

# Modeling and Interpretation of Pressuremeter Test Results with Artificial Neural Networks

Mohammad Emami · S. Shahaboddin Yasrobi

Received: 9 April 2013 / Accepted: 17 December 2013 / Published online: 29 December 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** In this paper, three types of artificial neural network (ANN) are employed to prediction and interpretation of pressuremeter test results. First, multi layer perceptron neural network is used. Then, neuro-fuzzy network is employed and finally radial basis function is applied. All applied networks have shown favorable performance. Finally, different models have been compared and network with the most outstanding performance in two stages is determined. Contrary to conventional behavioral models, models based neural network do not demonstrate the effect of input parameters on output parameters. This research is response to this need through conducting sensitivity analysis on the optimal structure of proposed models.

**Keywords** Pressuremeter · Prediction · Interpretation · Artificial neural network · Sensitivity analysis

---

M. Emami (✉) · S. S. Yasrobi  
Department of Civil Engineering, Faculty of Engineering,  
Tarbiat Modares University, Intersection of Jalale  
Ale Ahmad and Shahid Chamran Highways,  
P.O. Box 14155-143, Tehran, Iran  
e-mail: r\_m\_emami@yahoo.com

S. S. Yasrobi  
e-mail: yasrobis@modares.ac.ir

*Present Address:*

M. Emami  
Unit 6, No. 16, 3th St., Kamali Boulvared, Ashrafi  
Esfehani, St. Pounak, P.O. Box 1476964171, Tehran, Iran

## 1 Introduction

In-situ tests play important role in any geotechnical investigation. Pressuremeter test can be considered one of the most important in situ tests. This test is capable to properly estimate the ground strength and compressibility under measured deformations of the probe.

During the past decades, increasing interest has been shown in the development of a satisfactory formulation for the stress–strain relationships of engineering soils that incorporates a concise statement of nonlinearity, inelasticity and stress dependency based on a set of assumptions and proposed failure criteria. In spite of the considerable complexities of these constitutive models, and due to an inadequate understanding of the mechanisms and all factors involved, it is not possible to capture the complete material response along all complex stress paths and densities. Furthermore, the degree of complexity of these constitutive models (in many cases) inhibits their incorporation into general purpose numerical codes, thus restricting their usefulness in engineering practice Shahin and Jaksa (2005).

Artificial neural network (ANN) offers a fundamentally different approach for modeling soil behavior. ANN is an oversimplified simulation of the human brain and composed of simple processing units referred to as neurons. It is able to learn and generalize from experimental data even if they are noisy and imperfect. This ability allows this computational

system to learn constitutive relationships of materials directly from the result of experiments. Unlike conventional models, it needs no prior knowledge, or any constants and/or assumptions about the deformation characteristics of the geomaterials. Other powerful attributes of ANN models are their flexibility and adaptivity, which play an important role in material modeling Hagan and Menhaj (1994).

When conventional models cannot reproduce a new set of experimental results, a new constitutive model or a set of new constitutive equations, needs to be developed. However, trained ANN models can be further trained with the new data set to gain the required additional information needed to reproduce the new experimental results. These features ascertain the ANN model to be an objective model that can truly represent natural neural connections among variables, rather than a subjective model, which assumes variables obeying a set of predefined relations Hagan et al. (1996).

In recent times, ANNs have been applied successfully to many prediction tasks in geotechnical engineering, as they have the ability to model nonlinear relationships between a set of input variables and corresponding outputs. Shahin et al. (2001) give a comprehensive list of the applications of ANNs in geotechnical engineering. A review of the literature reveals that ANNs have been used successfully in pile capacity prediction, modeling soil behavior, site characterization, earth retaining structures, settlement of structures, slope stability, design of tunnels and underground openings, liquefaction, soil permeability and hydraulic conductivity, soil compaction, soil swelling and classification of soils. Najjar and Ali (1998) used neural networks to characterize the soil liquefaction resistance utilizing field data sets representing various earthquake sites from around the world.

Goh (1994a; 1995b) presented a neural network to predict the friction capacity of piles in clays. Chan et al. (1995) developed a neural network as an alternative to pile driving formulae. Lee and Lee (1996) utilized neural networks to predict the ultimate bearing capacity of piles. Kumar and Kumar (2006) applied the artificial neural network model to predict the lateral load capacity of piles in clay. Teh et al. (1997) proposed a neural network for estimating the static pile capacity determined from dynamic stress-wave data for precast reinforced concrete piles with a

square section. Goh (1994a) developed a neural network for the prediction of settlement of a vertically loaded pile foundation in a homogeneous soil stratum. Shahin et al. (2000) carried out similar work for predicting the settlement of shallow foundations on cohesionless soils.

Ellis et al. (1995) developed an ANN model for sands based on grain size distribution and stress history. Penumadu and Zhao (1999) also used ANNs to model the stress–strain and volume change behavior of sand and gravel under drained triaxial compression test conditions. Banimahd et al. (2005) utilized artificial neural network for stress–strain behavior of sandy soils. Lee et al. (2003) presented an approach to estimate unsaturated shear strength using artificial neural network. Moosavi et al. (2006) utilized artificial neural networks for modeling the cyclic swelling pressure of mudrock. Kim and Bae (2004) presented a neural network based prediction of ground surface settlements due to tunneling. Kanungo et al. (2006) done a comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas.

So far, ANNs have been applied to the constitutive modeling of rocks Kanungo et al. (2006), clays Ghaboussi and Sidarta (1998), clean sands Nabney (1999), gravels Buckley and Hayashi (1994a) and residual soils Hagan and Menhaj (1994). It has also been shown that ANN can generalize traditional constitutive laws well (e.g. a hyperbolic model) by considering their descriptive parameters Shahin and Jaksa (2005). Despite their good performance on the available data ANN models give no clue on the way inputs affect the output and are therefore considered as a *\_black box* class of model. The lack of interpretability of ANN models has inhibited them from achieving their full potential in real world problems Moosavi et al. (2006) as the credibility of the artificial intelligence paradigm frequently depends on its ability to explain its conclusion Shahin and Jaksa (2005).

Therefore, for verification of such models, as well as the accuracy measuring of ANN based models with available data, a methodology should be adopted to extract the meaningful rule from the trained networks, which are comparable with trends inferred from experiments. Lu et al. (Moosavi et al. 2006) reviewed the methods that have been introduced to acquire the

knowledge contained in a trained ANN [e.g. fuzzy logic Lee et al. (2000) and principal component analysis Shahin and Jaksa (2005)] and stated that these methods could not determine the effect of each input parameter on the output variable, in terms of magnitude and direction. They defined input sensitivity based on the first order partial derivative between the ANN output variable and the input parameters in mathematical term for the sensitivity analysis of spool fabrication productivity problems. Hashem (Shahin and Jaksa 2005) has also obtained this formulation. As will be discussed later, this formulation cannot be used in the sensitivity analysis of ANN constitutive soil models.

## 2 Pressuremeter Test Results

The pressuremeter test is one of the most important in situ tests for determination of the stress–strain behavior of subsurface layers. The pressuremeter consists of two main elements: a radially expandable cylindrical probe, which is placed inside the borehole at the desired test depth and a monitoring unit, which remains on the ground surface. These parameters can be obtained from pressuremeter test results:

- Pressuremeter Modulus  $E_m$ .
- Undrained shear strength  $C_u$  for clays
- Shear modulus
- Pressuremeter limit pressure of the ground

There are several different kinds of pressuremeters (Briaud 1992; Clarke and Gambin 1998) that differ mainly by the way the probe is placed in the ground.

- The pre-bored or predrilled pressuremeter (PBP).
- The self-bored pressuremeter (SBP).
- The cone pressuremeter, either pushed or driven in place.
- The pushed Shelby tube pressuremeter.
- The Predrilled pressuremeter, that is the Menard pressuremeter (MPM), which is a volume-displacement pressuremeter, has been used in this project. It is based on expansion of the central test section of the probe that is placed inside the borehole and the pressure–volume variation during testing is recorded. The Menard pressuremeter, which has been used in this project, is shown in Fig. 1.



**Fig. 1** A view of used Menard pressuremeter instrument

### 2.1 Analysis of Pressuremeter Tests According to ASTM D 4719

The procedure to perform the pressuremeter test data is as follow:

1. Volume and Pressure loss calibration. The pressure difference between cell and core  $P_c$  based on calibration tests.
2. Determination of hydrostatic pressure between cell and control unit  $P_\delta$

$$P_\delta = H \times \gamma_1 \quad (1)$$

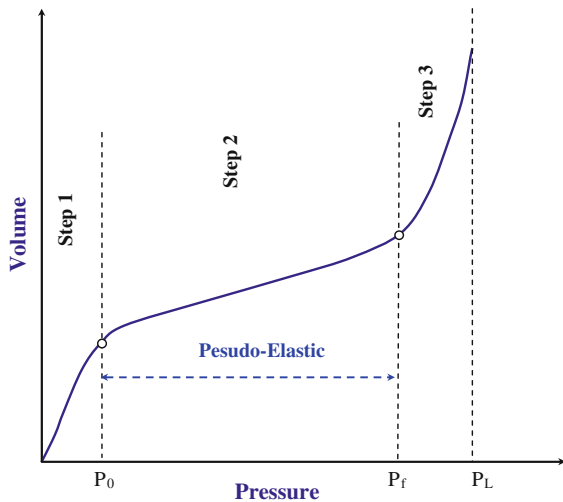
$H$ : Depth of cell placement with respect to control unit (in m)

$\gamma_1$ : Unit weight of fluid (in  $\text{KN}/\text{m}^3$ )

3. Pressuremeter test data correction based on above-mentioned items (1–3):

$$P_G = P_{\text{pressuremeter}} + P_\delta - P_c \quad (2)$$

4. Drawing pressure–volume change curve such as Fig. 2.
5. Determination of  $\Delta P$ ,  $\Delta V$  and  $V_m$  from Fig. 2.  
 $\Delta P$ : Corrected pressure difference in the linear part of the graph  
 $\Delta V$ : Corrected volume difference in the linear part of the graph



**Fig. 2** Pressure–volume change curve in pressuremeter test

$V_m$ : Corrected volume reading in the central portion of the  $\Delta V$  volume increase

6. Based on the values of item 5 and  $V_0$  (primary volume of the probe at ground surface), pressuremeter, Modulus Menard ( $E_M$ ) is calculated::

$$E_p = E_m = 2(1 + \nu)(V_0 + V_m) \frac{\Delta P}{\Delta V} \quad (3)$$

The pressuremeter Modulus ( $E_M$ ) is related to the Oedometer Modulus ( $E_s$ ) with following relation:

$$E_s = E_{oed} = E_p / \alpha \quad (4)$$

The values of  $\alpha$  is presented in Table 1.

In very compact embankment the values of  $\alpha$  may be more than one  $\alpha = 1$ . According to above Table  $\alpha$  is equal to 0.67 in marls.  $\alpha$  is equal to  $\alpha = 0.67$  in Normally-Consolidated clays and  $\alpha = 1$  for Over-Consolidated clays.

The values of pressuremeter modulus obtained from pressuremeter tests are converted to Young modulus as follow:

Elasticity Modulus – Young Modulus :

$$E = \frac{E_p(1 + \nu)(1 - 2\nu)}{(1 - \nu)} \quad (5)$$

**Table 1** The values of  $\alpha$  for various soil types (Menard and Rousseau 1962)

Soil type	Clay	Silt	Sand	Gravel
$\alpha$	0.67	0.5	0.34	0.25

$\nu$ : Poisson’s ratio ( $\nu = 0,33$  usually in pressuremeter theory).

### 3 Database

In this research, the results of approximately 500 conducted pressuremeter tests on the soils by Pajohesh Omran Rahvar Ltd (2006–2007) is employed. The number of tests decreased to 400 due to lack of accuracy and also high changes in the range of pressuremeter modules. The tests have been carried out on the soils of Northwest Iran (Tabriz), South Iran (Kharg Island) and Northeast Iran (Mashhad). The pressuremeter instrument used is a pre-bored Menard MPM type. Tests performed according to ASTM-D4719 represented acceptable accuracy. An example of physical and density properties of soils is presented at Table 2. In addition, a number of soil grading curves are shown at Figs. 3 and 4.

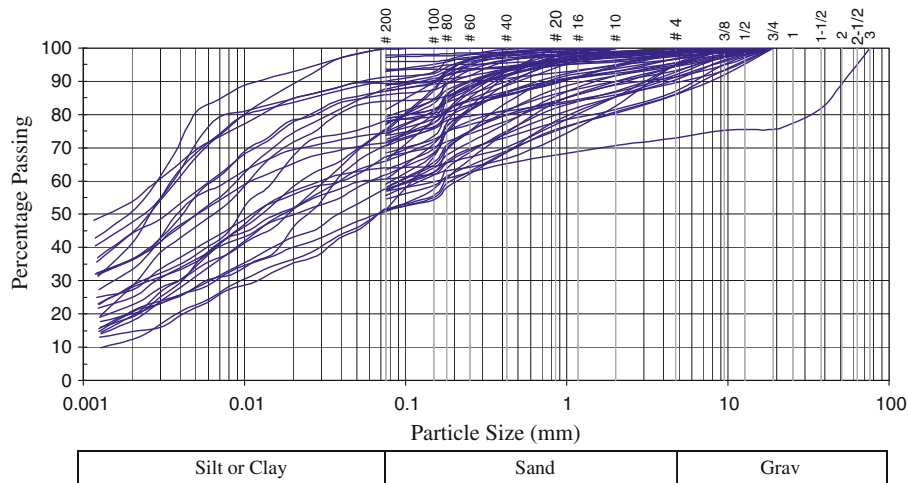
Bank information used in the current study is divided into two different categories to predict volume change–pressure curve obtained from pressuremeter test and to interpret pressuremeter module. Due to using applied pressure as one of the input parameters, the number of data at the prediction stage is more than that of at the interpretation stage. 500 data resulting from pressuremeter tests are employed at the stage of prediction. A number of volume change–pressure charts of pressuremeter tests are presented at Fig. 5.

### 4 Model Development

In this study, three types of artificial neural network (ANN) are used to predict and interpret the

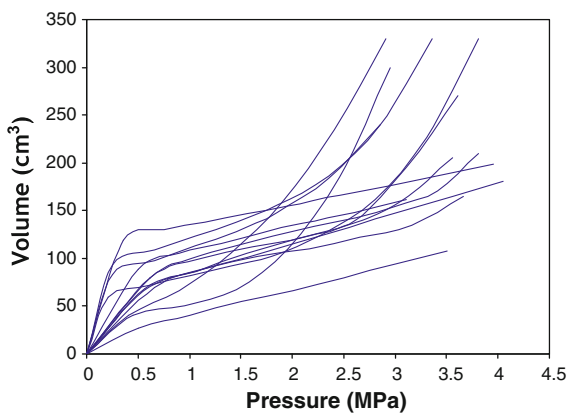
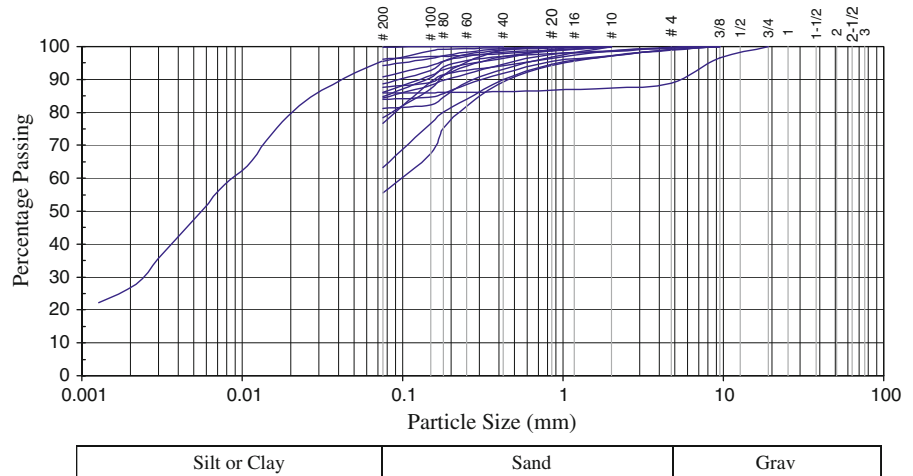
**Table 2** Physical and density properties of soils

$N_{SPT}$	Grain maximum size ( $D_{max}$ ) (mm)	Normal specific weight ( $kN/m^3$ )	Water content (%)	Depth (m)
12	9.5	18.2	18	6.75
32	0.425	17.7	14.5	7.75
33	24.5	18.1	15.6	8.75
28	38.1	18.5	14.2	9.25
36	24.5	17.8	18.5	10.75
23	4.75	17.0	23.2	11.75
49	3.35	16.0	21.3	13.25
30	24.5	17.5	18.1	14.75



**Fig. 3** Particle size distribution of ML soils

**Fig. 4** Particle size distribution of CL soils



**Fig. 5** Volume change-pressure charts of pressuremeter tests

pressuremeter test. First, multi layer perceptron neural network, one of the most applicable neural networks, is used. Two types of perceptron neural network with one and two hidden layers, respectively, along with different neurons in the hidden layers are used. Perceptron networks are employed both to obtain volume change–pressure charst at the prediction stage and to estimate the Menard pressuremeter module (EM) at the interpretation stage. Physical and density properties of the soils have been used as model input parameters. 6 and 5 input parameters have been used in the prediction stage of the network architecture and interpretation stage, respectively. One parameter is considered in both stages as output parameter. Hidden

**Table 3** The boundaries of input and output parameters

Limits	Inputs						Output	
	$N_{SPT}$	$\gamma_m$ (KN/m <sup>3</sup> )	W (%)	H (m)	$D_{max}$ (mm)	P (KPa)	$\Delta V$ (cm <sup>3</sup> )	$E_m$ (KPa)
Min.	10	1.53	10	3	0.25	100	10	10,000
Max.	50	1.88	26	30	38.1	3,000	120	40,000

layers with different number of neurons are tested in both one and two layers networks so as to select the most proper network architecture. It has been shown that a three-layer perceptron with differential transfer functions and sufficient number of neurons in hidden layer can approximate any nonlinear relationship Shahin and Jaksa (2005). Consequently, one hidden layer is used in the present study. The neural network toolbox of MATLAB7.4, a popular numerical computation and visualization software, is used for training and testing of the MLPs. Transfer functions of networks are selected by trial and error. Boundaries of inputs and outputs are given in Table 3.

The Levenberg–Marquardt algorithm is used to train the networks. This method, which is an approximation of Newton's method, has been shown to be one of the fastest algorithms for training moderate size MLPs Hagan and Menhaj (1994). In order to apply this algorithm a static training approach is utilized, i.e. This approach, which has been used satisfactorily by Ghaboussi et al. (Hagan and Menhaj 1994) does not have the potential drawback in that errors accumulate excessively upon each step during the training process Nabney (1999). However it lacks similarity of training and prediction phases Nabney (1999).

To improve generalization of MLPs, training should be stopped when overfitting is started. Overfitting makes MLPs memorize training patterns in such a way that they cannot generalize well to new data. In the present study, cross validation technique are used as the stopping criterion. In this technique, the database is divided into three sets: training, validation and testing. The training set is used to update networks' weights. During this process, the error on the validation set is monitored. When the error on the validation set begins to increase, the training should be stopped because it is considered to be the best point of generalization. Finally, testing data is fed into the networks to evaluate their performance. In this study, the results of 250, 70 and 80 pressuremeter tests are used for training, validation and testing of networks,

respectively. In order to choose the best structure of the model the performances of MLPs with different number of hidden neurons are studied.

The other neural network examined in this research is neuron-fuzzy network. In this network, 6 and 5 physical and density parameters are used as input parameters. Also, one parameter is used as output parameter. This networks utilized a fuzzy system with backpropagation algorithm. In order to update the connection weights involved in the fuzzy neurons (FNs), some learning and adaptation mechanisms for the FN models that were proposed in the last section are presented in this section. Like the least square error functions used in the conventional BP algorithm for multilayered feedforward neural networks (MFNNs), the generic performance index used here is also expressed as a squared error between the output of the fuzzy neuron and a desired value. The learning procedure for the free parameters in such a neural network is considered on the basis of the elements of the set of the training patterns. Given a set of input and desired output pairs the adaptive weight learning rule performs an optimization process such that the output error function, defined as the summation of the square of the errors between the desired and the real outputs of the network, is minimized.

Finally, radial basis function neural network is applied to evaluate the success of two previous networks. The number of input and output parameters is similar to the neuron-fuzzy and perceptron networks. Radial basis function (RBF) neural networks have recently been studied intensively. The RBF neural network has the universal approximation ability, therefore, the RBF neural network can be used for the interpolation problem. A Gaussian radial basis function, an unnormalized form of the Gaussian density function, is highly nonlinear, and it provides some good characteristics for incremental learning, and has many well-defined mathematical features. Gaussian neural networks, which have been found to be powerful scheme for learning complex input–

**Table 4** The structural details of neural network models

Networks	Inputs	Output
<i>Prediction</i>		
MLP1	$N_{SPT}$ , $\gamma_m$ (KN/m <sup>3</sup> ), $W$ (%), $D_{max}$ , $H$ (m), $P$ (KPa)	$\Delta V$ (cm <sup>3</sup> )
MLP2	$N_{SPT}$ , $\gamma_m$ (KN/m <sup>3</sup> ), $W$ (%), $D_{max}$ , $H$ (m)	$\Delta V$ (cm <sup>3</sup> )
NF1	$N_{SPT}$ , $\gamma_m$ (KN/m <sup>3</sup> ), $W$ (%), $D_{max}$ , $H$ (m), $P$ (KPa)	$\Delta V$ (cm <sup>3</sup> )
NF2	$N_{SPT}$ , $\gamma_m$ (KN/m <sup>3</sup> ), $W$ (%), $D_{max}$ , $H$ (m)	$\Delta V$ (cm <sup>3</sup> )
RBF1	$N_{SPT}$ , $\gamma_m$ (KN/m <sup>3</sup> ), $W$ (%), $D_{max}$ , $H$ (m), $P$ (KPa)	$\Delta V$ (cm <sup>3</sup> )
RBF2	$N_{SPT}$ , $\gamma_m$ (KN/m <sup>3</sup> ), $W$ (%), $D_{max}$ , $H$ (m)	$\Delta V$ (cm <sup>3</sup> )
<i>Interpretation</i>		
MLPI	$N_{SPT}$ , $\gamma_m$ (KN/m <sup>3</sup> ), $W$ (%), $D_{max}$ , $H$ (m)	$E_m$ (KN/m <sup>3</sup> )
ANFISI	$N_{SPT}$ , $\gamma_m$ (KN/m <sup>3</sup> ), $W$ (%), $D_{max}$ , $H$ (m)	$E_m$ (KN/m <sup>3</sup> )
RBF1	$N_{SPT}$ , $\gamma_m$ (KN/m <sup>3</sup> ), $W$ (%), $D_{max}$ , $H$ (m)	$E_m$ (KN/m <sup>3</sup> )

output mapping, have been used in learning, identification, equalization, and control of nonlinear dynamic systems.

The term radial basis function derives from the fact that these functions are radially symmetric; that is, each node produces an identical output for inputs that lie at a fixed radial distance from the center. In other words, a radial basis function  $\varphi(x-c_i)$  has the same value for all neural inputs  $x$  that lie on a hypersphere with the center  $c_i$ .

It has been shown that the Gaussian RBF neural networks are capable of uniformly approximating arbitrary continuous functions defined on a compact set to satisfy a given approximation error. This approximation process is usually carried out by a learning phase where the number of hidden nodes and the network parameters are appropriately adjusted so that the approximation error is minimized. There are a variety of approaches for using the Gaussian networks. Most of them start by breaking the problem into two stages: learning in the intermediate stage, that is, adjusting the center and variance parameters, followed by learning or adjusting the weight parameters of the linear combiners in the output stage. Learning in the intermediate stage is typically performed using the clustering algorithm, while learning in the output stage is a supervised learning. Once an initial solution is found using this approach, a supervised learning algorithm is sometimes applied to both stages simultaneously to update the parameters of the network.

In what follows, the optimal structure of each model based on their performance in error index is introduced. Finally, the volume change-pressure chart

resulting from optimal structures of neural network models at the prediction stage have been compared with the empirical results. In order to evaluate the generalization ability of the optimal structure of neural network at the prediction stage, this model is evaluated with the inexperienced data. The structural details of neural network models are presented at Table 4.

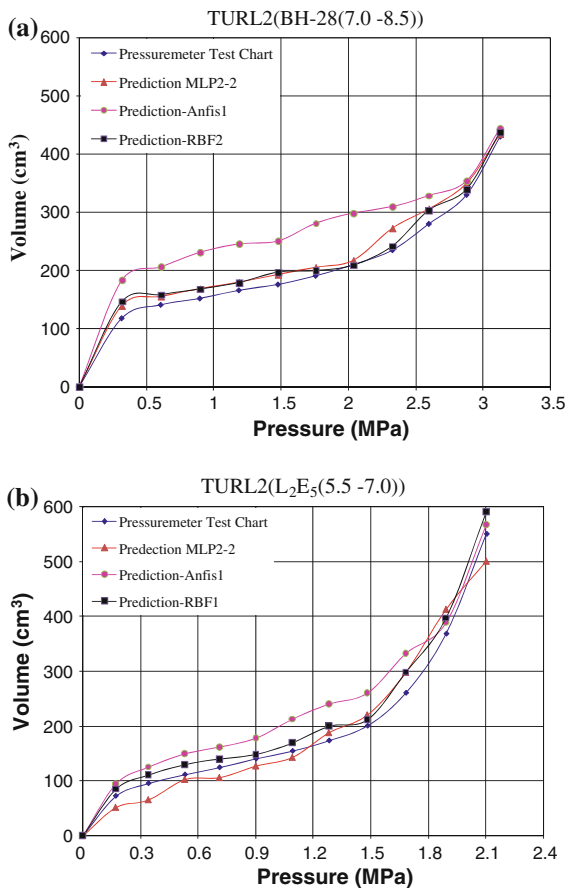
This section deals with the comparison of efficiency of the optimal structures of three neural network models. Table 5 shows the error index for three models have been shown for training, test and validation subsets. As can be seen clearly, MLP2 network with two hidden layers, 15 neurons in each layer proved to show the best performance compared to other three models. However, the MLP1 network with two hidden layers, 20 neurons in each layer showed the best performance in the training subset. MLP2 network has been assigned the lowest values of other error index. Therefore, the mentioned network can be considered as the most successful network to predict the results of pressuremeter test.

The prediction of the change volume-pressure chart resulting from pressuremeter test for the optimal structures of perception, neuron-fuzzy and radial basis function networks are shown in the Fig. 6a, b. These charts are plotted based on the comparison with empirical charts. As can be seen, the obtained charts predict the soil behavior with satisfactory accuracy. Also, neuron-fuzzy network showed the least accuracy in this stage.

This section discusses the comparison of the performance of optimal structures for three used neural network models. Table 6 shows the error index

**Table 5** The comparison of the optimal structures of models based on the error index at prediction stage

Networks	Num. of neurons or MFs		R	RMSE	MAE	MAXAE	SDAE	SSE
<i>Validation subset</i>								
MLP1-2	20	20	0.94	0.061	0.033	0.37	0.056	0.75
MLP2-2	15	15	0.95	0.059	0.036	0.30	0.048	0.73
<i>Training subset</i>								
MLP1-2	20	20	0.99	0.022	0.007	0.27	0.021	0.48
MLP2-2	15	15	0.98	0.037	0.018	0.32	0.028	1.3
NF1	3		0.94	0.050	0.042	0.55	0.055	2.5
RBF2	250		0.98	0.033	0.018	0.32	0.027	1.1
<i>Testing subset</i>								
MLP1-2	20	20	0.90	0.080	0.057	0.30	0.045	1.35
MLP2-2	15	15	0.94	0.062	0.036	0.33	0.051	0.61
NF1	3		0.90	0.080	0.065	0.50	0.065	3.7
RBF2	250		0.92	0.072	0.042	0.46	0.053	1.8

**Fig. 6** **a** The obtained charts from the neural network models. **b** The obtained charts from the neural network models

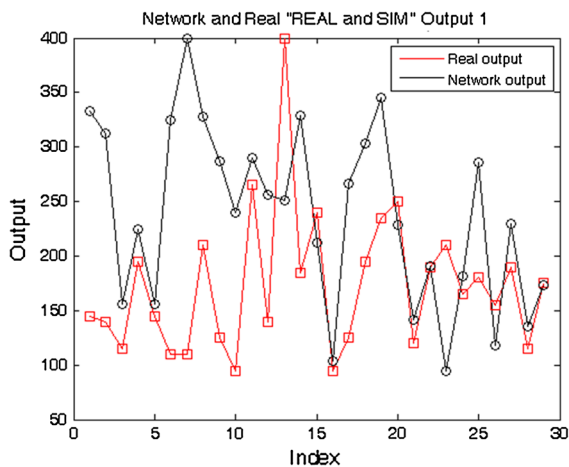
of three models for training, test and validation subsets. As can be seen clearly, MLP2 network with two hidden layers, 15 neurons in each layer represented the best performance compared to other three models. However, the RBF1 network showed the best performance in the training subset. MLP1 network has been assigned the lowest values in other error index. Therefore, the mentioned network can be considered as the most successful network to predict the results of pressuremeter test.

In order to evaluate the generalization ability of the optimal structure of neural network, this model is evaluated with the inexperienced data. For this purpose, the pressuremeter tests conducted in the second phase of geotechnical investigation of urban train of Tabriz are employed. pressuremeter tests have been conducted on soils similar to the first phase. Due the availability of 30 pressuremeter tests, the results of simulation have been compared with the results of the best neural network model. According to the previous section, the MLP-11 model with 15 neurons in the hidden layer is selected as the most proper model. The obtained chart is shown at Fig. 7. As can be seen, the network represents acceptable performance in the simulation of inexperienced data. Considering the fact that the pressuremeter test is an in situ test, the prediction of pressuremeter module ( $E_p$ ), as the most important output of pressuremeter test, has been done with satisfactory error by the proposed neural network.



**Table 6** The comparison of optimal structures of models based on the error index at interpretation stage

Networks	Num. of neurons or MFs	R	RMSE	MAE	MAXAE	SDAE	SSE
<i>Validation subset</i>							
MLP-I1	15	0.96	0.060	0.042	0.26	0.037	0.43
MLP-I2	11                      11	0.94	0.062	0.052	0.18	0.051	0.45
<i>Training subset</i>							
MLP-I1	15	0.96	0.073	0.051	0.16	0.041	2.0
MLP-I2	11                      11	0.96	0.068	0.052	0.21	0.043	2.4
NFI	3	0.96	0.061	0.050	0.18	0.050	2.2
RBF1	400	0.98	0.035	0.022	0.19	0.025	0.82
<i>Testing subset</i>							
MLP-I1	15	0.95	0.064	0.053	0.16	0.044	0.56
MLP-I2	11                      11	0.96	0.068	0.051	0.19	0.052	0.55
NFI	3	0.92	0.088	0.068	0.29	0.063	1.5
RBF1	400	0.87	0.10	0.085	0.38	0.072	3.3



**Fig. 7** Comparison between simulation and empirical result

Accordingly, the neural network model showed acceptable performance in the interpretation of pressuremeter test.

**5 Sensitivity Analysis**

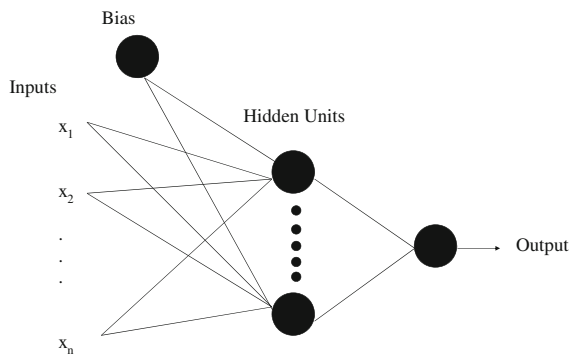
In this section, final structures of the models are investigated at each stage including feedforward neural network with one hidden layer, 15 neurons in each layer and an output layer, 1 neuron in each layer. And as such, indeterminate and sensitivity analysis have been conducted on the neural network model to

evaluate the efficiency and sensitivity of the parameters.

Although neural network is a robust method to learn unknown and complex relation from input space into output space, contrary to mathematic models do not automatically explain the effects of input parameters on the output parameters and the way that model output is determined Lu et al. (2001). In order to resolve this problem, numerous studies have been so far conducted to explain the governing equation of the neural network and the effects of input parameters on the output parameters. In this section, first the definition of the indeterminate and sensitivity analysis is presented.

The relation of output derivative in relation to input is proposed for optimal structures of the network at each stage. Then, the effect of five main input variables including normal specific weight, water content, depth of, the number of standard penetration test and the grains maximum size is discussed. This research is based on the statistical analysis of the values of the relative derivative of the outputs in relation to specified inputs at 200 points located in the five-dimensional space of inputs. These points are derived by using normal distribution function.

As mentioned above, the sensitivity of neural network output to input parameters is defined by output derivative in relation to specified input. In this section, the relation of the output relative derivative in relation to independent inputs is presented. Lu et al. (2001) determined the output relative derivative of a



**Fig. 8** Scheme of used MLP network

network in which the input were totally independent from output Lu et al. (2001). However, since at least one input is the output of the model in the last seconds in the behavior models based on the neural network, this relation cannot be used directly and should be modified. Figure 8 shows schematically the neurons of input, hidden and output layers and their connection in a MLP network that can be used as a soil behavior model.

The study of relating the uncertainties of the model output to different sources of that of the model output is called sensitivity analysis (Saltelli et al. 2004). Thus, the uncertainties analysis focuses to determine the quantity of these uncertainties.

When the relation between inputs and outputs has been defined, indeterminate and sensitivity analysis can be conducted. For this purpose, Monte Carlo analysis that deals with the distribution function of input and output parameters may be used. The Monte Carlo analysis involves the following procedure:

1. A distribution like normal distribution should be provided for each input parameters  $x_1, x_2, \dots$

Note: It is assumed that random variable  $x$  with mean  $\mu$  and standard deviation  $\sigma$  has normal distribution and its density function is as follows:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad (6)$$

and is described as  $x \sim N(\mu, \sigma)$ .

2. It is assumed that parameters are independent from each other.
3. The distribution of each input parameter is derived by using sampling methods like random sampling

methods This leads to create a matrix that its rows and columns correspond to the number of samples and input parameters, respectively.

$$\begin{bmatrix} x_1^1 & x_2^1 & \dots \\ x_1^2 & x_2^2 & \dots \\ \vdots & \vdots & \dots \\ x_1^N & x_2^N & \dots \end{bmatrix} \quad (7)$$

4. Then, derived data will be applied to the model to obtain an output matrix as follows:

$$\begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^N \end{bmatrix} \quad (8)$$

Mean, standard deviation, safety boundaries, etc. can be determined for output parameters using obtained results. After indeterminate analysis, sensitivity analysis may be conducted on the model to determine the input parameter which has the most importance to create indeterminate in the network output.

It is evident that output derivative  $y_i$  in relation to input  $x_i$ , i.e.  $\frac{\partial y}{\partial x_i}$ , can be a proper mathematic definition of model output sensitivity to its input. However, the simple derivative method may not properly describe the out sensitivity to the input. In order to clarify this issue, suppose a linear model is given by  $y = \sum_{i=1}^n \Omega_i X_i$ , where  $\Omega_i$  are fixed factors and  $X_i$  are model variables. The derivative relation is  $\frac{\partial y}{\partial X_i} = \Omega_i$ . In the event of constant  $\Omega_i$  factors, it is concluded that the output sensitivity to all input parameters is the same and consequently the inputs have the same importance for output. However, the values of standard deviation of each input may be different. In order to solve this problem, Sigma-Normalized Derivatives method may be used.

$$S_{x_i}^\sigma = \frac{\sigma_{x_i} \partial y}{\sigma_y \partial x} \quad (9)$$

If a factor is considered to be fixed,  $X_{\sim i} = x_i^*$ , in reality, an effective change sources on the output variance is fixed. Therefore, output variance is always smaller than total output variance  $V(y)$  when one parameter is fixed  $V(y|X_i = x_i^*)$ .

$$V(y|X_i = x_i^*) \leq V(y) \tag{10}$$

Accordingly, it is expected that conditional variance  $V(y|X_i = x_i^*)$  may be an approximation of relative importance of  $X_i$  in a way that the effect of  $X_i$  increases with decrease in the  $V(y|X_i = x_i^*)$ . However, output sensitivity is affected by  $x_i^*$ .

In the event of calculating the mean of variance on the possible  $x_i^*$ , the dependence effect to the  $x_i^*$  will be removed. Conditional variance on the different  $x_i^*$  is as follows:

$$E_{X_i}(V(y|X_i)) \tag{11}$$

If difference between conditional variance mean on the all possible  $x_i^*$  and unconditional variance is equal to the output mean variance on the different values of  $x_i^*$ .

$$V(y) - E_{X_i}(V(y|X_i)) = V_{X_i}(E(y|X_i)) \tag{12}$$

When  $X_i$  is an important parameter for model, the value of  $E_{X_i}(V(y|X_i))$  and  $V_{X_i}(E(y|X_i))$  will be small and large, respectively. Using the above equation, we may have:

$$V_{X_i}(E(y|X_i)) \leq V(y) \tag{13}$$

Variance  $V_{X_i}(E(y|X_i))$  is called the first order effect of  $X_i$  on the output. Thus, sensitivity index of first order of  $X_i$  parameter on the output  $y$  is expressed as follows:

$$S_i = \frac{V_{X_i}(E(y|X_i))}{V(y)} \tag{14}$$

where  $S_i$  is a number between zero and one. The proximity of  $S_i$  to zero indicates the more importance of mentioned parameter.

### 5.1 Output Derivative of Multi Layer Perceptron Neural Network in Relation to Input

Before, applying required equations to determine the output relative derivative in relation to inputs, the parameters used are introduced:

- $O_k$ : output of  $k$ th neuron in the output layer
- $x_i$ :  $i$ th input to network
- $w_{ik}^j$ : weight of connections between  $i$ th neuron in the  $j$ th layer and  $k$ th neuron in  $(j + 1)$  th layer
- $b_i^j$ : bias of  $i$ th neuron in the  $j$ th layer
- $net_i^j$ : weighted sum of input to  $i$ th neuron in the layer  $j$ th

$h_i^j$ : output of  $i$ th neuron in the  $j$ th layer

It is evident that  $h_i^j = f(net_i^j)$ , where  $f()$  is activity function of neuron at  $j$ th layer.

If the number of network hidden layer and neurons in the output layer is considered to be one and also the number of hidden layer and input to be  $m$  and  $n$ , respectively, the output derivative in relation to input  $x_i$  may be determined using the following equations:

The output of second layer, output layer with one neuron and linear transfer function, may be described by the following equation:

$$\begin{aligned} O_1 &= f(net_1^2) = net_1^2 = b_1^2 + \sum_{j=1}^m w_{j1}^1 h_j^1 \\ &= b_1^2 + \sum_{j=1}^m w_{j1}^1 f(net_j^1) \\ &= b_1^2 + \sum_{j=1}^m w_{j1}^1 f\left(b_j^1 + \sum_i^n w_{ij}^0 x_i\right) \end{aligned} \tag{15}$$

Since the transfer function of hidden layer is in the form of tangent hyperbolic, therefore equation (15) is written as follows:

$$O_1 = b_1^2 + \sum_{j=1}^m \left\{ w_{j1}^1 \cdot \text{tansig}\left(b_j^1 + \sum_i^n w_{ij}^0 x_i\right) \right\} \tag{16}$$

The derivative of the only network output,  $O_1$ , in relation to each of input,  $x_i$ , can be determined using chain rule.

$$\frac{\partial O_1}{\partial x_i} = \frac{\partial O_1}{\partial h_j^1} \frac{\partial h_j^1}{\partial net_j^1} \frac{\partial net_j^1}{\partial x_i} \tag{17}$$

$$\frac{\partial O_1}{\partial h_j^1} = w_{j1}^1 \tag{18}$$

$$\frac{\partial h_j^1}{\partial net_j^1} = \frac{\partial \text{tansig}(net_j^1)}{\partial net_j^1} = 1 - (h_j^1)^2 \tag{19}$$

$$\frac{\partial net_j^1}{\partial x_i} = w_{ij}^0 \tag{20}$$

Equation 17 defines the relative derivative in a model based the multi layer perceptron network as output absolute sensitivity to independent input. In effect, it shows the expected changes in the output in relation to change unit, while other variables are assumed to be constant Lu et al. (2001). Since each

input has different units and changes range in the real problems, absolute sensitivity may not be proper to compare the importance of input and output variables. Therefore, the values of absolute sensitivity should be modified based on the change range of input so as to obtain a proper parameter to compare the effects of inputs on the output. Lu et al. (2001) defined the output relative sensitivity to input by product of sensitivity values in a proportion of changes range (10 % of changes range) Lu et al. (2001). Thus, the output relative sensitivity of the behavioral model based on the independent input is:

$$\left(\frac{\partial O_i}{\partial X_p}\right)_r = \frac{\text{Max}(X_p) - \text{Min}(X_p)}{n} \left(\frac{\partial O_i}{\partial X_p}\right) \quad (21)$$

where  $n$  is the coefficient of change range (for example 10 for 10 % of changes range).

Since the weights are constant after training and according to equations (17–20), the output relative derivative in relation to each input is function of the values of networks output at that moment and the values of sensitivity at the previous moment. Thus, we may have:

$$\left(\frac{\partial O_i}{\partial X_p}\right) = F\left(\text{Inputs}, \frac{\partial O_i}{\partial X_p}\right) \quad (22)$$

In this study, the sensitivity analysis has been conducted to the five input parameters including normal specific weight, water content, test depth, the number of standard penetration test and grains maximum size. In this analysis, 200 points located in the five-dimensional space of input parameters correspondent with normal distribution are selected by Simlab 3.0 Software.

For these points which have its own input values, the values of derivative of the volume change at the stage of prediction and pressuremeter module at the stage of interpretation are calculated for optimal structure. The statistical characteristics of these values are presented at Tables 7 and 8.

According to what mentioned above and also variability of the sensitivity values, a method is required for sensitivity analysis that simultaneously shows dispersion and the probability of different values of output sensitivity to input. And as such, statistical method of relative sensitivity proposed by Loo et al. (2001) has been used in the current paper. In this method, five statistical percents (D10, D25, D50,

**Table 7** Statistical characteristics of absolute sensitivity of volume change to the inputs

Input	Max.	Min.	Mean.	Standard deviation
$\gamma$	4.045	0.361	0.950	0.535
W	3.201	-1.218	1.482	0.984
H	2.068	-11.78	-1.593	1.188
$N_{\text{SPT}}$	6.422	-17.34	-1.427	1.387
$D_{\text{max}}$	2.584	-1.925	-0.659	0.620

**Table 8** Statistical characteristics of absolute sensitivity of pressuremeter module to the inputs

Input	Max.	Min.	Mean.	Standard deviation
$\gamma$	0.555	-0.482	0.052	0.257
W	0.166	-1.294	-0.360	0.337
H	0.452	0.246	0.384	0.057
$N_{\text{SPT}}$	1.795	1.104	1.585	0.160
$D_{\text{max}}$	0.256	-0.416	-0.034	0.200

D75 and D90) of the values of output relative sensitivity are determined to input. This method is capable to evaluate the effects of increase and/or decrease of each input on the output and to determine general trend governing throughout total input space based on the random samples. The explanations of obtained results are as follows Lu et al. (2005):

D10: represents a value for the relative sensitivity that 90 and 10 %s of the values are larger and smaller than it, respectively. Therefore, the positive value implies that the probability of positive value for relative sensitivity is more than 90 %. In other words, the probability of increasing the output by increase in the input is more than 90 %.

D10: shows a value for the relative sensitivity that 90 and 10 %s of the values are smaller and larger than it, respectively. Therefore, the negative value indicates that the probability of decreasing the output by increase in the input is more than 90 %.

The explanation of D25 and D75 is similar to D10 and D90.

D50: If this value is on the zero line, it will show that the probability of decreasing or increasing the output by increase in the input is 50 %.

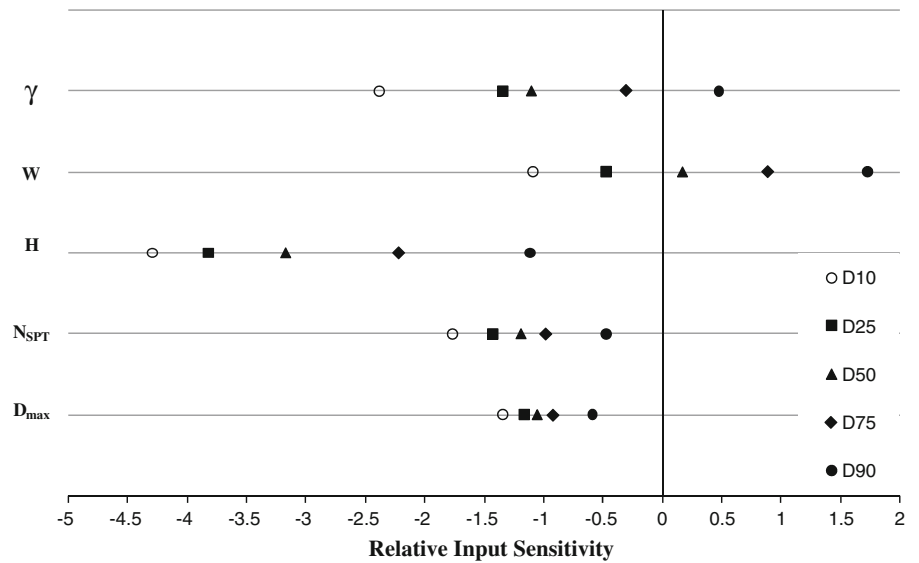
According to distance of statistical class from zero line and the values of the statistical percent, the impact factor of each variable on the output can be compared.

The mean values of relative sensitivity of the volume change and pressuremeter factor to inputs are discussed at Table 9.

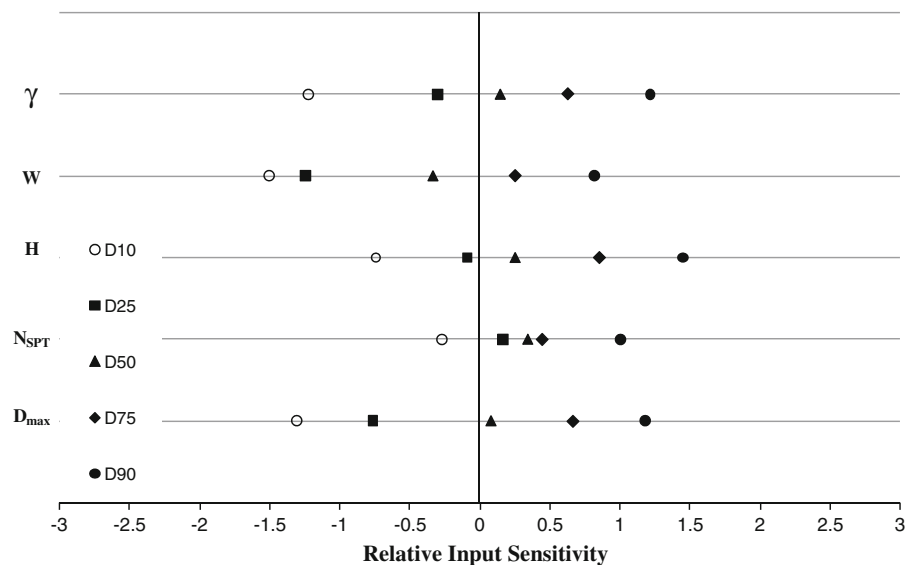
**Table 9** Relative sensitivity of the volume change and pressuremeter module to inputs

Output	$\Delta V$					$E_p$				
	$\gamma$	W	H	$N_{SPT}$	$D_{max}$	$\gamma$	W	H	$N_{SPT}$	$D_{max}$
Relative mean	-0.928	0.246	-2.920	-1.165	-1.015	0.0957	-0.399	0.346	0.341	-0.025

**Fig. 9** The sensitivity analysis of neural network at the prediction stage



**Fig. 10** The sensitivity analysis of neural network at the interpretation stage



In this section, the values of statistical percent relative sensitivity of the volume change for five inputs of MLP2 network with 15 neurons in the hidden layer are presented at Fig. 9. As can be seen clearly, more than 75 % of the values of relative sensitivity for normal specific weight ( $\gamma_m$ ) are negative. This indicates the reduction in the volume change due to the increase of normal specific weight.

Comparison the effects of five input parameters on the volume change revealed that based on the distance of statistical class from zero line, the values statistical percent (Fig. 9) and the values of relative mean (Table 9), test depth (H) has the most effect.

In this section, the values of statistical percent relative sensitivity of the pressuremeter module for five inputs of MLP2 network with 15 neurons in the hidden layer are shown at Fig. 10.

Comparison the effects of five input parameters on the pressuremeter module showed that based on distance of statistical class from zero line, values statistical percent (Fig. 10) and the values of relative mean (Table 9), test depth (H) and the number of standard penetration test ( $N_{SPT}$ ) have the most effect. However, due to positive values more than 90 % of the number of standard penetration test ( $N_{SPT}$ ) in the relative sensitivity, this parameter is selected as the most effective input parameters on the pressuremeter module.

## 6 Conclusion

In this paper, three types of artificial neural network (ANN) are employed to prediction and interpretation of pressuremeter test results. First, multi layer perceptron neural network, one of the most applicable neural networks, is used. Then, neuro-fuzzy network, combination of neural-phase network is employed and finally radial basis function, a successful network in solving nonlinear problems, is applied.

Of all neural network models, multi layer perceptron neural network proved to be the most effective. However, other applied networks have shown favorable performance. Finally, different models have been compared and network with the most outstanding performance in two stages is determined. In order to evaluate the network in the prediction stage, pressure–volume change charts resulting from simulation of optimal structure of each model have been compared with empirical results. Also in for the purpose of

assessment the capability of model generalization, the performance of mentioned network against unexperienced data has been compared with empirical results.

Contrary to conventional behavioral models, models based neural network do not demonstrate the effect of input parameters on output parameters. This research is response to this need through conducting sensitivity analysis on the optimal structure of proposed models. Also, derivation of governing equation for neural network model give more assurance to user to employ such models and consequently this facilitates the application of models in the engineering practices. These general rules were compared with those inferred from engineering experience as a second way of verification for the model. A consistent response was observed in terms of the dominant trends governing the model and the model output when compared to the available data. The insight into the behavior of ANN constitutive models by the approach presented here gives the user more confidence in model predictability and hence facilitates the incorporation of such models into engineering practice.

**Acknowledgments** The authors wish to thank the reviewers for their useful comments and suggestions.

## References

- American Society for Testing and Materials. Annual book of ASTM standards. D04-08, ASTM International Pub
- Banimahd M, Yasrobi SS, Woodward PK (2005) Artificial neural network for stress–strain behavior of sandy soils: knowledge based verification. *Comput Geotech* 32:377–386
- Ghaboussi J, Sidarta DE (1998) New nested adaptive neural networks (NANN) for constitutive modeling. *Comput Geotech* 22:29–52
- Hagan MT, Menhaj MB (1994) Training feedforward networks with the marquardt algorithm. *IEEE Trans Neural Netw* 5(6):989–992
- Hagan MT, Demuth HB, Beale MH (1996) *Neural network design*. PWS Publishing, Boston, MA
- Kanungo DP, Arora MK, Sarkar S, Gupta RP (2006) A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. *Eng Geol* 85:347–366
- Lee SJ, Lee SR, Kim YS (2003) An approach to estimate unsaturated shear strength using artificial neural network and hyperbolic formulation. *Comput Geotech* 30:489–503
- Lu M, Abourizk SM, Hermann UH (2001) Sensitivity analysis of neural networks in spool fabrication productivity studies. *J Comp Civil Eng* 15(4):299–308

- Moosavi M, Yazdanpanah MJ, Doostmohammadi R (2006) Modeling the cyclic swelling pressure of mudrock using artificial neural networks. *Eng Geol* 87:178–194
- Nabney T (1999) Efficient training of RBF networks for classification. *Proc 9th ICANN* 1:210–215
- Shahin MA, Jaksa MB (2005) Neural network prediction of pullout capacity of marquee ground anchors. *Comput Geotech* 32:153–163
- Shahin MA, Jaksa MB, Maier HR (2001) Department of Civil and Environmental Engineering, Adelaide University. Artificial neural network applications in geotechnical engineering