A Committee Neural Network for Prediction of Normalized Oil Content from Well Log Data: an Example from South Pars Gas Field, Persian Gulf

Ali Kadkhodaie-Ilkhchi[1]; M.Reza Rezaee[1,2*]; Hossain Rahimpour-Bonab[1]

1. School of Geology, University College of Science, University of Tehran, Tehran, Iran
2. Department of Petroleum Engineering, Curtin University of Technology, Perth, Australia

Abstract

Normalized oil content (NOC) is an important geochemical factor for identifying potential pay zones in hydrocarbon source rocks. The present study proposes an optimal and improved model to make a quantitative and qualitative correlation between NOC and well log responses by integration of neural network training algorithms and the committee machine concept. This committee machine with training algorithms (CMTA) combines Levenberg-Marquardt (LM), Bayesian regularization (BR), gradient descent (GD), one step secant (OSS), and resilient back-propagation (RP) algorithms. Each of these algorithms has a weight factor showing its contribution in overall prediction. The optimal combination of the weights is derived by a genetic algorithm. The method is illustrated using a case study. For this purpose, 231 data composed of well log data and measured NOC from three wells of South Pars Gas Field were clustered into 194 modeling dataset and 37 testing samples for evaluating reliability of the models. The results of this study show that the CMTA provides more reliable and acceptable results than each of the individual neural networks differing in training algorithms. Also CMTA can accurately identify production pay zones (PPZs) from well logs.

Keywords: Normalized oil content, neural network, committee machine with training algorithms, genetic algorithm, well log data, South Pars Gas Field.

---

* Corresponding author: Tel: +618-9266 7980
E-mail addresses: r.rezaee@curtin.edu.au
akadkhoda@khayam.ut.ac.ir (A. Kadkhodaie-Ilkhchi), r.rezaee@curtin.edu.au (MR. Rezaee), rahimpor@khayam.ut.ac.ir (H. Rahimpour-Bonab)
1 Address: School of Geology, University College of Science, University of Tehran, Iran
Tel. /Fax: +98 21 66491623.
2 Department of Petroleum Engineering, Curtin University of Technology, ARRC Building, 26 Dick Perry Avenue, Kensington, WA 6151, Australia.
Tel: +618 9266 7980; Fax: +618 9266 7063

1. Introduction

Normalized oil content (NOC) which is a measure of oil (in mg) produced from one gram of total organic carbon (TOC) at 300 ºC, is a useful parameter for identifying potential pay zones in organic matter bearing intervals. This parameter is measured by Rock-Eval pyrolysis which is a time consuming and expensive method. To date, numerous researchers have tried to make a qualitative and quantitative correlation between well log responses and organic richness of rocks. Among them Beers (1945), Swanson (1960), Fertle (1988), Schmoker (1981), and Hertzog et al. (1989) used gamma-ray spectral log to identifying organic rich rocks. Schmoker and Hester (1983) proposed the use of the density log for estimating organic matter content. Meyer and Nederlof (1984) used a combination of resistivity, density, and sonic logs to discriminate qualitatively between source and non-source rocks. Passey et al. (1990) invented ΔlogR method which employs the separation between sonic and resistivity logs for identifying and calculating total organic carbon. Huang and Williamson (1996) applied neural network modeling for source rock characterization. Kamali and Mirshady (2004) used the ΔlogR and neuro-fuzzy techniques for determination of total organic carbon from well log data.

Committee machine approach which is a new type of neural network can be used to approximate NOC data from well logs. It has a parallel structure that produces a final output by combining the results of individual experts using an optimization technique (Haykin, 1991; Sharkey, 1996; Chen and Lin, 2006). The experts can be empirical formula, neural network, a decision tree, or another type of algorithm (Sharkey, 1996). Genetic algorithm (GA) is an effective optimization technique based on the principles of natural selection and genetics (Holland, 1975). They are often described in biological terms. Potential solutions are called chromosomes. A set of chromosomes is called a population and a problem to be solved is represented by a fitness function. Genetic operators such as crossover and mutation are operators used to create a new population. (Reformat, 1997). More details about GAs can be found in Lucasius and Kateman (1993, 1994), Goldberg (1989) and Huang et al. (2001).

In this research, GA will be applied in construction of a committee for predicting normalized oil content (NOC) from well log data and identifying potential pay zones in

Upper Permian to Lower Triassic Dalan and Kangan Formations (Kadkhodaie-Ilkhchi et al, 2006), South Pars Gas Field, Persian Gulf.

2. The methodology: Committee machine with training algorithms (CMTA)

The proposed methodology, CMTA, consists of four steps: (1) selection of appropriate inputs among the available well log data; (2) Designing back-propagation networks with different training algorithms; (3) Construction of CMTA; and (4) Generalization of the constructed CMTA. The methodology described in this study provides an improved and novel model for predicting NOC parameter in two ways. They are, in use of committee machine concept for predicting NOC parameter and thus reaping the benefit of all of the work, and in use of genetic algorithms for determining the contributions (weights) of individual algorithms used in constructing CMTA. It is clear that many components of the method described in this study are based on other researcher's works which are not novel in their own right. For example, neural network training algorithms or GAs are well known techniques. Overall, the integrated technique described in this study can be considered as an efficient and instrumental way for predicting NOC parameter from well log responses.

2.1. Selection of appropriate inputs

This step of the work plays an important role in model construction. Normally, the inputs with stronger relationships with output can provide more accurate predictions than weaker ones. Relationships between available well log data and NOC are shown in figures 1a-h. Comparisons show that thermal neutron porosity (TNPHI), bulk density (FDC), sonic transit time (DT), and the ratio of true resistivity to flushed zone resistivity (RT/Rxo) have a stronger relationship with NOC, whereas, this relationship is weaker for RT, Rxo, GR, and PEF data. The used well log data are displayed in figure 2.

In order to selection of the appropriate inputs for designing neural networks with different training algorithms, a simple three layered neural network with default parameters was designed for NOC estimation using Matlab software. In input layer, several groups of well log data were considered (152 data points for training and 42 data points for validation). In each run, performance of constructed model in the test data (37 data points) was measured. Results show that selecting DT, TNPHI, FDC, and RT/Rxo data in input layer will be associated with the minimum MSE (Table 1).

The justification based on physical relationships between input used and output data can be stated as below:

Normally, hydrogen index in organic matter is high due to high hydrogen content. Thus, neutron porosity increases in the organic rich intervals. The sonic transit time (DT) is a function of formation lithology, porosity and distribution models of fluids (water, gas, oil, kerogen, etc.). NOC tends to increase the apparent DT value. Organic matters have a low density (about 1 gr/cm^3) and their concentration tends to decrease the bulk density of the rock. Generally, organic rich rocks have high true resistivity than other rocks. Specially, once kerogen becomes mature and generates hydrocarbon filling in voids and fractures.

2.2. Designing networks with different training algorithms

A back-propagation neural network is a supervised training technique that sends the input values forward through the network then computes the difference between calculated output and corresponding desired output from the training dataset. The error is then propagated backward through the net, and the weights are adjusted during a number of iterations named epochs. The training stops when the calculated output values best approximate the desired values (Bhatt and Helle, 2002). Depending on the method used for updating weights and bias values, several training algorithms have been developed. In this study, five of the most common training algorithms are used. A very brief description and some references to each training algorithms are provided in this section.

Levenberg-Marquardt (LM) is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization whose details of computation and process can be find in Boadu (1997, 1998), Bishop (1995) and Burney et al. (2004). It is very fast, but it requires a lot of memory to run.

Bayesian regularization (BR) is a network training function that updates the weight and bias values according to Levenberg-Marquardt optimization. It minimizes a combination of squared errors and weights, and then determines the correct combination so as to produce a network that generalizes well. More details about Bayesian regularization are given in MacKay (1992), Demuth and Beale (2002), and Aggarwal et al. (2005).

Gradient descent (GD) is a network training function that updates weight and bias values according to gradient descent. More description can be found in Baird and Moore (1999) and Kononen (2005).

One step secant (OSS) is a network training function that updates weight and bias values according to the one step secant method. More details are given in Battiti (1992).

Resilient back-propagation (RP) is a network training function that updates weight and bias values according to the resilient back-propagation algorithm (Riedmiller and Braun, 1993).

2.3. Construction of CMTA

Generally, a committee machine consists of a group of experts which combines the outputs of each system and thus reaps the benefits of all of the work, with little additional computation. So, performance of the model can be better than best single network (Haykin, 1991; Sharkey, 1996; Chen and Lin, 2006). A schematic diagram of a committee machine is shown in figure 3. There are different ways of combining the experts in the combiner. The simple ensemble averaging method is most popular (Naftaly et al., 1997, Chen and Lin, 2006). Proper combination of contribution (weight) of individual experts in a committee machine can be obtained by a GA.

In this study, experts of committee machine are different training algorithms of back propagation neural network (CMTA). The section below describes the fundaments of our CMTA with regard to the works of Bates and Granger (1969), Haykin (1991), Geman et al. (1992), Naftaly et al. (1997), Huang et al. (2001), Ligtenberg and Wansink (2001), Bhatt and Helle (2002), Lim (2005), and Chen and Lin (2006).

Assumption is that there are N training algorithms with output vector o_i which are used to predict target vector T. The prediction error can be written as

$$e_i = o_i - T\ , \tag{1}$$

The sum of the squared error for the i^{th} network o_i is

$$E_i = \xi[(o_i - T)^2] = \xi[e_i^{\,2}], \tag{2}$$

in which $\xi[.]$ is the expectation. The average error for each of the algorithms acting alone is

$$E_{avg} = \frac{1}{N}\sum_{i=1}^{N} E_i = \frac{1}{N}\sum_{i=1}^{N}\xi[{e_i}^2], \tag{3}$$

Applying the averaging method, output vector o_i of the CMTA is

$$O_{CMTA} = \frac{1}{N}\sum_{i=1}^{N} o_i \ . \tag{4}$$

Therefore, the CMTA has the prediction squared error:

$$E_{CMTA} = \xi[(O_{CMTA} - T)^2] = \xi[(\frac{1}{N}\sum_{i=1}^{N} o_i - T)^2] = \xi[(\frac{1}{N}\sum_{i=1}^{N} e_i)^2]. \tag{5}$$

Considering Cauchy's inequality:

$$(a_1 b_1 + a_2 b_2 + ... + a_n b_n) \le (a_1^2 + a_2^2 + ... + a_n^2).(b_1^2 + b_2^2 + ... + b_n^2) \tag{6}$$

and applying it to the E_{CMTA}

$$E_{CMTA} = \xi[(\frac{1}{N}\sum_{i=1}^{N} e_i)^2] \le \frac{1}{N}\sum_{i=1}^{N}\xi[{e_i}^2] = E_{avg}. \tag{7}$$

which indicates that the CMTA gives more accurate and reliable estimations than that of any one of the individual training algorithms.

2.4. Generalization of the constructed CMTA

In this part of research, a CMTA was used for overall prediction of NOC by combination of the results obtained from different training algorithms of neural network. As the inputs of the mentioned CMTA are individual neural networks so, at the first stage, several neural networks were designed to learn the relationships between NOC data and well log responses. Afterward, the CMTA were constructed using two methods including simple averaging and weighted averaging. In the simple averaging method, the outputs estimated from individual neural networks were simply averaged to produce final estimation of NOC data. In weighted averaging method, the results estimated from individual neural network experts were multiplied by a weight factor showing its contribution in overall prediction. The GA was used to obtain weight coefficients from training data. Then, they were applied to the test data (Eq. 8).

Following is the equation used for final estimation of NOC by CMTA:

$$NOC_{CMTA} = \sum_{i=1}^{N} w_i .NOC_i \tag{8}$$

where N is the total number of the algorithms used, w_i is the weight coefficient of algorithm i and NOC_i is the estimated NOC from algorithm i.

3. Case study

3.1. Data preparation and processing

The data sets used in this study for models construction and evaluation came from three wells of South Pars Gas Field. Two hundred thirty-one data composed of well log data and NOC (from Rock-Eval pyrolysis) were used. One hundred ninety-four training data points were used for models construction and 37 samples from the third well for testing the developed models. Well log data were processed and bad hole intervals were removed. FDC values ranged from 2.04 to 2.83 g/cm^3 (average: 2.46). DT varied from 51.18 to 79.30 μs/ft (average: 64.17). TNPHI was between 0.021 and 0.158 pu (average: 0.098), and RT/Rxo varied from 1.25 from 69.78 (average: 21.02). The target parameter, NOC, was between 8.0 and 517 mg Oil/g TOC (average: 159.44).

3.2. Predicting NOC by CMTA

As the experts of our CMTA are various training algorithms, first a back propagation neural network was designed in Matlab software environment. In order to design the networks with different training algorithms it was necessary to set optimal parameters of each one including number of hidden layers, number of neurons in hidden layers, training epochs, and transfer functions. These parameters were determined by trial and error. Specifying inadequate number of training epochs or training data may lead to under-training. For example, stopping too early means the ANN has not yet learnt all the information from the training data. Another major pitfall of neural network is over-training in which the network only memorizes the training set and loses its ability to generalize to new data. The result is a network that performs well on the training set but performs poorly on out-of-sample test data and later during actual trading (Tetko et al., 1995). Adding more hidden layers involves adding activation (using the outputs of the previous hidden layer) and error correction calculations (using the derivative of the transfer function) for each layer. Both situations are likely to result in sub-optimal operational performance of an ANN model. It is for this reason that the available data were divided in three separate data sets: a: training set (194 data points, b: cross-validation set (42 data point), and c: validation set (37 data point). The minimization of the training error is stopped as soon as the cross-validation error starts to increase. This point is considered to lie between under-training and over-training an ANN. An example of over-training and under-training problems is shown in figure 4. Followings are the optimum parameters of the networks designed:

Four neurons corresponding to well log data including DT, TNPHI, FDC, and RT/Rxo were considered in input layer, respectively. The network included one hidden layer. The output layer included one neuron for NOC data. Number of neurons in hidden layers is 7 for LM, 9 for BR and OSS, 4 for GD, and 8 for RP algorithm. Tansigmoid transfer function was selected from layer one to two and Purelin transfer function from layer two to three for all of the networks. The mean squared errors (MSE) of training and validation data for algorithms used are shown graphically in figure 5. According to figure 5, RP has the minimum MSE for both training and validation data.

Adjusted weight and bias values, after a specific number of epochs, for different training algorithms including bias from layer 1 to layer 2 (b{1}), weights from layer 1 to layer 2 (W{1}), bias from layer 2 to layer 3 (b{2}), weights from layer 2 to layer 3 (W{2}) and MSE are shown in Table 2. The architecture of the constructed networks is shown in figure 6.

After training procedure, the CMTA was constructed. Determining the number of the algorithms to be combined in the committee machine is necessary for obtaining accurate results. For this purpose, numerous cases of the algorithms combination were considered in constructing CMTA. The combinations ranged from two to the entire training algorithms. The mentioned CMTAs were first constructed by applying simple averaging method. In this approach, any one of the training algorithms has equal contribution in constructing CMTA.

In the next step, a genetic algorithm was used to obtain appropriate weight coefficients of CMTA in training data. The fitness function which should be minimized by GA was defined as MSE of training data predictions (Eq. 9):

$$MSE_{CMTA} = \sum_{i=1}^{m} 1/m((\sum_{i=1}^{N} w_i . NOC_i) - NOC_{measured})^2 \quad (9)$$

where m is the number of training data (152 samples), w_i, N, and NOC_i are same as those of Eq. (8). Parameter settings for GA are described in below.

Initial population size is 30 which specifies number of individuals in each generation and initial range is [0, 1] which specifies the range of the vectors in the initial population. The crossover function is *scattered* and its fraction is 0.88. Mutation function is *Gaussian* that adds a random number, or mutation, from a Gaussian distribution, to each entry of the parent vector. Parameters controlling the mutation are specified as the *scale value* of 0.9 and *shrink value* of 1. The *scale value* controls the standard deviation of the mutation at the first generation. *Shrink value* controls the rate at which the average amount of mutation decreases. The standard deviation decreases linearly so that its final value equals 1.

After running the GA, optimized weight coefficients were applied in CMTA to produce the final output.

3.3. Results and discussion

The performance of the simple averaging and the weighted averaging CMTA constructed from combining two, three, four, and all of the training algorithms are shown in Table 3. According to Table 3, the simple averaging CMTA using LM, BR, RP, and OSS methods has produced the minimum error whereas, combination of the all the algorithms used (LM, BR, GD, OSS, and RP) in weighted averaging CMTA is associated with the minimum error. So, the GA optimized case was selected for overall estimation of NOC. According to figure 7a, after 80 generations the mean and best fitness values were fixed in 0.024 and 0.022, respectively. Figure 7b shows best, worst and mean scores within mentioned 80 generations. The GA derived values for w_1, w_2, w_3, w_4, and w_5 corresponding to LM, BR, GD, OSS, and RP estimations are 0.187, 0.241, 0.082, 0.076, and 0.413, respectively. Figure 8 shows the diagram of CMTA designed in this study. Overall estimation of NOC by CMTA for testing data (37 samples) was calculated as below:

$$NOC_{CMTA} = 0.187\times NOC_{LM} + 0.241\times NOC_{BR} + 0.082\times NOC_{GD} + 0.076\times NOC_{OSS} + 0.413\times NOC_{RP} \quad (10).$$

Table 4 shows the comparison of MSE for 37 testing data points from well C using different algorithms including LM, BR, GD, OSS, RP, simple averaging CMTA (all algorithms), and GA optimized CMTA (all algorithms). Considering crossplots of figure 9a-e and Table 4, among the five neural network algorithms used, RP has provided the smallest error (MSE=2.078) and R^2 value of 0.703 for the test samples. In the meanwhile, GD is associated with highest error (MSE=2.291). Applying averaging method for construction of CMTA using all algorithms has provided MSE of 1.981 and R^2 value of 0.725 (figure 9f) which shows some improvement in comparison with individual training algorithms. MSE of the GA optimized CMTA using all algorithms for the test data is 1.860 which corresponds to the R^2 value of 0.751 (figure 10). This indicates that CMTA has had a significant improvement for the estimation of NOC from well log data. Namely, CMTA performs better than any one of the individual training algorithms acting alone for NOC predicting problem. Also it has provided better results than constructed CMTA by simple averaging method. It might be noticed that in our case study the weighted averaging committee machine performed better than simple averaging method whereas; in some cases it may not be so. For example, if the weighted averaging CM in the best possible conditions provides the equal weights for all of the experts used, then the simple averaging committee machine will be preferred. However, as a general rule it can be said that CMs provide better results than simple averaging methods to solve a problems (Cauchy's inequality, Eq. (6)).

Generally, the zones with NOC>100 are considered as potential pay zone (PPZ). Figure 11 is a graphical illustration showing a comparison between PPZs determined from measured (11b) and CMTA predicted NOC (11c) (zones in black color). According to figure 11, irrespective of the interval between 2819.50 and 2816.60, there is a good agreement between measured and predicted PPZs. Specially, once the NOC is around the value of 100, CMTA can identify PPZs successfully. In figure 12a, the results of generalization of CMTA for the forth well of the South Pars Gas Field which has no core data is shown. Predicted PPZs based on CMTA in this well are shown in figure 9b (black zones).

Table 1 Performance of the neural network for predicting NOC in the test data using several sets of input well log data (in this table only 9 of top sets is shown).

Inputs	MSE
DT	14.53
DT, TNPHI	8.01
DT, FDC, TNPHI	6.39
DT, FDC, TNPHI, GR	7.80
DT, FDC, TNPHI, PEF	7.25
DT, FDC, TNPHI, RT	4.81
DT, FDC, TNPHI, Rxo	4.97
DT, FDC, TNPHI, RT/Rxo	4.60
DT, FDC, TNPHI, RT/Rxo, PEF, GR	7.11

Table 2 Adjusted network parameters for different training algorithms

Algorithm	No. of neurons in layers 1, 2, 3	W {1}	b {1}	W {2}	b {2}	Epochs	MSE
LM	4, 7, 1	0.113 -12.925 -0.710 0.039	-8.158	-0.822	-0.043	7	0.101
		0.005 29.640 0.901 0.028	-7.720	-0.457			
		0.087 -12.722 1.574 -0.046	-7.538	-0.181			
		-0.025 12.240 -4.577 -0.030	12.83	-0.051			
		0.097 -19.379 3.012 0.011	-11.60	0.817			
		-0.133 -9.173 2.839 0.004	0.927	0.192			
		0.016 -0.106 3.178 -0.055	-4.557	-0.342			
BR	4, 9, 1	0.064 27.099 -0.707 -0.033	-6.059	0.810	-0.455	12	0.087
		-0.125 -15.021 2.761 0.029	3.845	-0.687			
		0.011 23.109 0.778 0.050	-7.730	-0.755			
		0.097 29.102 0.261 0.006	-10.27	0.525			
		0.041 -19.495 -3.398 0.041	5.857	0.443			
		0.047 -23.521 -2.929 0.035	5.480	0.308			
		0.086 -22.903 -0.984 0.039	-1.352	0.506			
		-0.117 -13.506 -1.928 -0.037	13.11	0.326			
		-0.071 1.264 3.673 0.044	-8.518	0.766			
GD	4, 4, 1	-4.931 -2.451 4.648 -14.988	-7.946	- 7.571	49.669	9	0.126
		2.226 -1.411 7.437 -20.480	11.44	- 66.06			
		8.260 -2.728 -2.169 -8.0755	2.605	16.341			
		0.649 -3.859 -7.778 -0.0141	6.270	65.669			
OSS	4, 9, 1	-0.092 -24.851 2.500 -0.017	5.202	-0.422	-0.569	15	0.130
		0.051 -29.817 0.589 0.033	-5.025	0.942			
		-0.095 -16.468 3.380 -0.025	1.754	0.901			
		-0.071 -18.196 -1.616 0.049	9.094	-0.543			
		0.050 25.299 2.197 -0.037	-9.547	0.917			
		-0.049 -19.390 4.120 0.029	-6.745	0.359			
		-0.125 15.590 1.895 0.030	-0.104	-0.890			
		0.063 -13.617 3.905 -0.039	-9.194	0.199			
		0.104 -13.135 2.045 -0.044	-6.607	-0.213			

RP	4, 8, 1	-0.038 -27.587 -2.317 -0.026 0.098 -16.527 1.873 0.039 -0.082 16.256 2.996 0.036 0.023 -21.142 2.248 -0.046 0.120 -13.028 1.654 0.035 -0.053 25.738 -3.314 0.012 0.102 -15.624 -3.825 0.006 -0.098 -0.080 -4.287 -0 .025	13.942 -12.556 -3.627 -3.767 -11.536 7.7739 5.4584 15.435	0.644 -0.473 0.507 0.319 -0.571 0.204 0.209 0.319	-0.633	18	0.057

Table 3 Performance of the constructed CMTA (simple averaging and weighted averaging) by combining different number of the training algorithms. In this table, the best results obtained from combining two, three, four, and all of algorithms are shown

Algorithms used (Best combinations)	**Performance of CMTA (MSE) and corresponding weights to their algorithms used**	
	Simple averaging	**Weighted averaging (GA optimized)**
BR, RP	2.034 *(w_1=w_2=0.50)*	2.016 *(w_1=0.431; w_2=0.569)*
LM, BR, RP	1.975 *(w_1=w_2=w_3=0.333)*	1.932 *(w_1=0.211; w_2=0.385; w_3=0.403)*
LM, BR, RP, OSS	1.958 *(w_1=w_2=w_3=w_4=0.25)*	1.927 *(w_1=0.126; w_2=0.321; w_3=0.480; w_4=0.073)*
LM, BR, GD, OSS, RP	1.981 *(w_1=w_2=w_3=w_4=w_5=0.20)*	1.860 *(w_1=0.187; w_2=0.241; w_3=0.082; w_4=0.076; w_5=0.413)*

Table 4 Comparison of MSE for test well data using different algorithms.

Algorithm	**MSE**	**Rank**
LM	2.160	5
BR	2.148	4
GD	2.291	7
OSS	2.255	6
RP	2.078	3
CMTA (simple averaging)	1.981	2
CMTA (GA optimized)	1.860	1

4. Conclusions

In this paper, a committee machine with training algorithms (CMTA) of back propagation neural network were developed for the estimation of NOC from well log data in South Pars Gas Field. Among the different algorithms used resilient back propagation (RP) is associated with the smallest error (MSE=2.078). In CMTA, each algorithm has a weight coefficient which was obtained by simple averaging method and genetic algorithm. In simple averaging method, combination of RP, BR, LM, and OSS algorithms was produced the minimum MSE (1.958) whereas, in weighted averaging method, combination of all the training algorithms including LM, BR, GD, OSS, and RP was the best case (MSE=1.086). The GA derived weights for w_1 (LM), w_2 (BR), w_3 (GD), w_4 (OSS), and w_5 (RP) are 0.187, 0.241, 0.082, 0.076, and 0.413, respectively. The CMTA is expected to provide improved and more accurate results when there are multiple ways to solve a problem, as our research demonstrated it. Similarly, CMTA was successful to identify potential pay zones from well logs.

5. Acknowledgements

The vice-president of Research and Technology of the University of Tehran provided financial support for this research, which we are grateful. We also extend our appreciation to the P.O.G.C (Pars Oil and Gas Company of Iran) for sponsoring, data preparation, and permission to publish this paper.

5. References

Aggarwal, K., Singh, Y., Chandra, P., Manimala, P., 2005. Bayesian Regularization in a Neural Network Model to Estimate Lines of Code Using Function Points. J. Comput. Sci. 1 (4), 505-509.

Baird, L., Moore, A., 1999. Gradient descent for general reinforcement learning. Advances in Neural Information Processing Systems 11, 968-974.

Bates, J.M., Granger, C.W.J., 1969. The combination of forecast. Operations Research Quarterly 20, 451–468.

Battiti, R., 1992. First and second order methods for learning: Between steepest descent and Newton's method. Neural Computation 4 (2), 141-166.

Beers, R.F., 1945. Radioactivity and organic content of some Paleozoic shales. AAPG Bulletin 26, 1–22.

Bhatt, A., Helle, H.B., 2002. Committee neural networks for porosity and permeability prediction from well logs. Geophys. Prospect. 50, 645–660.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Clarendon Press, Oxford, p. 670.

Boadu, F.K., 1997. Rock properties and seismic attenuation: neural network analysis. Pure Appl. Geophys. 149, 507–524.

Boadu, F.K., 1998. Inversion of fracture density from field seismic velocities using artificial neural networks. Geophysica 63, 534 545.

Burney, S.M.A., Jilani, T.A., Ardil, C., 2004. Levenberg-Marquardt algorithm for Karachi sock exchange share rates forecasting. Transactions on Engineering, Computing and Technology 3, 1305-1313.

Chen, C.H., Lin, Z.S., 2006. A committee machine with empirical formulas for permeability prediction. Journal of Computers & Geosciences 32, 485-496.

Demuth, H., Beale, M., 2002. Neural Network Toolbox, User's Guide (Version 4), the Mathworks Inc.

Fertle, H., 1988. Total organic carbon content determined from well logs. SPE Formation Evaluation 15612, 407– 419.

Geman, S., Bienenstock, E., Doursat, R., 1992. Neural networks and the bias/variance dilemma. Neural Computation 4, 1–58.

Goldberg, D.E., 1989. Genetic algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA.

Haykin, S., 1991. Neural Networks: A Comprehensive Foundation. Prentice-Hall, Englewood Cliffs, NJ p. 842.

Hertzog, R., Colson, L., Seeman, B., O'Brian, M., Scott, H., 1989. Geochemical logging with spectrometry tools. SPE Formation Evaluation 4, 153– 162.

Holland, J.H., 1975. Adaptation in Natural and Artificial Systems. Ann Arbor, University of Michigan Press.

Huang, Y., Gedeon, T.D., Wong, P.M., 2001. An integrated neural-fuzzy-genetic-algorithm using hyper-surface membership functions to predict permeability in petroleum reservoirs. Engineering Applications of Artificial Intelligence 14, 15-21.

Huang, Z., Williamson, M.A., 1996. Artificial neural network modelling as an aid to source rock characterization. Marine and Petroleum Geology 13, 227-290.

Kadkhodaie-Ilkhchi, A., Rezaee, M.R., Moallemi, S.A. 2006. A fuzzy logic approach for the estimation of permeability and rock types from conventional well log data: an example from the Kangan reservoir in Iran Offshore Gas Field, Iran. Journal of Geophysics and Engineering 3, 356-369.

Kamali, M.R., Mirshady, A.A., 2004. Total organic carbon content determined from well logs using ΔlogR and neuro fuzzy techniques. J. Pet. Sci. Eng. 45, 141–148.

Kononen, V., 2005. Gradient descent for symmetric and asymmetric multiagent reinforcement learning. Web Intelligence and Agent Systems 3, 17-30.

Ligtenberg, J.H, Wansink, A.G., 2001. Neural Network Prediction of Permeability in the EL Garia Formation, Ashtart Oilfield, Offshore Tunisia, Journal of Petroleum Geology 24, 389-404.

Lim, J-S., 2005. Reservoir properties determination using fuzzy logic and neural networks from well data in offshore Korea. J. Pet. Sci. Eng. 49, 182-192.

Lucasius, C.B., Kateman, G., 1993. Understanding and using genetic algorithms: part 1. Concepts, properties and context. Chemometrics and Intelligent Laboratory System 19, 1231.

Lucasius, C.B., Kateman, G., 1994. Understanding and using genetic algorithms: part 2. Representation, configuration and hybridization. Chemometrics and Intelligent Laboratory System 20, 92145.

MacKay, D. J. C., 1992. A practical Bayesian framework for backpropagation networks. Neural Computation 4, 448-472.

Meyer, B.L., Nederlof, M.H., 1984. Identification of source rocks on wireline logs by density/resistivity and sonic transit time/resistivity cross plots. AAPG Bulletin 68, 121–129.

Naftaly, U., Intrator, N., Horn, D., 1997. Optimal ensemble averaging of neural networks. Network: Computation in Neural Systems 8, 283–296.

Passey, O.R., Moretti, F.U., Stroud, J.D., 1990. A practical modal for organic richness from porosity and resistivity logs. AAPG Bulletin 74, 1777–1794.

Reformat, M., 1997. Application of Genetic Algorithms in Control Design for Advanced Static VAR Compensator at ac/dc Interconnection. University of Manitoba Press, p. 129.

Riedmiller, M., Braun, H., 1993. A direct adaptive method for faster back-propagation learning: The RPROP algorithm. Proc. IEEE Conf. on Neural Networks.

Schmoker, J.W., 1981. Determination of organic-matter content of Appalachian Devonian shales from gamma-ray logs. AAPG Bulletin 65, 2165–2174.

Schmoker, J.W., Hester, T.C., 1983. Organic carbon in Bakken Formation, United States portion of Williston Basin. AAPG Bulletin 67, 2165–2174.

Sharkey, A.J.C., 1996. On combining artificial neural nets. Connection Science 8, 299-314.

Swanson, V.E., 1960. Oil yield and uranium content of black shales. USGS professional paper 356-A, p. 1–44.

Tetko I.V., Livingstone, D.J., Luik, A.I., 1995. Neural network studies: 1. Comparison of overfitting and overtraining. Journal of Chemical Information & Computer Sciences, 35: 826-833.

Figure captions

Figure 1 Crossplots showing the relationship between NOC and TNPHI (a), FDC (b), DT (c), RT/Rxo (d), RT (e), Rxo (f), GR (g), and PEF (h) in the well A (Kangan Formation).

Figure 2 Display of the used well log data in the well A.

Figure 3 A schematic diagram of a committee machine (Haykin, 1991).

Figure 4 Graph showing MSE for training and validation data against training epochs for a complex network trained by BR algorithm (this does not indicate optimum network for BR). According to the figure, after 10 epochs, MSE decreases for training data and increases for validation data. This epoch is a boundary between over-training and under-training. Stopping earlier means that a network does not take full advantage of the information content of the input signals, and stopping later means that the networks loses its capability to generalize.

Figure 5 Graphs showing the mean squared error (MSE) of training and validation data predictions by training algorithms.

Figure 6 A schematic diagram showing architecture of the constructed networks with different training algorithms.

Figure 7 (a) Plot showing mean and best fitness values for fitness function after 80 generations. **(b)** Best, worst and mean scores within 80 generations.

Figure 8 Diagram showing the CMTA designed in this study.

Figure 9 Crossplots showing the correlation coefficient between measured and predicted NOC from LM (a), BR (b), GD (c), OSS (d), RP (e), and averaging on based CMTA (f).

Figure 10 (a) Crossplots showing the correlation coefficient between measured and predicted NOC from genetic algorithm optimized CMTA at the test well. **(b)** Graph showing a comparison between measured and CMTA predicted NOC at the test well.

Figure 11 Graphical illustrations showing stair diagram of measured NOC at the test well **(a)**, PPZs determined from measured data **(b)**, and PPZs predicted from CMTA **(c)**, PPZs are displayed by black colors.

Figure 12 (a) Stair diagram showing the predicted NOC from CMTA at the forth well of the study field. **(b)** A graphical illustration showing PPZs determined from generalization of CMTA to the forth well, PPZs are displayed by black colors.

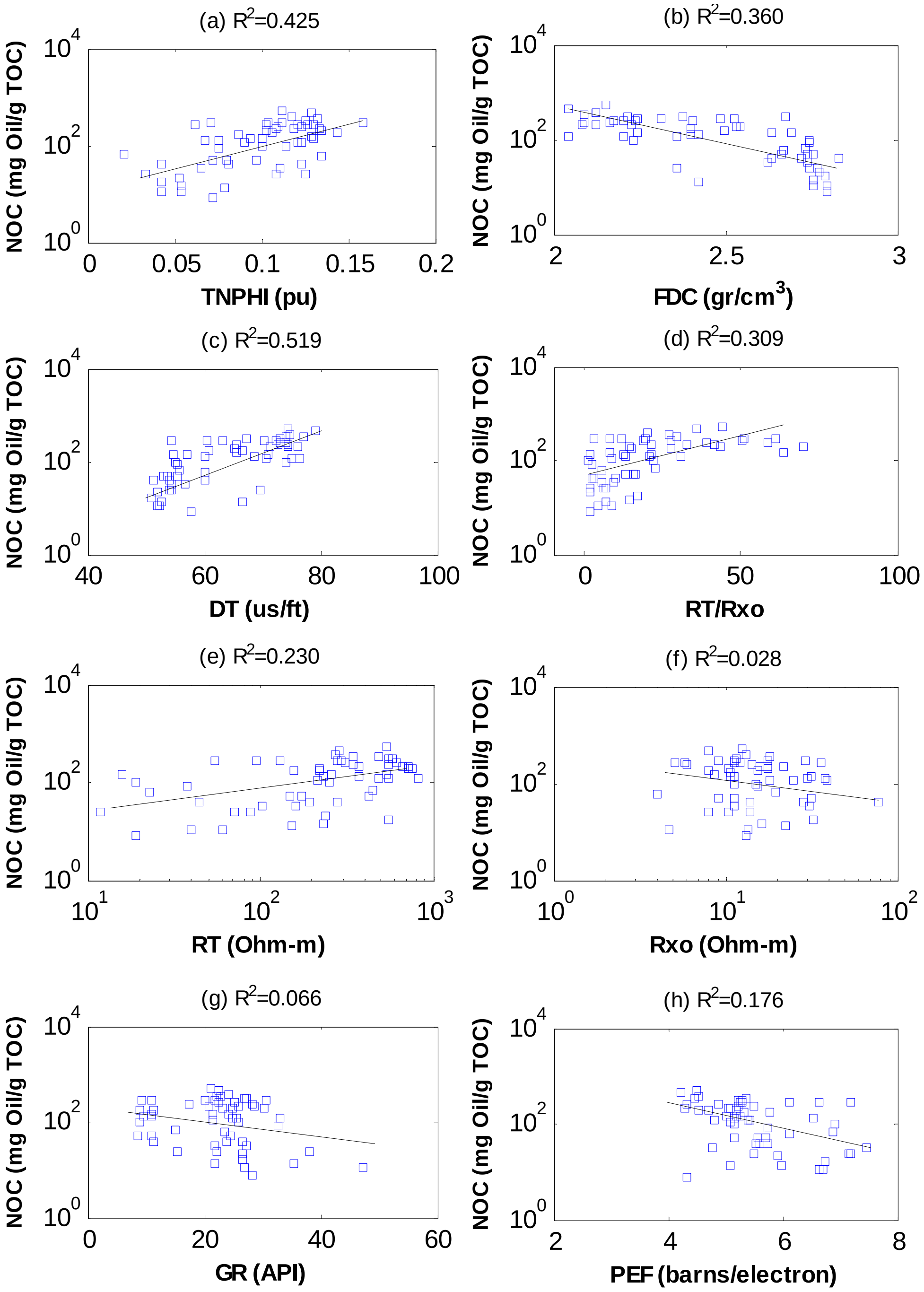

Figure 1 Cross-plots showing the relationship between NOC and TNPHI (a), FDC (b), DT (c), RT/Rxo (d), RT (e), Rxo (f), GR (g), and PEF (h) in the well A (Kangan Formation).

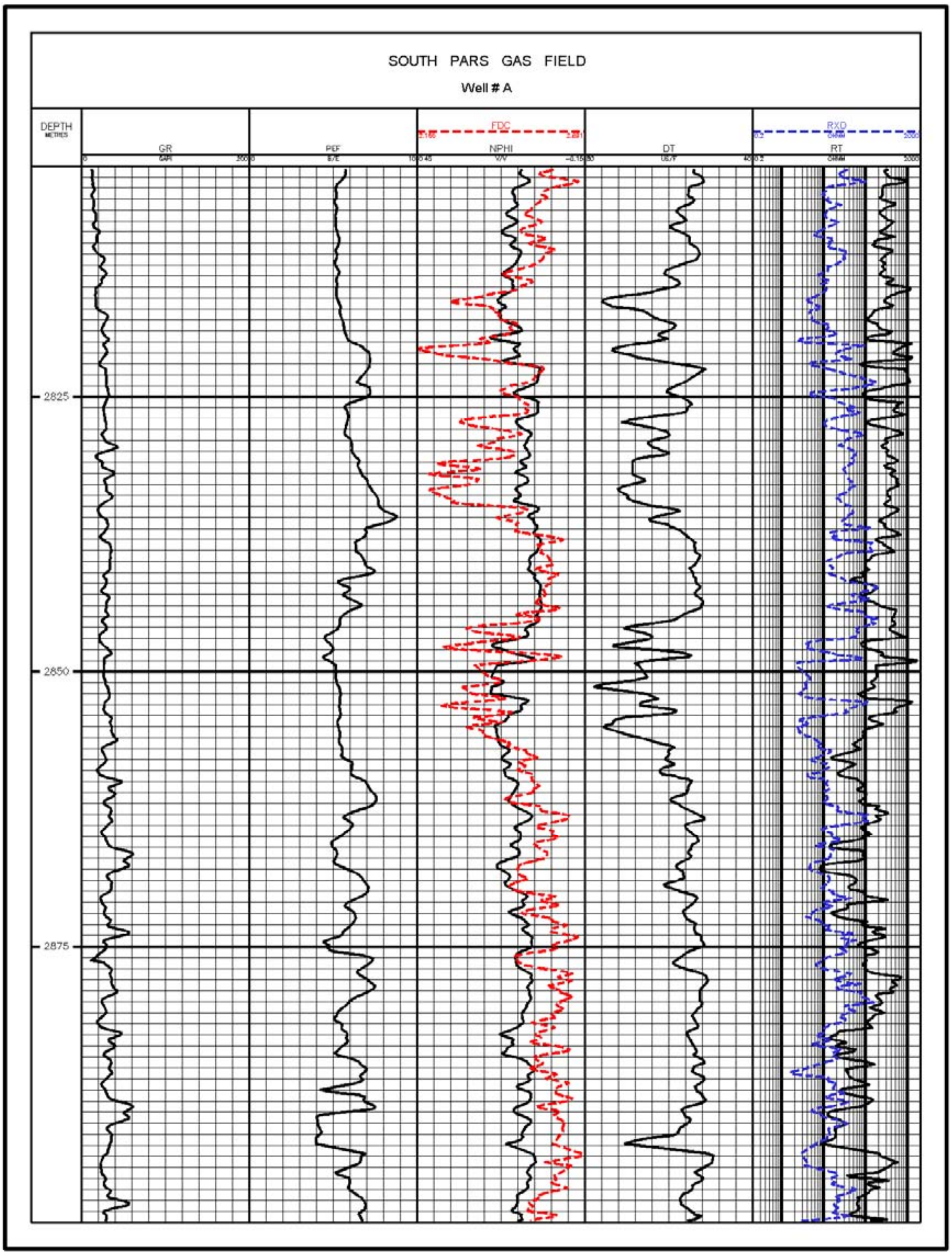

Figure 2 Display of the used well log data in the well A.

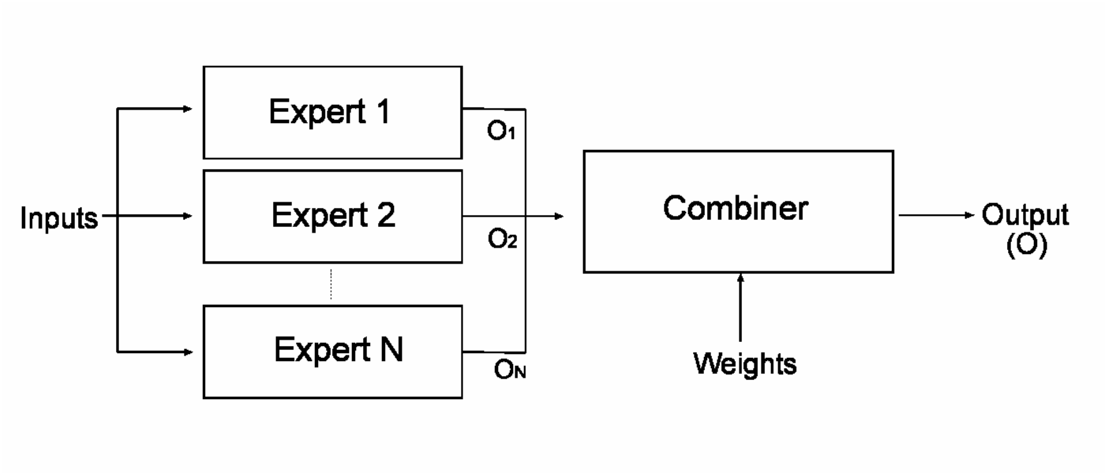

Figure 3 A schematic diagram of a committee machine (Haykin, 1991).

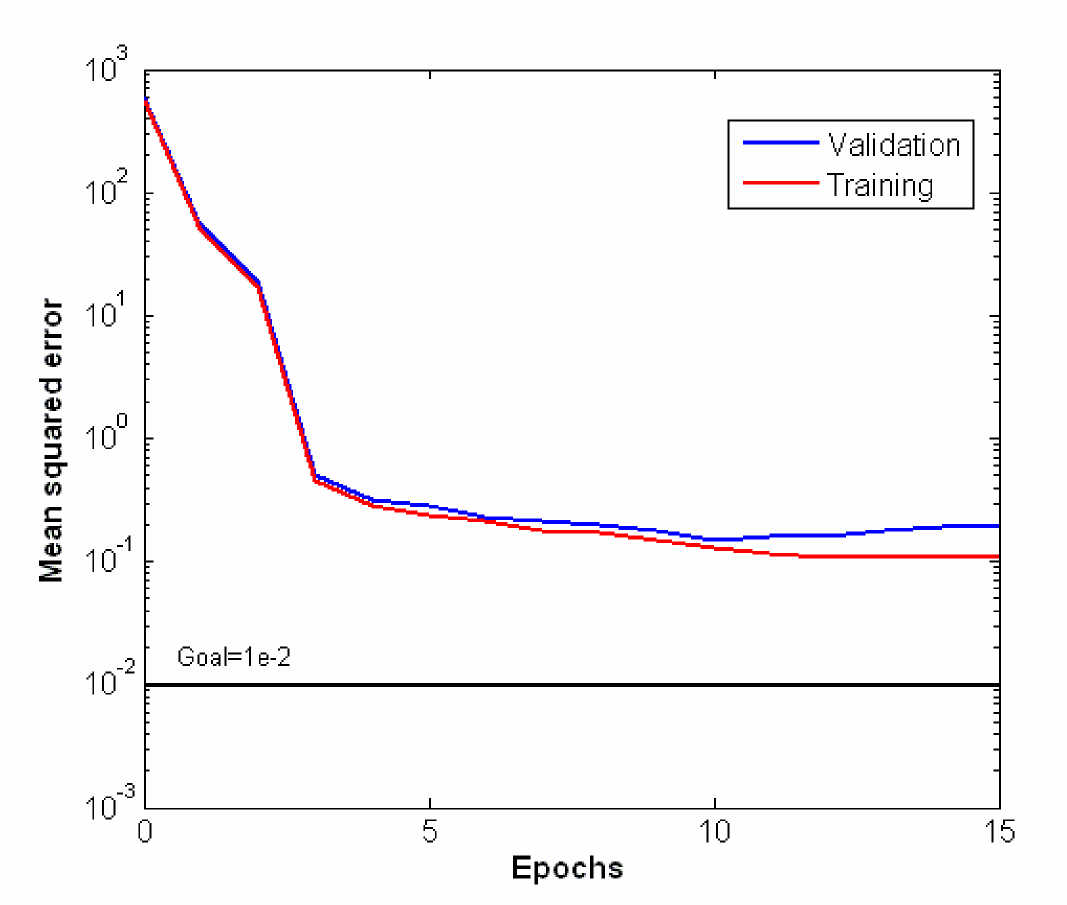

Figure 4 A graph showing MSE for training and validation data against training epochs for a complex network trained by BR algorithm (this does not indicate optimum network for BR). According to the figure, after 10 epochs, MSE decreases for training data and increases for validation data. This epoch is a boundary between over-training and under-training. Stopping earlier means that a network does not take full advantage of the information content of the input signals, and stopping later means that the networks loses its capability to generalize.

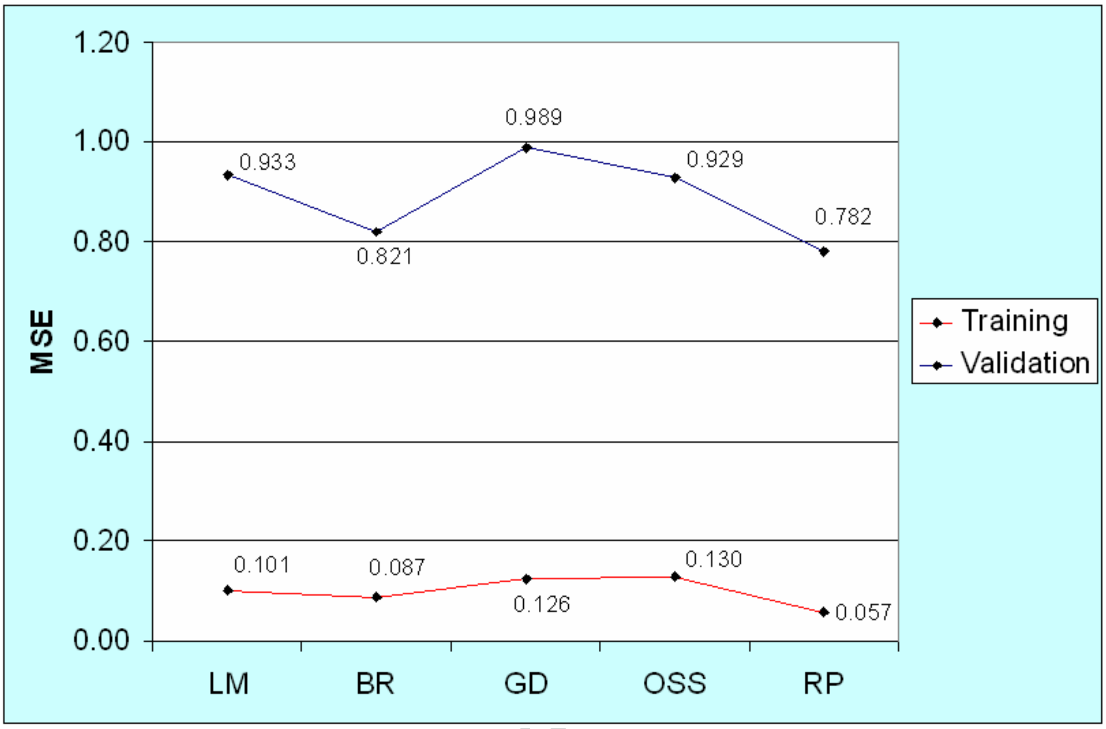

Figure 5 Graphs showing the mean squared error (MSE) of training and validation data predictions by training algorithms.

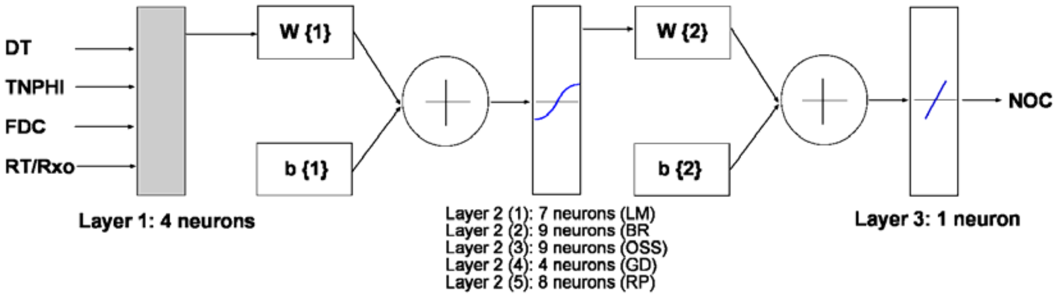

Figure 6 A schematic diagram showing architecture of the constructed networks with different training algorithms.

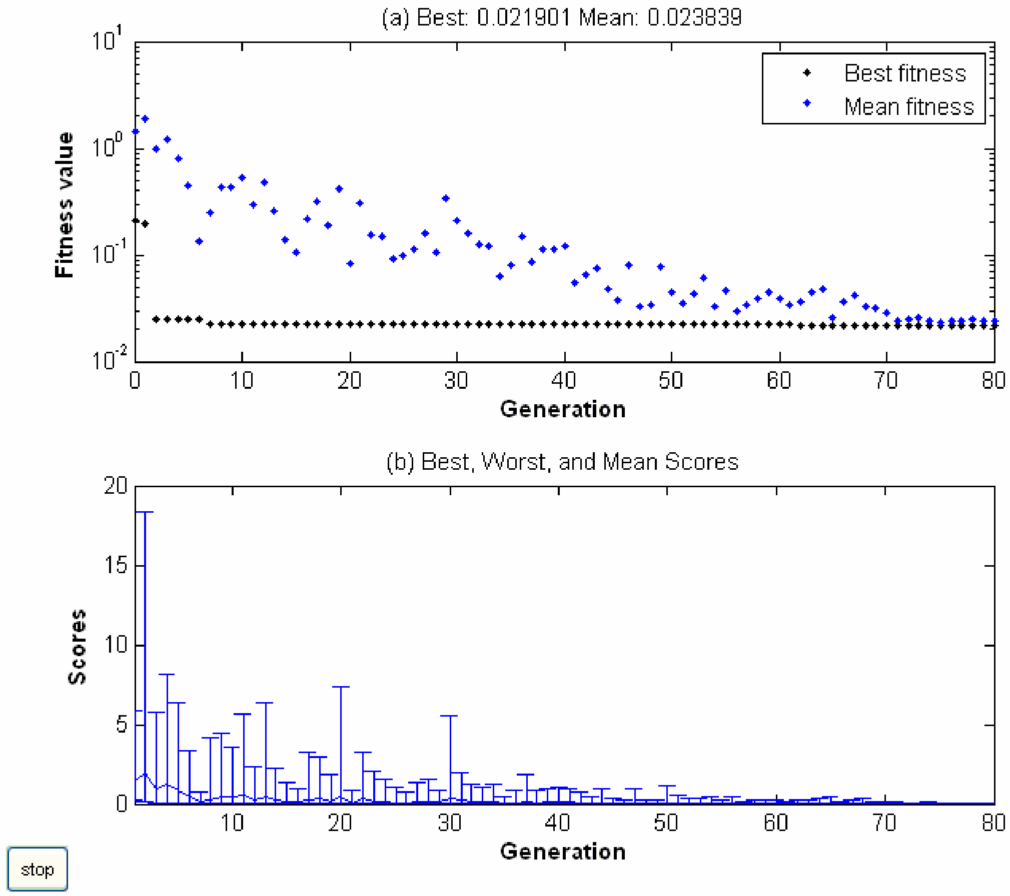

Figure 7 (a) Plot showing mean and best fitness values for fitness function after 80 generations. **(b)** Best, worst and mean scores within 80 generations.

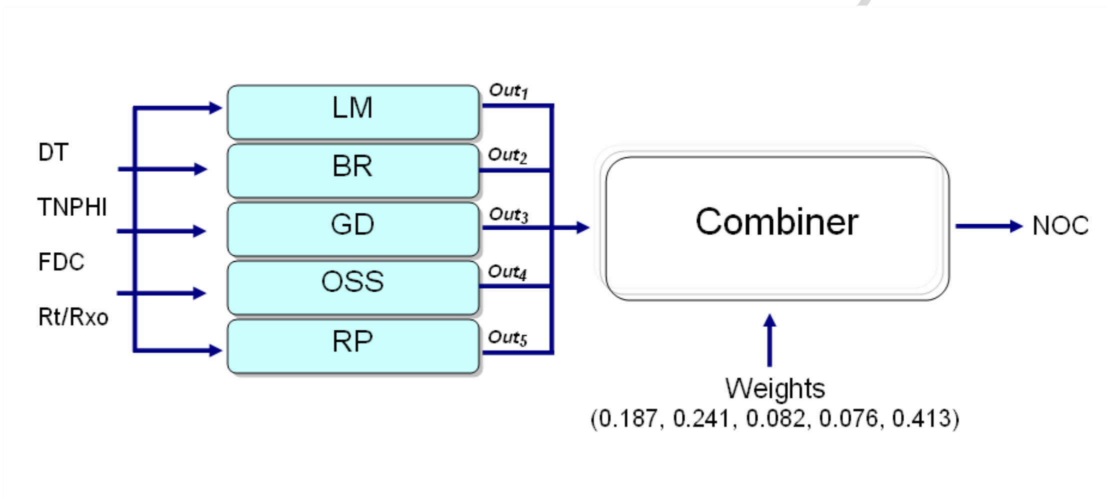

Figure 8 Diagram showing the CMTA designed in this study.

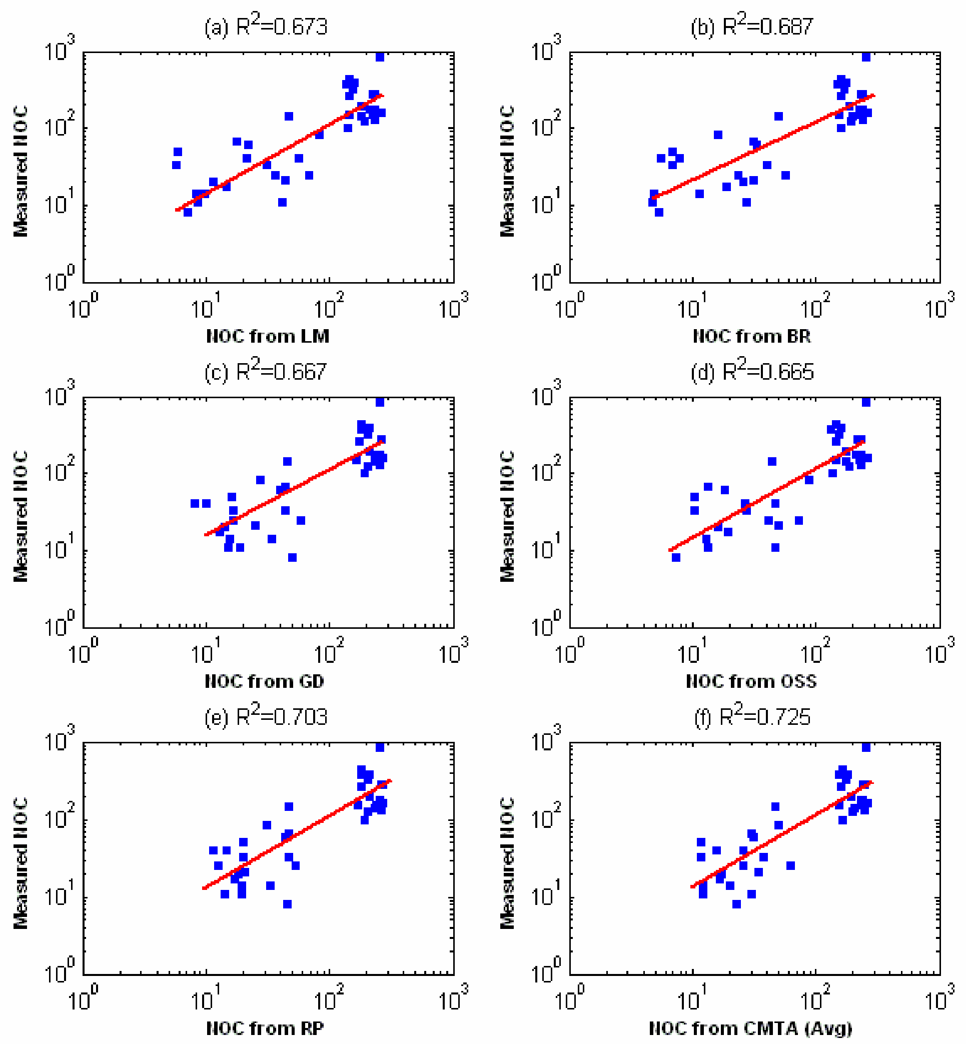

Figure 9 Cross-plots showing the correlation coefficient between measured and predicted NOC from LM (a), BR (b), GD (c), OSS (d), RP (e), and averaging on based CMTA (f).

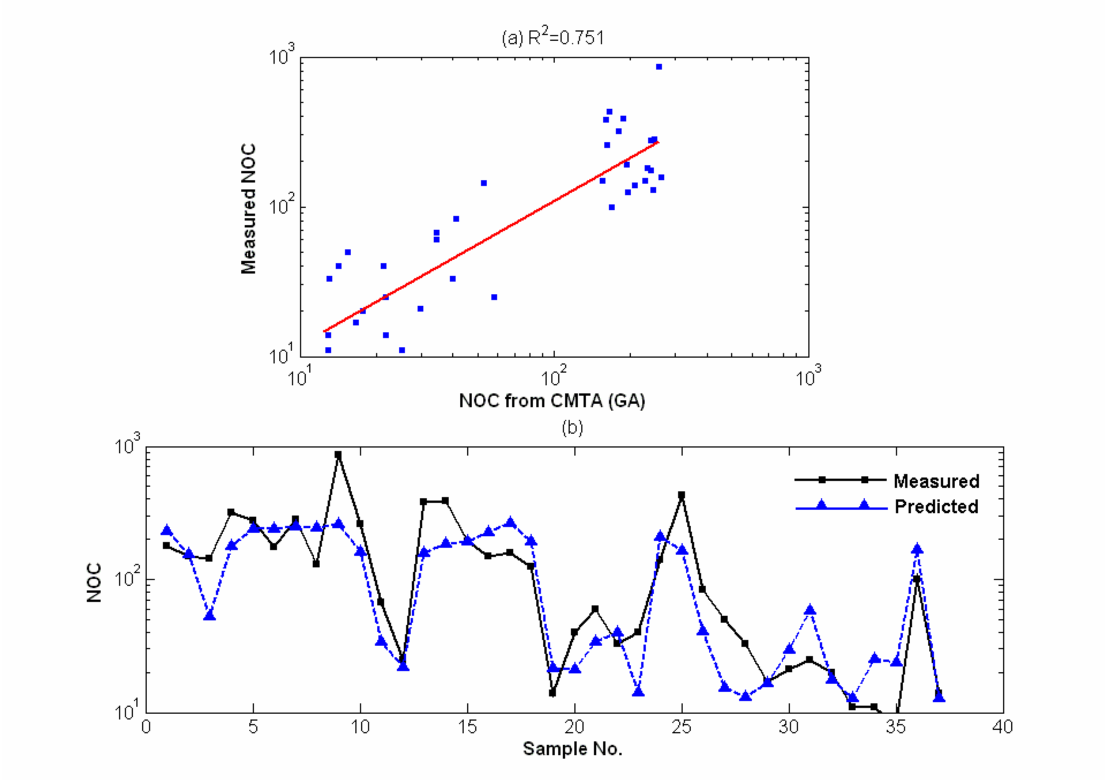

Figure 10 (a) Cross-plots showing the correlation coefficient between measured and predicted NOC from genetic algorithm optimized CMTA at the test well. **(b)** Graph showing a comparison between measured and CMTA predicted NOC at the test well.

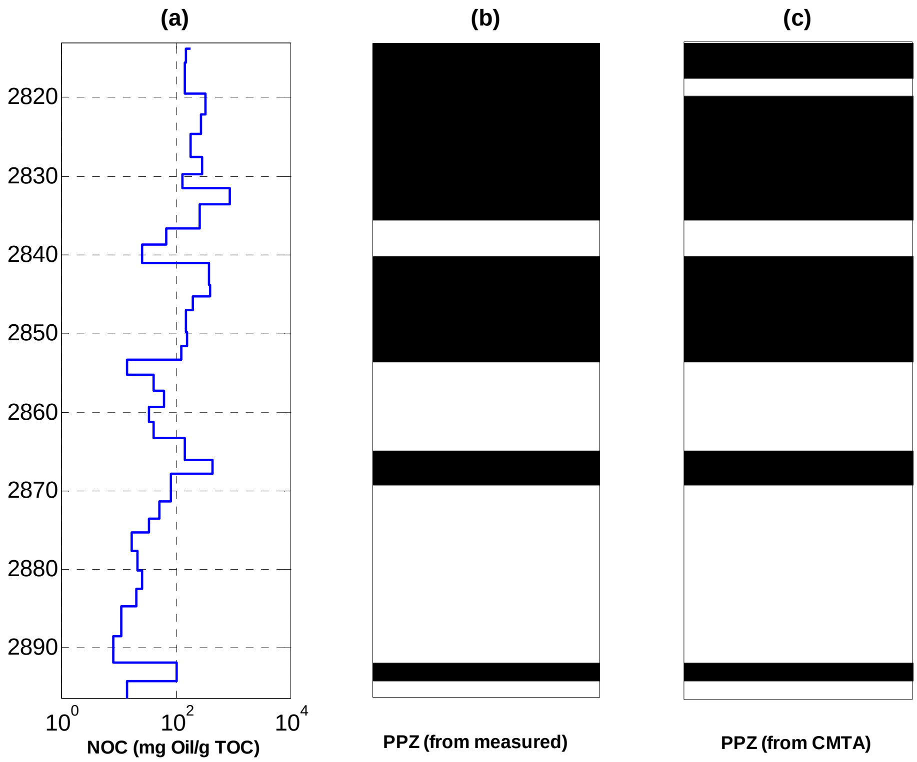

Figure 11 Graphical illustrations showing stair diagram of measured NOC at the test well **(a)**, PPZs determined from measured data **(b)**, and PPZs predicted from CMTA **(c)**, PPZs are displayed by black colors.

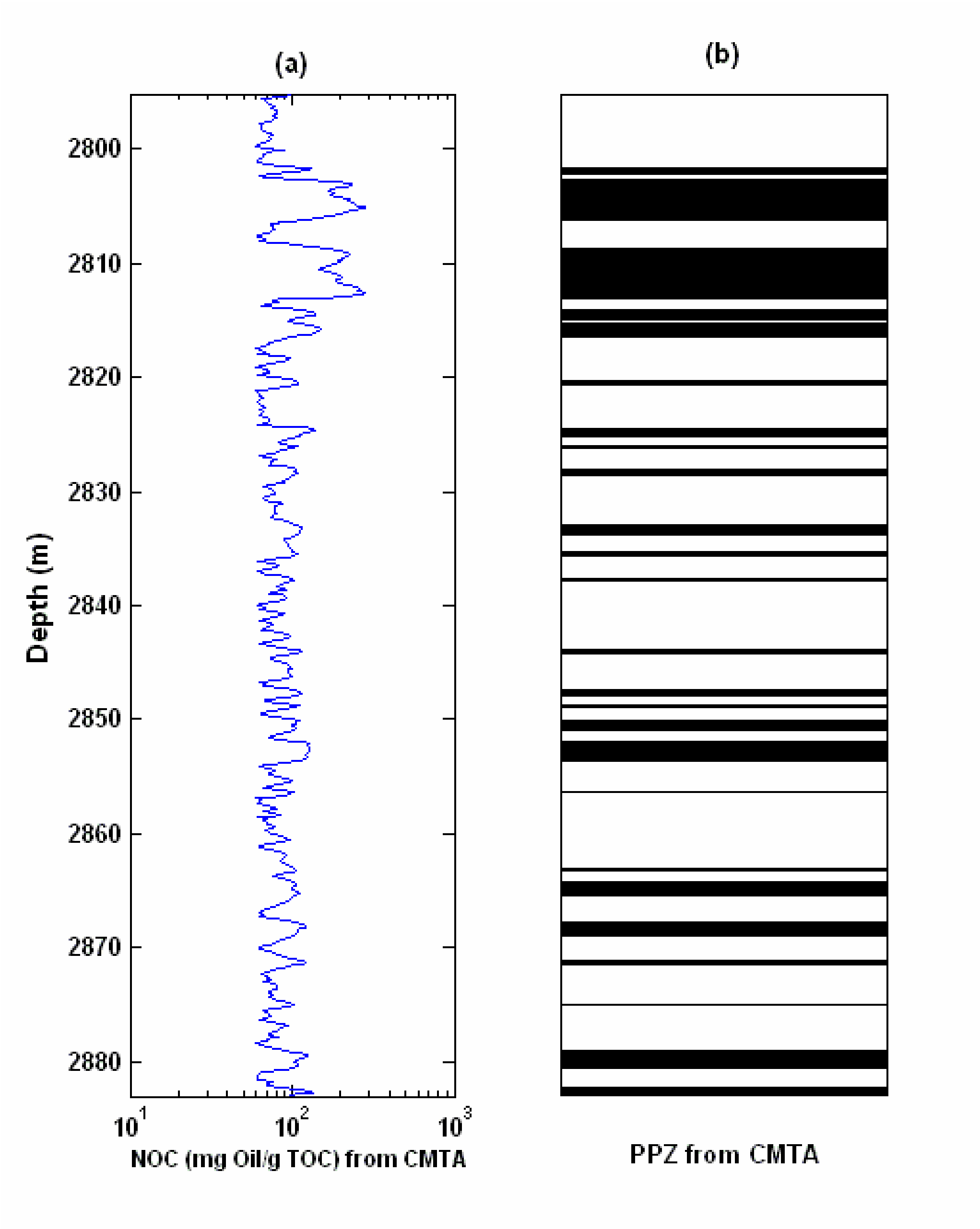

Figure 12 (a) Diagram showing the predicted NOC from CMTA at the forth well of the study field. **(b)** A graphical illustration showing PPZs determined from generalization of CMTA to the forth well, PPZs are displayed by black colors.