LITHOLOGY PREDICTION BY ARTIFICIAL NEURAL NETWORKS AND PREPARATION OF INPUT DATA ON UPPER MIOCENE SEDIMENTS

Marko **Cvetković**, Josipa **Velić**

Faculty of Mining, Geology and Petroleum Engineering, Department of Geology and Geological Engineering, Zagreb, Croatia

e-mail: marko.cvetkovic@rgn.hr

ABSTRACT

Well logs of spontaneous potential and of shallow and deep resistivity were pre-processed for the purpose of lithology prediction. Several artificial neural networks were trained on data from one well for the purpose of lithology prediction in a second. Data were taken from Upper Miocene intervals from two wells in the Kloštar field. Two learning approaches and three methods of prediction were applied. The results show that the best approach for inter-well lithology prediction is by training the neural network on an entire well interval, instead of training a separate neural network for each interval and subsequently employing them as an ensemble for the prediction.

Keywords: artificial neural networks, lithology prediction, Miocene, Croatia

1. Introduction

Well log data from the Kloštar oil and gas field, located in the southwest part of the Pannonian Basin, in the Sava Depression (Fig. 1), were used to test the application of neural networks for lithology prediction in the sandstone and marl intervals of Upper Miocene sediments. One dimensional (1D) lithology prediction, e.g., training of the artificial neural network (ANN) and prediction, has been proven successful only if both are made in the same well and on the same formation (Cvetković et al., 2009). However, for practical purposes this is unsatisfactory and hence, the success of two dimensional (2D) prediction

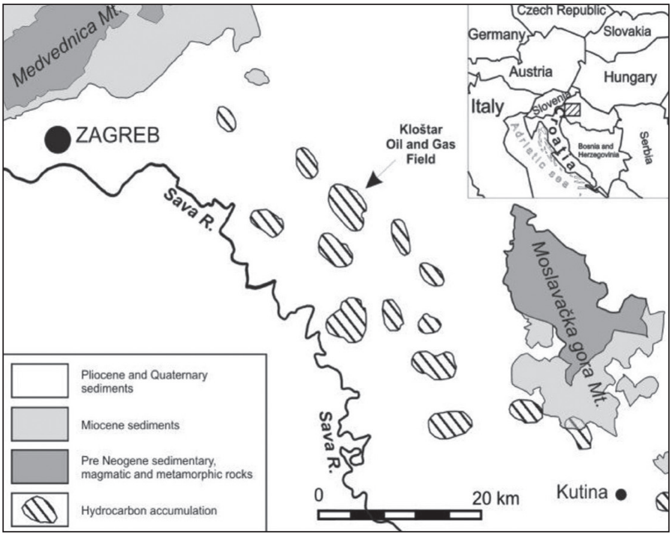

Fig. 1 – Location of Kloštar oil and gas Field (Velić et al., 2011)

is presented, which is performed on a much larger depth interval and on two wells. Additionally, two types of learning procedures are introduced: one for the entire well interval and the other for each clastic Neogene formation individually.

2. Basic geological settings

Analysis with the ANN is limited to just the clastic Upper Miocene intervals (Fig. 2). These are represented by the sediments of the Prkos, Ivanić-Grad, Kloštar Ivanić and Široko Polje Formations.

The lithology can be described as an interchange of marls and sandstones in the Prkos, Ivanić-Grad, Kloštar Ivanić and Široko Polje Formations with sporadic occurrences of breccias and biogenic limestones in the oldest parts of the Prkos Formation. Furthermore, lignites, sand and clays can found in the youngest interval of the Široko Polje Formation.

The youngest clastic interval of the Lonja Formation (Pliocene, Pleistocene and Holocene) was not included in the analysis because of its small thickness (<150 m) in the selected locations.

3. Artificial neural networks

The basic architecture of ANNs consists of neurons that are organised into layers. A neuron is a basic

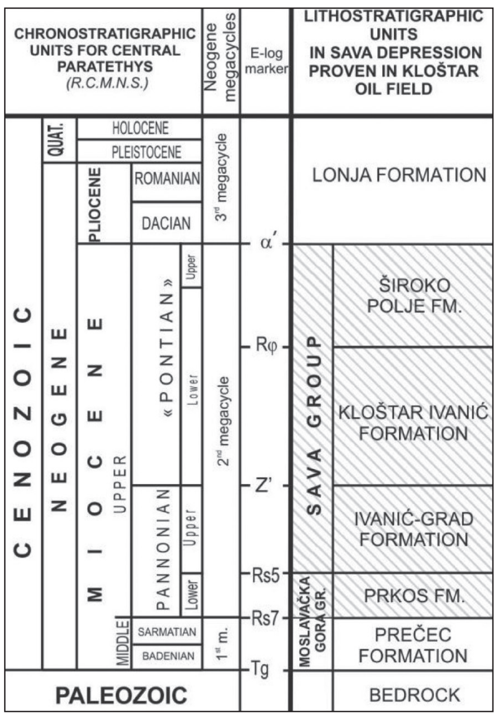

Fig. 2 – Formal lithostratigrapic and chronostratigraphic units and E-log markers valid for the Sava Depression (Velić, 2002) with pointed out Formations used for ANN prediction of lithology

element of a network that is mathematically presented as a point in space towards which signals are transmitted from surrounding neurons or inputs (Fig. 3).

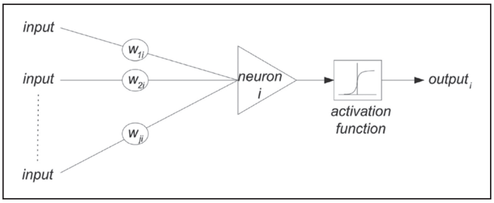

Fig. 3 – Basic architecture of an artificial neuron

The value of a signal on the activity of a neuron is determined by a weighting factor multiplied by a corresponding input signal. The total input signal is determined as a summation of all products of weighting factors multiplied by the corresponding input signals, given by

$$u_i = \sum_{j=1}^{n} (w_j \cdot input_j) \qquad (1)$$

where n represents the number of inputs for the neuron i. If the total input signal has a value greater than the sensitivity threshold of a neuron, then it will have an output of maximum intensity. Alternatively, if a neuron is inactive, it has no output. The value of the output is given by

$$output_i = F(u_i \cdot t_i) \qquad (2)$$

where F represents the activation function and t_i is the targeted output value of neuron i. One can find detailed descriptions of the basics of neural networks and methods in McCulloch and Pitts (1943), Rosenblatt (1958) and Anderson and Rosenfeld (1988).

For this analysis, only a multilayer perceptron (MLP) network was used. The MLP network is based on a backpropagation algorithm that has one or more hidden layers. In addition, the network can use two-phase learning with algorithms, such as the conjugate gradient descent (Gorse et al., 1997), quasi-Newton (Bishop, 1995), Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1963), quick propagation (Fahlman, 1988) and delta-bar-delta (Jacobs, 1988), all of which use a program of packages that allow such training. The MLP is most successfully applied in classification and prediction problems (Rumelhart et al., 1986) and is the most often used neural network in solving geological problems. Neural networks have been successfully applied in petroleum geology problems, such as determining reservoir properties (e.g., lithology and porosity) from well logs (Bhatt, 2002) and well-log correlations (Luthi, Bryant, 1997). In the Croatian part of the Pannonian Basin, only a few research projects on petroleum geology have been performed. In these studies, clastic facies were determined from well logs (Malvić, 2006, Cvetković et al., 2009) and porosity was predicted based on well and seismic data (Malvić, Prskalo, 2007).

4. Preparation of well log data for ann analyses

Any analysis with geomathematical or statistical methods requires that the numerical values for the same variables have a constant interval. The main problem is the use of spontaneous potential (SP) values for the ANN analysis, or more precisely, the inconsistency in mV values for the sandstone variable. Values of the SP curve can differ by more than 70% for the marl line, which determines the border

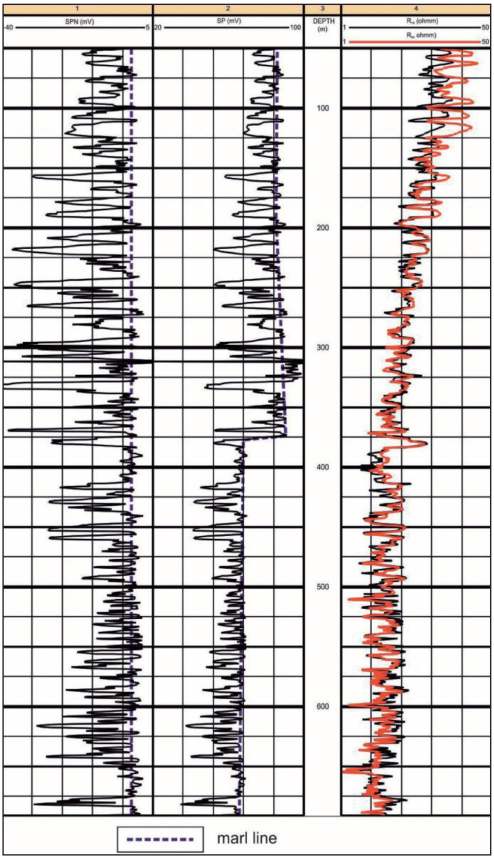

Fig. 4 – Presentation of the differences SP values along with the representation of the presumed position of the marl line in SP and SPN logs

between the sandstones and marls (Fig. 4).

Such differences can be caused by:

– heat distribution in the subsurface,

– the salinity ratio between mud and formation water and,

– settings on the logging tool.

To solve the issue, a normalised SP curve (SPN) should be made prior to performing any ANN analysis.

5. Methodology and analysis results

Initially, for the purpose of the ANN training, the lithology was manually determined in one well (Klo-105) based on well log curves of normalised spontaneous potential (SPN), shallow resistivity (R16) and deep resistivity (R64). These were also used as input data for the ANN analysis.

Two types of data preparation were used. In the first approach, the data were split for each formation and training was performed for each formation individually. In the second, the entire well interval was used for ANN training.

The ANN learning process splits the input data into two groups: the test and validation groups. The first set of data is used for learning purposes and the second is used for testing the prediction possibilities of the trained ANN. Accordingly, two sets of performance values are given, one for the training and one for the test. Performance values range from 0–100, where 100 represents a perfect match or a prediction accuracy of 100%.

The ANN that was most successful for each formation and for the entire well interval, together with its properties, is presented in Table 1.

Table 1 – Overview of the most successful neural networks for each formation and for the whole Miocene clastic interval

	Type	Train	Test
Whole interval	MLP 3-10-3	97.38641	97.37336
Široko Polje Formation	MLP 3-40-2	99.43228	99.23664
Kloštar Ivanić Formation	MLP 3-19-2	94.02460	93.26923
Ivanić-Grad Formation	MLP 3-10-2	99.65636	100.0000
Prkos Formation	MLP 3-33-2	97.66082	97.05882

The data presented in Table 1 show that all neural networks had very low training and test errors; less than 7% of maximum training error in the case of the Kloštar Ivanić Formation. The next step was to employ the trained networks on a new set of data; in this case, on the Miocene clastic interval of well Klo-106. Three types of lithology predictions were applied:

Prediction by a single ANN for the entire clastic Miocene interval, labelled LITH A in Fig. 5

Prediction by a single ANN individually for each formation, labelled LITH B in Fig. 5

Prediction by an ensemble of ANNs trained on each formation for the prediction for the entire clastic Miocene interval, labelled LITH C in Fig. 5.

Even though all the ANNs had very low training errors, only the one that was trained with the entire Miocene clastic interval (LITH A in Fig. 5A) successfully predicted the lithology in well Klo-106. The individual networks for each Formation, as well as the ensemble, were only partially successful in the Kloštar Ivanić Formation (LITH B and C in Fig. 5B).

6. Conclusions

For any possibility of the employment of ANNs for 2D lithology prediction, the SP curve for the input data has to be normalised. Inter-well prediction of lithology via ANNs is possible and successful only in cases where the ANN is trained on a large interval. Possibly, ANNs trained on each formation were over-trained and lost their ability to successfully predict lithology on new data, even though their training and testing errors were minimal.

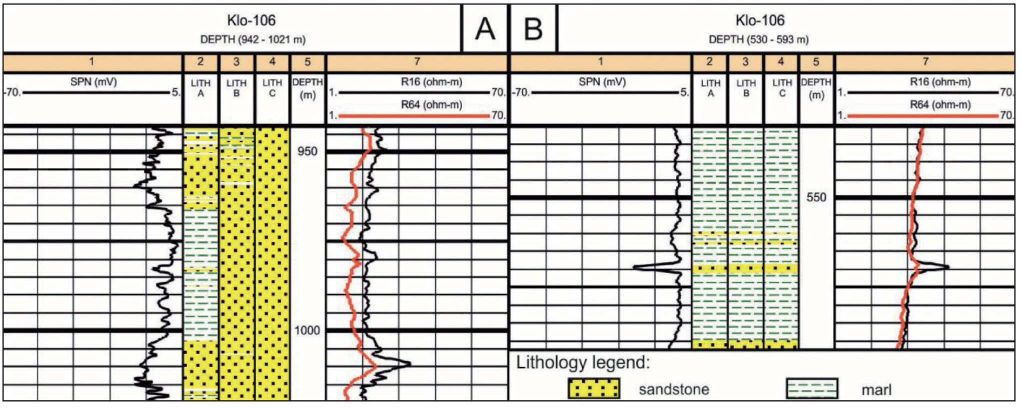

Fig. 5 – Well log section showing the poor results of lithology prediction with ANN's trained on a single formation (LITH B) and by an ensemble (LITH C) in Ivanić-Grad Formation (A) and successfulness of all three in Kloštar Ivanić Formation (B) (Cvetković, Velić, 2012)

Acknowledgements

This work is part of a research project "Stratigraphical and geomathematical researches of petroleum geological systems in Croatia" (project no. 195-1951293-0237), financed by the Ministry of Science, Education and Sports of the Republic of Croatia.

The authors would also like to thank INA-Industry of Oil Plc., E&P of Oil and Gas (Zagreb) for the data and Schlumberger and Synergy for supplying academic licenses for Petrel and Interactive Petrophysics to the Faculty of Mining, Geology and Petroleum Engineering in Zagreb.

References

Anderson, J.A., Rosenfeld, E. (1988): Neurocomputing**:** Foundations of Research. MIT Press, Cambridge, 729 p.

Bhatt, A. (2002): Reservoir properties from well logs using neural networks: doctoral dissertation, NTNU, Trondheim, 151 p.

Bishop, C.M. (1995): Neural Networks for Pattern Recognition, Oxford University Press, New York, 504 p.

Cvetković, M., Velić, J., Malvić, T. (2009): Application of neural networks in petroleum reservoir lithology and saturation prediction. Geologia Croatica, **62/2**, 115–121.

Cvetković, M., Velić, J. (2012): Successfulness of inter well lithology prediction on Upper Miocene sediments with artificial neural networks. In: Malvić, T., Geiger, J., Cvetković, M. (Eds.): Geomathematics as geosciences. Hrvatsko geološko društvo, Zagreb, 13–20.

Fahlman, S.E. (1988): Faster Learning Variations on Back-propagation: An Empirical Study. In: Proceedings of the 1988 Connectionist Models Summer School, Los Altos, 38-51.

Gorse, D., Shepherd, A., Taylor, J.G. (1997): The new ERA in supervised learning. Neural Networks, **10/2**, 343–352.

Jacobs, R.A. (1988): Increased rates of convergence through learning rate adaptation. Neural Networks, **1**, 295–307.

Levenberg, K. (1944): A method for the solution of certain problems in least squares. Quardt. Appl. Math., **2**, 164–168.

Luthi, S.M., Bryant, I.D. (1997): Well-log correlation using a back-propagation neural network. Math. Geol., **29/3**, 413–425.

Malvić, T. (2006): Clastic facies prediction using neural network (case study from Okoli field). Nafta, **57/10**, 415–431.

Malvić, T., Prskalo, S. (2007): Some benefits of the neural approach in porosity prediction (Case study from Beničanci field). Nafta, **58/9**, 455–467.

Marquardt, D.W. (1963): An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Indust. Appl. Math., **11/2**, 431–441.

McCulloch, W.S., Pitts, W. (1943): A logical calculus of ideas immanent in nervous activity. B. Math. Biol., **5/4**, 115–133.

Rosenblatt, F. (1958): The perceptron: A probabilistic model for information storage and organization in the brain. Psychol. Rev., **65/6**, 386–408.

Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986): Learning internal representations by error propagation. In: Rumelhart, D.E., McClelland, J.L. (Eds.): Parallel Distributed Processing, **1**. MIT Press, Cambridge, 381–362.

Velić, J., Weisser, M., Saftić, B., Vrbanac, B., Ivković, Ž. (2002): Petroleum-geological characteristics and exploration level of the three Neogene depositional megacycles in the Croatian part of the Pannonian basin. Nafta, **53/6–7**, 239–249.

Velić, J., Malvić, T., Cvetković, M. (2011): Palinspastic reconstruction of synsedimentary tectonics of Neogene and Quaternary sediments in the Kloštar Field (Sava Depression, Pannonian Basin, Croatia). Zeitschrift der Deutschen Gesellschaft für Geowissenschaften, **162/2**, 193–203.