
VAF	1%	11%	14%	12%	6%	2%	10%	13%	11%	6%
PI	1%	13%	13%	15%	7%	3%	12%	13%	13%	7%
RSR	8%	33%	37%	35%	25%	13%	30%	36%	32%	24%
MAPE	9%	15%	53%	23%	30%	9%	17%	50%	22%	31%
WI	0%	3%	5%	3%	2%	0%	3%	4%	3%	2%
MAE	1%	2%	4%	3%	2%	1%	2%	3%	3%	2%
MBE	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%

Fig. 11 Error matrix for lateral dimension.

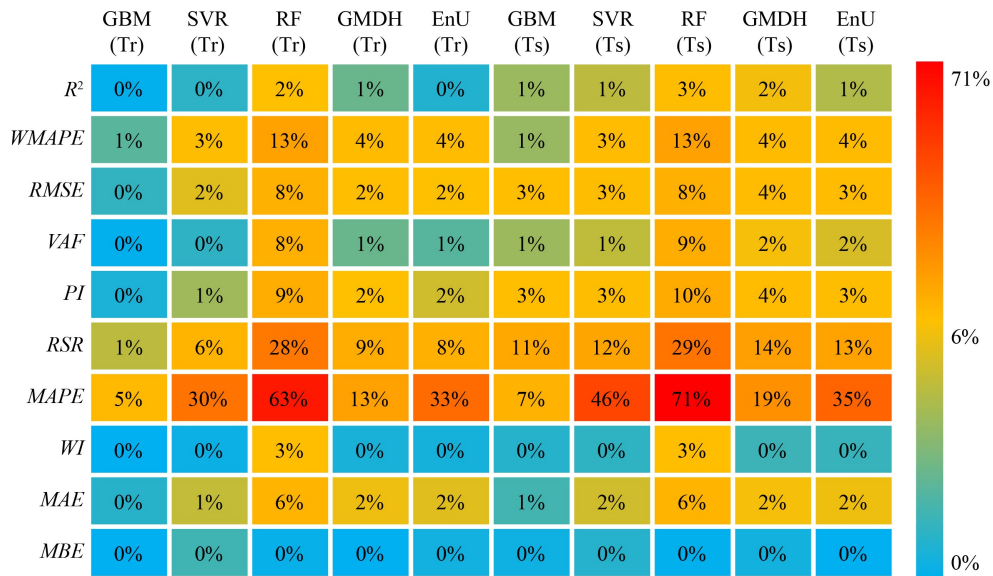|	GBM (Tr)	SVR (Tr)	RF (Tr)	GMDH (Tr)	EnU (Tr)	GBM (Ts)	SVR (Ts)	RF (Ts)	GMDH (Ts)	EnU (Ts)
R^2	0%	0%	2%	1%	0%	1%	1%	3%	2%	1%
WMAPE	1%	3%	13%	4%	4%	1%	3%	13%	4%	4%
RMSE	0%	2%	8%	2%	2%	3%	3%	8%	4%	3%
VAF	0%	0%	8%	1%	1%	1%	1%	9%	2%	2%
PI	0%	1%	9%	2%	2%	3%	3%	10%	4%	3%
RSR	1%	6%	28%	9%	8%	11%	12%	29%	14%	13%
MAPE	5%	30%	63%	13%	33%	7%	46%	71%	19%	35%
WI	0%	0%	3%	0%	0%	0%	0%	3%	0%	0%
MAE	0%	1%	6%	2%	2%	1%	2%	6%	2%	2%
MBE	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Fig. 12 Error matrix for longitudinal dimension.

developed by Akaike to assess the generalized capacity of models. This parameter has been used to assess the power of models in engineering-related challenges [53,54]. The AIC criteria are as follows:

$$AIC = N \times \ln\left((RMSE)^2\right) + 2K, \qquad (29)$$

The number of datasets in testing or training is N, while the number of fitting parameters is K. Table 4 and Fig. 13 show the AIC values for GBM, SVR, RF, GMDH, and EnU for the testing and training datasets. This value of the best model is the lowest. As a result, GBM has the lowest generalization potential when compared to other models based on these parameters.

5 Conclusions

In this paper, data comprising 3000 measurements were used for predicting the rock strain in the lateral and longitudinal dimensions. To forecast strain, four models (GBM, SVR, RF, and GMDH) and EnU were used. Seventy and thirty percent of the main dataset from the whole dataset were used for model training and testing, respectively. The predicted results were examined using the stress-strain curve, Young's modulus (E), Poisson's ratio (μ), rank analysis, the actual vs. predicted curve, error matrix and the AIC values. In each analysis, the GBM (R^2 = 0.9931; 0.9816 and 0.9998; 0.9884) was shown to be more accurate than the SVR, RF, and GMDH models, as well as EnU, in both training and

Table 4   AIC values of models

models	train (x)	test (x)	train (y)	test (y)
GBM	−18415.1	−7089.486	−23364.93	−6386.733
SVR	−12564.4	−5603.029	−17173.15	−6196.313
RF	−12159.1	−5309.413	−10848.1	−4589.758
GMDH	−12406	−5521.362	−15642.85	−5984.094
EnU	−13833.6	−6038.649	−16058.97	−6107.009
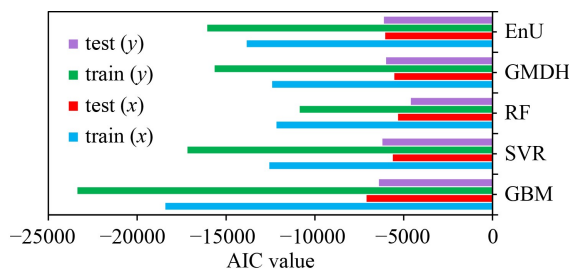
Fig. 13   AIC values for all models in x and y dimensions.

testing of lateral and longitudinal dimension, according to the experimental data. EnU (R^2 = 0.9503; 0.9484 and 0.9974; 0.9869) also performed well, but it is in next level of GBM. The R^2 values are given here in the sequence of training and testing phases for the lateral and longitudinal dimension, respectively. Overall, based on existing experimental datasets, the proposed GBM model can be used as a promising model to predict rock strain. The study's future path could involve a full evaluation of the proposed GBM and hybrid models of additional optimization algorithm and neural networks, as well as Deep Learning. EnU also performed well but it is dependent on the results of other models. In the future, further improvement must be made in forecasting the strain in the rock sample using RF and GMDH models.

References

1. Bruning T, Karakus M, Nguyen G D, Goodchild D. An experimental and theoretical stress−strain−damage correlation procedure for constitutive modelling of granite. International Journal of Rock Mechanics and Mining Sciences, 2019, 116: 1−12

2. Yu P, Pan P Z, Feng G, Wu Z, Zhao S. Physico-mechanical properties of granite after cyclic thermal shock. Journal of Rock Mechanics and Geotechnical Engineering, 2020, 12(4): 693−706

3. Zhang G, Gao F, Wang Z, Yue S, Deng S. Quantifying the progressive fracture damage of granite rocks by stress–strain, acoustic emission, and active ultrasonic methods. Journal of Materials in Civil Engineering, 2021, 33(12): 04021353

4. Zhao K, Yu X, Zhou Y, Wang Q, Wang J, Hao J. Energy evolution

of brittle granite under different loading rates. International Journal of Rock Mechanics and Mining Sciences, 2020, 132: 104392

5. Duan K, Ji Y, Wu W, Kwok C Y. Unloading-induced failure of brittle rock and implications for excavation-induced strain burst. Tunnelling and Underground Space Technology, 2019, 84: 495−506

6. Goswami S, Anitescu C, Chakraborty S, Rabczuk T. Transfer learning enhanced physics informed neural network for phase-field modeling of fracture. Theoretical and Applied Fracture Mechanics, 2020, 106: 102447

7. Samaniego E, Anitescu C, Goswami S, Nguyen-Thanh V M, Guo H, Hamdia K, Zhuang X, Rabczuk T. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. Computer Methods in Applied Mechanics and Engineering, 2020, 362: 112790

8. Nguyen-Thanh V M, Anitescu C, Alajlan N, Rabczuk T, Zhuang X. Parametric deep energy approach for elasticity accounting for strain gradient effects. Computer Methods in Applied Mechanics and Engineering, 2021, 386: 114096

9. Zhuang X, Guo H, Alajlan N, Zhu H, Rabczuk T. Deep autoencoder based energy method for the bending, vibration, and buckling analysis of Kirchhoff plates with transfer learning. European Journal of Mechanics. A, Solids, 2021, 87: 104225

10. Anitescu C, Atroshchenko E, Alajlan N, Rabczuk T. Artificial neural network methods for the solution of second order boundary value problems. Computers, Materials & Continua, 2019, 59(1): 345−359

11. Guo H, Zhuang X, Rabczuk T. A deep collocation method for the bending analysis of Kirchhoff plate. Computers, Materials & Continua, 2019, 59(2): 433−456

12. Bardhan A, Kardani N, GuhaRay A, Burman A, Samui P, Zhang Y. Hybrid ensemble soft computing approach for predicting penetration rate of tunnel boring machine in a rock environment. Journal of Rock Mechanics and Geotechnical Engineering, 2021, 13(6): 1398−1412

13. Asteris P G, Mamou A, Hajihassani M, Hasanipanah M, Koopialipoor M, Le T T, Kardani N, Armaghani D J. Soft computing based closed form equations correlating L and N-type Schmidt hammer rebound numbers of rocks. Transportation Geotechnics, 2021, 29: 100588

14. Kardani N, Bardhan A, Roy B, Samui P, Nazem M, Armaghani D J, Zhou A. A novel improved Harris Hawks optimization algorithm coupled with ELM for predicting permeability of tight carbonates. Engineering with Computers, 2021

15. Kardani N, Zhou A, Shen S L, Nazem M. Estimating unconfined compressive strength of unsaturated cemented soils using alternative evolutionary approaches. Transportation Geotechnics, 2021, 29: 100591

16. Kardani N, Bardhan A, Samui P, Nazem M, Asteris P G, Zhou A. Predicting the thermal conductivity of soils using integrated approach of ANN and PSO with adaptive and time-varying acceleration coefficients. International Journal of Thermal Sciences, 2022, 173: 107427

17. Kardani N, Bardhan A, Gupta S, Samui P, Nazem M, Zhang Y, Zhou A. Predicting permeability of tight carbonates using a hybrid

machine learning approach of modified equilibrium optimizer and extreme learning machine. Acta Geotechnica, 2022, 17: 1239−1255

18. Kaloop M R, Bardhan A, Kardani N, Samui P, Hu J W, Ramzy A. Novel application of adaptive swarm intelligence techniques coupled with adaptive network-based fuzzy inference system in predicting photovoltaic power. Renewable & Sustainable Energy Reviews, 2021, 148: 111315

19. Laghaei M, Baghbanan A, Hashemolhosseini H, Dehghanipoodeh M. Numerical determination of deformability and strength of 3D fractured rock mass. International Journal of Rock Mechanics and Mining Sciences, 2018, 110: 246–256

20. Gundewar C S. Government of India Ministry of Mines INDIAN BUREAU OF MINES Controller General Indian Bureau of Mines Application of Rock Mechanics in Surface and Underground Mining. New Delhi: Indian Bureau of Mines, 2014

21. Zhang W G, Li H R, Wu C Z, Li Y Q, Liu Z Q, Liu H L. Soft computing approach for prediction of surface settlement induced by earth pressure balance shield tunneling. Underground Space, 2021, 6(4): 353−363

22. Zhou J, Qiu Y, Armaghani D J, Zhang W, Li C, Zhu S, Tarinejad R. Predicting TBM penetration rate in hard rock condition: A comparative study among six XGB-based metaheuristic techniques. Geoscience Frontiers, 2021, 12(3): 101091

23. Pradeep T, Bardhan A, Samui P. Prediction of rock strain using soft computing framework. Innovative Infrastructure Solutions, 2021, 7(1): 37

24. Kardani N, Samui P, Kim D, Zhou A. Smart phase behavior modeling of asphaltene precipitation using advanced computational frameworks: ENN, GMDH, and MPMR. Petroleum Science and Technology, 2021, 39(19−20): 804−825

25. Pradeep T, Bardhan A, Burman A, Samui P. Rock strain prediction using deep neural network and hybrid models of ANFIS and Meta-Heuristic optimization algorithms. Infrastructure, 2021, 6(9): 129

26. Li N, Nguyen H, Rostami J, Zhang W, Bui X N, Pradhan B. Predicting rock displacement in underground mines using improved machine learning-based models. Measurement, 2022, 188: 110552

27. Lawal A I, Kwon S. Application of artificial intelligence to rock mechanics: An overview. Journal of Rock Mechanics and Geotechnical Engineering, 2021, 13(1): 248−266

28. Breiman L. Random forests. Machine Learning, 2001, 45(1): 5−32

29. Cutler A, Cutler D R, Stevens J R. Ensemble Machine Learning. Boston: Springer, 2012, 157−175

30. Louppe G. Understanding Random Forests from Theory to Practice. University of Liège, 2014

31. Gong M, Bai Y, Qin J, Wang J, Yang P, Wang S. Gradient boosting machine for predicting return temperature of district heating system: A case study for residential buildings in Tianjin. Journal of Building Engineering, 2020, 27: 100950

32. Konstantinov A V, Utkin L V. Interpretable machine learning with an ensemble of gradient boosting machines. Knowledge-Based Systems, 2021, 222: 106993

33. Zhou J, Li E, Yang S, Wang M, Shi X, Yao S, Mitri H S. Slope stability prediction for circular mode failure using gradient boosting machine approach based on an updated database of case

histories. Safety Science, 2019, 118: 505−518

34. Natekin A, Knoll A. Gradient boosting machines, a tutorial. Frontiers in Neurorobotics, 2013, 7

35. Awad M, Khanna R. Efficient Learning Machines. Berkeley: Apress, 2015, 67−80

36. Fan G F, Yu M, Dong S Q, Yeh Y H, Hong W C. Forecasting short-term electricity load using hybrid support vector regression with grey catastrophe and random forest modeling. Utilities Policy, 2021, 73: 101294

37. Roohi R, Emdad H, Jafarpur K. Toward a realistic reconstruction and determination of blood flow pattern in complex vascular network: 3D, non-Newtonian, multi-branch simulation based on CFD and GMDH algorithm. International Communications in Heat and Mass Transfer, 2021, 122: 105185

38. Ebtehaj I, Bonakdari H, Zaji A H, Azimi H, Khoshbin F. GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs. Engineering Science and Technology, an International Journal, 2015, 18(4): 746−757

39. Isah B W, Mohamad H, Ahmad N R, Harahap I S H, Al-Bared M A M. Uniaxial compression test of rocks: Review of strain measuring instruments. In: IOP Conference Series: Earth and Environmental Science, Volume 476, 2nd International Conference on Civil & Environmental Engineering. Langkawi, 2019.

40. Friedman J H. Greedy function approximation: A gradient boosting machine. Annals of statistics, 2001, 29(5): 1189−1232

41. Chen Y, Zheng W, Li W, Huang Y. Large group activity security risk assessment and risk early warning based on random forest algorithm. Pattern Recognition Letters, 2021, 144: 1−5

42. Jin Z, Shang J, Zhu Q, Ling C, Xie W, Qiang B. RFRSF: Employee turnover prediction based on random forests and survival analysis. In: International Conference on Web Information Systems Engineering. Cham: Springer, 2020, 503−515

43. Ko C N, Lee C M. Short-term load forecasting using SVR (support vector regression)-based radial basis function neural network with dual extended Kalman filter. Energy, 2013, 49: 413−422

44. Koopialipoor M, Nikouei S S, Marto A, Fahimifar A, Jahed Armaghani D, Mohamad E T. Predicting tunnel boring machine performance through a new model based on the group method of data handling. Bulletin of Engineering Geology and the Environment, 2019, 78(5): 3799−3813

45. Nourani V, Abdollahi Z, Sharghi E. Sensitivity analysis and ensemble artificial intelligence-based model for short-term prediction of NO_2 concentration. International Journal of Environmental Science and Technology, 2021, 18(9): 2703−2722

46. Wong F S. Slope reliability and response surface method. Journal of geotechnical Engineering, 1985, 111(1): 32−53

47. Srinivasulu S, Jain A. A comparative analysis of training methods for artificial neural network rainfall-runoff models. Applied Soft Computing, 2006, 6(3): 295−306

48. Willmott C J. Spatial Statistics and Models. Dordrecht: Springer, 1984, 443−460

49. Chai T, Draxler R R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature Geoscientific Model Development, 2014, 7(3): 1247−1250

50. Villeneuve M C, Heap M J, Kushnir A R L, Qin T, Baud P, Zhou G, Xu T. Estimating *in situ* rock mass strength and elastic modulus of granite from the Soultz-sous-Forêts geothermal reservoir (France). Geothermal Energy, 2018, 6: 1–29

51. Domede N, Parent T, Sellier A. Mechanical behaviour of granite: A compilation, analysis and correlation of data from around the world. European Journal of Environmental and Civil Engineering, 2019, 23(2): 193−211

52. Pradeep T, GuhaRay A, Bardhan A, Samui P, Kumar S, Armaghani D J. Reliability and prediction of embedment depth of sheet pile walls using hybrid ANN with optimization techniques. Arabian Journal for Science and Engineering, 2022, 1−19

53. Guven A, Kişi Ö. Estimation of suspended sediment yield in natural rivers using machine-coded linear genetic programming. Water Resources Management, 2011, 25(2): 691−704

54. Ayoubloo M K, Azamathulla H M, Jabbari E, Zanganeh M. Predictive model-based for the critical submergence of horizontal intakes in open channel flows with different clearance bottoms using CART, ANN and linear regression approaches. Expert Systems with Applications, 2011, 38(8): 10114−10123