Geologia Croatica	62/2	115–121	6 Figs.	2 Tabs.		Zagreb 2009

Application of neural networks in petroleum reservoir lithology and saturation prediction

Marko Cvetković[1], Josipa Velić[1] and Tomislav Malvić[1,2]

[1] Department of Geology and Geological Engineering, Faculty of Mining, Geology and Petroleum Engineering, Pierottijeva 6, 10000 Zagreb, Croatia; (marko.cvetkovic@rgn.hr)

[2] Development Department, Oil & Gas Exploration and Production, INA-Industrija nafte d.d., Šubićeva 29, 10000 Zagreb, Croatia; (tomislav.malvic@ina.hr)

doi: 10.4154/gc.2009.10

Geologia Croatica

ABSTRACT

The Kloštar oil field is situated in the northern part of the Sava Depression within the Croatian part of the Pannonian Basin. The major petroleum reserves are confined to Miocene sandstones that comprise two production units: the Lower Pontian I sandstone series and the Upper Pannonian II sandstone series. We used well logs from two wells through these sandstones as input data in the neural network analysis, and used spontaneous potential and resistivity logs (R_{16} and R_{64}) as the input in network training. The first analysis included prediction of lithology, which was defined as either sandstone or marl. These two rock types were assigned categorical values of 1 or 0 which were then used in numerical analysis. The neural network was also used to predict hydrocarbon saturation in selected wells. The input dataset was extended to depth and categorical lithology. The prediction results were excellent, because the training and prediction dataset showed little disagreement between the true and predicted values. At present, this study represents the best and most useful application of neural networks in the Croatian part of the Pannonian Basin.

Keywords: Kloštar field, neural network, prediction, sandstone, hydrocarbon saturation, Croatia

1. INTRODUCTION

We used data from the Kloštar oil field in Croatia to test the application of neural networks as a part of a project with a Croatian oil company (INA). The only other geomathematical tools used on this oil filed were several interpolation methods used for porosity mapping (BALIĆ et al., 2008). The field is located approximately 35 km east of the Croatian capital of Zagreb (Fig. 1). This particular field was selected because it is part of a joint research project between the Faculty of Mining, Geology and Petroleum Engineering, and a Croatian oil company. We used neural network analysis to predict reservoir lithology of the I and II sandstone series as either shale or sandstone, as well as the hydrocarbon saturation of the sandstone intervals. These intervals are generally represented by clastic, brackish to freshwater deposits that are characteristic of the Upper Pannonian and Lower Pontian succession throughout the Croatian part of the Pannonian Basin (LUČIĆ et al., 2001).

2. PETROLEUM GEOLOGY SETTINGS

The Dinaric oriented (NW–SE) Križ structure (including the Kloštar field) is located at the most northwestern part of Moslavačka Gora Mountain. Despite many available well data, the borders of the stratigraphic units are commonly not precisely defined, mostly because of a lack of palaeontological samples and complex tectonics resulting in many tectonic blocks. At favorable locations, stratigraphic boundaries are determined based on available well data, including cores, mud chips, and logs. The following five units are defined and de-

Figure 1: Geographic position of the Kloštar field in Croatia.

scribed below, based on these boundaries: Palaeozoic, Middle Miocene, Upper Miocene, Pliocene, and Quaternary sediments.

The **basement of the Tertiary system** represents the core of the Križ structure. It includes extrusive igneous and metamorphic rocks, including granites of Palaeozoic age (VRAGOVIĆ & MAYER, 1980). This formation is a buried hill, formed by radial tectonic movements and denudation processes that occurred before the Miocene epoch. The rocks are weathered and fractured. In structurally favourable places, hydrocarbon accumulations are confirmed.

The **Middle Miocene** (Badenian, Sarmatian) unit unconformably overlies the Palaeozoic igneous–metamorphic complex. The basal part is characterized by coarse-grained conglomerates, conglomeratic sandstones, and sandstones often intercalated with shale. It is overlain by dark-gray, sandy, bituminized marls, partially intercalated with light-gray, fine-grained sandstones. Miocene beds form economic hydrocarbon reservoirs at the southern and eastern parts of the Kloštar structure.

Upper Miocene (Pannonian, Pontian) strata are well documented throughout the field area. Lower Pannonian strata conformably overlie Sarmatian bituminized marls. Upper Pannonian sediments are represented by predominantly brown or dark-gray calcareous marls and sandstones, located in the southwestern part, and partially saturated by hydrocarbons. They are defined as the II sandstone series. Lower Pontian sediments are represented by dark-gray, massive marls and sandy marls to the south and east of the field area. Sandstones, mostly arenites with minor proportions of marl or clay, are dominant in the northwest and are partially saturated with hydrocarbons. The Lower Pontian sandstones are also called the II sandstone series.

The Upper Pontian succession is monotonous, consisting of soft sandy or clayey sediments and the proportion of sand increases upward.

The **Pliocene** deposits, i.e., Dacian and Romanian, are also locally known as the Paludina beds. These sediments are characterized by the alternation of clays and medium- and coarse-grained sands.

Quaternary deposits consist predominantly of yellowish sandy clays with abundant lime concretions. The average thickness ranges between 10 and 15 m.

The Kloštar structure is a faulted anticline with Dinaric strike (NW–SE). The geological history appears interesting, as shown by investigations of structural settings and the tectonic evolution of the western part of the Sava Depression (VELIĆ, 1979, 1980, 1983). The field structure was formed in the Middle Miocene, when intensive Badenian and Sarmatian uplifting events resulted in formation of a NW to SE-oriented anticline of 7×2 km. Later, in the Late Miocene, this structure was differentiated in two smaller parts: the northern, which was uplifted during the Pontian, and the southern, which was only activated during the Upper Pontian. The recent structural shape was tectonically created during the Pliocene and Quaternary, when the main phase of hydrocarbon migration probably occurred.

3. ARTIFICIAL NEURAL NETWORKS

A neuron is a basic element of a network that is mathematically presented as a point in space toward which signals are transmitted from surrounding neurons (Fig. 2).

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

$$u_i = \sum_{j=1}^{n} (w_{ji} \cdot input_j),$$

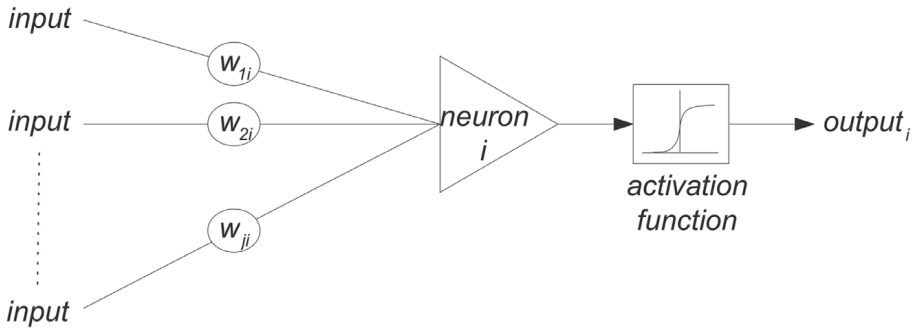

Figure 2: Artificial neuron model.

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, a neuron is inactive and has no output. Value of the output is given by

$$output_i = F(u_i \cdot t_i),$$

where F represents the activation function and t_i, the targeted output value of neuron i. One can find a more detailed description of neural network basics and methods in MCCULLOCH & PITTS (1943), ROSENBLATT (1958) and ANDERSON & ROSENFELD (1988).

The basic architecture of a neural network consists of neurons divided into layers. Three is a minimum number of layers which a neural network has to have. These layers are the input layer, the hidden layer and the output layer. The input layer is for accepting signals, or in our case, input variables. These data are transferred to the hidden layer where it is processed by the activation function belonging to the neurons within it. Data which proved to be significant in the analysis in the hidden layer neurons is sent to the output layer as the resulting data for the predicted variable. The number of hidden layers can be more than one or strictly one, depending on the type of the neural network. For example, the multi layer perceptron (MLP) neural network is basically designed to be able to have more than one hidden layer and to perform better with two or three hidden layers than with one. In opposition to the multi layer perceptron, the radial basis function (RBF) neural network can only have one hidden layer but the number of neurons within this layer is much larger than the number of neurons in the single hidden layer of the MLP.

For this analysis, we used the two aforementioned types of neural networks, the supervised learning-multilayer perceptron and the radial basis function neural network. The MLP network is based on a back propagation algorithm which calculates the error surface gradient in each step of the analysis. In the following step, the weight factors are adjusted according to the earlier calculated error surface gradient so the error minimizes and a new error surface is calculated. Also, the network can utilize a two-phase learning with second learning algorithms such as 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) which basically work on similar principles to the back propagation algorithm with a somewhat different approach. The greatest advantage of these aforementioned algorithms over the back propagation is that they are significantly faster but sometimes the standard back propagation algorithm gives the best results. For more information in Croatian on these learning algorithms please refer to the geostatistical dictionary of MALVIĆ et al. (2008).

The MLP is more successfully applied in classification and prediction problems (RUMELHART et al., 1986), and is the most often used neural network in solving geological problems. The RBF network is also a commonly used neural network but is more successfully and frequently applied in solving classification problems than in solving prediction 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 correlation (LUTHI & BRYANT, 1997). In the Croatian part of the Pannonian Basin, only a few petroleum geology research projects have been performed. In these studies, clastic facies were determined from well logs (MALVIĆ, 2006) and porosity was predicted based on well and seismic data (MALVIĆ & PRSKALO, 2007).

4. DATA ANALYSIS

We used well log data from two wells as input data for the neural network analysis. The data consist of the most basic well logs because measurements were taken in 1956 (well Klo–A) and 1957 (well Klo–B). Although these logs were taken some 50 years ago, we can successfully use them for interpretation (Fig. 3). Available data for analysis included resistivity (R_{16} and R_{64}) and spontaneous potential (SP) logs. Resolution of the data was 10 cases (measurements) per metre of well log.

We performed two types of analysis. First we predicted the lithology, followed by the hydrocarbon saturation. We carried out the analysis such that the neural network was first trained on a specific interval of well log data (overseen learning), and afterward we used the trained neural network to predict the value of desired parameters for the intervals on which the neural network was not trained.

All neural network analysis were made using StatSoft STATISTICA 7.1.

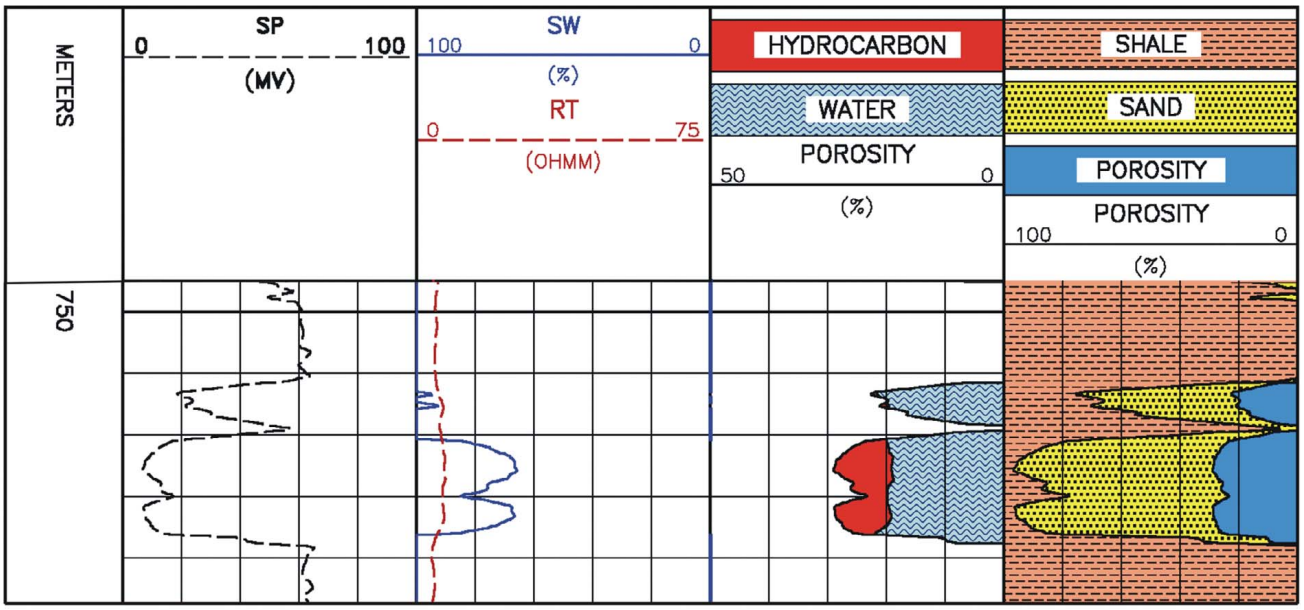

Figure 3: Part of the Klo–B well log representing the interval for the I sandstone series.

4.1. Lithology prediction

To predict lithology, we manually determined the lithological component by distinguishing layers of marl and sandstone from well logs (BASSIOUNI, 1994) on wells Klo–A and Klo–B. The neural network was trained on the first set of data, which includes intervals that correspond to the I sandstone series, and the prediction was made on the intervals that correspond to the II sandstone series and vice versa. Input data for training of the neural network, resistivity (R_{16}, R_{64}) and SP logs, were used. We defined lithology as a categorical variable (1 for sandstone, 0 for marl). For training of the neural network, we used manually determined lithology. Results described in the tables represent the success of the neural network training. Values are shown as training error and prediction error. The program in which the analyses were made automatically divides the training dataset into two parts, the length of which is user defined. The first set is used for training of the neural network; the second, for testing the neural network's ability to predict cases. This kind of data distribution minimizes the risk of overtraining. The training procedure is stopped when user-defined conditions have been met (final number of iterations or desired amount of error) or when the program decides that further training will no longer yield better results. Results of neural network training for the prediction of lithology are given in Table 1 and expressed as two error values in percentages. The training error corresponds to the previously mentioned dataset that was used for training the neural network. The selection error describes the neural network's success for predicting the values on unknown data. Two neural networks were trained for each well, one MLP and one RBF network. Here the prediction could only be done on the deeper or shallower parts of the well log in the same well; cross-prediction or 2D neural network analysis did not yield satisfactory results here because of the different values of SP logs in the two wells.

Table 1: Neural network parameters for lithology prediction

Neural network type and properties[a]	Well	Training error[b]	Selection error[b]
RBF 3–31–1	Klo-A	0.152942	0.172753
MLP 3–4–6–3–1	Klo-A	0.31438	0.133478
RBF 3–13–1	Klo-B	0.156621	0.149185
MLP 3–6–4–2–1	Klo-B	0.255012	0.214935

[a] Neural network type and properties correspond to the type of network and number of neurons per layer where first and last number represent the properties of the input and output layer. Values between these represent the number and properties of the hidden layers.

[b] Error value ranges from 0 to 1, where 0 represents 100% success of prediction, i.e., no error.

Table 1 shows that both of the neural networks have been successfully trained on the corresponding interval. The anomaly is shown in Table 1, where the MLP network showed a high training error but low selection error. Initially one might presume that the RBF network with significantly lower training error is more successful in predicting the unknown data interval, but this is not the case. The MLP network had slightly better results than the RBF, so we conclude that the value of the selection error is a much more reliable indicator of success of a neural network than the training error is. Thus, when neural network training is finished, the best network parameters are the ones with the smallest selection errors. The relationship between manually determined lithology and lithology gained from neural network analysis is shown in Figs. 4 and 5.

4.2. Hydrocarbon saturation prediction

As opposed to lithology prediction, hydrocarbon saturation prediction uses cross-prediction. The neural network is trained on one well log interval, much larger than in the former prediction case. Training had been done on one well, Klo–A, and prediction had been performed on another well,

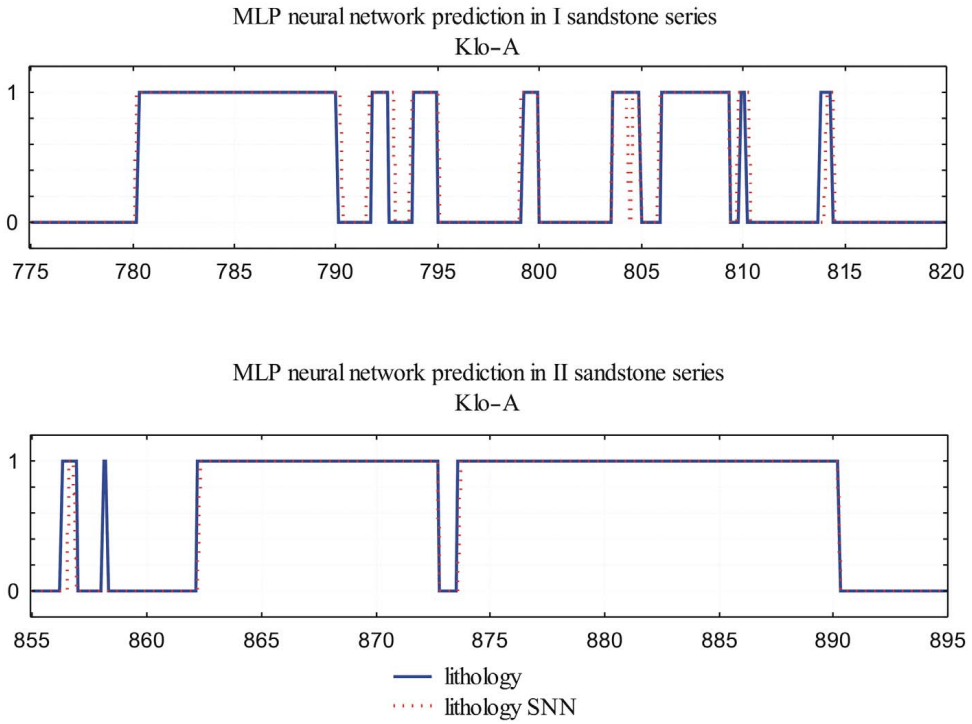

Figure 4: Comparison of MLP neural network predicted (dotted line) and manually determined data (solid line) in lithology prediction analysis for well Klo–A. The diagram's abscissa represents vertical depth, and the ordinate represents the value of lithology expressed as either marl (0) or sandstone (1).

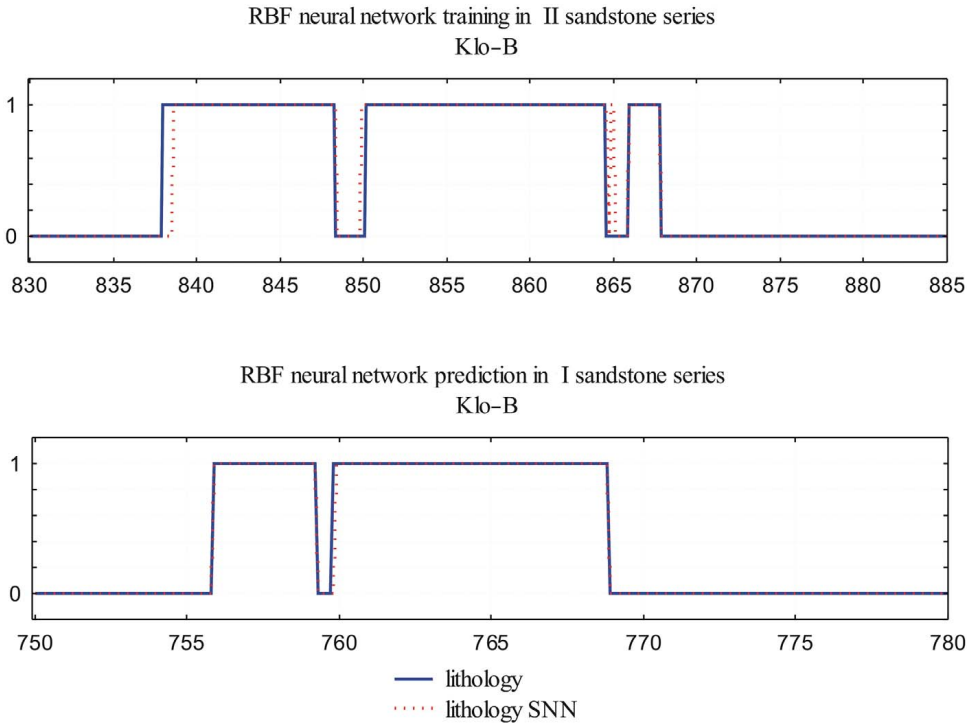

Figure 5: Comparison of RBF neural network predicted (dotted line) and manually determined data (full line) in lithology prediction analysis for well Klo–B. The diagram's abscissa represents vertical depth, and the ordinate represents the value of lithology expressed as either marl (0) or sandstone (1).

Klo–B. For input data we used resistivity logs (R_{16}, R_{64}), SP logs, corresponding data about well depth and lithology, and the hydrocarbon saturation value. Hydrocarbon saturation was manually determined from resistivity log R_{64}. The corresponding data well depth gave better results in this neural network performance than in lithology prediction, where it had little or no effect. Hydrocarbon saturation value, as well as lithology, was defined as a categorical value. Here "1" stands for positive hydrocarbon accumulation and "0" for

Table 2: Neural network parameters for hydrocarbon saturation prediction

Neural network type and properties[a]	Training error[b]	Selection error[b]
MLP 5–6–8–1	0.056897	0.091173

[a] Neural network type and properties correspond to the type of network and number of neurons per layer where the first and last numbers represent the properties of the input and output layer. Values between these represent the number and properties of the hidden layers.

[b] Error value ranges from 0 to 1, where 0 represents 100% success of prediction, i.e., no error.

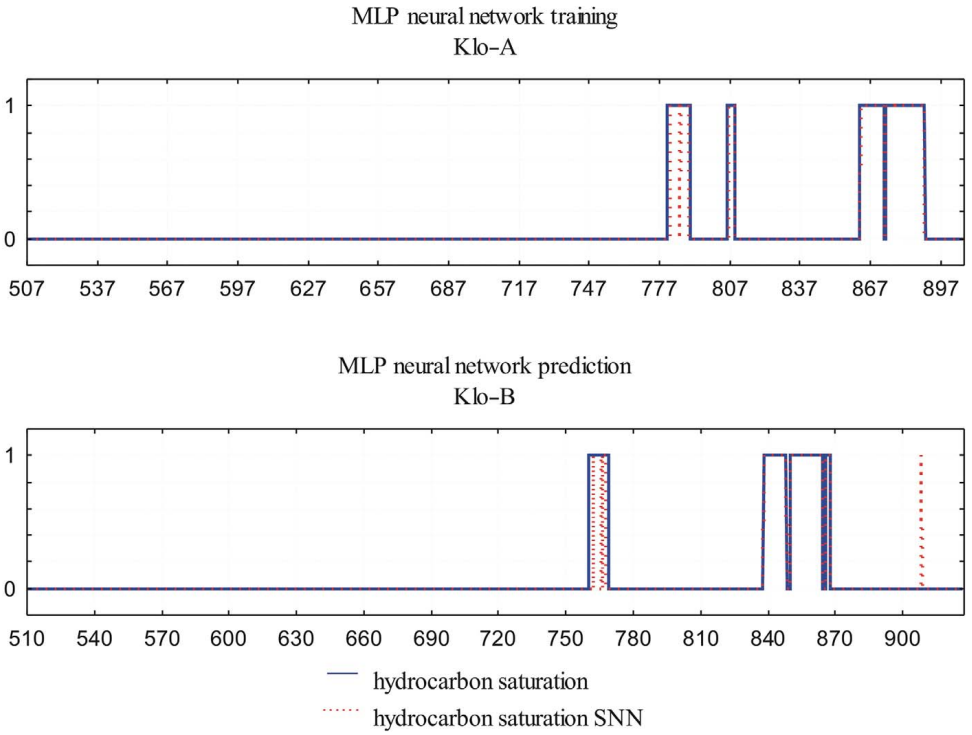

Figure 6: Comparison of MLP neural network predicted (dotted line) and manually determined data (solid line) in hydrocarbon saturation analysis in wells Klo–A and Klo–B. The diagram ordinate represents hydrocarbon saturation (0 or 1), whereas the abscissa represents the corresponding data depth.

negative. For this analysis, only the MLP neural network was used because an RBF network was characterized by a high selection and training error.

Neural network parameters are shown in Table 2. Relationships between manually determined and neural network predicted hydrocarbon saturation are shown in Fig. 6.

5. DISCUSSION

Generally, for all neural network analyses, the more input cases and more input variables used, the more successful the results and the better prediction will be.

Prediction of the lithology has proven reliable only when the extrapolation of data was within one well interval. In this analysis the most significant value was the SP log. With the input of resistivity logs alone, correspondence between true and predicted values was not satisfactory. Also, the input logs were recorded in 1956 and 1957.

A problem that appeared in this study, which prohibited cross-prediction of lithology, was different SP log values for wells Klo–A and Klo–B. The R_m for Klo–A was 72 Ωm for mud temperature of 13°C and 625 Ωm for Klo–B with a mud temperature of 1°C. This problem led to unsuccessful prediction in the shallower part of the well log data interval, where electric properties of mud on wells Klo–A and Klo–B were significantly different, and to successful prediction in the deeper part of the well log data interval (>850 m of depth), where values were similar on both corresponding well log data intervals. This problem probably occurred because of the different electrical properties of mud, influenced by different temperature values and the composition of the mud itself. In deeper segments of the well, temperature and electrical properties of mud on both wells were normalized, and therefore cross-prediction on deeper intervals was possible. One solution to this problem for shallower intervals could be introducing lithology descriptive variables that are not as dependent on mud properties as the SP log. For example, gamma ray logs as well as other well logs that define reservoir properties, such as compensated neutron and density logs, could be used to obtain better neural network performance.

6. CONCLUSIONS

In this study, we trained several artificial neural networks with the task of predicting the lithology of Upper Pannonian sediments (II sandstone series) and Lower Pontian deposits (I sandstone series), as well as hydrocarbon saturation within these beds. Sandstone facies are adequate media for statistical and neural network analysis. Our analysis of sandstone reservoirs of the Kloštar field by neural tools yielded the following results:

When determining the lithological component in wells Klo–A and Klo–B with RBF and MLP neural networks, we achieved excellent correspondence between true and predicted values.

Prediction of hydrocarbon saturation in well Klo–B with a neural network trained in well Klo–A gave excellent correspondence between true and predicted values.

Our results show the great potential of neural networks' application in petroleum geology research, where they could be used to quickly acquire results from well logs, to obtain vertical and lateral correlation of such logs, and to solve other petroleum geology problems.

REFERENCES

ANDERSON, J.A. & ROSENFLED, E. (1988): Neurocomputing**:** Foundations of Research.– MIT Press, Cambridge, 729 p.

BALIĆ, D., VELIĆ, J. & MALVIĆ, T. (2008): Selection of the most appropriate method for sandstone reservoirs in the Kloštar oil and gas field.– Geol. Croat., 61/1, 27–35.

BASSIOUNI, Z. (1994): Theory, Measurement, and Interpretation of Well Logs.– Society of Petroleum Engineers, Richardson, 372 p.

BHATT, A. (2002): Reservoir properties from-well logs using neural networks.– Unpubl. PhD Thesis, NTNU, Trondheim, 151 p.

BISHOP, C.M. (1995): Neural Networks for Pattern Recognition, 1st edition.– Oxford University Press, New York, 504 p.

FAHLMAN, S.E. (1988): Faster Learning Variations on Back-propagation: An Empirical Study.– 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.

LUČIĆ, D., SAFTIĆ, B., KRIZMANIĆ, K., PRELOGOVIĆ, E., BRITVIĆ, V., MESIĆ, I. & TADEJ, J. (2001): The Neogene evolution and hydrocarbon potential of the Pannonian Basin in Croatia.– Mar. Petrol. Geol. 18/1, 133–147.

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.

MALVIĆ, T., CVETKOVIĆ, M. & BALIĆ, D. (2008): Geomatematički rječnik [Geomathematical dictionary].– Hrvatsko geološko društvo (HGD), Zagreb, 74 p.

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, MIT Press, Cambridge, 1, 381–362.

VELIĆ, J. (1979): O razlikovanju neotektonskih, struktura u zapadnom dijelu Savske depresije [*Distinguishing the neotectonic structures in western part of Sava Depression* – in Croatian].– Geol. vjesnik, 31, 175–183.

VELIĆ, J. (1980): Geološka građa zapadnog dijela Savske depresije [*Geological settings of the western part of Sava Depression* – in Croatian].– Unpubl. PhD Thesis, Faculty of Mining, Geology and Petroleum Engineering, Zagreb, 139 p.

VELIĆ, J. (1983): Neotektonski odnosi i razvitak zapadnog dijela Savske potoline [*Neotectonic relations and development of the Western part of Sava Depression* – in Croatian].– Acta Geol. 13/2, 26–65.

VRAGOVIĆ, M. & MAJER, V. (1980): Prilozi za poznavanje metamorfnih stijena Zagrebačke gore, Moslavačke gore i Papuka (Hrvatska) [*Samples for recognition of metamorphic rocks of Zagrebačka gora, Moslavačka gora and Papuk mountain (Croatia)* – in Croatian].– Geol. vjesnik, 31, 295–308.

Manuscript received December 16, 2008
Revised manuscript accepted May 15, 2009
Available online June 19, 2009