Natural Resources Research, Vol. 30, No. 6, December 2021 (© 2021)
https://doi.org/10.1007/s11053-021-09955-w

Original Paper

Prediction of Mechanical Properties of Coal from Non-destructive Properties: A Comparative Application of MARS, ANN, and GA

Abiodun Ismail Lawal,[1,2] Gafar O. Oniyide,[2] Sangki Kwon,[1,6] Moshood Onifade,[3] Ekin Köken,[4] and Nafiu O. Ogunsola[5]

Received 24 June 2021; accepted 12 September 2021
Published online: 6 October 2021

Rock properties are useful for safe operation and design of both surface and underground mines including civil engineering projects. However, the cost and time required to perform detailed assessments of rock properties are high. In addition, rock properties are required in numerical modeling. Different models have been proposed for quick and easy assessments of rock properties but majority of these models are regression-based, which are incapable of capturing inherent variabilities in rock properties. Therefore, this study proposed three different soft computing models (i.e., double input–single output ANN, multivariate adaptive regression spline, genetic algorithm) for accurate prediction of several mechanical properties of coal and coal-like rocks. The performances of the proposed models were statistically evaluated using various indices and they were found to predict rock properties suitably with very strong statistical indices. The proposed models were validated further using external datasets aside from those used in the model development to test the generalization potential of the models. The Pearson's correlation coefficients for the validation were close to 1, indicating that the proposed models can be used to assess geo-mechanical properties of coal, shale, and coal-bearing rocks.

KEY WORDS: Coal, Rock properties, MARS, Soft computing, Statistical indices.

INTRODUCTION

Rock properties are important in both surface and underground mine designs. They are used in many numerical software for rock stability analysis or mine designs. These rock properties could be either physical (e.g., density, p-wave) or mechanical (e.g., uniaxial compressive strength (UCS), tensile strength (TS), Young modulus (YM)). Measuring physical properties of rocks are generally considered less tedious compared to measuring mechanical properties, which is time consuming, tedious and expensive. During the preliminary stages of surface or underground mine and as the work progresses, a quick estimation of these properties is often re-

[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.
[3]Department of Civil and Mining Engineering, University of Nambia, P.O. Box 3624 Ongwediva, Namibia.
[4]Nanotechnology Engineering Department, Abdullah Gul University, 38100 Kayseri, Turkey.
[5]Department of Mineral Resources and Energy Engineering, Jeonbuk National University, Jeonju-si, Jeollabuk-do, South Korea.
[6]To whom correspondence should be addressed; e-mail: kwonsk@inha.ac.kr, ailawal@futa.edu.ng

quired. However, the constraint of getting enough core samples, cost and time implications has made geotechnical/rock engineers to depend mainly on existing empirical models. Some of the empirical models used to predict UCS and YM are presented in Table 1. More details of such empirical models in predicting the UCS of rocks are summarized in the study by Aladejare et al. (2021). Although the coefficients of determination (R^2) of these models may be satisfactory, they usually fail when subjected to datasets obtained from similar rock types.

Empirical models, like those in Table 1, can be categorized based on the number of parameters involved in the model developments into either one parameter or multiple parameters. Both categories are based on linear or multiple regression analyses. Simple regression analyses are unable to capture nonlinear relationships between model parameters. However, rocks are highly heterogeneous, and their relationships are commonly nonlinear. The quest for more accurate means of estimating rock properties has made researchers utilize various soft computing (SC) methods. For instance, Dehghan et al. (2010) proposed an artificial neural network (ANN) model to predict UCS and YM using some non-destructive rock properties and point load strength (PL) index with 30 datasets. Manouchehrian et al. (2012) also used ANN to predict UCS based on void ratio and mineralogical composition with 44 datasets. Mohamad et al. (2018) employed particle swarm optimization (PSO)—ANN based on combined destructive and non-destructive rock properties with 38 datasets. Singh et al. (2017) used an adaptive neuro-fuzzy inference system (ANFIS) to predict UCS, TS, PL and YM with 90 datasets using density, porosity and p-wave velocity (P_{WV}) as the input parameters. Roy and Singh (2019) proposed ANN and ANFIS models to assess the deformation properties of coal using selected mechanical properties, such as UCS, TS, shear strength (SS) and P_{WV}.

In addition, Sun et al. (2021) developed ensemble learning models to predict the strength of coal-grout materials. They had earlier investigated the strengths of grouting concretes, cemented paste backfill, reinforced concrete beams and YM of jet grouted coalcretes using beetle antennae optimized neural network and intelligent models (Sun et al., 2019a, 2019b, 2020; Zhang et al., 2020). Studies by Jalali et al. (2017), Sharma et al. (2017), Aboutaleb et al. (2018) and Ren et al. (2019) also predicted UCS of sedimentary rocks using SC methods. More detailed of SC techniques used to predict rock properties are documented by Lawal and Kwon (2020).

Although several models are available to predict rock properties, most commonly UCS, the easy-to-use empirical models are not as reliable as established in the literature (Singh et al., 2017). In addition, accurate SC models cannot be easily used in geotechnical/rock engineering designs because they are not in the form of simple mathematical models (Lawal & Kwon, 2020) that can be used in practice. Hence, there is still heavy reliance on empirical models for quick estimation of rock properties. It is also important to mention that it is extremely difficult to get core samples of some weak rocks like coal because their physico-mechanical properties are significant in their production planning and safety. Therefore, this type of rocks requires accurate models for estimating their mechanical properties from simple non-destructive physical properties.

In this study, the MARS (multivariate adaptive regression spline) and double input–single output ANN ((DISO–ANN) are proposed for the prediction of UCS, TS, SS, and YM of coal. The proposed MARS model has never been used for the prediction of several properties of coal to the best of the knowledge of the authors. In addition, the genetic algorithm (GA) model has not been widely used in predicting several coal properties. The proposed MARS and ANN models are advantageous over many other SC methods (such as PSO) because they do not require prior expression of models in the form of mathematical equations (Lawal et al., 2021a). Therefore, the novelty of this study relies on the fact that the adopted models have not been used previously in predicting several properties of coal. In fact, the ANN that has been used previously to predict UCS is largely black box in nature. Therefore, the proposed models are believed to be useful in various surface or underground coal designs or as a means of estimating the coal strength parameters.

MATERIALS AND METHOD

Statistical Description of the Datasets

The physico-mechanical properties of coal that are crucial in the design and safety of underground and surface mines are investigated using various SC methods. This is necessary because the time and cost required for detailed assessments of the rock prop-

Table 1. Some empirical models for rock properties prediction

Models	Rock type	R^2	Authors
YM = $4.9718P_{WV} - 7151$, MPa	Coal	0.974	Khandelwal and Singh (2009)
YM = $(P_{WV} - 1.7528)/0.0937$, GPa	Carbonate	0.86	Yasar and Erdogan (2004)
YM = $2.06P_{WV}^{2.78}$, GPa	Limestone, sandstone, and marble	0.92	Moradian and Behnia, (2009)
YM = $0.0114P_{WV} + 3.7059$, GPa	Limestone	0.76	Kurtulus et al. (2016)
YM = $0.076P_{WV}^{3.23}$, GPa	Shale	0.99	Horsrud (2001)
UCS = $0.0375P_{WV} - 50.969$, MPa	Grainstone, Wackestone-Mudstone, Boundstone, Gypsum, and Silty Marl	0.67	Heidari et al. (2017)
UCS = $0.026P_{WV} - 20.47$, MPa	Marl	0.91	Abdolazim and Rassoul (2015)
UCS = $16S_{WV}^{1.6}$, MPa	Limestone, Sandstone, and Siltstone	0.82	Uyanik et al. (2019)
UCS = $10.4BTS + 18.2$, MPa	Flint	0.63	Aliyu et al. (2019)
UCS = $17.6Is_{(50)} + 13.5$, MPa	Flint	0.88	Aliyu et al. (2019)

erties are very tedious, and geo-engineers often require quick assessments of rock properties. Hence, they rely largely on empirical models. The performances of most of the empirical models are not too reliable to estimate properties and majority of the available empirical models is not specific to coal. Among these properties, some are easy to perform like the P_{WV} and density, which require non-destructive tests, while mechanical properties like UCS, TS, SS, and YM are tedious to assess and yet they are crucial in the design of mines and in numerical modeling as material properties. Therefore, this study has proposed different SC models that are mathematically based for the computation of four mechanical properties of rocks from the simple tests based on the existing data in the literature.

The fragility of coal makes coring it so demanding and the hard ones reported in the literature are often used for models. For instance, Roy and Singh (2019) predicted the deformation properties of coal by combining some coal properties using the datasets obtained from Khandelwal and Singh (2009). The data used for this study were those available in Verma et al. (2012) from three coal mines in India (i.e., Neeljay, Ghugus and Sasti surface mines), which are subsidiaries of western coal field limited (WCL). Experiments were conducted on 15 core samples of coal obtained in accordance with the International Society for Rock Mechanics (ISRM, 1981) standard for each location. The PUNDIT was used for the P_{WV} test, which was conducted in accordance with ISRM (1978) procedure. Other tests, such as Brazilian tensile strength and triaxial test, were also conducted based on relevant test standards. The descriptive statistics of the 45 coal samples from the three mines are presented in Figure 1. The mean values of P_{WV}, ρ, UCS, TS, SS, and YM are 1903.868, 1.577, 22.398, 2.338, 4.542, and 2305.724, respectively, while their standard deviations are 71.702, 0.179, 5.59, 0.433, 0.911, and 69.347, respectively. Because the standard deviations of density and tensile strength are close to zero, there are no wide variations in these two properties for the samples from the three mines. The variation in tensile strength is not also wide. However, P_{WV} and YM of the rock samples deviate strongly from their respective means (Fig. 1). Because there are variations in the properties exhibited by rock samples, simple linear or multiple regressions may be unable to capture the relationships among the variables (Gevrey et al., 2003; Roy & Singh, 2019). Hence, models that are capable of capturing the variability in rock properties are proposed in this study. The explanation and procedures of each of the three proposed SC-based models are presented in the next section.

Artificial Neural Network (ANN)

ANN is a SC method that emulates artificially the functionality of human neurons. It consists of a series of interconnectivity just like neurons in a human brain. The technique has been applied in different fields of sciences and engineering for solving different complicated problems that are usually impossible to solve using various conventional methods. This is possible in ANN because it can solve problems in which no visible relationships exist between variables (Lawal & Kwon, 2020). However, the method has been criticized to be a black-box approach as relationships between the predicted variable(s) and predictors cannot be deciphered

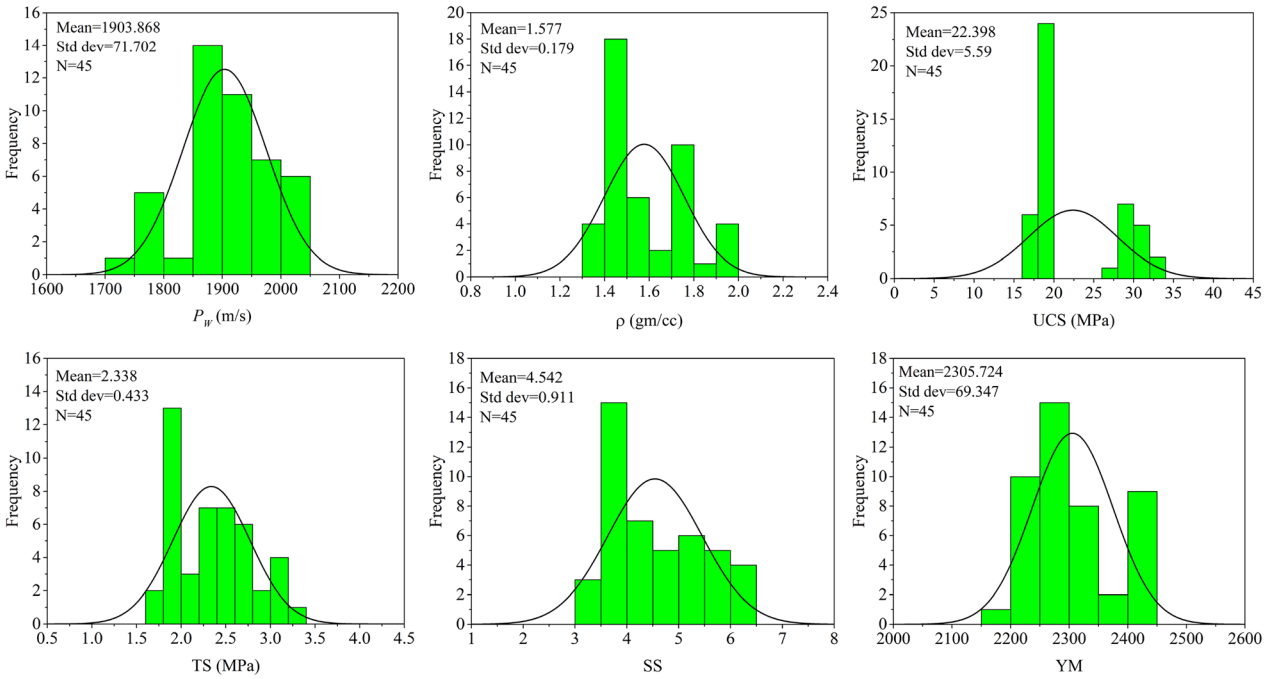

Figure 1. Statistical descriptions of the model parameters.

easily. Yet, it remains one of the best SC tools and it has gained popularity among researchers including in the mining/geotechnical engineering fields. Therefore, ANN was adopted in this study for predicting mechanical properties of coal from easy-to-do/perform non-destructive tests with the aim of eliminating the black-box nature of the ANN and to make it a white-box ANN. This is achieved in two ways, namely by establishing the mathematical relationships between the predicted variable(s) and predictors and by the weighted average method for formulating the influence of the model predictors on the predicted variable.

The ANN method proposed here was implemented in MATLAB environment. The datasets were first pre-processed by normalizing them within the range from − 1 to 1.[1] This is paramount in order to eliminate bias among the model variables and to avoid over-fitting of the network (Lawal & Idris, 2019). In addition, it should be noted that all the model parameters used in all the developed models are their normalized values. The pre-processed data were then imported into MATLAB. The datasets were then partitioned randomly into training, testing and validation datasets. Then, the number of neurons in each layer, most importantly the neurons in the hidden layer, was set because the number of neurons in the input and out layers are dictated by the number of input and output variables; however, the number of neurons in the hidden layer is usually based on trial-and-error. Therefore, different numbers of neurons ranging between 2 and 4 were tried in this study and only a 3-layer ANN structure was used throughout. Then, activation functions were specified for the hidden and output layers. The hyperbolic tangent (tanh) was used in the hidden layer and the purelin was used in the output layer throughout. Thereafter, the training was conducted using feedforward backpropagation algorithm with Levenberg-Marquardt training function. It is also important to state that DISO–ANN structure was used. The output of various number of hidden layers tried for each properties and the selected optimum ANN structures for each coal properties are presented in Table 2 and Figure 2. From Table 2 and Figure 2, the 2–4–1 DISO–ANN was the best for the UCS or TS and 2–3–1 was the best for the SS or YM. The overall performances of the selected DISO–ANN structure for each of the rock properties are

[1] $\Omega_{norm} = 2(\Omega - \Omega_{min})/ (\Omega_{max} - \Omega_{min}) - 1$, where Ω is the actual variable to be normalized while Ω_{max} and Ω_{min} are its maximum and minimum values and Ω_{norm} is its normalized form.

presented in Figure 3. The selected structures for each of the coal properties showed remarkable performances.

As a way of unlocking the black-box nature of the ANN structure as previously stated, the mathematical relationships were established to enhance easy applicability of the proposed model. This was achieved by extracting the weights and biases from the trained DISO–ANN structure and then combining them with the transfer functions. For the cases of UCS and TS, there were four neurons in the hidden layer, and the obtained relationships were:

$$\mathrm{UCS} = 7.505 \sum_{i=1}^{4} x_i + 341.9596 \tag{1}$$

$$\mathrm{TS} = 0.735 \sum_{i=1}^{4} y_i + 22.2679 \tag{2}$$

where the unknown values of x_i and y_i in Eqs. 1 and 2 are presented, respectively, as:

$$\begin{cases} x_1 = 4.4872 \tanh(-3.7878 P_{\mathrm{WV}} + 4.9306\rho + 2.2509) \\ x_2 = 5.13798 \tanh(4.0925 P_{\mathrm{WV}} - 5.1749\rho - 2.3202) \\ x_3 = -0.6315 \tanh(15.6855 P_{\mathrm{WV}} - 41.2701\rho - 30.7462) \\ x_4 = -43.1699 \tanh(-0.7168 P_{\mathrm{WV}} - 0.0773\rho + 3.5498) \end{cases} \tag{3}$$

$$\begin{cases} y_1 = 27.7294 \tanh(0.35012 P_{\mathrm{WV}} + 0.1898\rho - 2.26805) \\ y_2 = 0.28935 \tanh(13.8504 P_{\mathrm{WV}} - 5.0728\rho - 1.3884) \\ y_3 = -0.4789 \tanh(3.7115 P_{\mathrm{WV}} - 18.8725\rho - 14.2089) \\ y_4 = -0.2955 \tanh(-3.6214 P_{\mathrm{WV}} + 105.3787\rho + 92.3417) \end{cases} \tag{4}$$

Similarly, the obtained mathematical transformation of the 2–3–1 DISO–ANN structures for the SS and YM predictions were:

$$\mathrm{SS} = 1.59 \sum_{i=1}^{3} z_i + 12.3548 \tag{5}$$

$$\mathrm{YM} = 122.47 \sum_{i=1}^{3} w_i + 2327.6325 \tag{6}$$

where the unknown z_i and w_i in Eqs. 5 and 6 are presented, respectively, as:

$$\begin{cases} z_1 = -8.9795 \tanh(-0.16504 P_{\mathrm{WV}} + 0.0549\rho + 0.6801) \\ z_2 = 0.9653 \tanh(-3.4289 P_{\mathrm{WV}} + 15.2137\rho + 12.18408) \\ z_3 = 0.49303 \tanh(2.6874 P_{\mathrm{WV}} - 19.3276\rho - 16.2618) \end{cases} \tag{7}$$

$$\begin{cases} w_1 = 0.1411 \tanh(16.0274 P_{\mathrm{WV}} - 1.7596\rho - 15.7498) \\ w_2 = 0.6646 \tanh(-33.4303 P_{\mathrm{WV}} + 24.7252\rho - 3.3203) \\ w_3 = -1.6886 \tanh(-0.3103 P_{\mathrm{WV}} - 0.0385\rho - 0.06501) \end{cases} \tag{8}$$

Multivariate Adaptive Regression Spline (MARS)

The MARS proposed by Friedman (1991) is a nonparametric regression method, which can be perceived as a hybrid linear model that addresses nonlinearity and variables' interactions automatically. It helps in obtaining a better fit than linear models by generating the kink using hinge functions. The general form of the MARS model is the addition of the sum of weighted basis functions, thus:

$$\Upsilon = \sum_{i=1}^{k} c_i B_i(x) \tag{9}$$

where c_i is the constant coefficient and $B_i(x)$ is the basis function that can take different forms, namely a constant 1 at the intercept, a hinge function that is of the form max (u, x-u) or max (x-u, u) (where u is the knots) and a product of two or more hinge function to define interaction between multiple model variables. There are two important processes in the MARS model, namely forward pass and backward pass. In the forward pass, the MARS model is initiated with the constant term and a pair of basis function is added subsequently. When a pair of basis function is added, the maximum error reduction is evaluated and a pair of basis function with the lowest error is selected. However, in the forward pass, the performance of training datasets is usually good but when tested/validated, they give poor performance (i.e., the forward pass over-fits the model). This necessitates the backward pass, in which pruning of ineffective terms takes place. While the forward pass adds terms in pairs, the backward pass discards one part of the pair and typically the product will be seen with single term. The backward pass does its function using the generalized cross-validation (GCV) to compare the

Table 2. Different trials of DISO–ANN structures

DISO–ANN structures		Training	Testing	Validation	Overall
		UCS			
2–1-1	R	0.7487	0.9888	0.8067	0.7578
2–2-1		0.9942	0.9994	0.9566	0.9955
2–3–1		0.9991	0.9995	0.9834	0.9992
2–4–1		**0.9998**	**0.9999**	**0.9989**	**0.9998**
		TS			
2–1-1	R	0.8809	0.7077	0.9394	0.8617
2–2-1		0.9876	0.9934	0.9962	0.9898
2–3–1		0.9928	0.9817	0.8674	0.9809
2–4–1		**0.9963**	**0.9982**	**0.9955**	**0.9970**
		SS			
2–1-1	R	0.7068	0.9303	0.7424	0.7497
2–2-1		0.9901	0.9980	0.9826	0.9912
2–3–1		**0.9982**	**0.9962**	**0.98068**	**0.9975**
2–4–1		0.9964	0.9979	0.9988	0.9973
		YM			
2–1-1	R	0.9501	0.9797	0.8878	0.9402
2–2-1		0.998	0.99929	0.99335	0.9974
2–3–1		**0.9991**	**0.9998**	**0.9992**	**0.9991**
2–4–1		0.9988	0.9997	0.9999	0.9992

Rows in bold represent the DISO–ANN structures selected for each of the investigated properties

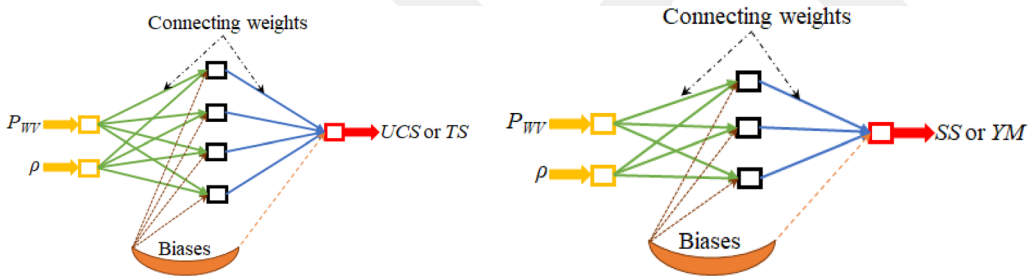

Figure 2. Selected DISO–ANN structures.

performance of the model subsets for it to pick the best subset. The formula for the GCV is:

$$\mathrm{GCV} = \mathrm{RSS}/\left(N(1 - \mathrm{ENP}/N)^2\right) \qquad (10)$$

where RSS stands for the measured residual sum of squares, ENP stands for the effective number of parameters, and N stands for the number of observations.

Generally, the MARS model has also gained popularity among researchers and it has been used to solve various engineering problems. It was also adopted in this study to predict the most relevant mechanical properties of coal. The proposed MARS model was implemented in MATLAB environment using the same number of datasets as in the ANN model. That is, the number of datasets used from the ANN training, testing and validation were also used in the MARS model. The datasets were imported to MATLAB and the model parameters were first defined using *aresparams* command and the maximum function and the degree of interaction were defined at this point. The piecewise-linear method was adopted. Afterward, the model was built using the *aresbuild* command. Then, after the model was trained, the results obtained with the MARS information, which comprises the number of basis function selected by the model, their corresponding

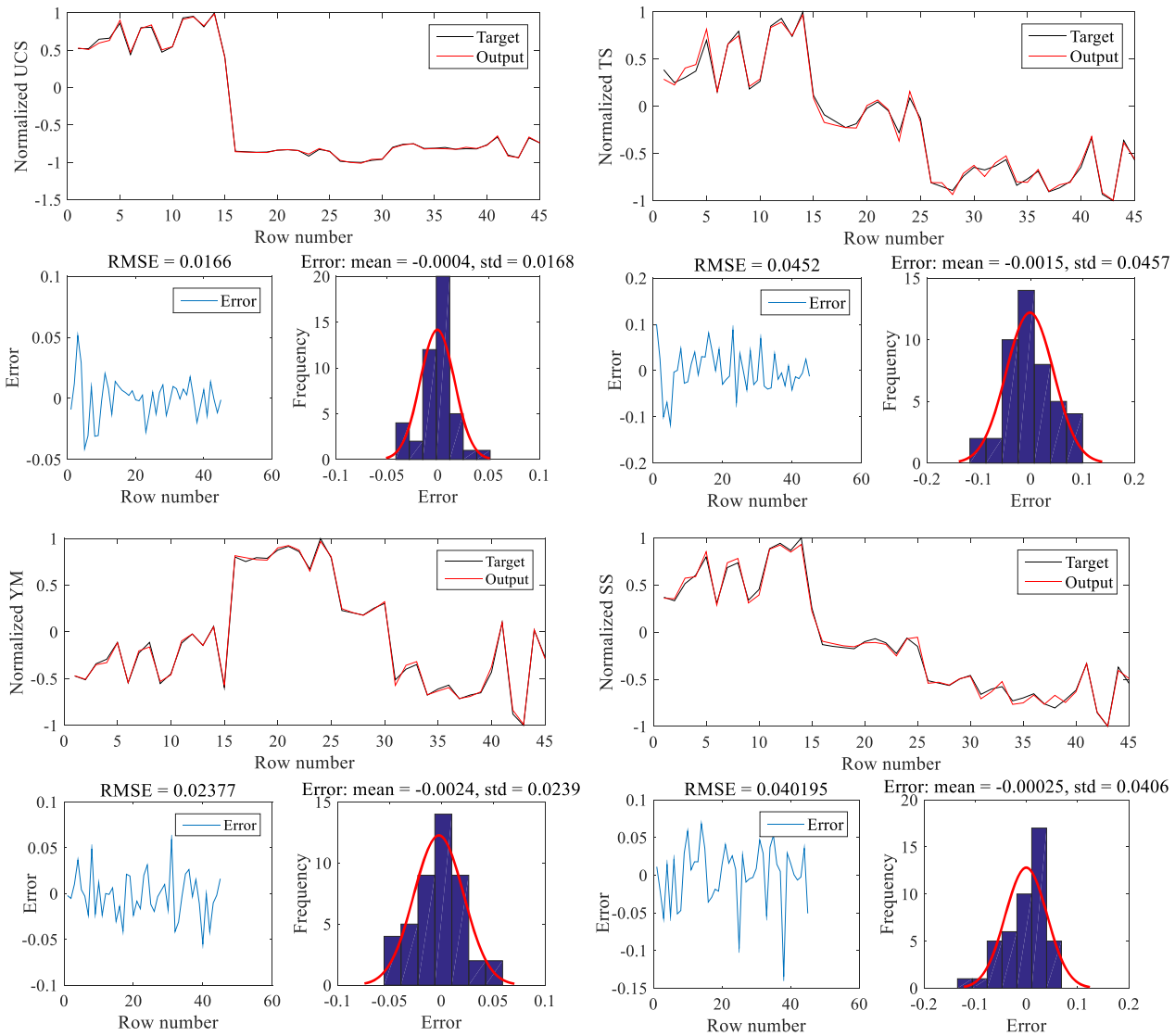

Figure 3. Overall performances of the selected DISO–ANN structures.

equations (Table 3) and their respective coefficients, were extracted together with the following final model equations:

$$\begin{aligned} UCS_{MARS} = {} & 2.6091 - 0.75068BF1 + 16.966BF2 \\ & - 14.651BF3 - 4.5733BF4 + 3.7838BF5 + \ldots \\ & 2.4975BF6 + 57.966BF7 + 55.961BF8 \\ & - 1.6973BF9 + 13.169BF10 + 7.5418BF11 - \ldots \\ & 6.8792BF12 - 71.955BF13 - 11.521BF14 \\ & + 9.8996BF15 - 63.433BF16 - 2.5517BF17 + \ldots \\ & 20.801BF18 + 8.4267BF19 \end{aligned} \tag{11}$$

$$\begin{aligned} TS_{MARS} = {} & -3.0892 - 13.25BF1 + 145.26BF2 \\ & - 188.04BF3 + 48.258BF4 - 79.647BF5 - \ldots \\ & 76.828BF6 + 73.456BF7 + 11.122BF8 \\ & - 11.821BF9 + 19.667BF10 - 159.27BF11 + \ldots \\ & 88.745BF12 - 13.744BF13 - 166.97BF14 \\ & + 198.84BF15 - 287.39BF16 + 15.222BF17 - \ldots \\ & 3.7894BF18 - 14.153BF19 + 210.32BF20 \end{aligned} \tag{12}$$

Table 3. Basis functions for each property

BF	BF equations for the corresponding rock properties	
	UCS	*TS*
BF1	$\Gamma(0, x_2 + 0.44262)$	$\Gamma(0, -0.44262 - x_2)$
BF2	BF1 * $\Gamma(0, x_1 - 0.25556)$	$\Gamma(0, x_2 + 0.44262)$ * $\Gamma(0, x_1 - 0.25556)$
BF3	BF1 * $\Gamma(0, 0.25556 - x_1)$	$\Gamma(0, x_2 + 0.44262)$ * $\Gamma(0, 0.25556 - x_1)$
BF4	$\Gamma(0, x_1 - 0.14348)$	$\Gamma(0, x_1 - 0.14348)$
BF5	$\Gamma(0, 0.14348 - x_1)$	$\Gamma(0, 0.14348 - x_1)$
BF6	$\Gamma(0, x_2 + 0.63934)$	BF1 * $\Gamma(0, x_1 - 0.4792)$
BF7	$\Gamma(0, -0.63934 - x_2)$ * $\Gamma(0, x_1 - 0.10687)$	BF1 * $\Gamma(0, 0.4792 - x_1)$
BF8	$\Gamma(0, -0.63934 - x_2)$ * $\Gamma(0, 0.10687 - x_1)$	$\Gamma(0, x_2 + 0.63934)$
BF9	$\Gamma(0, x_2 - 0.081967)$	$\Gamma(0, -0.63934 - x_2)$
BF10	$\Gamma(0, 0.081967 - x_2)$ * $\Gamma(0, x_1 - 0.4792)$	BF4 * $\Gamma(0, x_2 + 0.57377)$
BF11	$\Gamma(0, 0.081967 - x_2)$ * $\Gamma(0, 0.4792 - x_1)$	BF4 * Γ $(0, -0.57377 - x_2)$
BF12	BF4 * $\Gamma(0, x_2 + 0.57377)$	BF9 * $\Gamma(0, x_1 - 0.10687)$
BF13	BF4 * $\Gamma(0, -0.57377 - x_2)$	$\Gamma(0, x_2 + 0.7377)$
BF14	$\Gamma(0, 0.4792 - x_1)$	BF13 * $\Gamma(0, x_1 - 0.25556)$
BF15	BF14 * $\Gamma(0, x_2 + 0.67213)$	BF13 * $\Gamma(0, 0.25556 - \times 1)$
BF16	BF14 * $\Gamma(0, -0.67213 - x_2)$	BF5 * $\Gamma(0, -0.67213 - x_2)$
BF17	$\Gamma(0, x_2 + 0.7377)$	$\Gamma(0, 0.069212 - x_1)$
BF18	$\Gamma(0, -0.7377 - x_2)$	$\Gamma(0, x_1 - 0.069212)$ * $\Gamma(0, 0.081967 - x_2)$
BF19	BF17 * $\Gamma(0, 0.10687 - x_1)$	$\Gamma(0, x_1 - 0.069212)$ * $\Gamma(0, -0.40984 - x_2)$
	SS	*YM*
BF1	$\Gamma(0, x_2 + 0.27869)$	$\Gamma(0, -0.27869 - x_2)$
BF2	BF1 * $\Gamma(0, 0.091538 - x_1)$	BF1 * $\Gamma(0, x_1 - 0.15837)$
BF3	$\Gamma(0, -0.27869 - x_2)$ * $\Gamma(0, x_1 - 0.58971)$	BF1 * $\Gamma(0, 0.15837 - x_1)$
BF4	$\Gamma(0, -0.27869 - x_2)$ * $\Gamma(0, 0.58971 - x_1)$	$\Gamma(0, x_1 - 0.25556)$
BF5	$\Gamma(0, x_2 + 0.63934)$	$\Gamma(0, -0.80328 - x_2)$
BF6	$\Gamma(0, x_1 - 0.25556)$	$\Gamma(0, x_1 - 0.1059)$
BF7	$\Gamma(0, 0.25556 - x_1)$	$\Gamma(0, 0.1059 - x_1)$
BF8	$\Gamma(0, -0.63934 - x_2)$ * $\Gamma(0, x_1 - 0.58971)$	BF7 * $\Gamma(0, 0.081967 - x_2)$
BF9	$\Gamma(0, -0.63934 - x_2)$ * $\Gamma(0, 0.58971 - x_1)$	BF4 * $\Gamma(0, x_2 + 0.67213)$
BF10	$\Gamma(0, -0.80328 - x_2)$	BF4 * $\Gamma(0, -0.67213 - x_2)$
BF11	$\Gamma(0, x_2 + 0.80328)$ * $\Gamma(0, x_1 - 0.58971)$	$\Gamma(0, -0.0040932 - x_1)$
BF12	$\Gamma(0, x_2 - 0.081967)$	$\Gamma(0, x_1 + 0.0040932)$ * $\Gamma(0, x_2 + 0.80328)$
BF13	$\Gamma(0, 0.081967 - x_2)$	$\Gamma(0, x_1 + 0.0040932)$ * $\Gamma(0, -0.80328 - x_2)$

▫ x_1 and x_2 stand for P_{WV} and ρ, respectively, and Γ stands for maximum (max)

$$\begin{aligned} SS_{MARS} = {} & 4.3647 - 2.6835BF1 + 3.7504BF2 \\ & - 3.6614BF3 + 5.2786BF4 - 4.3539BF5 + ... \\ & 0.58805BF6 - 1.9778BF7 + 18.333BF8 \\ & - 3.4446BF9 + 2.085BF10 + 0.81903BF11 + ... \\ & 6.7507BF12 - 6.4767BF13 \end{aligned} \quad (13)$$

$$\begin{aligned} YM_{MARS} = {} & 0.47818 - 2.9819BF1 + 8.1552BF2 \\ & + 20.443BF3 + 4.4688BF4 - 11.587BF5 - ... \\ & 6.1601BF6 + 2.3383BF7 - 10.863BF8 \\ & - 1.8315BF9 + 24.439BF10 - 2.5998BF11 + ... \\ & 2.8594BF12 + 143.56BF13 \end{aligned} \quad (14)$$

Genetic Algorithm (GA)

GA is a population-based metaheuristic algorithm based on the principle of evolution theory. It was proposed by Holland (1975) to mimic the Darwinian evolutional theory of survival of the fittest. Descriptively, GA consists of various elements, such as chromosomes, fitness selection, inversion, and mutation (Holland, 1975). The chromosomes are known to take the form of binary string and are considered to be a position in the solution space. The chromosomes are processed using genetic operators by replacing its population iteratively. The fitness function assigns a value to the entire chromosomes in the population. The chromosomes are selected based on their fitness value for further processing. Random locus is chosen and it changes

the subsequence among the chromosomes to create off-springs by crossover operator while in mutation, bits of chromosomes are flipped randomly based on probability (Katoch et al., 2021; Lawal & Kwon, 2020; Michalewicz, 1992; Michalewicz & Schoenauer, 1996).

GA has gained popularity among researchers in various fields including engineering. Therefore, it was also proposed in this study to predict rock properties. The basic steps of the GA algorithm implementation are chronicled in the pseudocode below for the classic GA algorithm (Katoch et al., 2021).

Input
Population size, n
Maximum number of iterations, max
output*:*
Overall best solution, Y_{best}
start
Create initial population of n chromosomes Y_i *(i=1,2,...n)*
Define the iteration counter to zero (t=0)
Compute the fitness value of each chromosomes
while *(t<max)*
Pick a pair of chromosomes from initial population based on fitness
Use crossover operation on selected chromosomes with crossover probability
Use mutation on the offspring with mutation probability
Replace old population with newly generated population
Increase the current iteration t by 1
end while
Return the best solution, Y_{best}
End

The GA described above was implemented in this study using the MATLAB GA optimization tool with the same number of datasets used in the other proposed models to enable fair comparison. Just before employing the GA model, a nonlinear relationship between the model dependent and independent variables was formulated as:

$$Y = a_o + a_1 P_{WV} + a_2 \rho + a_3 P_{WV}\rho + a_4 P_{WV}^2 + a_5 \rho^2 \quad (15)$$

where a_0 to $a5$ are the unknown coefficients used to generate the objective function, which was linked to the GA optimization interface in the MATLAB. GA was then used to compute the unknown coefficients in the equation. The model parameters were then set as presented in Table 4.

Based on the parameters in Table 4, the model was initiated, and the results obtained when the stopping criterion, which in this case is either the maximum number of generation or the attainment of the error tolerance, is satisfied. The convergence curves showing the plots of the fitness against the number of generations for each of the parameters are presented in Figure 4 and the final coefficients obtained for each of the parameters after the convergence were:

$$UCS_{GA} = 23.9918 + 16.5674P_{WV} - 6.7896\rho + 21.8088P_{WV}\rho - 9.2688P_{WV}^2 - 10.8992\rho^2 \quad (16)$$

$$TS_{GA} = 2.5827 + 1.4257P_{WV} - 0.26743\rho + 1.9615P_{WV}\rho - 0.9464P_{WV}^2 - 0.7926\rho^2 \quad (17)$$

$$SS_{GA} = 4.8633 + 1.4023P_{WV} + 0.0535\rho + 1.8312P_{WV}\rho - 0.1812P_{WV}^2 - 1.5228\rho^2 \quad (18)$$

Table 4. Selected GA parameter settings

GA parameters	Selected
Population type	Double vector
Population size	50
Creation function	Constraint dependent
Fitness scaling function	Rank
Selection function	Stochastic uniform
Crossover function	Constraint dependent
Mutation function	Constraint dependent

$$\mathrm{YM_{GA}} = 2351.3354 - 179.3854P_{\mathrm{WV}} + 158.2887\rho - 194.9337P_{\mathrm{WV}}\rho + 140.8538P_{\mathrm{WV}}^2 + \ldots 7.3679\rho^2 \quad (19)$$

From Figure 4, the best fit and mean fit were not obvious for UCS, TS, and SS until the optimum generation was reached, but the case of YM is different. Hence, the GA model maybe more suitable for the YM than for the other geo-mechanical properties.

RESULTS AND DISCUSSION

Model Comparison

The proposed models were compared for each of the investigated geo-mechanical properties of coal. The obtained results for UCS using DISO–ANN, MARS and GA together with the ideal fit line are presented in Figure 5. The R^2 obtained using DISO–ANN was 0.9995, MARS 0.9918 and GA 0.8232. The R^2 is proportional to the closeness of the predicted values to the ideal fit line (Fig. 5a). Similarly, the measured and the predicted values at each data point are plotted in Figure 5b, and the GA prediction seemed to deviate largely from the measured values while the predictions of DISO–ANN and MARS were closer to the measured values.

Figure 6a depicts the plots of the predicted TS using the three adopted methods against the measured values. The plots of the model predictions and the measured values against the sample row numbers are also presented in Figure 6b. The R^2 values of the DISO–ANN, MARS, and GA models were 0.994, 0.966, and 0.776, respectively. Again, the predictions of the DISO–ANN were closer to the ideal fit line; hence, it had the highest R^2. This is also confirmed by Figure 6b as the predictions of the DISO–ANN were the closest to the measured values followed by the MARS models and then the GA model.

The predicted SS were also compared with the measured values and the sample row number. In Figure 7a, the R^2 values of the DISO–ANN, MARS, and GA were 0.995, 0.981, and 0.831, respectively. Comparing the models predictions to the measured values, the DISO–ANN model predictions were again closer to the measured SS values (Fig. 7b). The GA model seemed to deviate from the measured values and the ideal fit line most, and as a result it had the lowest R^2. The low performances recorded for the GA model predictions for UCS, TS, and SS can be traced to the divergence of the best fitness and the mean fitness in Figure 4.

Finally, the outcomes of the comparison of the three proposed models with the measured values of the YM are also presented in Figure 8. The R^2 values of the DISO–ANN, MARS and GA were 0.998, 0.949 and 0.951, respectively (Fig. 8a). All the three models also showed very close predictions to the measured values when plotted with the sample row number (Fig. 8b). However, the DISO–ANN and GA models performed better than the MARS model here. The good performance of the GA model can be traced to Figure 4 where the best fitness and mean fitness revealed similar pattern. In general, the DISO–ANN outperformed the other two models in all the predicted geo-mechanical properties.

Performance Assessments of the Models

The performances of the proposed DISO–ANN, MARS and GA models were assessed further using various statistical indicators, such as Pearson's *r*, root-mean-square error (RMSE), mean absolute percentage error (MAPE) and variance accounted for (VAF). The equations for computing the statistical indicators are common in the literature (Lawal, 2020; Lawal et al., 2021b, 2021c; Onifade et al., 2019) and are not repeated here. Each of these statistical indicators has some physical meanings or value ranges that explain how good model predictions are. The closer a model's Pearson's *r* value is to 1, VAF to 100%, RSME and MAPE to 0, the better the model is. Based on this, the DISO–ANN satisfied the physical explanation of the statistical indicators and turned out to be the best among the proposed models for predicting UCS, TS, SS, and YM as presented in Table 5. It was followed by the MARS

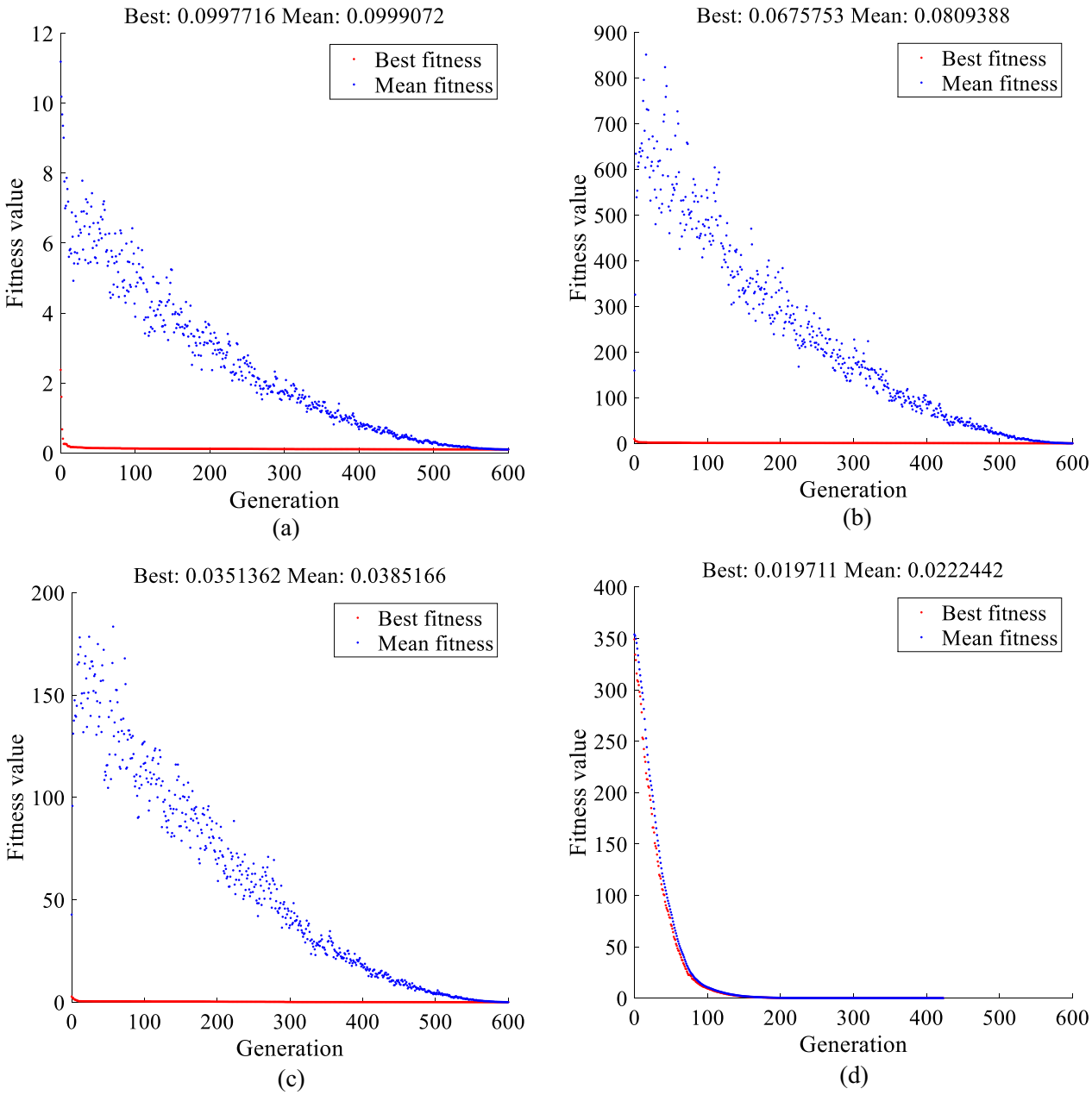

Figure 4. Convergence curve of the GA optimization algorithm for **a** UCS **b** TS **c** SS and **d** YM.

model and then the GA model. However, the performance of the GA model seemed to be good for the YM prediction.

In addition, the performances of the proposed models were assessed further using Taylor's diagram for easy identification of the best model, which has the highest Pearson's correlation coefficient and relatively the least standard deviation. For the four rock properties (i.e., UCS, TS, SS, and YM), the obtained Taylor's diagram is shown in Figure 9. The measured datasets have standard deviations of zero and Pearson's r of 1. For the rock properties studied, the DISO–ANN, annotated as ANN in Figure 9, yielded the best prediction performance, followed by the MARS and GA. However, it should be mentioned that the methods of DISO–ANN and MARS can be evaluated together because there is no a clear superiority over their prediction capability.

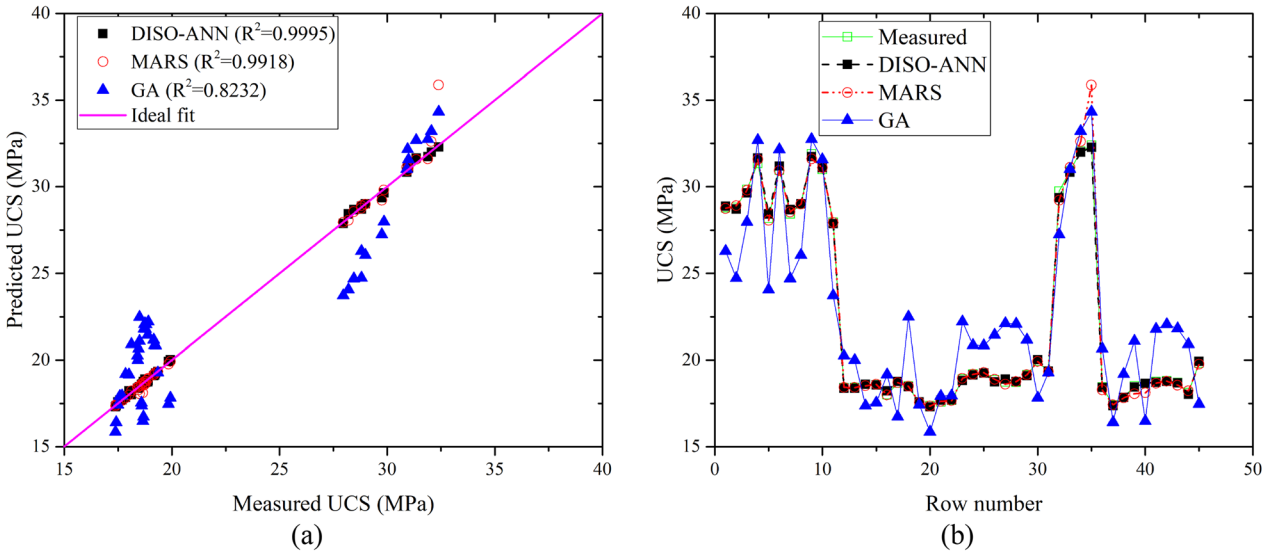

Figure 5. Model comparison for UCS.

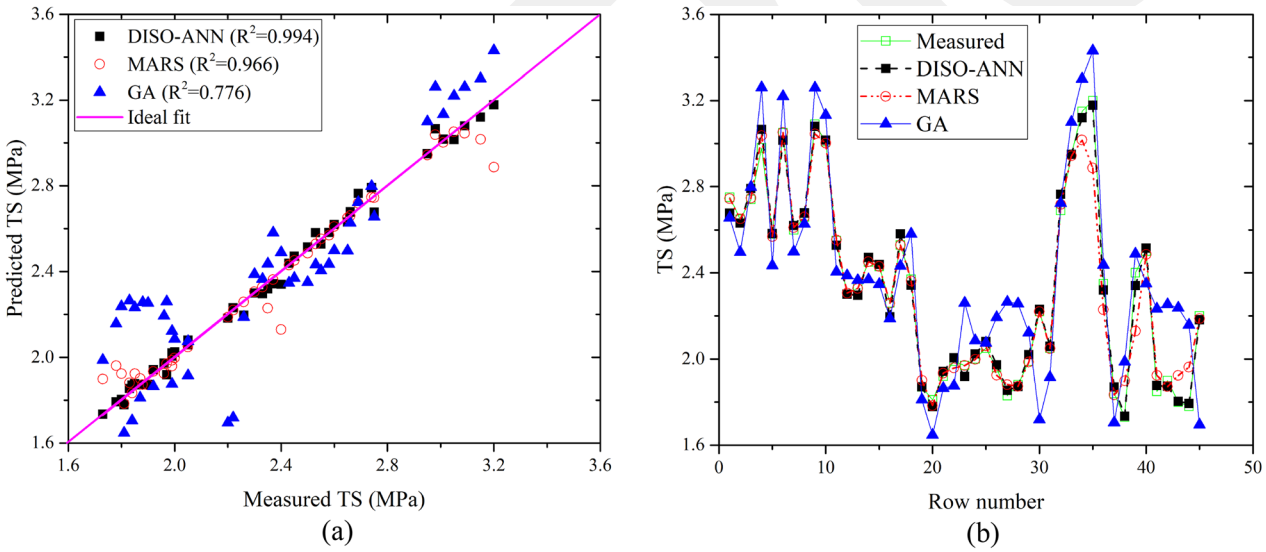

Figure 6. Model comparison for TS.

Models' Implementation Example

The datasets obtained from Khandelwal and Singh (2009) were used to validate the proposed models to establish further the possible effectiveness of their predictions. The datasets used for the purpose were never part of those used for the model developments. The datasets used for the validation comprises coal, shale and sandstone samples. However, some of the values of the rock properties presented in Khandelwal and Singh (2009) fell outside the range of those used in developing the model. For instance, P_{WV} maybe within the range, while density could be outside the range or one of

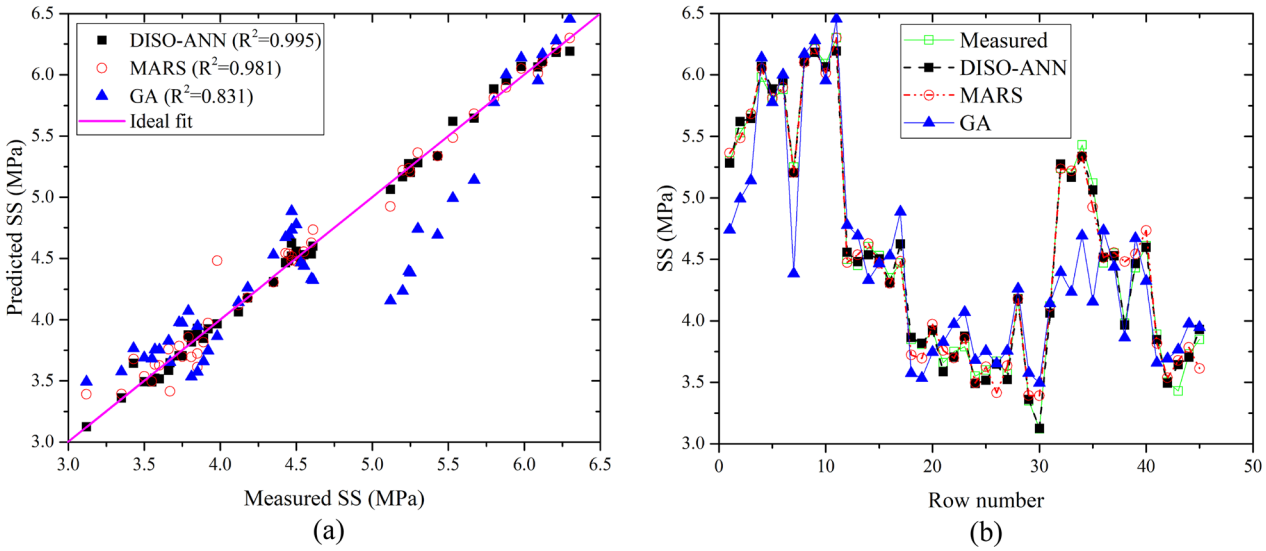

Figure 7. Model comparison for SS.

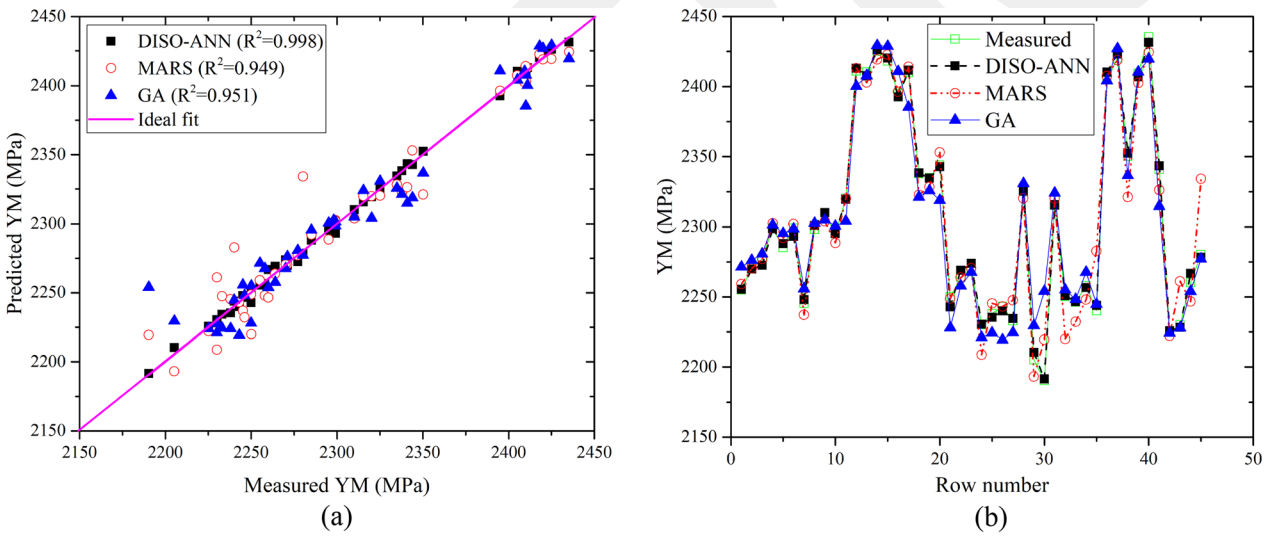

Figure 8. Model comparison for YM.

the investigated properties. However, the gap between them is not too significant and therefore we used the 12 datasets made available by Khandelwal and Singh (2009) to validate our model. The rock samples used by Khandelwal and Singh (2009) were also from an Indian coal mine. For this comparison, the best model selected previously, namely the DISO–ANN, was used and the outcomes for each of the investigated parameters are presented in Figures 10 and 11. The values of Pearson's *r* obtained for the four coal properties predicted were greater than 0.97 and very close to 1, which is the optimum value of Pearson's *r*. Therefore, the proposed model can still give reasonable predictions of the datasets for rock samples aside from those used in the model development. Hence, the proposed model can be considered general and useful for better geo-mechanical properties investigation.

Table 5. Statistical evaluation of the models' performances

	UCS				TS		
Pearson's r	0.9998	0.9959	0.9073	Pearson's r	0.9970	0.9831	0.8839
RMSE	0.1248	0.5495	2.3449	RMSE	0.0332	0.0818	0.2218
MAPE	0.3885	0.6908	9.5245	MAPE	1.1114	1.8457	8.3176
VAF	99.9491	99.0203	83.4732	VAF	99.3990	96.3721	70.9993
	SS				YM		
Pearson's r	0.9975	0.9906	0.9117	Pearson's r	0.9991	0.9742	0.9750
RMSE	0.0639	0.1237	0.3766	RMSE	2.91108	15.5015	15.2937
MAPE	1.1069	1.9851	6.2721	MAPE	0.31295	0.6675	0.6860
VAF	99.4971	98.1350	82.6339	VAF	99.8215	94.8814	94.8712

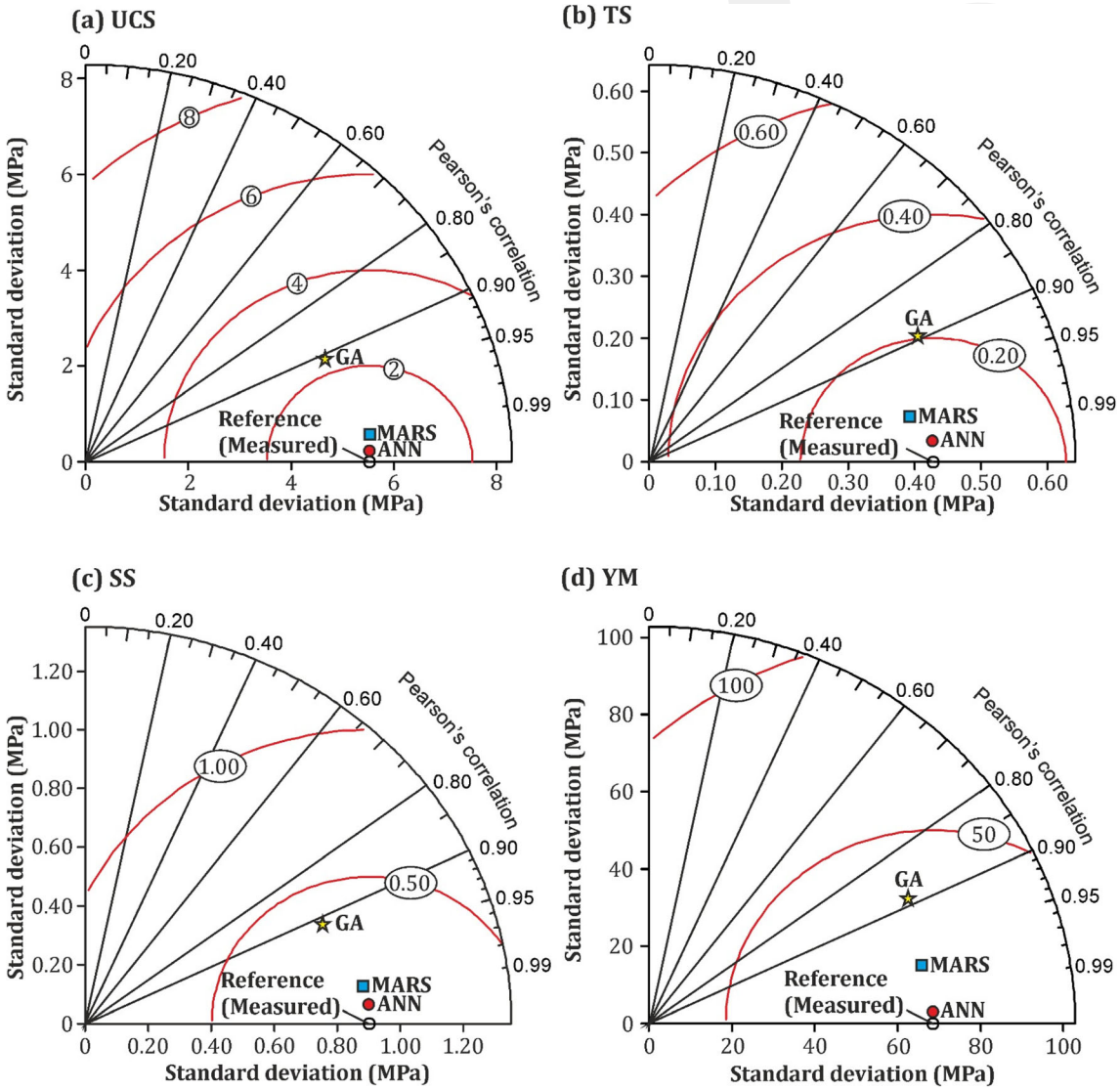

Figure 9. Comparison of the proposed models using Taylor's diagram.

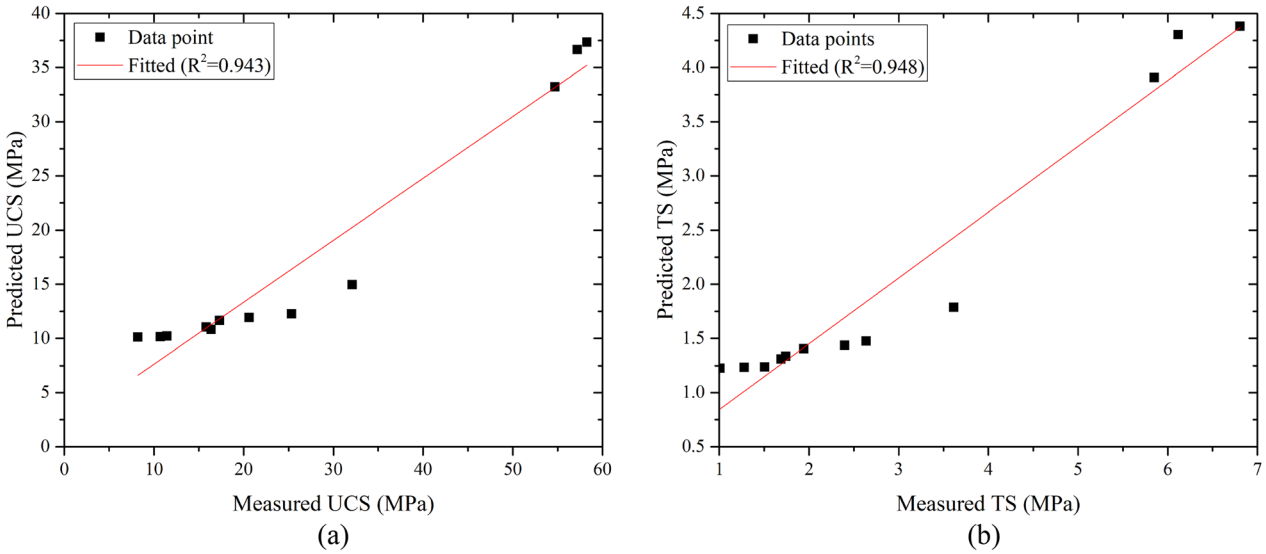

Figure 10. Validations of the proposed models using new datasets for **a** UCS and **b** TS.

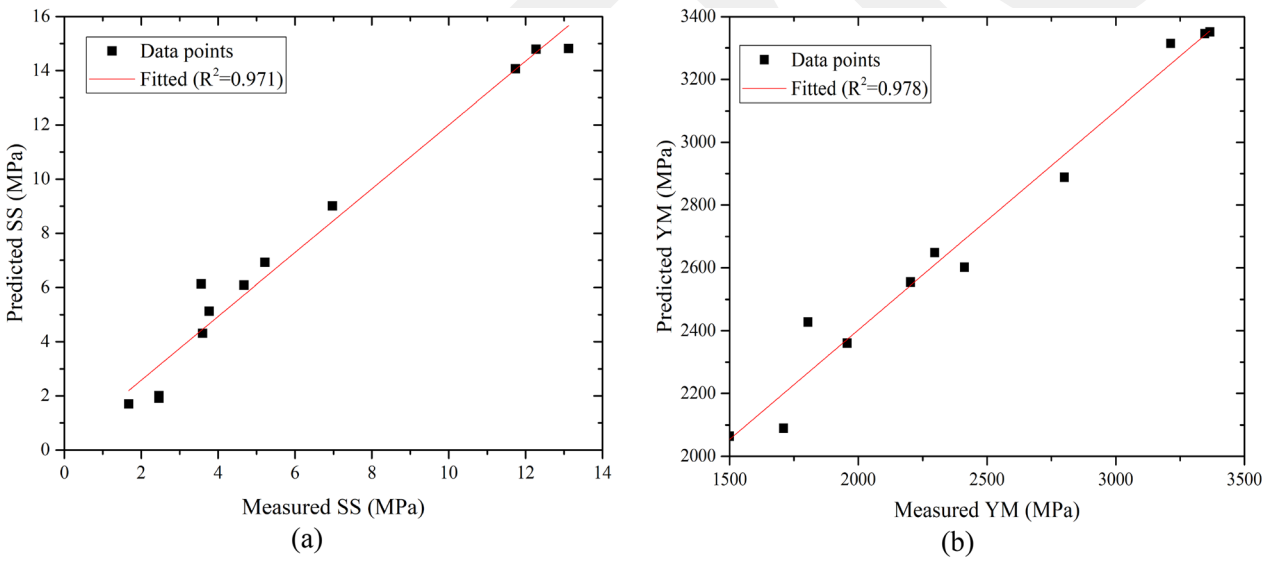

Figure 11. Validations of the proposed models using new datasets for **a** UCS and **b** TS.

Sensitivity Analysis

The influence of each of the input parameters into the models was investigated using the weighted average method. This method is peculiar to the ANN and one of the means for unlocking the black box nature of the ANN according to Gevrey et al. (2003). The method involves multiplication of the input and output weights of the ANN structure having in mind that each input parameter has the equivalent assigned weight. The formula for the computation of the weighted average method is (Garson, 1991; Lawal et al., 2021b):

$$R_i(\%) = \frac{\sum_{h=1}^{mh} \frac{|W_i| \times |W_o|}{\sum_{i=1}^{mi} (|W_i| \times |W_o|)}}{\sum_{h=1}^{mh} \sum_{i=1}^{mi} \frac{|W_i| \times |W_o|}{\sum_{i=1}^{mi} (|W_i| \times |W_o|)}} \times 100 \qquad (20)$$

where $R_i(\%)$ stands for the contributions (in percentage) of the model independent variables into the model, W_i stands for the connecting weights between the input and hidden layer, W_o represents the weights connecting the hidden and output layers, *mi* stands for the number of input variables, and *mh* represents the number of hidden layer neurons. Based on Eq. 20, the obtained percentage contributions of P_{WV} and ρ into the DISO–ANN model for the UCS, TS, SS, and YM are presented in Figure 12. From Figure 12, P_{WV} had the highest influence on the UCS and YM while the coal ρ influenced the TS and SS the most. Hence, if P_{WV} is high, it indicates that the strength of a rock will be high. However, the higher the density is, the stiffer the resistance to tensile or shear strength is. The density also influenced both UCS and YM but not as high as in the case of TS and SS.

CONCLUSIONS

Different empirical models have been proposed in predicting coal properties to minimize cost and time; but, majority of the existing models is based on regression, which is unable to capture inherent variabilities in rock properties. Hence, the need for reliable models in predicting mechanical properties of coal from simple non-destructive tests necessitated this study. Datasets from the literature were used to develop DISO–ANN, MARS and GA models for the prediction of UCS, TS, SS, and YM. The performances of the models were compared with the measured values and were assessed statistically using various statistical indices. In addition, the models were validated using datasets that were not part of those used in developing the models. It was found that the proposed models are capable of predicting mechanical properties of coal to a high degree of accuracy, and the relationships between the best fitness and mean fitness in GA dictated its prediction capability.

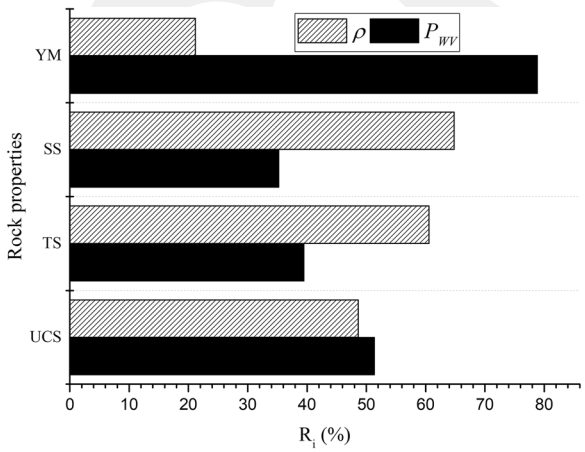

Figure 12. Percentage contribution of the input variables into the models.

ACKNOWLEDGMENTS

This work was supported by the Korea Research Fellowship Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2019H1D3A1A01102993) and Inha University Research Grant (2021).

REFERENCES

Abdolazim, A., & Rassoul, A. (2015). Empirical correlation of physical and mechanical properties of marly rocks with P wave velocity. *Arabian Journal of Geosciences, 8*, 2069–2079.

Aboutaleb, S., Behnia, M., Bagherpour, R., & Bluekian, B. (2018). Using non-destructive tests for estimating uniaxial compressive strength and static Young's modulus of carbonate rocks via some modeling techniques. *Bulletin of Engineering Geology and Environment, 77*, 1717–1728.

Aladejare, A. E., Alofe, E. D., Onifade, M., Lawal, A. I., Ozoji, T. M., & Zhang, Z.-X. (2021). Empirical estimation of uniaxial compressive strength of rock: Database of simple, multiple, and artificial intelligence-based regressions. *Geotechnical and Geological Engineering, 39*, 4427–4455.

Aliyu, M. M., Shang, J., Murphy, W., Lawrence, J. A., Collier, R., Kong, F., & Zhao, Z. (2019). Assessing the uniaxial compressive strength of extremely hard cryptocrystalline flint. *International Journal of Rock Mechanics and Mining Sciences, 113*, 310–321.

Dehghan, S., Sattari, G., Chelgani, S. C., & Aliabadi, M. (2010). Prediction of uniaxial compressive strength and modulus of elasticity for Travertine samples using regression and artificial neural networks. *Mining Science and Technology, 20*(1), 41–46.

Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics, 19*(1), 1–67.

Garson, G. D. (1991). Interpreting neural network connection weights. *Artificial Intelligence Expert, 6*(4), 47–51.

Gevrey, M., Dimopoulos, I., & Lek, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling, 160*(3), 249–264.

Heidari, M., Mohseni, H., & Jalali, S. H. (2017). Prediction of uniaxial compressive strength of some sedimentary rocks by fuzzy and regression models. *Geotechnical and Geological Engineering, 36*, 401–412.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor: The University of Michigan Press.

Horsrud, P. (2001). Estimating mechanical properties of shale from empirical correlations. *SPE Drilling and Completion, 16*(2), 68–73. https://doi.org/10.2118/56017-PA.

ISRM. (1978). Suggested method for determining sound velocity. *International Journal Rock Mechanics and Mining Sciences Geomechanics Abstract, 15*(2), 53–58.

ISRM. (1981). *Suggested method for rock characterization ISRM commission on testing methods* (p. 211). Pergamon Press, Oxford.

Jalali, S. H., Heidari, M., & Mohseni, H. (2017). Comparison of models for estimating uniaxial compressive strength of some sedimentary rocks from Qom Formation. *Environmental Earth Sciences, 76*, 1–15.

Katoch, S., Chauhan, S. S., & Kumar, V. (2021). A review on genetic algorithm: Past, present, and future. *Multimedia Tools and Applications, 80*, 8091–8126.

Khandelwal, M., & Singh, T. N. (2009). Correlating static properties of coal measures rocks with P-wave velocity. *International Journal of Coal Geology, 79*(1–2), 55–60.

Kurtulus, C., CakIr, S., & Yogurtcuoglu, A. C. (2016). Ultrasound study of limestone rock physical and mechanical properties. *Soil Mechanics and Foundation Engineering, 52*, 348–354.

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. *Scientific African, 8*, 1–10.

Lawal, A. I., & Idris, M. A. (2019). An artificial neural network-based mathematical model for the prediction of blast-induced ground vibrations. *International Journal of Environmental Studies, 77*(2), 318–334.

Lawal, A. I., & Kwon, S. (2020). Application of artificial intelligence in rock mechanics: An overview. *Journal of Rock Mechanics and Geotechnical Engineering, 13*, 248–266.

Lawal, A. I., Kwon, S., Hammed, O. S., & Idris, M. A. (2021a). Blast-induced ground vibration prediction in granite quarries: An application of Gene expression programming, ANFIS, and Sine Cosine algorithm optimized ANN. *International Journal of Mining Science and Technology, 31*, 265–277.

Lawal, A. I., Kwon, S., & Kim, G. Y. (2021b). Prediction of the blast-induced ground vibration in tunnel blasting using ANN, moth-flame optimized ANN, and gene expression programming. *Acta Geophysica, 69*, 161–174.

Lawal, A. I., Kwon, S., & Kim, G. Y. (2021c). Prediction of an environmental impact of tunnel blasting using artificial neural network, particle swarm and Dragonfly optimized artificial neural networks. *Applied Acoustics, 181*, 1–15.

Manouchehrian, A., Sharifzadeh, M., & Moghadam, R. H. (2012). Application of artificial neural networks and multivariate statistics to estimate UCS using textural characteristics. *International Journal of Mining Science and Technology, 22*(2), 229–236.

Michalewicz, Z. (1992). *Genetic algorithms + data structures = evolution programs*. Springer-Verlag.

Michalewicz, Z., & Schoenauer, M. (1996). Evolutionary algorithms for constrained parameter optimization problems. *Evolutionary Computation, 4*(1), 1–32.

Mohamad, E. T., Armaghani, D. J., Momeni, E., Yazdavar, A. H., & Ebrahimi, M. (2018). Rock strength estimation: A PSO-based BP approach. *Neural Computing and Applications, 30*, 1635–1646.

Moradian, Z. A., & Behnia, M. (2009). Predicting the uniaxial compressive strength and static Young's modulus of intact sedimentary rocks using the ultrasonic test. *International Journal of Geomechanics, 9*, 14–19.

Onifade, M., Lawal, A. I., Aladejare, E. A., Bada, S., & Idris, M. A. (2019). Prediction of gross calorific value of solid fuels from their proximate analysis using soft computing and regression analysis. *International Journal of Coal Preparation and Utilization*. https://doi.org/10.1080/19392699.2019.1695605.

Ren, Q., Wang, G., Li, M., & Han, S. (2019). Prediction of rock compressive strength using machine learning algorithms based on spectrum analysis of geological hammer. *Geotechnical and Geological Engineering, 37*, 475–489.

Roy, D. H., & Singh, T. N. (2019). Predicting deformational properties of Indian coal: Soft computing and regression analysis approach. *Measurement, 149*, 1–13.

Sharma, L. K., Vishal, V., & Singh, T. N. (2017). Developing novel models using neural networks and fuzzy systems for the prediction of strength of rocks from key geomechanical properties. *Measurement, 102*, 158–169.

Singh, R., Umrao, R. K., Ahmad, M., Ansari, M. K., Sharma, L. K., & Singh, T. N. (2017). Prediction of geomechanical parameters using soft computing and multiple regression approach. *Measurement, 99*, 108–119.

Sun, Y., Li, G., Zhang, J., Sun, J., & Xu, J. (2020). Development of an ensemble intelligent model for assessing the strength of cemented paste backfill. *Advances in Civil Engineering, 2020*, 1643529.

Sun, Y., Li, G., Zhang, N., Chang, Q., Xu, J., & Zhang, J. (2021). Development of ensemble learning models to evaluate the strength of coal-grout materials. *International Journal of Mining Science and Technology, 31*(2), 153–162.

Sun, Y., Zhang, J., Li, G., Ma, G., Huang, Y., Sun, J., Wang, Y., & Nener, B. (2019a). Determination of Young's modulus of jet grouted coalcretes using an intelligent model. *Engineering Geology, 252*, 43–53.

Sun, Y., Zhang, J., Li, G., Wang, Y., Sun, J., & Jiang, C. (2019b). Optimized neural network using beetle antennae search for predicting the unconfined compressive strength of jet grouting coalcretes. *International Journal of Numerical and Analytical Methods in Geomechanics, 43*(4), 801–813.

Uyanik, O., Sabbag, N., Uyanik, N. A., & Oncu, Z. (2019). Prediction of mechanical and physical properties of some sedimentary rocks from ultrasonic velocities. *Bulletin of Engineering Geology and Environment, 78*(8), 6003–6016.

Verma, D., Kainthola, A., Singh, R., & Singh, T. N. (2012). Assessment of geo-mechanical properties of some Gondwana coal using P-wave velocity. *International Research Journal of Geology and Mining (IRJGM), 2*(9), 261–274.

Yasar, E., & Erdogan, Y. (2004). Correlating sound velocity with the density, compressive strength and Young's modulus of carbonate rocks. *International Journal of Rock Mechanics and Mining Sciences, 41*(5), 871–875.

Zhang, J., Sun, Y., Li, G., Wang, Y., Sun, J., & Li, J. (2020). Machine-learning-assisted shear strength prediction of reinforced concrete beams with and without stirrups. *Engineering with Computers*. https://doi.org/10.1007/s00366-020-01076-x.