Cost and Time Management of Reservoir Characterization through the Advantages of Expert Systems: Implemented on Conventional and Modern Well Logging Data

Payam Parvasi[1], Javad Ghiasi-Freez[2], Amir Hatampour[1]
[1] Islamic Azad University of Dashtestan, Borazjan, Iran
[2] Iranian Central Oil Fields Company, National Iranian Oil Company (NIOC), Tehran, Iran

Abstract

Implementation of expert systems for solving engineering problems has led to a revolution in optimization and minimization of industrial projects. The present study integrates the advantages of genetic algorithm, as an evolutionary model, and neural networks to minimize the risk of being stuck in a local minimal point when searching for the best values of neural network model including weights and biases. This strategy was capable of significantly improving the accuracy of a neural network by optimizing network parameters. The optimized network is employed to find an applicable relationship between a set of conventional well logs, which are used for rock evaluation, and two outputs of nuclear magnetic resonance (NMR) log, named free fluid index and permeability. NMR outputs have a valuable application in reservoir characterization. Necessary training and testing data for the research are obtained from one of the carbonate reservoir rocks of Iran. Conventional well logs are the easily accessible tools compared to NMR log's parameters which is an expensive tool and unavailable for all wells of a hydrocarbon field; therefore, finding a reasonable relationship between conventional logs and NMR will help petroleum engineers to estimate NMR outputs in the case which are not available. The performance of optimized neural network model was verified through unseen test data. The results show that GA-based network satisfies the necessary accuracy needed for studying reservoir rock properties during rock modeling.

Keywords: *Well Logging, Reservoir Engineering, Free Fluid Porosity, Permeability, Intelligent Systems, Evolutionary Models.*

1. Introduction

T Nuclear magnetic resonance (NMR) log is one of the potent tools that made a revolution in measurement of reservoir properties that had been impossible to measure before by logging instrument. Permeability, free and bulk fluid of rock that previously measured through time consuming laboratory measurements can be identified by NMR log without the necessity of expensive operation of coring. However, conventional log sets, such as neutron, sonic, resistivity, and density logs, which have been run in all drilled wells, running a NMR log in all drilled wells of a field is a costly task and usually impossible due to economical limitations. NMR log is only run in limited number of drilled wells of field, which has more than dozen of wells. Also, it is impossible to run NMR log in wells that drilled some years ago and already have been cased. The present paper aims to employ the capabilities of supervised intelligent systems for solving these limitations and problems using conventional well logs. Since conventional well logs have been run in all wells of each field, developing a quantitative formulation between conventional well log data and outputs of NMR log can be a reasonable methodology to have a view from the responses of NMR in each well. Studying the previous literatures shows that intelligent models have been diversely used in prediction of reservoir properties [1-6]. The goal of the present research is not only introduction of an intelligent model for prediction of NMR log's outputs but also it tries to benefit from the advantage of genetic algorithm (GA) optimization model in designing the neural network model. The performance of the optimized network is studied in terms of prediction error and convergence time of optimization model.

2. NMR logging

The importance and application of Nuclear Magnetic Resonance (NMR), as a modern well logging equipment, is not only limited to reservoir evaluation, while it is widely used in physics, chemistry, biology, and medicine. Combining the permanent magnets and pulsed radio frequencies with concept of logging led to an important logging tool that today known as NMR log. An applicable instrument of NMR logging was introduced by taking the benefits of a medicine magnetic resonance imaging (MRI) [7]. The fundamental of all NMR logging tools is the same. Two time-based parameters of longitudinal relaxation time (T_1) and transverse relaxation time (T_2) are obtained from the spin-echo trains and include the most crucial outputs of NMR log that lead to further properties of the formation. T_1 and T_2 indicate the time in which protons relax longitudinally and transversely, respectively, related to the transmitted magnetic field. In fact, T_2 is the most important parameters that can be directly converted to porosity. The T_2 plot includes movable and immovable fluids of the rock, which are separated based on a cut-off value of T_2. Only fluids are visible for transmitted magnetic field on NMR tool and this fact redounds to one of the advantages of NMR log compared to conventional logs such as sonic, neutron, and bulk-density logs. NMR porosity is lithology-independent and does not need to be calibrated with lithology in different zone and intervals of a well. Three groups of invaluable information about reservoir condition can be obtained from NMR raw data, including pore size distribution of a formation, fluids properties of pore spaces and finally quantities of these fluids. Porosity and pore size distribution are theoretically related to permeability in a direct relationship. Therefore, permeability and movable fluids (free fluid index) of the formation can be estimated using aforementioned raw data.

3. Intelligent models

Nowadays, the capabilities, and beneficial performance of expert systems are obvious in the world of engineering. Various branches of science and engineering have employed intelligent models to decrease cost and time, increase efficiency and accuracy of measurements and practical operations. These reasons introduced intelligent models as potent and efficient methodologies for solving engineering problems and various researchers employed different types of expert systems, including neural networks and fuzzy models, and optimization algorithms such as genetic algorithm, particle swarm optimization, and ant colony optimization, in their studies [8-12].

3.1. Neural Network Model

Neural networks are computational non-linear algorithms used for image and signal processing, pattern recognition, classification issues, and etc. Up to now, different types of structures have been introduced for neural network models. Multilayer perceptron is one of these structures. Fig. 1 shows a perceptron with n inputs and a bias.

A multilayer perceptron shows the non-linear relationship between input vector and output vector, which is possible due to a connection between the neurons of each node in consecutive layers. The outputs of neurons are multiplied to the weight coefficient and used as inputs of the activation function. A multilayer perceptron uses training algorithms to predict the hidden relationship of inputs and output in the form of a non-linear function. The non-linear function includes a set of weighting coefficients and biases. Various training algorithms have been introduced among them back-propagation can be mentioned as the most important one. However, back-propagation algorithm suffers from different problems such as low speed of convergence during training step and also sticking in local minima, known as under training, and also memorizing the training data, known as over-training. Several methods have been introduced to overcome these weaknesses of back-propagation training algorithm [13]. Recently, the researchers suggested optimization algorithms for training the neural network models. In the current research GA is used and a brief introduction of these model is presented below.

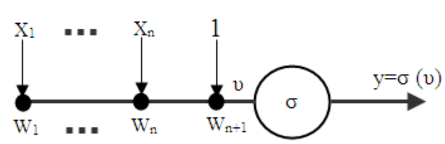

Fig. 1. A perceptron with n inputs and a bias

3.2. Genetic Algorithm

GAs were introduced by Holland in 1975 as an optimization model, which mimic their behavior from the natural evolution processes, including survival of the fittest, reproduction, crossover, and mutation. A simple *GA* cycle consists of three processes. These include selection, genetic operation, and replacement. The population comprises a group of chromosomes. The chromosomes are the candidates for the solution. The fitness evaluation unit determines a fitness value of all chromosomes. The better parent solutions are reproduced, and the next generation of the solutions (children) is generated by applying the crossover and mutation (genetic operators). The crossover operator chooses parents at random and produces children. The mutation operator randomly changes the values of the elements in a string. New generations of solutions are evaluated, and such a *GA* cycle repeats until a desired termination criterion is obtained [14]. Fig. 2 illustrates a schematic flowchart of GA procedure.

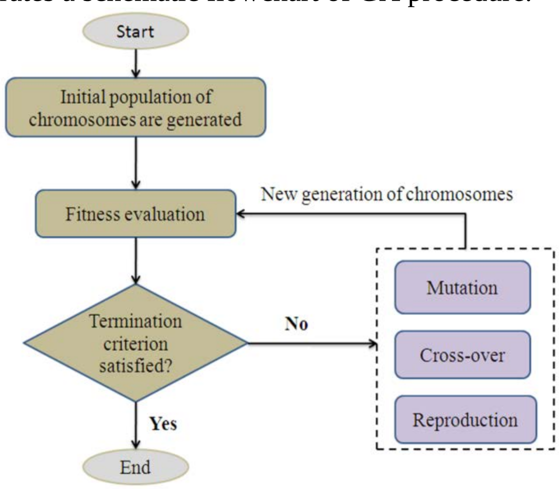

Fig. 2. Schematic diagram of a GA procedure 3. Tables, Figures and Equations

4. The optimized neural network model designing

Weights and biases of the neural network model should be optimized by GA optimization technique. The optimization of NN model by optimization techniques is discussed in terms of mathematical concepts below.

If the n^{th} layer of a NN model includes R numbers of inputs and M numbers of neuron, then the weights and biases matrixes of this network are as below.

$$W^n = \begin{bmatrix} (w_1^n)^T \\ (w_2^n)^T \\ \vdots \\ (w_m^n)^T \end{bmatrix} \qquad (1)$$

$$B^n = \begin{bmatrix} b_1^n \\ b_2^n \\ \vdots \\ b_m^n \end{bmatrix} \qquad (2)$$

Where $w_m^n = \left[w_{m,1}^n, w_{m,2}^n, \ldots, w_{m,R}^n\right]^T$ and it is vector of weights that connect the m^{th} neuron of M^{th} layer to the inputs of that layer. The vector of parameters of this layer can be shown as below.

$$X^n = \begin{bmatrix} w_1^n \\ \vdots \\ w_M^1 \\ \hline b_1^n \\ \vdots \\ b_m^n \end{bmatrix} \qquad (3)$$

For all layers, the matrixes of weights and biases are defined as the same. The matrixes of all layers form the entire parameters that should be optimized by an optimization algorithm. In simple word, for a network with L layer, the matrix of X variables obtains from below matrix.

$$X = \begin{bmatrix} X^1 \\ X^2 \\ \vdots \\ X^l \end{bmatrix} \qquad (4)$$

In fact, this matrix is the same with the vector of Eq.(1), while the optimized values of its variables is determined by the optimization technique. To get this goal, first N vectors of location vector (X_i) are randomly generated. The neural network is run so that its parameters (weights and biases) are set based on numerical values of optimization variables. The obtained error of the network is considered as the fitness value of variables vector of the network.

5. Developing the optimized network model

The optimized neural network models were trained using a training matrix of conventional well logs and desired responses of NMR log. The correspondence depth of logs and NMR responses were correlated to each other. A brief explanation of data processing is discussed below.

5.1. Selection of Appropriate Inputs

It is necessary to select appropriate well logs for obtaining most accurate intelligent models. The inputs should be sensitive to quantity and type of entrapped fluid in pore spaces. The available well log data were full-suite conventional porosity log of three wells. Each set includes neutron porosity (NPHI), bulk density (RHOB), sonic transit time (DT), effective porosity (PHIE), true formation resistivity (RT), photoelectric factor (PEF) and gamma ray log (GR). The corresponding NMR responses are also available for these wells. The cross plot analysis was used to select the most suitable inputs for designing an intelligent model. Table 1 represent the coefficient of determination between the conventional logs and two desired outputs of NMR log.

Table 1. Cross-plot analysis results (R^2) for free fluid porosity and permeability

Output / Input	FFP	Permeability
RT	0.091	0.023
PHIE-HILT	0.569	0.240
NPHI	0.461	0.317
PEF	0.004	0.026
RHOB	0.465	0.125
GR	0.053	0.048
DT	0.589	0.194

Inputs with higher R^2 value have deeper effect on the selected outputs; therefore the arrangement of four logs of neutron porosity (NPHI), bulk density (RHOB), sonic transit time (DT), effective porosity (PHIE) were selected as the most suitable inputs of optimized neural networks.
Selecting porosity logs as the most effective inputs of the optimized networks is compatible with petrophysical facts. Porosity logs, such as neutron, density and sonic log, obtain valuable information about the quantity of pore space, which is equivalent to quantity of fluid. Also, the difference of primary and secondary porosity is distinguishable by combination of sonic and neutron logs. Furthermore, PHIE is a good indicator of connected pore spaces; therefore, it is reasonable to see high correlation between the PHIE values and both permeability and free fluid index.

5.2. Preparation of Training Matrix

The available well log data of three wells were used to prepare the training matrix of the optimized network model. The matrix includes 6 columns and 796 rows. Number of rows is equal to number of data points that used. The first 4 rows of the matrix include the numerical value of employed conventional porosity logs, including effective porosity, sonic transmitted, neutron porosity, and bulk density logs. The two last columns are desired outputs of optimized NN model.
Normalization of data is an importance task when working with NN models. This task obviously improves the performance of the model. As shown in Table 2 the tolerance of input and output data is high and this fact affected the results of training procedure. In the present study, input and output data were normalized in the range of (-1, 1) and (0, 1), respectively, using the equations 5 and 6.

$$I_n = 2\left(\frac{I - I_{\min}}{I_{\max} - I_{\min}}\right) - 1$$

$$O_n = \frac{O - O_{\min}}{O_{\max} - O_{\min}} \qquad (6)$$

Where I_n and O_n is the normalized value of input and output variable, respectively; I and O denotes the actual value, I_{min}, O_{min}, I_{max} and O_{max} represent the minimum and maximum values of entire data set of the respective input or output, respectively. The input data were normalized in the range of (-1, 1) because the variations of TANSIG function occur in this range. Moreover, it reduces the search space of optimization algorithm. Table 2 represents the range of used data for training the optimized NN. In validation step, the test data should be normalized based on these ranges.

Table 2. The range of training data used for optimization of weights and biases.

DT	NPHI	RHOB	PHIE-HILT	FFP	K
45.38 to 78.92	0.008 to 0.23	2.09 to 2.94	0.001 to 0.296	0.001 to 0.136	0.001 to 177

5.3. Implementing the Optimized Network Model

A weak trained neural network model utilizing a good training matrix undoubtedly will fail in validation procedure. In addition, a weak training matrix will conclude to a network performing poor in test step. The combination of a good training matrix and accurate designing of network structure are two crucial parts of developing an intelligent model that guarantees the accurate performance of the model for future applications. In this study, we have tried to cover both of these facts. To design the optimized structure of the NN model, three optimization algorithms were used and the weights and biases of NN model were optimized.

All the necessary codes of this study were written in MATLAB program environment. First, the aforementioned algorithm of GA as well as NN were codified in two separate *m file*. The NN model is the objective function of optimization algorithms and its unknown variable, including weights and biases of neurons, are optimized using a prepared training matrix of conventional porosity logs and NMR responses.

Increasing the number of outputs in a network model causes some complexities and inaccuracies in model construction; therefore in the current research, two separated models were trained for prediction of FFP and K. The structure of the models is the same. The only difference is in number of output type and number of neurons in hidden layer. There are four inputs and one output for each model.

5.4. Extracting the Optimized Weights and Biases

The NN code was introduced to the optimization algorithm to extract its optimized weights and biases. Number of neurons in hidden layer of NN model was adjusted on 11. This value was obtained by a try-and-error procedure in which 25 network models were run and number of neurons was changed from 5 to 30. The mean square error (MSE) for training data was determined for each model. The results show that the minimal MSE is obtained if number of hidden neurons is 11.

After running the optimization algorithm, it takes the NN model as its objective function and tries to optimize it in a way in which the MSE is minimal. Here, the developed MSE function is mathematically introduced. The computational formulation of a three-layered neural network, which contains four nodes in input layer, one node in output layer, and i nodes in hidden layer, is as follow.

$$net_{i\times 1} = IW_{i\times 4} \times \begin{bmatrix} NPHI \\ RHOB \\ DT \\ PHIE \end{bmatrix} + Ib_{i\times 1} \quad (7)$$

$$out_{i\times 1} = \frac{2}{1+e^{-2net_{i\times 1}}} - 1 \quad (8)$$

$$O_{NN} = OW_{1\times i} \times out_{i\times 1} + b_{1\times 1} \quad (9)$$

Where, IW and OW are matrices of weight factors corresponding to hidden layer and output layer, respectively; Ib and b are matrices of bias values corresponding to hidden layer and output layer; and O_{NN} is estimated value of output. Equation 8 is TANSIG transfer function which is embedded in the nodes of hidden layer. In above formulation, conventional porosity well log data (NPHI, RHOB, DT, and PHIE) are used as inputs, while nuclear magnetic resonance log parameters including free fluid porosity and permeability are used as outputs. To extract the optimum values of weights and biases, following mean square error function was introduced into optimization algorithm.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(NMR_{response} - O_{NN})^2 \quad (10)$$

Number of data set in training matrix, initial population of swarms and number of iterations change the time of investigations of optimization algorithm; larger number of aforementioned variables increases the time. In the last step, the obtained values are substituted in the main formula of NN models, which had been written in Matlab Environment. It is worth mentioning that this process was done two times; one for permeability and one time for free fluid porosity.

6. Performance of GA-based neural network

Weights and biases of the neural network model were optimized by a genetic algorithm. To get the best results, the training data set was divided into two parts, named training and validation data point including 70 and 30 percent of data set, respectively. First, the model was trained with 70 percent of data and then its performance was evaluated by the remained 30 percent. It leads to the best performance of optimization algorithm. After finding the minimal MSE of the GA, the obtained weights were used for predicting FFP and permeability in testing data points. Fig. 3 illustrates the correlation coefficient of measured and predicted values of FFP and permeability in the test data. The results demonstrate reliable performance of the proposed model. The MSE of the proposed model for FFP and permeability is 3.75% and 4.3 md, respectively. These values are acceptable from the stand point of petroleum engineering in reservoir characterization.
The GA model takes 87.37 second in its best run to find the convergence point of weights and biases.

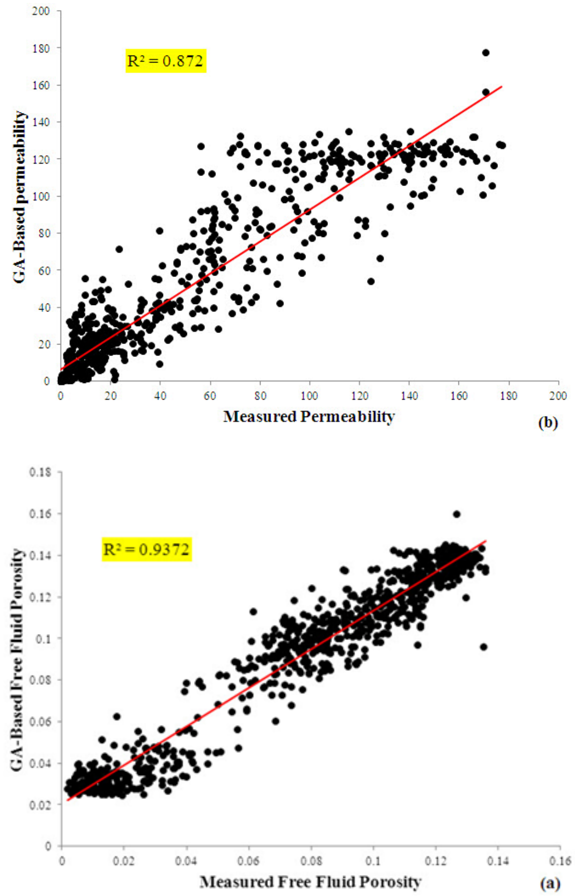

Fig. 3. Crossplots showing the correlation coefficients between measured and predicted results from GA-based neural network for (a) FFP and (b) permeability.

The obtained values of FFP and permeability in the test points were compared to actual values as shown in Fig. 5 and Fig. 6. The figures show that the proposed model was predicted the trend of increment or reduction in the data. The porous and permeable intervals can be easily distinguished even based on predicted values. This fact shows that the predicted values of FFP and permeability can be used in reservoir modeling with high degree of accuracy.

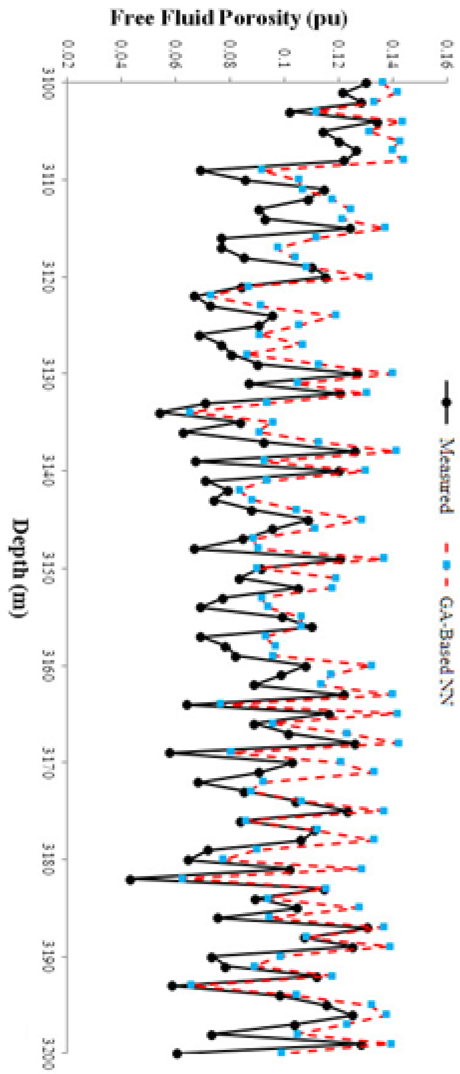

Fig. 4. A comparison between measured and GA-based predicted FFP versus depth in the test data.

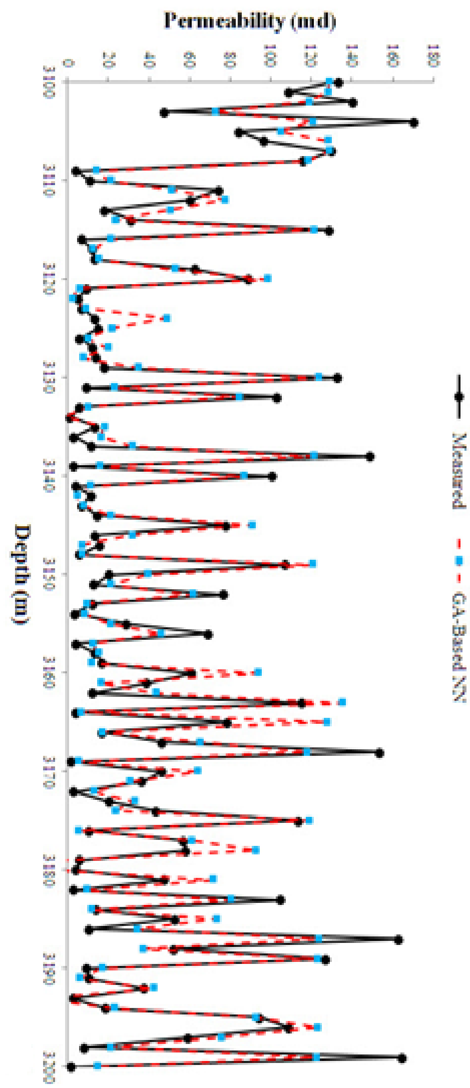

Fig. 5. A comparison between measured and GA-based predicted permeability versus depth in the test data

7. Conclusions

In the present research, GA optimization algorithm was used for optimizing the weights and biases of a three layered neural network model. The results show that GA optimization algorithm was performed reliable in finding the NN parameters. Importance of FFP and permeability as two crucial reservoir characteristics motivated the researchers to find the most accurate techniques; therefore, the performance and accuracy of optimization model was investigated in detail. Considering two features of accuracy and convergence time shows that GA-based model was an accurate model for optimization of NN parameters. A statistical analysis of predicted values of FFP and permeability uncovers that in all models most of predicted values of FFP are smaller than actual value, while in permeability prediction it is revers. This fact should be considered when the predicted values are employed for modeling tasks. In another word, the developed model based on predicted data will show the reservoir more porous and less permeable than its actual case. The test data are normalized to the range of (-1, 1) using the maximum and minimum data of training matrix before introducing to the optimized trained network and consequently the output of the network is conditioned back to real rang of data using maximum and minimum of training matrix for responses of NMR log. In other words, the estimated output is converted to real values by software before representing in the specific blank part of the interface.

Acknowledgments

The authors would like to appreciate Islamic Azad University of Dashtestan for providing the technical and financial supports of this research.

References

[1] S. D. Mohaghegh, "Virtual-intelligence applications in petroleum engineering: part 1- Artificial Neural Network" Journal of Petroleum Technology, Vol. 52, pp. 64-73, 2000.
[2] S. R. Ogilvie, S. Cuddy, C. Lindsay, A. Hurst, "Novel methods of permeability prediction from NMR tool data. Dialog" magazine of the London Petrophysical Society, pp. 15-22, 2002.
[3] C. H. Chen, Z. S. Lin, "A Committee Machine with Empirical Formulas for Permeability Prediction" Computers and Geosciences, Vol. 32, pp. 485–496, 2006.
[4] J. Ghiasi-Freez, A. Hatampour, "A Fuzzy Logic Model for Predicting Dipole Shear Sonic Imager Parameters From Conventional Well Logs" Petroleum Science and Technology, Vol. 31, pp. 2557–2568, 2013.
[5] M. H. Afshari, "Application of Ant Colony Algorithm to Pipe Network optimization" Iranian Journal of Science & Technology, Transaction B, Engineering, Vol. 31, pp 487-500, 2007.
[6] O. Velez-Langs, "Genetic algorithms in oil industry: An overview" Petroleum Science and Engineering, Vol. 47, pp. 15– 22, 2005.
[7] G. R. Coates, M. Miller, M. Gillen, C. Henderson, "The MRIL In Conoco 33-1 An Investigation Of A New Magnetic Resonance Imaging Log, 32nd Annual SPWLA Logging Symposium Transactions, p. 24. 1991.
[8] S. K. Nanda, D. P. Tripathy, S. K. Nayak, S. Mohapatra, "Prediction of Rainfall in India using Artificial Neural Network (ANN) Models" I.J. Intelligent Systems and Applications,Vol. 12, pp. 1-22, 2013.
[9] M. H. Afshari, "Application of Ant Colony Algorithm to Pipe Network optimization" Iranian Journal of Science & Technology, Transaction B, Engineering, Vol. 31, pp 487-500, 2007.
[10] A. Gupta, P. R. Sharma, "Application of GA for Optimal Location of FACTS Devices for Steady State Voltage Stability Enhancement of Power System" I.J. Intelligent Systems and Applications, Vol. 03, pp. 69-75, 2014.
[11] M. B. Gueye, A. Niang, S. Arnault, S. Thiria, M. Crépon, "Neural approach to inverting complex system: Application to ocean salinity profile estimation from surface parameters" Computers and Geosciences, Vol. 72, pp. 201-209, 2014.
[12] S. Mirzaei, A. Nikpey, P. Zarbakhsh, " Recognition of Control Chart Patterns Using Imperialist Competitive Algorithm and Fuzzy Rules Approach" I.J. Intelligent Systems and Applications, Vol. 10, pp. 67-76, 2014.
[13] S. Haykin, Neural networks: A comprehensive foundation. Englewood Cliffs, NJ: Prentice-Hall. 1991.
M. Saemi, M. Ahmadi, A. Y., Varjani, "Design of neural networks using genetic algorithm for the permeability estimation the reservoir" Petroleum Science and Engineering. Vol. 59, pp. 97–105, 2007.

Payam Parvasi is currently assistant professor at Azad University (Dashtestan Branch). He gained his PhD on Chemical engineering from Shiraz University. His MSc degree is on Chemical engineering from Shiraz University. His BSc degree is on Chemical engineering from Kashan University of Technology. Parvasi's research interests are in modeling of chemical processes.

Javad Ghiasi-Freez is a petroleum geologist working on the importance and implementation of intelligent systems in reservoir characterization and geoscience. He holds an MSc in petroleum engineering from Shahrood University of Technology, Iran. Up to now, he has presented several articles related to integrated intelligent models in international periodicals and conferences. He is an active referee of several ISI journals. In addition, application of Image analysis and pattern recognition in geosciences is his interest.

Amir Hatampour holds a BS degree from the Petroleum University of Technology and a MS degree from Amirkabir University of Technology (Tehran Polytechnic), both in petroleum exploration engineering and is pursuing a PhD on intelligent reservoir characterization.