CrossMark

ORIGINAL ARTICLE

Feasibility of PSO-ANN model for predicting surface settlement caused by tunneling

Mahdi Hasanipanah[1] · Majid Noorian-Bidgoli[2] · Danial Jahed Armaghani[3] ·
Hossein Khamesi[1]

Abstract The potential surface settlement, especially in urban areas, is one of the most hazardous factors in subway and other infrastructure tunnel excavations. Therefore, accurate prediction of maximum surface settlement (MSS) is essential to minimize the possible risk of damage. This paper presents a new hybrid model of artificial neural network (ANN) optimized by particle swarm optimization (PSO) for prediction of MSS. Here, this combination is abbreviated using PSO-ANN. To indicate the performance capacity of the PSO-ANN model in predicting MSS, a pre-developed ANN model was also developed. To construct the mentioned models, horizontal to vertical stress ratio, cohesion and Young's modulus were set as input parameters, whereas MSS was considered as system output. A database consisting of 143 data sets, obtained from the line No. 2 of Karaj subway, in Iran, was used to develop the predictive models. The performance of the predictive models was evaluated by comparing performance prediction parameters, including root mean square error (RMSE), variance account for (VAF) and coefficient correlation (R^2). The results indicate that the proposed PSO-ANN model is able to predict MSS with a higher degree of accuracy in comparison with the ANN results. In addition, the results of sensitivity analysis show that the horizontal to vertical stress ratio has slightly higher effect of MSS compared to other model inputs.

Keywords Tunneling · Surface settlement · PSO-ANN · Hybrid model

✉ Mahdi Hasanipanah
  Hasanipanah.m@grad.kashanu.ac.ir

  Majid Noorian-Bidgoli
  noriyan@kashanu.ac.ir

  Danial Jahed Armaghani
  danialarmaghani@gmail.com

  Hossein Khamesi
  khamesi_hossein@yahoo.com

1 Department of Mining Engineering, University of Kashan, Kashan, Iran

2 Division of Mining Engineering, Engineering Department, University of Kashan, Kashan, Iran

3 Young Researchers and Elite Club, Science and Research Branch, Islamic Azad University, Tehran, Iran

1 Introduction

With the increasing population and urbanization in urban areas, as well as the growing demand for public transportation, the requirement for metro tunnels has been significantly increased. In subway tunnel excavations, it is necessary to estimate and control surface settlements observed after excavation that may cause damages to the surface structures [1]. Based on previous researches [2, 3], many geotechnical and geometrical parameters, such as cohesion, Poisson's ratio, Young's modulus, angle of internal friction and face support pressure, have been considered in predicting the values of the MSS.

Surface settlements are influenced by three main groups of factors, i.e., excavation and support method, tunnel geometry and ground properties. In the first group, the excavation and support methods are including Excavation, such as NATM and TBM, excavation type (full face or sequential mining) and Support, such as anchoring, shotcrete, steel sets and lining. In the second group, tunnel geometry factors are including worksite conditions, depth, diameter, number of tunnels and distance between tunnels.

In the third group, ground properties are including elasticity modulus, unit weight, cohesion, friction angle, Poisson's ratio, groundwater and permeability [4].

Previously, empirical and analytical methods as well as numerical analysis by finite difference (FD) and finite element (FE) methods were developed to predict the values of MSS [5, 6]. For instance, a method for estimating surface settlement above tunnels constructed in soft ground was developed in the study conducted by Schmidt [7]. Attewell and Farmer [8] evaluated ground disturbance caused by shield tunneling in a stiff, over-consolidated clay. In the other study, Ocak [9] proposed a new equation for estimating the transverse settlement curve of twin tunnels. He demonstrated that the proposed equation can estimate the transverse settlement with degree of confidence in the Otogar–Kirazli metro case studies. Atkinson and Potts [10] investigated the influence of the depth of burial and crown settlement on the surface settlement above shallow tunnels driven in soft ground. Hamza et al. [11] studied the ground movements due to construction of cut-and-cover structures and slurry shield tunnel of the Cairo Metro. Chi et al. [12] indicated the application of the conjugate gradient method for the back-analysis of tunneling-induced ground movement. They established semi-empirical equations to predict the tunneling-induced ground movement in the silty clay and silty sand of Taipei basin. Chou and Bobet [13] used twenty-eight tunnels to evaluate predictions from an analytical solution for shallow tunnels in saturated ground. As a result, comparisons between predictions and observations from actual tunnels indicated good agreement, generally within 15 % difference. In the other study of analytical solutions, Park [14] applied elastic solutions to predict the tunneling-induced undrained ground movements for shallow and deep circular tunnels in soft ground. He showed a good agreement of the predicted ground deformations with field observations for tunnels in uniform clay. Short-term surface settlements for twin tunnels, located between the Esenler and Kirazlı stations on the Istanbul Metro line, were predicted by Ercelebi et al. [15]. For this purpose, they used three different methods, including FE, semi-theoretical (semi-empirical) and analytical methods to predict surface settlement caused by tunneling. Their results indicated that the FE method can be used as a reliable method to predict short-term settlement.

Apart from empirical models, in recent years, artificial intelligence (AI) methods, such as artificial neural network (ANN), fuzzy inference system and support vector machine (SVM), have been developed for solving problems of rock and geotechnical engineering [16–19]. In the field of MSS prediction, these models have been widely used and developed. Ocak and Seker [1] used three different methods, including ANN, SVM and Gaussian processes (GP) to estimate surface settlement. They concluded that the GP is a

more precise method than the ANN and SVM models. In addition, a comprehensive study for prediction of MSS by ANN and multiple regression was presented by Mohammadi et al. [3]. The results of their research demonstrated that the ANN method can be regarded as a more reasonable predictive technique in predicting MSS.

Recently, the use of combination of evolutionary algorithms, such as particle swarm optimization (PSO) and imperialist competitive algorithm (ICA) with ANN has been highlighted in the field of rock engineering [20, 21]. The results indicated that such algorithms are useful to design the ANN. Nevertheless, as long as author's knowledge, evolutionary algorithms have not been used and proposed for MSS prediction. In this research, a combination of PSO and ANN was proposed to predict MSS induced by tunneling along the line 2 of Karaj subway. In fact, PSO algorithm is utilized to incorporate ANN for its optimization propose.

2 Theory and methods

2.1 Artificial neural network (ANN)

One of the subsystems of AI systems is an ANN. The ANN model has been developed since the 1960s. Generally, the structure of an ANN, which is inspired by the human brain, consists of a group of computational units called neurons or nodes. These neurons are highly interconnected with each other. In addition, the capability of these neurons for performing mass parallel distributed processing is proved by many researches [22–24]. A typical ANN consists of three layers, namely input, hidden and output layers. The mentioned neurons are placed in these layers and linked to each other by weights. On the other hand, problem effective and objective variables are placed in the input and output layers, respectively [25]. Theoretically, there are no restrictions on the No. of hidden layers and No. of neurons in the hidden layers and can be determined based on trial-and-error procedure [26]. To construct an ANN model, in the first step, ANNs require training to learn and consequently map a relationship from the data. There are many algorithms to train the network, such as Levenberg–Marquardt (LM), conjugate gradient and scaled conjugate gradient algorithms [27]. The selection of the best algorithm depends on the given problem, the purposes of the performed network such as classification and prediction, the number of datasets and so on. In the second step, to check the performance capacity of the constructed model, the rest of datasets are used for testing [28]. Although, ANN is used as a quick solution for engineering problems, it has a number of disadvantages: slow learning rate and getting trapped in local minima [29, 30].

Fig. 1 Location of the line 2 of Karaj Subway
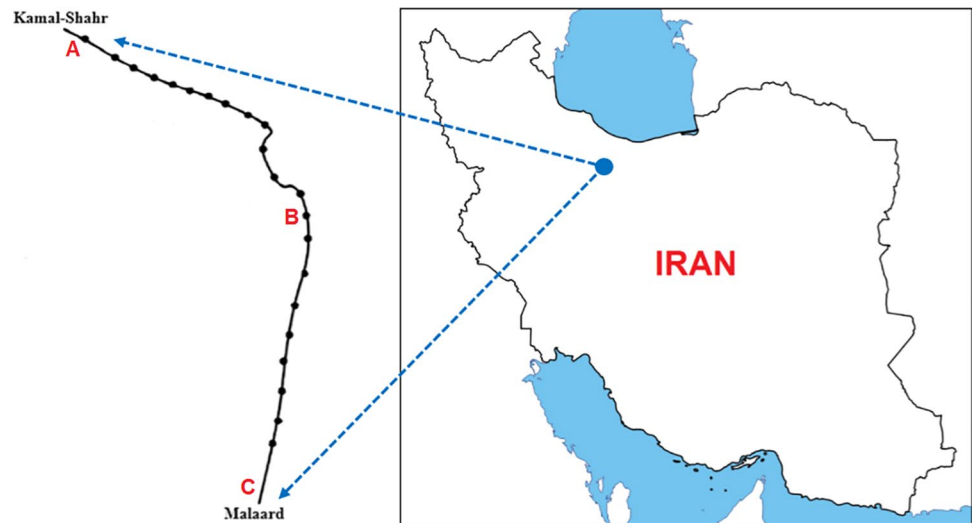
2.2 Particle swarm optimization (PSO)

PSO which was first introduced by Kennedy and Eberhart [31] is a simple and powerful optimization technique inspired by social behavior of bird flocking or fish schooling. In the PSO algorithm, a number of simple particles are placed in the search space of n-dimensional problem or function [32, 33]. A potential solution can be represented by each particle and the particles evaluate the objective function at their current position. The next location of each particle is determined by combining some aspects of their own current and best position with those of other swarm particles, with some random perturbations [34]. Eventually, the swarm can be expected to move close to the optimum of fitness function [35].

Using Eqs. (1) and (2), the position and velocity of the particles can be determined and updated.

$$X_i^{k+1} = X_i^k + V_i^{k+1} \tag{1}$$

$$V_i^{k+1} = V_i^k + c_1\, r_1 \left(p_{\text{best},i}^k - X_i^k \right) + c_2\, r_2 \left(g_{\text{best}}^k - X_i^k \right) \tag{2}$$

where X_i^k is the n-dimensional vector that represents the position of particle i in the search space at iteration k. V_i denotes the velocity of this particle. The velocity vector derives the optimization process by reflecting both the experimental knowledge of the particle and socially shared information from the particle's neighborhood [36] by introducing distance of the particle from its own best position and swarm best position. The best position the particle has visited and found by the swarm so far are represented by $p_{\text{best},i}$ and g_{best}, respectively, in Eq. (2). Furthermore, r_1 and r_2 are random values in the range of zero to one, c_1 and c_2 are positive acceleration constants. The fitness function

f measures how close the corresponding solution is to the optimum by calculating $p_{\text{best},i}$ and g_{best}. Due to this fact, objective function plays an integral role in this problem. Considering the minimization problem, the personal and global best positions at the next iteration are defined as:

$$p_{\text{best},i}^{k+1} = \begin{cases} X_i^{k+1}, & f\left(X_i^{k+1}\right) < f\left(p_{\text{best},i}^k\right) \\[2mm] P_{\text{best},i}^k, & f\left(X_i^{k+1}\right) \geq f\left(p_{\text{best},i}^k\right) \end{cases} \tag{3}$$

$$\begin{aligned} g_{\text{best}}^k \in \left\{ p_{\text{best},0}^k, \cdots, p_{\text{best},n_s}^k \right\} &\Big| f\left(g_{\text{best}}^k\right) \\ &= \min\left\{ f\left(p_{\text{best},0}^k\right), \cdots, f\left(p_{\text{best},n_s}^k\right) \right\} \end{aligned} \tag{4}$$

where n_s denotes the total number of particles in the swarm. The particles continue to move in the search space, with their position being updated at each iteration until the stopping condition is met.

3 Case study and data collection

In this research, datasets were collected from Karaj Subway (line No. 2), in Iran. Karaj is one of the large cities in Iran with 1.4 million inhabitants. Due to the increasing population and urbanization in this city, construction of a new subway system is necessary and crucial. Constructing of the operational line No. 2 of Karaj Subway was started in February 2007 with a total length of 27 km. Shape of tunnel is horseshoe and tunnel has 7.8 m height and 8.4 m width. Tunnel depth change 7–14 m. This project connects the Kamal-Shahr and Malaard, in northwestern and south of Karaj city, respectively (see Fig. 1). Based on many parameters, i.e., geotechnical analysis and economic
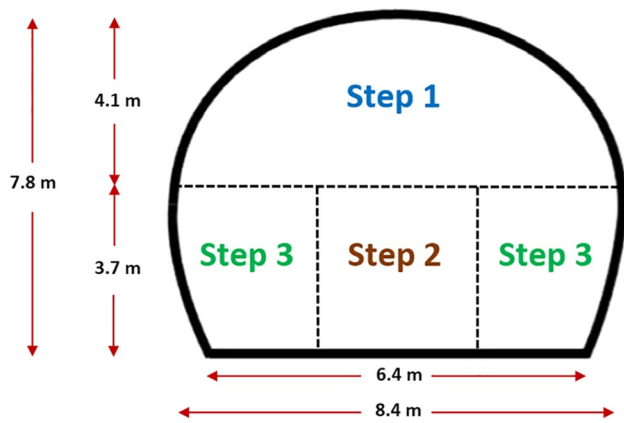
Fig. 2 The steps of tunnel excavation in the line 2 of Karaj Subway using NATM

Table 1 The range of measured parameters for MSS prediction

Variable category	Parameter	Unit	Min	Max
Input	Horizontal to vertical stress ratio	–	0.3	0.6
	Cohesion	kPa	10	45
	Young's modulus	MPa	50	200
Output	Max surface settlement	mm	3.7	7

studies, the tunnels have been designed and built in two phases, i.e., first and second, 14.5 and 12.5 km, respectively (see Fig. 1). According to Fig. 1, AB and BC are the first and second phases, respectively. Both the first and second phases are excavated using New Austria tunnel method (NATM). According to NATM, the tunnel excavation, in this project, was designed in three sections, as shown in Fig. 2. Based on Fig. 2, the heading was excavated in the step 1. Afterwards, the steps 2 and 3 were excavated, respectively. Moreover, it is observed that the tunnel has a horseshoe shape with 7.8 m height and 8.4 m width with the lining. After excavating the step 1, the exposed area is supported using steel fiber–reinforced shotcrete.

In this research, a group of datasets, including 143 datasets, was collected from the laboratory and in situ tests. In this regard, the values of horizontal to vertical stress ratio (coefficient of earth pressure), cohesion and Young's modulus were measured and considered as input parameters. To determine the coefficient of earth pressure, in situ horizontal stress and in situ vertical stress tests were conducted. In addition, the values of MSS were carefully measured and considered as output parameter. The range of the mentioned parameters to construct the predictive models, for all of 143 data sets, is given in Table 1. To measure MSS, the settlement markers were installed, grouted about 100 cm into the ground, placed approximately at intervals of 25 m along the
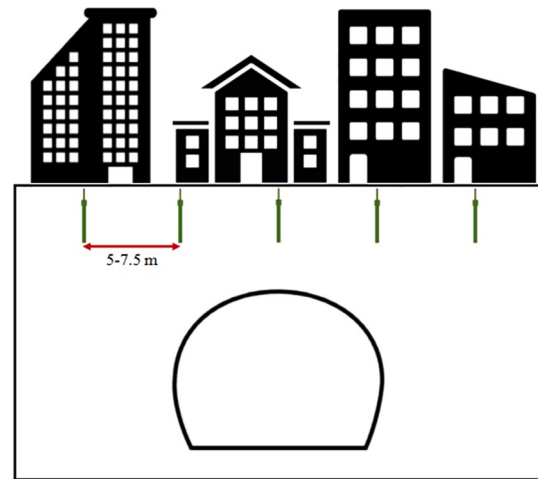

Fig. 3 Schematic diagram of settlement marker location in the line 2 of Karaj Subway

tunnel alignment, and the surface settlements were measured. In addition, in each transverse section, three or five surface settlement markers, which are arranged approximately at intervals of 5–7.5 m, were installed, as depicted in Fig. 3.

4 Prediction of MSS

In this section, the modeling procedures of ANN and hybrid PSO-ANN models for MSS prediction are described. These models are constructed with the MatLab environment using MatLab2013b. To develop the models, the datasets have been divided into two groups: training and testing datasets. Previous researchers have recommended various percentages for the testing datasets [37–39]. In the present study, 80 and 20 % of whole datasets were used for model developments and checking the performance of the developed models, respectively. Selection of the random training and testing data was carried out by a MatLab code written by authors.

4.1 Prediction of MSS by ANN model

In this part, an attempt has been made to estimate MSS using ANN procedure. In the first stage of this modeling procedure, the prepared database was normalized to simplify the design procedure as follows:

$$X_{norm} = (X - X_{min}) / (X_{max} - X_{min}) \tag{5}$$

where X and X_{norm} are the measured and normalized values, respectively. X_{max} and X_{min} are the maximum and minimum values of the X. Note that, to achieve a reasonable solution,

Table 2 Several equations for determination of the no. of hidden node by previous investigators

Heuristic	References
$\leq 2 \times N_i + 1$	Hecht-Nielsen [45]
$(N_i + N_0)/2$	Ripley [50]
$\frac{2 + N_0 \times N_i + 0.5 N_0 \times (N_0^2 + N_i) - 3}{N_i + N_0}$	Paola [51]
$2N_i/3$	Wang [52]
$\sqrt{N_i \times N_0}$	Masters [53]
$2N_i$	Kaastra and Boyd [54]
	Kannellopoulas and Wilkinson [41]

N_i number of input neuron, N_0 number of output neuron

it is recommended that the numeric values of input and output parameters be normalized [17–21].

In the next stage of ANN modeling, the prepared database should be divided into training and testing datasets for model developments and also model evaluations. Here, testing datasets are utilized to evaluate the performance capacity of the developed models. In ANN modeling, selection of the ANN training algorithm and also the determination of the network architecture are the most difficult tasks [40, 41]. Among all ANN training algorithms, as mentioned before, LM was selected and utilized to train the ANN systems. Many researchers highlighted the efficiency of the LM algorithm, among other training algorithms, in solving engineering problems (e.g., [42–44]). On the other hand, as mentioned by many scholars (e.g., [45–47]) an ANN network with only one hidden layer can estimate almost all problems. In addition, developing an ANN model with one hidden layer is of attention because of its beneficial effect on decreasing the complexity of a model and as a consequence the likelihood of model overfitting. Hence, in this study, all proposed artificial intelligent (AI) models were designed using one hidden layer.

In the next stage of ANN design, number of hidden nodes (N_h) in a hidden layer should be determined. Sonmez et al. [48] and Sonmez and Gokceoglu [49] stated that the number of hidden node(s) has a deep impact on the performance prediction of an ANN model. In this regard, previous researchers proposed several equations for determining the N_h as shown in Table 2. Based on this table, the upper limit for the N_h is $2N_i + 1$, where N_i is the number of input parameters. Considering the presented equations in Table 2 and the prepared datasets, in this study, a range of 1–7 for the number of hidden nodes can solve MSS problem. It seems that the proper N_h should be obtained using the trial-and-error procedure. For this purpose, a series of ANN models were designed using the mentioned parameters. The performance prediction of the constructed models was checked using both coefficient of determination (R^2) and root mean square error (RMSE) criteria as presented

Table 3 R^2 and RMSE values of the constructed ANN models

Model No.	Nodes in hidden layers	Run 1 Training R^2	Run 1 Training RMSE	Run 1 Testing R^2	Run 1 Testing RMSE	Run 2 Training R^2	Run 2 Training RMSE	Run 2 Testing R^2	Run 2 Testing RMSE	Run 3 Training R^2	Run 3 Training RMSE	Run 3 Testing R^2	Run 3 Testing RMSE	Run 4 Training R^2	Run 4 Training RMSE	Run 4 Testing R^2	Run 4 Testing RMSE	Run 5 Training R^2	Run 5 Training RMSE	Run 5 Testing R^2	Run 5 Testing RMSE
1	1	0.903	0.067	0.918	0.077	0.92	0.068	0.912	0.073	0.899	0.071	0.911	0.07	0.882	0.074	0.906	0.072	0.908	0.071	0.91	0.072
2	2	0.909	0.065	0.921	0.071	0.923	0.065	0.901	0.078	0.91	0.068	0.918	0.065	0.913	0.069	0.915	0.071	0.911	0.069	0.917	0.071
3	3	0.921	0.064	0.926	0.066	0.924	0.064	0.92	0.067	0.891	0.073	0.925	0.065	0.908	0.07	0.918	0.066	0.923	0.067	0.931	0.069
4	4	0.929	0.066	0.932	0.067	0.939	0.063	0.94	0.064	0.935	0.066	0.933	0.066	0.927	0.068	0.926	0.067	0.931	0.066	0.937	0.065
5	5	0.934	0.065	0.937	0.065	0.931	0.065	0.938	0.064	0.92	0.067	0.937	0.064	0.938	0.064	0.938	0.065	0.933	0.065	0.932	0.065
6	6	0.922	0.067	0.93	0.068	0.928	0.065	0.91	0.069	0.919	0.068	0.928	0.065	0.936	0.063	0.933	0.066	0.937	0.063	0.917	0.066
7	7	0.915	0.069	0.922	0.07	0.925	0.068	0.917	0.068	0.936	0.064	0.935	0.065	0.924	0.065	0.93	0.066	0.929	0.065	0.92	0.067

Table 4 Results of PSO-ANN models for various number of particles in predicting MSS

Model No.	No. of particle	PSO-ANN result				Rank value				Total Rank
		Training		Testing		Training		Testing		
		R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	
1	25	0.882	0.091	0.845	0.091	1	1	3	3	8
2	50	0.887	0.086	0.891	0.087	2	2	6	4	14
3	75	0.920	0.075	0.881	0.081	8	7	5	6	26
4	100	0.928	0.066	0.862	0.106	10	11	4	2	27
5	150	0.894	0.085	0.927	0.068	3	3	11	10	27
6	200	0.898	0.082	0.945	0.057	4	4	12	12	32
7	250	0.922	0.074	0.915	0.068	9	8	7	10	34
8	300	0.913	0.076	0.924	0.065	6	6	10	11	33
9	350	0.918	0.073	0.917	0.077	7	9	8	7	31
10	400	0.930	0.069	0.920	0.071	11	10	9	8	38
11	450	0.902	0.080	0.927	0.070	5	5	11	9	30
12	500	0.938	0.060	0.917	0.082	12	12	8	5	37

in Table 3. In this table, each hidden node is run five times. It is well established that a constructed model with lower RMSE and higher R^2 values is of advantage. Based on the obtained results, run 2 of the ANN model No. 4 with $N_h = 4$ indicates higher R^2 and lower RMSE values compared to other constructed models. So, an architecture of $(3 \times 4 \times 1)$ was selected and introduced for solving an MSS problem by ANN model. More discussions regarding the evaluation of the ANN model will be given later.

4.2 Prediction of MSS by PSO-ANN model

As mentioned in Sect. 1, in this study, an attempt has been made to increase the performance prediction of the ANN model by incorporating PSO algorithm to develop a predictive model with a higher degree of accuracy for MSS prediction. In this system, PSO is performed for minimization of a cost function by adjusting the weights and biases. The followings are the modeling procedure of the hybrid PSO-ANN model in predicting MSS.

4.2.1 Swarm size

The number of particle or swarm size has a significant impact on the performance capacity of the hybrid PSO-ANN technique. Considering the results of previous studies, there is no any specific way to determine proper swarm size. Therefore, it is well known to obtain swarm size considering parametric study using trial-and-error method (e.g., [55, 56]). Table 4 presents the results of PSO-ANN models for various numbers of particles together with their RMSE and R^2 values. In these analyses, iteration number of 100 and architecture of $2 \times 5 \times 1$ were considered. In addition, based on literature's suggestions [55, 56], velocity

coefficients of 2 ($C_1 = C_2 = 2$) and inertia weight of 0.25 were utilized in all PSO-ANN models of this study.

As depicted in Table 4, selecting the best swarm size is very difficult. To overcome this problem, a ranking technique introduced by Zorlu et al. [57] was used. According to the mentioned technique, each performance index (RMSE or R^2) was ordered in its class and the best performance index was assigned the highest rating. For example, values of 0.882, 0.887, 0.920, 0.928, 0.894, 0.898, 0.922, 0.913, 0.918, 0.930, 0.902 and 0.938 were achieved for R^2 of training datasets of models 1–12, respectively, and values of 1, 2, 8, 10, 3, 4, 9, 6, 7, 11, 5 and 12 were assigned to their ranks, respectively. Additionally, in the case of RMSE and also testing datasets, this procedure was applied. Afterwards, for each PSO-ANN model, the ratings of the RMSE and R^2 for both training and testing datasets were summed up (total rank). According to the total rank results, PSO-ANN model No. 10 with a swarm size of 400 shows the highest total rank value. Hence, 400 was chosen as the optimum number of particle or swarm size in predicting MSS.

4.2.2 Termination criteria

The defined termination criteria in this study are considered as maximum number of iterations (I_{Max}). An usual way for determining the I_{Max} is to compare the network result in various iteration numbers. Previous researchers [32, 58] suggested various I_{Max} values for solving different engineering problems. For instance, I_{Max} values of 400, 400 and 450 were recommended for solving the problems in the studies conducted by Jahed Armaghani et al. [32], Gordan et al. [58] and Tonnizam Mohamad et al. [59], respectively. Therefore, another parametric study was conducted on the swarm size values used in the previous stage to find I_{Max}
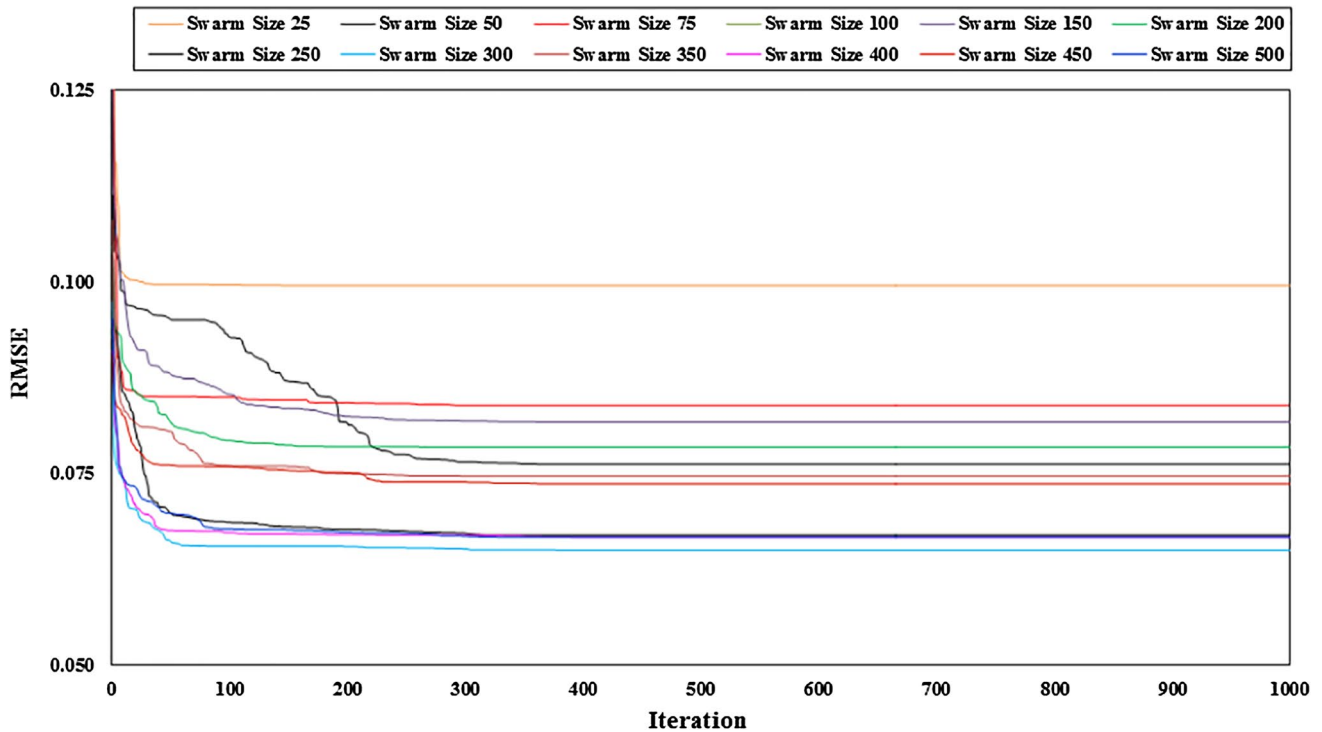
Fig. 4 Results of PSO-ANN network for determining the I_{Max}

as displayed in Fig. 4. Here, performance prediction of the network was checked using RMSE results. In obtaining I_{Max}, iteration number of 1000, $C_1 = C_2 = 2$, inertia weight of 0.25 and architecture of $2 \times 5 \times 1$ was applied. As shown in Fig. 4, after iteration No. of 300, there are no significant changes in the network results for all swarm size values. Hence, I_{Max} of 300 was chosen in the modeling process of this study in predicting MSS.

4.2.3 Network architecture

In this step of PSO-ANN design (which is the last step of that), using the obtained PSO parameters from the previous steps, 5 PSO-ANN models were trained like ANN design section. The performance prediction of these models was also considered based on RMSE and R^2 results as presented in Table 5. As a result, the best PSO-ANN model for the MSS prediction is obtained as run No. 3 considering both results of RMSE and R^2. Model details about the evaluation of the developed PSO-ANN model are discussed in the following section.

5 Results and discussion

In this study, two non-linear AI models, i.e., ANN and PSO-ANN were developed to predict MSS caused by

tunneling. To evaluate the accuracy level of the aforementioned models, results of training (114 datasets) and testing (29 datasets) datasets, based on 80 % and 20 % of whole datasets, were considered and these results were compared to the measured MSS values. Three of the most well-known performance indices, namely RMSE, R^2 and variance account for (VAF) were used/computed to check the performance of the predictive models:

$$\text{RMSE} = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} \left[(x_i - x_p)^2 \right]} \tag{6}$$

$$\text{VAF} = \left[1 - \frac{\text{var}(x_i - x_p)}{\text{var}(x_i)} \right] \times 100 \tag{7}$$

$$R^2 = \frac{\left[\sum_{i=1}^{n} (x_i - x_{\text{mean}})^2 \right] - \left[\sum_{i=1}^{n} (x_i - x_p)^2 \right]}{\left[\sum_{i=1}^{n} (x_i - x_{\text{mean}})^2 \right]} \tag{8}$$

where x_i is the measured value, x_p is the predicted value, x_{mean} is mean of the measured value, 'var' is the sign for the variance and n is the number of data sets. If RMSE is zero, VAF is 100 (%) and R^2 is one, the model will be excellent. Results of models performance indices for developed models are presented in Table 6. Based on Table 6, the lowest values of RMSE and the highest value of VAF and R^2 are

Table 5 R^2 and RMSE values of the constructed PSO-ANN models

Network result																			
Run 1				Run 2				Run 3				Run 4				Run 5			
Training		Testing		Training		Testing		Training		Testing		Training		Testing		Training		Testing	
R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
0.964	0.049	0.957	0.053	0.953	0.054	0.956	0.064	0.973	0.043	0.968	0.052	0.965	0.045	0.946	0.065	0.968	0.046	0.955	0.054

obtained from the PSO-ANN model. For instance, RMSE equal to 0.04 and 0.05, for training and testing datasets, respectively, reveal that PSO-ANN model can predict MSS with high accuracy level. Furthermore, the relationships between the best datasets of ANN and PSO-ANN models in predicting MSS and the measured MSS for training and testing datasets are displayed in Figs. 5 and 6, respectively. Results of developed ANN model based on R^2 values are obtained at 0.939 and 0.940 for training and testing datasets, respectively, whereas values of 0.973 and 0.968 are achieved for R^2 of the selected PSO-ANN model. This indicates the superiority of the predictive PSO-ANN model compared to the proposed ANN predictive model. Note that the mentioned comparison was performed using normalized datasets for both measured and predicted values.

6 Sensitivity analysis

To determine the relative influence of the each input parameter on the output parameter, sensitivity analysis was performed using the cosine amplitude method [60]. This method is formulated in the following equation:

$$R_{ij} = \frac{\sum_{k=1}^{n}(x_{ik} \times x_{jk})}{\sqrt{\sum_{k=1}^{n} x_{ik}^2 \sum_{k=1}^{n} x_{jk}^2}} \tag{9}$$

where x_i and x_j represent input and output parameters, respectively, and n is the number of all data sets. R_{ij} is in the range of [0–1] and for the most influential parameter, R_{ij} will be equal 1. In the present paper, horizontal to vertical stress ratio, cohesion and Young's modulus were selected as input parameters, while, output parameter is MSS. The strengths of the relations between input and output parameters are given in Table 7. As can be seen from Table 7, horizontal to vertical stress ratio is the most influential parameter on MSS in this research.

7 Conclusion

Especially in urban areas, MSS prediction with a high degree of accuracy is very necessary. For this purpose, a new application of PSO-ANN model was proposed for predicting MSS caused by tunneling along the line 2 of Karaj subway. To check the performance capacity of PSO-ANN model, a pre-developed ANN model was applied. In this regard, 143 groups of datasets were prepared in 114 and 29 datasets for training and testing datasets, respectively. The values of horizontal to vertical stress ratio, cohesion and Young's modulus were taken as input parameters, while MSS was considered as an output parameter. To evaluate the authenticity and accuracy of the developed models,

Table 6 Performance prediction of the developed ANN and PSO-ANN models

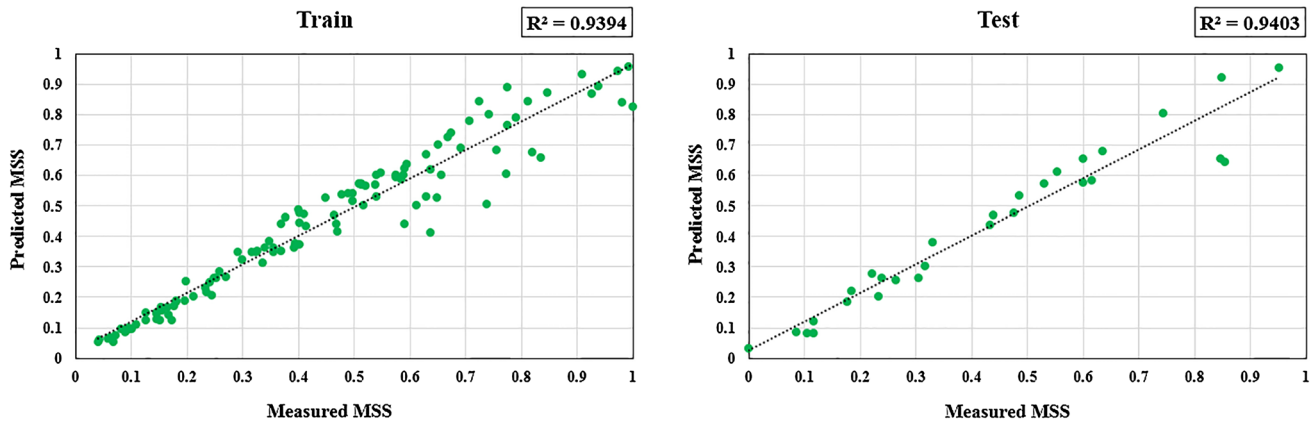| Predictive model	Training			Testing		
	R^2	VAF (%)	RMSE	R^2	VAF (%)	RMSE
ANN	0.939	93.94	0.063	0.94	94.02	0.064
PSO-ANN	0.9725	97.21	0.04	0.968	96.77	0.05

Fig. 5 R^2 values of the selected ANN datasets for training and testing
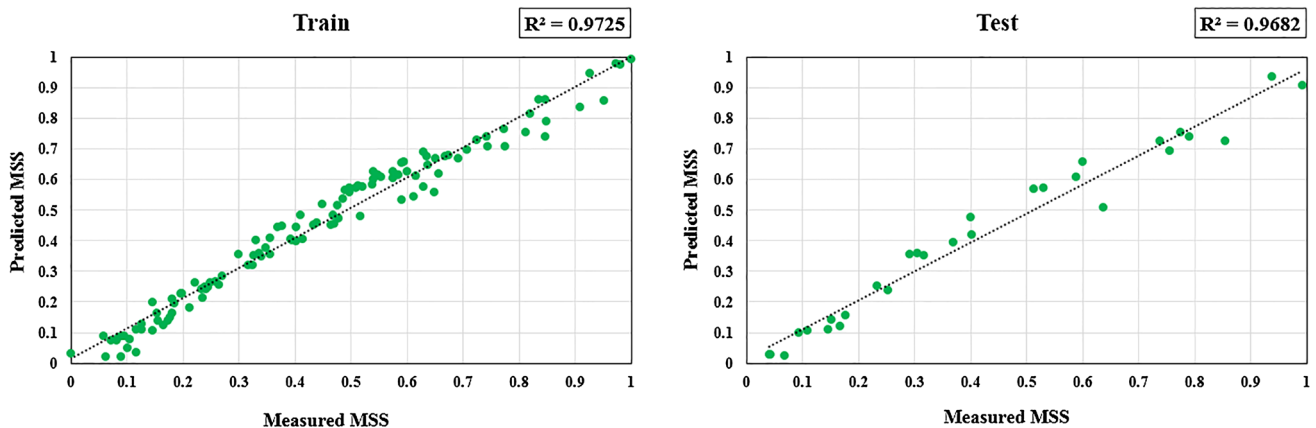
Fig. 6 R^2 values of the selected PSO-ANN datasets for training and testing

three performance indices, namely RMSE, VAF and R^2 were applied. The results revealed that PSO-ANN model can perform better than the ANN model for prediction of MSS. The R^2 equal to 0.9725 and 0.968 for training and testing datasets, respectively, indicate the high conformity of the PSO-ANN model in predicting MSS, while these values were obtained at 0.939 and 0.94 for ANN model, for training and testing datasets, respectively. Moreover, sensitivity analysis was carried out with input and output parameters and it was found that horizontal to vertical stress ratio has the strongest effect, based on considered datasets in this case study, on the MSS.

Table 7 Strengths of relation between input and output parameters

Input parameter	R_{ij}
Horizontal to vertical stress ratio	0.953
Cohesion	0.895
Young's modulus	0.891

References

1. Ocak I, Seker SE (2013) Calculation of surface settlements caused by EPBM tunneling using artificial neural network,

SVM, and Gaussian processes. Environ Earth Sci 69(5):1673–1683. doi:10.1007/s12665-012-2002-7

2. Chakeri H, Unver B (2014) A new equation for estimating the maximum surface settlement above tunnels excavated in soft ground. Environ Earth Sci 71:3195–3210. doi:10.1007/s12665-013-2707-2

3. Mohammadi SD, Naseri F, Alipoor S (2014) Development of artificial neural networks and multiple regression models for the NATM tunnelling-induced settlement in Niayesh subway tunnel. Bull Eng Geol Environ, Tehran. doi:10.1007/s10064-014-0660-2

4. Ocak I (2013) Interaction of longitudinal surface settlements for twin tunnels in shallow and soft soils: the case of Istanbul Metro. Environ Earth Sci 69:1673–1683. doi:10.1007/s12665-012-2002-7

5. Farias MM, Junior AM, Assis AP (2004) Displacement control in tunnels excavated by the NATM: 3-D numerical simulations. Tunn Undergr Space Technol 19:283–293

6. Karakus M, Ozsan A, Basarir H (2007) Finite element analysis for the twin metro tunnel constructed in Ankara Clay-Turkey. Bull Eng Geol Environ 66:71–79

7. Schmidt B (1969) A method of estimating surface settlement above tunnels constructed in soft ground. Can Geotech J 20:11–22

8. Attewell PB, Farmer IW (1974) Ground disturbance caused by shield tunneling in a stiff. Can Geotech J 11:380–395

9. Ocak I (2014) A new approach for estimating the transverse surface settlement curve for twin tunnels in shallow and soft soils. Environ Earth Sci 72:2357–2367. doi:10.1007/s12665-014-3145-5

10. Atkinson JH, Potts DM (1977) Subsidence above shallow tunnels in soft ground. ASCE Geotechnical Eng Div, pp 59–64

11. Hamza M, Ata A, Roussin A (1999) Ground movements due to construction of cut-and-cover structures and slurry shield tunnel of the Cairo Metro. Tunn Undergr Sp Tech 14(3):281–289

12. Chi SY, Chern JC, Lin CC (2001) Optimized Back-Analysis for Tunneling-Induced Ground Movement Using Equivalent Ground Loss Model. Tunn Undergr Sp Tech 16:159–165

13. Chou WI, Bobet A (2002) Predictions of ground deformations in shallow tunnels in clay. Tunn Undergr Sp Tech 17:3–19

14. Park KH (2005) Analytical solution for tunneling-induced ground movement in clays. Tunn Undergr Sp Tech 20(3):249–261. doi:10.1016/j.tust.2004.08.009

15. Ercelebi SG, Copur H, Ocak I (2011) Surface settlement predictions for Istanbul Metro tunnels excavated by EPB-TBM. Environ Earth Sci 62(2):357–365. doi:10.1007/s12665-010-0530-6

16. Yagiz S, Gokceoglu C, Sezer E, Iplikci S (2009) Application of two non-linear prediction tools to the estimation of tunnel boring machine performance. Eng Appl Artif Intel 22(4):808–814

17. Yagiz S, Gokceoglu C (2010) Application of fuzzy inference system and nonlinear regression models for predicting rock brittleness. Expert Syst Appl 37(3):2265–2272. doi:10.1016/j.eswa.2009.07.046

18. Amiri M, Bakhshandeh Amnieh H, Hasanipanah M, Mohammad Khanli L (2016) A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure. Eng Comput. doi:10.1007/s00366-016-0442-5

19. Ghasemi E (2016) Particle swarm optimization approach for forecasting backbreak induced by bench blasting. Neural Comput Appl. doi:10.1007/s00521-016-2182-2

20. Jahed Armaghani D, Hajihassani M, Mohamad ET, Marto A, Noorani SA (2014) Blasting-induced flyrock and ground vibration prediction through an expert artificial neural network based on particle swarm optimization. Arab J Geosci 7:5383–5396

21. Jahed Armaghani D, Hasanipanah M, Mohamad ET (2015) A combination of the ICA-ANN model to predict air overpressure resulting from blasting. Eng Comput. doi:10.1007/s00366-015-0408-z

22. Simpson PK (1990) Artificial neural system: foundation, paradigms, applications and implementations. Pergamon, New York

23. Haykin S (1999) Neural networks, 2nd edn. Prentice-Hall, Englewood Cliffs

24. Hasanipanah M, Jahed Armaghani D, Khamesi H, Bakhshandeh Amnieh H, Ghoraba S (2015) Several non-linear models in estimating air-overpressure resulting from mine blasting. Eng Comput. doi:10.1007/s00366-015-0425-y

25. Monjezi M, Bahrami A, Yazdian Varjani A (2010) Simultaneous prediction of fragmentation and flyrock in blasting operation using artificial neural networks. Int J Rock Mech Min Sci 47:476–480

26. Monjezi M, Hasanipanah M, Khandelwal M (2013) Evaluation and prediction of blast-induced ground vibration at Shur River Dam, Iran, by artificial neural network. Neural Comput Appl 22:1637–1643

27. Fausett LV (1994) Fundamentals of neural networks: architecture, algorithms and applications. Prentice-Hall, Englewood Cliffs

28. Khandelwal M, Singh TN (2009) Prediction of blast-induced ground vibration using artificial neural network. Int J Rock Mech Min Sci 46:1214–1222

29. Wang XG, Tang Z, Tamura H, Ishii M, Sun WD (2004) An improved backpropagation algorithm to avoid the local minima problem. Neurocomputing 56:455–460

30. Adhikari R, Agrawal RK (2011) Effectiveness of PSO based neural network for seasonal time series forecasting. Indian International Conference on Artificial Intelligence (IICAI). Tumkur, India, pp 232–244

31. Kennedy J, Eberhart RC (1995) Particle swarm optimization. Proc. IEEE International Conference on Neural Networks (Perth, Australia), IEEE Service Center, Piscataway, p 1942–1948

32. Jahed Armaghani D, Raja SNSB, Faizi K, Rashid ASA (2015) Developing a hybrid PSO–ANN model for estimating the ultimate bearing capacity of rock-socketed piles. Neural Comput Appl. doi:10.1007/s00521-015-2072-z

33. Khamesi H, Torabi S, Mirzaei-Nasirabad H, Ghadiri Z (2015) Improving the Performance of Intelligent Back Analysis for Tunneling Using Optimized Fuzzy Systems: case Study of the Karaj Subway Line 2 in Iran. J Comput Civ Eng. doi:10.16.1061/(ASCE)CP.1943-5487.0000421

34. Poli R, Kennedy J, Blackwell T (2007) Particle swarm optimization. Swarm Intell 1(1):33–57

35. Shi Y, Eberhart R (1998) Parameter selection in particle swarm optimization. Proceedings of the seventh annual conference on evolutionary. Springer, New York, pp 591–600

36. Das MT, Dulger LC (2009) Signature verification (SV) toolbox: application of PSO-NN. Eng Appl Artif Intell 22(4):688–694

37. Swingler K (1996) Applying neural networks: a practical guide. Academic Press, New York

38. Looney CG (1996) Advances in feed-forward neural networks: demystifying knowledge acquiring black boxes. IEEE Trans Knowl Data Eng 8(2):211–226

39. Nelson M, Illingworth WT (1990) A Practical Guide to Neural Nets. Addison–Wesley, Reading MA

40. Hush DR (1989) Classification with neural networks: a performance analysis. Proceedings of the IEEE International Conference on Systems Engineering. Dayton, OH, USA, pp 277–280

41. Kanellopoulas I, Wilkinson GG (1997) Strategies and best practice for neural network image classification. Int J Remote Sens 18:711–725

42. Maulenkamp F, Grima MA (1999) Application of neural networks for the prediction of the unconfined compressive strength

(UCS) from Equotip hardness. Int J Rock Mech Min Sci 36:29–39

43. Ornek M, Laman M, Demir A, Yildiz A (2012) Prediction of bearing capacity of circular footings on soft clay stabilized with granular soil. Soil Found 52:69–80

44. Ceryan N, Okkan U, Kesimal A (2013) Prediction of unconfined compressive strength of carbonate rocks using artificial neural networks. Environ Earth Sci 68:807–819

45. Hecht-Nielsen R (1987) Kolmogorov's mapping neural network existence theorem. In: Proceedings of the First IEEE International Conference on Neural Networks, San Diego, pp 11–14

46. Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal Approximators. Neural Networks 2:359–366

47. Baheer I (2000) Selection of methodology for modeling hysteresis behavior of soils using neural networks. J Comput Aid Civil Infrastruct Eng 5:445–463

48. Sonmez H, Gokceoglu C, Nefeslioglu HA, Kayabasi A (2006) Estimation of rock modulus: for intact rocks with an artificial neural network and for rock masses with a new empirical equation. Int J Rock Mech Min Sci 43:224–235

49. Sonmez H, Gokceoglu C (2008) Discussion on the paper by H. Gullu and E. Ercelebi, "A neural network approach for attenuation relationships: an application using strong ground motion data from Turkey. Eng Geol 97:91–93

50. Ripley BD (1993) Statistical aspects of neural networks. In: Barndoff- Neilsen OE, Jensen JL, Kendall WS, editors. Networks and chaos-statistical and probabilistic aspects. London: Chapman & Hall, pp. 40–123

51. Paola JD (1994) Neural network classification of multispectral imagery. MSc thesis, The University of Arizona, USA

52. Wang C (1994) A theory of generalization in learning machines with neural application. PhD thesis, The University of Pennsylvania, USA

53. Masters T (1994) Practical neural network recipes in C++. Academic Press, Boston MA

54. Kaastra I, Boyd M (1996) Designing a neural network for forecasting financial and economic time series. Neurocomputing 10:215–236

55. Hajihassani M, Jahed Armaghani D, Sohaei H, Mohamad ET, Marto A (2014) Prediction of airblast-overpressure induced by blasting using a hybrid artificial neural network and particle swarm optimization. Appl Acoust 80:57–67

56. Momeni E, Jahed Armaghani D, Hajihassani M, Amin MFM (2015) Prediction of uniaxial compressive strength of rock samples using hybrid particle swarm optimization-based artificial neural networks. Measurement 60:50–63

57. Zorlu K, Gokceoglu C, Ocakoglu F, Nefeslioglu HA, Acikalin S (2008) Prediction of uniaxial compressive strength of sandstones using petrography-based models. Eng Geol 96(3):141–158

58. Gordan B, Jahed Armaghani D, Hajihassani M, Monjezi M (2015) Prediction of seismic slope stability through combination of particle swarm optimization and neural network. Eng Comput. doi:10.1007/s00366-015-0400-7

59. Tonnizam Mohamad E, Jahed Armaghani D, Momeni E, Alavi Nezhad Khalil Abad SV (2014) Prediction of the unconfined compressive strength of soft rocks: a PSO-based ANN approach. Bull Eng Geol Environ. doi:10.1007/s10064-014-0638-0

60. Yang Y, Zang O (1997) A hierarchical analysis for rock engineering using artificial neural networks. Rock Mech Rock Eng 30:207–222