Maximum surface settlement prediction in EPB TBM tunneling using soft computing techniques

H Samadil J Hassanpour¹ and E Farrokh²
¹ School of Geology, College of Science, University of Tehran, Tehran, Iran
² Department of Mining Engineering, Amirkabir University of Technology, Tehran, Iran

E-mail: hassanpour@ut.ac.ir

Abstract. Earth pressure balance (EPB) TBMs are commonly used for soft ground tunneling in urban areas. In metro tunnels’ excavation, designing a comprehensive monitoring system to control surface settlement is essential to prevent damage to surface structures. The present study aims to develop new prediction models to estimate the ground surface settlement using two soft computing techniques, SVM and ANN-MLP, and a multiple variable regression model to develop the new empirical formulas. The TBM operational parameters collected from the Tehran metro line 6, South extension (TML6-SE) project have been applied to confirm the provided models. In the data analysis process, the relationships between various parameters (torque index, thrust index, and earth pressure) and the ground surface settlement are investigated. Moreover, several statistical evaluation criteria are implemented to evaluate the performance of the developed models. The results show that the predicted values are in good agreement with the real data. The results can be used for similar ground and TBM tunneling conditions.

1. Introduction

One of the main issues in damage control for infrastructures adjacent to tunneling projects in densely populated urban areas is the assessment of ground settlement caused by EPBM along with the tunnel progress. The conservative estimation of the surface settlement is a logical task to reduce the cost and risk management of any tunneling project.

During the last decades, many researchers have proposed theoretical and empirical methods based on the geological conditions and tunnel operational parameters for surface settlement prediction [1], [2], [3], [4]. Due to the presence of complex geological features and different machines’ specifications, these developed analytical models are not applicable in all projects.

Three parameters of machine operating characteristics, geological conditions of the environment, and geometric characteristics of the tunnel can affect the surface settlement value, which directly impacts the chamber pressure [5]. The machine specifications include thrust force, cutterhead torque, and penetration rate. Also, the geological conditions consist of cohesion, internal friction angle, and elasticity modulus of materials. One note is that the geological characterization usually involves uncertainty.

The relationship between EPB machine parameters and ground settlement is a nonlinear issue that requires complex evaluation. Hence, nonlinear methods such as artificial intelligence-based soft computing techniques have been developed to estimate ground subsidence with the help of real data. A powerful methodology using several artificial intelligence techniques is considered, which can
implicitly detect the complicated nonlinear relationship between independent and dependent variables. Artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), and gene expression programming (GEP) have been implemented successfully to solve complex problems such as estimating maximum surface settlement (MSS) with the least error, rapid calculation, and high agreement with the previous investigation [6], [7]. Suwansawat and Einstein (2006) have performed the ANN method based on fifty collected data from the MRTA project [8]. Darabi et al. (2012) proposed a back-propagation technique to assess the tunnel-induced surface settlement in Tehran subway line 3 [9]. The principal component analysis (PCA) and ANFIS methods have been developed by Bouayad and Emeriault (2017) according to 95 collected datasets to evaluate the ground settlement [10]. Chen et al. (2019) have employed three neural network methods, including general regression neural network (GRNN), back-propagation (BP), and the radial basis function (RBF) based on a total number of two hundred datasets obtained from the Changsha metro line 4 [11]. Fattah et al. (2011) have investigated the shape of settlement caused by tunneling in cohesive ground by different approaches; including, analytical and empirical solutions, and numerical solution by the finite element method [12]. In the another study by the Fattah et al. (2012), the boundary element technique was performed as a practical problem-solving for lined tunnel problem in clayey soil [13]. The primary purpose of this study is to introduce new formulas using support vector machine (SVM), artificial neural network (ANN) multiple variable regression (MVR) techniques to estimate the maximum surface settlement value using 55 outputs obtained from Tehran metro Line 6 -Southern Extension (TML6-SE). The obtained results of the mentioned approaches are compared with statistical indices to choose the best method for estimating the maximum ground displacement.

2. Project overview
Tehran Metro Line 6, with a length of more than 38 km, is the longest in the metropolis of Tehran, which connects the northwest and southeast parts. Due to the significant increase in population, it was decided to add a 6.8 km section to the ending part. In this research, the area between chainage 2+150 and 3+000 is studied. As shown in Figure 1, this area is mainly composed of clayey silt and silty clayey sand with gravel.

In this range, the tunnel overburden varies between 11 and 20 m, with an average of 19 meters. The tunnel with a diameter of 9.19 meters is under construction by a refurbished EPB-TBM manufactured by HERRENKNECHT company.

Ground settlement monitoring system was performed using five points of vertical displacement gauge in the excavation area of Line 6 tunnels (Figure 2). In general, the maximum surface displacement that occurred was 14.3 mm and the average displacement above the tunnel axis was 7.34 mm.

3. Inputs and output
According to previous literature and regression analysis of available parameters in TML6-SE project, the factors leading to surface settlement can be divided into three major categories:
1) operational parameters of EPB-TBM (Tunnel Boring Machine);
2) tunnel geometry, and
3) earth pressure data.

It has been recognized that the EPB-TBM operational parameters, including torque index and thrust index, have the most impact on the maximum surface settlement. Dullmann (2014) showed that the torque index and thrust index can be obtained from equations 1 and 2. In these indices, the other performance parameters like penetration rate and cutting surface are taken into account as the geometric characteristics [14]. Tq and P are the cutterhead torque and penetration rate, respectively. MSP is the specific torque showing the torque per penetration rate. SCF is the specific contact cutting force in kN/m2.

\[
\text{Torque Index (TI)} = \frac{\text{MSP}}{A}, \quad \text{MSP} = \frac{Tq}{P} \times \text{MN.m/(mm/rev)/m}^2 \tag{1}
\]
\[
\text{Thrust Index (FI)} = \frac{P}{\text{SCF}}, \quad \text{(mm/rev)/(kN/m}^2) \tag{2}
\]
Although the tunnel’s overburden is the only geometric factor that needs to be investigated, there is no meaningful correlation between the maximum surface settlement and the overburden depth in this project. However, the earth support pressure in the chamber indicates a significant effect on maximum ground surface settlement. The face pressure monitoring is provided by the pressure cells installed in the front chamber.

The geotechnical properties of each geological engineering unit are shown in Table 1. As the study area in this paper is mostly composed of ET-4 engineering geological unit, the mechanical parameters of soil at the tunnel face are not used in the soft computing techniques.

The database collected in this project includes TBM performance parameters such as face pressure and machine operating parameters. Information of the geological parameters and EPB operational parameters for 533 concrete segmental rings are organized in a database. All these parameters have been collected during the construction phase and obtained from the daily operating records and the TBM data logger. The histograms and distribution curves of the earth support pressure and EPB operating...
parameters as input factors and the maximum surface settlement as a target value recorded in the database are shown in Figure 3.

### Table 1. TML6-SE average soil properties [15].

| Engineering geological units | Cohesion Ave. (kN/m²) | Internal friction angle Ave. (°) | Sat. Unit Weight (kN/m³) |
|-----------------------------|------------------------|---------------------------------|--------------------------|
| ET-2                        | 25                     | 25                              | 18                       |
| ET-3                        | 45                     | 23                              | 18                       |
| ET-4                        | 40                     | 15                              | 16.5                     |

![Distribution curve and frequency histogram of determined parameters in the database.]

4. Material and methods

Statistical methods such as multiple variable regression (MVR) and artificial intelligence models such as SVM and ANN have been applied for predicting the maximum surface settlement using the collected data. In this section, a brief introduction of the mentioned models is provided. Afterward, the models are developed using various computational algorithms as follows.
4.1. Development of SVM model

Support vector machine (SVM) is a powerful supervised algorithm that attempts to divide the datasets into two classes correctly to find the maximum marginal hyperplane (MMH) in multidimensional space [16]. This technique was performed in both regression and classification due to their ability to handle multiple continuous and categorical variables. To make powerful, flexible, and accurate results from SVM, the kernel function is implemented to convert non-separable problems into separable problems by adding more dimensions. In the present paper, the radial basis function (RBF) kernel has been performed, which maps input space in an infinite dimensional space. The following equation determines the mathematical formula of RBF:

\[ K(X, X') = \exp\left(-\gamma \|X - X'\|^2\right) \]

\[ \gamma = \frac{1}{\rho^2} \]  

\[ \|X - X'\|^2 \] may be recognized as the Euclidean distance square between two characteristic vectors. During the SVM model development to predict the best and most effective correlation with measured data, the unit weight (\(\gamma\)) is set by 0.1. The maximum margin hyperplane (MMH) and margins for an SVM trained with samples from two classes are shown in Figure 4. In particular, 80% of the data were randomly selected for training the generalized models, and the remaining datasets were applied for the verification phase.

4.2. Development of ANN model

Neural networks consist of input, output, and hidden layers. The input layers receive input data from an adjusted dataset to calculate the target parameter. The calculation of the output is based on the weighting of the input factors according to their importance degree in the learning model. The hidden layer is responsible for performing calculations with designated inputs and transforming modified results to output nodes with the transfer function [17].

Normalizing the prepared dataset within the range of 0 – 1 is the first step to perform the ANN. Determining the model architecture is necessary to minimize the statistical indices, then the optimum value of 3-4-1 is considered in this research. This particular arrangement means that the ANN has three layers in total, three neurons in the input level, similar to the number of input parameters, one hidden layer with four neurons, followed by one neuron in the output layer that eventually generates the MSS value. The ANN network and its structure are shown in Figure 5.

4.3. Multiple Regression Analysis

Multi-Variables Regression (MVR) is a statistical technique for interpreting the variance of the scalar response by using given explanatory variables and determines the relative contribution of each predictor in the total variance explained. The independent variables can be continuous or categorical. Generally, this method has been widely used for accurate prediction and can be employed for presenting a mathematical relationship between independent and dependent variables. The MVR approach can
specify the independent variables which have a significant impact on dependent variables. The multiple linear regression equation, defined as follow:

\[ Y = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_p \]  

(4)

Where \( Y \) is the predicted value of independent variables, \( X_1 \) to \( X_P \) are distinct and independent variables, \( b_0 \) is the intercept, and \( b_1 \) to \( b_P \) is the estimated regression coefficients.

Figure 5. The structure of ANN based on the determined database.

5. Implementation of methods

In this study, evolutionary algorithms of soft computing techniques are presented to develop more accurate and efficient models for predicting the maximum surface settlement (MSS). A new algorithm for these methods, ANN and SVM, is also generated to determine the performance analysis by considering the optimal neural network architecture. Each network is trained using 45 samples received from the TML6-SE. The remaining twelve samples are also randomly considered to validate the network’s performance. Moreover, an experimental formula was obtained through statistical analysis.

5.1. Developing new empirical equation

In this study, the multi-variables regression was applied to introduce an empirical equation that relates the surface settlement as a function of operational parameters and geological conditions. Empirical equations have great importance during the tunneling, which is obtained according to the actual data along the tunnel. For this purpose, machine specifications (applied thrust index ((mm/rev)/(kN/m²)) and torque index (MN.m/(mm/rev)/m²)) and earth pressure (bar) were assumed as independent variables and the monitored maximum settlement as a dependent variable. The regression analysis has been performed to evaluate the influence of each input variable on the maximum settlement by SPSS software version 22. An empirical equation for maximum settlement prediction defined as follows:

\[ S = -0.318 + 167.212 \times TI - 3.396 \times FI - 3.517 \times s \]  

(5)

5.2. Comparing the results of developed models

In this research, the statistical method and artificial intelligence models have been developed to predict the maximum surface settlement caused by EPBM in urban areas. The ANN progress based on defined architecture and the input functions is summarized in Table 2. Figure 6 illustrates the prediction results of the SVM and ANN-MLP models for both the training and testing stages. The diagrams indicate a similar trend between measured and prediction variables.

A comparison between the measured and predicted results using equation (5) is shown in Figure 7. The most reliable results are obtained by performing the multiple regression analysis and have good agreement with the recorded performance data. As indicated in the marginal histogram, the measured and calculated values of the ground settlement show the same regime. Comparison of predicted target values using proposed models and output values measured for 12 datasets from the testing phase is shown in Figure 8.
Table 2. ANN progress based on the defined algorithm.

| Epoch       | Time | Performance | Gradiant | mu       | Validation checks |
|-------------|------|-------------|----------|----------|------------------|
| 500 iteration | 11 s | 0.00345     | 1.85E-06 | 1.00E-11 | 0                |

Figure 6. Results of models; Comparison of measured and predicted values of settlement.

Figure 7. Results of models; a: Comparison of measured and predicted values, b: Correlation between measured and predicted values.

5.3. Performance analysis
Performance analysis aims to display the validity of the predicted techniques. To assess the deviation between the predicted and actual values of models, the statistical indices such as mean absolute deviation (MAD), mean squared error (MSE), root mean square error (RMSE), relative root mean square error (rRMSE) and mean absolute percentage error (MAPE) were used. The formulas of mentioned statistical indices are available in probabilistic publications.
The statistical indices of the predicted results are presented in Table 3 (a). The developed artificial intelligence methods have good performance for regression of the obtained results, proving the high efficiency and reliable prediction ability of the model. The proposed models are feasible and admissible for predicting the maximum surface settlement. Also, the statistical indices of the MVR models are presented in Table 3 (b).

**Table 3 (a).** Comparison between models based on statistical indices values in training and testing stages.

| Model | MAD | MSE  | RMSE | rRMSE | MAPE   |
|-------|-----|------|------|-------|--------|
| SVM   | 0.0457 | 0.0113 | 0.1063 | 0.0477 | 0.1791 |
| ANN   | 3.24E-11 | 7.29E-09 | 8.54E-05 | 0.0202 | 8.83E-06 |
| SVM   | 0.066 | 0.216 | 0.465 | 0.015 | 2.422 |
| ANN   | 0.003 | 0.480 | 0.693 | 0.141 | 0.278 |

**Table 3 (b).** Evaluation of the MVR model based on the statistical indices values.

| Model | MAD | MSE  | RMSE | rRMSE | MAPE |
|-------|-----|------|------|-------|------|
| MVR   | 0   | 0.079 | 0.282 | 0.042 | 0.506 |

5.4. Sensitivity analysis of the developed model

Sensitivity analysis is defined as a method to investigate the output uncertainty in a mathematical model related to the uncertainty in input variables. This process significantly focuses on quantification and propagation of the uncertainty. The sensitivity analysis results could be presented in a tornado or spider diagram, which shows the effect of change in each uncertain input on output variables.

In this study, the outputs recalculation under alternative assumptions could be beneficial to determine the impact of EPB operational parameters and earth pressure along the tunnel under sensitivity analysis to indicate the most effective variables on the surface settlement.

The surface settlement variation versus the variations in the input parameters is presented in Figure 9. As a result, the thrust index and torque index have the least effect on the ground settlement, while the earth pressure is the most effective parameter on the surface displacement.
Figure 9. Sensitivity analysis based on empirical equation; (a): Spider diagram of settlement value by changing the values of input parameters, (b): Tornado graph.

6. Conclusion

The predictability of maximum ground settlement in mechanized tunneling is one of the most important factors in risk management. In this study, the prediction of maximum surface settlement (MSS) has been made by three models using soft computing technology (SVM, ANN-MLP) and statistical methods. In order to investigate the predictive efficiency of the proposed models, a dataset of 55 data records gathered from Tehran Metro Line 6 Southern Extension Sector (TML6-SE) has been used. In the conducted analyses, torque index, thrust index, and chamber pressure are selected as the input parameters to predict MSS.

The empirical formulas based on the collected data from the extension sector have been developed for the MSS evaluation. Statistical indices such as MAPE, RMSE, MSE, MAD, and rRMSE were utilized to evaluate the accuracy of the empirical results, which were developed through a MVR method. The values of these parameters are in the normal ranges which indicates the model’s accuracy.

To assess the deviation between the model outputs and measured values obtained from the developed SVM and ANN models, the statistical indices such as MAPE, RMSE, MSE, MAD, and rRMSE have been utilized. The results of statistical indices illustrate an admissible and reliable accuracy of the models.

The SVM, ANN, and MVR methods are given to demonstrate the applicability of SVM in the MSS, due to its high ability in prediction with a small dataset. However, using a combination of several models to obtain reliable and effective estimation is recommended.

7. References

[1] Sun S, Pei H and Zhang S 2006 Analysis of face stability and ground settlement in EPB shield tunneling for the Nanjing metro Geo. Soci. Lon. IAEG Paper number 274

[2] Attewell P.B and Hurrell M.R 1985 Settlement development caused by tunnelling in soil Ground Eng 18 17–20

[3] Celestino T.B, Gomes R.A.M.P and Bortolucci A.A 2000 Errors in ground distortions due to settlement trough adjustment. Tun. Under. Spa. Tech. 15 97–100

[4] Fattah M.Y, Shlash K.T and Salim N. M 2013 Prediction of settlement trough induced by tunneling in cohesive ground. Acta. Geotec. 8 167–79

[5] Suwansawat S 2002 Earth pressure balance (EPB) shield tunneling in Bangkok: Ground Response and Prediction of Surface Settlements Using Artificial Neural Networks Ph.D Thesis Massachusetts Institute of Technology Cambridge MA

[6] Chiorboli M.A and Marcheselli P.P 1996 Analysis and control of subsidence due to earth pressure shield tunneling in Passante Ferroviario di Milano Proceedings of the 96th International Conference on North American Tunneling Balkema Rotterdam 97–106
[7] Ocak I and Seker S 2013 Calculation of surface settlements caused by EPBM tunneling using artificial neural network, SVM, and Gaussian processes Envir. Ear. Sci. Springer-Verlag Berlin Heidelberg

[8] Suwansawat S and Einstein H.H 2006 Artificial neural networks for predicting the maximum surface settlement caused by EPB shield tunneling. Tunn. Under. Spa. Tech. 21 133–50

[9] Darabi A, Ahangari K, Noorzad A and Arab A 2012 Subsidence estimation utilizing various approaches – A case study: Tehran No. 3 subway line Tunn. and Under. Spa. Tech. 31 117–27

[10] Bouayad D and Emeriault F 2017 Modeling the relationship between ground surface settlements induced by shield tunneling and the operational and geological parameters based on the hybrid PCA/ANFIS method Tunn. and Under. Spa. Tech. 68 142–52

[11] Chen R, Zhang P, Kang Z.H, Zhong Z.H & Liu Y 2019 Prediction of maximum surface settlement caused by earth pressure balance (EPB) shield tunneling with ANN methods Soi. and Found. 59 284–95

[12] Fattah M.Y, Shlash K.T and Salim N.M 2011 Settlement trough due to tunneling in cohesive ground. Indi. Geotec. Jour. 41 64-75

[13] Fattah M.Y, Shlash K.T and al-Soud M.S 2012 Boundary element analysis of a lined tunnel problem. Inter. Jour. Eng. (IJE) Trans. B: App. 25 87-94

[14] Dullmann J 2014 Performance optimization and wear forecast in hydro-shield tunneling in soft ground Ph.D Thesis Ruhr University Bochum

[15] Sabir International Company 2018 Geological and engineering geological report for Tehran metro line 6 southern extension sector project Report (In persian)

[16] Christopher J.C and Burges A 1998 Tutorial on support vector machines for pattern recognition Dat. Mini. Know. Disc. 121–67

[17] Hecht N.R 1987 Kolmogorov’s mapping neural network existence theorem Proceedings of the First International Conference on Neural Networks San Diego USA 11–4