An Optimized Neural Network for Predicting Settlements during Tunneling Excavation

GEORGE J. TSEKOURAS[1], JOHN KOUKOULIS[2], NIKOS E. MASTORAKIS[1]
[1]Department of Electrical Engineering and Computer Science, Hellenic Naval Academy, Terma Hatzikyriakou, Piraeus
[2]Dipl. Civil Engineer, Free Lancer, 8 Ploutarxou, Palaio Faliro, Athens
GREECE
tsekouras@snd.edu.gr, cv94068@hotmail.com, mastor@wseas.org, http://www.snd.edu.gr

Abstract: - Several empirical and analytical relations exist between different tunnel characteristics and surface and subsurface deformation, while numerical analyses (mainly using finite difference programs) have also been applied with satisfactory results. In the last years, the solution of some soil mechanics problems has been derived using the approach of the applied computational intelligence methods, especially the artificial neural networks (ANN). The objective of this paper is to describe an optimized artificial neural network (ANN) method in order to estimate the settlements of roof, face and walls during tunneling excavation. The ANN method uses as input variables the overload factor, the placement of the temporary lining ring behind the face, the thickness of the shotcrete, the modulus of elasticity for the surrounding rock-mass and the position where the measurement of the settlement takes place referring to the tunnel axis. The respective settlements are being calculated using different ANN's. For each ANN an optimization process is conducted regarding the values of crucial parameters such as the number of neurons, the time parameter and the initial value of the learning rate, etc. using the respective values of a pre-chosen evaluation set. The success of each ANN in predicting the respective settlements is measured by the correlation index between the experimental and predicted values for the evaluation set. Finally, the ANN with the closest to 1 correlation index is specified. A sensitivity analysis for different parameters (the input variables, the population of input vectors, etc.) is also presented showing the reliability of the proposed method.

Key-Words: - Artificial neural network, tunneling excavation, face settlement, roof settlement, walls settlement, parameters optimization, settlement prediction

1 Introduction

The computation of the settlements during the excavation of a tunnel is very important for the geotechnical tunneling design and has been extensively studied by a large number of researchers giving satisfactory solutions. The estimation of the ground and the underground deformation has been mainly approached by the following four ways [1]:

- *Stochastic and empirical methods*: The mathematical model for predicting the settlements of a "stochastic medium" was suggested by Litwinisyn [2]. Later, the surface settlement was determined empirically using the Gaussian probability curve by Peck [3] and O'Reilly and New [4]. The design parameters to be used in the last curve were compiled from previous field measurements in tunnel conditions with different soil conditions [3-5]. The stochastic theory for predicting longitudinal surface settlements was formed simultaneously [6]. In the last decade typical convergence curves for tunnel walls were given with respect to the distance from the tunnel face [7-8].
- *Analytical methods*: Kirsch gave an analytical solution for un-shield circular tunnel in big depths with specific modulus of elasticity and Poisson ratio for the surrounding rock-mass [7, 9]. This method was adopted the elastoplastic behavior using the Mohr-Coulomb criterion, giving as a result the convergence-confinement curves [7, 9-10]. Other researchers have developed methods based on closed form solutions, such as a two-dimensional analysis of ground deformations in an initially isotropic, homogenous incompressible medium giving as a result the correspondent strain field [11]. The previous method was extended for different values of Poisson's ratio [12-13], neglecting the distortion component [14] and studying surface troughs and lateral deformations [15].
- *Numerical methods*: The numerical methods have been increasingly applied to problems, involving the prediction of ground stresses and settlements during tunneling excavation. Many 2D and 3D finite elements analyses have been performed for soil and rock tunneling excavation

[16-19], providing the general form of the deformation field. But many “questions” have been appeared, such as the reliability of the method in conjunction with the accuracy of in-situ material properties.

- *Laboratory experiments*: Results of model tunnel tests in cohesive and non-cohesive materials were the base for the study of the mechanism of ground settlements [20-23] and for the validation of the numerical solutions [23].

Except these classical approaches, computational intelligence has been applied during the last decade. Kim et al were pioneers in the application of ANN methods for predicting the ground surface settlements due to tunneling, implementing an extended research for the selection of the proper input parameters [24]. Suwansawat and Einstein have used artificial neural networks successfully for predicting the maximum surface settlement caused by earth pressure balance shield tunneling [1, 25]. Similar problems have also been solved using ANN’s, such as estimation of ground surface settlement induced by deep excavation [26], settlement prediction of shallow foundations [27], estimation of consolidation settlement caused by groundwater drawdown [28], determination of preconsolidation pressure [29], estimation of the uplift capacity of suction caissons [30], determination of soil profiles [31], of soil models [32] and of soil parameters [33].

In this paper a new method for the estimation of the settlements of roof, face and walls during tunneling excavation is presented using an artificial neural network (ANN) based on back-propagation training algorithm. Its basic advantages are:

- the use of a small number of input variables, which are (a) the overload factor N_S, (b) the placement of the temporary lining ring behind the face L, (c) the thickness of the shotcrete t, (d) the modulus of elasticity for the surrounding rock-mass E_r and (e) the position x of the measurement point referring to the tunnel axis,
- the automatic optimization process of the ANN’s parameters based on the performance of the R^2 index of the evaluation set, such as the number of neurons, the initial values and the time parameters of momentum term and training rate, the kind and the parameters of activation functions,
- the sensitivity analysis for the population of the training set, for the set of the input variables, for the values set of each input variable etc.

The proposed method is applied for estimating the settlements of roof, face and walls during tunneling excavation for several types of uniform, isotropic rock-mass and various depths of tunnels. For the purposed of this study, the results of geotechnical finite difference analyses were used. These analyses were performed by Spyropoulos [34] via the usage of the finite difference program FLAC3D ver. 2.00. It is mentioned that a different, more complicated 3D model could be used, but the main intention of this paper is to analyze the optimized ANN method, to elaborate the respective problems of the optimization process and to mention the sensitivity of the methodology.

In section 2 the ANN method is presented in detail. In section 3, the method is analytically applied for the estimation of the settlements of the face. A sensitivity analysis for the population of the training set is also carried out. In sections 4 and 5 the settlements of the roof and of the walls are determined respectively. In section 6 a sensitivity analysis for the kind of the training process (stochastic or serial), for the random initialization of weights of ANN, for the set of the input variables and for the values of one characteristic input variable is carried out. Section 7 concludes the ANN performance, while in appendix the formation of the input data is described.

2 Proposed ANN Method for the Estimation of the Settlements during Tunneling Excavation

The estimation of the settlements during tunneling excavation is achieved by applying an ANN method through the optimization of the respective parameters of the back-propagation training algorithm. This method has the following flow chart, shown in Figure 1. The basic steps of the ANN optimization method are:

- *Data selection*: The results of the application of the finite difference method, using the code FLAC3D ver. 2.00 [34], have been used as input data for the proposed ANN methods. The input variables are:

(a) the *overload factor* N_S,

$$N_S = \frac{2p_o}{\sigma_{cm}} \quad (1)$$

where p_o is the respective isotropic geostatic pressure in the tunnel’s layer before tunneling excavation takes place and σ_{cm} is the uniaxial compressive strength of the surrounding rock-mass,

(b) the *placement of the temporary lining behind the face L*, as it is presented in Fig. 2 (in m),

(c) the *thickness of the shotcrete lining ring t* (in cm),

(d) the *modulus of elasticity for the surrounding*

rock-mass E_r (in MPa) and
(e) the *position* x (in m) where the measurement of the settlement takes place referring to the tunnel axis.

The output variables can be:
(a) the settlement of the roof s_z (in mm),
(b) the settlement of the face s_x (in mm) and
(c) the settlement of the walls s_y (in mm).

It is mentioned that a different ANN will be formed for each one of the three aforementioned settlements, because different ANNs for each output have better performance than one global ANN with all outputs [35].

● *Data preprocessing*: Generally, data are examined for normality, in order to modify or delete the values that are obviously wrong (*noise suppression*). In order to avoid saturation problems [36], the input and the output values are normalized as shown by the following expression:

$$\hat{x} = a + \frac{b-a}{x_{\max} - x_{\min}} \left(x - x_{\min} \right) \quad (2)$$

where $\hat{x}$ is the normalized value for variable x, $x_{\min}$ and $x_{\max}$ are the lower and the upper values of variable x, a and b are the respective values of the normalized variable.

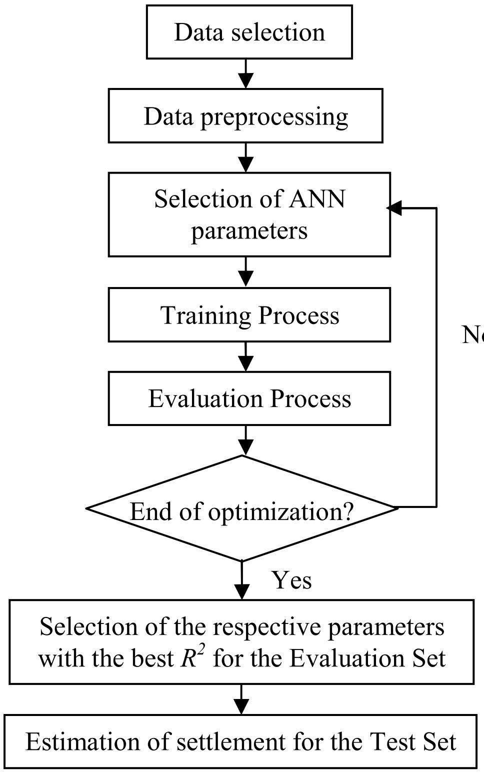


Fig. 1 Flowchart of the ANN optimization method.

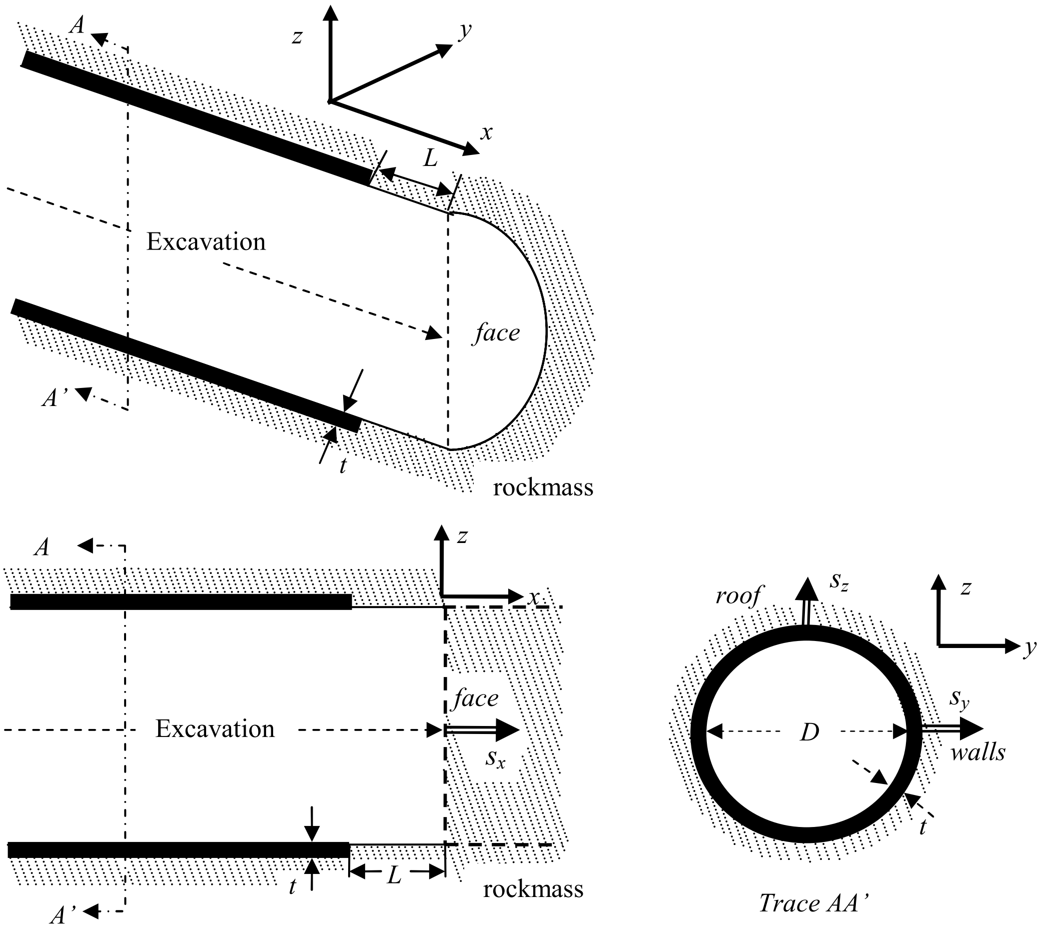


Fig. 2 Presentation of the co-ordinates system of a tunnel during its excavation.

● *Selection of ANN parameters*: The ANN parameters (like the number of neurons etc.) are not already known, but they can be specified empirically through trials. In this paper, the parameters are examined thoroughly. For example, all artificial neural networks with one hidden layer are formed with different number of neurons from 2 to 15. The one with the best R^2 index is selected. In this part of the study the number of neurons of the hidden layer, the initial values & the time parameters of the training rate and the momentum term and the activation functions with their parameters are defined.

● *Training Process*: A multilayer feed-forward neural network is adopted using the stochastic back-propagation algorithm with training rate and momentum term. The neurons in the network can be divided into three layers: input, hidden and output layer (see Fig. 3). According to Kolmogorov's theorem [36], an ANN can solve a problem using one hidden layer, if the last one has the proper number of neurons. In this study one hidden layer is used, but the number of neurons needs to be properly selected. This has forced to the examination of the various combinations of the critical ANN parameters. It is clarified that the number of neurons at the output layer is equal to the number of output variables, while the input nodes correspond to the input variables. The basic structure of an ANN is presented in Fig. 3a, while the latter's basic sub-steps are as follows:

(a) *Initialization*: Connection weights are equal to small random values between [-0.1,0.1] according to uniform distribution.

(b) *Training set's presentation*: During current epoch *ep* all patterns of the training set are presented randomly. For each vector (c) and (d) steps are realized. It is clarified that, in order to converge more rapidly than in the conventional method, both the training rate $\eta(ep)$ and the momentum term $a(ep)$ are changing their values at the beginning of each epoch *ep*:

$$\eta(ep)=\eta(ep-1)\cdot\exp(-1/T_\eta) \quad (3)$$

$$a(ep)=a(ep-1)\cdot\exp(-1/T_a) \quad (4)$$

where T_η, $\eta_0=\eta(0)$, T_a, $a_0=a(0)$ are respectively the time parameters and the initial values of both the training rate and the momentum term.

(c) *Forward pass calculations*: The *n*-th training pattern is defined as $\{\vec{x}_{in}(n),\vec{t}(n)\}$, where $\vec{x}_{in}(n)$ is the input vector -consisted of the normalized values of the input variables- with dimension q_{in} and $\vec{t}(n)$ the respective desired normalized output vector with dimension q_{out}. The activation signal of the *k*-neuron of the ℓ - layer is:

$$u_k^{(\ell)}(n)=\sum_{v=0}^{p_{\ell-1}} w_{kv}^{(\ell)}(n)y_v^{(\ell-1)}(n) \quad (5)$$

where $w_{kv}^{(\ell)}(n)$ is the weight between the ℓ - layer's *k*- neuron and the $(\ell-1)$ - layer's *v*- neuron, $p_{\ell-1}$ is the total number of neurons for the $(\ell-1)$ - layer and $y_v^{(\ell-1)}(n)$ is the output of the *v*-respective neuron (see Fig. 3b). For $v=0$, the threshold value is defined as $\theta_k=w_{k0}$, while $y_0^{(\ell-1)}=-1$. The activation function $f(x)$ can be linear $h_1\cdot x+h_2$, hyperbolic tangent $\tanh(h_1\cdot x+h_2)$, or hyperbolic sigmoid $\sinh(h_1\cdot x+h_2)$ for each layer (the parameters h_1 and h_2 should be defined just like the other ANN parameters).

The neuron's output is:

$$y_k^{(\ell)}(n)=f(u_k^{(\ell)}(n)) \quad (6)$$

In the input layer one demands:

$$y_v^{(0)}(n)=x_v(n)\,,\ \forall v \quad (7)$$

where $x_v(n)$ is the *v*-th element of input vector $\vec{x}_{in}(n)$.

In the output layer $L^{/}$ one determines:

$$y_k^{(L^{/})}(n)=o_k(n)\,,\ \forall k \quad (8)$$

where $o_k(n)$ is the *k*-th element of the output vector $\vec{o}(n)$, estimated by the ANN. The error of the output *k*-neuron is:

$$e_k(n)=t_k(n)-o_k(n) \quad (9)$$

where $t_k(n)$ is the *k*-th element of the desired normalized output vector $\vec{t}(n)$.

(d) *Reverse pass calculations*: The weight is calculated by the delta-rule:

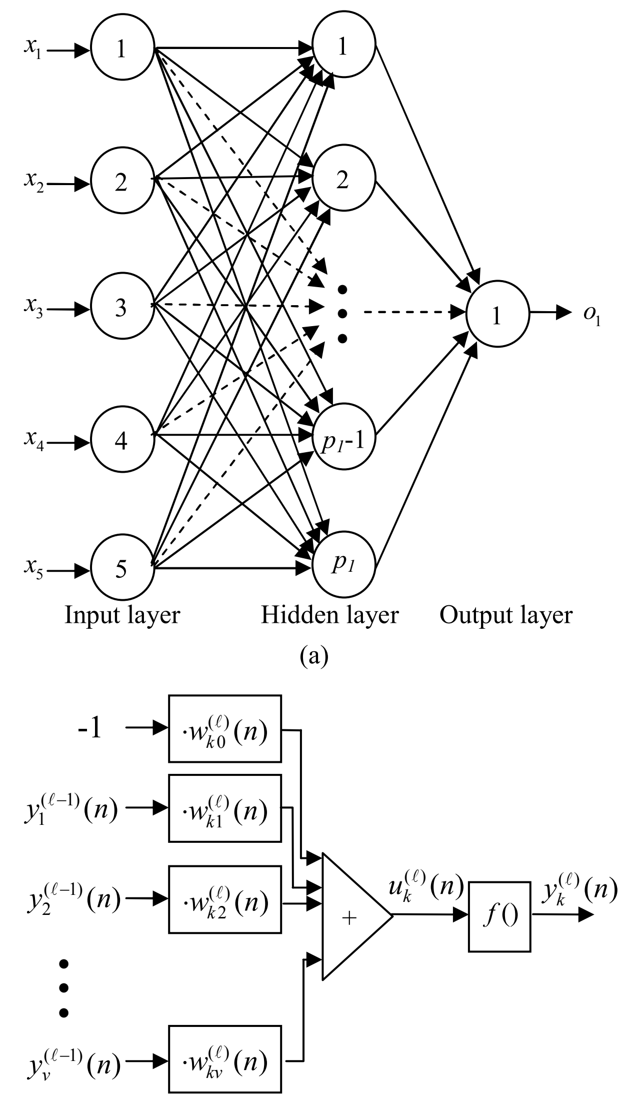
$$w_{kv}^{(\ell)}(n+1)=\begin{cases} w_{kv}^{(\ell)}(n)+ \\ \eta(ep)\cdot\delta_k^{(\ell)}(n)\cdot y_v^{(\ell-1)}(n)+ \\ \alpha(ep)\cdot[w_{kv}^{(\ell)}(n)-w_{kv}^{(\ell-1)}(n-1)] \end{cases} \quad (10)$$

where $\delta_k^{(\ell)}(n)$ is the local descent of the *k*-neuron determined for the output layer and for the hidden one respectively as:

$$\delta_k^{(L)}(n)=e_k^{(L)}(n)\cdot f^{/}\left(u_k^{(L)}(n)\right) \quad (11)$$

$$\delta_k^{(\ell)}(n)=f^{/}\left(u_k^{(\ell)}(n)\right)\cdot\sum_i \delta_i^{(\ell+1)}(n)w_{ik}^{(\ell+1)}(n) \quad (12)$$

(e) *Stopping criteria*: The steps (b) to (d) are repeated continuously until the weights to be stabilized or the respective error function not to be improved or the maximum number of epochs to be exceeded.



(a)

(b)

Fig. 3 (a) Typical structure of an ANN (5-p_1-1) where there are 5 neurons in input layer, p_1 neurons in hidden layer and 1 neuron in output layer, (b) Enlargement of the ℓ - layer's k- neuron during the presentation of the n-th training pattern.

Analytically the weights criterion is defined with the following expression:

$$\left|w_{kv}^{(\ell)}(ep) - w_{kv}^{(\ell)}(ep-1)\right| < \ell imit_1, \forall k, v, \ell \qquad (13)$$

where $\ell imit_1$ has a proper value and ep is the current epoch of training algorithm.

The error function is the root mean square error $RMSE_{tr}$ for the training set (where the respective population equals to m_1) according to:

$$RMSE_{tr} = \sqrt{\frac{1}{m_1 \cdot q_{out}} \sum_{m=1}^{m_1} \sum_{k=1}^{q_{out}} e_k^2(m)} \qquad (14)$$

and the respective criterion is:

$$\left|RMSE_{tr}(ep) - RMSE_{tr}(ep-1)\right| < \ell imit_2 \qquad (15)$$

where $\ell imit_2$ is the respective limitation value. Practically, it is an early stopping criterion.

The maximum number of epochs' criterion is:

$$ep \geq \max\_epochs \qquad (16)$$

If one of the above criteria comes true, the main core of back propagation algorithm stops. Otherwise the number of epochs is increased by one, the whole process returns to step (b) and the training rate and the momentum term are re-calculated by eq. (3) and (4). The criterions' purposes are: (i) to avoid the over-fitting problem and (ii) to enable the convergence of the algorithm.

● *Evaluation Process*: After the convergence of the training algorithm, the R^2 correlation index between the desired and the estimated values of the under study settlement s for the evaluation set is calculated. It is noted that:

$$R^2 = r_{s-\hat{s}}^2 = \frac{\left(\sum_{i=1}^{n}\left(\left(s_i - \overline{s}_{real}\right)\cdot\left(\hat{s}_i - \overline{s}_{est}\right)\right)\right)^2}{\sum_{i=1}^{n}\left(s_i - \overline{s}_{real}\right)^2 \cdot \sum_{i=1}^{n}\left(\hat{s}_i - \overline{s}_{est}\right)^2} \qquad (17)$$

where s_i is the desired value of the settlement, $\overline{s}_{real}$ the mean desired value of the respective data set (training, evaluation or test), $\hat{s}_i$ the estimated value, $\overline{s}_{est}$ the mean estimated value of the data set, n the population of the respective data set. The mean absolute error MAE is also calculated:

$$MAE = \sum_{i=1}^{n}\left|s_i - \hat{s}_i\right| \Big/ n \qquad (18)$$

It is mentioned that the desired settlements are the respective ones from the FLAC3D program.

● *End of Optimization Loop*: In this section the end of the optimization loop is checked. If all possible combinations of the under study ANN parameters are examined, then the next step will follow, otherwise a new selection of the ANN parameters will take place and the training process will be continued.

● *Selection of the ANN parameters with the best R^2 index for the Evaluation Set*: From all examined combinations the one with the biggest R^2 index for the evaluation set is chosen as the best one with the respective ANN parameters and the finally estimated weights.

● *Estimation of Settlement for the Test Set*: The parameters and weights of the previous step consist the final proposed Artificial Neural Network, which can be used for the estimation of the settlement for the unknown test set.

3 Application of the Proposed ANN Method for the Face's Settlement during Tunnel Excavation

Following, the proposed method presented in section 2 is applied for the estimation of the face settlement during the excavation of a tunnel. The input vector $\vec{x}_{in}(n)$ is formed with the normalized input variables of the overload factor $\hat{N}_S$, the placement of the temporary lining ring behind the face $\hat{L}$, the thickness of the shotcrete $\hat{t}$, the modulus of elasticity for the surrounding rock-mass $\hat{E}$ and the position $\hat{x}$ where the measurement of the settlement takes place referring to the tunnel axis:

$$\vec{x}_{in} = \left[\hat{N}_S \quad \hat{L} \quad \hat{t} \quad \hat{E} \quad \hat{x} \right]^T \tag{19}$$

The output vector $\vec{t}(n)$ is formed by the normalized output variable $\hat{s}_x$ of the face settlement:

$$\vec{t} = \left[\hat{s}_x \right] \tag{20}$$

The available vectors are 2475 (see appendix) and they are separated to two sets randomly: the *training* set with the percentage p% of the vectors population and the *evaluation* set with the rest ones. The test set is not defined, because there are not any experimental values.

There are several crucial ANN parameters to be selected, such as:

- the number of neurons of the hidden layer, which ranges from 2 to 15 with incremental step 1,
- the initial value $a_0 = a(0)$ and the time parameter T_a of the momentum term, which get values from the sets {0.1,0.2,…,0.9} and {500, 1000,…,3000} respectively,
- the initial value $\eta_0 = \eta(0)$ and the time parameter T_η of the training rate, which get values from the sets {0.1, 0.2, …, 0.9} and {500, 1000, …, 3000} respectively,
- the type and the parameters of the activation functions of the hidden and the output layers, where the type can be *hyperbolic tangent*, *linear or hyperbolic sigmoid*, while the h_1 & h_2 parameters get values from the set {0.1,0.2,…, 1.5}.

The parameters of the stopping criteria are defined after a few trials as max_epochs=7000, $\ell imit_1=10^{-4}$, $\ell imit_2=10^{-4}$.

The application of the abovementioned method in Visual Fortran 6.0 gives the capability to realize all possible combinations of the values of the crucial parameters. In this study the respective combinations run into 88,668,600 for each case of training – evaluation sets, which practically can not be examined. This forced the authors to apply the proposed optimization process gradually through consecutive steps in order to determine the values of the ANN's parameters.

As a first step, the number of neurons varies from 2 to 15, while the remaining parameters are assigned with fixed values ($a_0 = 0.6$, $T_a = 1000$, $\eta_0 = 0.6$, $T_\eta = 1000$, activation functions in both layers: hyperbolic tangent, h_1=0.7, h_2=0.0). Simultaneously, the method is executed for different cases of training-evaluation sets, which means that different percentage p from the available vectors is used for the formation of the training set. The percentage p takes values from the set {10%, 20%, …, 90%} and nine scenarios are implemented for each set of neurons (2 to 15), where each scenario is a different case of the neurons' optimization process for percentage p. The best results of the R^2 index of the evaluation set are generally given for p=80%, as it is shown in Fig.4. The last percentage is used for the following steps as the best one. In Fig. 5 the R^2 index for the training and the evaluation set with p=80% are presented, where the R^2 index of the evaluation set keeps step with the respective one of the training set. With the neurons numbered from 5 to 8 and from 10 to 11 the R^2 index for the evaluation set has big values (the biggest is for 11), while for bigger values it rapidly decreases.

As a second step the initial value a_0 and the time parameter T_a of the momentum term change simultaneously in the respective regions, while the neurons are 11 and the other parameters are constant. In Fig. 6 it is clear that the results of the R^2 index for the evaluation set are satisfactory for $a_0 \geq 0.6$ and $T_a \geq 1000$. The respective improvement of the R^2 index from the proper calibration of the parameters a_0, T_a is significant from 0.79 to 0.96. The best result is given for $a_0 = 0.9$, $T_a = 3000$. It is mentioned that R^2 decreases dramatically for $a_0 \leq 0.5$. As a third step the initial value η_0 and the time parameter T_η of the training rate change simultaneously in the respective regions, while the other parameters remain constant (neurons=11, $a_0 = 0.9$, $T_a = 3000$, activation functions in hidden & output layers: hyperbolic tangent, h_1=0.7, h_2=0.0). In Fig. 7 it is clear that the results of the R^2 index for the evaluation set are satisfactory for $\eta_0 \geq 0.5$ and $1000 \leq T_\eta \leq 2500$. The best result is given for $\eta_0 = 0.5$, $T_\eta = 2500$. It is mentioned that R^2 decreases dramatically for $\eta_0 \leq 0.3$ and $T_\eta \leq 1000$.

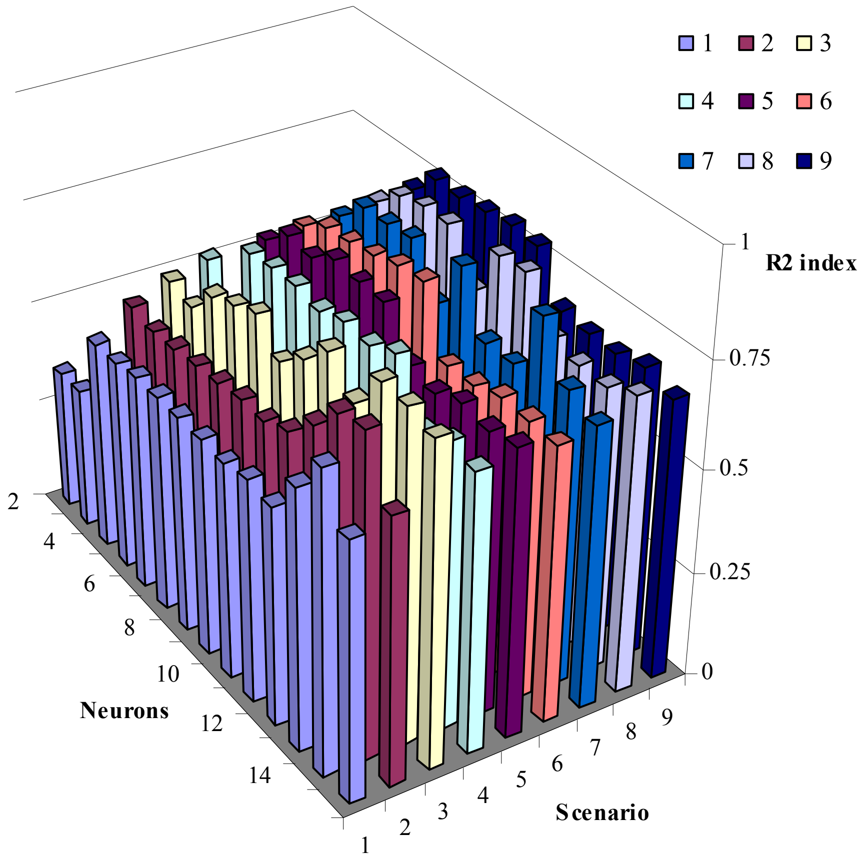


Fig. 4 R^2 index for the evaluation set for different percentage p of training set (neurons: 2 to 15, $a_0 = 0.6$, $T_a = 1000$, $\eta_0 = 0.6$, $T_\eta = 1000$, activation functions in both layers: hyperbolic tangent, h_1=0.7, h_2=0.0).

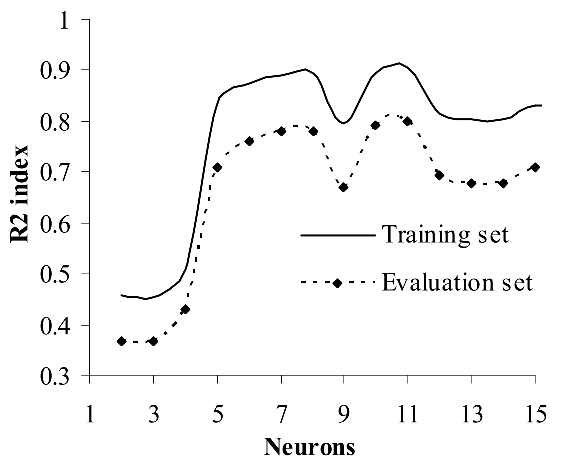


Fig. 5 R^2 index for the evaluation & the training set (p=80%), neurons:2-15, $a_0 = 0.6$, $T_a = 1000$, $\eta_0 = 0.6$, $T_\eta = 1000$, activation functions in both layers: hyperbolic tangent, h_1=0.7, h_2=0.0.

Similarly it is found that the ANN gives better results using as an activation function hyperbolic tangent in both layers with parameters $h_1 = 0.6 - 0.8$ and $h_2 = 0$.

The final calibration of the ANN model is realized for 10 to 11 neurons, $a_0 = 0.8 - 0.9$, $T_a = 1500 - 3000$, $\eta_0 = 0.5 - 0.6$, $T_\eta = 1000 - 2500$, activation functions in both layers: hyperbolic tangent with parameters h_1=0.6-0.8, h_2=0. In this way a local minima is avoided. The best result for the R^2 index of the evaluation set is 98.57% and is given for an ANN with 11 neurons in the hidden layer, $a_0 = 0.9$, $T_a = 3000$, $\eta_0 = 0.5$, $T_\eta = 2500$, h_1=0.7 and h_2=0 using hyperbolic tangent as the activation function in both layers.

In Fig. 8 the settlements of the face from the execution of the FLAC3D program and the respective estimated settlements for the training and evaluation sets are presented.

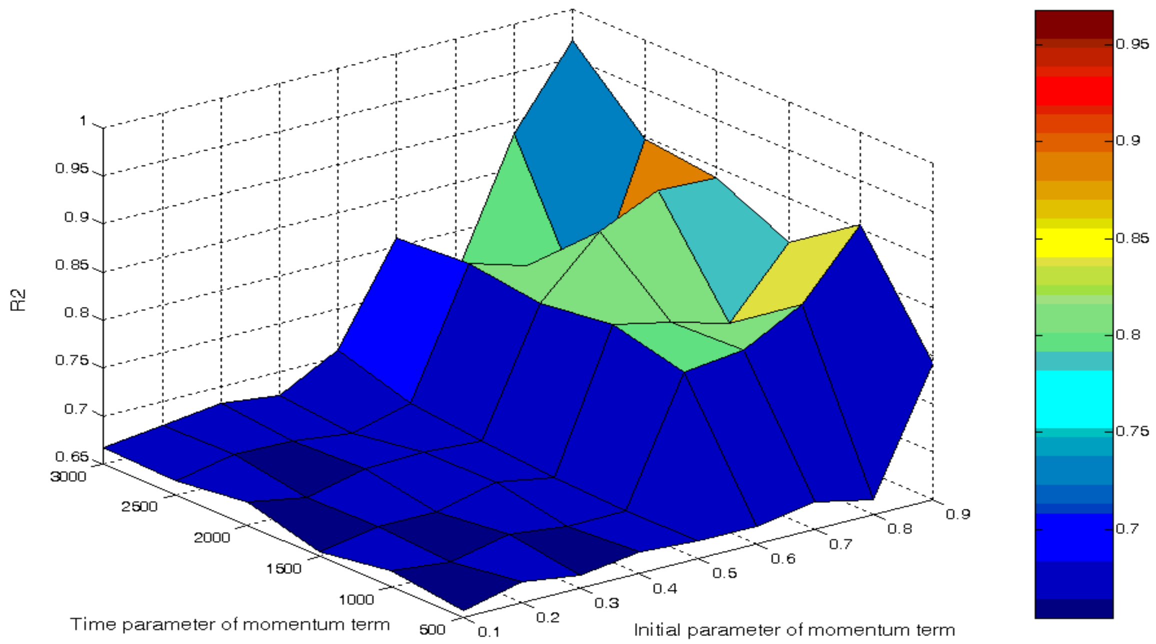


Fig. 6 R^2 index for the evaluation set for 11 neurons, $a_0 = \{0.1, 0.2, \ldots, 0.9\}$, $T_a = \{500, 1000, \ldots, 3000\}$, $\eta_0 = 0.6$, $T_\eta = 1000$, activation functions in both layers: hyperbolic tangent, h_1=0.7, h_2=0.0.

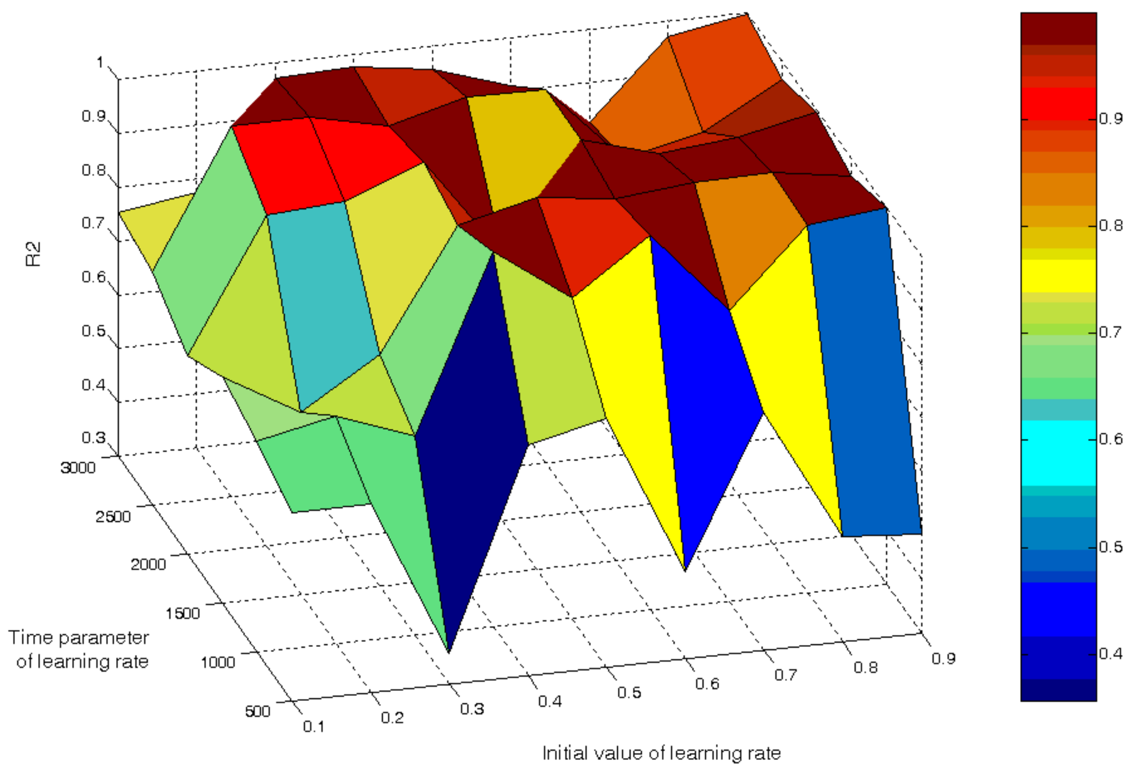
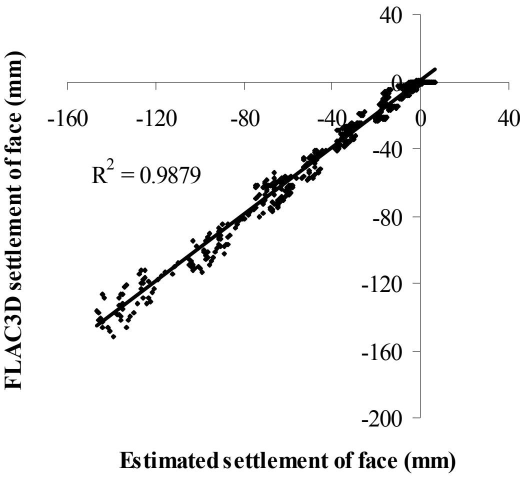
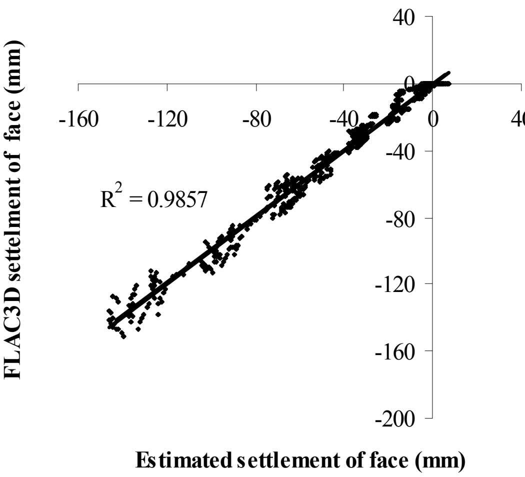


Fig. 7 R^2 index for the evaluation set for 11 neurons, $a_0 = 0.9$, $T_a = 3000$, $\eta_0 = \{0.1, 0.2, \ldots, 0.9\}$, $T_\eta = \{500, 1000, \ldots, 3000\}$, activation functions in both layers: hyperbolic tangent, h_1=0.7, h_2=0.0.



(a) Training set



(b) Evaluation set

Fig. 8 The settlements of the face from FLAC3D program and the estimated settlements of the face during tunneling excavation for (a) the training and (b) the evaluation set.

After the proposed optimization process, the basic results are the following:

- the respective mean value of the absolute error *MAE* is 2.25 mm for the training set and 2.21 mm for the evaluation set,
- the respective root mean square error *RMSE* is 3.23 mm for the training set and 3.09 mm for the evaluation set,
- the R^2 correlation index between the estimated and the settlements from FLAC3D program is 98.79% for the training set and 98.57% for the evaluation set.

It is observed that RMSE, MAE and R^2 are slightly better for the evaluation set than ones for the training set. It is mentioned that this is not a paradox, because the training and evaluation sets have not any common members. So, if RMSE and MAE are slightly smaller for the evaluation set than ones for the training set, the respective R^2 for the evaluation set will not be larger necessarily than the respective one for training set.

4 Application of the Proposed ANN Method for the Roof's Settlement during Tunnel Excavation

The proposed method is applied for the estimation of the settlement of the roof during the excavation of a tunnel. The input vector $\vec{x}_{in}(n)$ is given by eq. (19), while the output vector $\vec{t}(n)$ is formed by the normalized output variable $\hat{s}_z$ of the settlement of the roof:

$$\vec{t} = \left[\hat{s}_z\right] \tag{21}$$

The available vectors are 7650 (see appendix) and they are separated into two sets randomly: the *training* set with the *p*% of the vectors population and the *evaluation* set with the rest ones. The parameter *p* and the respective crucial ANN parameters range in the same regions, as it is presented in the case of the face's settlement (see § 3). After the gradual application of the proposed optimization process the best R^2 index of the evaluation set is given by an ANN with 11 neurons in the hidden layer, $a_0 = 0.7$, $T_a = 2000$, $\eta_0 = 0.7$, $T_\eta = 2500$, $h_1 = 0.7$ and $h_2 = 0$ using hyperbolic tangent as the activation function in hidden and output layers. This ANN has been trained using 20% of the available input vectors (*p*=20%). It is noted that the available vectors for the roof's settlement is quite different from the one of the face's settlement and the respective variation of R^2 index for the evaluation set for different percentage *p* is narrower than the respective one of Fig. 4.

In Fig. 9 the settlements of the roof from the execution of the FLAC3D program and the respective estimated settlements for the evaluation set are presented.

The basic results are the following, which are quite satisfactory:

- the respective mean value of the absolute error *MAE* is 4.52 mm for the training set and 4.73 mm for the evaluation set,
- the respective root mean square error *RMSE* is 6.41 mm for the training set and 6.74 mm for the evaluation set,
- the R^2 correlation index between the estimated and the settlements from FLAC3D program is 96.62% for the training set and 96.08% for the evaluation set.

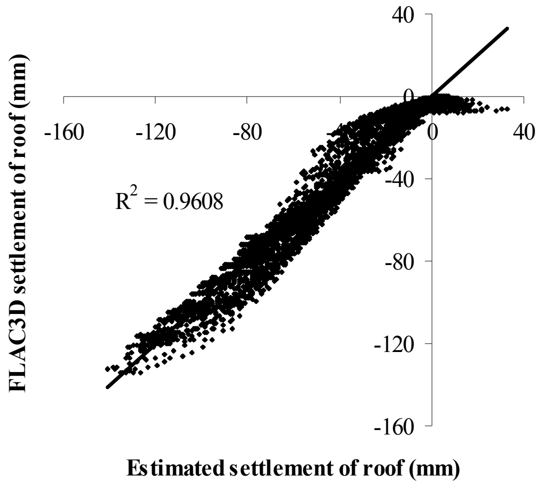


Fig. 9 The settlements from FLAC3D program and the estimated settlements of the roof during tunneling excavation for the evaluation set.

The results are slightly worse than the respective ones from face settlement, which indicates a less accurate ANN. There is probably a physical mechanism that is not accounted for by the input variables (i.e. the occurrence of plasticity close to the roof and walls).

5 Application of the Proposed ANN Method for the Side Walls' Settlement during Tunnel Excavation

The proposed method is applied for the estimation of the settlement of the side walls during the excavation of a tunnel. The input vector $\vec{x}_{in}(n)$ is given by eq. (19), while the output vector $\vec{t}(n)$ is formed by the normalized output variable $\hat{s}_y$ of the settlement of the walls:

$$\vec{t} = \left[\hat{s}_y\right] \tag{21}$$

The available vectors are 7650 (see appendix) and they are separated into the training set and the evaluation set randomly, while the parameter ***p*** and the respective crucial ANN parameters range in the same regions, as it is presented in § 3, 4. After the gradual application of the proposed optimization process the best R^2 index of the evaluation set is given by an ANN with 11 neurons in the hidden layer, $a_0 = 0.5$, $T_a = 2500$, $\eta_0 = 0.6$, $T_\eta = 1500$, $h_1 = 0.7$ and $h_2 = 0$ using hyperbolic tangent as the activation function in hidden and output layers. This ANN has been trained using 80% of the available input vectors (p=80%).

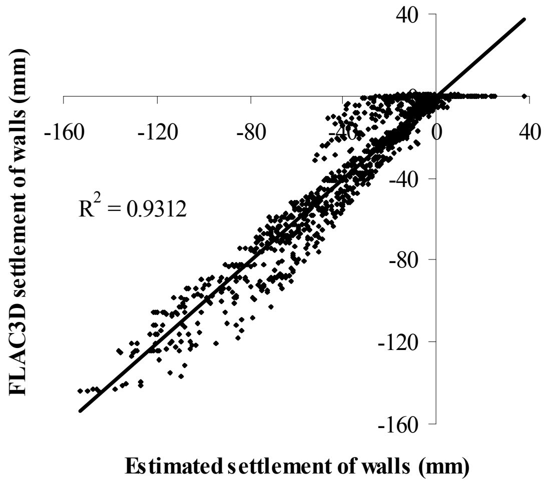


Fig. 10 The settlements from FLAC3D program and the estimated settlements of the walls during tunneling excavation for the evaluation set.

In Fig. 10 the settlements of the side walls from the execution of the FLAC3D program and the respective estimated settlements for the evaluation set are presented.

The basic results are the following, which are quite satisfactory:

- the respective mean value of the absolute error *MAE* is 5.83 mm for the training set and 5.98 mm for the evaluation set,
- the respective root mean square error *RMSE* is 8.67 mm for the training set and 8.98 mm for the evaluation set,
- the R^2 correlation index between the estimated and the settlements from FLAC3D program is 93.49% for the training set and 93.12% for the evaluation set.

The results are slightly worse than the respective ones from the previous two cases. The reason should probably be a physical mechanism that is not accounted for by the input variables.

6 Sensitivity Analysis for the Proposed ANN Method

In this section a sensitivity analysis is carried out. The main cases, which are examined in this sensitivity analysis, concern the following:

- the random selection of the input vectors per epoch and the random initialization of the weights,
- the difference between the random selection of the input vectors per epoch (stochastic training) and the serial presentation of the input vectors per epoch (serial training),

- the removal of vectors with specific values of one input variable from the training set, which is synopsized as the range of the values set of the input variables,
- the omission of one variable from the set of the input variables.

6.1 Random Selection of the Input Vectors per Epoch & Random Initialization of the Weights

During the optimization process the weights were initialized randomly, but for every different case of parameters the initial weights were the same. Here, the proposed ANN method is applied 10 times consecutively for the estimation of the side wall's settlement using the parameters of the best combination (see § 5) with different initialization of weights. During the training process the input vectors are selected randomly, as it was already happened. The respective results for R^2 indexes for the two sets are presented in Table 1. It is noticed that the values of the R^2 indexes have small variations showing that the behavior of the proposed ANN method is stable.

TABLE 1
10 EXECUTIONS OF THE PROPOSED ANN FOR THE SIDE WALL'S SETTLEMENT

Trial	Number of epochs	R^2 for training set	R^2 for evaluation set
1	5285	92.50%	92.20%
2	5486	93.50%	92.80%
3	5273	94.80%	94.30%
4	5408	93.90%	93.60%
5	5505	93.30%	92.80%
6	5153	93.70%	93.40%
7	5398	93.30%	93.00%
8	5447	92.80%	92.60%
9	5458	93.50%	93.30%
10	5291	92.60%	92.40%

6.2 Difference between Stochastic and Serial Training

According to the proposed ANN method the vectors of the training set are selected randomly (stochastic training). In this section, the ANN is modified, so that the presentation order of the vectors is the same per epoch (serial training), which is easier to be programmed than the stochastic training. In Fig. 11 the mean absolute error *MAE* of the side wall's settlement using ANN with the parameters of the best combination (see § 5) is presented for the serial training and for the stochastic one (1st and 3rd trial of Table 1). The convergence of the *MAE* for the serial training is smoother and slower than the respective one of the stochastic training. The results of ANN with serial training are the following:

- the respective mean value of the absolute error *MAE* is 6.96 mm for the training set and 7.25 mm for the evaluation set,
- the respective root mean square error *RMSE* is 9.99 mm for the training set and 10.35 mm for the evaluation set,
- the R^2 correlation index between the estimated and the settlements from FLAC3D program is 81.6% for the training set and 82.1% for the evaluation set.

The comparison between the two methods of the training shows the superiority of the stochastic training with which the danger of the over-fitting is suppressed [35].

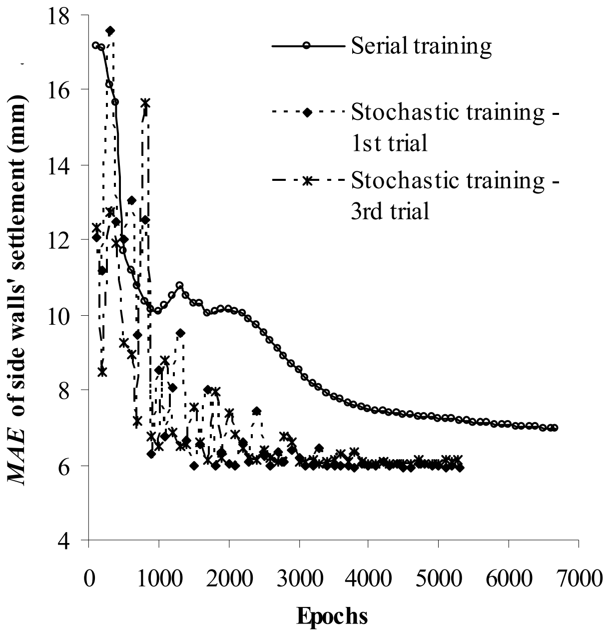


Fig. 11 *MAE* of the side wall's settlement using ANN with the parameters of the best combination for the serial training and for the stochastic training (1st and 3rd trial of Table 1).

6.3 Range of Values' set of Input Variables

The values set of each input variable of the training set should cover the whole range of values of the respective input variable of the evaluation and test sets. Otherwise, instability phenomena and poor convergence will appear.

In this study, the proposed ANN method is applied for the estimation of the face's settlement examining the following four different scenarios for the input variable of the overload factor N_S: (a) Scenario 1: $N_S \neq 5$, (b) Scenario 2: $N_S > 4$, (c) Scenario 3: $N_S < 7$, (d) Scenario 4: N_S takes all available values of the region [3.5, 7.0]. The training set is formed using 80% of the available input vectors, while the rest ones forms the evaluation set (p=80%). In Table 2 the population of the members (input vectors) of the training set and of the evaluation set is presented. It is noted that there are 495 members for each value of N_S, so for the first scenario one value of N_S has been omitted and the rest members (2475-495=1980) are divided according to the percentage 80% (1584 for the training set and 396 for the evaluation set). The 495 omitted members are added in the evaluation set.

For all scenarios the number of neurons varies from 2 to 15, while the remaining parameters are assigned with fixed values ($a_0 = 0.6$, $T_a = 1000$, $\eta_0 = 0.6$, $T_\eta = 1000$, activation functions in both layers: hyperbolic tangent, h_1=0.7, h_2=0.0). The respective results for the R^2 indexes of the two sets are presented in Table 2. From these results, it can be seen that the systematic omission of specific values of one important input variable gives poor results. The results of scenario 1 are worse than those of scenario 3, even though the population of the training set is the same (scenario 1: $N_S \neq 5$, scenario 3: $N_S \neq 7$). The reason is that the meaningfulness of the value "5" for N_S is larger than the respective one of the second case. The results of scenario 2 are the worst because of the absence of the low part of the values set. The respective range is limited between 5 and 7. The population of the training set is not the most crucial parameter, as it has been proved by the comparison of scenarios 1 and 3. The respective optimized population of the hidden layer's neurons can also be changed significantly (see scenario 1 & 3 in Table 2), because the ANN is supplied with a different training set and the respective optimization process forms a different structure of ANN.

TABLE 2
DIFFERENT SCENARIOS FOR THE VALUES SET OF THE OVERLOAD FACTOR N_S FOR THE TRAINING SET USING THE PROPOSED ANN METHOD FOR THE FACE'S SETTLEMENT WITH THE OPTIMIZED NUMBER OF NEURONS

Scenario	Neurons	Number of members of training set	Number of members of evaluation set	R^2 for training set	R^2 for evaluation set
1	14	1584	495+396=891	62.0%	63.9%
2	11	1188	2*495+297=1287	41.1%	47.3%
3	15	1584	495+396=891	70.1%	66.2%
4	11	1980	495	90.4%	79.9%

6.4 Omission of One Input Variable

The omission of one crucial input variable can give poor convergence. For the purpose of this study, the proposed ANN method is applied for the estimation of the face's settlement omitting the modulus of elasticity for the surrounding rock-mass E_r. The parameter E_r is preferred than the overload factor N_S, because it is already known the significance of the last one from § 6.3.

For this application the training set is formed using 80% of the available input vectors (p=80%). The number of neurons varies from 2 to 15, while the remaining parameters are assigned with the fixed values at § 6.3. The optimized population of the hidden layer's neurons is 13 against 11 of the basic scenario. The respective results of this application of the ANN method are the following:

- the mean value of the absolute error *MAE* is 9.0 mm for the training set and 9.3 mm for the evaluation set instead of 8.8 mm and 9.1 mm respectively from the basic scenario,
- the R^2 correlation index between the estimated and the settlements from FLAC3D program is 68.3% for the training set and 72.2% for the evaluation set instead of 90.4% and 79.9% respectively from the basic scenario.

It is observed that the results are slightly better for the evaluation set than ones for the training set, because of the no common members between the training and evaluation sets.

If the same process repeats for any other input variable, the omission gives poorer results than the respective one with the omission of the modulus of elasticity for the surrounding rock-mass E_r (the R^2 indexes are lower than 40% for all cases for both sets). The latter proves the necessary use of all input variables.

7 Conclusions

This paper describes an optimized artificial neural network method in order to estimate the settlements of roof, face and walls during tunneling excavation.

The ANN method uses as input variables the overload factor, the placement of the temporary lining ring behind the face, the thickness of the shotcrete, the modulus of elasticity for the surrounding rock-mass and the position of the measurement point referring to the tunnel axis.

For each kind of settlement (roof, face or walls) a different ANN is formed using an optimization process regarding the values of crucial parameters, such as the number of the hidden layer's neurons, the time parameter and the initial value of the learning rate and of the momentum term, the kind and the parameters of the activation functions. The implementation of the optimization process is realized gradually based on the correlation index R^2 between the experimental and estimated values for the evaluation set. Finally the ANN with the closest to 1 correlation index is selected.

After the proposed optimization process, the basic results are the following with respect to the face, roof and walls settlements:

- the mean value of the absolute error *MAE* is 2.3 mm, 4.7 mm and 6.0 mm respectively,
- the root mean square error *RMSE* is 3.2 mm, 6.7 mm and 9.0 mm respectively,
- the R^2 correlation index between the estimated and the settlements from FLAC3D program is 98.6%, 96.1% and 93.1% for the evaluation set respectively.

Finally, a sensitivity analysis for the proposed ANN method is carried out. Specifically the following cases have been examined:

- The differences between the stochastic training and serial one have been studied showing the superiority of the random presentation of the input vectors per epoch (stochastic training).
- The random selection of the input vectors per epoch and the random initialization of the weights have also been examined showing the stability of the proposed ANN method.
- The effects of the omission of specific values of the input variables have been studied, from which the necessity of the values' set of each input variable of the training set to overlap the values' set of the respective input variable of the evaluation set has been proved.
- The omission of one input variable from the input vectors have been examined, from which the necessity of all input variables has been evident.

It is mentioned that the ANN could use different parameters than the five pre-chosen variables or in situ material properties as input variables. After the optimized training process the ANN's structure might be different than the proposed one, but the ANN could estimate the respective settlement. In any case the proposed optimized ANN can be self-adjusted, if it is supplied with different input variables, and the most suitable structure will be formed according to the best R^2 correlation index of the evaluation set. The reliability of the ANN depends from the kind of the final chosen input variables and from the accuracy of the respective measurements.

Acknowledgements

The authors would like to thank Dr. John Spyropoulos for providing the necessary data, Dr Maria-Katerina Nikolinakou from MIT for her kindly thoughts and Professor M. Kavvadas from NTUA for his guidance.

Appendix

For the purposes of this study, the results of geotechnical finite difference analyses were used. These analyses were performed by Spyropoulos [34] via the usage of the finite difference program FLAC3D ver. 2.00.

The input data has been estimated using Hoek-Brown criterion with equivalent Mohr-Coulomb parameters for different isotropic uniform rock-mass qualities. The basic assumptions are:

- the cross-section diameter of tunnel D is 11.00 m,
- the special weight of rock-mass γ is 24 kN/m^3,
- the overhead thickness of rock-mass H is between 40 m and 1800 m,
- the excavation is performed with conventional methods,
- the Geological Strength Index *GSI* is between 10 and 60,
- the factor of the rock quality m_i is 10,
- the uniaxial compressive strength of the rock-mass σ_{cm} ranges in the region [0.6,40] (in MPa),
- the Poisson's ratio of rock-mass v_r is 0.333,
- the horizontal pressure factor K_0 is 0.50,
- the modulus of elasticity for the surrounding rock-mass E_r ranges in the region $[500, 9000]$ (in MPa),
- the overload factor N_S has values from the set $\{3.5, 4.0, 5.0, 6.0, 7.0\}$,
- the equivalent cohesion c ranges in the region [200,2000] (in kPa),
- the equivalent friction angle ϕ ranges in the region [18^o, 32^o],
- the equivalent diastolic angle ψ is $\phi/5$,
- the temporary support system is closed shotcrete

rings,

- the Poisson's ratio of shotcrete v_s is 0.25,
- the modulus of elasticity for the shotcrete E_s is 10 GPa,
- the thickness of the shotcrete t has values from the set $\{10, 20, 30\}$ (in cm),
- the placement of the temporary lining ring behind the face L has values from the set $\{1.0, 3.0, 5.0\}$ (in m),
- the position x of the measurement point referring to the tunnel axis ranges in region $[-23, 33]$ (in m). Practically the negative values of x correspond behind the face of the tunnel, while the positive values correspond in front of the face, where the excavation has not taken place. Virtually, the position x of the measurement point, referring to the tunnel axis, is extended two diameters behind the tunnel face and three diameters in front of the face.

In ANN the following 34 discrete values of the position x has been used: {-23.000, -22.000,-21.000, -20.000, -19.000, -18.000, -17.000, -16.000,-15.000, -14.000, -13.000, -12.000, -11.000, -10.000, -9.000, -8.000, -7.000, -6.000, -5.000, -4.000, -3.000,-2.000, -1.000, 0.000, 0.992, 2.223, 3.784, 5.722, 8.145, 11.173, 14.959, 19.691, 25.606, 33.000} (in m). For the estimation of the face settlement only the eleven positive values has been used. For the estimation of the side walls and roof settlements the total group of the values has been used. In [34] 225 scenarios has been developed, so 225*11= 2475 input vectors have been formed for the settlement of the face, while 225*34= 7650 input vectors have been formed for the settlements of the wall and of the roof.

It is noted that the analytical datasets can be found in [34], while the software ANN code of this method has been registered in [37].

References:

[1] S. Suwansawat. *Earth Pressure Balance Shield Tunneling in Bangkok: Ground Response and Prediction of Surface Settlements Using Artificial Neural Networks*. Doctoral Thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, USA, June 2002.

[2] J. Litwinsizyn. Application of the equation of stochastic processes to mechanics of loose bodies. *Arch. Mech.* Stosow, Vol. 8, 1956, pp. 396-411.

[3] R.B. Peck. Deep excavations and tunnelling in soft ground. *Proc. 7th Int. Conf. on Soil Mechanics and Foundation Engineering*, Mexico City, State of the Art Volume, 1969. pp. 225-290.

[4] M.P. O'Reilly, B.M. New. Settlement above tunnels in the United Kingdom – their magnitude and prediction. London, *Tunneling*, Vol. 88, 1982, 231-241.

[5] E.J. Cording, W.H. Hansmire. Displacement around tunnels in soil. *General report 5th Pan American Conference on Soil Mechanics and Foundation Engineering*, Buenos Aires, Session IV, 1975, pp. 571-632.

[6] P.B. Attewell, J.P. Woodman. Predicting the dynamics of ground settlement and its dericatives caused by tunnelling in soil. *Ground Engineering*, Vol. 15, No. 8, 1982, pp. 13-22, 36.

[7] M. Kavvadas. *Design of underground constructions*. School of Civil Engineering, National Technical University of Athens, March 2000 (in Greek).

[8] M. Panet. *Le calcul des tunnels par la methode convergence–confinement*. Presses de L'ecole Nationale des Ponts et Chaussees, Paris, 1995.

[9] E. Hoek, P.K.Kaiser, W.F. Bawden. *Support of underground excavations*. Balkema, 1994.

[10] P.P. Oreste. Analysis of structural interaction in tunnels using the convergence-confinement approach. *Tunnelling and Underground Space Technology*, Vol. 18, 2003, pp. 347–363.

[11] C. Sagaseta. Analysis of unstrained soil deformations due to ground loss. *Geotechnique*, Vol. 37, No. 3, 1987, pp. 301-320.

[12] A. Verruijit, J.R. Booker. Surface settlements due to deformation tunnel in an elastic half plane. *Geotechnique*. Vol. 46, No. 4, 1996, pp. 753-756.

[13] C. Gonzalez, C. Sagaseta. Patterns of soil deformations around tunnels. Application to the extension of Madrid Metro. *Computer and Geotechniques*, Vol. 28, 2001, pp. 445-468.

[14] N. Loganathan, H.G. Poulos. Analytical prediction for tunnelling-induced ground movements in clays. *Journal of Geotechnical and Geoenvironmental Engineering*, Vol. 124, No. 9, 1988, pp. 846-856.

[15] F. Pinto. *Analytical methods to interpret ground deformations due to soft ground tunnelling*. SM Thesis, Massachusetts Institute of Technology, USA, 1999.

[16] G.W. Clough, E. Leca. With focus of use of finite element methods for soft ground

tunneling. *Review Paper in Tunnels et Micro-Tunnels en Terrain Meuble-du Chantier a la Theorie*, Paris, 1989, pp. 531-573.

[17] H. Akagi, K. Komiya. Finite element simulation of shield tunneling processes in soft ground. Proc. Int. *Symposium on geotechnical Aspects of Underground Construction in Soft Ground*, London, Balkema, pp. 447-452.

[18] K.M. Lee, R.K. Rowe. An analysis of the three-dimensional ground movements: the Thunder Bay Tunnel. *Canadian Geotech. J.* Vol. 28, 1991, pp. 25-41.

[19] E. Eberhardt. Numerical modelling of three dimensional stress rotation ahead of an advancing tunnel face. *International Journal of Rock Mechanics & Mining Sciences*, Vol. 38, 2001, pp. 499-518.

[20] J.H. Atkinson, E.T. Brown, D.M. Potts. Collapse of shallow unlined tunnels in dense sand. *Tunnel & Tunneling*, May 1975. pp. 81-87.

[21] J.H. Atkinson, D.M. Potts. Subsidence above shallow tunnels in soft ground. *Proc. ASCE Geotechnical Engineering Division*, Vol. 103, No. 4, 1997, pp. 307-325.

[22] T. Kimura, R.J. Mair. Understanding ground movements caused by tunneling. *Proc. Of Conf. on Underground Engineering*, 1981, paper no. D21G11, London.

[23] T. Nomoto, S. Imamura, T. Hagiwara, O. Kusakabe, N. Fujii. Shield tunnel construction in centrifuge. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, Vol. 125, No. 4, 1999, pp. 289-300.

[24] C.Y. Kim, G.J. Bae, S.W. Hong, C.H. Park, H.K. Moon, H.S. Shin. Neural network based prediction of ground surface settlements due to tunneling. *Computer and Geotechniques*, Vol. 28, 2001, pp. 517-547.

[25] S. Suwansawat, H.H. Einstein. Artificial neural networks for predicting the maximum surface settlement caused by EPB shield tunneling. *Tunneling and Underground Space Technology*, Vol. 21, 2006, pp.133-16.

[26] S.S. Leu, H.Ch. Lo. Neural-network-based regression model of ground surface settlement induced by deep excavation. *Automation and Construction*, Vol. 13, 2003, pp. 279-289.

[27] M.A. Shahin, H.R. Maier, M.B. Jakes. Predicting settlement of shallow foundations using neural networks. *Journal of Geotechnical and Geoenvironmental Engineering*, Vo.128, No. 9, 2002, pp. 785-793.

[28] T. Kerch, Y.G. Hu, C.H. Wu. Estimation of consolidation settlement caused by groundwater drawdown using artificial neural networks. *Advances in Engineering Software*, Vol. 34, 2003, pp. 559-568.

[29] S. Celik, O. Tan. Determination of preconsolidation pressure with artificial neural network. *Civil Engineering and Environmental Systems*, Vol. 22, No. 4, 2005, pp. 217-231.

[30] M.S. Rahman, J. Wang, W. Deng, J.P. Carter. A neural network model for the uplift capacity of suction caissons. *Computer and Engineering*, Vol. 28, 2001, pp. 269-287.

[31] N. Caglar, H. Arman. The applicability of neural networks in determination profiles. *Bull Eng. Geol. Environ*. Vol. 66, 2007, pp. 295-301.

[32] Q. Fu, Y.M.A. Hashash, S. Jung, J. Ghaboussi. Integration of laboratory testing and constitutive modeling of soils. *Computer and Geotechnics*. Vol. 34, 2007, pp. 330-345.

[33] Z. L. Wang, Yong-chi Li, R.F. Chen. Correction of soil parameters in calculation of embankment settlement using a BP network back-analysis model. *Engineering Geology*, Vol. 91, 2007, pp. 168-177.

[34] John Spyropoulos. *Investigation of face stability in deep tunnels.* Doctoral Thesis, School of Civil Engineering, National Technical University of Athens, 2007.

[35] H.S. Hippert, C.E. Pedreira, R.C. Souza. Neural networks for short-term load forecasting: A review and evaluation. *IEEE Trans. on Power Systems,* Vol. 16, No. 1, 2001, pp. 44-55.

[36] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 1994.

[37] George J. Tsekouras. *Application of artificial neural networks' method in tunneling problems.* Diploma Thesis, School of Civil Engineering, National Technical University of Athens, 2004.