INTERNATIONAL JOURNAL OF OPTIMIZATION IN CIVIL ENGINEERING
Int. J. Optim. Civil Eng., 2019; 9(2):343-353

TUNNEL BORING MACHINE PENETRATION RATE PREDICTION BASED ON RELEVANCE VECTOR REGRESSION

H. Fattahi*, †
Department of Mining Engineering, Arak University of Technology, Arak, Iran

ABSTRACT

key factor in the successful application of a tunnel boring machine (TBM) in tunneling is the ability to develop accurate penetration rate estimates for determining project schedule and costs. Thus establishing a relationship between rock properties and TBM penetration rate can be very helpful in estimation of this vital parameter. However, this parameter cannot be simply predicted since there are nonlinear and unknown relationships between rock properties and TBM penetration rate. Relevance vector regression (RVR) is one of the robust artificial intelligence algorithms proved to be very successful in recognition of relationships between input and output parameters. The aim of this paper is to show the application of RVR in prediction of TBM performance. The model was applied to available data given in open source literatures. In this model, uniaxial compressive strengths of the rock (UCS), the distance between planes of weakness in the rock mass (DPW) and rock quality designation (RQD) were utilized as the input parameters, while the measured TBM penetration rates was the output parameter. The performances of the proposed predictive model was examined according to two performance indices, i.e., coefficient of determination (R^2) and mean square error (MSE). The obtained results of this study indicated that the RVR is a reliable method to predict penetration rate with a higher degree of accuracy.

Keywords: relevance vector regression; tunnel boring machine; rock properties; penetration rate.

Received: 10 August 2018; Accepted: 15 December 2018

1. INTRODUCTION

Tunnel boring machine (TBM) penetration rate assessment is an important issue for schedule and cost planning in mechanical tunneling construction projects. Therefore, proper

*Corresponding author: Department of Mining Engineering, Arak University of Technology, Arak, Iran
†E-mail address: h.fattahi@arakut.ac.ir (H. Fattahi)

estimation of TBM penetration rate is helpful for risk reduction in tunneling projects [1]. In recent years, many researchers devoted their works to predict penetration rate. In this paper, the well–known research works are addressed. McFeat-Smith, Tarkoy [2] presented different relations to predict the penetration rate for different types of machines in different geological conditions. Cassinelli et al. [3] used a rock structure rating system for correlation with TBM performance. Nelson [4] studied TBM performance at several tunneling projects mainly in sedimentary rock formations by comparing the instantaneous penetration rate achieved with different rock properties. Tarkoy [5] developed an empirical relationship between total hardness and TBM rate of penetration. Barton [6,7] reviewed a wide range of TBM tunnels to establish the database for estimating penetration rate, utilization and advance rate. Moradi, Farsangi [8] estimated the advance rate in rock TBM tunneling using the risk matrix method. Besides these theoretical and empirical models, soft computing methods have been used to predict the penetration rate. Mikaeil et al. [9], Fattahi [10], Gholamnejad, Tayarani [11], Khademi Hamidi et al. [12], Acaroglu [13], Ghasemi et al. [14], Acaroglu et al. [15], and Grima et al. [16] employed soft computing methods for the prediction of TBM penetration rate.

However, neural networks may result in very poor generalization or even over-fitting when parameters involved in modeling are not chosen wisely. Support vector machine used for regression, the so called support vector regression (SVR), is a suitable machine learning methodology introduced in the early 1990s [17] and has been successfully used for regression tasks in the recent years even for TBM penetration rate estimation [18,19]. However, even this capable network suffers from numerous limitations including parameters and kernel selection which may have significant effect on its prediction efficiency [20,21]. Relevance vector machine based regression (RVR) is a Bayesian sparse kernel technique used for regression having most of the SVR characteristics while avoiding its limitations [22]. It typically leads to much sparser models and correspondingly faster performance on test data as well as a sophisticated generalization error [23,24]. However, RVR has not yet been used for prediction of TBM penetration rate in any kinds of mechanical tunneling construction project.

In this study, the RVR is proposed for indirect prediction of TBM penetration rate. The goodness of RVR model was evaluated by using the data available in the literature. Finally, a statistical error analysis has been performed on the modeling results to investigate the effectiveness of the proposed method.

2. RELEVANCE VECTOR REGRESSION

RVR is a probabilistic model whose functional form is equivalent to that of support vector regression (SVR). It achieves comparable recognition accuracy to the SVR, yet provides a full predictive distribution, and also requires substantially fewer kernel functions [25,26]. RVR is based on Bayesians approach in which a prior is introduced over the model weights and each weight is administrated by one hyperparameter. The most probable value of each hyper parameter is iteratively evaluated from the data. The model is sparser since the posterior distributions of some proportion of the weights are set to zero.

Consider a given training set of M regression data points $\left\{\left(x_m, y_m\right)\right\}_{m=1}^{M}$, where $x_m \in R^M$ is the input data to the actual plant and $y_m \in R$ is the output data of the actual plant and is assumed to contain Gaussian noise ε with mean 0 and variance σ^2. In high dimensional feature space z, the outputs of an extended linear model can be expressed as a linear combination of the response of a set of M basis functions as follows:

$$y\left(x,\omega\right)=\sum_{m=1}^{M}\omega_m\varphi_m\left(x\right)+\varepsilon=\omega^T\varphi+\varepsilon \quad (1)$$

Now, the predicted output $\hat{y}$ of the true value y is

$$\hat{y}\left(x,\omega\right)=\sum_{m=1}^{M}\omega_m\varphi_m\left(x\right)=\omega^T\varphi \; where \;\; \omega\in z \quad (2)$$

In the above nonlinear function estimation model, ω_m is the weight vector and $\varphi_m\left(.\right)$ is an arbitrary basis function (or kernel). In the present work, RBF is used as the kernel function because of its ability to reduce computational complexity of the training process. The vector form of $\omega=\left[\omega_1\cdots\omega_M\right]^T$ and the responses of all kernel function $\varphi\left(x\right)=\left[\varphi_1\left(x\right)\cdots\varphi_M\left(x\right)\right]^T$ maps the input data into a high dimensional feature space z. Hence, the obtained error signal could be stated as

$$\varepsilon_m=y_m-\hat{y}_m=N\left(0,\sigma^2\right) \quad (3)$$

The objective of relevance vector regression is to find the finest value of such that $\hat{y}\left(x,\omega\right)$ makes good predictions for unknown input data. For the RVR model in equation (2) let $\alpha=\left[\alpha_1\cdots\alpha_M\right]^T$ be the vector of M independent hyperparameters, each associated with one model weight or kernel function.

The Gaussian prior distributions of the RVR framework are chosen as

$$p\left(\frac{\omega_m}{\alpha_m}\right)=\prod_{m=1}^{M}\left(\frac{\alpha_m}{2\pi}\right)^{\frac{1}{2}}\exp\left\{-\frac{\alpha_m\omega_m^2}{2}\right\} \quad (4)$$

Here, α_M is the hyperparameter that governs each weight ω_m. The likelihood function of independent training targets $y=y_m, m=1,\ldots,M$ can be stated as

$$p\left(\frac{y}{\omega},\sigma^2\right)=\prod_{m=1}^{M}p\left(\frac{y_m}{\omega},\sigma^2\right)=\frac{e^{-\frac{\left\|(y-\hat{y})^2\right\|}{2}\sigma^2}}{\sqrt{\left(2\pi\sigma^2\right)^M}} \quad (5)$$

The above likelihood function is enhanced by the prior in equation (4) defined over each weight to reduce the complexity of the model and to avoid over fitting. Now, using Bayes' rule, the posterior distribution over model weights could be calculated as follows:

$$p\left(\frac{\omega}{y},\alpha,\sigma^2\right)=\frac{p\left(\frac{y}{\omega},\sigma^2\right)p\left(\frac{\omega}{\alpha}\right)}{p\left(\frac{y}{\alpha},\sigma^2\right)} \quad (6)$$

The posterior distribution in equation (6) is a Gaussian distribution function,

$$p\left(\frac{\omega}{y},\alpha,\sigma^2\right)=N\left(\mu,\sigma^2\right) \quad (7)$$

whose covariance and mean are respectively given by

$$\sum=\left(\sigma^{-2}\varphi^T\varphi+A\right)^{-1}, \quad (8)$$

$$\mu=\sigma^{-2}\sum\varphi^T y \quad (9)$$

with $A=diag\left\{\alpha\right\}$.

Marginalization of the likelihood distribution over the training targets given by equation (5) can be obtained by integrating out the weights to acquire the marginal likelihood for the hyperparameters.

$$p\left(\frac{y}{\alpha},\sigma^2\right)=\int p\left(\frac{y}{\omega},\sigma^2\right)p\left(\frac{\omega}{\alpha}\right)d\,\omega=N\left(0,C\right) \quad (10)$$

Here, the covariance is given by $C=\sigma^2 I+\varphi A^{-1}\varphi^T$ .In equations (8) and (9), the only unknown variables are the hyperparameters α. The values of these hyperparameters are estimated using the framework of type II maximum likelihood [27].

$$p\left(\frac{y}{\alpha},\sigma^2\right)=-\frac{1}{2}\left(M\ \log 2\pi+\log\left|C\right|+y^T C^{-1}y\right) \quad (11)$$

Logarithm is included in equation (11) to reduce computational complexity.

Maximization of the logarithmic marginal likelihood in equation (11) over α leads to the most probable value α_{MP} which provides the maximum a posteriori (MAP) estimate of the weights.

The ambiguity about the optimal value of the weights, given by (6), is used to express ambiguity about the predictions made by the model, i.e., given an input x^*, the probability distribution of the corresponding output y^* is given by the predictive distribution

$$p\left(\frac{y^*}{x^*},\hat{\alpha},\hat{\sigma}^2\right)=\int p\left(\frac{y^*}{x^*},\omega,\hat{\sigma}^2\right)p\left(\frac{\omega}{y},\hat{\alpha},\hat{\sigma}^2\right)d\,\omega \tag{12}$$

which has the Gaussian form

$$p\left(\frac{y^*}{x^*},\hat{\alpha},\hat{\sigma}^2\right)=N\left(Y^*,\sigma^{*2}\right) \tag{13}$$

The mean and variance of the predicted model are, respectively,

$$Y^*=\varphi^T(\mathrm{x}^*)\mu \text{ and } \sigma^{*2}=\hat{\sigma}^2+\varphi^T(\mathrm{x}^*)\sum\varphi(\mathrm{x}^*) \tag{14}$$

Maximizing the logarithmic marginal likelihood in (11) leads the optimal values of many of the hyperparameters α_m typically infnite yielding a posterior distribution in (6) of the corresponding weights ω_m that tends to be a delta function peaked to zero. Thus, the corresponding weights are deleted from the model along with its accompanying kernel function. Hence, very few data points corresponding to nonzero weights build the RVR model and are called the relevance vectors. This results in better sparseness of RVR model than SVR model. Thus, the computation time for prediction using RVR model is reduced significantly. In this paper, the RVR model is used for prediction of TBM penetration rate.

3. DATABASE INFORMATION

The main scope of this work is to implement the above methodology in the problem of TBM penetration rate prediction. Dataset applied in this study for determining the relationship among the set of input and output variables are gathered from open source literature [11].
Database were obtained from three different TBM projects: 1) The Queens Water Tunnel #3, Stage 2, USA. This project is intended to improve fresh water distribution throughout the New York City, USA. The tunnel, about 7.5 km long, was excavated with a high power TBM. 2) The Karaj-Tehran water transfer tunnel, Iran. The 30-km Karaj-Tehran tunnel is the longest water transfer project in Iran and is now being excavated using double shield TBM’s. 3) The Gilgel Gibe II hydroelectric project, Ethiopia. This is a 25-km tunnel that allows power to be generated by the exploitation of the elevation drop between the basin created by the Gilgel Gibe I dam on the Gilgel Gibe river and the river Omo [11]. In the

present study, 185 data sets were collected. Partial dataset used in this study are presents in Table 1. Also, Table 2 shows statistical description of datasets used in this study.

Table 1: Partial dataset used in this study [11]

No.	**Tunnel station**	**Input parameters**			**Output parameter**
		UCS (MPa)	RQD (%)	DPW (m)	Measured Penetration rate (m/h)
1	Queens (USA)	173.1	99.81	1.6	2.17
2	Queens (USA)	159.6	97.35	0.4	2.26
3	Queens (USA)	137.2	99.81	1.6	2.2
4	Queens (USA)	118.3	99.28	0.8	2.22
5	Gilgel Gibe II (Ethiopia)	158.6	73.58	0.1	2.07
6	Gilgel Gibe II (Ethiopia)	140	66.26	0.08	1.58
7	Karaj-Tehran (Iran)	30	56.25	0.25	14.43
8	Karaj-Tehran (Iran)	75	60	0.3	8.69
9	Karaj-Tehran (Iran)	60	62.5	0.3	10.3
10	Gilgel Gibe II (Ethiopia)	75	49.32	0.06	2.89

Table 2**:** Statistical description of dataset utilized for construction of model

Parameter	**Min**	**Max**	**Average**
UCS (MPa)	30	199.7	142.96
RQD (%)	40.6	99.88	91.04
DPW (m)	0.05	2	0.87
Penetration rate (m/h)	1.27	14.43	2.33

4. DATA PROCESSING

To start the training, inputs and output data should be normalized for increasing the efficiency of networks in recognition of the relationships between inputs and output data. Normalization is also really helpful in increasing the accuracy of prediction and scaling the data to minimize the biasing of the networks. Data normalization can also reduce the consuming time of training. It is especially useful for modeling those applications where input data are in different scales [21,28]. There are many normalization techniques conventionally used to scale up the data including Z–Score normalization, Min–Max normalization, sigmoid normalization, statistical column normalization, etc. However, for the purpose of this study, Min–Max normalization method was used. This was due to the capability of Min–Max normalization in maintaining the variation of each feature after normalization. Beside, this normalization method can preserve all of the relationships in the data [28]. Min–Max normalization equation is expressed as below:

$$x_M = 2\left(\frac{x - x_{\min}}{x_{\max} - x_{\min}}\right) - 1 \qquad (15)$$

where x is the original value of the dataset, x_M is the mapped value, and x_{max} (x_{min}) denotes the maximum (minimum) raw input values, respectively.

In addition to the normalization, mean square error (MSE) and coefficient of determination (R^2) are two conventional criteria considered to assess the efficiency of the networks. The MSE is calculated using the following equation:

$$MSE = \frac{1}{n}\sum_{k=1}^{n}(t_k - \hat{t}_k)^2 \tag{16}$$

where t_k be the actual value and $\hat{t}_k$ be the predicted value of the k^{th} observation and n is the number of samples used for training or testing the network. MSE is routinely used as a criterion to show the discrepancy between the measured and estimated values of the network. Coefficient of determination, R^2, is also calculated as

$$R^2 = 1 - \frac{\sum_{k=1}^{n}(t_k - \hat{t}_k)^2}{\sum_{k=1}^{n} t_k^{\,2} - (\sum_{k=1}^{n} \hat{t}_k^{\,2} / n)^2} \tag{17}$$

R^2 is widely used as a representation of the initial uncertainty of the model. The best network model which is unlikely to build, would have MSE=0 and R^2 =1.

5. RESULTS AND DISCUSSION

In this study, RVR model was utilized to build a prediction model for the prediction of TBM penetration rate from available data, using MATLAB environment. A dataset that includes 185 data points was employed in current study, while 148 data points (80%) were utilized for constructing the model and the remainder data points (37 data points) were utilized for model performance evaluation. In this model, uniaxial compressive strengths of the rock (UCS), the distance between planes of weakness in the rock mass (DPW) and rock quality designation (RQD) were utilized as the input parameters, while the measured TBM penetration rates was the output parameter.

In RVR model, hyper parameter estimation is carried out by expectation maximization (EM) updates on the objective function [22,26]. For this RVR model, radial basis function (RBF) kernel is used with the width parameter estimated automatically by the learning procedure [22,26] which improves generalization ability and reduces computational complexity of the training process. Thus, unlike in SVR there is no necessity for computationally expensive determination of regularization parameter by cross validation technique. Also in the RVR model confidence intervals, likelihood values and posterior probabilities could be explicitly encoded easily.

After modeling, a comparison between estimated values of penetration rate by the RVR model and measured values for 185 data sets at training and testing phases is shown in Fig.

1. As shown in Fig. 1, the results of the RVR model in comparison with actual data show a good precision of the RVR model.

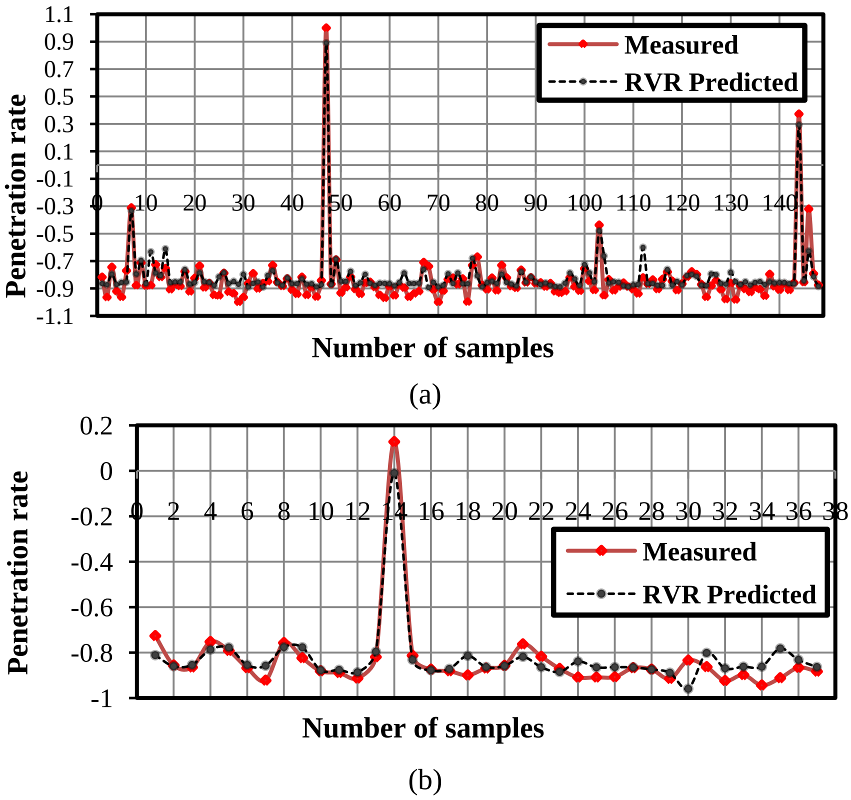

Figure 1. Comparison between measured and estimated penetration rate for a) training datasets, b) testing datasets

Furthermore, a correlation between estimated values of penetration rate by the RVR model and measured values for 185 data sets at training and testing phases is shown in Fig. 2.

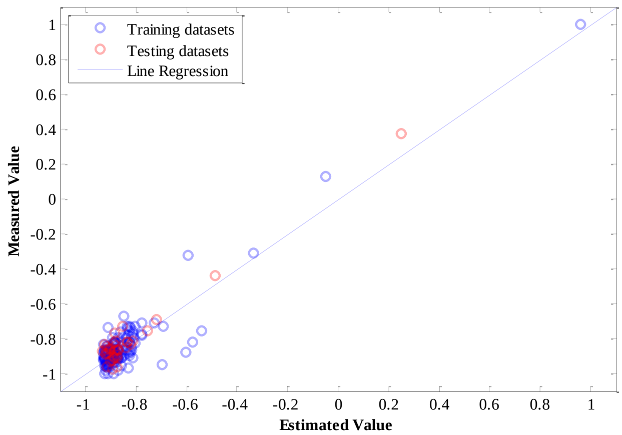

Figure 2. Correlation between measured and estimated penetration rate for training and testing datasets

Also, performance analysis of the RVR model for predicting penetration rate is shown in Table 3. The performance indices obtained in Table 3 indicate the high performance of the RVR model that can be used successfully to the estimation of the penetration rate.

Table 3: Performance analysis of the RVR model for predicting penetration rate

Description		R^2	MSE
RVR model	Training	0.0050	0.932
	Testing	0.0028	0.976

Also, relative error (error percentage) for data point (training and testing samples) is assessed and revealed in Fig. 3 Relative error for most data points is located in range of [-18% 14%], which is an acceptable value.

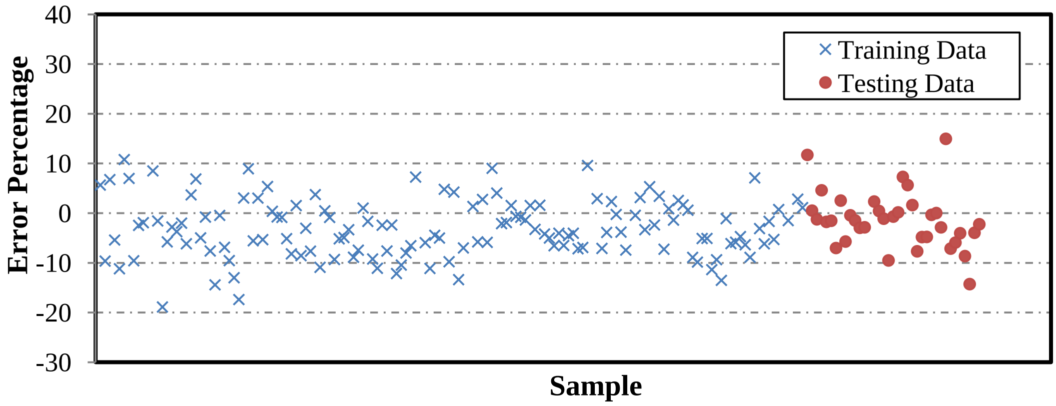

Figure 3. Relative error (error percentage) of RVR model in estimating the penetration rate

Eventually, we compared our results with the results obtained by Gholamnejad, Tayarani [11]. This comparison is demonstrated in Table 4. Table 4 contains two methods, including artificial neural network and the RVR model suggested in this study.

As can be seen, the RVR model indicate better results relative to previously published model. As presented in Table 4, the RVR model with MSE= 0.0028 and R^2= 0.976 is found to be the best predictive model. However, the best previously published model (artifitial neural network) has the MSE = 0.1082 and R^2= 0.939.

Table 4: Comparison of performance of the proposed model and previously presented model

Description		MSE	R^2
RVR model	Training	0.0050	0.932
(Proposed in this study)	Testing	0.0028	0.976
Artificial neural network model	Training	0.0083	----
(Gholamnejad, Tayarani [11])	Testing	0.1082	0.939

6. CONCLUSION

Prediction of TBM penetration rate is one of the most important concerns in estimating cost and time of a tunnel project. There are various techniques of nonlinear analysis utilized for estimating the TBM penetration rate. In this study, the RVR technique has been used for estimating the TBM penetration rate. It is observed that the UCS, DPW and RQD have major effect on the TBM penetration. So, the model was generated based on relevant properties. The following conclusions can be drawn:

- The RVR with MSE= 0.0028 and R^2= 0.976 is a reliable system modeling technique for predicting TBM penetration rate with highly acceptable degree of accuracy and robustness.
- Comparison between the developed model and previously presented model reveals the superiority of the RVR in prediction of TBM penetration rate.

This study shows that the RVR approach can be applied as a powerful tool for modeling of some problems involved in tunnel engineering.

REFERENCES

1. Koopialipoor M, Nikouei SS, Marto A, Fahimifar A, Jahed Armaghani D, Mohamad ET. Predicting tunnel boring machine performance through a new model based on the group method of data handling, *Bull Eng Geology Envir* 2018. doi:10.1007/s10064-018-1349-8.
2. McFeat-Smith I, Tarkoy PJ. Assessment of tunnel boring machine performance, *Tunn Tunnel Int* 1979; **11**(10): 33-7.
3. Cassinelli F, Cina S, Innaurato N. Power consumption and metal wear in tunnel-boring machines: analysis of tunnel-boring operation in hard rock: In: Tunneling 82, *Proceedings of the 3rd International Symposium*, Brighton, 7–11 June 1982, P73–81. Publ London: IMM, 1982. Paper presented at the International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts.
4. Nelson P. Tunnel boring machine performance in sedimentary rock, 1983.
5. Tarkoy P. Practical geotechnical and engineering properties for tunnel-boring machine performance analysis and prediction transportation research record 1087, *Transport Res Board, National Res Council* 1987: 62-78.
6. Barton, N. TBM performance estimation in rock using QTBM, *Tunn Tunneling Int* 1999; **31**: 41-8.
7. Barton NR. *TBM Tunnelling in Jointed and Faulted Rock,* CRC Press, 2000.
8. Moradi MR, Farsangi MAE. Application of the risk matrix method for geotechnical risk analysis and prediction of the advance rate in rock TBM tunneling, *Rock Mech Rock Eng* 2014; **47**(5): 1951-60.
9. Mikaeil R, Naghadehi MZ, Ghadernejad S. An extended multifactorial fuzzy prediction of hard rock TBM penetrability, *Geotech Geol Eng* 2018; **36**(3): 1779-1804.
10. Fattahi H. Adaptive neuro fuzzy inference system based on fuzzy C–means clustering algorithm, a technique for estimation of TBM peneteration rate, *Int J Optim Civil Eng* 2016; **6**(2): 159-71.

11. Gholamnejad J, Tayarani N. Application of artificial neural networks to the prediction of tunnel boring machine penetration rate, *Min Sci Tech* 2010; **20**(5): 727-33.
12. Khademi Hamidi J, Shahriar K, Rezai B, Bejari H. Application of fuzzy set theory to rock engineering classification systems: an illustration of the rock mass excavability index, *Rock Mech Rock Eng* 2010; **43**(3): 335-50.
13. Acaroglu O. Prediction of thrust and torque requirements of TBMs with fuzzy logic models, *Tunn Undergr Sp Tech* 2011; **26**(2): 267-75.
14. Ghasemi E, Yagiz S, Ataei M. Predicting penetration rate of hard rock tunnel boring machine using fuzzy logic, *Bull Eng Geology Envir* 2014; **73**(1): 23-35.
15. Acaroglu O, Ozdemir L, Asbury B. A fuzzy logic model to predict specific energy requirement for TBM performance prediction, *Tunn Undergr Sp Tech* 2008; **23**(5): 600-8.
16. Grima MA, Bruines P, Verhoef P. Modeling tunnel boring machine performance by neuro-fuzzy methods, *Tunn Undergr Sp Tech* 2000; **15**(3): 259-69.
17. Vapnik V. *The Nature of Statistical Learning Theory*, John Wiley and Sons, 1998.
18. Mahdevari S, Shahriar K, Yagiz S, Shirazi MA. A support vector regression model for predicting tunnel boring machine penetration rates, *Int J Rock Mech Min Sci* 2014; **72**: 214-29.
19. Fattahi H, Babanouri N. Applying optimized support vector regression models for prediction of tunnel boring machine performance, *Geotech Geol Eng* 2017; **35**(5): 2205-17.
20. Bishop CM. Pattern recognition and machine learning. In., p. p. 746. Springer, 2006.
21. Gholami R, Moradzadeh A, Maleki S, Amiri S, Hanachi J. Applications of artificial intelligence methods in prediction of permeability in hydrocarbon reservoirs, *J Pet Sci Eng* 2014; **122**: 643-56.
22. Tipping ME. Sparse Bayesian learning and the relevance vector machine, *J Machine Learn Res* 2001; **1**(Jun): 211-44.
23. Liu X, Zhang XH, Yuan J. Relevance vector machine and fuzzy system based multi-objective dynamic design optimization: A case study, *Expert Syst Appl* 2010; **37**(5): 3598-3604.
24. Caesarendra W, Widodo A, Yang BS. Application of relevance vector machine and logistic regression for machine degradation assessment, *Mech Syst Signal Pr* 2010; **24**(4): 1161-71.
25. Bishop CM, Tipping ME. Variational relevance vector machines. In: *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence* 2000, pp. 46-53, Morgan Kaufmann Publishers Inc.
26. Nisha MG, Pillai G. Nonlinear model predictive control with relevance vector regression and particle swarm optimization, *J Control Theory App* 2013; **11**(4): 563-9.
27. Berger JO. *Statistical Decision Theory and Bayesian Analysis*, Springer Science & Business Media, 2013.
28. Jayalakshmi T, Santhakumaran A. Statistical normalization and back propagation for classification, *Int J Comput Theory Eng* 2011; **3**(1): 1793-8201.