A general approach for porosity estimation using artificial neural network method: a case study from Kansas gas field

Sagar Singh[1], Ali Ismet Kanli[2*] and Selcuk Sevgen[3]

1 Department of Earth Sciences, Indian Institute of Technology Roorkee, 247667 Roorkee, India

2 Department of Geophysical Engineering, Faculty of Engineering, Istanbul University, Avcilar Campus, 34320 Istanbul, Turkey (kanli@istanbul.edu.tr)

3 Department of Computer Engineering, Faculty of Engineering, Istanbul University, Avcilar Campus, 34320 Istanbul, Turkey

* Corresponding author

Received: February 19, 2015; Revised: May 9, 2015; Accepted: August 4, 2015

ABSTRACT

This study aims to design a back-propagation artificial neural network (BP-ANN) to estimate the reliable porosity values from the well log data taken from Kansas gas field in the USA. In order to estimate the porosity, a neural network approach is applied, which uses as input sonic, density and resistivity log data, which are known to affect the porosity. This network easily sets up a relationship between the input data and the output parameters without having prior knowledge of petrophysical properties, such as pore-fluid type or matrix material type. The results obtained from the empirical relationship are compared with those from the neural network and a good correlation is observed. Thus, the ANN technique could be used to predict the porosity from other well log data.

Keywords: porosity estimation, artificial neural network, well log data, Kansas gas field

1. INTRODUCTION

Porosity, which is one of the most important parameters for reservoir characterization, is defined as the fraction of the total volume of a rock that is not occupied by the solid constituents. High-precision estimation of porosity is one of the challenging tasks for the oil industries.

In order to reliably characterize petroleum reservoirs, lithofacies have to be correctly determined. In general, lithofacies can be identified by making direct observation of the underground cores which are small rock samples retrieved from the borehole at selected depths of the reservoir rocks.. Well data give precise information on reservoir properties at specific field locations with high vertical resolution. On the other hand, drilling along with the coring itself is a very expensive process, therefore a lower cost method, which

Stud. Geophys. Geod., 60 (2016), 130–140, DOI: 10.1007/s11200-015-0820-2

provides the same accuracy, is desirable. Several studies have been conducted to estimate the porosity by making use of electrical logs (*Hearst, 2000*; *Al-Qahtani, 2000*; *Helle et. al., 2001*). Characterizing a reservoir is a very complex task due to its inherent heterogeneity. Distinct geological ages, depositional environments and different characteristics of rocks can be given as causes of this heterogeneity (*Arabani, 2002*). This heterogeneous structure makes it difficult to explicitly quantify a spatial relationship among variable reservoir characteristics because of its nonlinear nature.

The integration of well-log and seismic data is an essential process for reservoir characterization. This fact has been well documented in the literature, see *Gastaldi et al. (1997)*, *Russell et al. (1997)*, *Hampson et al. (2001)*. Some techniques such as map-based geostatistical methods and multi-attribute transforms have been used for porosity estimation, see *Pramanik et al. (2004)*. Using genetic algorithms is another approach that was also applied to this task, see *Dorrington and Link (2004)*.

Computer based intelligence methods like the artificial neural networks, genetic algorithms, etc., easily handle these nonlinearity problems in an efficient way (*Ouenes, 2000*; *Nikravesh and Aminzadeh, 2001*; *Nikravesh et al., 2003*; *Aminian et. al., 2005*; *Kaydani et. al., 2012*). These computational methods are able to provide more reliable values of reservoir properties because these methods are independent of the inherent uncertainties present in the borehole (*Nikravesh et. al., 2003*).

Neural networks have been successfully used to estimate porosity and permeability form well logs, see *Bhatt and Helle (2002)* and *Helle et al. (2001)*. *Russell (2004)*, has examined the relationship between seismic attributes and reservoir parameters such as porosity. He also discussed the application of the probabilistic neural network to porosity classification. He showed the disadvantage of linear methods, which is that they can solve only linear problems, whereas the disadvantage of the multi-layer perceptron approach is that the final answer is dependent of this initial guess of the weights. As with linear methods, the solution to the weights does not depend on an initial guess, but, unlike linear methods, the basis function approach can solve nonlinear problems. The advantage of using smooth regression to estimate the effective porosity by combining seismic attributes and well log data over other techniques is that it will be able to tell us how well we can predict the porosity using any model that is likely to be operating. It can delineate the complex non-linear relationships if there are any. Similar to any other statistical and mathematical models, ANN models have also some disadvantages. Having a large number of input variables is one of the most common problems for their development because they are not engineered to eliminate superfluous inputs. It is critical to devise a systematic features election scheme that provides guidance on choosing the most representative features for estimation of petrophysical parameters (*Iturrarán-Viveros and Parra, 2014*).

In this study, we applied a well known method which is called back-propagation artificial neural network (BP-ANN) to estimate the reliable porosity values from the well log data taken from the well in Finney county of Kansas gas field (Fig. 1) located in USA. The porosities were computed using combined neutron-density empirical relationship from real neutron (N_ϕ) and density (ρ_B) log data. Then, these porosities were validated by ANN technique applied to a set of well-log data (sonic, density and resistivity log).

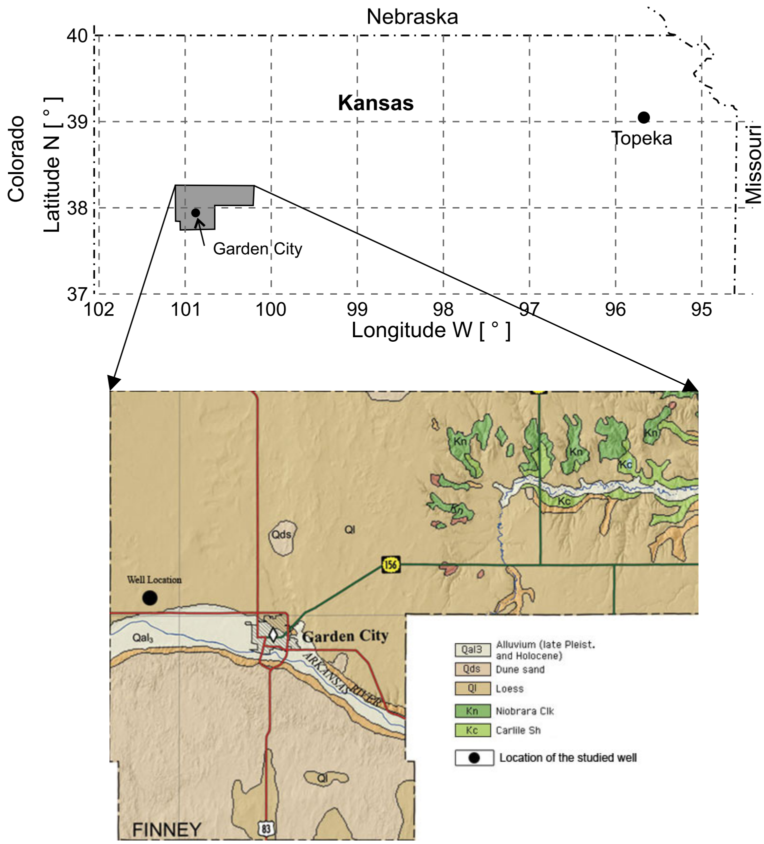

Fig. 1. Location of the study area and simplified geological map of Finney County (from Kansas Geological Survey, *http://www.kgs.ku.edu/General/Geology/County/def/finney.html*).

2. WELL LOG DATA

Porosity is affected by numerous parameters inside the borehole including fluid type and rock matrix, therefore exact calculation for the porosity using electrical logs is quite difficult. Normally it is obtained either using the wireline log data or through coring of the samples, which are time consuming and costly methods. Porosity is usually measured by using density logs and sonic logs but those are not direct measurements as well.

2.1. Sonic log

The propagation speed of a wave in the formation is calculated using the time spent to travel through the certain thicknesses of the formation and it is known as the sonic log (Fig. 2a). Porosity determined by using sonic log is given as

$$\phi_s = \frac{\Delta t_m - \Delta t}{\Delta t_m - \Delta t_f} \ , \tag{1}$$

where ϕ_s is the sonic porosity, Δt_m is the transit time of the wave inside the rock matrix, Δt is the transit time of the wave inside the rock material and Δt_f is the transit time of the wave inside the pore fluid of the borehole (*Wyllie et. al., 1956*).

2.2. Density log

The density-porosity values can be determined from the density log data (RHOB) (Fig. 2b) as (*Serra, 1984a*).

$$\phi_d = \frac{\rho_m - \rho}{\rho_m - \rho_f} \ , \tag{2}$$

where ϕ_d is the density-porosity, ρ_m is the matrix density, ρ is the bulk density and ρ_f is the fluid density.

2.3. Resistivity log

A single log cannot resolve the petrophysical properties of the formation by itself. Instead, combinations of different logs are used (*Helle et. al., 2001*). Resistivity logs which are the best indicators of the pore fluids are used for supporting the other logs in oil wells (Fig. 2c). Using resistivity log, porosity is determined from

$$\frac{R_o}{R_w} = \frac{a}{\phi^m} \ , \tag{3}$$

where ϕ is the porosity, R_o is the resistivity of the formation in clean porous aquifer, R_w is the resistivity of the interstitial water, a is the coefficient between 0.6 and 2.0 depending on the lithology, and m is the cementation factor which varies between 1 and 3 according to the type of sediment, the shape of pore, the type of porosity and the extent to the degree of compactation (*Serra, 1984b*).

2.4. Combined density-neutron porosity

The combination of density and neutron logs (Fig. 3) is commonly used to determine porosity which is free from the lithology influences to a great extent. Each of the individual log records an apparent porosity that is true only when the zone lithology matches. Limestone-equivalent porosity is a good choice for both the neutron and the density logs, because the calcite has properties that functions as an intermediate between the dolomite and the quartz. And the effects of the dolomite and the quartz tend to be cancelled out by averaging the apparent neutron- and density-porosities of a zone

$$\phi = \frac{\phi_d + \phi_n}{2} \ . \tag{4}$$

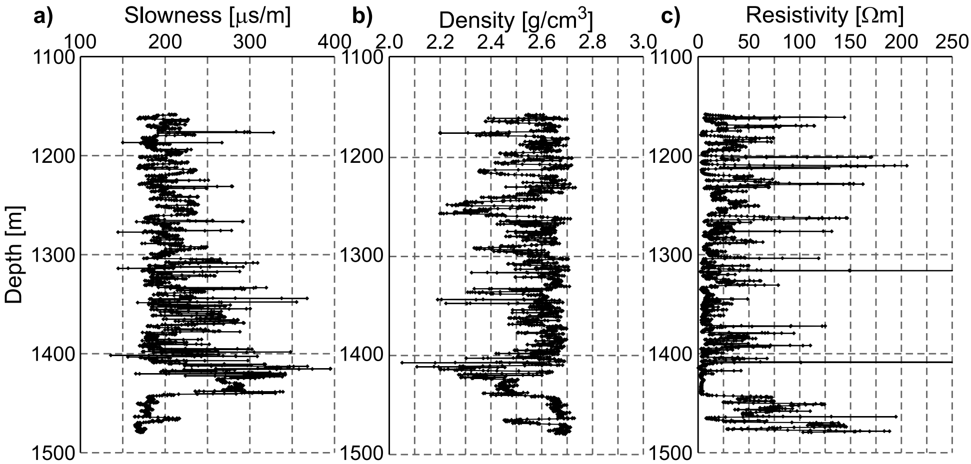

Fig. 2. **a)** Sonic log, **b)** density log, **c)** resistivity log from the studied well shown in Fig. 1.

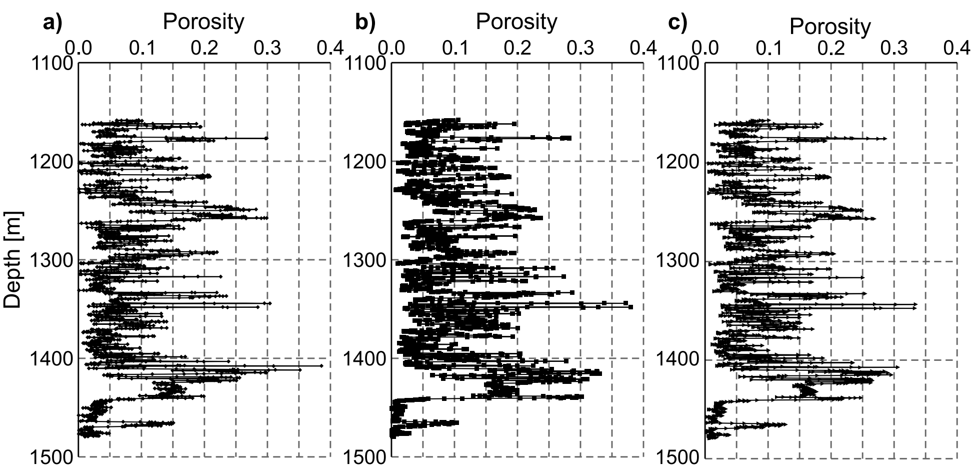

Fig. 3. **a)** Density-porosity log, **b)** neutron-porosity hydrogen index (HI) log, and **c)** combined porosity log from the studied well shown in Fig. 1.

Equation (4) is the combination of the density- and neutron-porosity used in this study. In the above equation, ϕ is the calculated porosity, ϕ_d is the density-porosity and ϕ_n is the neutron-porosity.

3. GEOLOGY OF THE STUDY AREA

The study area in Kansas gas field is shown in Fig. 1. The rocks which are found in Finney County are all sedimentary originated and their history ranges from the Upper

Cretaceous ages to the recent times. The oldest rocks on the surface layer in this area belong to the Upper Cretaceous ages and they comprise some parts of the Greenhorn limestone, Carlile shale, and the Niobrara formation. The Tertiary deposits of silt, sand, and gravel (Ogallala formation) that overlie the Cretaceous beds along most of the Finney County can be observed only on some parts of the Arkansas valley, along the Pawnee valley and, on the northern uplands of the Pawnee valley. Clay is seen above the Cretaceous beds in the northwestern and the southern Finney County.

Information about the unexposed rocks that lie beneath the Finney County has been obtained from the test holes that are drilled during the course of the investigation, from the logs of oil and gas testing wells, and from the exposures of these rocks in the nearby areas. They include shales and sandstones from the Cretaceous ages which lie under the Greenhorn limestone all over this area. Besides, they also include Paleozoic limestones and shales with lesser amounts of sandstone, gypsum, anhydrite, and salt that are found beneath the Cretaceous deposits (*Latta, 1944*).

4. NEURAL NETWORK DESIGN

In machine learning, artificial neural networks (ANN) are branches of the statistical learning algorithms inspired by the biological neural networks (the central nervous systems of animals, in particular the brain) and they are used to estimate or approximate nonlinear functions that depend on a large number of inputs which are generally unknown. The artificial neural networks are generally presented as systems of interconnected "neurons" which can compute values from inputs, and which are capable of machine learning and pattern recognition as a result of their adaptive nature. Learning of the neural network can be accomplished by making use of the supervised or unsupervised algorithms. The supervised training requires a set of known input-output data patterns (or training patterns), while the latter requires only the input patterns (*Wong and Nikravesh, 2001*). Neural networks are increasingly popular in geophysics. Since they are universal approximators, these tools can approximate any continuous function with an arbitrary precision. Hence, they are accepted to have important contributions in finding solutions in various geophysical applications (*Baan, 2000*).

4.1. Supervised networks and feed forward networks

Supervised neural networks are trained to produce the essential outputs in response to the sample inputs, making them particularly well-suited in the modeling and the controlling of a dynamic system, in the classification of the noisy data and in the prediction of the future events. In this study, feed forward neural network is used.

Feed forward networks have one-way connections from the input to the output layers. They are most commonly used for the prediction, the recognition of the pattern, and the nonlinear function fitting.

In this study, a three layered neural network model is used; the input layer, the hidden layer and the output layer (Fig. 4).

For a general neural network the input set is given as

$$X = \left[x_1 , x_2 , \ldots , x_n \right] , \tag{5}$$

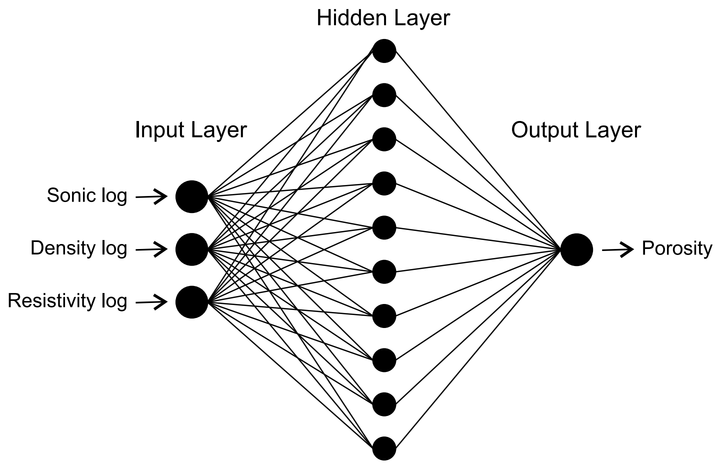

Fig. 4. The neural network design used in this study.

where x_1, x_2, …, x_n are input units. Functionally each input unit x_i is first weighed and then multiplied with the weight function w_i and then summed up. Thus, the net input in the input layer is given by (*Tanner, 1995*)

$$net = \sum_{i=1}^{n} x_i w_i \quad . \tag{6}$$

Then, activation function or transfer function is determined. A sigmoid function is used as the activation function in this study. However, it is required to add the bias value θ before the computation of activation function (*Taner, 1995*) which serves as a threshold for the activation function and therefore the output will be given by

$$O_j = \frac{1}{1 + \exp\left(-\frac{net_j + \theta_j}{\theta_0} \right)} \ , \tag{7}$$

where θ_0 is the sigmoid shaping factor. For higher values of θ_0, sigmoid will have a low slop but for the lower values, it will be steep.

The back-propagation technique for error is used which is a type of supervised method in estimating the problems of ANN. The difference between the measured and the estimated results should be below a determined level.

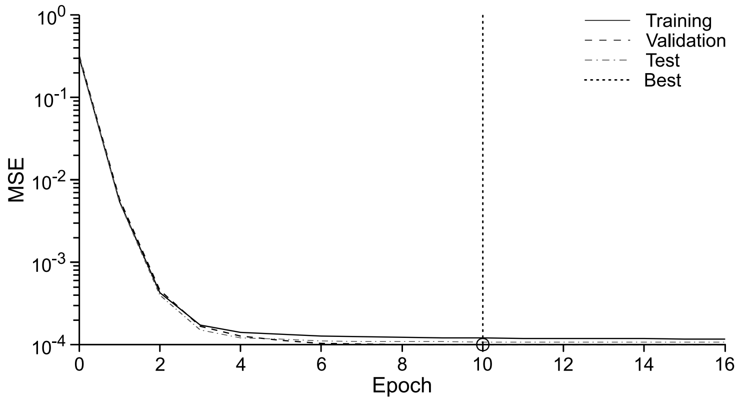

Fig. 5. Performance of the neural network. *MSE*: mean square error. The best validation performance is 0.00010196 at epoch 10.

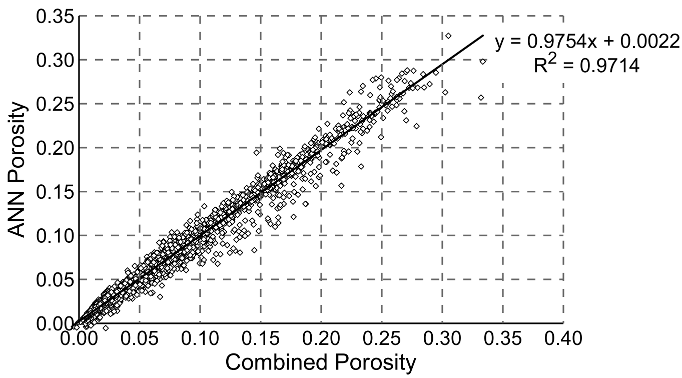

Fig. 6. Regression fit between the empirically derived combined porosity and the porosity determined using the artificial neural network (ANN).

In this study, some parts of the input data are used in training (70%), some are used in testing (15%) and the rest are used in validating. For the purpose of training, the 'trainlm' network function of MATLAB is used which updates the weight and bias values in accordance with the Levenberg-Marquardt optimization. The network validation (15%) is performed simultaneously with the network training in each course and when the validation data error begins to rise, the training stops. The mean square error (*MSE*) curve in terms of the number of the training courses (epochs) for training the data shows that the network has arrived to the best learning and the lowest error after ten iterations (Fig. 5). The correlation in both the training and testing stages is $R^2 > 0.97$. According to these results, the efficiency of the artificial neural network in the estimation of the porosity is very high.

5. RESULTS

ANN process requires three steps; the training, the testing and the validation of the network. In the training step, the network tries to acquire knowledge (information) from the given data set and it establishes a relationship between the input data set and the output parameter. After the training process, the network is tested by other data sets and finally, it is tested by the validation step in order to check the reliability of the process.

The neural network is in the process of training until it reaches global minimum error value between the empirically derived porosities and the ANN porosities and is given as the *MSE*. If Y_i are the combined density-neutron porosity values and Y_i' are the ANN calculated porosity values, then the *MSE* of n such predictions is given as (*Wackerly and Scheaffer, 2008*)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i' - Y_i\right)^2 . \tag{8}$$

This error is then propagated backwards to adjust itself to the weights and the biases of the network. When the training stage stops after *MSE* reaches its global minimum value, then the final weight and bias values is stored and can be applied for other wells.

The performance of the well is shown in the Fig. 5, which indicates that global minimum of the error is found at tenth epoch (iteration).

The regression fit between the empirically derived combined porosity and the porosity obtained from the artificial neural network is also shown in the Fig. 6. Since this correlation is very high ($R^2 > 0.97$), the same results can also be used to determine the porosity in the other wells in case the lithology does not change.

Figure 7 shows that the match between empirically derived combined porosity and the ANN porosity is nearly perfect at all depths. Thus, we believe that well-log data can be used in the ANN method to estimate porosities. Cross validation such as a blind test can better validate the ANN technique if porosity values are available.

6. CONCLUSIONS

The porosity generated from the combined density and neutron log data is compared with the one generated from the neural network approach. The high degree of the correlation between the observed porosity and the ANN derived porosity demonstrate the potential of the ANN method for the reservoir characterization problems. The neural net approach does not require an underlying mathematical model or an assumption of linearity among the variables. Its limitation is the amount of effort required to select the representative collection of training data, which is common for all models with real data. This method requires minimum of the computing time which saves time and money spent on the core sampling without any prior knowledge of matrix material or pore-fluid. Neural network designed in the study can be used to reliably estimate the porosity values in other wells of the field in case the lithology does not change and real well log data is not available.

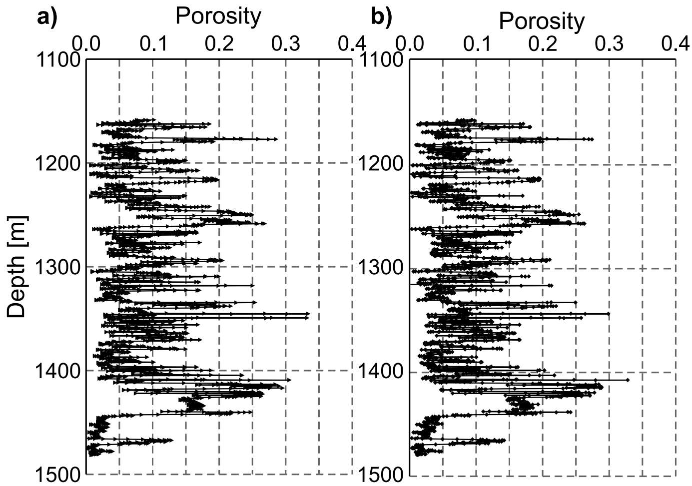

Fig. 7. **a)** Empirical derived combined porosity and **b)** ANN porosity of the studied well shown in Fig. 1.

Acknowledgments: We would like to thank Dr. G. Randy Keller and anonymous reviewers for their constructive comments in the preparation of the final form of the paper.

References

Al-Qahtani F.A., 2000. *Porosity Distribution Prediction Using Artificial Neural Networks*. MSc Thesis, Morgantown Virginia University, West Virginia.

Aminian K. and Ameri S., 2005. Application of artificial neural networks for reservoir characterization with limited data. *J. Petrol. Sci. Eng.*, **49**, 212–222.

Arabani M.S. and Bidhendi M., 2002. Porosity prediction from wireline logs using artificial neural networks: a case study in north-east of Iran. *Iranian Int. J. Sci.*, **3**, 221–233.

Baan M. and Jutten C., 2000. Neural networks in geophysical applications. *Geophysics*, **65**, 1032–1047.

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

Dorrington K.P. and Link C.A., 2004. Genetic-algorithm/neural-network approach to seismic attribute selection for well-log prediction. *Geophysics*, **69**, 212–221.

Gastaldi C., Biguenet J. and Pazzis L.D., 1997. Reservoir characterization from seismic attributes: an example from the Peciko field (Indonesia). *Leading Edge*, **16**, 263–266.

Hearst J.R., Nelson P.H. and Paillet F.L., 2000. *Well Logging for Physical Properties*. John Wiley and Sons Ltd., New York.

Helle H.B., Bhatt A. and Ursin B., 2001. Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study. *Geophys. Prospect.*, **49**, 431–444.

Hampson D.P., Schuelke J.S. and Quierin J.A., 2001. Use of multiattribute transforms to predict log properties from seismic data. *Geophysics*, **66**, 220–236.

Iturrarán-Viveros U. and Parra J.O., 2014. Artificial neural networks applied to estimate permeability, porosity and intrinsic attenuation using seismic attributes and well-log data. *J. Appl. Geophys.*, **107**, 45–54.

Kaydani H., Mohebbi A. and Baghaie A., 2012. Neural fuzzy system development for the prediction of permeability from wireline data based on fuzzy clustering. *Petrol. Sci. Technol.*, **30**, 2036–2045.

Latta F.B., 1944. *Geology and Ground-Water Resources of Finney and Gray Counties, Kansas*. Kansas Geological Survey Bulletin **55** (http://www.kgs.ku.edu/General/Geology/Finney/index.html).

Nikravesh M. and Aminzadeh F., 2001. Past, present and future intelligent reservoir characterization trends. *J. Petrol. Sci. Eng.*, **31**, 67–79.

Nikravesh M., Aminzadeh F. and Zadeh L.A., 2003. *Soft Computing and Intelligent Data Analysis in Oil Exploration*. Developments in Petroleum Sciences **51**. Elsevier, Amsterdam, The Netherlands.

Ouenes A., 2000. Practical application of fuzzy logic and neural networks to fractured reservoir characterization. *Comput. Geosci.*, **26**, 953–962.

Pramanik A.G., Singh V., Vig R., Srivastava A.K. and Tiwary D.N., 2004. Estimation of effective porosity using geostatistics and multiattribute transforms: a case study. *Geophysics*, **69**, 352–372.

Russell B., Hampson D., Schuelke J. and Quirein J., 1997. Multiattribute seismic analysis. *Leading Edge*, **16**, 1439–1443.

Russell B.H. 2004. *The Application of Multivariate Statistics and Neural Networks to the Prediction of Reservoir Parameters Using Seismic Attributes*. PhD Thesis, University of Calgary, Calgary, Alberta, Canada.

Serra O. 1984a. Formation density measurements (the gamma-gamma log or density log). In: Serra O. (Ed.), *Fundamentals of well-log interpretation.* Developments in Petroleum Sciences **15A**. Elsevier, Amsterdam, The Netherlands, 195–204.

Serra O. 1984b. The measurement of resistivity. In: Serra O. (Ed.), *Fundamentals of well-log interpretation.* Developments in Petroleum Sciences **15A**. Elsevier, Amsterdam, The Netherlands, 51–76.

Taner M., 1995. Neural networks and computation of neural network weights and biases by the generalized delta rule and back-propagation of errors. *Rock Solid Images* (http://www.rocksolidimages.com/pdf/neural_network.pdf).

Wackerly D. and Scheaffer W., 2008. *Mathematical Statistics with Applications*. 7th Edition. Thomson Brooks/Cole, Duxbury, MA.

Wong P.M. and Nikravesh M., 2001. Introduction: field applications of intelligent computing techniques. *J. Petrol. Geol.*, **24**, 381–387.

Wyllie M.R.J., Gregory A.R. and Gardner L.W., 1956. Elastic wave velocities in heterogeneous and porous media. *Geophysics*, **21**, 41–70.