Forecasting Surface Settlement Caused by Shield Tunneling Using ANN-BBO Model and ANFIS Based on Clustering Methods

Hadi Fattahi*, Zohreh Bayatzadehfard;
Department of Mining Engineering, Arak University of Technology, Arak, Iran.

Received: 25 Nov 2016 Revised: 22 August 2017

Abstract

Maximum surface settlement (MSS) is an important parameter for the design and operation of earth pressure balance (EPB) shields that should be determined before tunneling. Artificial intelligence (AI) methods are accepted as a technology that offers an alternative way to tackle highly complex problems that cannot be modeled in mathematics. They can learn from examples and they are able to handle incomplete data and noisy data. The adaptive network–based fuzzy inference system (ANFIS) and hybrid artificial neural network (ANN) with biogeography-based optimization algorithm (ANN-BBO) are kinds of AI systems that were used in this study to make a prediction model for the MSS caused by EPB shield tunneling. Two ANFIS models were implemented, the ANFIS-subtractive clustering method (ANFIS-SCM) and ANFIS-fuzzy c-means clustering method (ANFIS-FCM). The estimation abilities offered using three models were presented by using field data from the Bangkok Subway Project in Thailand. In these models, depth, distance from shaft, ground

Corresponding author: h.fattahi@arakut.ac.ir

[Downloaded from isharat.khu.ac.ir on 2026-10-29]

[DOR: 20.1001.1.22286837.1397.12.5.3.5]

[DOI: 10.18869/acadpub.jeg.12.5.55]

water level from tunnel invert, average face pressure, average penetrate rate, pitching angle, tail void grouting pressure and percent tail void grout filling were utilized as the input parameters, while the MSS was the output parameter. To compare the performance of models for MSS prediction, the coefficient of correlation (R^2) and mean square error (MSE) of the models were calculated, indicating the good performance of the ANFIS-SCM model.

Keywords: Maximum surface settlements, EPB shield tunneling, Adaptive network-based fuzzy inference system, Artificial neural network, Biogeography-based optimization algorithm.

1. Introduction

Urban population growth and quick economic development have been increasing the necessity for underground space utilization. Tunneling plays an important role in the underground engineering, providing a solution for human needs with minimum surface impacts [1]. Of all tunneling methods, Earth Pressure Balance (EPB) shield tunneling is considered to be a suitable tunneling method when surface settlements must be avoided by controlling face stability and underground water inflow [2]. Some studies have been done in this area, which it is referred to some of them. Shao and Lan [3] presented an optimization control method based on the particle swarm optimization algorithm for the screw conveyor rotating speed when considering tunnel face stability. Hu et al. [4] developed an EPB control model with the real time measured data by using the adaptive network-based fuzzy inference system (ANFIS).

[Downloaded from isharat.khu.ac.ir on 2026-10-29]

[DOR: 20.1001.1.22286837.1397.12.5.3.5]

[DOI: 10.18869/acadpub.jeg.12.5.55]

It has been established that the maximum surface settlement (MSS) depends on various factors, including tunnel geometry (tunnel depth (m), distance from launching station (m)), geological conditions (geology at tunnel crown, geology at tunnel invert, ground water level from tunnel invert (m)) and shield operation factors (face pressure (kPa), penetration rate (mm/min), pitching angle (º), tail void grouting pressure (bar), percent tail void grout filling) [5-7]. The MSS caused by EPB shield tunneling is one of the important parameters that must be predicted quickly. However, relatively little quick research has been done in this area. There are several methods for estimation of MSS, for example empirical and analytical relations, 2-D and 3-D numerical analyses, statistical methods and artificial intelligence methods. For example in the field of artificial intelligence methods for prediction of MSS, Suwansawat and Einstein [7] utilized artificial neural networks (ANNs) model. In terms of MSS modeling, although previous studies are valuable, offering new models with accurate results can eliminate many field problems relating to this scope. In this paper, the application of artificial intelligence methods for data analysis named ANFIS-subtractive clustering method (ANFIS-SCM) and ANFIS-fuzzy c-means clustering method (ANFIS-FCM) and hybrid ANN with biogeography-based optimization (ANN-BBO) to estimate the MSS are demonstrated. In these models (ANFIS-SCM, ANFIS-FCM, ANN-BBO), depth (m), distance from shaft (m), invert to WT (m), average face pressure, average penetration, pitching (º), grouting pressure (bar) and grout filling (%) are utilized as the input parameters, while the MSS (mm) is the output parameter. The estimation abilities offered using AI

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

models are presented by using field data of Bangkok Subway Project in Thailand that this is the first phase of an integrated transportation plan for Bangkok, to be implemented in conjunction with other schemes, by the Mass Rapid Transit Authority (MRTA) [7].

The main scope of this paper is the application and comparison between three models (ANFIS-SCM, ANFIS-FCM, ANN-BBO) for estimation of MSS caused by EPB shield tunneling and investigation of the performance and convergence of AI models.

2. Description of selected models

Several AI techniques employed in this study include ANN, BBO, ANFIS-FCM and ANFIS-SCM. A brief overview of these techniques is presented here.

2.1 Adaptive network-based fuzzy inference system (ANFIS)

In this section, we present the basic theory of ANFIS model. Both ANN and fuzzy logic (FL) are used in ANFIS architecture [8-11]. ANFIS consists of if–then rules and couples of input–output. Also for ANFIS training, learning algorithms of a neural network are utilized [12] ,[13]. To better understanding ANFIS, an example with two inputs (x and y) and one output (f) is shown in Figure 1.

The architecture of ANFIS consists of five layers (Figure 1), and a brief introduction of the model is as follows.

Layer 1: Each node i in this layer generates a membership grades of a linguistic label. For instance, the node function of the ith node that is defined as given in Equation (1),

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

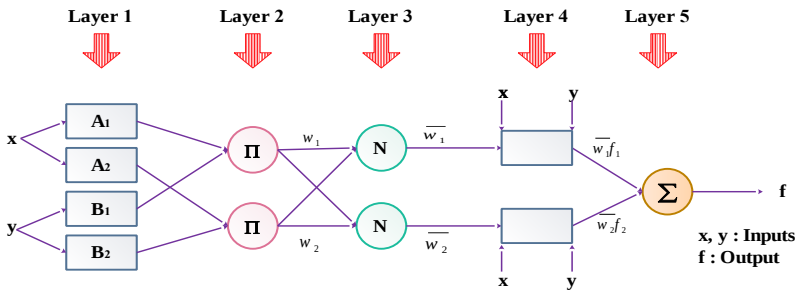


Figure 1. ANFIS architecture for two-input [13]

$$Q_i^1 = \mu_{Ai}(x) = \frac{1}{1+\left[\left(\frac{x-v_i}{\sigma_i}\right)^2\right]^{b_i}} \tag{1}$$

where, x is the input to node i, and Ai is the linguistic label (small, large,…) associated with this node; and $\{\sigma_i, V_i, b_i\}$, is the parameter set that changes the shapes of the membership function (MF). Parameters in this layer are referred to as the "premise parameters".

Layer 2: Each node in this layer calculates the "firing strength" of each rule via multiplication (Eq. 2).

$$Q_i^2 = W_i = \mu_{Ai}(x).\mu_{Bi}(y) \qquad i = 1,2 \tag{2}$$

Layer 3: The ith node of this layer calculates the ratio of the ith rule' firing strength to the sum of all rules firing strengths that is defined as shown in Eq.3.

$$Q_i^3 = \overline{W}_i = \frac{w_i}{\sum_{j=1}^{2} w_j}, \qquad i = 1,2 \tag{3}$$

For convenience, outputs of this layer will be called "normalized firing" strengths.

Layer 4: Every node i in this layer is a node function which is defined in

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

Eq.4,

$$Q_i^4 = \bar{W}_i f_i = \bar{W}_i (p_i x + q_i y + r_i) \tag{3}$$

where, $\bar{W}_i$ is the output of layer 3. Parameters in this layer will be referred to as "consequent parameters".

Layer 5: The single node in this layer is a circle node labeled R that computes the "overall output" as the summation of all incoming signals that is defined as shown in Eq.5.

$$Q_i^5 = Overall\ Output = \sum \bar{W}_i f_i = \frac{\sum w_i f_i}{\sum w_i} \tag{5}$$

For a given data set, different ANFIS models can be constructed, using different identification methods. SCM and FCM are two methods utilized in this study to identify the antecedent MFs. The ANFIS-SCM combines the subtractive clustering method and ANFIS. The ANFIS-FCM is the combination of a fuzzy c-means clustering method and ANFIS.

Clustering methods are extremely important for explorative data analysis. Two types of these methods are described below.

2.1.1. Subtractive clustering method (SCM)

The subtractive clustering method (SCM) is suggested by Chiu [14]. It clusters data points in an unsupervised way by measuring the potential of data points in the feature space. When there is not a clear idea how many clusters there should be utilized for a given data set, it can be used for estimating the cluster centers and the number of clusters. The SCM assumes that each data point is a potential cluster center and calculates the potential for each data point based on the density of surrounding data points. Then the data point with the highest potential is selected as the first cluster center, and

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

the potential of data points near the first cluster center (within the influential radius) is destroyed. Then data points with the highest remaining potential as the next cluster center and the potential of data points near the new cluster center is destroyed [15]. The influential radius is critical for determining the number of clusters. A smaller radius leads to many smaller clusters in the data space, which results in more rules, and vice versa. The it is significant to select a proper influential radius for clustering the data space [16].

2.1.2. Fuzzy C-means clustering method (FCM)

Fuzzy c-means method (FCM) is suggested by Bezdek [17]. The FCM partitions a collection of n vector X_i, $i = 1, 2, ..., n$, into c fuzzy groups, and finds a cluster center in each group such that a cost function of dissimilarity measure is minimized. The steps of FCM algorithm are therefore, first described in brief.

Step 1: Chose the cluster centers c_i, $i = 1, 2, ..., c$, randomly from the n points $\{X_1, X_2, X_3, ..., X_n\}$.

Step 2: Compute the membership matrix U using Eq.6,

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}}\right)^{2/m-1}} \tag{6}$$

where, $d_{ij} = \|c_i - x_j\|$, is the Euclidean distance between ith cluster center and jth data point, and m is the fuzziness index.

Step 3: Compute the cost function according to the Eq. 7. Stop the process if it is below a certain threshold.

$$J(U, c_1, ..., c_2) = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} . \sum_{j=1}^{n} \mu_{ij}^m d_{ij}^2 \tag{7}$$

[Downloaded from isharat.khu.ac.ir on 2026-10-29]

[DOR: 20.1001.1.22286837.1397.12.5.3.5]

[DOI: 10.18869/acadpub.jeg.12.5.55]

Step 4: Compute new c fuzzy cluster centers c_i, $i = 1, 2, ..., c$, using the Eq. 8.

$$c_i = \frac{\sum_{j=1}^{n} \mu_{ij}^m x_j}{\sum_{j=1}^{n} \mu_{ij}^m} \tag{8}$$

go to step 2.

2.2. Hybrid artificial neural network with biogeography-based optimization (ANN-BBO)

2.2.1 Artificial neural network (ANN)

Artificial Neural Networks (ANNs) are parallel information processing methods, which can express nonlinear relationships and complex numbers of input–output training patterns from the experimental data. ANNs provide a nonlinear mapping between outputs and inputs by its intrinsic ability [18,19]. The success in obtaining a reliable and robust network depends on the correct data preprocessing, correct architecture selection, and correct network training choice [20]. The ANN is trained by performing optimization of weights for each node interconnection and bias terms; until the values output at the output layer neurons are as close as possible to the actual outputs.

In this regard, the data are split into two sets, a training data set and a testing data set. The model is produced using only the training data. The testing data are utilized to estimate the accuracy of the model performance. In training a ANN, the objective is to find an optimum set of weights. When the number of weights is higher than the number of available data, the error in-fitting the non-trained data initially decreases, but then increases as the network becomes over-trained. In contrast, when the number of weights is

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

smaller than the number of data, the over-fitting problem is not crucial [21].

2.2.2. ANN-BBO model

ANN training has traditionally been carried out using a BP algorithm. However, this approach has some drawbacks, such as local minimum trapping, over-fitting, and weight interference, which have complicated ANN training. In contrast, the optimization algorithms have balanced exploration and exploitation capabilities; therefore, it does not get stuck in local minima [22-24]. In the present study, the BBO [25] is proposed for optimizing the weights of ANN.

Figure 2 shows the formation of two connection weight matrixes ω_1 and ω_2. The former represents the connection weight matrix between the input layer and hidden layer, and the latter represents the connection weight matrix between the hidden layer and the output layer. The total weights can be defined as

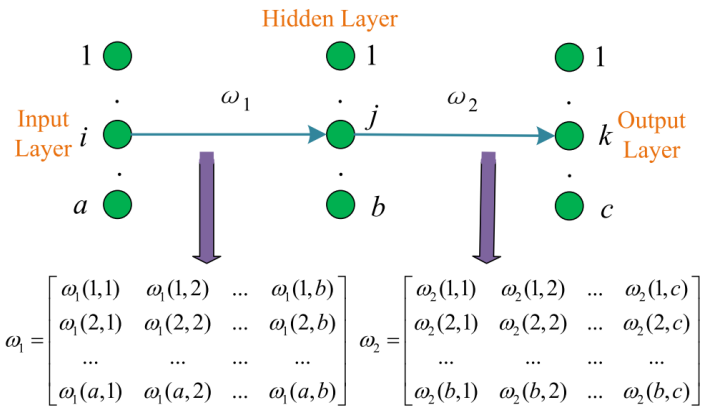
$$\omega = \left[\nu(\omega_1), \nu(\omega_2)\right] = \left[\omega_1(1,1), ..., \omega_1(a,b), \omega_2(1,1), ..., \omega_2(b,c)\right] \quad (9)$$

where ν represents the vectorization operation. a, b, and c denote the number of input, hidden, and output neurons. ω represents the weights that need to be trained [26].

[Downloaded from isharat.khu.ac.ir on 2026-10-29]

[DOR: 20.1001.1.22286837.1397.12.5.3.5]

[DOI: 10.18869/acadpub.jeg.12.5.55]



$$\omega_1 = \begin{bmatrix} \omega_1(1,1) & \omega_1(1,2) & \dots & \omega_1(1,b) \\ \omega_1(2,1) & \omega_1(2,2) & \dots & \omega_1(2,b) \\ \dots & \dots & \dots & \dots \\ \omega_1(a,1) & \omega_1(a,2) & \dots & \omega_1(a,b) \end{bmatrix} \quad \omega_2 = \begin{bmatrix} \omega_2(1,1) & \omega_2(1,2) & \dots & \omega_2(1,c) \\ \omega_2(2,1) & \omega_2(2,2) & \dots & \omega_2(2,c) \\ \dots & \dots & \dots & \dots \\ \omega_2(b,1) & \omega_2(b,2) & \dots & \omega_2(b,c) \end{bmatrix}$$

Figure 2. The two connection weight matrix within a feed forward neural network

The fitness function (FF) is defined in the following four steps:

(1) The outputs of hidden neurons are as follows:

$$y_i = A_H\left(\sum_{i=1}^{a} \omega_1(i,j)x_i\right) \quad j = 1,2,\dots,b \tag{10}$$

where x_i represents the input of ith input neuron and y_j the output of jth hidden neuron. A_H is the activation function of hidden layer in the form of

$$A_H(x) = \frac{1}{\exp(-x)+1} \tag{11}$$

(2) The outputs of output neurons are as follows

$$O_k = A_O\left(\sum_{j=1}^{b} \omega_2(j,k)y_j\right) \quad k = 1,2,\dots,c \tag{12}$$

here, A_O represents the activation function of output layer, usually a linear function.

(3) The error between output and target values is calculated as mean squared error (MSE)

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

$$E_l = MSE\left(\sum_{k=1}^{c}\left(O_k - T_k\right)\right) \quad l = 1, 2, ..., N_s \tag{13}$$

where T_k represents the kth realistic value and N_S represents the sample number.

(4) The FF is deduced as the averaged MSE of all samples

$$f\left(\omega\right) = \sum_{l=1}^{N_S} E_l \tag{14}$$

The task was the minimization of this f(ω), that is, to force the output results to approximate to corresponding realistic class types. In closing, the training procedure was transformed into an optimization problem, in which the average MSE between the output and the target was treated as the FF and the weights/biases of ANN were regarded as the variables. BBO was used to solve this optimization problem. The flowchart of the ANN-BBO model is shown in Figure 3.

3. Study area and data

The main aim of this work is to implement the above methodologies in the problem of MSS estimation. The dataset applied in this study for determining the relationship among the set of inputs and output variables are gathered from open source literature [7]. The collected data sets used to construct the database are from Bangkok Subway Project in Thailand.

This project consisting of 20 km of twin tunnels is classified into two main tunnel sections namely, the north tunnel section and the south tunnel section.

[Downloaded from isharat.khu.ac.ir on 2026-10-29]

[DOR: 20.1001.1.22286837.1397.12.5.3.5]

[DOI: 10.18869/acadpub.jeg.12.5.55]

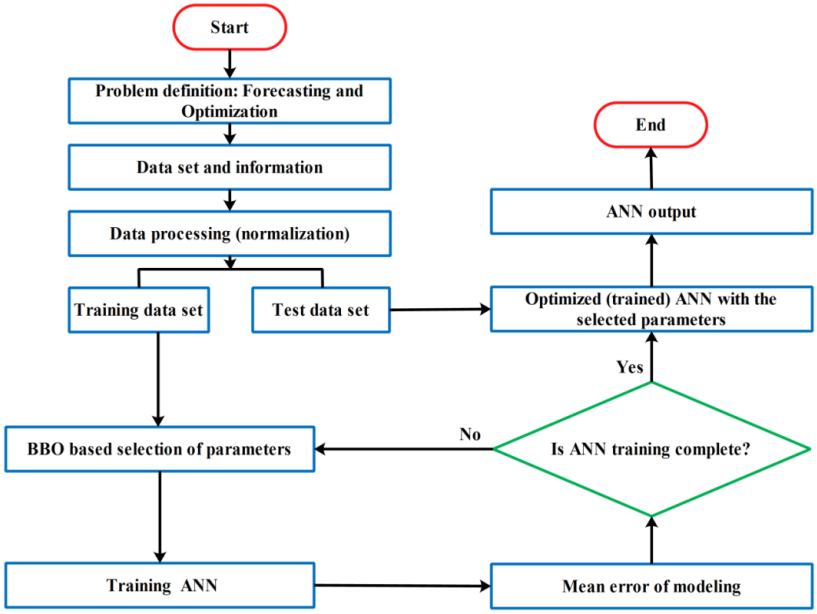


Figure 3. Flowchart of the ANN-BBO model

Each data set contains the parameters of depth (m), distance from shaft (m), ground water level from tunnel invert (invert to WT) (m), average face pressure (kPa), average penetrate rate (mm/min), pitching angle (º), grouting pressure (bar), grout filling (%), geology at tunnel crown, geology at tunnel invert and measured MSS (mm). In this paper, we have used all parameters except geology at the tunnel crown and geology at the tunnel invert. Partial dataset used in this study are presented in Table 1. Also, descriptive statistics of the all data sets are shown in Table 2. All data (49 data sets) were divided into two subsets: 80% of the total data (39 cases) was allotted to training data of the models construction and 20% of the total data (10 cases) was selected for test data used to assess the reliability of the developed models.

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

Table 1. Partial dataset used for constructing the AI models [7]

Case No.	Depth (m)	Distance from shaft (m)	Invert to WT (m)	Average face pressure (kPa)	Average penetrate rate (mm/min)	Pitching angle (deg)	Grouting pressure (bar)	Grout filling (%)	Surface settlement (mm)
							Input parameters		Output parameter
1	18.2	33.6	0.65	34.5	33.5	-0.07	3.03	92	-60.5
2	18.61	58.8	0.24	32	42.4	-1.01	3.03	100	-51.4
3	18.7	62.4	0.15	31	41.65	-1.05	3.03	100	-47.9
4	19.21	82.8	-0.36	54.5	34.45	-1.38	7.4	122	-31.9
5	19.63	99.6	-0.78	84.5	32.55	-0.88	5.6	116	-15.9
6	20.17	121.2	-1.32	100.5	30	-1.12	5.3	110	-13.5
7	21.1	158.4	-2.25	131	26.4	-1.11	2.5	121	-15.7
8	22.06	196.8	-3.21	123	29.75	-1.11	2.5	117	-16.8
9	23.09	223.2	-4.24	65.5	40.65	-1.14	2.5	119	-21.5
10	23.22	252	-4.37	14.5	48.9	-0.45	2.5	127	-43.8

Table 2. Statistical description of dataset utilized for construction of three models

Parameter	Average	Min	Max
Depth (m)	22.05	17.89	24.82
Distance from shaft (m)	1320.27	33.60	3055.20
Invert to WT (m)	-3.20	-5.97	0.96
Average face pressure (kPa)	54.73	14.50	131.00
Average penetrate (mm/min)	42.63	20.10	76.85
Pitching angle (deg)	0.05	-1.38	1.43
Grouting pressure (bar)	2.78	2.30	7.40
Grout filling (%)	125.96	70.00	224.00
surface settlement (mm)	-28.09	-60.50	-6.25

4. Pre-processing of data and performance criteria

The actual data is often incomplete, inconsistent and noisy. Pre-processing methods used for data normalization. This work ensures that the raw data retrieved from database is perfectly suitable for modeling. In this

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

study, all data samples are normalized to adapt to the interval [0, 1] according to the following linear mapping function (Eq. 15),

$$x_m = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{15}$$

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

Furthermore, to consider the performances of the AI models, MSE and correlation coefficient (R^2) were chosen to be the measure of accuracy. MSE and R^2 could be defined, respectively, as follows,

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

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

Let t_k be the actual value and $\hat{t}_k$ be the predicted value of the kth observation and n be the number of observations.

5. Estimation of maximum surface settlement (MSS) using ANFIS models

In this section, ANFIS was utilized to build a prediction model for estimation of MSS from available data, using MATLAB environment. Two ANFIS models were implemented, SCM and FCM. Figure 4 displays the fuzzy architecture of the ANFIS.

[Downloaded from isharat.khu.ac.ir on 2026-10-29]

[DOR: 20.1001.1.22286837.1397.12.5.3.5]

[DOI: 10.18869/acadpub.jeg.12.5.55]

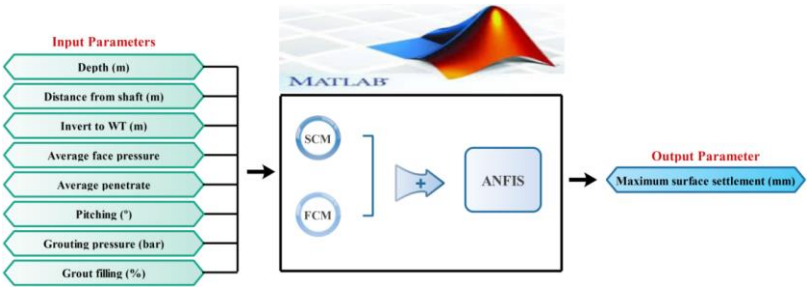


Figure 4. Architecture of the ANFIS based on the SCM and FCM.

A dataset that includes 49 data points was employed in the current study, while 39 data points (80%) were utilized for constructing the model and the remainder data points (10 data points) were used for assessment of degree of accuracy and robustness. The characterizations of the ANFIS models are given in Tables 3 and 4. Also, the MFs of the input parameters for different models are illustrated in Figures. 5 and 6.

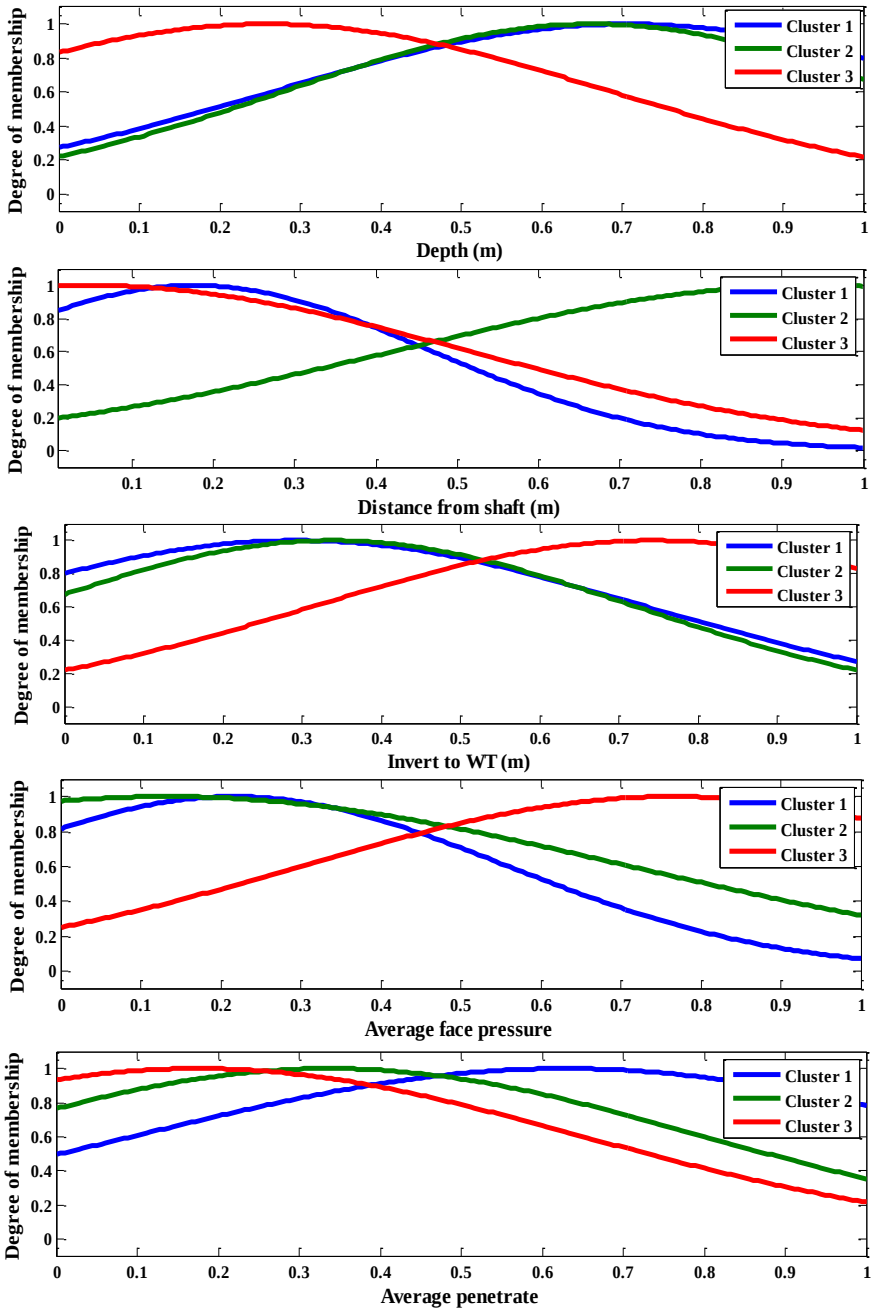
Table 3. Characterizations of the ANFIS models

ANFIS parameter	ANFIS–SCM	ANFIS–FCM
MF type	Gaussian	Gaussian
Output MF	Linear	Linear
Number of nodes	65	47
Number of linear parameters	27	18
Number of nonlinear parameters	48	32
Total number of parameters	75	50
Number of training data pairs	39	39
Number of testing data pairs	10	10
Number of fuzzy rules	3	2

Table 4. The optimal parameters of the ANFIS models

parameter	ANFIS models
Error goal	0
The initial step size	0.01
Step size decrease rate	0.6
Step size increase rate	1.1

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]



[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

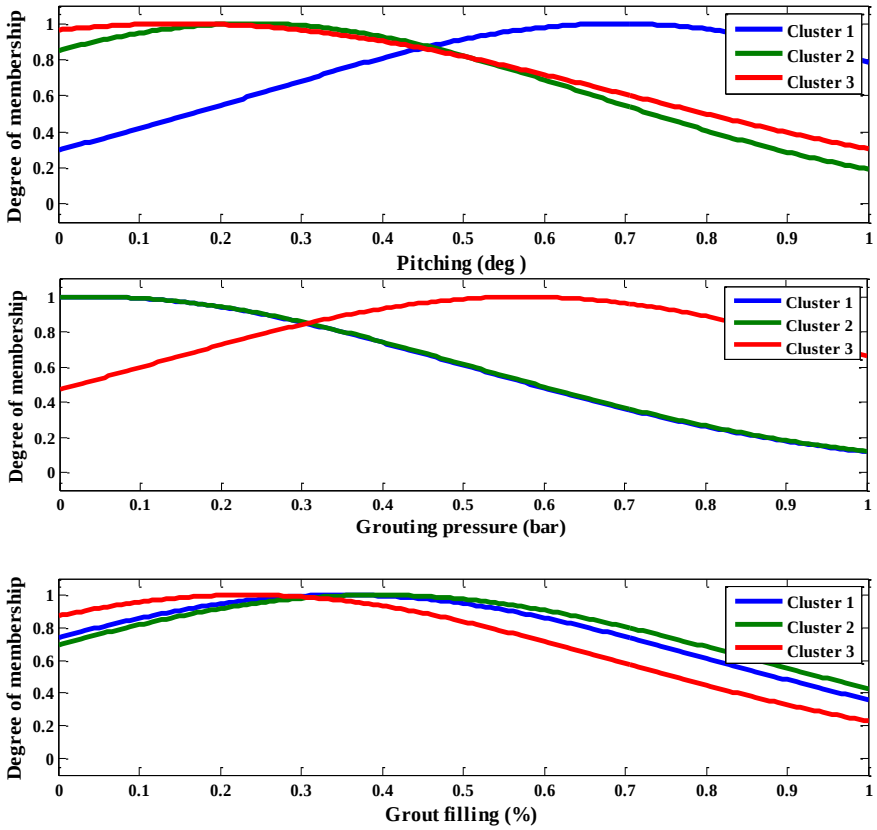
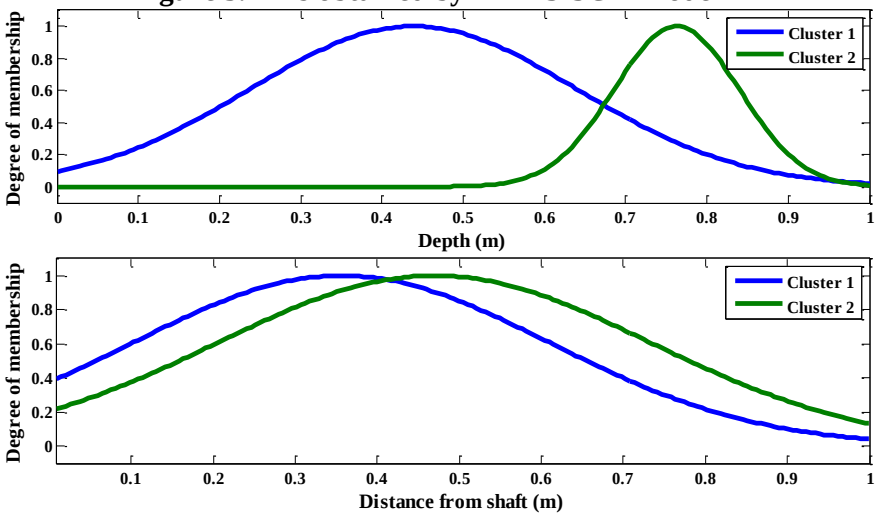
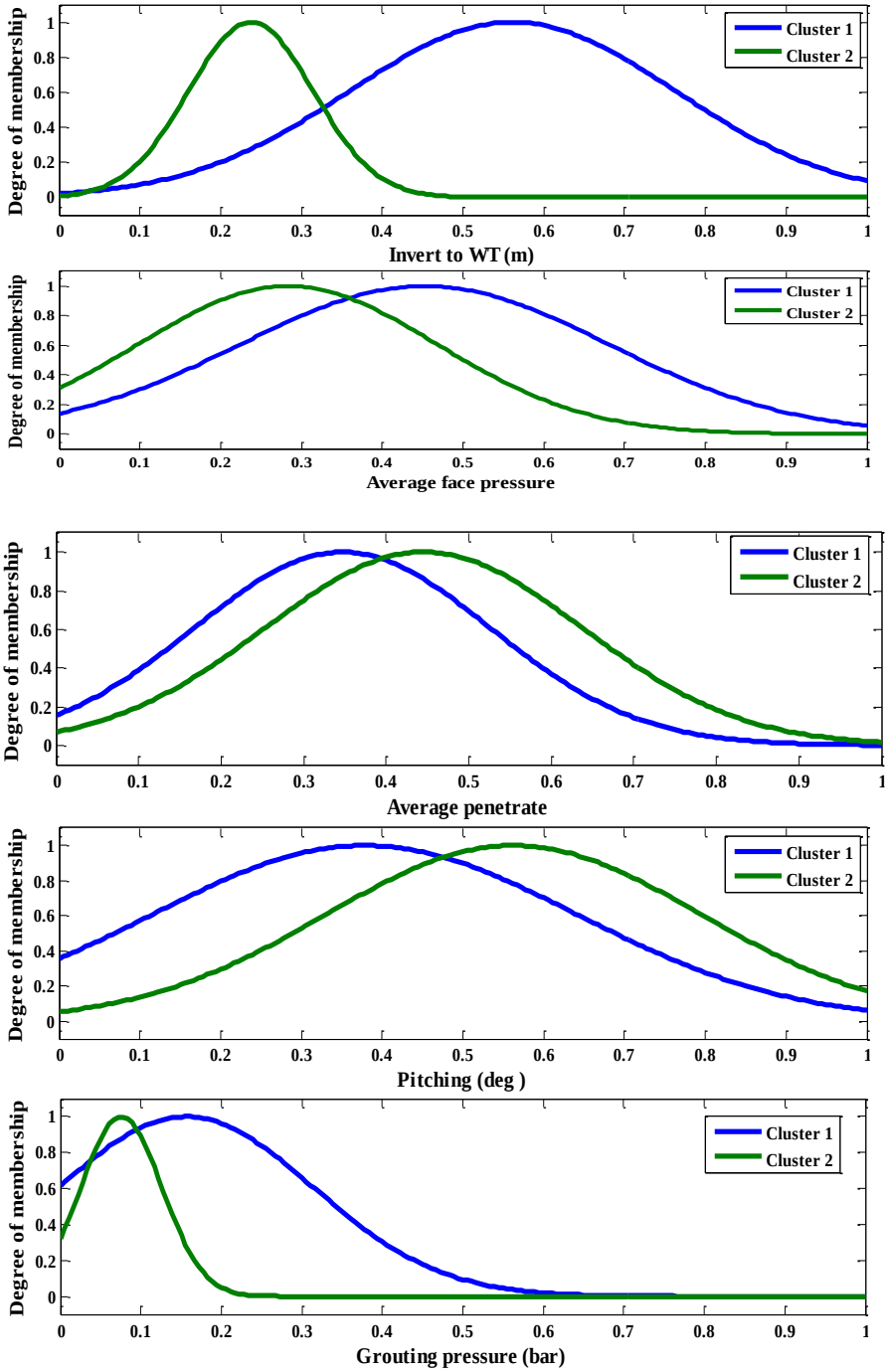


Figure 5. MFs obtained by ANFIS-SCM model



[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]



[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

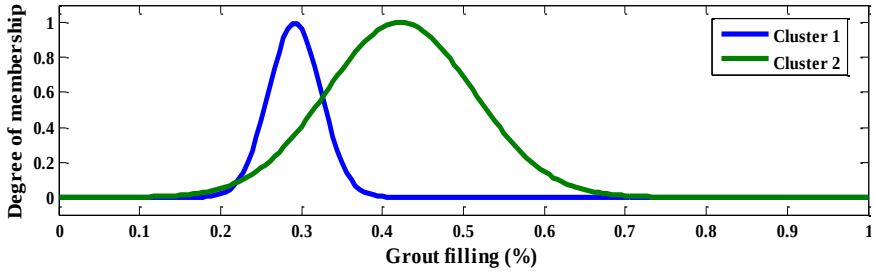


Figure 6. MFs obtained by ANFIS-FCM model.

5.1. Training and validation models

Part of the sensitivity analysis of ANFIS-SCM and ANFIS-FCM models are shown in Tables 5 and 6. A comparison between the results of two models for testing and training datasets is presented in Table7. Based on Table 7, the ANFIS-SCM model can obviously be considered as the best model. The point is that, in this model, MSE and R^2 (Figures. 7 and 8) values for training and testing datasets are minimum and maximum respectively, which means the complex relation between inputs and output were fully captured.

Table 5. Part of the sensitivity analysis of the ANFIS-SCM model

Influence radius	The number of periodic training process	Step size decrease rate	Step size increase rate	R^2_{Train}	MSE_{Train}	R^2_{Test}	MSE_{Test}
1.3	400	0.6	1.1	0.9627	0.0015	0.4341	0.1432
1.2	200	0.5	1.1	0.9545	0.0019	0.5080	0.0584
1.2	100	0.6	1.2	0.9488	0.0021	0.6076	0.0513
1.6	150	0.1	1.5	0.8382	0.0069	0.6244	0.0442
1.5	100	0.6	1.2	0.8384	0.0068	0.6264	0.0440
1.1	100	1.9	0.1	0.9318	0.0029	0.6284	0.0455
1.3	100	0.1	2	0.9469	0.0022	0.7122	0.0387
1.3	100	0.6	1.2	0.9447	0.0023	0.8571	0.0263
1.3	100	0.3	1.2	0.9429	0.0024	0.8794	0.0290
1.3	100	0.6	1.1	0.9431	0.0024	0.8806	0.0287

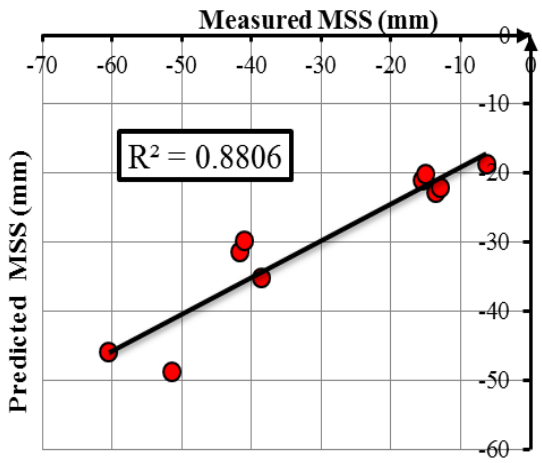
[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

Table 6. Part of the sensitivity analysis of the ANFIS-FCM model

Number of clusters	The number of periodic training process	Step size decrease rate	Step size increase rate	R^2_{Train}	MSE_{Train}	R^2_{Test}	MSE_{Test}
3	50	0.6	1.3	0.8655	0.0057	0.5917	0.0894
3	60	0.6	1.7	0.8595	0.0059	0.6265	0.0725
3	150	0.3	1.5	0.9219	0.0033	0.6863	0.0409
3	100	0.3	1.3	0.9240	0.0032	0.6965	0.0394
2	100	0.6	1.1	0.8746	0.0053	0.7181	0.0476
2	1000	0.1	0.9	0.8684	0.0056	0.7375	0.0471
2	1000	0.6	1.1	0.9022	0.0041	0.7936	0.0452
2	400	0.6	1.1	0.9018	0.0041	0.8023	0.0434
2	200	0.6	1.1	0.9010	0.0042	0.8095	0.0419
2	150	0.6	1.1	0.9003	0.0042	0.8202	0.0403

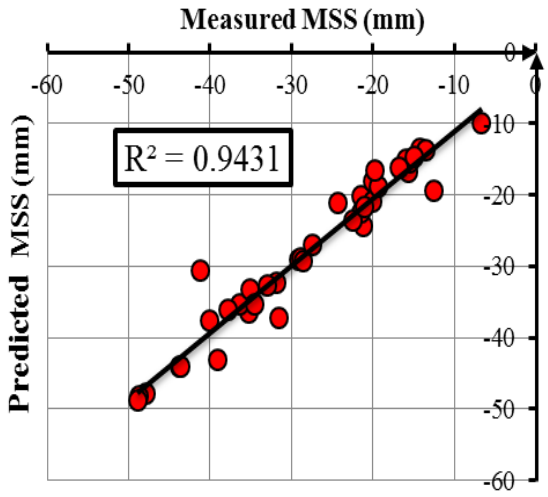
Table 7. A comparison between the results of ANFIS models

ANFIS model		MSE	R^2
ANFIS–SCM	Training	0.0024	0.9431
	Testing	0.0287	0.8806
ANFIS–FCM	Training	0.0042	0.9003
	Testing	0.0403	0.8202



(a)

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]



(b)

Figure 7. Correlation between measured and predicted values of MSS by ANFIS-SCM (a) training data, (b) testing data.

6. Estimation of maximum surface settlement (MSS) using ANN-BBO model

In this section, ANN was utilized to build a prediction model for estimation of MSS from available data, using MATLAB environment. Although ANNs are able to map input to output patterns directly and to use all influential parameters in model prediction, they still have some shortcomings such as a slow rate of learning and getting trapped in local minima [27], [28]. To overcome these problems, it is used BBO to better regulate the weights and biases of the ANN model. 80% (39 data points) of the datasets were assigned for training purposes, while 20% (10 data points) were used for testing the network performance.

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

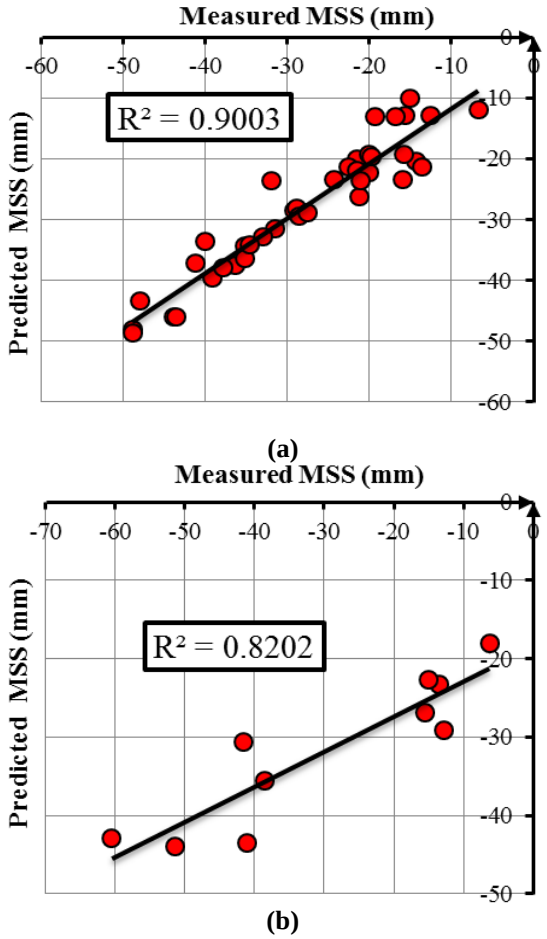


Figure 8. Correlation between measured and predicted values of MSS by ANFIS-FCM (a) training data, (b) testing data.

While all of these affect ANN performance, increased attention has been especially directed to finding the best architecture. This is justified not only by the fact that it is directly associated with the model performance but also because there is no theoretical background as to how this architecture will be found or what it should look like. The most typical method followed is a

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

repetitive trial-and-error process, during which, a large number of different architectures are examined and compared to one another (see Table 8). The optimal network for this study having one input layer with eight inputs, one hidden layer with nine neurons, and sigmoid hyperbolic tangent (tansig) activation function. The output layer has one neuron with a sigmoid hyperbolic logarithm (logsig) activation function. The architecture of the ANN-BBO model is shown in Figure 9. Also the optimal control parameters used for running BBO is demonstrated in Table 9. Correlations between measured and predicted values of MSS for training and testing phases are depicted in Figure10.

Table 8. Part of the sensitivity analysis of the ANN-BBO model

Model architecture	Activation functions	Maximum Number of Iterations	Population size	R^2_{Train}	MSE_{Train}	R^2_{Test}	MSE_{Test}
8-3-6-1	TanSig-TanSig-TanSig	50	30	0.4009	0.0302	0.4366	0.1159
8-6-3-1	TanSig-TanSig-LogSig	50	50	0.5874	0.0526	0.4785	0.0723
8-3-5-1	TanSig-LogSig-LogSig	50	200	0.6105	0.0428	0.5325	0.0832
8-4-5-1	LogSig-TanSig-LogSig	150	50	0.6426	0.0174	0.5961	0.0486
8-5-1	TanSig-LogSig	100	50	0.6541	0.0252	0.3414	0.0884
8-3-1	LogSig-TanSig	500	100	0.7568	0.0204	0.5150	0.0554
8-10-1	TanSig-TanSig	300	50	0.5329	0.0351	0.5502	0.0402
8-8-1	LogSig-LogSig	50	100	0.7970	0.0209	0.7979	0.0246
8-9-1	TanSig-LogSig	1000	50	0.8002	0.0149	0.8009	0.0231
8-9-1	TanSig-LogSig	700	50	0.8118	0.0105	0.8028	0.0268

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

Table 9. The optimal control parameters used for running BBO.

Definition	Value
Number of habitats (population size)	50
Highest number of repeat algorithm steps	700
Migration formula coefficient	0.9
Percentage of mutation	0.1
Percentage of old population that is directly transferred to the new population	0.2

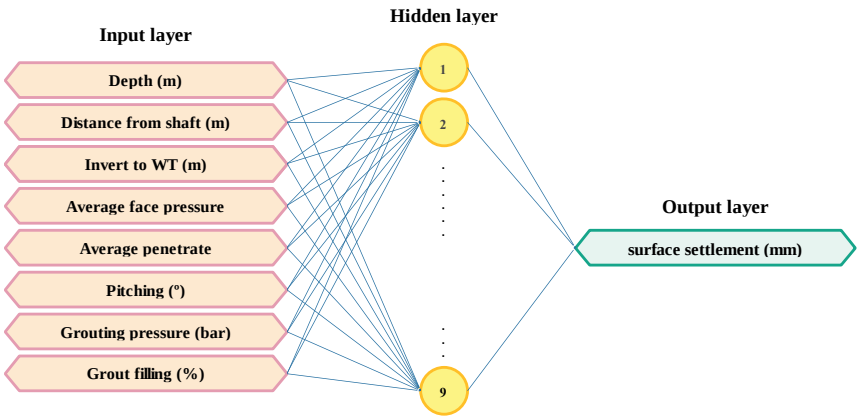


Figure 9. Architecture of ANN-BBO model.

To visually consider results in Tables 5, 6 and 8 in detail, performance of models compared with measured data in the training phase and testing phase are displayed in Figures. 11 and 12. These figures imply that complex relationships and structure among the data were highly captured by ANFIS-SCM model in this case of study. It is also concluded from Figures. 11 and 12, Tables 5 and 8 that the error of the training phase and testing phase are negligible which make this model as an accurate and valid model.

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

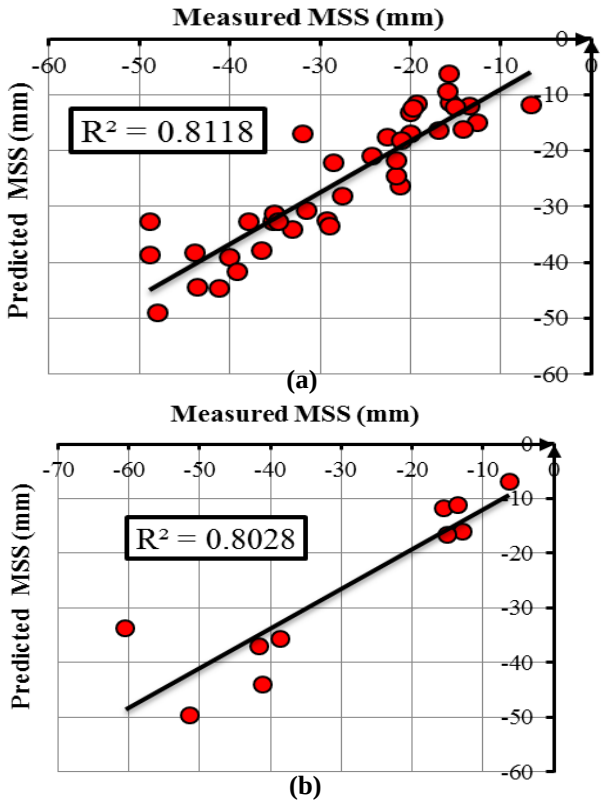
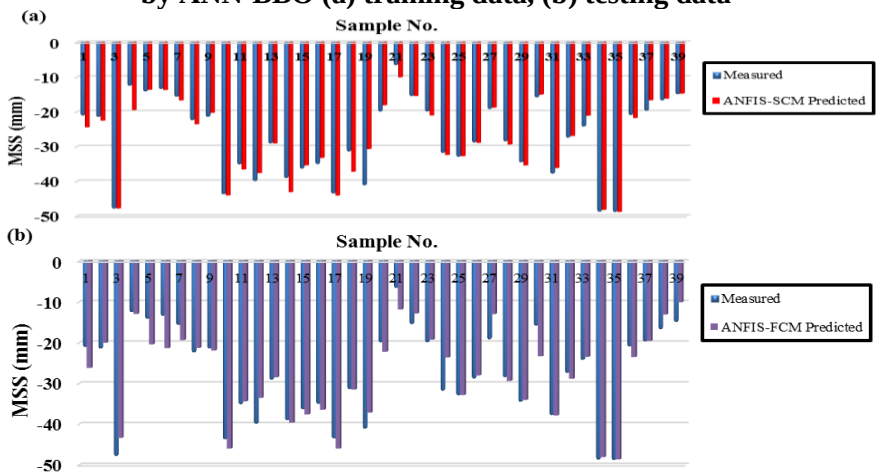


Figure 10. Correlation between measured and predicted values of MSS by ANN-BBO (a) training data, (b) testing data



[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

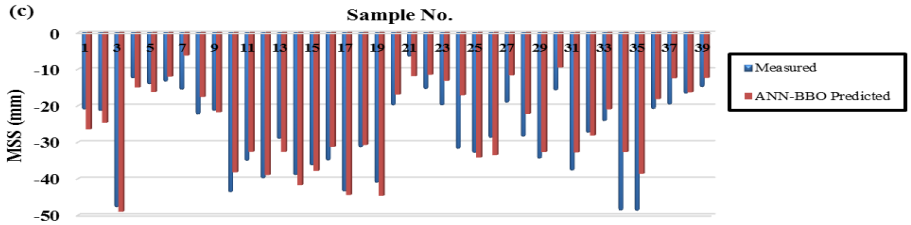


Figure 11. Bar chart for measured and predicted values of MSS in training phase: a ANFIS–SCM, b ANFIS–FCM, c ANN-BBO

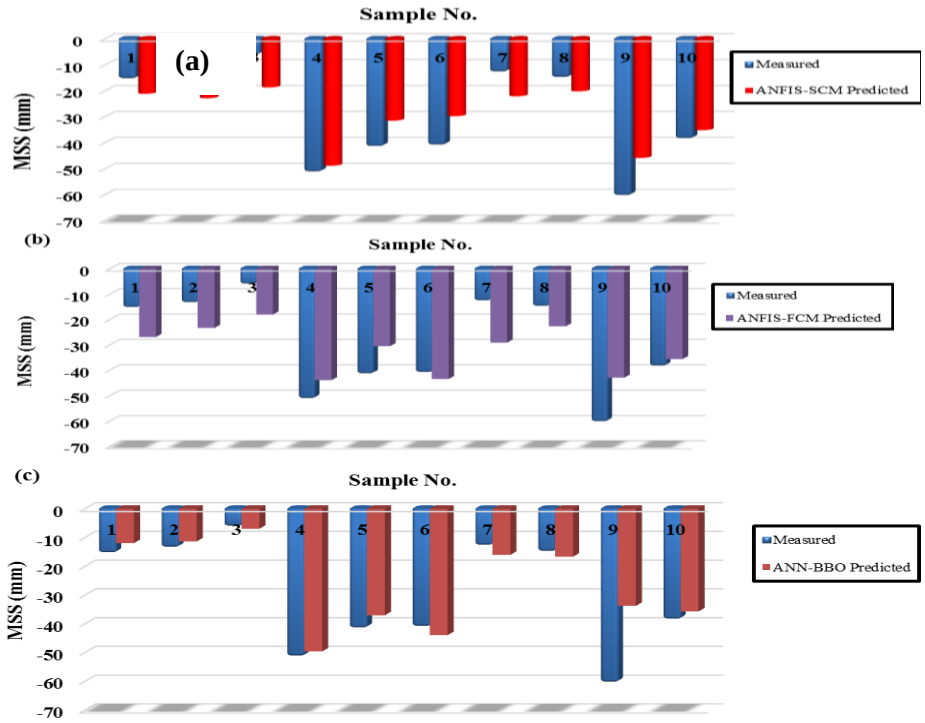


Figure 12. Bar chart for measured and predicted values of MSS in testing phase: a ANFIS–SCM, b ANFIS–FCM, c ANN-BBO

7. Conclusions

An attempt was made to study the performance of several AI methods for forecasting the MSS caused by EPB shield tunneling. The forecasting methods that have been investigated include the ANFIS–FCM, ANFIS-SCM and

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

ANN-BBO. The field data from the Bangkok Subway Project in Thailand were employed to develop various models investigated in this study. Two standard statistical performance evaluation measures are adopted to evaluate the performances of various models developed. The obtained results indicate that the AI methods are powerful tools to model the MSS caused by EPB shield tunneling. The results represent that the best performance can be obtained by ANFIS-SCM, in terms of different evaluation criteria during the training and testing phases. Also, ANN-BBO model is able to obtain the better forecasting accuracy in terms of different evaluation measures during the validation phase during both the training phase and the testing phase. The prediction by ANFIS-FCM model during the validation phase are inferior to the results during the training phase. Therefore, the results of the study are highly encouraging and to be suggested that ANFIS-SCM and ANN-BBO approaches are promising in modeling MSS caused by EPB shield tunneling, and this may provide valuable reference for researchers and engineers who apply AI methods for modeling MSS caused by EPB shield tunneling forecasting.

References

1. Huayong Y., Hu S., Guofang G., Guoliang H., "Electro-hydraulic proportional control of thrust system for shield tunneling machine. Automat Constr 18 (7) (2009) 950-956.
2. Peila D., Picchio A., Chieregato A., "Earth pressure balance tunnelling in rock masses: Laboratory feasibility study of the conditioning process", Tunn Undergr Sp Tech 35 (2013) 55-66.
3. Shao C., Lan D., "Optimal control of an earth pressure balance shield with tunnel face stability", Automat Constr 46 (2014) 22-29.

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

4. Hu S., Guofang G., Huayong Y., "Control model of earth pressure balance for shield tunneling", J China Coal Society 33 (3) (2008) 343-346.
5. Finno R. J., Clough G. W., "Evaluation of soil response to EPB shield tunneling" J Geotech Eng 111 (2) (1985) 155-173.
6. Clough G. W., Leca E., "EPB shield tunneling in mixed face conditions", J Geotech Eng 119 (10) (1993) 1640-1656.
7. Suwansawat S., Einstein H. H., "Artificial neural networks for predicting the maximum surface settlement caused by EPB shield tunneling", Tunn Undergr Sp Tech 21 (2) (2006) 133-150.
8. Boyacioglu M. A., Avci D., "An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: the case of the Istanbul stock exchange", Expert Syst Appl 37 (12) (2010) 7908-7912.
9. Avci E., "Comparison of wavelet families for texture classification by using wavelet packet entropy adaptive network based fuzzy inference system", Appl Soft Comput 8 (1) (2008) 225-231.
10. Avci E., Akpolat Z. H., "Speech recognition using a wavelet packet adaptive network based fuzzy inference system", Expert Syst Appl 31 (3) (2006) 495-503.
11. Avci E., Turkoglu I., Poyraz M., "Intelligent target recognition based on wavelet adaptive network based fuzzy inference system. In: Iberian Conference on Pattern Recognition and Image Analysis", Springer, (2005) 594-603.
12. Avci E., Hanbay D., Varol A., "An expert discrete wavelet adaptive network based fuzzy inference system for digital modulation recognition", Expert Syst Appl 33 (3) (2007) 582-589.
13. Jang J. S. R., "ANFIS: Adaptive-network-based fuzzy inference system", IEEE T Syst Man Cyb 23 (3) (1993) 665-685.
14. Chiu S. L., "Fuzzy model identification based on cluster estimation", J Intell Fuzzy Syst 2 (3) (1994) 267-278.

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

15. Jassar S., Liao Z., Zhao L., "Adaptive neuro-fuzzy based inferential sensor model for estimating the average air temperature in space heating systems", Build Environ 44 (8) (2009) 1609-1616.
16. Wei M., Bai B., Sung A. H., Liu Q., Wang J., Cather M. E., "Predicting injection profiles using ANFIS", Inform Sciences 177 (20) (2007) 4445-4461.
17. Bezdek J. C., "Fuzzy mathematics in pattern classification", Cornell university, Ithaca (1973).
18. Hornik K., Stinchcombe M., White H., "Multilayer feedforward networks are universal approximators", Neural Networks 2 (5) (1989) 359-366.
19. Ahmadi M. H., Aghaj S. S. G., Nazeri A., "Prediction of power in solar stirling heat engine by using neural network based on hybrid genetic algorithm and particle swarm optimization", Neural Comput Appl 22 (6) (2013) 1141-1150.
20. García-Pedrajas N., Hervás-Martínez C., Muñoz-Pérez J., "COVNET: a cooperative coevolutionary model for evolving artificial neural networks", IEEE T Neural Networ 14 (3) (2003) 575-596.
21. Hossein M., Ahmadi M. A., Ghare S. S., "Prediction of Power in Solar Stirling Heat Engine by Evolving Particle Swarm Optimization and Neural Network", Int J Comput Appl 34 (1) (2011).
22. Uzlu E., Kankal M., Akpınar A., Dede T., "Estimates of energy consumption in Turkey using neural networks with the teaching-learning-based optimization algorithm", Energy 75 (2014) 295-303.
23. Toğan V., "Design of planar steel frames using teaching–learning based optimization", Eng Struct 34 (2012) 225-232.
24. Togan V., "Design of pin jointed structures using teaching-learning based optimization", Structural Engineering and Mechanics 47 (2) (2013) 209-225.
25. Simon D., "Biogeography-based optimization", IEEE T Evolut Comput

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]

12 (6) (2008) 702-713.

26. Zhang Y., Phillips P., Wang S., Ji G., Yang J., Wu J., "Fruit classification by biogeography-based optimization and feedforward neural network", Expert Systems 33 (3) (2016) 239-253.
27. Hajihassani M., Armaghani D. J., Marto A., Mohamad E. T., "Ground vibration prediction in quarry blasting through an artificial neural network optimized by imperialist competitive algorithm", Bull Eng Geology Envir 74 (3) (2015) 873-886.
28. Tsou D., MacNish C., "Adaptive particle swarm optimisation for high-dimensional highly convex search spaces. In: Evolutionary Computation", CEC'03. The 2003 Congress on, IEEE, (2003) 783-789.

[Downloaded from isharat.khu.ac.ir on 2026-10-29]
[DOR: 20.1001.1.22286837.1397.12.5.3.5]
[DOI: 10.18869/acadpub.jeg.12.5.55]