Journal of Rock Mechanics and Geotechnical Engineering 7 (2015) 573–583

Contents lists available at ScienceDirect

Journal of Rock Mechanics and Geotechnical Engineering

journal homepage: www.rockgeotech.org

Full length article

Prediction of roadheaders' performance using artificial neural network approaches (MLP and KOSFM)

Arash Ebrahimabadi [a,*], Mohammad Azimipour [b], Ali Bahreini [c]

[a] Department of Mining, Islamic Azad University, Qaemshahr Branch, Qaemshahr, Iran
[b] Department of Water and Wastewater Engineering, Faculty of Water Engineering and Environment, Shahid Beheshti University, Tehran, Iran
[c] Department of Petroleum, Faculty of Engineering, Science & Research Branch, Islamic Azad University, Tehran, Iran

ARTICLE INFO

Article history:
Received 25 March 2015
Received in revised form
4 June 2015
Accepted 9 June 2015
Available online 4 August 2015

Keywords:
Artificial neural network (ANN)
Performance prediction
Roadheader
Instantaneous cutting rate (ICR)
Tabas coal mine project

ABSTRACT

Application of mechanical excavators is one of the most commonly used excavation methods because it can bring the project more productivity, accuracy and safety. Among the mechanical excavators, roadheaders are mechanical miners which have been extensively used in tunneling, mining and civil industries. Performance prediction is an important issue for successful roadheader application and generally deals with machine selection, production rate and bit consumption. The main aim of this research is to investigate the cutting performance (instantaneous cutting rates (ICRs)) of medium-duty roadheaders by using artificial neural network (ANN) approach. There are different categories for ANNs, but based on training algorithm there are two main kinds: supervised and unsupervised. The multi-layer perceptron (MLP) and Kohonen self-organizing feature map (KSOFM) are the most widely used neural networks for supervised and unsupervised ones, respectively. For gaining this goal, a database was primarily provided from roadheaders' performance and geomechanical characteristics of rock formations in tunnels and drift galleries in Tabas coal mine, the largest and the only fully-mechanized coal mine in Iran. Then the database was analyzed in order to yield the most important factor for ICR by using relatively important factor in which Garson equation was utilized. The MLP network was trained by 3 input parameters including rock mass properties, rock quality designation (RQD), intact rock properties such as uniaxial compressive strength (UCS) and Brazilian tensile strength (BTS), and one output parameter (ICR). In order to have more validation on MLP outputs, KSOFM visualization was applied. The mean square error (MSE) and regression coefficient (R) of MLP were found to be 5.49 and 0.97, respectively. Moreover, KSOFM network has a map size of 8×5 and final quantization and topographic errors were 0.383 and 0.032, respectively. The results show that MLP neural networks have a strong capability to predict and evaluate the performance of medium-duty roadheaders in coal measure rocks. Furthermore, it is concluded that KSOFM neural network is an efficient way for understanding system behavior and knowledge extraction. Finally, it is indicated that UCS has more influence on ICR by applying the best trained MLP network weights in Garson equation which is also confirmed by KSOFM.

1. Introduction

Performance prediction of roadheaders plays a major role in successful application of these machines. Over the last few years, some researchers have made attempts to set up accurate models for predicting roadheaders' cutting performance. This subject is a vital matter, for an accurate evaluation of machine performance can remarkably reduce the job costs and enhance productivity of the project. Predicting performance is an essential task for efficient application of roadheader and is mainly related to evaluation of instantaneous cutting rate (ICR) for various cutting conditions. ICR is defined as the amount of rock excavated per time of cutting (tons or cubic meter/cutting hour). According to Rostami et al. (1994), there are some factors affecting roadheader performance which include: (1) intact rock characteristics, (2) rock mass parameters, (3) machine specifications, and (4) operational parameters.

Sandbak (1985) and Douglas (1985) utilized a rock mass classification to describe variations in roadheaders' advance rates at San Manuel Copper Mine. Gehring (1989) developed equations for predicting cutting rate of axial and transverse type roadheaders as

* Corresponding author. Tel.: +98 11 42155025, +98 11 42155030.
E-mail addresses: Arash.xer@gmail.com, A.Ebrahimabadi@Qaemshahriau.ac.ir (A. Ebrahimabadi).

Peer review under responsibility of Institute of Rock and Soil Mechanics, Chinese Academy of Sciences.

http://dx.doi.org/10.1016/j.jrmge.2015.06.008

$$ICR = \frac{719}{UCS^{0.78}} \tag{1}$$

for transverse type roadheaders, and

$$ICR = \frac{1739}{UCS^{1.13}} \tag{2}$$

for axial type roadheaders, where ICR and UCS are the instantaneous cutting rate (m^3/h) and the uniaxial compressive strength (MPa), respectively. Moreover, Bilgin et al. (1988, 1990, 1996, 1997, 2004) proposed a predictive model as

$$ICR = 0.28 \times 0.974^{RMCI} P \tag{3}$$

$$RMCI = UCS(RQD/100)^{2/3} \tag{4}$$

where P, $RMCI$ and RQD are the cutter head power (kW), the rock mass cuttability index and the rock quality designation (%), respectively. Copur et al. (1997, 1998) investigated the effects of machine's weight and power on cutting performance. Their studies led to a more precise model presented as below:

$$ICR = 27.511e^{0.0023RPI} \tag{5}$$

$$RPI = PW/UCS \tag{6}$$

where RPI is the roadheader penetration index, W is the roadheader weight (t), and e is the base of natural logarithm. Thuro and Plinninger (1999) proposed a predictive model based on specific destruction work (W_z, in kJ/m^3). The specific destruction work is the quantity of energy required for a rock sample to be destroyed. Its predictive equation is presented as follows:

$$ICR = 107.6 - 19.5 \ln W_z \tag{7}$$

Other models are based on specific energy (SE), which is defined as the required energy to cut a unit volume of rock material. One of the most reliable ways to determine the SE is to estimate it from core cutting test. McFeat-Smith and Fowell (1977, 1979) conducted broad laboratory and field investigations and proposed a predictive model based on SE for medium- and heavy-duty roadheaders.

Rostami et al. (1994) developed an accurate model involving cutting power and optimum SE achieved from linear cutting tests in laboratory, as presented below:

$$ICR = k\frac{P}{SE_{opt}} \tag{8}$$

where SE_{opt} is the optimum specific energy, and k is the coefficient of energy transfer. They stated that k changes in 0.45–0.55 for roadheaders and in 0.85–0.9 for tunnel boring machines (Rostami et al., 1994).

As mentioned above, notable attempts have been made to obtain precise prediction models, a brief history of which has been presented. Nevertheless, few investigations on performance prediction of roadheaders have been reported using artificial intelligences such as artificial neural network (ANN), fuzzy logic and neuro-fuzzy in recent years while these approaches have been extensively used in other rock mechanics and rock excavation issues (Grima and Babuska, 1999; Grima et al., 2000; Tiryaki, 2008; Yagiz and Karahan, 2011; Iphar, 2012). In this regard, the main aim of the current research work is to utilize ANN approach for predicting roadheader performance and understanding system behavior. The most extensively used neural network for prediction is multi-layer perceptron (MLP). Another neural network which is mainly utilized for knowledge extraction is Kohonen self-organizing feature map (KSOFM). These networks are discussed in more detail in Section 4. Salsani et al. (2014) employed ANN to model the relationship between the roadheader performance and the factors influencing the tunneling operations. Also, in this research a multiple variable regression (MVR) approach was applied and compared with ANN which has shown that ANN is significantly better (Salsani et al., 2014). It should be noted that the main drawback of previous models is that these models do not reflect the brittleness of rock (UCS/BTS) along with RQD for a given rock mass. In this work, we tried to incorporate these parameters into the model together.

This paper, at first, describes a brief history of roadheaders' performance prediction models and then presents a database established from the field data that consisted of roadheaders' cutting rates and rock properties in the galleries from the Tabas coal mine project (the largest and fully mechanized coal mine in Iran). Then, two different kinds of ANNs are investigated. Finally, using the data, an ANN model is developed for predicting the performance of medium-duty roadheaders and another model is proposed to establish relationships among the process variables.

2. Description of Tabas coal mine project

Tabas coal mine is located in central Iran near the city of Tabas in South Khorasan Province and is situated 75 km far from the southern Tabas (Figs. 1 and 2). The mine region is a portion of Tabas–Kerman coal field. The main part of this coal field is called "Parvadeh" which has the largest coal reserve by nearly 1.1 billion tons within the area of 1,200 km^2. The Parvadeh area is considered to be the main part for current and future coal extraction.

The coal seam expands from east toward west with thinner proportion of the seam. Seam thickness varies from 0.5 m to 2.2 m with the average thickness of 1.8 m. Shallow deposits are mined by room-and-pillar mining method while for deeper deposits, longwall mining method is used. The application of roadheaders was due in part to mechanized coal mining in Tabas coal mine. The mechanized coal mining needs rapid excavation of access galleries in which roadheaders can be employed efficiently. There are four DOSCO MD1100 roadheaders of 34 t in mass, with a 82-kW axial cutting head in order to excavate drift galleries in the mine.

As for the mechanism of rock or coal excavation by roadheaders, it is stated that the breaking process of cutting head can be divided

Fig. 1. General view of the location of Tabas coal mine project.

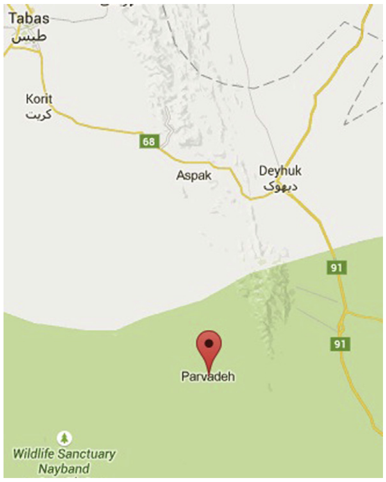

Fig. 2. The location of Tabas coal mine project in Parvadeh region.

into four stages as follows: (1) Stage of plastic deformation. The force of cutting pick on coal or rock increases gradually, leading the stress of coal or rock surrounding the contact to reach the yield limit firstly, resulting in plastic deformation. (2) Stage of cracking. When the tensile stress of coal or rock exceeds their ultimate tensile strength, with a further increase in the force with cutting picks, crack occurs. In this stage, rock strengths (UCS and BTS) play a major role in rock failure. Moreover, for jointed rocks, such as the conditions in Tabas mine, RQD is considered to be a crucial factor to assist rock breakage, leading to a higher cutting rate. (3) Stage of dense nucleus formation. With the crack expanding, broken coal or rock powder moves forward with the cutting picks, and dense nucleus emerges. Then a part of rock powder is ejected from the pick of blade surface, and dense nucleus volume decreases. (4) Stage of coal or rock breaking. With the further interaction of cutting picks and coal or rock, more rock powder becomes nucleus, and nucleus grows up; when the pressure exceeds a certain value, the coal or rock breaks, cutting load on picks decreases instantly, and a leapfrog cutting cycle is completed (Fig. 3).

Figs. 4 and 5 show the picture of roadheader employed and a view of rock formations in the tunnel, respectively. Table 1 lists the basic specifications of DOSCO MD1100 roadheaders (DOSCO Overseas Engineering Ltd.).

As seen in Table 1, a comprehensive database was prepared for 62 cutting cases in galleries and entries of the Tabas mine to be further analyzed to achieve accurate predictive models (Ebrahimabadi et al., 2011a).

3. Previous prediction models in Tabas coal mine

Many investigations on performance prediction of roadheaders have been carried out based on detailed field studies in the Tabas coal mine project (Ebrahimabadi et al., 2011b, 2012). Consequently, several models to predict the performance of roadheaders based on brittleness index were proposed. Rock mass brittleness index (RMBI) was developed to analyze the effect of rock mass properties on roadheaders cutting performance. Findings indicated that RMBI is highly correlated with ICR ($R^2 = 0.98$). Furthermore, through the broad analyses, another model was proposed for predicting pick or bit consumption rates (PCRs) ($R^2 = 0.94$). The prediction models were described as below (Ebrahimabadi et al., 2011a, b, 2012):

$$RMBI = e^{\frac{UCS}{BTS}}\left(\frac{RQD}{100}\right)^3 \tag{9}$$

$$ICR = 30.74RMBI^{0.23} \tag{10}$$

$$PCI = e^{RMBI}\left(\frac{RQD}{100}\right)^3 \tag{11}$$

$$PCR = 45.1PCI^{-0.15} \tag{12}$$

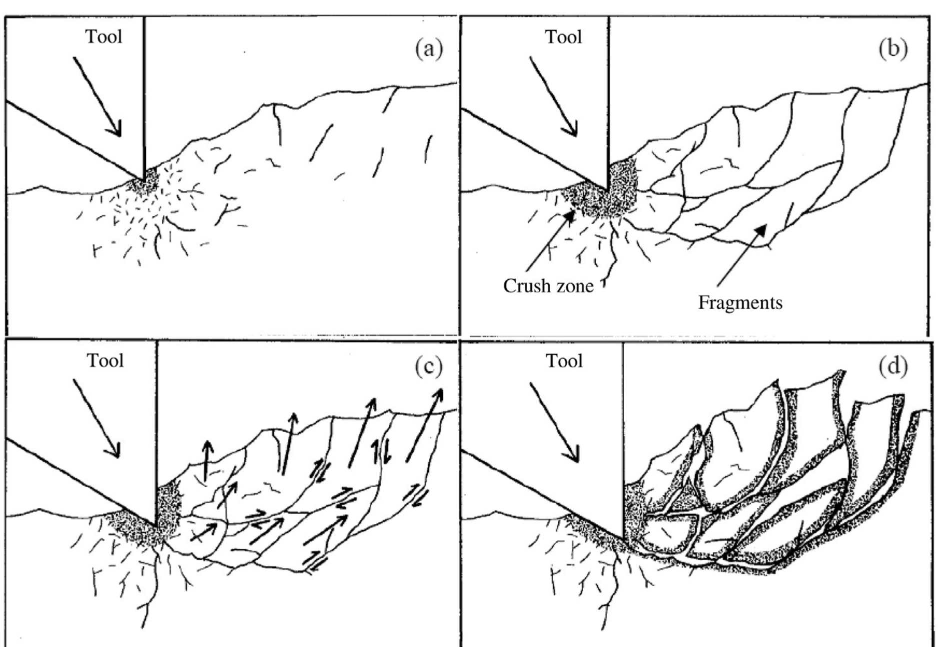

Fig. 3. The evolution of rock (or coal) breakage process.

Fig. 4. DOSCO MD1100 axial type roadheader (DOSCO Overseas Engineering Ltd.).

where *PCI* is the pick consumption index and *P* is equal to 82 kW (DOSCO MD1100 cutter head power).

It should be stated that the models (Eqs. (9–12)) were attained from the analysis of 42 cutting cases. After data collection from other 62 cutting cases, Eq. (10) has been improved as

$$ICR = 9.07 \ln RMBI + 29.93 \tag{13}$$

It should be noted that the aforementioned models have been yielded via statistical modeling. However, in this research, we aim to investigate the relation between the principal rock formation properties and roadheaders' cutting performance through ANN approach. We consider ANN as a powerful and new approach to develop more precise models.

4. Artificial neural network (ANN)

ANNs are mathematical features inspired from humans' brain biology (Mahdevari et al., 2012). They consist of simple interconnected processing elements called neurons under a prespecified topology (layers) (Benardos and Kaliampakos, 2004). Each neuron is connected to its neighbors with varying coefficients called weights (Hamed et al., 2004). The knowledge of an ANN is stored in its weights (Holubar et al., 2002; Hamed et al., 2004). Due to remarkable ability of the neural networks for deriving a general solution for complex systems, they can be used as patterns extraction and trends detection (Yilmaz and Kaynar, 2011). ANNs are classified into several types based on special characteristics. The most popular classification is based on learning algorithm which is the ability of the network to learn from its environments and to improve its performance (Haykin, 1994). The learning algorithm is a dynamic and iterative process which consists of modification of network's parameters in response to the received environmental signals (Moller, 1993). In the majority of cases, learning results in a change in the amount of weights (Khataee and Kasiri, 2011). The goal of learning is to minimize the error between the desired output (target) and network output (output). Two main algorithms are supervised and unsupervised. However, reinforcement learning algorithm is another one which is not studied in this paper. In supervised learning, a teacher that may be a training set of data is required. In this category, training a network includes presenting input and output data. When the network produces required outputs which are close to targets, it is considered to be complete. On the other hand, the network error has the least value. One of the most extensively used training algorithms is back-propagation as the workhorse of ANNs (Rumelhart et al., 1986). This algorithm has two phases (Meulenkamp and Grima, 1999): (1) Presenting an input pattern and calculating the output by network; and (2) Calculating the error and propagating backward this error from the outputs to inputs and finally updating the weights.

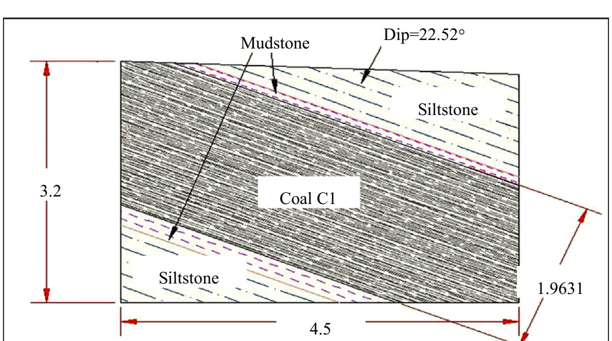

Fig. 5. Schematic cross-section of tunnels (all dimensions are in meter).

In unsupervised learning, despite the external influences on adjusting weights, there is an internal monitoring of performance (Haykin, 1994). The network uses this kind of algorithm to look for regularities or trends in the input signal. To perform unsupervised learning, a competitive learning rule may be utilized in which the wining neuron or the best matching unit will be chosen with "winner-takes-all" strategy (Haykin, 1994; Rustum et al., 2008).

For each training algorithm, there is a specified neural network. The most popular and widely used neural network that uses the supervised learning algorithm is MLP. Generally, this network contains three layers: input, hidden and output layers. The numbers of hidden neurons are determined in a trial-and-error approach. However, the numbers of neurons of input and output layers depend on the nature of the problem (Salari et al., 2005). Due to universal approximation theory, a network with a single hidden layer with a sufficiently large number of neurons can be used for any input–output structure interpretation (Salari et al., 2005) and it is sufficient for most related issues (Rahmanian et al., 2011). In an MLP network, the process of producing output has 4 phases: (1) entering the inputs, (2) multiplying inputs by weights (the result of this stage is called net), (3) passing through the net to transfer function, and (4) producing the outputs. This process is accomplished in hidden layer neurons. After that, the hidden neurons outputs are fed to output layer and the same process will be

Table 1
Typical specifications of DOSCO MD1100 roadheaders (DOSCO Overseas Engineering Ltd.).

Machine mass (base machine) (kg)	Total power (standard machine) (kW)	Power on cutting boom (standard machine) (kW)	Hydraulic system working pressure (MPa)	Tracking speeds—sumping/flitting (m/s)	Ground pressure (kg/m^2)	Machine size (m)		
						Length	Width	Height
34,000	From 157	82 (axial), 112 (transverse)	14	0.038/0.12	14,000	8.06	3	1.7

applied. To evaluate the performance of the models and find the best fitted one, various indices are proposed while in this survey mean square error (MSE) and R (regression coefficient) were chosen as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(T_i - O_i)^2 \tag{14}$$

$$R = \frac{\sum_{i=1}^{N}(T_i - \overline{T}_i)\left(O_i - \overline{O}_i\right)}{\sqrt{\sum_{i=1}^{N}(T_i - \overline{T}_i)^2 \sum_{i=1}^{N}\left(O_i - \overline{O}_i\right)^2}} \tag{15}$$

where T_i and O_i indicate experimental and network outputs, respectively; $\overline{T}_i$ and $\overline{O}_i$ are the average of mentioned data, respectively; N is the total number of data.

For unsupervised learning, there are two main models: KSOFM and Willshaw-von der Malsburg. KSOFM model has received more attention and therefore has been chosen in this paper. KSOFM is one of the most powerful and practical networks (Kohonen et al., 1996) which can be used at the same time to highlight groups of elements with similar characteristics (Grieu et al., 2006), and nonlinear interpolation and extrapolation (Kohonen et al., 1996), to reduce the amount of data by clustering, and to project the data nonlinearity onto a lower dimensional display (Hong et al., 2003). The term 'self-organizing' indicates the ability of learning and organizing information without applying the corresponding output values for the input patterns (Mukherjee, 1997). This neural network combines an input layer with a competitive one (Grieu et al., 2006; Ghaseminezhad and Karami, 2011) in which the input vectors that have n components ($x_1, x_2, \ldots, x_n$) are connected to each neuron (n) by a synaptic weight. Therefore, each neuron is represented by an n-dimensional weight $\boldsymbol{w}_i = [w_{i1}, \ldots, w_{in}]^{\mathrm{T}}$ (Hong et al., 2003; Ghaseminezhad and Karami, 2011). These weight vectors are also named prototype vectors (Garcia and Gonzalez, 2004) or reference vectors (Hong et al., 2003). In a self-organizing map, the neurons are placed at the nodes of a lattice that is usually one- or two-dimensional which is more common than higher dimensional maps (Haykin, 1994). This lattice of neurons called map can be illustrated as a rectangular, hexagonal or even irregular organization (Heikkinen et al., 2011). Due to better presentation of connections between neighboring neurons, the hexagonal one is more preferable (Hong et al., 2003; Heikkinen et al., 2011). Fig. 6 shows two illustrations of KSOFM map. Depending on the required application and details, size of the map (number of neurons) is varied. The number of neurons may vary from a few up to thousands which determines the mapping granularity and affects the accuracy and generalization capability of KSOFM (Llorens et al., 2008). To find the side length of any map, heuristic formula is applied which is proposed as follows (Garcia and Gonzalez, 2004):

$$M = 2\sqrt{N} \tag{16}$$

where M is the number of map units. In this paper, 62 samples were used. Therefore, there should be at least 39 map units.

There are two algorithms for training this kind of neural network: (1) sequential training in which samples are presented to the map one at a time, and the algorithm gradually moves the weight vectors towards them, and (2) batch training in which the data set is presented to the self-organizing map as a whole, and the new weight vectors are evaluated as the average of the data vectors (Vesanto, 1999). Training process depends on the attained errors. If the error is acceptable, training will stop. Two error criteria are normally used: the quantization error and the topographical error, as defined as follows (Garcia and Gonzalez, 2004; Rustum et al., 2008):

$$q_{\mathrm{e}} = \frac{1}{N}\sum_{i=1}^{N}||x_i - m_{\mathrm{c}}|| \tag{17}$$

$$t_{\mathrm{e}} = \frac{1}{N}\sum_{i=1}^{N}u(x_i) \tag{18}$$

where q_{e} is the quantization error; x_i is the i-th data sample or vector; m_{c} is the prototype vector of the best matching unit for x_i; $||\bullet||$ denotes the distance; t_{e} is the topographical error; and u is the binary integer, which is equal to 1 if the first and second best matching units for the argument of u are not adjacent units on the map, otherwise, $u = 0$.

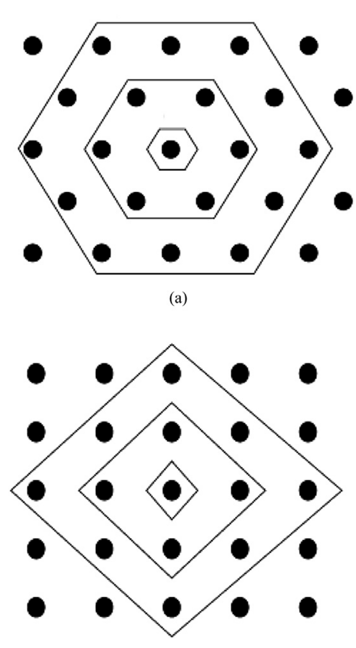

Fig. 6. (a) Hexagonal lattice, and (b) rectangular lattice (Vesanto, 1999).

Table 2
Rock geomechanical characteristics and roadheaders' performance for all cutting cases (Ebrahimabadi et al., 2011a, b, 2012).

Case No.	Representative UCS (MPa)	Representative BTS (MPa)	RQD (%)	Measured ICR (m^3/h)	Case No.	Representative UCS (MPa)	Representative BTS (MPa)	RQD (%)	Measured ICR (m^3/h)
1	14.8	3.8	19	22.2	32	14.7	3.8	19	15.7
2	15.2	3.9	19	25.3	33	17	4.2	20	18.7
3	15.2	3.8	20	24.8	34	15.7	4	19	16.8
4	15.4	4	19	23.7	35	16	3.9	20	18.3
5	15.3	3.9	19	23.2	36	16.2	3.8	21	26.4
6	15	3.9	19	22.8	37	16.7	3.9	22	25.3
7	15.6	3.9	20	27.1	38	17.2	3.9	22	28.5
8	14.5	3.7	20	25.7	39	17.3	4.2	21	25.4
9	16.2	4.2	18	25.6	40	19.2	4.5	24	29.4
10	15.4	4	18	20.2	41	15	3.9	19	22.4
11	16.9	4.2	20	28.2	42	21	3.8	20	36.4
12	14.3	3.8	19	24.4	43	22.3	3.9	20	37.7
13	15.5	4	19	26.4	44	25.6	4.2	19	40.4
14	17.2	4.1	20	25.7	45	26.7	4.3	19	41.1
15	23.9	4.7	27	41.5	46	26.7	4.3	19	41.1
16	27.2	5.3	28	46.2	47	27.2	4.3	19	41.4
17	20.1	4.5	23	32.2	48	27.6	4.4	19	41.6
18	14.1	3.6	19	16.7	49	28	4.4	19	41.8
19	15	3.8	20	17.4	50	19.2	3.7	21	34
20	14.4	3.8	19	16.8	51	25.1	4.1	19	40
21	14.8	3.9	18	17	52	27	4.3	19	41.3
22	14.7	3.8	19	17.5	53	27.1	4.3	19	41.3
23	15.7	4	19	16.8	54	27.4	4.3	19	41.5
24	16.4	4.3	18	16.7	55	27.5	4.3	19	41.5
25	15.1	3.9	19	16.1	56	27.6	4.3	19	41.6
26	14.5	3.7	19	17.7	57	27.7	4.4	19	41.7
27	15.1	3.9	19	16	58	27.9	4.4	19	41.8
28	15.2	4	19	17	59	28	4.4	19	41.8
29	14.4	3.7	19	14.6	60	28.1	4.4	19	41.9
30	15.6	3.9	20	19	61	27.9	4.4	19	41.8
31	14.5	3.7	18	17.7	62	28.2	4.4	19	41.9

5. ANN results

5.1. Multi-layer perceptron (MLP)

To find the best MLP network, the following stages should be carried out:

(1) Defining and selecting inputs and output(s) based on dataset.
(2) Preprocessing data. In this work data have fallen in the range of [−1, 1] by scaling them with respect to the minimum and the maximum of all the data (Mjalli et al., 2007).
(3) Dividing data into three subsets: training, validation and test. In this research, 60%, 20% and 20% of data were used for training, validation and test, respectively. The main reason of this division is to overcome overfitting which is one of the most important issues of neural networks (Rafiai and Jafari, 2011). This method is called early-stopping. To avoid random correlation, these subsets were randomly selected from all the data (Badalians Gholikandi et al., 2014).
(4) Creating different networks by varying numbers of hidden neurons and activation functions. The numbers of hidden neurons were varied from 1 to 20, and for activation functions, hyperbolic tangent and logarithmic sigmoids were taken for hidden layer; besides, linear function was utilized for the output layer. This indicates that for each neuron 6 networks with different combinations of activation functions and totally 120 networks were created, trained and tested (Badalians Gholikandi et al., 2014).
(5) Evaluating the performance of created networks by R and MSE. It should be noticed that, to calculate MSE, an inverse range scaling must be performed to return network outputs to their original scale which is necessary for comparing outputs and targets (Salari et al., 2005).

In this study, Levenberg–Marquardt (L–M) algorithm was used for training. This algorithm is a combination of Gauss–Newton algorithm and the steepest descent method, and it inherits the speed advantage of the former and the stability of the latter. Moreover, L–M algorithm is 10–100 times faster than usual gradient descent back-propagation method (Pendashteh et al., 2011) and proved to be the fastest and most robust algorithm (Charalambous, 1992).

In this study, Matlab software was used for simulation and modeling. Input parameters are UCS, BTS and RQD and output is ICR (Table 2). Fig. 7 represents the structure of defined networks. The best model which has the lowest MSE ($MSE = 5.49$) for all data was obtained from 3:14:1 structure (14 hidden neurons). The activation functions of this model were hyperbolic tangent (tansig) for both hidden and output layers. The MSE value and correlation

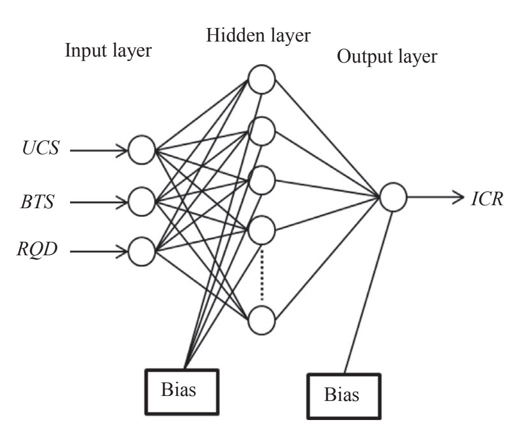

Fig. 7. The structure of MLP network.

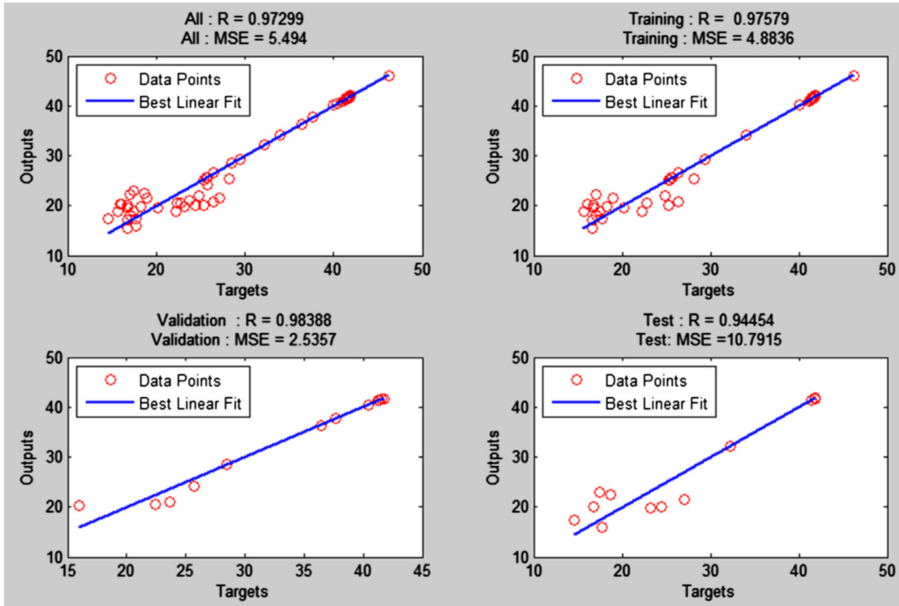

Fig. 8. *R* and *MSE* for all data and different data subsets.

(R) between outputs versus targets for all data and different data subsets are presented in Fig. 8. Fig. 9 shows the reduction of MSE of the best MLP model during learning. As it seems, the epoch (each iteration of learning) in which the best network is obtained is 39. The reason for stopping training network is validation checks where if the error of this data subset increases (for 6 times in this study), the training will be stopped and the weights of that network will be stored as the best network weights. Fig. 10 presents the comparison of the predicted values (outputs) versus measured values (targets), which shows that the selected network has a good level of accuracy for prediction of roadheader performance in most cases.

5.2. Relative importance

To find the relative importance of input variables for ICR, Garson equation (Eq. (19)) was applied which uses the weights of the best trained MLP network (Elmolla et al., 2010):

$$I_j = \frac{\sum_{m=1}^{N_h}\left[\left(\left|W_{jm}^{\mathrm{ih}}\right|\sum_{k=1}^{N_i}\left|W_{km}^{\mathrm{ih}}\right|\right)\left|W_{mn}^{\mathrm{ho}}\right|\right]}{\sum_{k=1}^{N_i}\left(\sum_{m=1}^{N_h}\frac{\left|W_{km}^{\mathrm{ih}}\right|}{\sum_{k=1}^{N_i}\left|W_{km}^{\mathrm{ih}}\right|}\left|W_{mn}^{\mathrm{ho}}\right|\right)} \tag{19}$$

where I_j is the relative importance of the j-th input variable for the output variable; N_i and N_h are the numbers of input and hidden neurons, respectively; and W is used for connection weights. The superscripts "i", "h" and "o" are for input, hidden and output layers, respectively, and the subscripts "k", "m" and "n" refer to input, hidden and output neurons, respectively (Delnavaz et al., 2010). Table 3 shows the best ANN model weights which were used in Eq. (19).

Table 4 shows the influence of each input variable which indicates that UCS has more influence on ICR. This result is the same

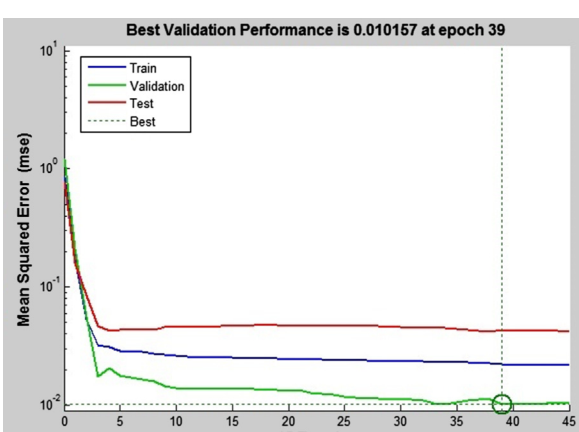

Fig. 9. MSE of the best MLP model during learning.

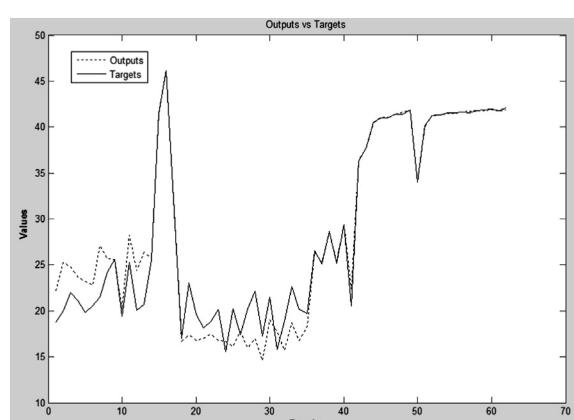

Fig. 10. Comparison of the predicted (outputs) versus measured (targets) values.

Table 3
Weights between input and hidden layers (W1) as well as weights between hidden and output layers (W2).

Neuron	W1, input variables				W2, output variables	
	UCS (MPa)	BTS (MPa)	RQD (%)	Bias	ICR (m^3/h)	Bias
1	1.9993	6.8575	−3.1087	6.9223	−2.117	1.4695
2	2.2266	1.0832	7.0414	−5.6436	0.8853	
3	−1.8969	−4.2569	4.878	5.3685	1.0379	
4	2.854	−2.3679	−6.3068	−3.2395	−1.6938	
5	−6.0247	3.4367	0.2686	3.4794	−1.8793	
6	−3.0023	−5.2419	−3.6634	2.1553	−0.4932	
7	−0.575	−4.9642	−6.2603	−0.1682	−1.7332	
8	0.8073	9.6966	−6.634	−2.5497	3.1359	
9	1.0488	3.9225	−6.8637	−0.406	0.8101	
10	−5.1908	−6.0878	−1.534	−0.5045	2.5344	
11	−3.8754	7.6857	−1.6847	−3.0635	−6.3317	
12	4.367	−7.913	7.0525	2.6689	−2.4061	
13	7.5232	−2.4736	2.4903	3.3329	1.7906	
14	2.705	−3.4497	7.6803	4.1297	3.3819	
15	2.5573	0.6353	6.7818	5.3491	−2.5927	
16	3.5321	−5.8252	1.0652	7.3234	1.3052	

as that in Salsani et al. (2014) in which sensitivity analysis was conducted to find the most effective parameter for ICR.

5.3. *Discussion on the effects of rock parameters on roadheaders' cutting performance*

To have a better understanding of system, three-dimensional figures were produced by the obtained best MLP network. For this purpose, two parameters were changed simultaneously and the other one was constant. The values of the parameters were considered as their average values. The average values of UCS, BTS and RQD are 19.6 MPa, 4 MPa and 19.7 MPa, respectively.

5.3.1. *The simultaneous effect of UCS and BTS*

As Fig. 11 shows, ICR increases with any increase in UCS and this is against the previous findings obtained from other researches. The relation between ICR and BTS also shows some variations, not allowing a consistent trend to be recognized. These results demonstrate that it is necessary to consider rock mass parameters, such as RQD, to gain more precise and realistic results.

5.3.2. *The simultaneous effect of UCS and RQD*

Considering aforementioned results and due to high degree of jointing and fracturing of existing rock formations, RQD would be expected to have a crucial effect on cutting performance. On the other hand, it is clearly known that UCS has a remarkable impact on ICR. Hence, it is highly recommended to simultaneously investigate both UCS and RQD to gain more accurate and realistic results. With this respect, the simultaneous effect of UCS and RQD was studied. By referring to Fig. 12, ICR increases as UCS increases, but the trend shows lower RQD; it means that high degree of jointing has a great impact on rock mass strength, leading to a significant rise in ICR, as demonstrated in Fig. 12. Furthermore, the vagueness in the relation between ICR and UCS was clarified.

Table 4
Relative importance of input variables.

Input variables	Importance (%)
UCS	39.97
BTS	28.64
RQD	31.39
Total	100

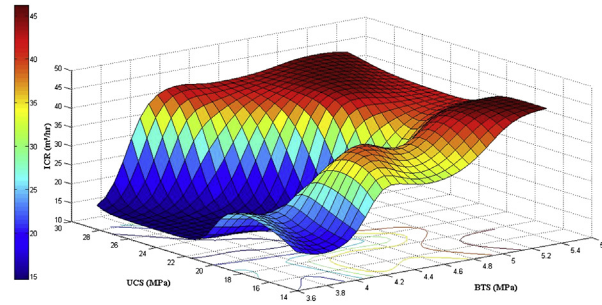

Fig. 11. Simultaneous effect of UCS and BTS.

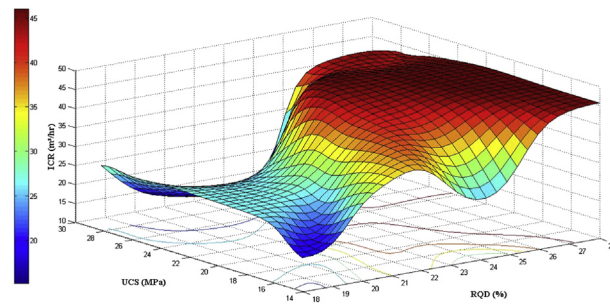

Fig. 12. Simultaneous effect of UCS and RQD.

5.3.3. *The simultaneous effect of BTS and RQD*

Fig. 13 represents the simultaneous effect of BTS and RQD on ICR. It shows that ICR increases with any increase in RQD and decreases as BTS increases along with some variations. It can be concluded that simultaneous consideration of BTS and RQD is not enough to give a sensible outcome for predicting ICR. Therefore, it is extremely recommended that UCS of rock should be taken into account through the analyses and interpretations.

5.4. *Kohonen self-organizing feature map (KSOFM)*

To find the regularities and correlations between input variables, the self-organizing feature map toolbox (CIS, 2005) has been utilized and applied in Matlab. Due to different magnitudes of inputs, all variables were normalized in which they were scaled with the variance of 1 and the mean of 0 (Cinar, 2005; Badalians Gholikandi et al., 2014). The map size during training was 8 × 5 which is sufficient by considering Eq. (16). Final quantization and topographic errors were 0.383 and 0.032, respectively.

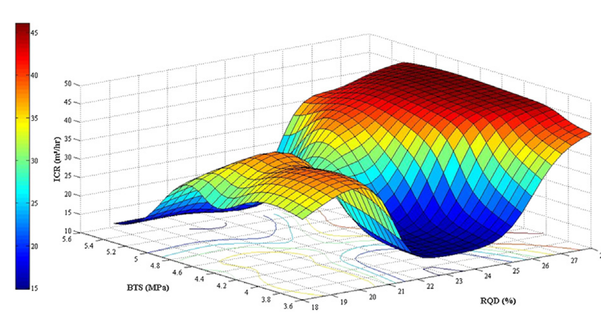

Fig. 13. Simultaneous effect of BTS and RQD.

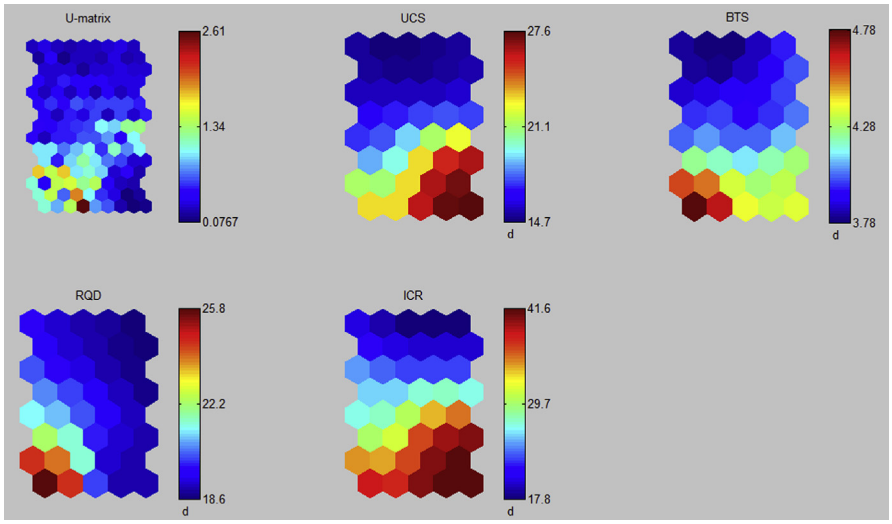

Fig. 14. Component planes of variables and U-matrix.

One of the most important features of KSOFM neural network is its visualization which plays an important role in analyzing results based upon the knowledge extraction (Hong et al., 2003). Two different kinds of visualizations are (Grieu et al., 2006): (1) component planes and scatter plots for determining relationships among process variables and (2) distance matrices and K-means algorithm for clustering visualization. To study the correlation and associations between variables, component planes were utilized in which each hexagon represents one map node and its colors tell the value of the component in that node (Hong et al., 2003). In each component, each hexagon corresponds to the same one in other component planes (Vesanto et al., 1999; Hong et al., 2003; Grieu et al., 2006). On the right side of each plane, there is a legend which indicates the range of each variable. Fig. 14 illustrates the component planes. In this figure, the U-matrix plane indicates the distances between neighboring data units (Lee and Scholz, 2006) in which the high values represent a cluster border and the areas of the low values represent clusters themselves (Tobiszewski et al., 2012).

By considering the component planes of UCS and ICR in Fig. 14, it can be indicated that UCS has more effect on ICR due to the similarity of their components. By looking at lower part of their components, it is noticed that by increasing UCS, ICR will also increase. This result is the same as that in Table 4 obtained by relative importance and has shown that UCS has more effect on ICR. As mentioned before, ICR increases with any increase in UCS; hence, there is a need to involve rock mass characteristics in the analyses. And that is the reason why RQD was taken into account. From Figs. 12 and 13, it can be seen that ICR increases with any increase in RQD. It is also contrary to previous findings. With this respect, it is true in higher values of RQD, while in lower values such as values in this study, this parameter also contributes to the increase in ICR. In this particular situation, as soon as the tips of pick hit the rock, a significant loosening in blocky rocks will be induced and it can facilitate rock removal instead of crushing and cutting the rock.

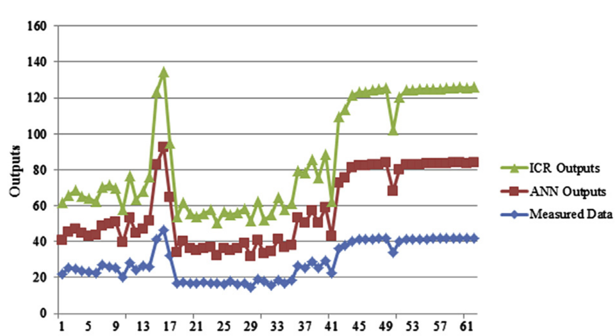

Fig. 15. Comparison of predicted ICR values achieved from previous model (ICR outputs) and new ANN model (ANN outputs) with measured data.

6. Comparison of previous and new ANN models

In order to represent the contribution of the new ANN model, as well as verifying the results of this model for a site-specific project (Tabas coal mine), the predicted ICR values achieved from previous model (Eq. (13)) and new ANN model are compared with measured values, as shown in Fig. 15.

As Fig. 15 shows, there is a slight difference between ANN results and measured data than values obtained from previous model. The MSEs for previous model and the new ANN model are 8.16 and 5.494, respectively. This result demonstrates that the ANN model has a higher accuracy to predict ICR of medium-duty roadheaders.

7. Conclusions

In this paper, a roadheader performance prediction model was developed using ANN approach in which the ICR can be predicted more precisely. MLP and KSOFM are the most widely used neural networks for supervised and unsupervised ones, respectively. To gain this goal, a database was primarily provided from roadheaders' performance and geomechanical characteristics of rock formations in tunnels and drift galleries in Tabas coal mine. Then the database was analyzed in order to yield the most important factor for ICR by using relative importance factor. It was indicated that UCS has more influence on ICR by applying the best trained MLP network weights in Garson equation. Moreover, from practical point of view and merely for such RQD values (less than

28%), ICR increases as RQD increases. The network was trained by 3 input parameters including rock mass properties, RQD, intact rock properties such as UCS and BTS, and one output parameter, ICR. This network has one hidden layer with 14 neurons and uses tansig as activation functions for both hidden and output layers. Besides, to have more validation on MLP outputs, KSOFM visualization was applied. MSE and regression coefficient (R) of MLP were found to be 5.49 and 0.97, respectively. Moreover, KSOFM network has a map size of 8 × 5 and final quantization and topographic errors were 0.383 and 0.032, respectively. The results showed that the neural network had a strong capability to predict and evaluate the performance of medium-duty roadheaders in coal measure rocks.

Conflict of interest

The authors wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References

Badalians Gholikandi G, Masihi H, Azimipour M, Abrishami A, Mirabi M. Optimizing stabilization of waste-activated sludge using Fered–Fenton process and artificial neural network modeling (KSOFM, MLP). Environmental Science and Pollution Research 2014;21(11):7177–86.
Benardos AG, Kaliampakos DC. Modelling TBM performance with artificial neural networks. Tunnelling and Underground Space Technology 2004;19(6):597–605.
Bilgin N, Dincer T, Copur H, Erdogan M. Some geological and geotechnical factors affecting the performance of a roadheader in an inclined tunnel. Tunnelling and Underground Space Technology 2004;19(6):629–36.
Bilgin N, Kuzu C, Eskikaya S, Özdemir L. Cutting performance of Jack hammers and roadheaders in Istanbul Metro drivages. In: Proceedings of the Word Tunnel Congress'97, Tunnels for People. Rotterdam: A.A. Balkema; 1997. p. 455–60.
Bilgin N, Seyrek T, Shahriar K. Roadheader performance in Istanbul, Golden Horn clean-up contributes valuable data. Tunnels and Tunnelling 1988;20(6):41–4.
Bilgin N, Seyrek T, Erdinc E, Shahriar K. Roadheaders glean valuable tips for Istanbul Metro. Tunnels and Tunnelling 1990;22(10):29–32.
Bilgin N, Yazici S, Eskikaya S. A model to predict the performance of roadheaders and impact hammers in tunnel drivages. In: Proceedings of Eurock'96. Rotterdam: A.A. Balkema; 1996. p. 715–20.
Charalambous C. Conjugate gradient algorithm for efficient training of artificial neural networks. Circuits, Devices and Systems 1992;139(3):301–10.
Cinar O. New tool for evaluation of performance of wastewater treatment plant: artificial neural network. Process Biochemistry 2005;40(9):2980–4.
Copur H, Ozdemir L, Rostami J. Roadheader applications in mining and tunneling. Mining Engineering 1998;50:38–42.
Copur H, Rostami J, Ozdemir L, Bilgin N. Studies on performance prediction of roadheaders based on field data in mining and tunneling projects. In: Proceedings of the 4th International Symposium on Mine Mechanization and Automation. Brisbane: Australia; 1997. A4-1–A4-7.
Delnavaz M, Ayati B, Ganjidoust H. Prediction of moving bed biofilm reactor (MBBR) performance for the treatment of aniline using artificial neural networks (ANN). Journal of Hazardous Materials 2010;179(1–3):769–75.
Douglas W. Roadheaders open new horizons at San Manuel. Engineering and Mining Journal 1985;186:22–5.
Ebrahimabadi A, Goshtasbi K, Shahriar K, Cheraghi Seifabad M. Predictive models for roadheader's cutting performance in coal measure rocks. Yerbilimleri 2011a;32(2):89–104.
Ebrahimabadi A, Goshtasbi K, Shahriar K, Cherghi Seifabad M. A model to predict the performance of roadheaders based on rock mass brittleness index. Journal of the South African Institute of Mining and Metallurgy 2011b;111(5):355–64.
Ebrahimabadi A, Goshtasbi K, Shahriar K, Cheraghi Seifabad M. A universal model to predict roadheaders' cutting performance. Archives of Mining Sciences 2012;57(4):1015–26.
Elmolla ES, Chaudhuri M, Eltoukhy MM. The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. Journal of Hazardous Materials 2010;179(1–3):127–34.
Garcia H, Gonzalez I. Self-organizing map and clustering for wastewater treatment monitoring. Engineering Applications of Artificial Intelligence 2004;17(3):215–25.
Gehring KH. A cutting comparison. Tunnels and Tunnelling 1989;21(11):27–30.
Ghaseminezhad MH, Karami A. A novel self-organizing map (SOM) neural network for discrete groups of data clustering. Applied Soft Computing 2011;11(4):3771–8.
Grieu S, Thiery F, Traore A, Nguyen TP, Barreau M, Polit M. KSOM and MLP neural networks for on-line estimating the efficiency of an activated sludge process. Chemical Engineering Journal 2006;116(1):1–11.
Grima MA, Babuska R. Fuzzy model for the prediction of unconfined compressive strength of rock samples. International Journal of Rock Mechanics and Mining Sciences 1999;36(3):339–49.
Grima MA, Bruines PA, Verhoef PNW. Modeling tunnel boring machine performance by neuro-fuzzy methods. Tunneling and Underground Space Technology 2000;15(3):259–69.
Hamed MM, Khalafallah MG, Hassanien EA. Prediction of wastewater treatment plant performance using artificial neural networks. Environmental Modelling and Software 2004;19(10):919–28.
Haykin S. Neural networks: a comprehensive foundation. Prentice Hall; 1994.
Heikkinen M, Poutiainen H, Liukkonen M, Heikkinen T, Hiltunen Y. Subtraction analysis based on self-organizing maps for an industrial wastewater treatment process. Mathematics and Computers in Simulation 2011;82(3):450–9.
Holubar P, Zani L, Hager M, Fröschl W, Radak Z, Braun R. Advanced controlling of anaerobic digestion by means of hierarchical neural networks. Water Research 2002;36(10):2582–8.
Hong YT, Rosen MR, Bhamidimarri R. Analysis of a municipal wastewater treatment plant using a neural network-based pattern analysis. Water Research 2003;37(7):1608–18.
Iphar M. ANN and ANFIS performance prediction models for hydraulic impact hammers. Tunnelling and Underground Space Technology 2012;27(1):23–9.
Khataee AR, Kasiri MB. Artificial neural network modeling of water and wastewater treatment processes. NOVA Science Publisher, Inc.; 2011. p. 9–11.
Kohonen T, Oja E, Simula O, Visa A, Kangas J. Engineering applications of the self-organizing map. Proceedings of the IEEE 1996;84(10):1358–84.
Laboratory of Computer and Information Science (CIS). SOM Toolbox 2.0. 2005.
Lee BH, Scholz M. Application of the self-organizing map (SOM) to assess the heavy metal removal performance in experimental constructed wetlands. Water Research 2006;40(18):3367–74.
Llorens E, Thiery F, Grieu S, Polit M. Evaluation of WWTP discharges into a Mediterranean river using KSOM neural networks and mass balance modelling. Chemical Engineering Journal 2008;142(2):135–46.
Mahdevari S, Torabi SR, Monjezi M. Application of artificial intelligence algorithms in predicting tunnel convergence to avoid TBM jamming phenomenon. International Journal of Rock Mechanics and Mining Sciences 2012;55:33–44.
McFeat-Smith I, Fowell RJ. Correlation of rock properties and the cutting performance of tunnelling machines. In: Proceedings of the Conference on Rock Engineering. London: Institution of Mining and Metallurgy; 1977. p. 581–602.
McFeat-Smith I, Fowell RJ. The selection and application of roadheaders for rock tunneling. In: Rapid Excavation and Tunnelling Conference Proceedings, Atlanta; 1979. p. 261–79.
Meulenkamp F, Grima MA. Application of neural networks for the prediction of the unconfined compressive strength (UCS) from Equotip hardness. International Journal of Rock Mechanics and Mining Sciences 1999;36(1):29–39.
Mjalli FS, Al-Asheh S, Alfadala HE. Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. Journal of Environmental Management 2007;83(3):329–38.
Moller MF. A scaled conjugate gradient algorithm for fast supervised learning. Neural Networks 1993;6(4):525–33.
Mukherjee A. Self-organizing neural network for identification of natural modes. Journal of Computing in Civil Engineering 1997;11(1):74–7.
Pendashteh AR, Fakhru'l-Razi A, Chaibakhsh N, Abdullah LC, Madaeni SS, Abidin ZZ. Modeling of membrane bioreactor treating hypersaline oily wastewater by artificial neural network. Journal of Hazardous Materials 2011;192(2):568–75.
Rafiai H, Jafari A. Artificial neural networks as a basis for new generation of rock failure criteria. International Journal of Rock Mechanics and Mining Sciences 2011;48(7):1153–9.
Rahmanian B, Pakizeh M, Mansoori SAA, Abedini R. Application of experimental design approach and artificial neural network (ANN) for the determination of potential micellar-enhanced ultrafiltration process. Journal of Hazardous Materials 2011;187(1–3):67–74.
Rostami J, Ozdemir L, Neil DM. Performance prediction: a key issue in mechanical hard rock mining. Mining Engineering 1994;46(11):1263–7.
Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533–6.
Rustum R, Adeloye AJ, Scholz M. Applying Kohonen self-organizing map as a software sensor to predict biochemical oxygen demand. Water Environment Research 2008;80(1):32–40.
Salari D, Daneshvar N, Aghazadeh F, Khataee AR. Application of artificial neural networks for modeling of the treatment of wastewater contaminated with methyl tert-butyl ether (MTBE) by UV/H_2O_2 process. Journal of Hazardous Materials 2005;125(1–3):205–10.
Salsani A, Daneshian J, Shariati S, Yazdani-Chamzini A, Taheri M. Predicting roadheader performance by using artificial neural network. Neural Computing and Applications 2014;24(7–8):1823–31.

Sandbak LA. Roadheader drift excavation and geomechanical rock classification at San Manuel, Arizona. In: Proceedings of the Rapid Excavation and Tunnelling Conference. New York; 1985. p. 902–16.
Thuro K, Plinninger RJ. Predicting roadheader advance rates. Tunnels and Tunnelling 1999;31:36–9.
Tiryaki B. Application of artificial neural networks for predicting the cuttability of rocks by drag tools. Tunnelling and Underground Space Technology 2008;23(3):273–80.
Tobiszewski M, Tsakovski S, Simeonov V, Namieśnik J. Chlorinated solvents in a petrochemical wastewater treatment plant: an assessment of their removal using self-organising maps. Chemosphere 2012;87(8):962–8.
Vesanto J. SOM-based data visualization methods. Intelligent Data Analysis 1999;3(2):111–26.
Vesanto J, Himberg J, Alhoniemi E, Parhankangas J. Self-organizing map in Matlab: the SOM Toolbox. In: Proceedings of the Matlab DSP Conference; 1999. p. 35–40.
Yagiz S, Karahan H. Prediction of hard rock TBM penetration rate using particle swarm optimization. International Journal of Rock Mechanics and Mining Sciences 2011;48(3):427–33.
Yilmaz I, Kaynar O. Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. Expert Systems with Applications 2011;38(5):5958–66.

Dr. Arash Ebrahimabadi is assistant professor and M.Sc. and doctoral student supervisor, and head of Department of Mining Engineering at Islamic Azad University, Qaemshahr Branch, Iran, where he specializes in interaction of mechanical miners and rock mass, engineering rock mechanics, mechanized tunneling and underground mining, rock cutting, petroleum geomechanics and mine closure optimization. He cooperates (as part time) with Department of Petroleum Engineering, and Department of Mining Engineering at Azad University, Science and Research Branch, Tehran South Branch and Central Tehran Branch, Tehran, Iran. He managed several research projects for the office of vice chancellor for research and technology of Islamic Azad University–Qaemshahr Branch. He has served as mine supervisor and technical advisor for seven years. To date, Dr. Ebrahimabadi has published more than 40 papers and one book entitled "*Roadheader's performance evaluation*". He is a referee and member of scientific committee for *Journal of Environmental Geology* as well as head of scientific committee for Mine Sciences Conference (Iran). He is a member of Iranian Mining Engineering Organization (IMEO), Iranian Society of Mining Engineering (IRSME) and Iranian Society for Rock Mechanics (IRSRM).