Middle-East Journal of Scientific Research 11 (4): 517-524, 2012
ISSN 1990-9233

Cyclic Swelling Estimation of Mudstone Using Adaptive Network-Based Fuzzy Inference System

Ramin Doostmohammadi

Mining Engineering Department, University of Zanjan, Zanjan, Iran

Abstract: Prediction of time dependent swelling behavior of weak rock in wetting and drying cycles is quite important for investigating the life time of structures in contact with these layers in arid and semi arid conditions. Previously proposed methods use information of previous cycles to predict the swelling behavior in the current cycle. This paper presents a new method which uses the information of the first wetting cycle in combination with Adaptive Network-Based Fuzzy Inference System (ANFIS) to predict the cyclic swelling pressure of mudstone. This facilitates the swelling pressure prediction in some wetting periods ahead. The data from laboratory tests on mudstone samples were used to calibrate and test the model. The ANFIS proves to be more effective in modeling the cyclic swelling pressure as opposed to the artificial neural networks (ANN) and multiple regression approach (MRA).

Key words: Adaptive network-based fuzzy inference system · Artificial neural network · Cyclic swelling behavior · Cyclic wetting and drying

INTRODUCTION

Swelling prediction of weak rocks is an essential criterion in the design of structures resting against such rocks. It is more complicated when the weak rock experiences wetting and drying periods which is pronounced in arid climates. The swelling strain and pressure of weak rock is increased in cycles gradually as reported by Pejon and Zuquette, Moosavi *et al*. and Doostmohammadi *et al*. [1-4]. Over years, consistent and accurate prediction of cyclic swelling behavior has not been achieved by using variety of methods ranging from MRA to ANN methods [5].

Huang *et al*. used regression approach to predict the maximum swelling behavior of shale in the second wetting period [6]. The maximum swelling strain of mudstone in various cycles was modeled by Pejon and Zuquette [1]. They fit a straight line to maximum axial swelling strain versus number of cycles. Moosavi *et al*. modeled the cyclic swelling behavior of mudstone using artificial neural networks [2]. In this model, the duration of a swelling cycle is divided to twelve equal points. The swell pressure of a desired point is predicted according to its four previous points. In a recent research, the time dependent swelling pressure of mudstone was predicted by Doostmohammadi *et al*. [5]. They used information of 3 previous cycles to predict the swelling pressure of the current cycle using ANN. In all the research, the swelling behavior of swallable rock is estimated using the information of instantaneous previous cycles and predictions of more than 1 new cycle is impossible. The first objective of the present study is to predict the swelling behavior of mudstone in some future cycles without considering the information of instantaneous previous cycles.

ANFIS is used to predict results in some engineering problems. However, its application to the geotechnical related studies are very limited. For instance, Ni *et al*. used the ANFIS for evaluating slope failure potential [7]. The mapping of cone penetration test (CPT) values into soil dynamic properties was performed by Romo and Garcia [8]. Shahin *et al*. predicted the settlement of shallow foundations on granular soils [9]. ANFIS as a new and powerful method to predict the cyclic swelling behavior of mudstone is introduced as a second objective of the current study.

Development of a Model for Cyclic Swelling Pressure: One of the most important steps in the model development for estimation of cyclic swelling behavior of weak rocks is identification of input parameters. The time, in which swelling pressure is determined, is introduced as

Corresponding Author: Ramin Doostmohammadi, Mining Engineering Department, University of Zanjan, Zanjan, Iran.
Tel: +98-241-5154015, Fax: +98-241-2283204.

the first input variable. Clay content, bulk density, specific gravity, dry density, texture, structure, initial moisture content, initial void ratio, porosity, degree of saturation, blue of methylene value, cataion exchange capacity, clay activity index, carbonate content and mercury porosity, as influencing parameters for swelling potential characterization, are required for introducing a reliable model. Determining all of these effective parameters is a difficult task and requires a lot of time and laboratory facilities. In the current research, the time dependent swelling pressure of mudstone in the first cycle (SP_{ini}) is used as a representative of all mentioned parameters in the second input location. The last input parameter, number of wetting periods (CN), is considered to extract the dynamic variation of swelling in cycles. The output parameter is time dependent swelling pressure (SP_f). These input/output arrangement improves the previous methods [2, 5] and allows determining the cyclic swelling pressure of mudstone without considering the information of former cycles.

Concept of ANFIS: The fuzzy logic approach is based on the linguistic of uncertain expression rather than numerical uncertainty. Since Zadeh proposed the fuzzy logic approach to describe complicated systems [10], it has since become popular and has been successfully applied in various engineering problems. Nonetheless, the main problem with this approach is that there is no systematic procedure for design of a fuzzy controller.

Basically a fuzzy inference system (FIS) is composed of five functional blocks (Figure 1).

- A *rulebase* containing a number of fuzzy if-then rules;
- A *database* which defines the membership functions of the fuzzy sets used in the fuzzy rules;
- A *decision-making unit* which performs the inference operation on the rules;
- A *fuzzification inference* which transforms the crisp inputs into degrees of match with linguistic values;
- A *defuzzification interface* which transforms the fuzzy results of the inference into a crisp output.

FIS implements a nonlinear mapping from its input space to the output space. This mapping is accomplished by a number of fuzzy if–then rules, each of which describes the local behavior of the mapping. The parameters of the if–then rules (referred to as antecedents or premises in fuzzy modeling) define a fuzzy region of the input space and the output parameters

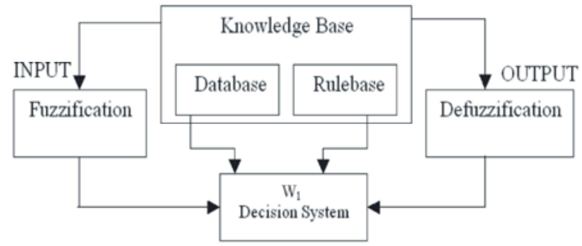

Fig. 1: The general structure of the fuzzy Inference System.

(also consequents in fuzzy modeling) specify the corresponding output. Hence, the efficiency of the FIS depends on the estimated parameters. The rule structure of a FIS makes it possible to incorporate human expertise about the system being modeled directly into the process to decide on the relevant inputs, the number of membership functions (MFs) for each input, etc. and the corresponding numerical data for parameter estimation.

Jang introduced a novel architecture and learning procedure for the FIS that uses a neural network learning algorithm for constructing a set of fuzzy if–then rules with appropriate membership functions from the stipulated input–output pairs [11]. This procedure of developing a FIS using the framework of adaptive neural networks is called an adaptive network-based fuzzy inference system.

ANFIS Architecture: ANFIS is a Sugeno-type FIS. The general structure of the ANFIS is presented in Figure 2.

It is assumed that the FIS has two inputs x and y and one output z. suppose that the rule base contains two fuzzy if-then rules of Takagi and sugeno's type [12]:

Rule 1: If x is A_1 and y is B_1 Then $f_1=p_1.x+q_1.y+r_1$ (1)
Rule 2: If x is A_2 and y is B_2 Then $f_2=p_2.x+q_2.y+r_2$ (2)

Where A_1, A_2 and B_1, B_2 are the membership functions for inputs x and y, respectively; p_1, q_1, r_1 and p_2, q_2, r_2 are the parameters of the output function. Figure 2(a) illustrates the fuzzy reasoning mechanism for this Sugeno model to derive an output function (f) from a given input vector [x,y]. The corresponding equivalent ANFIS architecture is presented in Figure 2(b), where nodes of the same layer have similar functions. The functioning of the ANFIS is as follows:

Layer 1: Each node in this layer generates membership grades of an input variable. The node output OP_i^1 is defined by

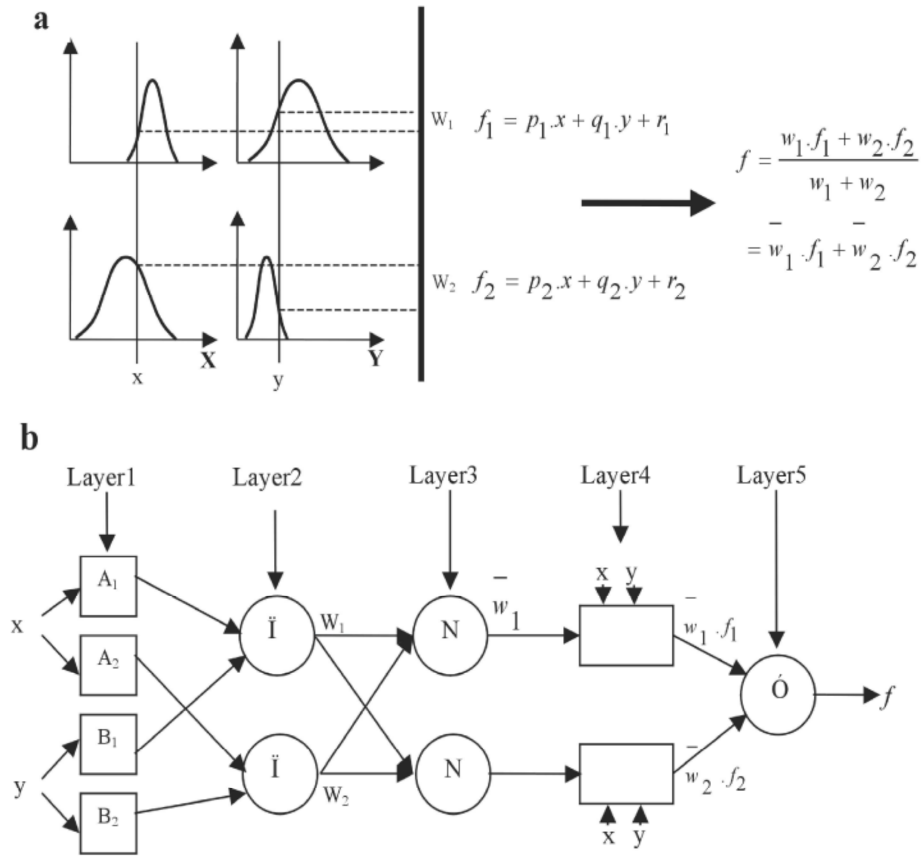

Fig. 2: Schematic of fuzzy and neurofuzzy paradigm: (a) fuzzy inference system and (b) equivalent ANFIS architecture

$$Op_i^1 = \mu_{Ai} \quad \text{for i=1, 2 or} \tag{3}$$
$$Op_i^1 = \mu_{A(i-2)} \quad \text{for i=3, 4} \tag{4}$$

Where x (or y) is the input to the node; A_i (or B_{i-2}) is a fuzzy set associated with this node, characterized by the shape of the MFs in this node and can be any appropriate functions that are continuous and piecewise differentiable such as Gaussian, generalized bell, trapezoidal and triangular shaped functions. Assuming a generalized bell function as the MF, The output Op_i^1 can be computed as:

$$OP_i^1 = \mu_{A_i}(x) = \frac{1}{1 + \left|\frac{x - c_i}{a_i}\right|^{2b_i}} \tag{5}$$

Where $\{a_i, b_i, c_i\}$ is the parameter set that changes the shape of the membership function with maximum equal to 1 and minimum equal to 0.

Layer 2: Every node in this layer multiplies the incoming signals, denoted as Π and the output Op_i^2 that represents the firing strength of a rule is computed as:

$$Op_i^2 = w_i = \mu_{A_i}(x).\ \mu_{B_i}(y) \quad i\text{=1, 2.} \tag{6}$$

Layer 3: The i^{th} node of this layer, labeled as N, computes the normalized firing strengths as:

$$OP_i^3 = \overline{w_i} = \frac{w_i}{w_1 + w_2} \quad i\text{=1,2.} \tag{7}$$

Layer 4: Node i in this layer computes the contribution of the i^{th} rule towards the model output, with the following node function:

$$OP_i^4 = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i) \tag{8}$$

Where $\overline{w_i}$ is the output of layer 3 and $\{p_i, q_i, r_i\}$ is the parameter set.

Layer 5: The single node in this layer computes the overall output of the ANFIS as:

$$OP_i^5 = Overall\_output = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{9}$$

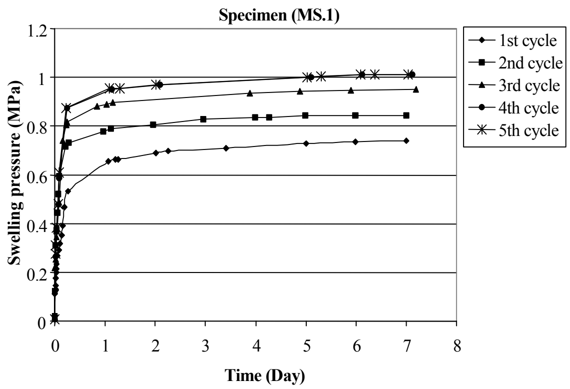

Fig. 3: Swelling pressure of specimen MS.1 in different cycles

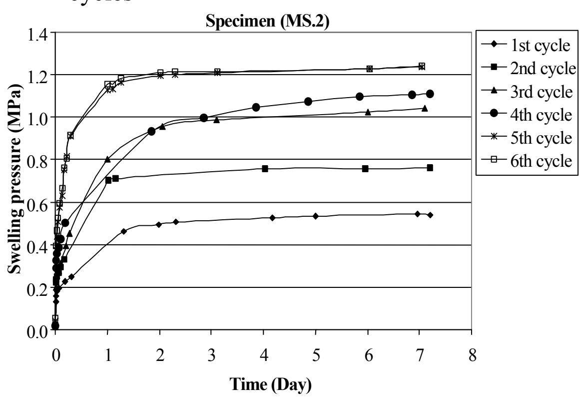

Fig. 4: Swelling pressure of specimen MS.2 in different cycles.

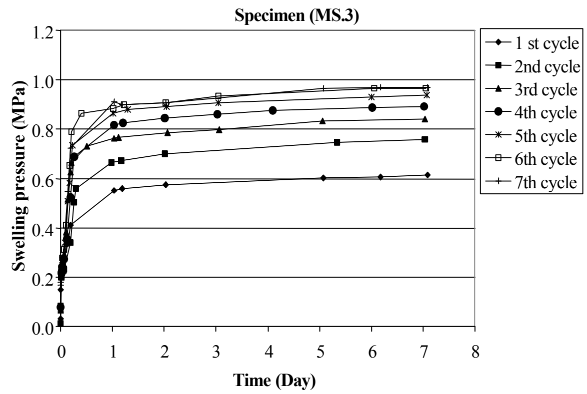

Fig. 5: Swelling pressure of specimen MS.3in different cycles.

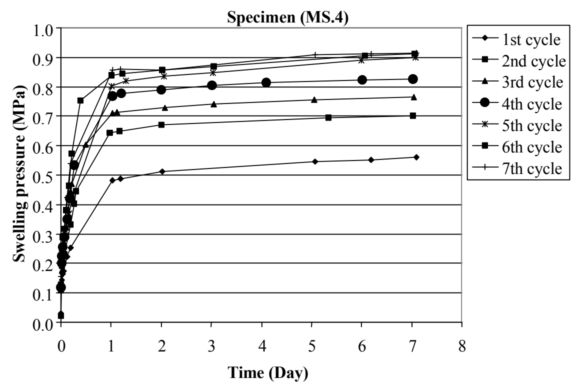

Fig. 6: Swelling pressure of specimen MS.4 in different cycles.

Estimation of Parameters: The parameters for optimization in an ANFIS are the premise parameters $\{a_i, b_i, c_i\}$, which describe the shape of the MFs and the consequent parameters $\{p_i, q_i, r_i\}$, which describe the overall output of the system. The basic learning rule of an adaptive network, the back propagation algorithm [13], which is based on the gradient descent rule, can be successfully applied to estimate these parameters. However, Jang argues that the gradient descent method is generally slow and is likely to get trapped in local minima [11]. Jang has proposed a faster learning algorithm, which combines the gradient descent method and the least squares estimate to identify parameters. A detailed description of the method can be found in Jang and Sun [14].

The Data Used for Model Development: The cyclic swell pressure tests in oedometric condition were performed to develop and assess the ANFIS model. The results of these tests on 4 mudstone samples taken across the Masjed-Soleiman Underground Hydro Electric Power Plant are shown in Figures 3, 4, 5 and 6.

Total data base includes 325 data points (Figures 3, 4, 5 and 6) are used for training (218 data points), checking (55 data points) and testing (52 data points) the ANFIS model. The training, checking and testing data points include 67%, 17% and 16% of the total data points are selected randomly from the total data.

The basic idea behind using a checking data set for model validation is that after a certain point in the training, the model begins overfitting the training data set. In principle, the model error for the checking data set tends to decrease as the training takes place up to the point that overfitting begins and then the model error for the checking data suddenly increases.

ANFIS Model Development: The ANFIS model for time dependent cyclic swelling pressure prediction is developed following the procedure described in Section 3. The FIS used in developing the ANFIS model, can be viewed as a partition in the multidimensional feature space, where the number of partitions in each dimension corresponds to the number of fuzzy sets and the corresponding

Table 1: Linguistic names of input variables

Input variables	Linguistic names
T	Low and high
SPini	Low and high
CN	Low and high

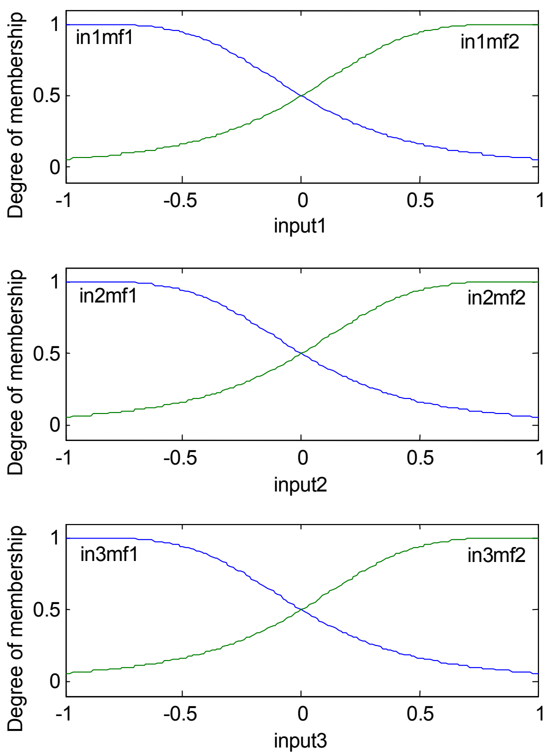

Fig. 7: The membership functions before training.

membership function that is defined in that dimension. Consequently, the input space partitioning plays a major role in the optimal architecture of the model. The number of MFs associated with each input variable is fixed by trial and error. In the present study, two triangular membership functions have been assigned to each input variable. Each input variable is classified into two fuzzy categories with linguistic attributes as given in Table 1. The initial values of the premise parameters are set in such a way that the centers of the MF are equally spaced along the range of each input variable. The model structure is implemented using the fuzzy logic toolbox of MATLAB software package.

The hybrid algorithm used in the present study for optimizing the parameters allows fast identification of parameters and substantially reduces the time needed to reach convergence. The minimum checking error is used as the stopping criterion to avoid overfitting. The ANFIS model has 50 parameters (32 linear and 18 nonlinear) and 8 fuzzy rules.

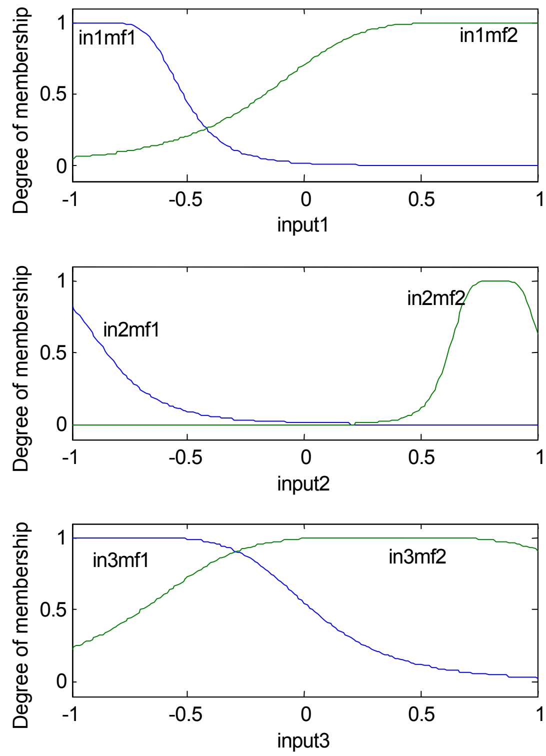

Fig. 8: The membership functions after training.

Performance Analysis and Testing the ANFIS Model: Figures 7 and 8 show the initial and final MFs before and after training, from which it can be seen that significant modifications have been done to the shapes of initial MFs through the learning process.

Comparison with the Real Swell Pressure Data: To test the ANFIS model, a matrix was used involving 4 swelling pressure signals in different wetting periods, randomly. The total test data points are 52 input/output pairs. These samples were not used in the training stage of the ANFIS model. Figures 9, 10, 11 and 12 show the real signal of swelling pressure and predicted amounts by ANFIS in different wetting periods. It can be observed that there is a good correspondence between them and the ANFIS is capable of modeling the swelling behavior in cycles.

Statistical comparisons between predicted and monitored swelling pressures were also performed. The root-mean-squared error (RMSE) and correlation coefficient (Corr) between the predicted and the monitored pressures are presented in Table 2. The RMSE of the ANFIS model for all data did not exceed 0.0994, which proves an accurate prediction.

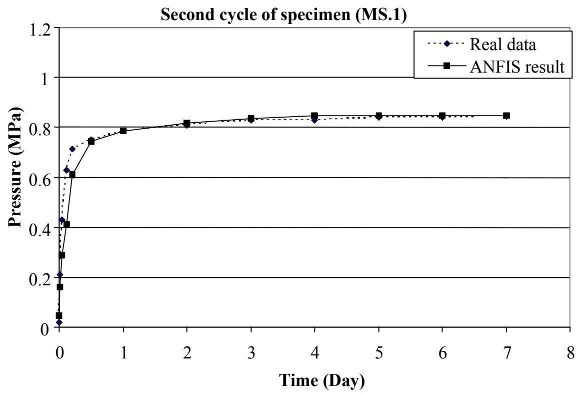

Fig. 9: The real and predicted trend of swelling behavior of specimen (MS.1) in the second cycle.

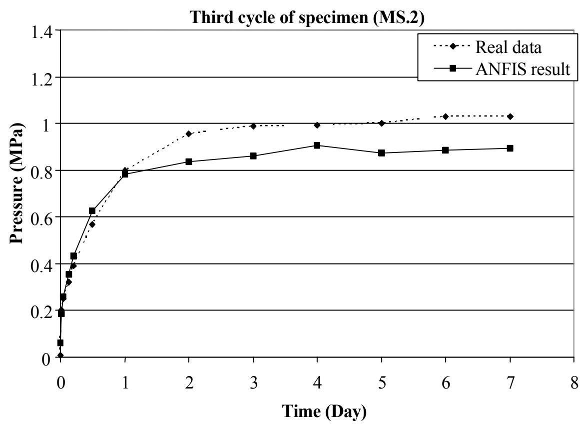

Fig. 10: The real and predicted trend of swelling behavior of specimen (MS.2) in the third cycle.

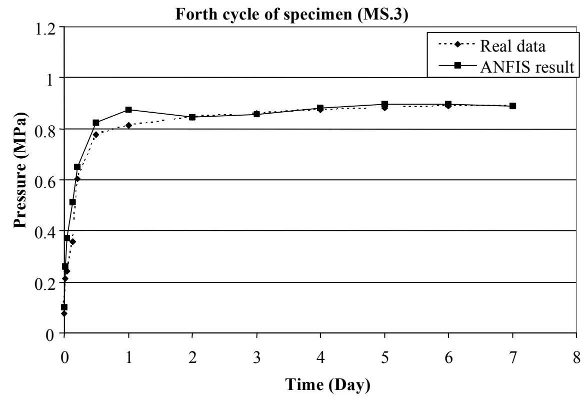

Fig. 11: The real and predicted trend of swelling behavior of specimen (MS.3) in the forth cycle.

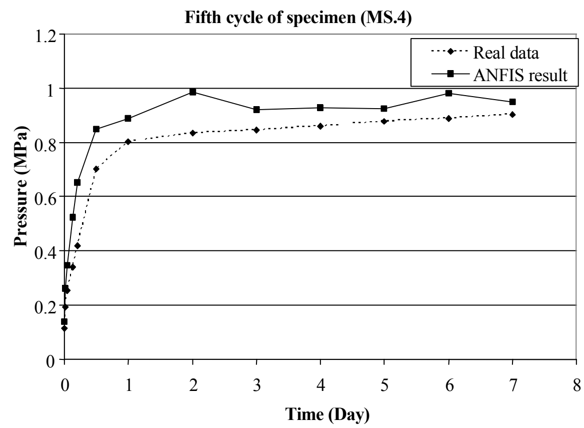

Fig. 12: The real and predicted trend of swelling behavior of specimen (MS.4) in the fifth cycle.

Table 2: Root mean square error (RMSE) and Correlation coefficient (Corr) between the measured pressure and those predicted by the ANFIS, MRA and ANN.

		RMSE	Corr
ANFIS	Training data	0.0885	0.9850
	Testing data	0.1450	0.9574
	All data	0.0994	0.9797
ANN	Training data	0.1396	0.9624
	Testing data	0.1701	0.9411
	All data	0.1455	0.9583
MR	Training data	0.2085	0.9134
	Testing data	0.1999	0.9174
	All data	0.2068	0.9142

It can be observed that except for the third cycle in specimen MS.2 (Figure 10), the fitting and testing errors for all the other specimens are very small. It should not be expected that the ANFIS produce very good results for all the training and testing samples particularly in the case that there might be conflicting data in the training or testing datasets. Looking into the training dataset (figure 4), we find that the difference of the second and the third cycles is much bigger than that of the first and the second cycles. Pejon and Zuquette have reported that the biggest difference of swelling potential is occurred between the first and the second cycles [1]. In our case, a sudden dissagregation and then a stress relief or easy access of water to the sample in the third cycle is supposed to be the reason for the unexpected increase in swelling pressure.

Comparison with Traditional Approaches: As described in introduction, using MRA and ANN have been so far the common methods for modeling cyclic swelling behavior of swelling rocks. Consequently, the real swelling pressure data was also compared with the swelling pressure that is predicted by MRA and ANN.

Comparison with MRA: MRA is usually used to summarize data and to study relations between variables.

The RMSE between the MRA results and the recorded pressures are presented in table 2. It can be noted that ANFIS model provides significant improvements in modeling the cyclic swelling behavior over MRA.

Comparison with ANN: The feed-forward multilayer perceptron (MLP) is used in the present study. One hidden layer is used and two nodes are used in the hidden layer. A Levenberg-Marquardt training combined with a Bayesian regularization is used as a learning rule for modeling the cyclic swelling behavior of mudstone as proposed by Moosavi *et al*. [2] and Doostmohammadi *et al*. [5]. A detailed algorithm of training stage has been discussed by MacKay, Hagan and Menhaj and also Foresee and Hagan [15-17]. The sigmoidal transfer function is used for the hidden and output layers. The output variables need to be scaled so that they remain within the range of sigmoidal function (-1, +1). Since it is desirable to do input scaling to allow the connection weights to have the same order of magnitude, each input is scaled to the range of -1 and +1 [18, 19].

The performance capability of the ANN model was examined based on the RMSE and Corr indexes. The result is shown in table 2. This comparison shows that the ANFIS model is an effective way of modeling the cyclic swelling pressure of mudstone with more accuracy.

CONCLUSIONS

Modeling cyclic swelling behavior of swellable rocks is a challenging job facing geotechnical engineers because good mathematical models can save them a significant amount of cost and time. In this paper an ANFIS model for cyclic swelling pressure estimation of mudstone is developed. This ANFIS learns the if–then rules between time dependent swelling pressure in various wetting cycles and swelling behavior of first cycle. It has been observed that ANFIS outperforms artificial neural networks, which has been found in the literature, to perform better than multiple regression models. In summary, ANFIS is a good choice and powerful tool for modeling cyclic swelling behavior of mudstone.

REFERENCES

1. Pejon, O.J. and L.V. Zuquette, 2002. Analysis of cyclic swelling of mudrocks. Engineering Geology, 67: 97-108.
2. Moosavi, M., M.J. Yazdanpanah and R. Doostmohammadi, 2006. Modeling the cyclic swelling pressure of mudrock using artificial neural network. Engineering Geology, 87: 178-194.
3. Doostmohammadi, R., M. Moosavi, T.H. Mutschler and C. Osan, 2007. Swelling pressure of mudstone under cyclic wetting and drying. 11th Congress of the International Society for Rock Mechanics, Lisbon, Portugal, pp: 443-446.
4. Doostmohammadi, R., M. Moosavi, T.H. Mutschler and C. Osan, 2009. Influence of Cyclic Wetting and Drying on Swelling Behavior of Mudstone in South West of Iran. Environmental Geology, 58: 999-1009.
5. Doostmohammadi, R., M. Moosavi and B.N. Araabi, 2008. A Model for Determining the Cyclic Swell-Shrink Behavior of Argillaceous Rock. Applied clay Science, 42: 81-89.
6. Huang, S.L., R.C. Speck and Z. Wang, 1995. The temperature effect on swelling of shales under cyclic wetting and drying. International J. Rock Mechanics and Mining Sciences and Geomechanics Abstracts, 32: 227-236.
7. Ni, S.H., P.C. Lu and C.H. Juang, 1996. A fuzzy neural network approach to evaluation of slope failure potential. Microcomputers in Civil Engineering, 11: 59-66.
8. Romo, M.P. and S.R. Garcia, 2003. Neurofuzzy mapping of CPT values into soil dynamic properties. Soil Dynamics and Earthquake Engineering, 23: 473-82.
9. Shahin, M.A., H.R. Maier and M.B. Jaksa, 2003. Settlement prediction of shallow foundations on granular soils using B-spline neurofuzzy models. Computers and Geotechnics, 30: 637-647.
10. Zadeh, L.A., 1965. Fuzzy sets. Information and Control, 8: 338-353.
11. Jang, J.S.R., 1993. ANFIS: Adaptive network based fuzzy inference system. IEEE Transactions on Systems Man Cybernetics, 23: 665-83.
12. Takagi, T. and M. Sugeno, 1985. Fuzzy identification of systems and its applications to modeling and control. Transactions on Systems Man Cybernetics, 15: 116-132.
13. Werbos, P., 1974. Beyond regression: new tools for prediction and analysis in the behavioral sciences. PhD thesis, Harvard University.
14. Jang, J.S.R. and C.T. Sun, 1995. Neurofuzzy modeling and control. The Proceedings of the IEEE, 83: 378-406.

15. MacKay, D.J.C., 1992. Bayesian interpolation, Neural Computation, 4: 415-447.
16. Hagan, M.T. and M.B. Menhaj, 1994. Training feedforward networks with the marquardt algorithm. IEEE Transactions on Neural Networks, 5: 989-993.
17. Foresee, F.D. and M.T. Hagan, 1997. Gauss-Newton approximation to bayesian regularization. Proceedings of International Joint Conference on Neural Networks., pp: 1930-1935.
18. Zurada, J.M., 1992. Introduction to artificial neural systems. West Publishing Company, Saint Paul, Minnesota.
19. Fausett, L., 1994. Fundamentals of neural networks. Prentice Hall International.