

Article

A Novel Combination of PCA and Machine Learning Techniques to Select the Most Important Factors for Predicting Tunnel Construction Performance

Jiangfeng Wang ¹, Ahmed Salih Mohammed ² , Elżbieta Macioszek ^{3,*} , Mujahid Ali ^{4,*} , Dmitrii Vladimirovich Ulrikh ⁵  and Qiancheng Fang ⁶

¹ College of Geosciences and Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450046, China; wangjiangfeng@ncwu.edu.cn

² Civil Engineering Department, College of Engineering, University of Sulaimani, Sulaymaniyah 46001, Iraq; ahmed.mohammed@univsul.edu.iq

³ Department of Transport Systems, Traffic Engineering and Logistics, Faculty of Transport and Aviation Engineering, Silesian University of Technology, Krasińskiego 8 Street, 40-019 Katowice, Poland

⁴ Department of Civil Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur 50603, Malaysia

⁵ Department of Urban Planning, Engineering Networks and Systems, Institute of Architecture and Construction, South Ural State University, 76 Lenin Prospect, 454080 Chelyabinsk, Russia; ulrikh@susu.ru

⁶ Institute of Architecture Engineering, Huanghuai University, Zhumadian 463000, China; hhufangqiancheng@126.com

* Correspondence: elzbieta.macioszek@polsl.pl (E.M.); mujahid_19001704@utp.edu.my (M.A.)

Abstract: Numerous studies have reported the effective use of artificial intelligence approaches, particularly artificial neural networks (ANNs)-based models, to tackle tunnelling issues. However, having a high number of model inputs increases the running time and related mistakes of ANNs.

The principal component analysis (PCA) approach was used in this work to select input factors for predicting tunnel boring machine (TBM) performance, specifically advance rate (AR). A reliable and precise forecast of TBM AR is desirable and critical for mitigating risk throughout the tunnel building phase. The developed PCAs (a total of four PCAs) were used with the artificial bee colony (ABC) method to predict TBM AR. To assess the created PCA-ANN-ABC model's capabilities, an imperialist competitive algorithm-ANN and regression-based methods for estimating TBM AR were also suggested. To evaluate the artificial intelligence and statistical models, many statistical evaluation metrics were evaluated and generated, including the coefficient of determination (R^2). The findings indicate that the PCA-ANN-ABC model (with R^2 values of 0.9641 for training and 0.9558 for testing) is capable of predicting AR values with a high degree of accuracy, precision, and flexibility. The modelling approach utilized in this study may be used to other comparable studies involving the solution of engineering challenges.

Keywords: PCA; ANN; Artificial Bee Colony Algorithm; TBM advance rate; hard rock condition



Citation: Wang, J.; Mohammed, A.S.; Macioszek, E.; Ali, M.; Ulrikh, D.V.; Fang, Q. A Novel Combination of PCA and Machine Learning Techniques to Select the Most Important Factors for Predicting Tunnel Construction Performance. *Buildings* **2022**, *12*, 919. <https://doi.org/10.3390/buildings12070919>

Received: 25 May 2022

Accepted: 26 June 2022

Published: 29 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In mechanized tunneling designs, estimating the tunnel boring machine (TBM) performance is considered a vital task before selecting the machine and conducting the project. It is important for the project schedule, management of relevant issues, and cost estimation of tunneling projects [1,2]. Many techniques/formulas have been proposed theoretically and empirically by previous investigators for predicting different factors related to TBM performance (e.g., penetration rate, PR and advance rate, AR) [3–8]. These empirical and theoretical techniques mostly used one or two predictors relevant to rock material and mass properties such as strength and joint conditions [9–11]. Hence, the mentioned techniques according to previous studies [12–14] are not good or strong enough to provide a suitable level of TBM performance prediction.

Aside from empirical and theoretical techniques, statistical-based models such as multiple and simple regression have been employed for assessing TBM performance [15–17]. In simple regression models/equations, researchers evaluated TBM performance such as PR using only one predictor, mainly from the rock mass and material properties (such as rock strength or a rock mass classification system) [18,19]. In multiple regression models/equations, the use of a minimum of two effective parameters on TBM performance was carried out by several researchers to predict TBM PR [15–17,20]. However, according to several studies such as [17,21], statistical-based techniques cannot always investigate complex systems. Additionally, the performance capacity of the mentioned models is not at a satisfactory level [9,12], while a satisfactory and acceptable degree of prediction is required for TBM performance in order to minimize the risks associated with tunneling project costs.

Several scholars have presented different artificial intelligence (AI)-based and machine learning (ML)-based techniques for solving TBM-related issues [2,13,22–25]. Some of the techniques proposed in this regard are support vector machine (SVM), particle swarm optimization (PSO), fuzzy inference system (FIS), and artificial neural network (ANN). These methods have been generally utilized to approximate the TBM field penetration index (FPI), TBM PR and TBM AR. The Athens Metro tunnel database was used by Benardos and Kaliampakos [26] to develop an ANN-based model to calculate TBM AR. In another project, Yagiz et al. [13] introduced an ANN approach for the estimation of the PR of TBM. On the other hand, in the case of the same datasets, Mahdevari et al. [15] intended an ML approach, namely support vector regression, to solve a TBM problem. In predicting FPI values, Feng et al. [27] and Adoko and Yagiz [28] tried to solve this problem by proposing deep learning and FIS techniques, respectively. A group of other authors tried to solve performance of TBM using other single intelligent approaches like group modeling of data handling, genetic-based and neuro-fuzzy [2,9,29–31]. On the other hand, some other scholars developed advanced intelligent techniques for this problem which are based on a combination of at least two different intelligent models [1,10,14,32,33]. In these combined models, there is a base technique and an optimization algorithm for improving prediction capacity of the base technique in estimating TBM performance. It is crucial to note that AI and ML approaches have been highly employed to solve difficulties in science and engineering [34–52].

After reviewing the related AI and ML works, it was found that the ANN model is one of the frequently used techniques in evaluating and predicting TBM performance. However, the ANN model is connected with several limitations that lead to a low-performance prediction for this model. These limitations include slow learning rate [53,54]. Therefore, this study aims to increase the prediction model's ability by training the ANN using the artificial bee colony (ABC) algorithm. The capability of the ABC-ANN model is compared with another hybrid ANN and statistical-based models to choose more accurate techniques in the area of TBM AR. It is significant to note that the principal component analysis (PCA) is applied to reduce the number of input parameters for predicting TBM AR. As far as the authors know, this is the first time using the PCA in the area of TBM performance to make a more straightforward technique. Then, the PCA-ANN-ABC model results will be discussed and compared at the end of this research.

2. Methodology Background

2.1. ANN

Among many ANN structures studied, the most common use is the hidden multi-layer feed-forward network structure. The main structure of a feed-forward (FF) model usually consists of three distinct layers [55]. Each layer is made up of nodes called neurons. In an FF model, data travels in only one direction, from input to output neurons, through the hidden nodes. Each artificial node of the network receives the signal and passes the signal through an activation function to estimate the output. Every neuron's output will be the input for the subsequent neuron. An ANN can be trained multiple times to improve its performance

capacity. During the data training, the architecture of an ANN and its connection weights are altered iteratively to reduce imperfections of the predicted data [56,57].

When the architecture of an FF neural network is defined, the parameters include the activation function, layer No. and the neuron No. per layer; in the next step, there is a need to adjust the ANN weights and biases. These parameters should be intended perfectly to receive the best prediction capacity (i.e., highest performance prediction). A complete explanations about ANNs and their process in performing prediction tasks can be found in previous investigations [58–61].

2.2. ABC

A colony-based optimization algorithm, namely ABC, was proposed by Tereshko [62] which works according to bee colony behavior in real life. In this optimization technique, there are two different behaviors for bee colonies: source abandonment and food source recruitment [62,63]. According to Tereshko [62], the basic components of ABC can be described in the following:

- Food source selection: In this stage, bee colonies will be searched for different sources of food and their characteristics. Some include nectar taste, the distance between bee colony and the hive, and energy richness [64].
- Employed forager: an employed forager works in the food source from which it currently feeds. The responsibility of the employed forager is to share the food supplier information with other bees waiting in the hive [63,64].

In the ABC algorithm, when there are values for several parameters such as the quantity of nectar and the position of the food source, an optimization process for finding the possible answers for the problem can be performed [64]. A random number of bees will be selected in the ABC algorithm, where they are searching within the search space for possible and practical solutions. Independent artificial bees work together and exchange information to find/optimize more accurate results/solutions [65]. Actually, these bees, using knowledge and sharing data, focus on promising areas and gradually leave low-hope areas. Overall, in each iteration of the ABC algorithm, the artificial bees will improve their capabilities in finding more accurate solutions [66]. The process of the ABC algorithm will continue until the best answers/solutions that the user can define from the beginning. The role of this algorithm in a combined ABC-ANN is actually to optimize the ANN weights and biases, which will increase the performance capacity of the hybrid system. The general process of the ANN technique optimized by the ABC algorithm is shown in Figure 1.

2.3. Principal Component Analysis (PCA)

PCA's main idea is to diminish the number of interdependent variables to preserve their variations in the set of remaining parameters. This result is achieved using a principal components transformation function on the main parameters. Principal components are not related to each other and are arranged to contain the highest variance of the main variables [67,68]. One of the most critical aspects of the PCA approach is choosing the number of PCs. This approach defines the accuracy according to the data and the desired results. Then the total changes in the number of principal components are selected based on the cumulative percentage. The expected accuracy is deemed to be 80% to 90% of the overall variations. The rate of incremental changes is defined from Equation (1), where l_k is the eigenvalue [67].

$$\text{Percent of variation} = \frac{l_k}{\sum_{j=1}^p l_k} \quad (1)$$

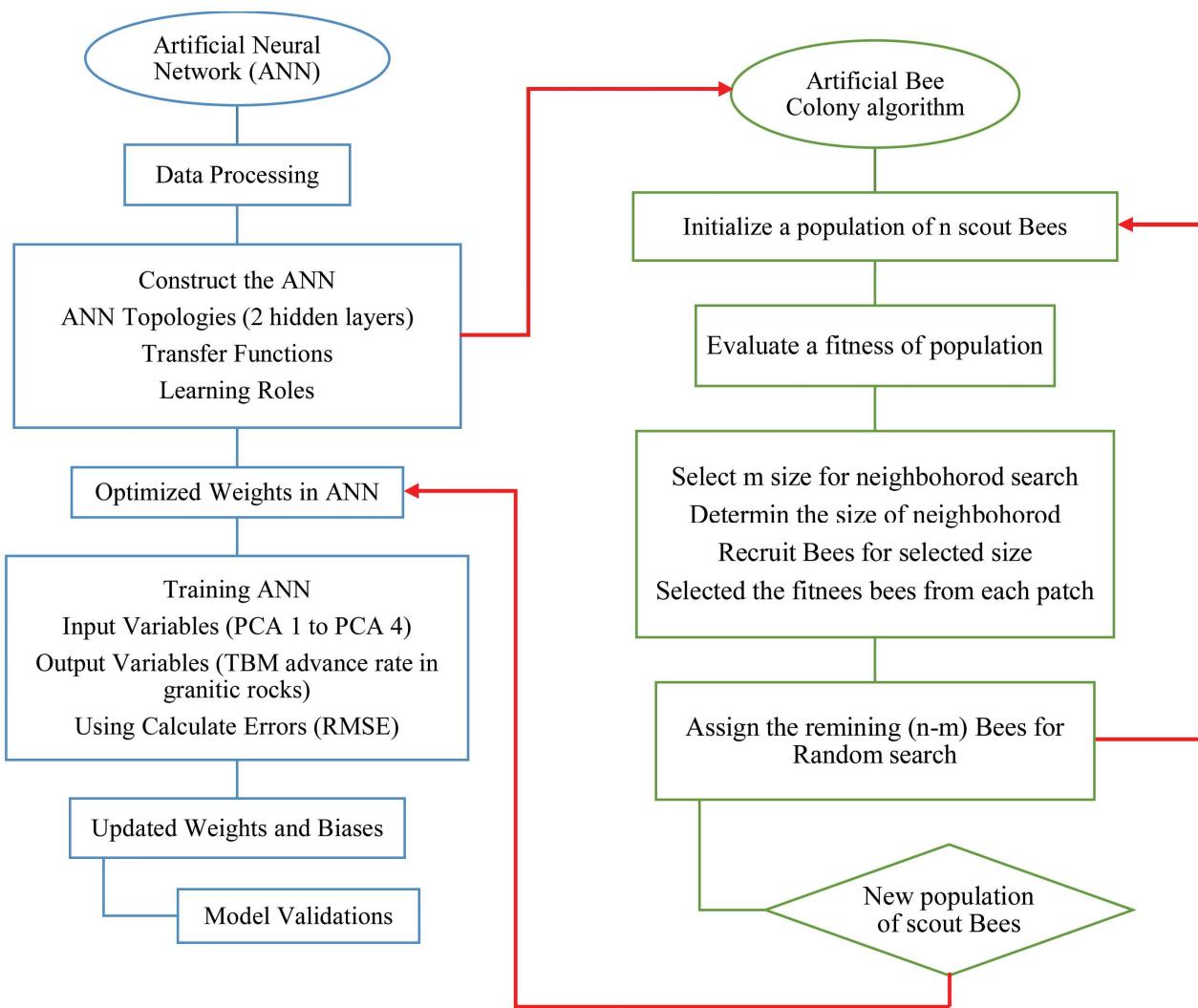


Figure 1. Flowchart of a hybrid ANN optimized with the Artificial Bee Colony algorithm.

3. Materials and Methods

3.1. Case Study and Data Collection

To predict TBM AR in this study, a tunnel site and project in Malaysia were selected to be studied. This tunnel aims to transfer water from one state (i.e., Pahang) to another (i.e., Selangor). The primary rock type in the research area was granite, and the tunnel has an overburden range of 100–1400 m. The total length of this tunnel is 44 km which was excavated using 2 different construction techniques namely (1) drilling and blasting and (2) mechanized excavation or TBM. In the case of mechanized excavation, three TBMs were used to excavate approximately 35 km of the tunnel with a diameter of 5.2 m.

In this study, the database was prepared based on the previous suggestions/investigations. The mentioned database, which comprises 1286 datasets, covers laboratory test results and field observations. In terms of laboratory tests, uniaxial compressive strength (UCS) and Brazilian tensile strength (BTS) tests were conducted and their results were recorded. In this case, more than 100 rock block samples were transferred to the laboratory and the samples for the mentioned tests were prepared and tested. These laboratory tests were carried out according to ISRM guidelines [69]. In the tunnel site, several rock mass properties, i.e., rock mass rating (RMR), rock quality designation (RQD), and weathering zone (WZ) were measured for each panel which is 10 m of the tunnel site. Various values were obtained for the mentioned parameters. For example, as presented in Table 1, minimum values of 10%, 45.4 MPa, 46, and 4.69 MPa were obtained for RQD, UCS, RMR and BTS

parameters, respectively, while their maximum values were measured as 95%, 193 MPa, 95 and 15.68 MPa. In terms of WZ, three zones were observed as fresh with a grade of 1 in the analysis, slightly weathered with a grade of 2 in the analysis, and moderately weathered with a grade of 3 in the analysis. As shown in Table 1, revolutions per minute (RPM) and thrust force per cutter (TFPC) were recorded by the TBMs in the tunnel and used in this study as two important machine parameters. The system output based on Table 1 is TBM AR or AR which is in the range of 0.20–2.57 m/h. It is significant to note that the database description in Table 1 is not considered as 1286 datasets, and it is only 1205 data samples. The difference in datasets is related to outliers (or unusual values in the data) that the authors attempted to identify and remove. This is considered as a mandatory task when there is a large amount of data, i.e., several data samples in this research. Outliers are able to increase the variability in the database, which produces an amount of modeling error (i.e., difference between measured AR and predicted AR) and decreases statistical power.

Table 1. Characteristics of the system inputs and output in this investigation.

Factor	Unit	Type	Max	Min	Average	STD
RQD	%	Input	95	10	53.79	27.85
UCS	MPa	Input	193	45.4	134.68	44.30
RMR	-	Input	95	46	72.74	15.71
BTS	MPa	Input	15.68	4.69	10.26	4.04
WZ	-	Input	3	1	1.68	0.69
TFPC	kN	Input	497.67	91.34	303.26	77.76
RPM	rev/min	Input	11.95	4.54	8.91	2.26
TBM AR	m/h	Output	2.57	0.20	1.08	0.57

Max: maximum; Min: minimum; STD: standard deviation.

A total of 81 data samples were identified as outliers from the whole data (i.e., inputs and output) through the use of boxplot rules [70]. Boxplots are a simple way to depict a five-number summary, including the lowest (min), highest (max), first (Q1), second (Q2), and third quartiles (Q3). This technique calculates the range (Max–Min) and IQR (Q3–Q1), which can help us distinguish normal data from outliers. The statistical information presented in Table 1 is based on the whole database after removing outliers.

In statistical terms, correlation or dependence is defined as any statistical association (causal or non-causal) between two arbitrary variables or bivariate data. Generally, any statistical association can be implied as a correlation; yet, it usually denotes the degree of a linear correlation between a pair of parameters [71]. A correlation matrix is demonstrated as a table that reflects the correlation coefficients between different variables in which each cell depicts the correlation between a pair of variables. A correlation matrix summarizes data that can be utilized as input data for more complex analyses and as a diagnostic tool for advanced analyses [72]. Table 2 and Figure 2 show the correlation matrix values for the input parameters.

Table 2. Correlation matrix for input variables.

Name	UCS	BTS	RQD	RMR	WZ	TFPC	RPM
UCS	1						
BTS	0.8	1					
RQD	0.71	0.67	1				
RMR	0.77	0.73	0.77	1			
WZ	−0.12	−0.11	−0.22	−0.23	1		
TFPC	−0.72	−0.67	−0.64	−0.72	−0.05	1	
RPM	−0.76	−0.78	−0.68	−0.69	0.02	0.68	1

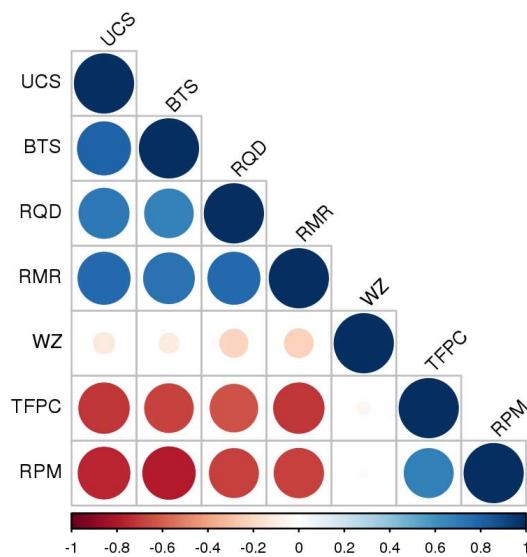


Figure 2. Correlation matrix for input variables.

3.2. Study Steps

The flowchart of the study and its different steps to solve problems related to TBM AR are presented in Figure 3. According to this flowchart, the model's inputs and output were measured in the laboratory and the tunnel site after investigating the data needed. Then, to propose a simpler model of interest and importance, the PCA technique was used to generate new input parameters. The new inputs are actually a combination of the other inputs presented in Table 1. After that, the modeling stage of this study was started by constructing the feed-forward ANN and ANN optimized by ABC models to predict TBM AR.

On the other hand, a multiple regression model and an ANN-imperialist competitive algorithm (ICA) technique were built for comparison purposes. The developed models were evaluated and the more accurate predicted model was selected and introduced for AR estimation. Finally, the effects of all model predictors were investigated and reported using another statistical technique.

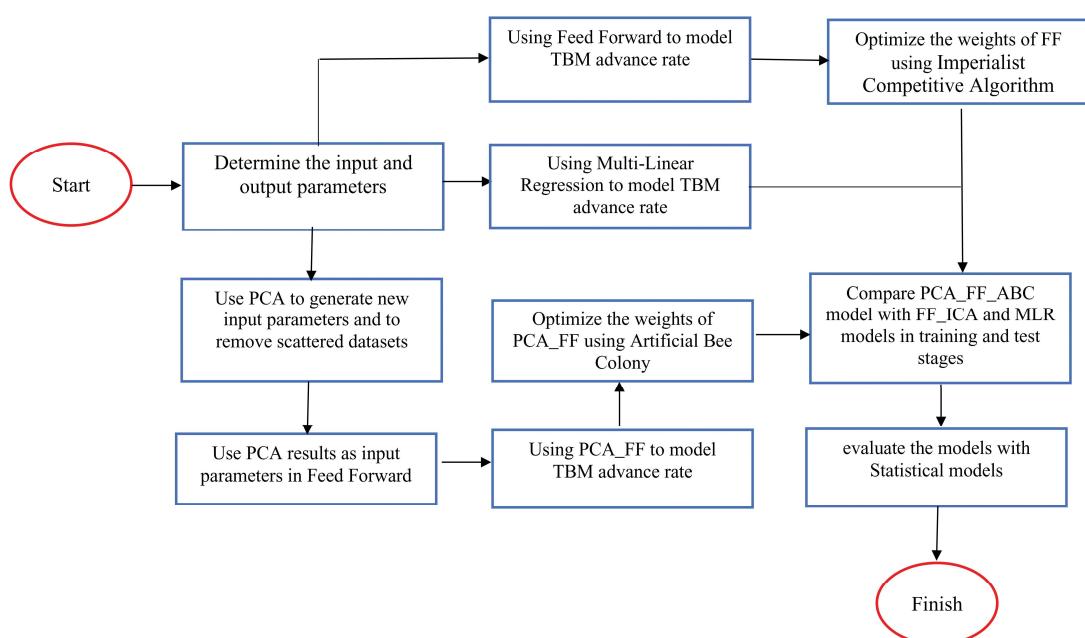


Figure 3. Flowchart of the study trend of a hybrid PCA and ANN optimized with the ABC algorithm.

3.3. Research Methodology

The number of input parameters used is seven, and the output parameter is the actual AR (m/h). Given that many input parameters in ANNs increase the error, PCA can be employed to orthogonalize the input variables relative to each other. The input density diagram is shown in Figure 4. MINITAB version 14.0 software was utilized to analyze the input parameters using PCA. The effects of the parameters are shown in Table 3. The graph of the effect of PCA parameters, which expresses the eigen analysis correlation matrix, is shown in Figure 5.

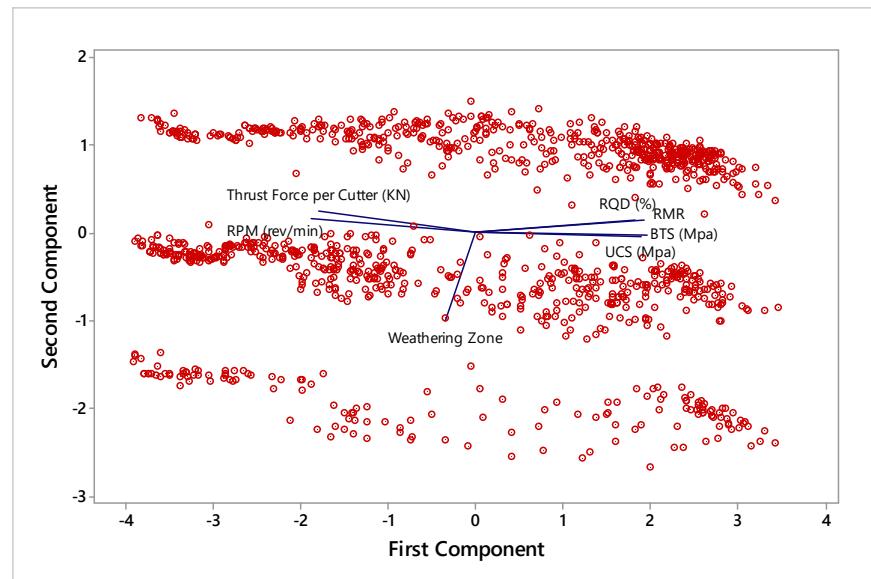


Figure 4. Density diagram of the input parameters using PCA.

Table 3. Eigen Analysis correlation matrix for establishing input variables by PCA.

Parameter	Inputs						
	PCA 1	PCA 2	PCA 3	PCA 4	PCA 5	PCA 6	PCA 7
Eigenvalue	4.6162	1.0666	0.3896	0.3151	0.2323	0.1937	0.1865
Proportion	0.659	0.152	0.056	0.045	0.033	0.028	0.027
Cumulative	0.659	0.812	0.867	0.913	0.946	0.973	1

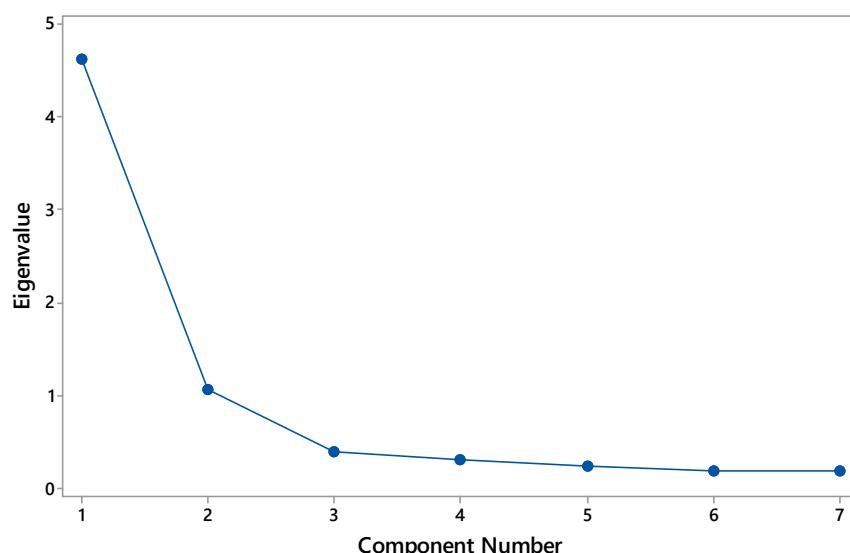


Figure 5. Eigen Analysis correlation matrix in displaying the PCA parameters.

According to Table 3, the conversion of seven input parameters into four variables of PCA 1 to PCA 4 using PCA results in approximately 91% of the data; due to data convergence, better results can be reached. The resulting input variables using PCA are listed in Table 4. Considering that the number of input parameters has decreased, PCA's resulting four input parameters were used to model the ANN (Table 4). Also, considering that AR output data's statistical behavior should be evaluated, its histogram is plotted in Figure 6. This indicates that 966 samples from the sample set are in the range of 0.2 to 1.56; approximately 80% of the data volume falls within this range. The probability plot diagrams for determining the normal distribution of the AR parameter (see Figure 6) show that their statistical behavior's AR output parameter follows the normal distribution.

Table 4. Relationship between the principal components and input parameters.

Variable	Unit	PCA 1	PCA 2	PCA 3	PCA 4	PCA 5	PCA 6	PCA 7
RQD	(%)	0.397	0.138	0.369	0.737	0.048	0.331	0.181
UCS	(MPa)	0.424	-0.031	-0.151	-0.159	-0.433	0.433	-0.629
RMR	(-)	0.417	0.128	0.317	-0.019	-0.338	-0.766	-0.092
BTS	(MPa)	0.412	-0.035	-0.517	-0.129	-0.293	0.056	0.675
WZ	(-)	-0.074	-0.941	0.07	0.23	-0.21	-0.083	0.009
TFPC	(kN)	-0.388	0.235	-0.499	0.573	-0.383	-0.202	-0.174
RPM	(rev/min)	-0.405	0.146	0.469	-0.185	-0.648	0.256	0.276

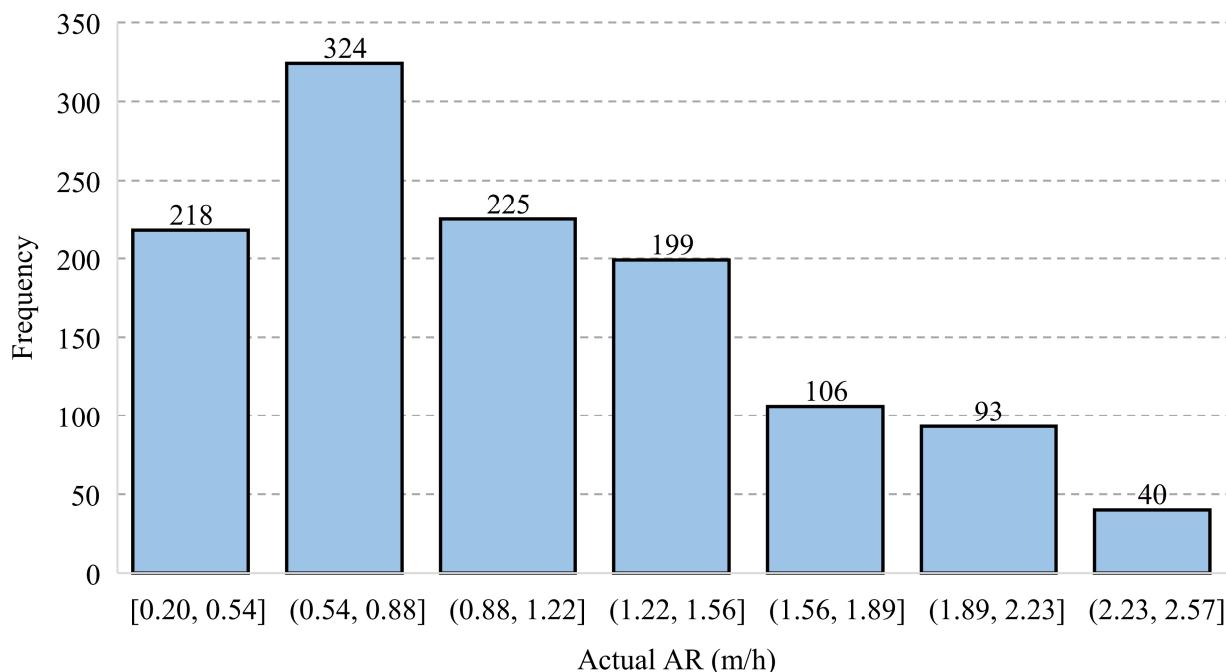


Figure 6. Histogram of the AR parameter.

In this research, the new feed-forward ANN is selected and used. To start the modeling, the data should be randomly classified into two groups to reduce the effects of excessive errors. Of the 1205 data samples, 70% (844 samples) were chosen and applied for training, and 30% (361 samples) were chosen for network testing. This is in line with some of the previous studies [73–76]. In the proposed model, four input parameters were obtained through the PCA technique. Thus, the trained ANNs in the input layer have four nodes, while in the output layer, it has only one node (Figure 7). Hidden two-layer networks have been used for modeling. For an ANN model, the No. of hidden layers and the No. of neurons in the hidden layers vary on the problem. Accordingly, the trial-and-error technique was applied to obtain the ideal structure (i.e., the structure that best

represents the data). A standard heuristic method for the maximum No. of nodes is given in Equation (2) [77]:

$$N_H \leq 2N_I + 1 \quad (2)$$

where: N_H is the No. of nodes in hidden layer, and N_I is the No. of model inputs.

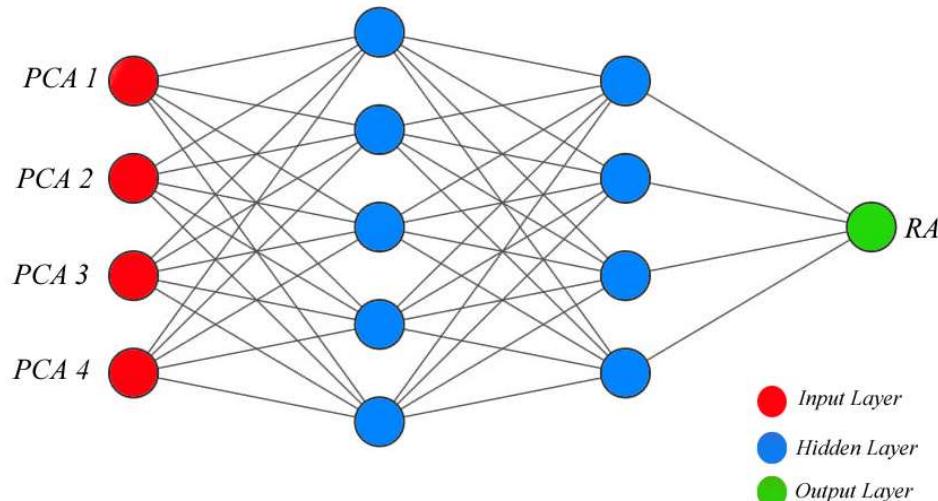


Figure 7. Topology of a feed-forward model with two hidden layers (4-5-4-1 topology).

Based on Equation (2), the total No. of N_H should be equal to or less than 9. Therefore, different architectures have a maximum of two hidden layers and a maximum of nine trained neurons. The employed architectures are given in Table 5 which include a total number of 20 models. In all models, the hyperbolic tangent stimulation function and the Levenberg–Marquardt training algorithm were used. In optimizing the ANN weights, the ABC algorithm as a strong optimization model was applied to optimize the ANN weights and ANN biases. ABC produces the least calculation error for the trained architecture. The properties of the ABC algorithm parameters are displayed in Table 6.

Table 5. The trained ANN architectures.

Num	Topology								
1	1-1	5	2-1	9	3-1	13	4-1	17	5-1
2	1-2	6	2-2	10	3-1	14	4-2	18	5-2
3	1-3	7	2-3	11	3-3	15	4-3	19	5-3
4	1-4	8	2-4	12	3-4	16	4-4	20	5-4

Table 6. The initialization parameters used in the ABC algorithm.

Number of Bees	Source Number of Bees	Max of Cycle Number	Onlooker Number
10	5	50	5

Among the trained models to determine the AR output parameter, the model with the 4-5-4-1 topology based on RMSE, R, and other statistical indices has been selected as the best model in Equations (3)–(5). The characteristics of the said model are provided in Table 7. The results of the best approach in the training and test sections are shown in Table 8. The statistical indices employed for the performance evaluation of the topologies include Root Mean Squared Error (RMSE), Average Absolute Error (AAE), and Variance Account Factor (VAF), which are defined according to Equations (3)–(5) [78].

$$\text{RMSE} = \left[\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2 \right]^{\frac{1}{2}} \quad (3)$$

$$\text{AAE} = \frac{\left| \sum_{i=1}^n \frac{(O_i - P_i)}{O_i} \right|}{n} \times 100 \quad (4)$$

$$\text{VAF} = \left[1 - \frac{\text{var}(O_i - P_i)}{\text{var}(O_i)} \right] \times 100 \quad (5)$$

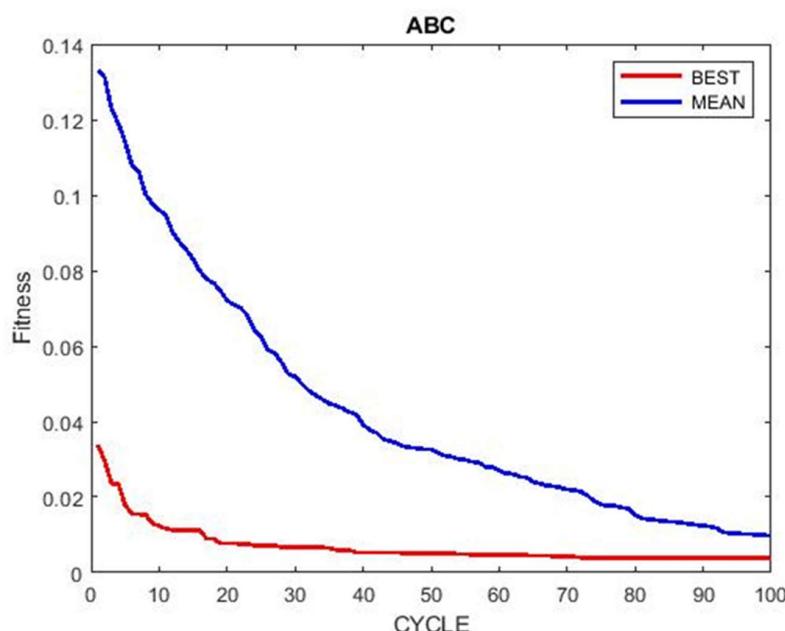
Table 7. Neural network characteristics as the best model.

Neural Network's Features					
Number of Inputs	Number of Outputs	Number of Hidden Layers	Number of Nodes in Hidden Layers	Transfer Function	Training Algorithm
5	1	2	5-4	tansig	Translim

According to Table 8, PCA-ABC-ANN 2L (5-4) networks have the lowest error in RMSE, AAE, and VAF indices and have the highest value of R^2 . For networks with the 4-5-4-1 topology, the values of R^2 in training, testing, and all modes for the two AR outputs are 0.9641, 0.9558, and 0.9617, respectively. The error criteria for training and test samples were computed using data rates in the main range of parameters and not in the normal range $[-1, +1]$, which is sometimes used in the literature. Figure 8 illustrates the PCA-ABC-ANN cost graph. The ANN performance is demonstrated in Figures 9 and 10 for the three phases of training, validation, and testing.

Table 8. Statistics of the best ANNs combined with the ABC algorithm using the 4-5-4-1 topology.

Step	Statistical Index	PCA-ABC-ANN 2L (5-4)
Train	RMSE	0.11
	AAE	0.12
	VAF%	96%
Test	RMSE	0.12
	AAE	0.13
	VAF%	96%

**Figure 8.** Cost graph for 100 replications of the PCA-ABC-ANN model as the best model.

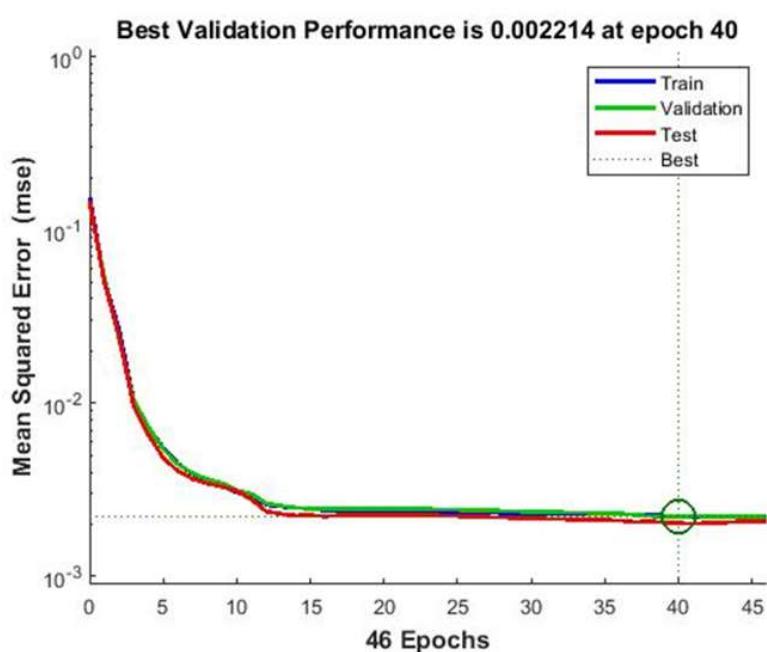


Figure 9. The MSE results of ANN together with the number of epochs.

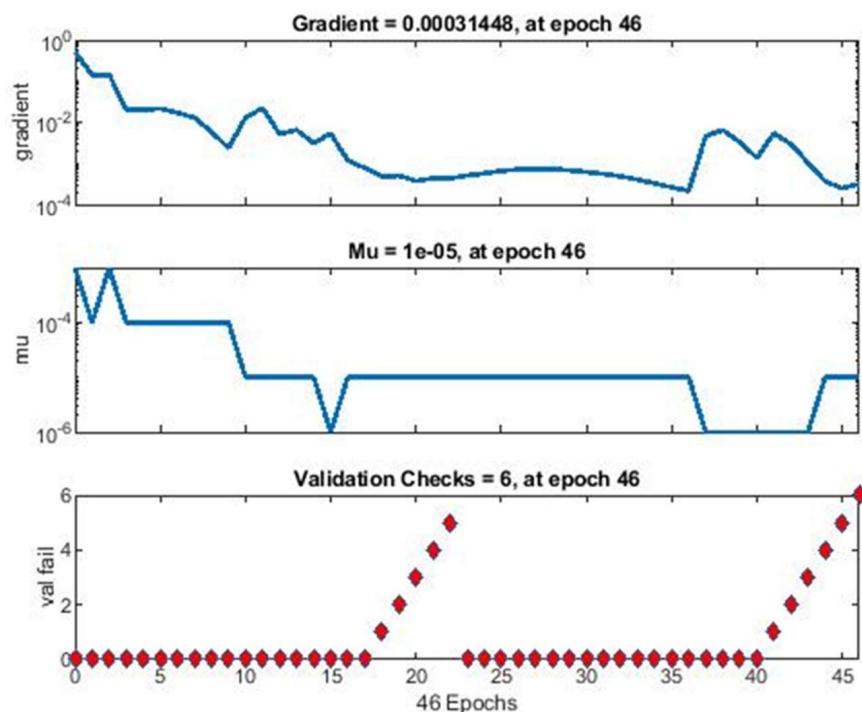


Figure 10. ANN results related to training phase.

The model's predicted values against their values from experiments for the training and test data are shown in Figures 11 and 12 to visualize the performance of the PCA-ABC-ANN 2L (5-4). As a result, the values predicted by the PCA-ABC-ANN model are very close to the line $y = x$, indicating a high accuracy capacity of this model.

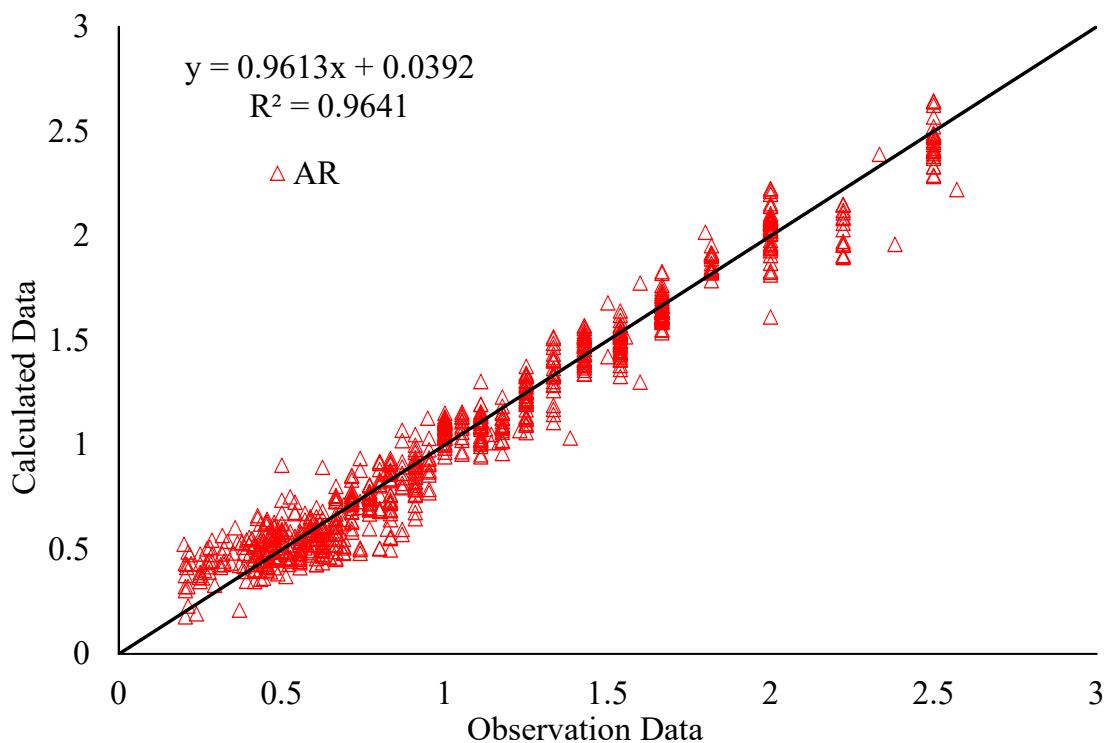


Figure 11. Predicted vs. experimental values of the AR output for the PCA-ABC-ANN model using training data.

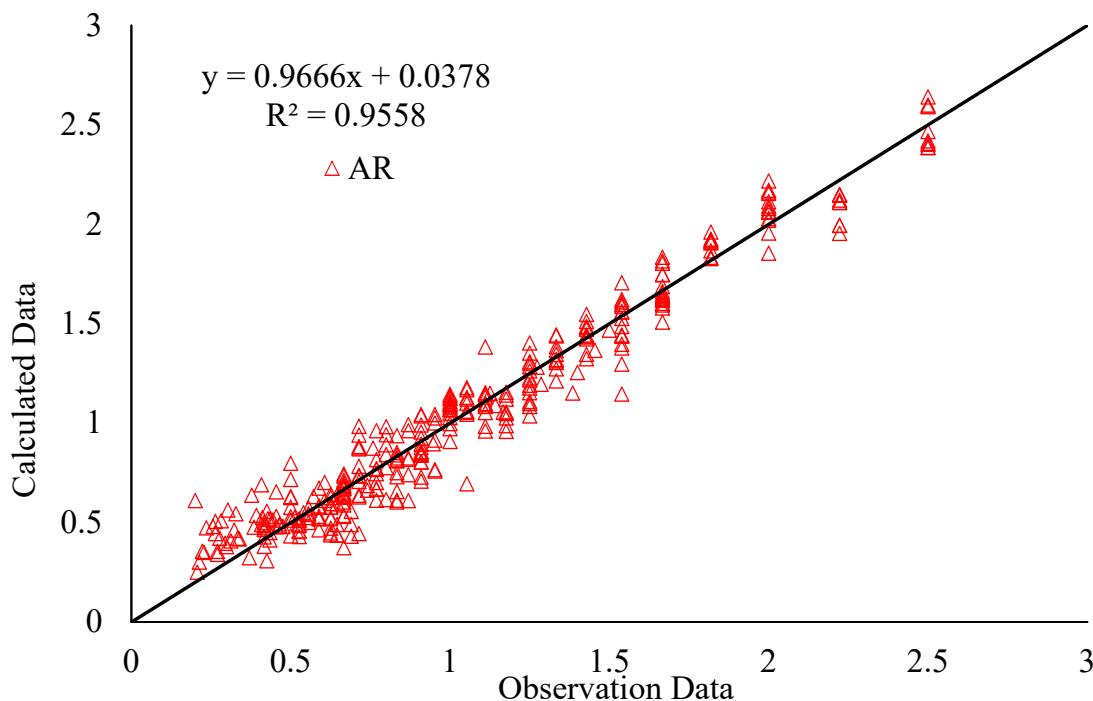


Figure 12. Predicted vs. experimental values of the AR output for the PCA-ABC-ANN model using testing data.

In most published articles on ANN models, it is common for the authors to provide the optimal ANN technique without any data about the proposed ANN weight values. Any structure without the ANN model's final weight values is of little value to experienced researchers and engineers. A proposed ANN architecture must be accompanied

by (quantified) weight values to be useful [63]. After finding a stable and final model in determining the AR output parameter, it is necessary to generate an information table for weights and biases so that the final modeling can be determined at any time, even with new data. Table 9 displays the ABC algorithm's final weight and bias for both hidden layers.

Table 9. Final bias and weight values of the developed PCA-ABC-ANN technique.

IW				b1	
1.0000	0.0048	-1.0000	-0.6979		-0.3815
-0.7198	0.0175	-0.0788	-0.6253		0.9675
-1.0000	0.1828	-0.8656	-0.4221		-0.6311
-1.0000	0.1674	1.0000	0.5205		0.7721
-0.4443	0.3115	1.0000	-0.5144		0.5210
LW1				b2	
-0.8369	0.6940	1.0000	-0.0327	0.1430	0.9120
-0.4156	0.0648	1.0000	0.4767	-0.3320	-0.4661
-0.0545	0.6210	-1.0000	-0.6569	0.3856	0.6196
1.0000	-1.0000	-0.0477	1.0000	0.7499	0.1558
LW2				b3	
0.3997	0.4148	-0.5161	-0.0175		0.1067

IW: Weight values for the input layer, LW1: Weight values for the first hidden layer, LW2: Weight values for the second hidden layer, b1: Bias values for the first hidden layer, b2: Bias values for the second hidden layer, b3: Bias values for the output layer.

3.4. Validation of the Developed Model

Multiple linear regression models were used to validate the ANN-based techniques. In the multiple regression approach, two or more individual variables have a significant influence on the dependent variable (Equation (6)) [79]:

$$y = f(x_1, x_2, \dots) \rightarrow y = a_0 + a_1x_1 + a_2x_2 + \dots \quad (6)$$

where y is the dependent or output parameter; x_1, x_2, \dots are the model inputs or predictors; and a_1, a_2, a_3, \dots are the coefficients related to each input [79–81]. A series of regression equations were investigated for input and output variables. These equations are shown in Table 10 with one to seven independent variables. Among all the equations, the most appropriate coefficients of the multiple linear regression model for the AR output parameter involve an equation with seven parameters presented in Equation (7). Also, the statistical indices for the best multiple linear regression model for all the samples are given in Table 11.

Table 10. Multiple linear regression equations with independent parameters.

Models	R ² %	Parameters	Equation Numbers
AR = 0.6405 – 0.002603 RQD – 0.001639 UCS – 0.006375 RMR – 0.00789 BTS – 0.00058 WZ + 0.002999 TFPC + 0.04854 RPM	91.79	7	(7)
AR = 0.6373 – 0.002600 RQD – 0.001639 UCS – 0.006364 RMR – 0.00788 BTS + 0.003002 TFPC + 0.04858 RPM	91.68	6	(8)
AR = 0.4662 – 0.002889 RQD (%) – 0.007356 RMR – 0.01365 BTS + 0.003159 TFPC + 0.05414 RPM	91.39	5	(9)
AR = 0.4266 – 0.009943 RMR – 0.01539 BTS + 0.003239 TFPC + 0.06152 RPM	90.66	4	(10)
AR = 1.7366 – 0.016681 RMR – 0.02506 BTS + 0.09089 RPM	82.45	3	(11)
AR = 0.5165 – 0.05625 BTS + 0.12752 RPM	73.29	2	(12)
AR = -0.7618 + 0.20618 RPM	67.04	1	(13)

Table 11. Statistical results of bee algorithms, colonial competition, and multiple linear regression model for all samples.

Step	Statistical Index	PCA-ABC-ANN 2L (5-4)	PCA-ICA-ANN 2L (5-4)	MLR 7
All	RMSE	0.11	0.16	0.16
	AAE	0.12	0.17	0.16
	VAF%	96%	92%	92%

For another evaluation of the model, the colonial competition algorithm is combined with an ANN. The colonial competition approach is a random population-based technique inspired by the idea of humans' socio-political evolution. In this algorithm, several colonial countries, together with their colonies, search for finding the best answers to solve problems [82,83]. The initial answers in ICA algorithm are also known as "country". Like ABC, in ICA, these countries try to improve themselves for finding better results/solutions. They are able to search within the search environment similar to the ABC technique. Since the main purpose of this study is not proposing a model based on ICA, the authors decided to stop explaining about this technique herein. The full description of this algorithm is available elsewhere [82,84–86].

For using ICA instead of ABC in the hybrid system, the same ANN with a 4-5-4-1 topology was used and the TBM AR values were predicted through the system. A series of PCA-ICA-ANN models with several values of important parameters in ICA (e.g., No. of countries and No. of imperialists) were built. The mentioned analyses and modeling were conducted using a trial-and-error process. According to the obtained results in this regard, the best parameters of ICA were investigated as presented in Table 12. It can be seen that the best PCA-ICA-ANN model has values of 500, 50 and 250 for the No. of countries, the No. of imperialists and No. of decades, respectively. In addition, RMSE, AAE, and VAF values related to the developed models in this study to estimate TBM AR are presented in Table 11.

Table 12. Characteristics of the imperialist competitive algorithm in the ANN model.

Models' Name	Employed Initialization Parameters in ICA		
	Number of Countries	Number of Imperialists	Number of Decades
PCA-ICA-ANN 2L (5-4)	500	50	250

The results of PCA-ABC-ANN, PCA-ICA-ANN, and MLR models are shown in Figures 13–15 for the AR prediction. According to these figures and Table 11, the results show that the ABC-optimized ANN model has higher accuracy than the ICA-optimized ANN and multiple linear regression models. It was found that by developing the PCA-ICA-ANN model, the performance capacity of the developed MLR equation can be improved for approximately 0.01 of R^2 . It means that the difference between these models is very low and there is no point in setting the PCA-ICA-ANN model by spending much effort and having sufficient AI knowledge. However, in case of the PCA-ICA-ANN model, the story is different. The more significant difference and higher performing capacity can be achieved by proposing this technique. R^2 of 0.9617 was obtained for the selected model and it is obviously better than the other proposed models in this study to predict TBM AR. The Taylor diagram was also used to evaluate and select the best model (Figure 16). According to the diagram, the accuracy of the PCA-ABC-ANN model is higher than the other models. This model is a combination of PCA technique and a hybrid ANN algorithm and can enjoy the advantages of all mentioned models in predicting TBM AR.

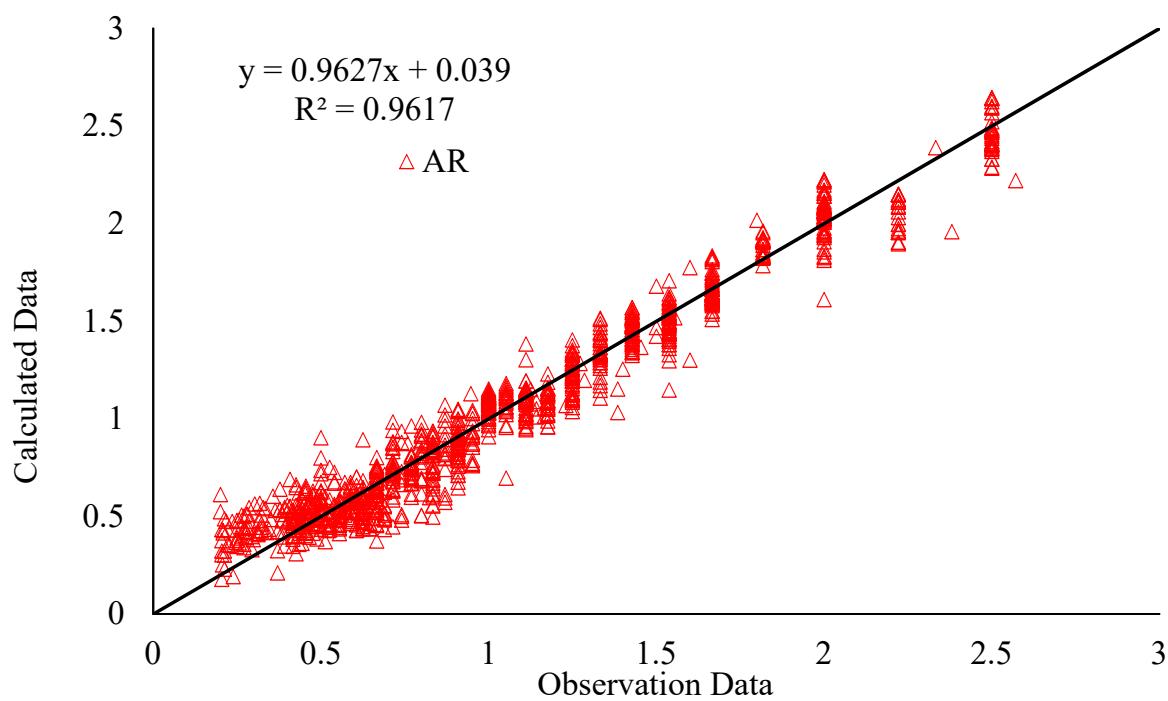


Figure 13. Predicted vs. experimental values of AR output for the PCA-ABC-ANN model using all data.

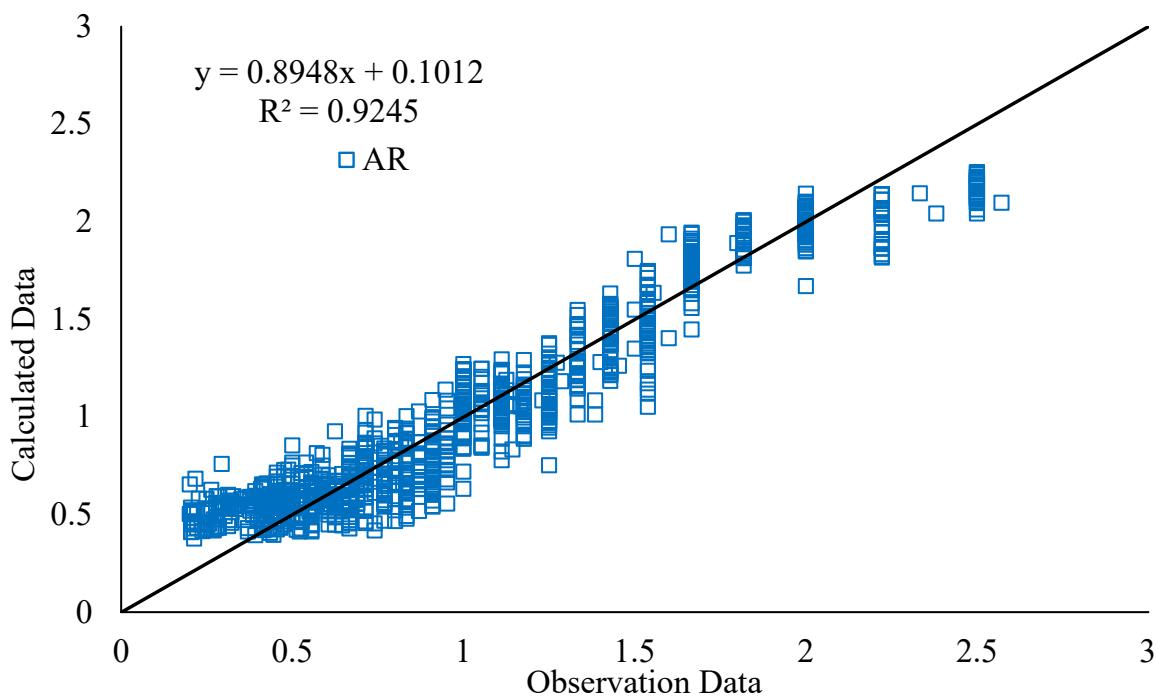


Figure 14. Predicted vs. experimental values of AR output for the PCA-ICA-ANN model using all data.

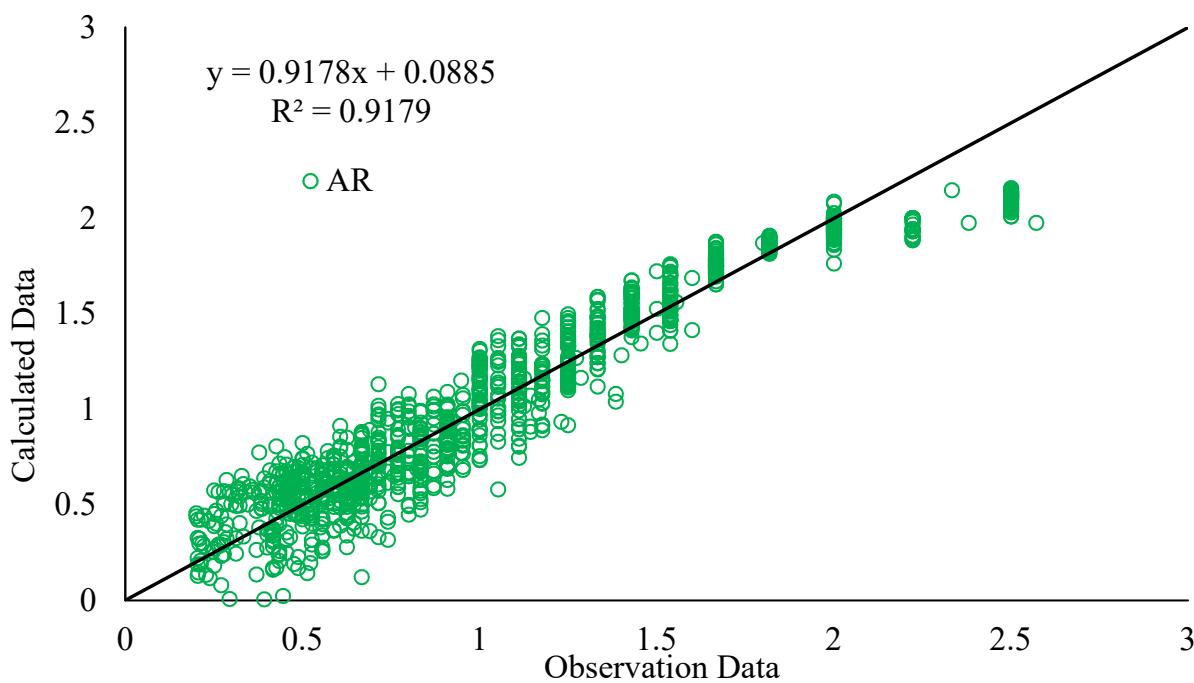


Figure 15. Predicted vs. experimental values of AR output for MLR model using all data.

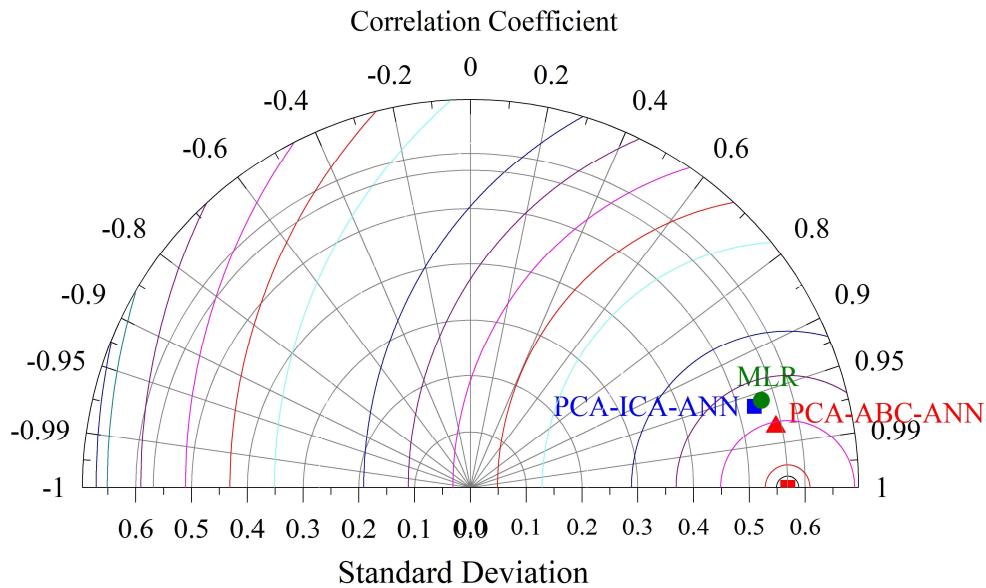


Figure 16. Taylor diagram visualization of the model performance in AR.

4. Sensitivity Analysis (SA)

Considering the PCA-ABC-ANN model exhibited superior performance compared to the PCA-ICA-ANN and regression models, an SA was conducted using the model to identify the relative impact of each predictor parameter on the TBM AR in granitic rock mass. In this way, Lek's proposed profile method [87,88] was applied in MATLAB version 2018 software. The method attempted to analyze each input while considering other values as constants. Details on the theory and implementation of the method are available in the literature [87,88]. In this work, a total number of 1205 data samples were used to analyze the effects of each PCA on the TBM AR. According to Figure 17, all four combined and generated parameters from the input data, including RQD, UCS, RMR, BTS, WZ, TFPC, and RPM, have a desirable effect on the network output. It is important to note that the

results obtained from this analysis are in line with some of the previous studies in this field [89–91].

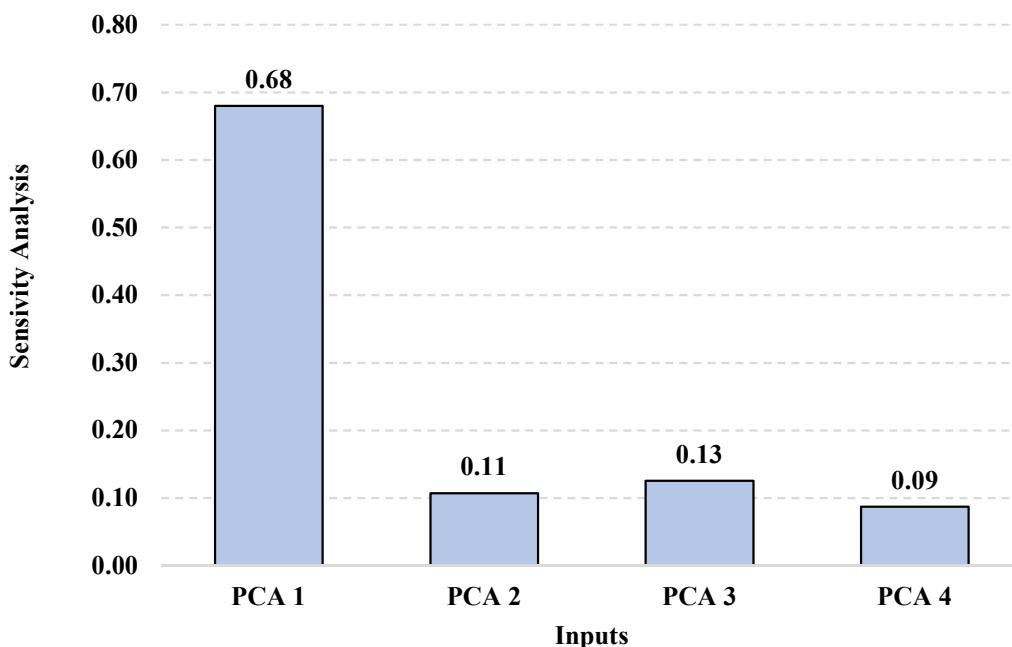


Figure 17. The relative influence of model predictors on the TBM AR.

5. Conclusions

The following conclusion remarks can be extracted from this study:

- Increasing the number of inputs in the ANN increases the number of errors. PCA reduces the inputs, which improves the results in the combined PCA-ANN-based models for determining AR output.
- After using the PCA results to reduce the ANN inputs, in order to optimize the weights of the ANN, the ABC algorithm was used to select the best structure for the ANNs and produce the least errors in the models. For this purpose, 20 models with different topologies were used, and the model with the 4-5-4-1 topology offered acceptable results.
- The PCA-ABC-ANN model combined with the Levenberg–Marquardt learning algorithm and the hyperbolic tangent transfer function were more capable and accurate in predicting TBM AR values. In this model, the R^2 values for TBM AR in the training and test stages were as 0.9641 and 0.9558, respectively, indicating the model's high accuracy. On training data, RMSE, AAE, and VAF% of the PCA-ABC-ANN model for TBM AR values were 0.11, 0.12 and 96%, respectively, while in testing data, they were 0.12, 0.13, and 96%. The statistical indices of VAF, AAE, and RMSE presented in Table 7 indicate the model's negligible error.
- To assess the accuracy of the ABC-optimized technique, it is compared with the ICA algorithm. For the PCA-ICA-ANN model on all data, RMSE, AAE, and VAF% for TBM AR values were 0.16, 0.17, and 92%, respectively.
- The authors have used a statistical model with seven input variables. For the MLR model on all data, RMSE, AAE, and VAF% for TBM AR values were 0.16, 0.16, and 92%, respectively.
- According to the evaluation results, the ABC algorithm received a higher accuracy level than the ICA algorithm followed by the MLR model.
- The modeling procedure introduced in this study regarding reducing the number of inputs using PCA can be implemented in the other similar fields.

- The models in this study were developed for a granitic rock mass, which includes simple geological conditions. Therefore, these models should be used in very similar conditions if very close performance is needed. Of course, the error is higher if different geological conditions are examined.

Author Contributions: Conceptualization, J.W. and A.S.M.; methodology, E.M. and M.A.; software, D.V.U. and Q.F.; validation, J.W., M.A. and E.M.; formal analysis, Q.F. and D.V.U.; investigation, J.W. and Q.F.; resources, A.S.M.; data curation, E.M.; writing—original draft preparation, A.S.M. and M.A.; writing—review and editing, J.W. and E.M.; visualization, Q.F. and D.V.U.; supervision, A.S.M., E.M. and M.A.; project administration, J.W. and E.M.; funding acquisition, E.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data of this research is available on special request from corresponding authors.

Acknowledgments: Project of tackling key problems of science and technology in Henan Province (222102320164); Key Scientific Research Project Plan of Henan Province colleges and universities (22B560009).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Armaghani, D.J.; Mohamad, E.T.; Narayanasamy, M.S.; Narita, N.; Yagiz, S. Development of hybrid intelligent models for predicting TBM penetration rate in hard rock condition. *Tunn. Undergr. Space Technol.* **2017**, *63*, 29–43. [[CrossRef](#)]
2. Zhou, J.; Yazdani Bejarbaneh, B.; Jahed Armaghani, D.; Tahir, M.M. Forecasting of TBM advance rate in hard rock condition based on artificial neural network and genetic programming techniques. *Bull. Eng. Geol. Environ.* **2020**, *79*, 2069–2084. [[CrossRef](#)]
3. Farmer, I.W.; Glossop, N.H. Mechanics of disc cutter penetration. *Tunn. Tunn.* **1980**, *12*, 22–25.
4. Snowdon, R.A.; Ryley, M.D.; Temporal, J. A study of disc cutting in selected British rocks. *Int. J. Rock Mech. Min. Sci. Geomech. Abstracts* **1982**, *19*, 107–121. [[CrossRef](#)]
5. Sanio, H.P. Prediction of the performance of disc cutters in anisotropic rock. *Int. J. Rock Mech. Min. Sci. Geomech. Abstracts* **1985**, *22*, 153–161. [[CrossRef](#)]
6. Rostami, J.; Ozdemir, L. A new model for performance prediction of hard rock TBMs. In Proceedings of the 1993 Rapid Excavation and Tunneling Conference, Boston, MA, USA, 13–17 June 1993; Society for Mining, Metallogy & Exploration, Inc.: Englewood, CO, USA, 1993; p. 793.
7. Yagiz, S. *Development of Rock Fracture and Brittleness Indices to Quantify the Effects of Rock Mass Features and Toughness in the CSM Model Basic Penetration for Hard Rock Tunneling Machines*; Colorado School of Mines: Golden, CO, USA, 2002.
8. Yang, H.; Wang, H.; Zhou, X. Analysis on the rock–cutter interaction mechanism during the TBM tunneling process. *Rock Mech. Rock Eng.* **2016**, *49*, 1073–1090. [[CrossRef](#)]
9. Armaghani, D.J.; Faradonbeh, R.S.; Momeni, E.; Fahimifar, A.; Tahir, M.M. Performance prediction of tunnel boring machine through developing a gene expression programming equation. *Eng. Comput.* **2018**, *34*, 129–141. [[CrossRef](#)]
10. Zhou, J.; Qiu, Y.; Zhu, S.; Armaghani, D.J.; Li, C.; Nguyen, H.; Yagiz, S. Optimization of support vector machine through the use of metaheuristic algorithms in forecasting TBM advance rate. *Eng. Appl. Artif. Intell.* **2021**, *97*, 104015. [[CrossRef](#)]
11. Yang, H.; Liu, J.; Liu, B. Investigation on the cracking character of jointed rock mass beneath TBM disc cutter. *Rock Mech. Rock Eng.* **2018**, *51*, 1263–1277. [[CrossRef](#)]
12. Grima, M.A.; Bruines, P.A.; Verhoef, P.N.W. Modeling tunnel boring machine performance by neuro-fuzzy methods. *Tunn. Undergr. Space Technol.* **2000**, *15*, 259–269. [[CrossRef](#)]
13. Yagiz, S.; Gokceoglu, C.; Sezer, E.; Iplikci, S. Application of two non-linear prediction tools to the estimation of tunnel boring machine performance. *Eng. Appl. Artif. Intell.* **2009**, *22*, 808–814. [[CrossRef](#)]
14. Armaghani, D.J.; Koopalipoor, M.; Marto, A.; Yagiz, S. Application of several optimization techniques for estimating TBM advance rate in granitic rocks. *J. Rock Mech. Geotech. Eng.* **2019**, *11*, 779–789. [[CrossRef](#)]
15. Mahdevari, S.; Shahriar, K.; Yagiz, S.; Shirazi, M.A. A support vector regression model for predicting tunnel boring machine penetration rates. *Int. J. Rock Mech. Min. Sci.* **2014**, *72*, 214–229. [[CrossRef](#)]
16. Oraee, K.; Khorami, M.T.; Hosseini, N. Prediction of the penetration rate of TBM using adaptive neuro fuzzy inference system (ANFIS). In Proceedings of the 2012 SME Annual Meeting & Exhibit 2012 (SME 2012): From Mine to Market, Seattle, WA, USA, 19–22 February 2012; pp. 297–302.

17. Jahed Armaghani, D.; Azizi, A. *Applications of Artificial Intelligence in Tunnelling and Underground Space Technology*; Springer Nature: Singapore, 2021; ISBN 9811610347. [CrossRef]
18. Yagiz, S. New equations for predicting the field penetration index of tunnel boring machines in fractured rock mass. *Arab. J. Geosci.* **2017**, *10*, 33. [CrossRef]
19. Delisio, A.; Zhao, J.; Einstein, H.H. Analysis and prediction of TBM performance in blocky rock conditions at the Lötschberg Base Tunnel. *Tunn. Undergr. Space Technol.* **2013**, *33*, 131–142. [CrossRef]
20. Farrokh, E.; Rostami, J.; Laughton, C. Study of various models for estimation of penetration rate of hard rock TBMs. *Tunn. Undergr. Space Technol.* **2012**, *30*, 110–123. [CrossRef]
21. Grima, M.A.; Verhoef, P.N.W. Forecasting rock trencher performance using fuzzy logic. *Int. J. Rock Mech. Min. Sci.* **1999**, *36*, 413–432.
22. Salimi, A.; Esmaeili, M. Utilising of linear and non-linear prediction tools for evaluation of penetration rate of tunnel boring machine in hard rock condition. *Int. J. Min. Miner. Eng.* **2013**, *4*, 249–264. [CrossRef]
23. Salimi, A.; Faradonbeh, R.S.; Monjezi, M.; Moormann, C. TBM performance estimation using a classification and regression tree (CART) technique. *Bull. Eng. Geol. Environ.* **2018**, *77*, 429–440. [CrossRef]
24. Ghasemi, E.; Yagiz, S.; Ataei, M. Predicting penetration rate of hard rock tunnel boring machine using fuzzy logic. *Bull. Eng. Geol. Environ.* **2014**, *73*, 23–35. [CrossRef]
25. Simoes, M.G.; Kim, T. Fuzzy modeling approaches for the prediction of machine utilization in hard rock tunnel boring machines. In Proceedings of the Conference Record of the 2006 IEEE Industry Applications Conference Forty-First IAS Annual Meeting, Tampa, FL, USA, 8–12 October 2006; Volume 2, pp. 947–954.
26. Benardos, A.G.; Kaliampakos, D.C. Modelling TBM performance with artificial neural networks. *Tunn. Undergr. Space Technol.* **2004**, *19*, 597–605. [CrossRef]
27. Feng, S.; Chen, Z.; Luo, H.; Wang, S.; Zhao, Y.; Liu, L.; Ling, D.; Jing, L. Tunnel boring machines (TBM) performance prediction: A case study using big data and deep learning. *Tunn. Undergr. Space Technol.* **2021**, *110*, 103636. [CrossRef]
28. Adoko, A.C.; Yagiz, S. Fuzzy inference system-based for TBM field penetration index estimation in rock mass. *Geotech. Geol. Eng.* **2019**, *37*, 1533–1553. [CrossRef]
29. Salimi, A.; Rostami, J.; Moormann, C.; Delisio, A. Application of non-linear regression analysis and artificial intelligence algorithms for performance prediction of hard rock TBMs. *Tunn. Undergr. Space Technol.* **2016**, *58*, 236–246. [CrossRef]
30. Koopalipoor, M.; Nikouei, S.S.; Marto, A.; Fahimifar, A.; Armaghani, D.J.; Mohamad, E.T. Predicting tunnel boring machine performance through a new model based on the group method of data handling. *Bull. Eng. Geol. Environ.* **2018**, *78*, 3799–3813. [CrossRef]
31. Minh, V.T.; Katushin, D.; Antonov, M.; Veinthal, R. Regression Models and Fuzzy Logic Prediction of TBM Penetration Rate. *Open Eng.* **2017**, *7*, 60–68. [CrossRef]
32. Zhou, J.; Qiu, Y.; Armaghani, D.J.; Zhang, W.; Li, C.; Zhu, S.; Tarinejad, R. Predicting TBM penetration rate in hard rock condition: A comparative study among six XGB-based metaheuristic techniques. *Geosci. Front.* **2020**, *12*, 101091. [CrossRef]
33. Zeng, J.; Roy, B.; Kumar, D.; Mohammed, A.S.; Armaghani, D.J.; Zhou, J.; Mohamad, E.T. Proposing several hybrid PSO-extreme learning machine techniques to predict TBM performance. In *Engineering with Computers*; Springer: Berlin/Heidelberg, Germany, 2021. [CrossRef]
34. Hajihassani, M.; Abdullah, S.S.; Asteris, P.G.; Armaghani, D.J. A Gene Expression Programming Model for Predicting Tunnel Convergence. *Appl. Sci.* **2019**, *9*, 4650. [CrossRef]
35. Asteris, P.G.; Armaghani, D.J.; Hatzigeorgiou, G.D.; Karayannis, C.G.; Pilakoutas, K. Predicting the shear strength of reinforced concrete beams using Artificial Neural Networks. *Comput. Concr.* **2019**, *24*, 469–488.
36. Apostolopoulou, M.; Armaghani, D.J.; Bakolas, A.; Douvika, M.G.; Moropoulou, A.; Asteris, P.G. Compressive strength of natural hydraulic lime mortars using soft computing techniques. *Procedia Struct. Integr.* **2019**, *17*, 914–923. [CrossRef]
37. Armaghani, D.J.; Hatzigeorgiou, G.D.; Karamani, C.; Skentou, A.; Zoumpoulaki, I.; Asteris, P.G. Soft computing-based techniques for concrete beams shear strength. *Procedia Struct. Integr.* **2019**, *17*, 924–933. [CrossRef]
38. Yang, H.Q.; Xing, S.G.; Wang, Q.; Li, Z. Model test on the entrainment phenomenon and energy conversion mechanism of flow-like landslides. *Eng. Geol.* **2018**, *239*, 119–125. [CrossRef]
39. Yang, H.Q.; Lan, Y.F.; Lu, L.; Zhou, X.P. A quasi-three-dimensional spring-deformable-block model for runout analysis of rapid landslide motion. *Eng. Geol.* **2015**, *185*, 20–32. [CrossRef]
40. Jian, Z.; Shi, X.; Huang, R.; Qiu, X.; Chong, C. Feasibility of stochastic gradient boosting approach for predicting rockburst damage in burst-prone mines. *Trans. Nonferrous Met. Soc. China* **2016**, *26*, 1938–1945.
41. Zhou, J.; Li, E.; Yang, S.; Wang, M.; Shi, X.; Yao, S.; Mitri, H.S. Slope stability prediction for circular mode failure using gradient boosting machine approach based on an updated database of case histories. *Saf. Sci.* **2019**, *118*, 505–518. [CrossRef]
42. Kardani, N.; Bardhan, A.; Samui, P.; Nazem, M.; Zhou, A.; Armaghani, D.J. A novel technique based on the improved firefly algorithm coupled with extreme learning machine (ELM-IFF) for predicting the thermal conductivity of soil. In *Engineering with Computers*; Springer: Berlin/Heidelberg, Germany, 2021. [CrossRef]
43. Zeng, J.; Asteris, P.G.; Mamou, A.P.; Mohammed, A.S.; Golias, E.A.; Armaghani, D.J.; Faizi, K.; Hasanipanah, M. The Effectiveness of Ensemble-Neural Network Techniques to Predict Peak Uplift Resistance of Buried Pipes in Reinforced Sand. *Appl. Sci.* **2021**, *11*, 908. [CrossRef]

44. Momeni, E.; Yarivand, A.; Dowlatshahi, M.B.; Armaghani, D.J. An Efficient Optimal Neural Network Based on Gravitational Search Algorithm in Predicting the Deformation of Geogrid-Reinforced Soil Structures. *Transp. Geotech.* **2020**, *26*, 100446. [[CrossRef](#)]
45. Hasanipanah, M.; Monjezi, M.; Shahnaz, A.; Armaghani, D.J.; Farazmand, A. Feasibility of indirect determination of blast induced ground vibration based on support vector machine. *Measurement* **2015**, *75*, 289–297. [[CrossRef](#)]
46. Asteris, P.G.; Rizal, F.I.M.; Koopialipoor, M.; Roussis, P.C.; Ferentinou, M.; Armaghani, D.J.; Gordan, B. Slope Stability Classification under Seismic Conditions Using Several Tree-Based Intelligent Techniques. *Appl. Sci.* **2022**, *12*, 1753. [[CrossRef](#)]
47. Asteris, P.G.; Lourenço, P.B.; Roussis, P.C.; Adami, C.E.; Armaghani, D.J.; Cavalieri, L.; Chalioris, C.E.; Hajihassani, M.; Lemonis, M.E.; Mohammed, A.S. Revealing the nature of metakaolin-based concrete materials using artificial intelligence techniques. *Constr. Build. Mater.* **2022**, *322*, 126500. [[CrossRef](#)]
48. Armaghani, D.J.; Mamou, A.; Maraveas, C.; Roussis, P.C.; Siorikis, V.G.; Skentou, A.D.; Asteris, P.G. Predicting the unconfined compressive strength of granite using only two non-destructive test indexes. *Geomech. Eng.* **2021**, *25*, 317–330.
49. Parsajoo, M.; Armaghani, D.J.; Mohammed, A.S.; Khari, M.; Jahandari, S. Tensile strength prediction of rock material using non-destructive tests: A comparative intelligent study. *Transp. Geotech.* **2021**, *31*, 100652. [[CrossRef](#)]
50. Asteris, P.G.; Mamou, A.; Hajihassani, M.; Hasanipanah, M.; Koopialipoor, M.; Le, T.-T.; Kardani, N.; Armaghani, D.J. Soft computing based closed form equations correlating L and N-type Schmidt hammer rebound numbers of rocks. *Transp. Geotech.* **2021**, *29*, 100588. [[CrossRef](#)]
51. Plevris, V.; Asteris, P.G. Modeling of masonry failure surface under biaxial compressive stress using Neural Networks. *Constr. Build. Mater.* **2014**, *55*, 447–461. [[CrossRef](#)]
52. Liao, J.; Asteris, P.G.; Cavalieri, L.; Mohammed, A.S.; Lemonis, M.E.; Tsoukalas, M.Z.; Skentou, A.D.; Maraveas, C.; Koopialipoor, M.; Armaghani, D.J. Novel Fuzzy-Based Optimization Approaches for the Prediction of Ultimate Axial Load of Circular Concrete-Filled Steel Tubes. *Buildings* **2021**, *11*, 629. [[CrossRef](#)]
53. Barkhordari, M.S.; Armaghani, D.J.; Mohammed, A.S.; Ulrikh, D.V. Data-Driven Compressive Strength Prediction of Fly Ash Concrete Using Ensemble Learner Algorithms. *Buildings* **2022**, *12*, 132. [[CrossRef](#)]
54. Lee, Y.; Oh, S.-H.; Kim, M.W. The effect of initial weights on premature saturation in back-propagation learning. In Proceedings of the IJCNN-91-Seattle International Joint Conference on Neural Networks, Seattle, WA, USA, 8–12 July 1991; Volume 1, pp. 765–770.
55. Wang, X.; Tang, Z.; Tamura, H.; Ishii, M.; Sun, W.D. An improved backpropagation algorithm to avoid the local minima problem. *Neurocomputing* **2004**, *56*, 455–460. [[CrossRef](#)]
56. Alavi Nezhad Khalil Abad, S.V.; Yilmaz, M.; Jahed Armaghani, D.; Tugrul, A. Prediction of the durability of limestone aggregates using computational techniques. *Neural Comput. Appl.* **2016**, *29*, 423–433. [[CrossRef](#)]
57. Koopialipoor, M.; Jahed Armaghani, D.; Haghghi, M.; Ghaleini, E.N. A neuro-genetic predictive model to approximate overbreak induced by drilling and blasting operation in tunnels. *Bull. Eng. Geol. Environ.* **2019**, *78*, 981–990. [[CrossRef](#)]
58. Nikoo, M.; Hadzima-Nyarko, M.; Karlo Nyarko, E.; Nikoo, M. Determining the natural frequency of cantilever beams using ANN and heuristic search. *Appl. Artif. Intell.* **2018**, *32*, 309–334. [[CrossRef](#)]
59. Bishop, C.M. *Neural Networks for Pattern Recognition*; Oxford University Press: Oxford, UK, 1995; ISBN 0198538642.
60. Haykin, S. *Neural Networks and Learning Machines*, 3rd ed.; Pearson Prentice Hall: Hoboken, NJ, USA, 2009.
61. Murlidhar, B.R.; Armaghani, D.J.; Mohamad, E.T. Intelligence Prediction of Some Selected Environmental Issues of Blasting: A Review. *Open Constr. Build. Technol. J.* **2020**, *14*, 298–308. [[CrossRef](#)]
62. Mohamad, E.T.; Noorani, S.A.; Armaghani, D.J.; Saad, R. Simulation of blasting induced ground vibration by using artificial neural network. *Electron. J. Geotech. Eng.* **2012**, *17*, 2571–2584.
63. Tereshko, V. Reaction-diffusion model of a honeybee colony's foraging behaviour. In *Proceedings of the International Conference on Parallel Problem Solving from Nature*; Springer: Berlin/Heidelberg, Germany, 2000; pp. 807–816.
64. Asteris, P.G.; Nikoo, M. Artificial bee colony-based neural network for the prediction of the fundamental period of infilled frame structures. *Neural Comput. Appl.* **2019**, *31*, 4837–4847. [[CrossRef](#)]
65. Karaboga, D.; Akay, B. A comparative study of artificial bee colony algorithm. *Appl. Math. Comput.* **2009**, *214*, 108–132. [[CrossRef](#)]
66. Zhou, J.; Koopialipoor, M.; Li, E.; Armaghani, D.J. Prediction of rockburst risk in underground projects developing a neuro-bee intelligent system. *Bull. Eng. Geol. Environ.* **2020**, *79*, 4265–4279. [[CrossRef](#)]
67. Le, L.T.; Nguyen, H.; Dou, J.; Zhou, J. A comparative study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for smart city planning. *Appl. Sci.* **2019**, *9*, 2630. [[CrossRef](#)]
68. Jolliffe, I.T. Principal components in regression analysis. In *Principal Component Analysis*; Springer: New York, NY, USA, 1986; pp. 129–155. [[CrossRef](#)]
69. Sadowski, L.; Nikoo, M.; Nikoo, M. Principal component analysis combined with a self organization feature map to determine the pull-off adhesion between concrete layers. *Constr. Build. Mater.* **2015**, *78*, 386–396. [[CrossRef](#)]
70. Ulusay, R.; Hudson, J.A. The complete ISRM suggested methods for rock characterization, testing and monitoring: 1974–2006. In *International Society for Rock Mechanics, Commission on Testing Methods*; ISRM Turkish Natl. Group: Ankara, Turkey, 2007; Volume 628.
71. Madar, V. Direct formulation to Cholesky decomposition of a general nonsingular correlation matrix. *Stat. Probab. Lett.* **2015**, *103*, 142–147. [[CrossRef](#)] [[PubMed](#)]

72. Jiang, B. Covariance selection by thresholding the sample correlation matrix. *Stat. Probab. Lett.* **2013**, *83*, 2492–2498. [[CrossRef](#)]
73. Harandizadeh, H.; Armaghani, D.J. Prediction of air-overpressure induced by blasting using an ANFIS-PNN model optimized by GA. *Appl. Soft Comput.* **2020**, *99*, 106904. [[CrossRef](#)]
74. Harandizadeh, H.; Armaghani, D.J.; Mohamad, E.T. Development of fuzzy-GMDH model optimized by GSA to predict rock tensile strength based on experimental datasets. *Neural Comput. Appl.* **2020**, *32*, 14047–14067. [[CrossRef](#)]
75. Harandizadeh, H.; Armaghani, D.J.; Khari, M. A new development of ANFIS–GMDH optimized by PSO to predict pile bearing capacity based on experimental datasets. *Eng. Comput.* **2021**, *37*, 685–700. [[CrossRef](#)]
76. Armaghani, D.J.; Harandizadeh, H.; Momeni, E. Load carrying capacity assessment of thin-walled foundations: An ANFIS–PNN model optimized by genetic algorithm. In *Engineering with Computers*; Springer: Berlin/Heidelberg, Germany, 2021. [[CrossRef](#)]
77. Bowden, G.J.; Dandy, G.C.; Maier, H.R. Input determination for neural network models in water resources applications. Part 1—background and methodology. *J. Hydrol.* **2005**, *301*, 75–92. [[CrossRef](#)]
78. Li, J.; Heap, A.D. *A Review of Spatial Interpolation Methods for Environmental Scientists*; Australian Gvernement: Canberra, Australia, 2008; ISBN 9781921498305.
79. Nikoo, M.; Torabian Moghadam, F.; Sadowski, Ł. Prediction of concrete compressive strength by evolutionary artificial neural networks. *Adv. Mater. Sci. Eng.* **2015**, *2015*, 849126. [[CrossRef](#)]
80. Mohamad, E.T.; Armaghani, D.J.; Mahdyar, A.; Komoo, I.; Kassim, K.A.; Abdullah, A.; Majid, M.Z.A. Utilizing regression models to find functions for determining ripping production based on laboratory tests. *Measurement* **2017**, *111*, 216–225. [[CrossRef](#)]
81. Gordan, B.; Armaghani, D.J.; Adnan, A.B.; Rashid, A.S.A. A New Model for Determining Slope Stability Based on Seismic Motion Performance. *Soil Mech. Found. Eng.* **2016**, *53*, 344–351. [[CrossRef](#)]
82. Atashpaz-Gargari, E.; Lucas, C. Imperialist competitive algorithm: An algorithm for optimization inspired by imperialistic competition. In Proceedings of the 2007 IEEE Congress on Evolutionary Computation, Singapore, 25–28 September 2007; pp. 4661–4667.
83. Sadowski, Ł.; Nikoo, M. Corrosion current density prediction in reinforced concrete by imperialist competitive algorithm. *Neural Comput. Appl.* **2014**, *25*, 1627–1638. [[CrossRef](#)] [[PubMed](#)]
84. Armaghani, D.J.; Hasanipanah, M.; Amnigh, H.B.; Mohamad, E.T. Feasibility of ICA in approximating ground vibration resulting from mine blasting. *Neural Comput. Appl.* **2018**, *29*, 457–465. [[CrossRef](#)]
85. Armaghani, D.J.; Hasanipanah, M.; Mohamad, E.T. A combination of the ICA-ANN model to predict air-overpressure resulting from blasting. *Eng. Comput.* **2016**, *32*, 155–171. [[CrossRef](#)]
86. Marto, A.; Hajihassani, M.; Jahed Armaghani, D.; Tonnizam Mohamad, E.; Makhtar, A.M. A novel approach for blast-induced flyrock prediction based on imperialist competitive algorithm and artificial neural network. *Sci. World J.* **2014**, *2014*, 643715. [[CrossRef](#)]
87. Lek, S.; Delacoste, M.; Baran, P.; Dimopoulos, I.; Lauga, J.; Aulagnier, S. Application of neural networks to modelling nonlinear relationships in ecology. *Ecol. Modell.* **1996**, *90*, 39–52. [[CrossRef](#)]
88. Gevrey, M.; Dimopoulos, I.; Lek, S. Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecol. Model.* **2003**, *160*, 249–264. [[CrossRef](#)]
89. Jahed Armaghani, D.; Azizi, A. A Comparative Study of Artificial Intelligence Techniques to Estimate TBM Performance in Various Weathering Zones. In *Applications of Artificial Intelligence in Tunnelling and Underground Space Technology*; Springer: Singapore, 2021; pp. 55–70. [[CrossRef](#)]
90. Jahed Armaghani, D.; Azizi, A. Empirical, Statistical, and Intelligent Techniques for TBM Performance Prediction. In *Applications of Artificial Intelligence in Tunnelling and Underground Space Technology*; Springer: Singapore, 2021; pp. 17–32. [[CrossRef](#)]
91. Yang, H.; Wang, Z.; Song, K. A new hybrid grey wolf optimizer-feature weighted-multiple kernel-support vector regression technique to predict TBM performance. In *Engineering with Computers*; Springer: Berlin/Heidelberg, Germany, 2022; Volume 38, pp. 2469–2482. [[CrossRef](#)]