Geomechanics and Engineering, *Vol. 33, No. 3 (2023) 279-289*
https://doi.org/10.12989/gae.2023.33.3.279

Soft computing based mathematical models for improved prediction of rock brittleness index

Abiodun I. Lawal[1,2a], Minju Kim[1b] and Sangki Kwon*[1]

[1]*Department of Energy Resources Engineering, Inha University Yong-Hyun Dong, Nam Ku, Incheon, Korea*
[2]*Department of Mining Engineering, Federal University of Technology, Akure, Nigeria*

(*Received December 15, 2021, Revised January 29, 2023, Accepted March 7, 2023*)

Abstract. Brittleness index (BI) is an important property of rocks because it is a good index to predict rockburst. Due to its importance, several empirical and soft computing (SC) models have been proposed in the literature based on the punch penetration test (PPT) results. These models are very important as there is no clear-cut experimental means for measuring BI asides the PPT which is very costly and time consuming to perform. This study used a novel Multivariate Adaptive regression spline (MARS), M5P, and white-box ANN to predict the BI of rocks using the available data in the literature for an improved BI prediction. The rock density, uniaxial compressive strength (σ_c) and tensile strength (σ_t) were used as the input parameters into the models while the BI was the targeted output. The models were implemented in the MATLAB software. The results of the proposed models were compared with those from existing multilinear regression, linear and nonlinear particle swarm optimization (PSO) and genetic algorithm (GA) based models using similar datasets. The coefficient of determination (R^2), adjusted R^2 (Adj R^2), root-mean squared error (RMSE) and mean absolute percentage error (MAPE) were the indices used for the comparison. The outcomes of the comparison revealed that the proposed ANN and MARS models performed better than the other models with R^2 and Adj R^2 values above 0.9 and least error values while the M5P gave similar performance to those of the existing models. Weight partitioning method was also used to examine the percentage contribution of model predictors to the predicted BI and tensile strength was found to have the highest influence on the predicted BI.

Keywords: brittleness index; MARS; punch penetration test; soft computing; uniaxial compressive strength

1. Introduction

Brittleness index (BI) is one of the key properties of rocks and it is of great interest in practical rock engineering as it concerns the safety of lives and property. For instance, BI is a major predictor of rockburst which has direct link with the stability of the host rock mass most especially in underground mines and tunnels. However, unlike other intact rock properties such as σ_c, point load, density with clear cut procedures in determining them, there is no general agreement among the researchers (Lawn *et al*. 1976, Cheng *et al.* 2015, Altindag and Guney 2010, etc.) regarding the definition and means of estimating the BI. BI is commonly estimated using the stress-strain relationships of the intact rock. However, Yagiz (2009) used the punch penetration test approach to provide a useful insight into the estimation of BI. Yagiz (2009) generated various regression models and discovered that the combination of rock density, σ_c and σ_t can give a BI with coefficient of determination (R^2) of 0.884. Some of the numerous existing BI estimation models based on the stress-strain curves are presented in Table 1 (Xia *et al*. 2019).

Asides the empirical models, the flexibility of soft computing techniques has also been employed in this aspect of rock mechanics just like the predictions of rock properties such as uniaxial compressive strength and shear strength properties among others (Armaghani *et al*. 2021, Fattahi and Hasanipanah 2021, Sun *et al*. 2021, Lawal *et al.* 2022, etc.). For example, Kaunda and Asbury (2016) adopted artificial neural network (ANN) to predict the BI based on various datasets obtained from Earth Mechanics Institute (EMI), Colorado School of Mines laboratory. However, the performance of their model is not very high as they obtained R^2 value of 0.706. Yagiz and Gokceoglu (2010) also used nonlinear multiple regression (MR) and Fuzzy inference system (FIS) to predict BI based on Yagiz (2009) experimental datasets.

The nonlinear MR performed better than the FIS with R^2 value of 0.898. Yagiz *et al.* (2018) proposed linear and nonlinear models based on the PSO and GA techniques using the datasets in Yagiz (2009). There is a little improvement in the performance of the models they proposed as compared with the original regression models in Yagiz (2009). Hussain *et al.* (2018) also conducted similar study based on stress ratio as against the punch penetration test (PPT) of Yagiz (2009). They proposed linear and nonlinear models using PSO and imperialism competitive algorithm (ICA) techniques. However, the major problem with the GA, PSO, ICA is that they require prior mathematical relationship between the dependent and independent variables. Sun *et al*. (2020) also used random forest (RF), support vector machine (SVM), K-nearest

*Corresponding author, Professor
E-mail: kwonsk@inha.ac.kr
[a]Ph.D.
E-mail: ailawal@futa.edu.ng

ISSN: 2005-307X (Print), 2092-6219 (Online)

Table 1 Empirical models for rock brittleness prediction based on stress-strain relationship

Equation	Authors	Parameters considered
$BI = \sigma_c / \sigma_t$	Altindag (2010)	σ_c and σ_t are the respective UCS and TS.
$BI = (\sigma_c - \sigma_t)/(\sigma_c + \sigma_t)$		
$BI = 0.5\sigma_c\sigma_t$		
$BI = \sqrt{0.5\sigma_c\sigma_t}$		
$BI = \sin\beta$	Hucka and Das (1974)	β stands for the friction angle
$BI = 0.5(E^{norm} + \mu^{norm})$ with $E^{norm} = 100(E-1)/7$ $\mu^{norm} = 100(\mu - 0.4)/-0.25$	Rickman *et al.* (2008)	E^{norm} and μ^{norm} are the normalized elastic modulus and Poison ratio
$BI = (3K - 5\lambda)/\lambda$	Huang *et al.* (2015)	K and λ stand for the respective bulk modulus and Lame's constant
$BI = E/\lambda$	Meng *et al.* (2015)	E is the elastic modulus
$BI = E.\rho$	Rybacki *et al.*(2016)	ρ is the rock density
$BI = (\lambda + 2G)/\lambda$	Guo *et al.* (2015)	G is the shear modulus
$BI = (\sigma_p - \sigma_r)/\sigma_p$	Rybacki *et al.*(2016)	σ_p and σ_r are the respective peak and residual strengths
$BI = \varepsilon_R/\varepsilon_p$	Hucka and Das (1974)	ε_R is the pre-peak recovery strain
$BI = (\varepsilon_f^p - \varepsilon_c^p)/\varepsilon_c^p$	Hajiabdolmajid and Kaiser (2003)	ε_f^p and ε_c^p are the respective plastic strains for breaking between grains and for shearing of new surfaces.
$BI = (\varepsilon_r - \varepsilon_p)/\varepsilon_p$		ε_p and ε_r are the respective peak and residual strains
$BI = W_e/W_p$	Andreev (1995)	We and Wp are the respective recoverable elastic energy and the bursting energy stored prior to fracturing.
$BI = W_{ini}/W_p$		W_{ini} energy at the beginning of fracture initiation
$BI = M/(M+E)$	Rybacki *et al.* (2016)	M is the secant modulus
$BI = H/E$		H is the hardening modulus
$BI = (M-E)/M$	Tarasov and Potvin (2013)	
$BI = 0.198\sigma_c - 2.174\sigma_t + 0.193\rho - 3.807$	Yagiz (2009)	ρ is the density, σ_c is the uniaxial compressive strength and σ_t is the tensile strength

neighbors (KNN), and ANN to predict the BI obtained from stress-strain relationship. Koopialipoor *et al*. (2019) proposed ANN and FA-ANN (Firefly Algorithm and ANN) in predicting estimated BI based on non-destructive rock properties. They do not make available mathematical models that could be of interest to the field engineers in their studies.

As evident in the previous paragraphs, several empirical models and soft computing (SC) models have been proposed in literature. These models are very important as there is no clear-cut experimental means for measuring BI asides the PPT which is expensive and time consuming to perform. The performance of the empirical models in predicting BI should be closed to the measured BI due to safety requirements. Therefore, in this study, Multivariate adaptive regression spline (MARS), M5P and white-box ANN approaches have been proposed for accurate prediction of BI. None of the existing studies have used Multivariate adaptive regression spline (MARS) and M5P methods to predict the BI. The advantage of this method is that it does not require prior relationships between the input and output variables. ANN on the other hand has previously been used but the white-box form of ANN is proposed in this study which is new in predicting BI.

1.1 Brief overview of the adopted methods

1.1.1 ANN model

Artificial neural network (ANN) is the most used artificial intelligence method in various fields of science and engineering most importantly for prediction purpose. Like its name implies, it artificially mimics the nature and manner of operation of the human neurons. Human being has thousands of highly interconnected neurons which make the information processing and interpretation possible. However, the neurons are independently placed in layers which are connected to the next layer via a series of interconnecting weights. The output of the previous layer is usually made to pass through the transfer function in the next layer. The transfer function could either be nonlinear (e.g., tansigmod or logsigmod) or linear (e.g., purelin). The learning and training processes are paramount in ANN

simulation to enable the network to learn from the series of datasets presented to it and apply the residual knowledge when a new set of datasets is presented to it. Many algorithms have been proposed to take the ANN through the training and learning stage, but the backpropagation (BP) training algorithm coupled with the Levenberg-Marquardt (LM) training function has been found most versatile. This is also adopted in this study. The general mathematical expression of the BP-ANN model for a single hidden layer is given in Eq. (1) (Lawal and Idris 2019, Lawal 2020). The detail procedures followed in implementing the ANN are presented in section 2.

$$Y_j = f\left(\theta_j + \sum_{i=1}^{n} w_{ji} x_i\right) \tag{1}$$

where Y_j stands for the model predicted value; θ_j represents the bias; n stands for the number of neurons in the inner lay er; w_{ji} represents the weight connecting the ith entry layer a nd the jth inner layer; x_i represents the model input value; a nd f stands for the transfer function.

1.1.2 Multivariate adaptive regression spline

MARS which is an adaptive modelling method was proposed by Friedman (1991) for solving non-linear relationships between variables. The mode of operation of MARS according to Leathwick *et al.* (2005) is piece-wise division of independent variables into linear segments to describe non-linear relationship between the dependent and targeted variables. MARS also offers flexible modelling method for high dimensional data coupled with less training time. The model generated using MARS can be easily interpreted and this is one of its advantages over other regression and classification techniques (Yuvaraj *et al.* 2013, Lawal *et al*. 2021a).

MARS model undergoes a supervised training method in which the independent and the targeted datasets are presented to it during training. The MARS model divides the training data into many splines at equal interval. MARS further divides the datasets into several subgroups and generates many knots which can be found between separate input parameters or separate intervals in the same input parameter, to separate the subgroups. The MARS model utilizes a smoothing spline to approximate the regression function named basis function (BF) for the general representation of the datasets in each subgroup (Friedman 1991, Yuvaraj *et al*. 2013). The characterisation of the datasets can be global or linear regression between any two knots with a distinct BF. The MARS predetermined the shortest distance between two neighbouring knots to forestall bias of the datasets in a subgroup and thereby prevents the overfitting or over-regression of the model.

The general mathematical form of the MARS model according to Friedman (1991) is presented in Eq. (2)

$$y = \hat{f}(x) = c_0 + \sum_{i=1}^{N} c_i B_i^{(q)}(x) \tag{2}$$

where y stands for the dependent variable, c_0 is constant of the equation while c_i represents a vector coefficients of the non-constant basis functions (i=1,2,…,N) and $B_i^{(q)}$ are the basis functions selected for inclusion in model of *qth* order given by Eq. (3)

$$B_i^q(x) = \prod_{t=1}^{k_i} [b_{ti}(x_{v(t,i)} - h_{ti})]_+^q \tag{3}$$

$B^q{}_i(x)$ is the vector of non-constant basis functions, i is the number of non-constant basis functions (1,2,…N), q is the power to which the split is raised for controlling the degree of smoothness of the total function estimate and is equal to 1 in the case while '+' means that the only positive result in the right hand side of Eq. (3) is valid; otherwise, the functions become zero (0).

The MARS model generally contains two basic steps or processes which are a forward process and a backward process. The basis functions are chosen to define Eq. (2) in the forward process while in the backward process, the deletion of the ineffective basis functions in the model takes place. The deletion is achieved using the generalized cross-validation (GCV) criterion (Craven and Wahba 1979) presented in Eq. (4).

$$GCV = \frac{\frac{1}{m}\sum_{i=1}^{m}[y_i - \hat{f}_B(x_i)]^2}{\left[1 - \frac{C(B)}{n}\right]^2} \tag{4}$$

where m stands for the number of input data, $f_B(x_i)$ is the basis function model, and $C(B)$ stands for the complexity penalty which varies proportional to the number of basis functions and it is defined as Eq. (5)

$$C(B) = (B+1) + dB \tag{5}$$

where B is the number of non-constant basis functions and d is the associated penalty to each basis function in the model. The d is otherwise known as the smooth parameters. More information about the MARS model can be found in (Friedman 1991).

1.1.3 M5P

M5P is a soft computing method introduced by Quinlan (1992) for solving regression problems based on decision tree learner. M5P algorithm allocates the linear regression functions at the terminal nodes and fits a multilinear regression model to individual subspace by separating the overall data space into many sub spaces. It works on the continuous class variable and not discrete classes with the ability to operate tasks with very high dimensionality. Piecewise information of individual linear model built to approximate nonlinear relationships of the datasets are revealed by M5P.

The details of splitting criteria for the M5P model tree is known based on the error computation at each node. The standard deviation of the class values arriving at the node is normally used to compute the error. Eq. (6) is used for the computation of standard deviation reduction (Sihag *et al*. 2019).

$$SDR = sd(K) - \sum \frac{|K_i|}{|K|} sd(K_i) \tag{6}$$

Table 2 The statistics of the adopted datasets

	ρ_r	σ_c	σ_t	BI
Min	20.5	9.5	2.3	10
Max	28.9	327	17.8	45
Skewness	-0.704	0.742	0.92	-0.046
Kurtosis	-0.259	0.705	1.229	-0.954
STD	2.123	70.259	3.411	9.413

where K stands for set of instances that reach the node; K_i stands for the subset of illustrations with the ith product of the possible set while *sd* stands for standard deviation.

2. Material and method

2.1 Data acquisition

The dataset obtained from Yagiz (2009) was used in developing the proposed models. The said dataset comprised the intact rock mass properties such as σ_c, σ_t and rock density (ρ) which were determined based on the recommended ASTM (1995) D3967, D2938, D4543 standards. The BI based on PPT was also obtained from the same literature. The datasets cut across all types of rocks including the sedimentary, metamorphic, and igneous rock types. In total, 48 datasets were obtained and used for the proposed models. This datasets is enough to produce reliable soft computing models as many authors (e.g., Dehghan *et al*. 2010, Ebrahimi *et al.* 2015, Akinwekomi and Lawal 2021, etc.) have used less than that amount of datasets to perform different reliable ANN models. The statistical evaluation of the dataset is also presented in Table 2. The minimum and maximum values of the density are very close. This can be attributed to the fact quartz content usually dominates the properties of rocks therefore there is no wide variations in the density of the same or similar rock types (Lawal *et al*. 2021b). The skewness and the kurtosis of the datasets revealed that they are not normally distributed but they have sufficient normality based on the general guideline that states that – or + 1 skewness and kurtosis values indicate sufficient normality. However, since the proposed models is soft computing and the distribution of the datasets may not have influence on the model performance, the soft computing method will still give a reasonable prediction.

2.2 Model developments

2.2.1 ANN model

To perform the proposed ANN model in this study, the adopted dataset were loaded into the MATLAB software and were automatically divided into training, testing and validation datasets in the ratio 0.7:0.15:0.15 respectively.

The number of neurons in the hidden layer and the number of iterations were defined. The network was then trained using both the BP and LM training function. While the number of neurons in the input and output layers are dictated by the number of dependent and independent variables (Lawal and Kwon 2020), the number of neurons in the hidden layer was varied between 1 and 10 to obtain the most suitable among them. The number of layers used were three because it has been found suitable for solving many engineering problems (Bishop 1995, Adegoke *et al.* 2019, Wang *et al*. 2020, Akinwekomi and Lawal 2021, Lawal *et al*. 2022) while the nonlinear transfer function was used in both the hidden and output layers. The performance of different variations of the number of hidden neurons based on their respective coefficient of correlation is presented in Table 3. From Table 3, 3-10-1 ANN architecture performed better than the remaining ANN architectures and therefore selected as the optimum ANN architecture (Fig. 1). The overall performance of the proposed ANN model is also presented in Fig. 2. It can be observed in Fig. 2 that the proposed ANN model is promising based on the closeness of the predicted output to the targeted output and the error values that are close to zero.

Table 3 Simulated ANN architectures

ANN architectures	Different phases			
	Training	Testing	Validation	Overall
	R			
3-1-1	0.9472	0.92306	0.93493	0.9408
3-2-1	0.94452	0.94498	0.98364	0.95176
3-3-1	0.97044	0.92164	0.95569	0.96221
3-4-1	0.96198	0.93359	0.98902	0.96078
3-5-1	0.98506	0.99295	0.94324	0.98006
3-6-1	0.97738	0.96701	0.99659	0.9792
3-7-1	0.98385	0.97126	0.99117	0.98225
3-8-1	0.98857	0.99017	0.99067	0.98693
3-9-1	0.85307	0.92714	0.96067	0.86766
3-10-1	0.98997	0.98779	0.99635	0.99145

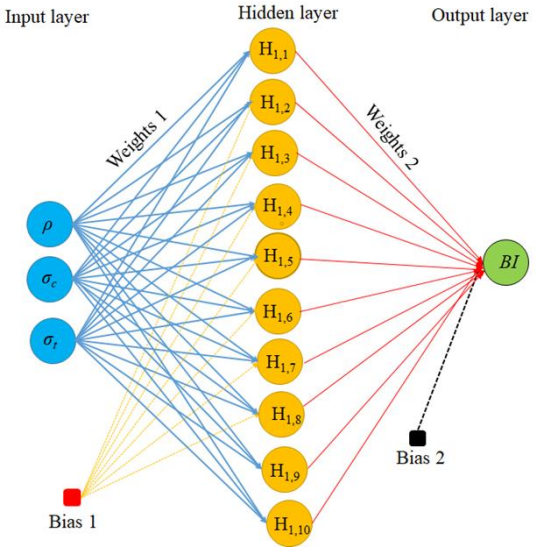

Fig. 1 Selected ANN architecture

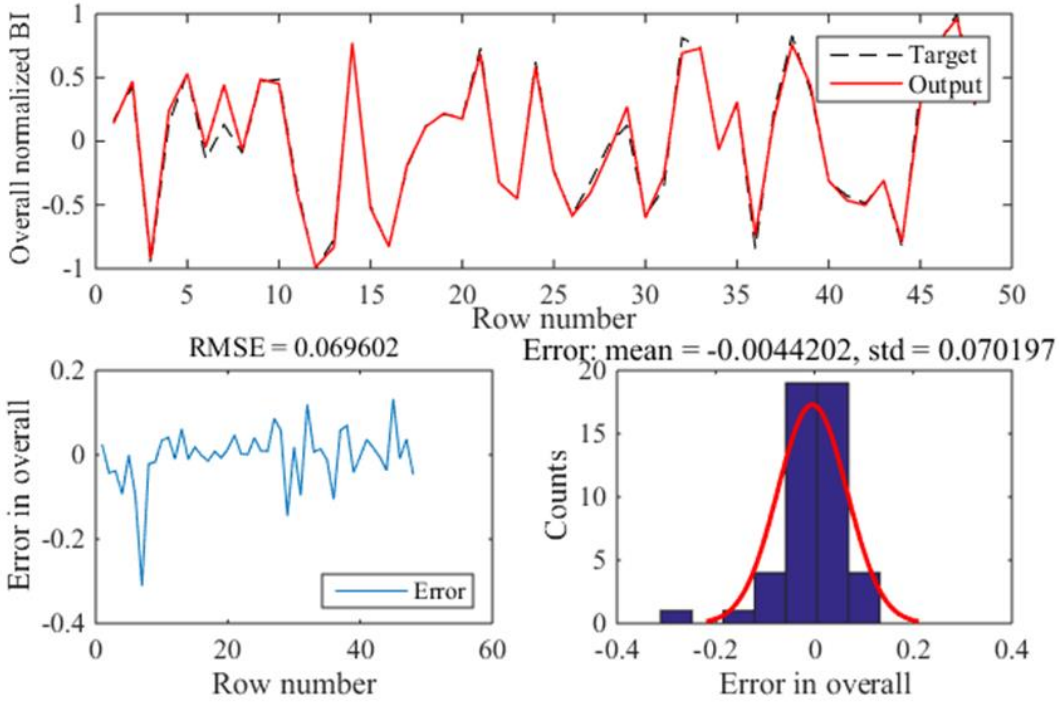

Fig. 2 Overall performance of the proposed ANN model

Table 4 Obtained weights and biases from the ANN simulation

Neurons	Weights			Biases		
	ρ_r^n	σ_c^n	σ_t^n	BI_m^n	b_1	b_2
1	7.290819	-7.10845	-2.76963	4.545885	-5.02503	0.528689
2	3.300231	-1.10186	-3.24294	-3.87509	-4.73859	
3	-5.76691	4.952261	2.774721	6.32796	4.03145	
4	6.162481	-3.0737	8.029862	-1.06319	-4.40158	
5	1.439778	-1.76678	8.788685	-1.11329	2.11846	
6	-1.22552	0.391079	0.006532	-8.3001	1.128402	
7	-3.20701	1.354774	3.964871	2.34039	-1.04862	
8	0.931197	6.425634	4.486773	1.849623	5.732838	
9	0.346029	6.93443	9.519873	-1.98765	9.337689	
10	1.73818	4.79948	-0.94923	1.811273	5.197305	

One of the unique features of the proposed ANN model is that it is not limited to just prediction as it is common in many of the existing ANN studies but also transformation of the obtained weights and biases (Table 4) from the optimum ANN architecture into a useful mathematical form.

The transformation of the weights and biases into a useful mathematical model can be termed transformation from the black-box ANN to the white box ANN. The obtained mathematical form of the simulated ANN is presented in Eq. (7) leveraging on Eq. (1).

where x_1 to x_{10} in Eq. (7) are given in Eq. (8).

$$BI = 17.5\tanh\left(\sum_{i=1}^{10} x_i + 0.5287\right) + 27.5 \tag{7}$$

$$\begin{cases} x_1 = 4.5459\tanh(7.2908\rho_r^n - 7.1084\sigma_c^n - 2.7696\sigma_t^n - 5.0250) \\ x_2 = -3.8751\tanh(3.3002\rho_r^n - 1.1019\sigma_c^n - 3.2429\sigma_t^n - 4.7386) \\ x_3 = 6.328\tanh(-5.7669\rho_r^n + 4.9523\sigma_c^n + 2.7747\sigma_t^n + 4.0314) \\ x_4 = -1.0632\tanh(6.1625\rho_r^n - 3.0737\sigma_c^n + 8.0299\sigma_t^n - 4.4016) \\ x_5 = -1.1133\tanh(1.4398\rho_r^n - 1.7668\sigma_c^n + 8.7887\sigma_t^n + 2.1185) \\ x_6 = -8.3001\tanh(-1.2255\rho_r^n + 0.3911\sigma_c^n + 0.0065\sigma_t^n + 1.1284) \\ x_7 = 2.3404\tanh(-3.2070\rho_r^n + 1.3548\sigma_c^n + 3.9649\sigma_t^n - 1.0486) \\ x_8 = 1.8496\tanh(0.9312\rho_r^n + 6.4256\sigma_c^n + 4.4868\sigma_t^n + 5.7328) \\ x_9 = -1.9876\tanh(0.346\rho_r^n + 6.9344\sigma_c^n + 9.5199\sigma_t^n + 9.3377) \\ x_{10} = 1.8113\tanh(1.7382\rho_r^n + 4.7995\sigma_c^n - 0.9492\sigma_t^n + 5.1973) \end{cases} \tag{8}$$

It should be noted that the normalized form of the model independent variables are used in Eq. (8) which can be

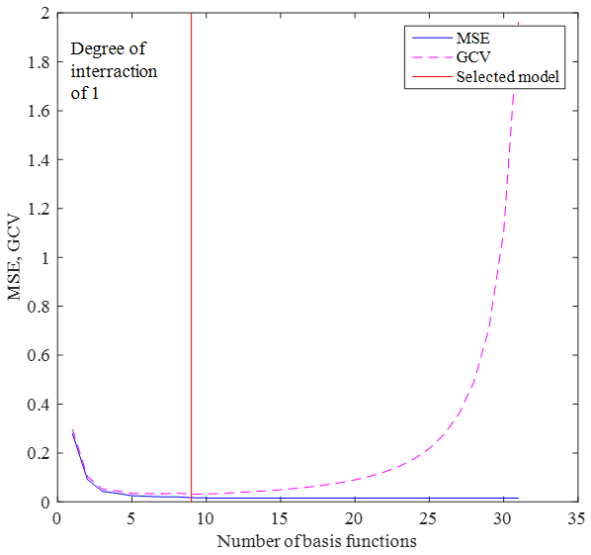

Fig. 3 Selection of the number of the basis functions

Table 5 Obtained basis functions and their associated equations

BF	Equations
BF1	max(0,0.42857 –ρ^n)
BF2	max(0, ρ^n -0.19048)
BF3	max(0,0.19048 – ρ^n)
BF4	max(0, σ^n_c +0.54961)
BF5	max(0,-0.54961 – σ^n_c)
BF6	max(0, σ^n_c +0.30394)
BF7	max(0,-0.058065 – σ^n_t)
BF8	max(0, σ^n_t +0.17419)

obtained using the formula in the footnote[1]. Therefore, with Eqs. (7) and (8), the result of the selected optimum ANN can be replicated.

2.2.2 Multivariate adaptive regression spline (MARS)

The MARS described above is used in the study to predict the BI. Since training phase is paramount to MARS, the dataset used in the training are the same with those used in training of the ANN model. Similarly, the same datasets used for testing and validation in ANN model are also used. The dataset was loaded into the MATLAB and the MARS is implemented using the "Adaptive Regression Splines toolbox for MATLAB". The piece-wise linear basis function was used while training the model and first order degree of interaction were tried. Some information about the BF is provided in Fig. 3. The MSE and GCV are plotted against the number of BF while the selected number of BF is indicated in Fig. 3. Based on the minimum MSE, the selected number of BF is eight (8) excluding the intercept which is constant as established in MARS model. The said eight (8) BFs are in principle displayed in Table 5. The transformation of the BF in Table 5 using the corresponding coefficients as obtained from the MATLAB into the required MARS model for the prediction of BI is presented in Eq. (9).

$$BI = -0.98904 + 1.2703BF1 + 0.68803BF2 - 1.3684BF3 + 2.7891BF4 - 1.8204BF5 - \ldots 1.322BF6 + 0.78152BF7 - 0.77413BF8 \quad (9)$$

2.2.3 M5P

M5P model described above is implemented in this study using the same datasets used for training, testing and validation of the ANN and MARS models. However, it was built in Weka software package. The already partitioned datasets were loaded to the software and the M5P simulation was carried out. The obtained M5P model tree with the smoothened linear models is presented in Fig. 4. B in Fig. 4 stands for σ_c. The weights of attributes related to the targeted output is expressed in Eqs. (10)-(12). The Eqs. (10)-(12) can be used in conjunction with Fig. 4. For example, when B (i.e., σ_c) is less than or equal to 102.5 MPa, Eq. (10) will be used. However, if B (i.e. σ_c) is greater than 102.5 MPa and 197.5 MPa respectively, Eq. (12) will the most suitable and if otherwise, Eq. (11).

LM num: 1

$$BI_{M5P} = 0.6987\rho_r + 0.0909\sigma_c + 0.0385\sigma_t - 1.1012 \quad (10)$$

LM num: 2

$$BI_{M5P} = 0.5099\rho_r + 0.0442\sigma_c - 0.1023\sigma_t + 13.501 \quad (11)$$

LM num: 3

$$BI_{M5P} = 0.5099\rho_r + 0.4932\sigma_t + 16.7856 \quad (12)$$

3. Results and discussion

3.1 Model comparison

The performance of the proposed models was compared with the measured values to determine which of the two models is closer to the measured BI. The comparison was made on the overall predictions of the soft computing based mathematical model presented in Eqs. (7)-(12). The outcome of the comparison is presented in Fig. 5. From Fig. 5(a), the prediction of the ANN model is very close to that of the measured BI as compared with the MARS and M5P models, although all the models show a very good prediction ability with strong coefficient of determination.

The R^2 values of the ANN, MARS and M5P models are 0.983, 0.904, and 0.811 as shown in Fig. 5(b) indicating that ANN model outperformed MARS and M5P models.

In addition to the comparison presented in Fig. 5, the performances of the proposed models are still depicted in Taylor's diagram as presented in Fig. 6. The Taylor's diagram comprises the standard deviation, root mean squared difference and the correlation coefficient. It enables easy understanding or identification of the model that is the closest to the measured values. From Fig. 5, ANN appears

[1] $X_{norm}=2(X-X_{min})/(X_{max}-X_{min})-1$, where X is the actual variable to be normalized and Xmin and Xmax are the minimum and maximum values X while Xnorm is the normalized value of X.

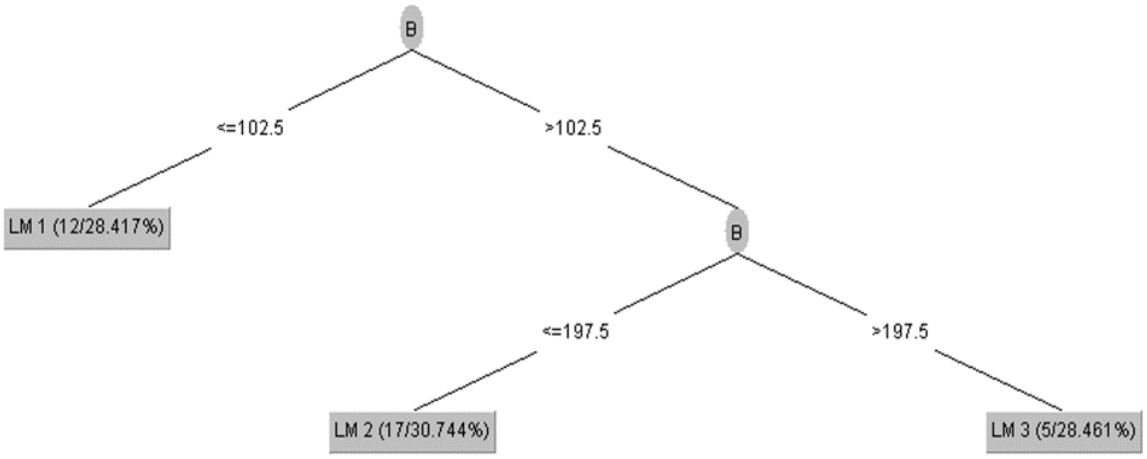

Fig. 4 M5P tree model for BI prediction

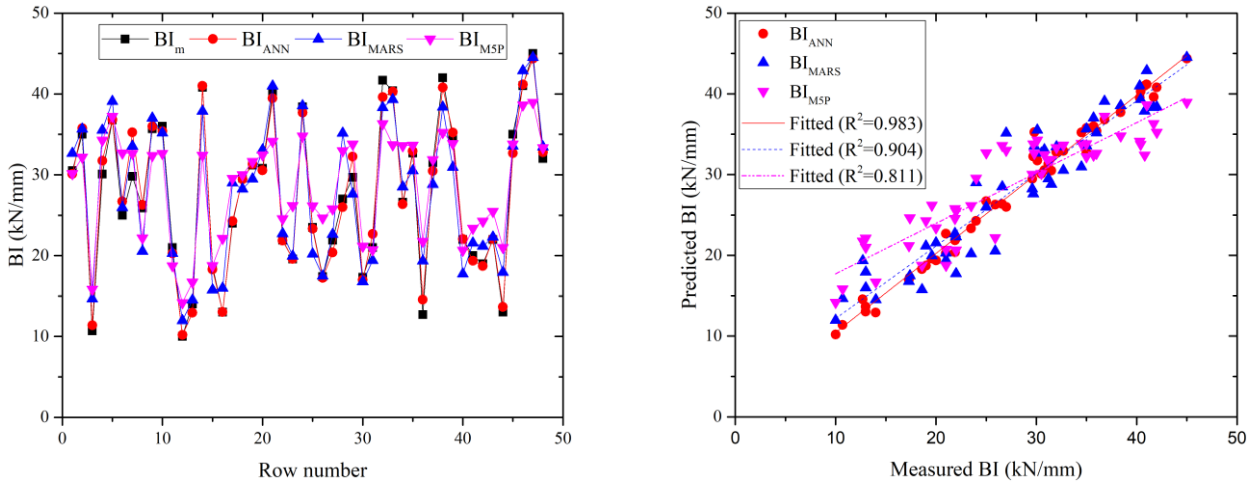

Fig. 5 Comparison of the proposed models with the measured values (a) BI and row number and (b) Predicted BI and Meas ured BI

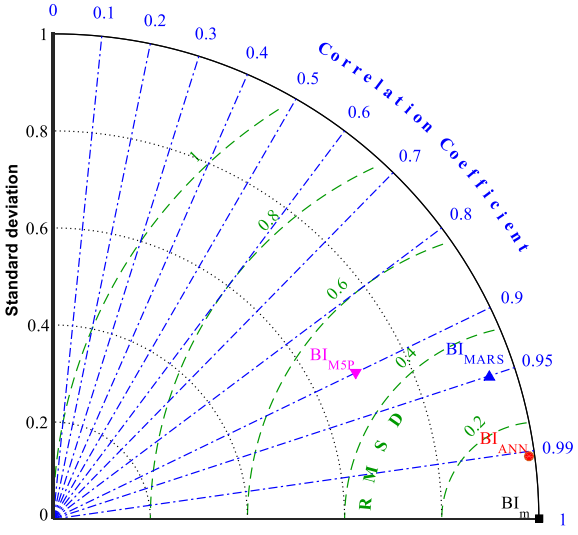

Fig. 6 Taylor's diagram comparing the models

to be the closest follow by MARS and M5P model. All the models' coefficient of correlations are above 0.9.

3.2 Comparison with the existing models

To further assess the performance of the proposed models, they were compared with the previous studies performed using the same kind of datasets adopted in this study. Originally, Yagiz (2009) proposed multiple linear regression relationship between the ρ, σ_c and σ_t and they later used different AI methods such as PSO and GA to propose another four equations tagged PSO linear and non-linear and GA linear and non-linear respectively. The performances of these existing models are compared statistically with the newly proposed models in this study to establish which of these models is closer to the measured values. The R^2, Adj R^2, RMSE, and MAPE are the performance indicators as shown in Table 4. The performances of the proposed models most importantly ANN model are better than the existing models developed based on the datasets adopted in this study.

The most accurate model will be the one with the R^2 and Adj R^2 values of 1, the percentage difference between the expected R^2 and Adj R^2 values for the ANN model are 1.7% for both R^2 and Adj R^2 while that of MARS are 9.6% and

Table 6 Comparison with the existing models

SI	Present			Y	YG	Ys			
	ANN	MARS	M5P	MR_L	MR_{NL}	PSO_L	GA_L	PSO_{NL}	GA_{NL}
R^2	0.983	0.904	0.811	0.885	0.899	0.800	0.883	0.858	0.898
Adj R^2	0.983	0.902	0.807	0.883	0.896	0.795	0.881	0.855	0.896
RMSE	1.219	2.920	4.636	3.165	2.975	4.206	3.250	3.518	3.05
MAPE	3.086	10.293	17.610	11.219	10.455	15.3	10.625	12.208	10.218

*SI-statistical indicators; MR_L-multiple regression linear; MR_{NL}- multiple regression nonlinear; PSO_L-PSO linear; GA_L-GA linear; PSO_{NL}-PSO nonlinear; ; GA_{NL}-GA nonlinear; Y- Yagiz (2009) YG- Yagiz and Gokceoglu (2010); Ys-Yagiz *et al*. (2018)

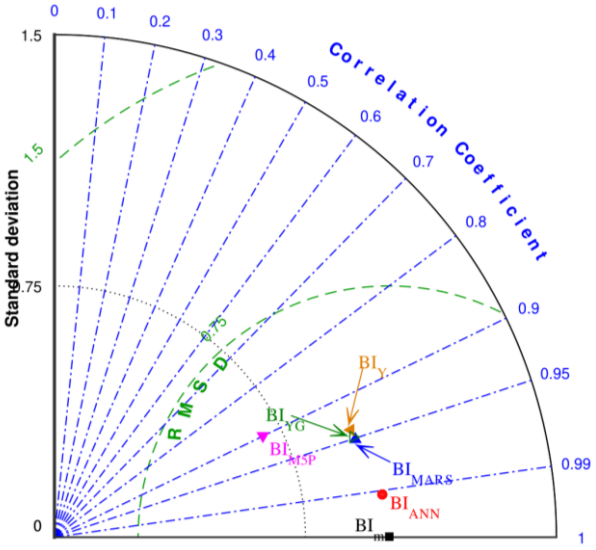

Fig. 7 Comparison of the proposed model with Yagiz (2009) (BIY) and Yagiz and Gokceoglu (2010) (BIYG)

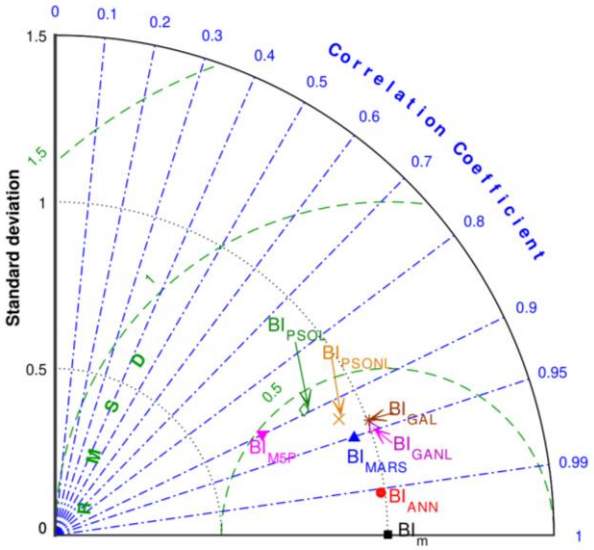

Fig. 8 C omparison of the proposed models with Yagzi *et al.* (2018) (BI_{PSOL}-BI predicted using PSO linear; BI_{GAL}-BI predicted using GA linear; BI_{PSONL}-BI predicted using PSO nonlinear and BI_{GANL} -BI predicted using GA nonlinear)

9.8% for the respective R^2 and Adj. R^2. For the M5P, the obtained percentage difference between the expected R^2 and Adj R^2 are 18.9% and 19.3% respectively. However, for the existing studies presented in Table 4, the best among them that is MR_{NL} has percentage difference of R^2 and Adj R^2 to be 10.2% and 10.4% respectively while the least performing model among them (PSO_L) has the respective percentage difference of R^2 and Adj R^2 values of 20% and 20.5% respectively. Figs. 7 and 8 also show the performances of the proposed models using the Taylor's diagram. The proposed models most importantly the ANN model has STD, RSMD and Pearson's coefficients that are close to those of the measured BI. MARS model also showed a promising performance when compared with the previous models. However, its performance is closer to the MR_{NL} and GANL models albeit it is slightly better than them.

It is also important to state here that the proposed ANN is white-box in nature in that suitable mathematical model form of the ANN has been proposed and the MARS model is also easy to implement. Hence, with the newly proposed models, more accurate prediction of the BI is guaranteed. Furthermore, the imperativeness of the proposed models cannot be overemphasized in that the BI is usually estimated from stress-strain curve while the only experimental means of obtaining BI is through Punch penetration test. The cost of performing this experiment is high and it is time consuming. The test apparatus is not also readily available as a result, models such as those presented in Table 6 have been made available for estimating the BI based on the available PPT experimental results, but BI requires model (s) that is (are) very close to the actual BI due to its import in rockburst prediction and prevention. Therefore, the proposed models will be very useful and practically adopted at this time that the experimental data of BI are not readily available.

3.3 Sensitivity analysis

The percentage contribution of each of the model independent parameters was investigated in this study. This was done using the ANN model as it is the best among the proposed models. This is also done as another way of unraveling the black-box feature of ANN (Gevrey *et al.* 2003). For this purpose, different methods have been proposed but the weights partitioning method has been adopted in this study. The method makes use of the input-hidden and output layers connecting weights of the ANN to establish the contributions of each of the model inputs into the predicted output. The formula for computing the

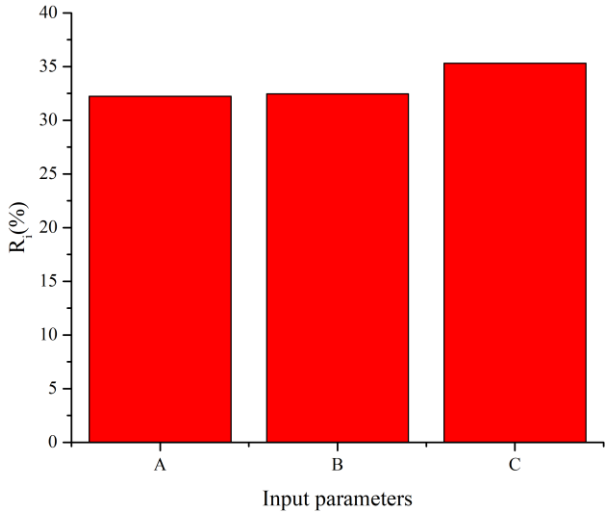

Fig. 9 Percentage contribution of input parameters

percentage contributions is presented in Eq. (13) (Garson 1991).

$$R_i(\%) = \frac{\sum_{h=1}^{nh} \frac{|W_{ih}| \times |W_{ho}|}{\sum_{i=1}^{ni} \left(|W_{ih}| \times |W_{ho}| \right)}}{\sum_{h=1}^{nh} \sum_{i=1}^{ni} \frac{|W_{ih}| \times |W_{ho}|}{\sum_{i=1}^{ni} \left(|W_{ih}| \times |W_{ho}| \right)}} \times 100 \quad (8)$$

where Ri(%) are the percentage contributions of each of the input variables, Wih is the input-hidden layer connecting weights, Who is the hidden-output layer connecting weight, ni is the neurons in the input layer, and nh is the neurons in the hidden layer. The obtained percentage contributions of the model parameters is presented in Fig. 9. In Fig. 9, A stands for rock density (ρ), B stands for σc while C stands for σ_t. The tensile strength has the highest influence on the predicted BI follow by σ_c and rock density.

4. Conclusions

The novel MARS, M5P, and ANN were proposed in this study to predict rock brittleness index based on the dataset from punch penetration test obtained from literature. The models' input parameters were the intact rock properties such as rock density, σ_c and σ_t. The MARS, M5P, and ANN models' results were compared with those of the existing linear regression, PSO and GA linear and non-linear models in the literature. The performance of the models was compared using R^2, Adj R^2, RMSE, and MAPE as the performance indicators. The ANN and MARS models' abilities to predict the BI of rocks were found to be higher, than the other models compared in this study. The ANN model followed by MARS model were found to be more suitable for BI prediction based on the outcomes of the performance indicators as they have highest R^2 and Adj R^2 values with least RMSE and MAPE values. Since there is no agreed procedure for laboratory determination of BI and the Punch penetration test seems to be the direct means of measuring BI of rocks. The proposed models will be of practical importance for accurate and timely prediction of BI. The sensitivity analysis conducted revealed that tensile strength influenced BI most.

Acknowledgments

This work was supported by Inha University Research Grant (2023).

References

Adegoke, M., Wong, H.T., Leung, A.C.S. and Sum, J. (2019), "Two noise tolerant incremental learning algorithms for single layer feed-forward neural networks", *J. Ambient Intell. Human Comput*., https://doi.org/10.1007/s12652-019-01488-8.

Akinwekomi, A.D. and Lawal, A.I. (2021), "Neural network-based model for predicting particle size of AZ61 powder during high energy mechanical milling", *Neural. Comput. Appl*., **33**, 17611-17619, https://doi.org/10.1007/s00521-021-06345.

Altindag, R. (2010), "Assessment of some brittleness indices in rock-drilling efficiency", *Rock Mech. Rock Eng*., **43**(3), 361-370. https://doi:10.1007/s00603-009-0057-x.

Altindag, R. and Guney, A. (2010), "Predicting the relationships between brittleness and mechanical properties (UCS, TS and SH) of rocks", *J. Sci. Res. Essay*, **5**, 35-39. https://doi.org/10.5897/SRE.9000753.

Andreev, G.E. (1995), *Brittle failure of rock materials: Test results and constitutive models*, Rotterdam: A. A. Balkema.

Armaghani, D.J., Mamou, A., Maraveas, C., Roussis, P.C., Siorikis, V.G., Skentou, A.D. and Asteris, P.G. (2021), "Predicting the unconfined compressive strength of granite using only two non-destructive test indexes", *Geomech. Eng*., **25**(4), 317-330. https://doi.org/10.12989/gae.2021.25.4.317.

ASTM (1995), *Standard practice for preparing rock core specimens and determining dimension and shape tolerances*. American Society for Testing and Materials. D4543.

ASTM (1995), *Standard test method for splitting tensile strength of intact rock core specimens*. American Society for Testing and Materials. D3967.

ASTM (1995), *Standard test method for unconfined compressive strength of intact rock core specimens*. American Society for Testing and Materials. D2938.

Bishop, C.M. (1995), *Neural network for pattern recognition*, 1st Ed. Oxford University Press.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984), *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software. ISBN 978-0-412-04841-8.

Cheng, W., Jin, Y. and Chen, M. (2015), "Reactivation mechanism of natural fractures by hydraulic fracturing in naturally fractured shale reservoirs". *J. Nat. Gas Sci Eng*., **23**, 431-439. https://doi:10.1016/j.jngse. 2015.01.031.

Craven, P. and Wahba, G. (1979). "Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation". *Numer. Math*., **31**, 317–403. https://doi.org/10.1007/BF01404567.

Dehghan, S., Sattari, G., Chelgani, S.C. and Aliabadi, M. (2010), "Prediction of uniaxial compressive strength and modulus of elasticity for Travertine samples using regression and artificial neural networks", *Min. Sci. Technol*., **20**, 41-46. https://doi.org/10.1016/S1674-5264(09)60158-7.

Ebrahimi, E., Monjezi, M., Khalesi, M.R. and Armaghani, D.J. (2015), "Prediction and optimization of back-break and rock

fragmentation using an artificial neural network and a bee colony algorithm", *Bull. Eng. Geol. Environ*., **75**, 27-36. https://doi.org/10.1007/s10064-015-0720-2.

Fattahi, H. and Hasanipanah, M. (2021), "Predicting the shear strength parameters of rock: A comprehensive intelligent approach", *Geomech. Eng*., **27**(5), 511-525. https://doi.org/10.12989/gae.2021.27.5.511.

Friedman, J.H. (1991), "Multivariate adaptive regression splines", *Ann. Stat*., **19**(1), 1-67. https://doi.org/10.1214/aos/1176347963.

Garson, G.D. (1991), "Interpreting neural network connection weights", *Artif. Intell. Exp.*, **6**, 47-51. https://doi.org/10.5555/129449.129452.

Gevrey, M., Dimopoulos, I. and Lek, S. (2003), "Review and comparison of methods to study the contribution of variables in artificial neural network models", *Ecol. Modell*., **160**(3), 249-264. https://doi.org/10.1016/S0304-3800(02)00257-0.

Guo, J.C., Luo, B., Zhu, H.Y., Wang, Y.H., Lu, Q.L. and Zhao, X. (2015), "Evaluation of fracability and screening of perforation interval for tight sandstone gas reservoir in western Sichuan Basin", *J. Nat. Gas Sci. Eng*., **25**, 77-87. https://doi.org/10.1016/j.jngse.2015.04.026.

Hajiabdolmajid, V. and Kaiser, P. (2003), "Brittleness of rock and stability assessment in hard rock tunnelling", *Tunn. Undergr. Sp. Technol*., **18**(1), 35-48. https://10.1016/S0886-7798(02)00100-1.

Huang, X.R., Huang, J.P., Li, Z.C., Yang, Q.Y., Sun, Q.X. and Wei, C. (2015), "Brittleness index and seismic rock physics model for anisotropic tight-oil sandstone reservoirs", *Appl. Geophys*., **12**(1), 11-22. https://10.1007/s11770-014-0478-0.

Hucka, V. and Das, B. (1974), "Brittleness determination of rocks by different methods". International *Int. J. Rock Mech. Min. Sci. Geomech. Abstr*., **11**, 389-392. https://doi.org/10.1016/0148-9062(74)91109-7.

Hussain, A., Surendar, A., Clementking, A., Kanagarajan, S. and Ilyashenko, L.K. (2018), "Rock brittleness prediction through two optimization algorithms namely particle swarm optimization and imperialism competitive algorithm". *Eng. Comput*., **35**, 1027-1035. https://doi.org/10.1007/s00366-018-0648-9.

Kaunda, R.B. and Asbury, B. (2016), "Prediction of rock brittleness using nondestructive methods for hard rock tunnelling", *J. Rock Mech. Geotech. Eng*., **8**(4), 533-540, http://dx.doi.org/10.1016/j.jrmge.2016.03.002.

Koopialipoor, M., Noorbakhsh, A., Ghaleini, E.N., Armaghani, D.J. and Yagiz, S. (2019), "A new approach for estimation of rock brittleness based on non-destructive tests", *J. Nondestruct. Eval*., **34**(4), 354-375. https://doi.org/10.1080/10589759.2019.1623214.

Lawal, A.I. (2020), "An artificial neural network-based mathematical model for the prediction of blast-induced ground vibration in granite quarries in Ibadan, Oyo State, Nigeria", *Sci. African*, **8**, e00413. https://doi.org/10.1016/j.sciaf.2020.e00413.

Lawal, A.I. and Idris, M.A. (2019), "An artificial neural network-based mathematical model for the prediction of blast-induced ground vibrations". *Int. J. Environ Std.*, **77**(2), 318-334. https://doi.org/10.1080/00207233.2019.1662186.

Lawal, A.I. and Kwon, S. (2020), "Application of artificial intelligence in rock mechanics: an overview", *J. Rock Mech. Geotech. Eng.,* **13**, 248-266. https://doi.org/10.1016/j.jrmge.2020.05.010.

Lawal, A.I., Kwon, S., Aladejare, A.E. and Oniyide, G.O. (2022), "Prediction of the static and dynamic mechanical properties of sedimentary rock using GPR, ANN, and response surface method", *Geomech. Eng*., **28**(3), 4547-4563. https://doi.org/10.12989/gae.2022.28.3.313.

Lawal, A.I., Oniyide, G.O., Kwon, S., Onifade, M., Köken, E, and Ogunsola, N.O. (2021a), "Prediction of mechanical properties of coal from non-destructive properties: A comparative application of MARS, ANN, and GA", *Nat. Resour. Res.*, **30**(6), 4547-4563.

Lawal, A.I., Kwon, S., Hammed, O.S. and Idris, M.A. (2021b), "Blast-induced ground vibration prediction in granite quarries: An application of Gene expression programming, ANFIS, and Sine Cosine algorithm optimized ANN", *Int. J. Min. Sci. Tech.,* **31**, 265-277.

Lawn, B.R., Jensen, T. and Arora, A. (1976), "Brittleness as an indentation size effect", *J. Mat. Sci*., **11**(3), 573-575. https://doi:10.1007/BF00540940.

Leathwick, J.R., Rowe, D., Richardson, J., Elith, J. and Hastie, T. (2005), "Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish", *Fresh W Biol*., **50**, 2034-2051. https://doi:10.1111/j.1365-2427.2005.01448.x

Meng, F.Z., Zhou, H., Zhang, C.Q., Xu, R.C. and Lu, J.J. (2015), "Evaluation methodology of brittleness of rock based on post-peak stress-strain curves", *Rock Mech. Rock Eng*., **48**(5), 1787-1805. https://doi:10.1007/s00603-014-0694-6.

Quinlan, J.R. (1992), "Learning with continuous classes", *Adams S (ed) Proceedings of AI'92. World Scientific*, Singapore.

Rickman, R., Mullen, M.J., Petre, J.E., Grieser, W.V. and Kundert, D. (2008), "A practical use of shale petrophysics for stimulation design optimization: All shale plays are not clones of the Barnett Shale", *SPE 115258 Proceeding of Annual Technical Conference, Society of Petroleum Engineers*, Denver, CO, USA.

Rybacki, E., Meier, T. and Dresen, G. (2016), "What controls the mechanical properties of shale rocks? – Part II: Brittleness", *J. Pet. Sci. Eng*., **144**, 39-58. https://doi:10.1016/j.petrol.2016.02.022.

Sihag, P., Karimi, S.M. and Angelaki, A. (2019), "Random forest, M5P and regression analysis to estimate the field unsaturated hydraulic conductivity". *App. W. Sci.*, **9**, 129. https://doi.org/10.1007/s13201-019-1007-8.

Sun, D., Lonbani, M., Askarian, B., Armaghani, D.J., Tarinejad, R., Pham, B.T. and Huynh, V.V. (2020). "Investigating the applications of machine learning techniques to predict the rock brittleness index", *Appl. Sci*., **10**, 1691. doi:10.3390/app10051691

Sun, H., Du, W.S. and Chi, L. (2021), "Uniaxial compressive strength determination of rocks using X-ray computed tomography and convolutional neural networks", *Rock Mech. Rock Eng*., **54**, 4225-4237. https://doi.org/10.1007/s00603-021-02503-1.

Tarasov, B. and Potvin, Y. (2013), "Universal criteria for rock brittleness estimation under triaxial compression", *Int. J. Rock Mech. Min. Sci*., **59**, 57-69. https://doi.org/10.1016/j.ijrmms.2012.12.011.

Wang, H., Cai, R., Zhou, B., Aziz, S., Qin, B., Voropai, N., Gan, L. and Barakhtenko, E. (2020), "Solar irradiance forecasting based on direct explainable neural network", *Energ. Convers. Manage*, **226**, 113487, https://doi.org/10.1016/j.enconman.2020.113487.

Xia, Y., Zhou, H., Zhang, C., He, S., Gao, Y. and Wang, P. (2019), "The evaluation of rock brittleness and its application: a review study", *Eur. J. Environ. Civ*., **22**(1), 239-279. https://doi.org/10.1080/19648189.2019.1655485.

Yagiz, S. (2009), "Assessment of brittleness using rock strength and density with punch penetration test", *Tunn. Undergr. Sp. Technol*., **24**(1), 66-74. https://doi:10.1016/j.tust.2008.04.002.

Yagiz, S. and Gokceoglu, C. (2010), "Application of fuzzy inference system and nonlinear regression models for predicting rock brittleness", *Exp. Syst. Appl*., **37**(3), 2265-2272. https://doi.org/10.1016/j.eswa.2009.07.046.

Yagiz, S., Ghasemi, E. and Adoko, A.C. (2018), "Prediction of rock brittleness using genetic algorithm and particle swarm optimization techniques", *Geotech. Geol. Eng*., **36**, 3767-3777,

https://doi.org/10.1007/s10706-018-0570-3.
Yuvaraj, P., Murthy, A.R, Iyer, N.R, Samui, P. and Sekar, S.K. (2013), "Multivariate adaptive regression splines model to predict fracture characteristics of high strength and ultra high strength concrete beams", *Tech Sci. Press, CMC*, **36**(1), 73-97.

CC