

Neural networks in petroleum geology as interpretation tools

Tomislav Malvić

INA-Oil Industry Plc.,
Sector for Geology and Reservoir Management,
Zagreb

Josipa Velić

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

Janina Horváth

Department of Geology and Paleontology,
University of Szeged, Szeged

Marko Cvetković

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

Three examples of the use of neural networks in analyses of geologic data from hydrocarbon reservoirs are presented. All networks are trained with data originating from clastic reservoirs of Neogene age located in the Croatian part of the Pannonian Basin. Training always included similar reservoir variables, i.e. electric logs (resistivity, spontaneous potential) and lithology determined from cores or logs and described as sandstone or marl, with categorical values in intervals. Selected variables also include hydrocarbon saturation, also represented by a categorical variable, average reservoir porosity calculated from interpreted well logs, and seismic attributes. In all three neural models some of the mentioned inputs were used for analyzing data collected from three different oil fields in the Croatian part of the Pannonian Basin. It is shown that selection of geologically and physically linked variables play a key role in the process of network training, validating and processing. The aim of this study was to establish relationships between log-derived data, core data, and seismic attributes. Three case studies are described in this paper to illustrate the use of neural network prediction of sandstone-marl facies (Case Study # 1, Okoli Field), prediction of carbonate breccia porosity (Case Study # 2, Beničanci Field), and prediction of lithology and saturation (Case Study # 3, Kloštar Field). The results of these studies indicate that this method is capable of providing better understanding of some clastic Neogene reservoirs in the Croatian part of the Pannonian Basin.

Key words: neural network, log data, seismic, Croatia, Pannonian Basin

Introduction

Generally, neural networks are modern interpretation tools with several purposes. In the early days of artificial intelligence Rosenblatt – an American

Addresses: T. Malvić: Šubićeva 29, 10000 Zagreb, Croatia, e-mail: tomislav.malvic@ina.hr

J. Velić, M. Cvetković: Pierottijeva 6, 10000 Zagreb, Croatia

J. Horváth: H-6722 Szeged, Egyetem u. 2–6, Hungary

Received: September 9, 2010; accepted: December 18, 2010

psychologist – developed a machine called a perceptron, based on memorizing the pattern of the human mind (Rosenblatt 1957). In 1958 Rosenblatt proposed the perceptron as a more general computational model than the McCulloch–Pitts units. McCulloch and Pitts (1943) developed the first artificial neuron, the Threshold Logic Unit (TLU) in 1943; later it was improved by Rosenblatt. The essential innovation was the introduction of numerical weights and a special interconnection pattern. In the original Rosenblatt model the computing units are threshold elements and the connectivity is determined stochastically. Learning takes place by adapting the weights of the network with a numerical algorithm (Minsky and Papert 1969; Rojas 1996).

Supervised and unsupervised trainable networks are used in many different fields such as geology; however, the application of a supervised network is more current because the unsupervised learning network can solve specific problems like clustering and pattern recognition problems, such as identifying spatial patterns (Horváth 2010). The geology and geophysics of hydrocarbon reservoirs, which include many indirect subsurface measurements, is an example where a supervised trainable neural network is widely applied (e.g. Chen et al. 1999; Huang and Williamson 1996; Luthi and Bryant 1997; McCormack 1991; Ouenes 2000). A considerable amount of data is available from different types of resistivity and radioactive logs as well as interpreted seismic sections. Such quantities make the process of learning or training, and the validating of networks, possible. How networks are constructed is critical. The user provides some examples for the neural network to learn, and then the network is tested or validated with another data set to check the success of training. Eventually the network is ready for processing of a new dataset with a target variable of unknown value. However, any new variable not included in the training set will be misclassified or not recognized. There are several neural algorithms, each with their advantages and limitations, which can be subdivided into two main groups: the feed back and feed forward networks. Networks are called feed forward when they do not have any feedback cycle in the construction of their spanning tree; otherwise they are called feedback ones. Feed forward networks are widely applied and fully connected perceptron artificial neural networks (ANN). There are several algorithms of such networks like Probabilistic Neural Networks (PNN), Radial Basis Function Networks (RBF), and Resilient Propagation Algorithm (RPA); some of them are applied in this work.

A trained neural geologic model for the same input always give the same output, but in each iteration this input is trained with different weighting coefficients for neurons, which practically means that errors and consequently output, will be different. At the same time neural algorithms require a large amount of data, which can be available in reservoirs of deterministic or, sometimes stochastic, types. According to the classification of Jensen et al. (2000) reservoirs can be considered as:

1. Completely deterministic reservoirs where the inter-well areas are well known and described by paleoenvironments. However, such localities are rare, and mostly include mature fields with many wells. These are of interest only in the development stage or when searching for "by-passed" oil;

2. Stochastic reservoirs are those where the architecture is mostly known at the macro-scale. However, the inter-well area cannot be described for certain. In this case improved interpolation methods, like geostatistics, are very desirable.

3. Random reservoirs are very rare and are mostly not a goal in geological modeling. They are defined only by a few wells and/or gravimetric and magnetometric data. They could be characterized by Monte Carlo sampling or by analogy of similar prospects.

This means that neural algorithms can be applied to well-studied reservoirs, where many previous results of geologic modeling are available; those algorithms are applied for advanced characterization of subtle traps or hydrocarbon reserves.

The important question was whether it was possible to apply neural algorithms for a better description of hydrocarbon fields in the Pannonian Basin. The reservoir description of such fields can consist of the description of clastic systems, which include more than 90% of the hydrocarbon pools in the Croatian part of the Pannonian Basin, where data has been collected. The term "better description" was used for the application of neural networks in (a) the prediction of reservoir lithology in parts without available electric logs or cores and (b) mapping of porosity with the expectation that the obtained maps will be more accurate than maps obtained with other interpolation techniques.

There are three Croatian hydrocarbon fields where neural analyses have been applied so far, with the purpose of describing reservoir facies (Fig. 1). The oldest analysis is of the Okoli Field (Malvić 2006) where neural networking has been applied to predict the vertical position of sandstone and marl intervals in two wells in the "c2" reservoir sequence (Lower Pontian age). A back-propagation algorithm was applied for log curves from two wells. The second target was the Beničanci Field where the reservoir is of massive type, represented by dolomite and limestone breccia. The neural analysis was performed in 14 wells for interval-averaged porosity and seismic attribute data by using a classical back-propagation algorithm. The last analysis was performed in the Kloštar Field, with the goal of predicting lithology and saturation in the 1st (Lower Pontian) and 2nd (Upper Pannonian) sandstone reservoir "series". Here two supervised network algorithms – multi-layer perceptron and radial basis function – were used.

Neural networks method description

The idea of a machine that could "learn" represents the prototype of a neural network. The scheme called perceptron included connections like those in an associative memory, representing the artificial neuron model (Fig. 2) that is

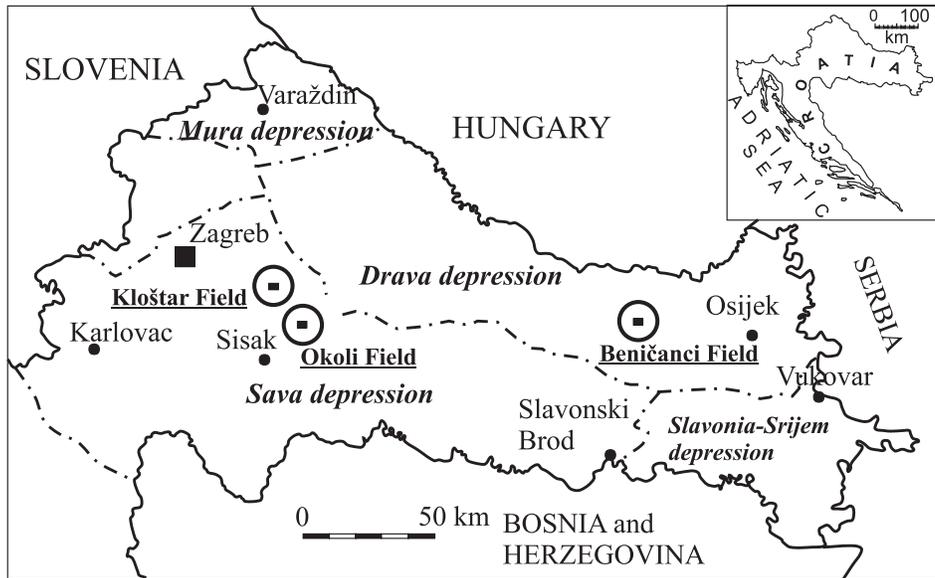


Fig. 1 Location of hydrocarbon fields where neural analyses had been applied in the Croatian part of the Pannonian Basin

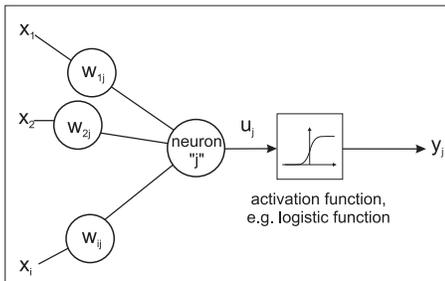


Fig. 2 The artificial neuron model

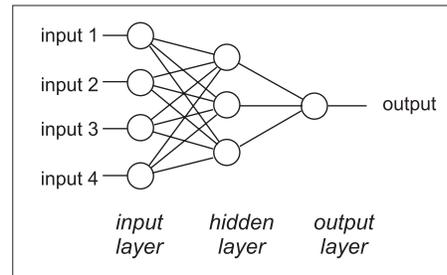


Fig. 3 Schematic organization of neural networks through layers

assembled from several inputs and a single output. Each input (x) is associated with related weight (w) added to the input value. Depending on the result (u_j), and conditions in activation function, the neuron could stay inactive or be activated (y). Neurons are organized through layers (Fig. 3). The first set of neurons makes an input layer. Such inputs are modified through hidden layers (that are not connected with information outside the network) and the result is given in the output layer. Expression (1) represents a set of calculations of the neuron, and equation (2) detects activation of the neuron.

$$u_j = \sum_{i=1}^n (x_i \times w_{ij}) \quad (1)$$

where:

- j – number of neurons
- i – number of inputs
- x_i – value of input "i"
- w_{ij} – previously determined weighting coefficient for input "i"
- u_j – value of output in neuron "j"

$$y_j = f \times (u_j + t_j) \quad (2)$$

where:

- f – activation function
- t_j – target value for neuron "j"
- y_j – layer output (or total output if it is the last layer)

Equation (1) implies previously determined weighting coefficients, value of activation hypothesis acceptance, number of layers and number of neurons in each layer. The network training (learning) means that values of the weighting coefficients and hypothesis acceptance are changed through this period. The very first networks were based only on perceptrons and the main problem was that the perceptron solves only a simple linear equation. It was a heavy restriction, which was overcome by the back error propagation paradigm (abbr. backprop). Such a network uses a large number of hidden layers and is named multiple layer perceptron (e.g. Anderson and Rosenfeld 1988).

Multi-layer networks and backpropagation learning algorithm

A multi-layer perceptron has an input layer of source nodes and an output layer of computation nodes and these two layers are connected by one or more layers of hidden neurons, which are not directly accessible. Usually the training of an MLP runs using a backpropagation algorithm that involves two training phases, the forward phase and backward phase (Werbos 1974; Rumelhart et al. 1986).

The term backpropagation algorithm means that network training includes determination of the difference between true and wanted network response using a training dataset. Calculation of differences, i.e. error, is backed in the network for optimizing training (Fig. 4). Such an error is determined for each neuron and used for adopting the existing weighting coefficient and activation value. This corrective procedure is called backpropagation. It is repeated as many times as required until the particular or total error is decreased below the limit or

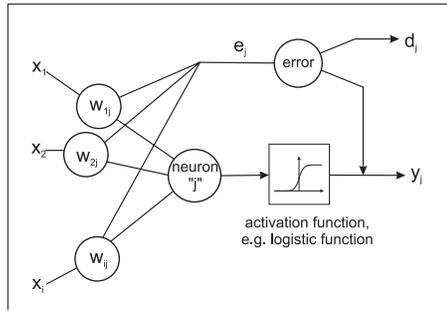


Fig. 4
Adoption of weighting coefficient and error decreasing

until maximal iteration is reached. Thereupon the trained network can be applied for processing new inputs. Backpropagation is the most popular paradigm that is applied to neural networks, where the weighting coefficients and errors are calculated

with equations (3) and (4).

$$[w_i]_{new} = [w_i]_{old} + [LR \times TF \times CT] + [MC \times previous \Delta w] \quad (3)$$

where:

- w_{new} – weighting coefficient of input (seismic attribute) in "i-th" iteration
- w_{old} – weighting coefficient of previous iteration
- Δw – difference between the two weighting coefficients (new and old)
- LR – Learning Rate; indicates in each iteration the use level of the transformation function and momentum coefficient. If LR=0 the transformation function is not used and entire network is based only on applying the Momentum Coefficient
- TR – Transfer Function; can be selected among several types. Here we used the sigmoid shape expressed as $f(x)=1/(1+e^x)$
- CT – Correction Term; value depends on differences between true (measured) and trained (by network) value
- MC – Momentum Coefficient; defines how large the influence of the result of previous iteration is in instantaneous calculation.

$$e_j = y_j \times (1 - y_j) \times (d_j - y_j) \quad (4)$$

where:

- y_j – true output
- d_j – desired response
- e_j – error rate

This algorithm is used in multi-layer networks, but often could be characterized by long-lasting training, which can be a limitation in a large dataset, such as in geophysics (McCormack 1991). This type of reduced learning rate resulted from the gradient descent method used in the backprop algorithm, which only allows sending information to previous layers, i.e. back.

The backpropagation learning algorithm is simple to implement and computationally efficient, in that its complexity is linear in the synaptic weights of the network. However, a major limitation of the algorithm is that it does not

always converge and can be excruciatingly slow, particularly when we have to deal with a difficult learning task that requires the use of a large network (Sandberg et al. 2001).

There are many additional algorithms that can be used in neural networks and which abolish the above-mentioned convergence problem, but most of them are considered as an extension of backpropagation algorithms or another gradient-based algorithm like Resilient Propagation Algorithm (RPA). They can be applied as primary (solely) or secondary learning algorithms. These are 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). The greatest advantage of these algorithms over backpropagation is that they are significantly faster, but not always better in terms of errors.

Resilient Propagation Algorithm

Today, there are several other algorithms that can increase training speed. One of them (Riedmiller and Braun 1993) is the Resilient Propagation Algorithm (RProp). The main difference with the standard backprop algorithm is expressed through using only partial derivations in the process of weighting coefficient adjustment. RProp uses so-called training through epochs, where weighting coefficients are adjusted after all the patterns or connections for input data are sets. RProp is 4–5 times faster than the standard backprop algorithm.

The speed of training can be easily explained by variables in equation (3) called momentum coefficient, learning rate and correction term. Look at an imaginary set of eight 1D porosity values 7.2, 7.0, 6.3, 5.7, 6.2, 6.5, 5.5, 5.2% that tend to a minimum at the end, also including one local minimum 5.7% in 4th place. The network will recognize these local minima if the momentum coefficient is set at an extremely sensitive level for the detection of local minima. But the general trend is recognized using the learning rate parameter. The correction term represents the differences between true and modeled values, which are calculated for each hidden layer, when the network tries to decrease these differences through the next iteration.

Radial Basic Function (RBF) networks

Another popular layered feed forward network is the radial basic function (RBF) network which has important universal approximation properties (Park and Sandberg 1993). RBF networks differ from multi-layer perceptrons in some fundamental respects (Sandberg et al. 2001). Hereinafter RBF networks are described by some differences:

The RBF networks, as opposed to the backpropagation model, have only one single hidden layer with a non-linear RBF activation function, and the output layer is always linear; the structure of the RBF model is shown in Figure 5.

RBF networks are local approximations, whereas MLP networks are global approximations.

The activation function of MLP computes the inner product between the input signal vector and the pertinent synaptic weight vector. On the other hand, in the RBF networks the activation function computes the Euclidean distance between the input signal vector and parameter vector of the network.

Radial basis functions are typically used to build up function approximations of the form (5).

$$y = \sum_i w_i g_i(x) = \sum_i w_i g(\|x - c_i\|) = w^T u \quad (5)$$

where:

- x_i – value of input "i"
- u_j – value of output in neuron "j"
- w_i – previously determined weighting coefficient for input "i"
- c_i – center vector of neuron "i"
- $g(\cdot)$ – radial basis function
- $\| \cdot \|$ – usually the Euclidean distance, but it can be replaced with the Mehalanobis distance.

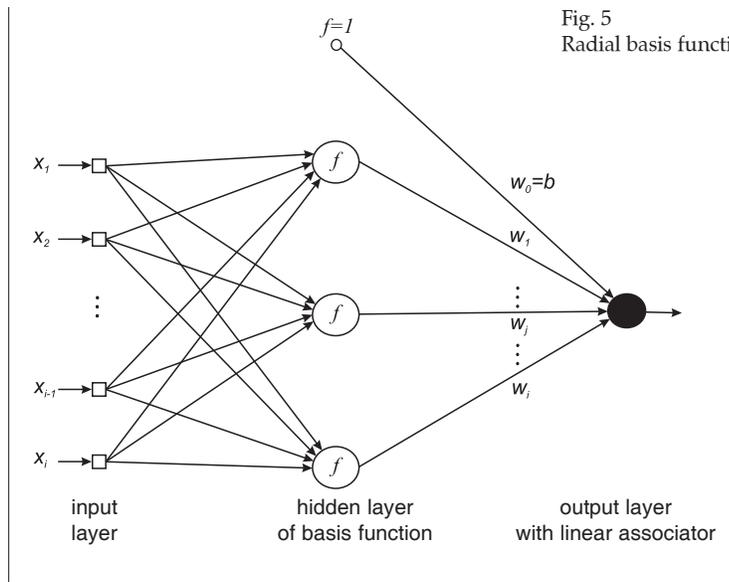


Fig. 5
Radial basis function network

The RBF network is frequently applied in solving classification problems rather 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 (e.g. Bhatt 2002; Malvić 2006; Cvetković et al. 2009), well-log correlation (Luthi and Bryant 1997), and mapping porosity supported by seismic attributes (e.g. Malvić and Prskalo 2007), etc.

Neural algorithm in reservoir characterization examples from Croatia

Three neural networks applied to analyzing of log and seismic data of three different oil and gas fields are presented here. Supervised neural algorithms had been applied for data analysis in all three fields.

Case Study # 1: Okoli Field

The first analyzed field is the Okoli Field, located in the northwestern part of the Sava Depression, which encompasses several sandstone series, each of which is divided into reservoir levels. The field structure is brachianticline with NW-SE strike, and faulted in its deeper parts. The reservoirs are of Upper Pontian ("b" and "c" series) and Pliocene ages ("p", "A", "B" and "a" series). Sandstone porosity varies between 14 and 37%, permeability 1.3 and $24.8 \times 10^{-3} \mu\text{m}^2$ and reservoir thickness ranges from 1 to 20 meters.

Particular intervals of the c2 reservoir are outlined. The entire "c" series was deposited over the major part of the field, and particular reservoirs are concordant (continuous), characterized by locally correlative marlstone levels at the top and bottom of sandstone ones. The following log curves were obtained from the B-1 well: GR (gamma ray), R16" and R64" (resistivity logs). In the B-2 well the selected log curves were: GR (gamma ray), PORE (effective porosity), PORT (total/true porosity), PORW (porosity in rock 100% saturated with water), R16" (resistivity), SANDSTONE (sand portion curve) and SHALE (marl portion curve). The newly calculated curve, in both cases, was called RESERVOIR. This curve represents an artificial "categorized" variable, defined with numerical values 0 and 1. Value 0 shows marlstone, and value 1 sandstone.

A backpropagation network was applied to the data from the Okoli Field. The network is described, in applied software, as a Completely Connected Perceptron type. In this type of network, training was divided into one cluster defined by 10 different sub-networks (called experts) with their own initial conditions. The values of 10 experts were selected here. The first expert is initiated by the random seed value of 0.01, and the next one with the previous value increased by 1. Each subsequent sub-network was built with an increasing number of hidden layers. Training options define the initial learning rule and learning rate. Here the local (minima) adaptive rule called Improved Resilient Propagation Plus (iRProp+) was selected. The initial learning rate was set at 0.01, later self-adjusted during training. Network training was completed on the basis of two criteria. The first is the error domain; when the domain value decreases below the limit, training is stopped in terms of what is used in this analysis. The second criterion is the number of maximum epochs in training; when this number is reached, training is completed, although the error is not below the limit. The results obtained by the cVision program (Fruhworth 2005) are represented by several tables and graphs (Malvić 2006), as in Figs 6 and 7. There are three main results expressed through so-called "machines". The Face machine describes the best network

obtained through the entire training, i.e. it is the measure of the network's successful work. The Best machine encompasses data of momentarily the best iteration, and the Trip machine includes a series of the most successful iterations obtained in the entire analysis.

In the B-1 well the learning process (L) is performed with 153 data with the value of 0, and 142 data with the value of 1. Validation (V) is carried out by using 48 data with the value of 0 and 50 data with the value of 1. In the B-2 well the learning process (L) is performed with 225 data of the value of 0, and 215 data with a value of 1. Validation (V) used 71 data of the value of 0 and 75 data of the value of 1.

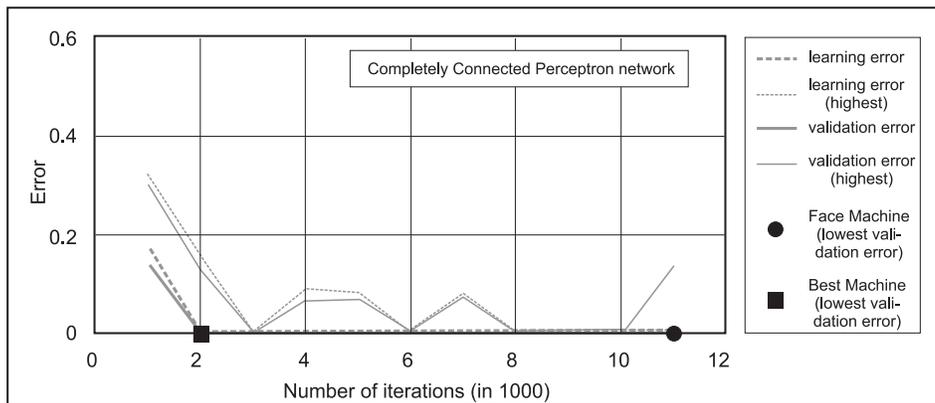


Fig. 6 Relations of errors in periods of training divided in learning (L) and validation (V) and position of Face and Best configurations (the symbols F, B in legend) for the B-1 well (from: Malvić 2006)

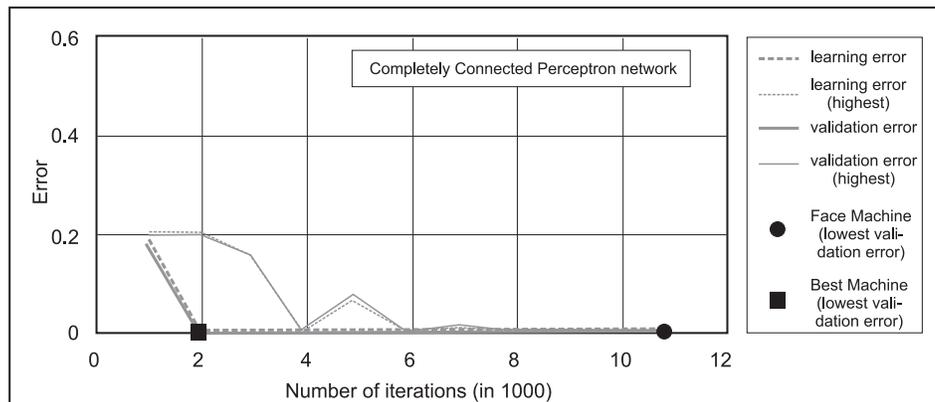


Fig. 7 Relations of errors in periods of training divided in learning (L) and validation (V) and position of Face and Best configurations (the symbols F, B in legend) for the B-2 well (from: Malvić 2006)

In all 393 data for learning and validation were used in the B-1 well, and 586 in the B-2 well.

The analytical target was the prediction of the true position of the values of 0 and 1 (i.e. sandstone and marlstone sequences). The success in prediction, in the B-1 well, of intervals that were not interpreted lithologically from e-log, but instead using neural networking, is visually presented in Table 1.

Table 1
 Prediction of lithology by trained neural network and comparison of existing results (masked for network). Theoretical probability of successful prediction was 88.1% and real probability was 52.5% (100% in sandstones, 6.66% in marlstones)

True facies	Predicted facies	True facies	Predicted facies	True facies	Predicted facies
Marlstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Marlstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Marlstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Marlstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Marlstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Marlstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Marlstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Marlstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Marlstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Marlstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Sandstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Sandstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Sandstone	Sandstone	Sandstone	Sandstone	Marlstone	Marlstone
Sandstone	Sandstone	Sandstone	Sandstone	Marlstone	Marlstone
Sandstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Sandstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Sandstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Sandstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Sandstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Sandstone	Sandstone	Sandstone	Sandstone	Marlstone	Sandstone
Sandstone	Sandstone	Marlstone	Sandstone		

Numerically a high correlation was obtained between predicted and true position of sandstone lithology in the reservoir; on the other hand, marlstone positions mostly do not correspond. Moreover, the total probability for facies prediction is high. In the B-1 well (3 log curves) it is $\geq 78.3\%$ and in the B-2 well (7 log curves) at least $\geq 82.1\%$. The Face machine is calculated relatively in the early period of network training (in the B-1 well this in the 2186th iteration, in the B-2 well in the 7626th iteration, of a total of 30,000).

Case Study # 2: Beničanci Field

The Beničanci Field is located in the eastern part of the Drava Depression. It is brachianticline with a W–E strike, connected to a "buried hill" above Paleozoic paleorelief. The reservoir is of massive type, lithologically represented by dolomitic and limestone breccia of Badenian age. The average porosity is 7.33% and initial water saturation 28.13%. The reservoir rocks are the result of Middle Miocene breccia depositional model that resulted from alluvial fan as well as tectonic activity. It is still one of the five main hydrocarbon reservoirs in Croatia, represented by carbonate breccia of Badenian age. The seismic amplitude, phase and frequencies are interpreted from the 3D seismic cube, and interpolated (from a total of 16,384 values) at 14 well locations with the newest reservoir mean porosity calculations. Those four variables at 14 locations represented the dataset used for the backpropagation network training with 10,000 iterations, searching for the correlation between attribute(s) and porosities. Such an obtained network, with the best weighting coefficient, had been used for porosity mapping in the southern part of the field (Fig. 8).

In this case the neural network uses a logistic activation function for the backpropagation algorithm. Furthermore, the network included two groups of parameters that dictated training, the network's low and the high sensitivity parameters. The low-sensitivity parameters were (a) the number of hidden

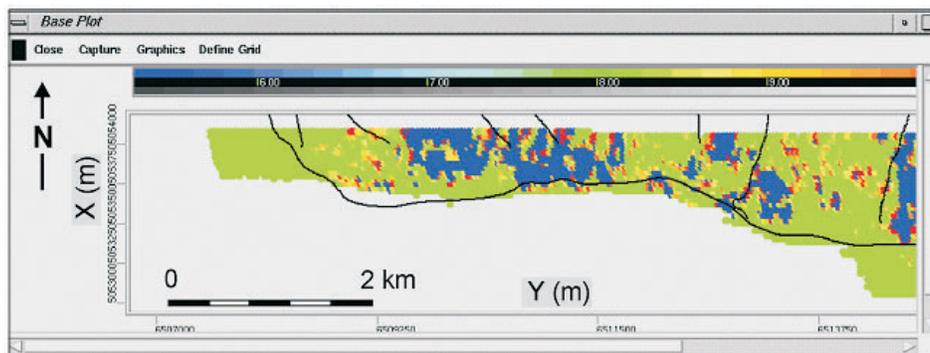


Fig. 8
The neural network porosity map of the Beničanci Field - color scale 5–10% (from: Malvić and Prskalo 2007)

layers, (b) the learning rate, (c) the momentum coefficient and (d) the number of iterations. The number of hidden layers varied between 5 and 25, but improved correlation (seismic-log porosity) only for 0.001–0.01 and significantly reduced the training. Values of learning rate and momentum coefficient are kept on defaults (0.9 and 0.6). Eventually, the number of iterations was 10,000, because experiment up to 30,000 did not improve results. The high-sensitivity parameters were (e) the number of seismic attributes, (f) the activation function and (g) the convergence criteria. The most important parameter was the number of included attributes, and the best training was obtained with all three attributes. Numerically the determination coefficient between amplitude + frequency + phase and porosity was $R^2=0.987$, with convergence ($\Sigma\epsilon^2$) of 0.329. The activation function was of the log-sigmoid shape. The convergence criteria ($\Sigma\epsilon^2$) had the role of stopping training if the value was the lower than the defined limit (which was 1). Also, the lower convergence indicated higher correlation.

Porosity could be predicted with any number of attributes, but the highest correlation was reached using all three attributes together. Estimated porosity ranged between 5.27–11.06%, but the estimation by the neural network had a tendency to group majority of cell data in a narrow range.

Case Study # 3: Kloštar Field

The Kloštar Field is situated in the northwestern part of the Sava Depression. It is structurally located in the Križ structure of Dinaric strike (NW–SE), along Mt. Moslavačka gora. The complete stratigraphic sequence includes the units of Paleozoic, Middle Miocene, Upper Miocene, Pliocene and Quaternary sediments. The main hydrocarbon reservoirs are found in Upper Pannonian ("II sandstone series") and Lower Pontian ("I sandstone series") sandstone, with calcareous marl as isolators. The field structure was formed between main strike slip faults in extensional phase from Badenian to Lower Pontian. Later, it was inverted to a faulted anticline, uplifted from the Upper Pontian onward, when the main faults changed displacement character.

Basic well log data were used from two wells (Klo-A and Klo-B) for the backpropagation neural analysis (Cvetković et al. 2009). Electric logs included resistivity (R16 and R64) and spontaneous potential (SP) with resolution of 10 measurements. Neural networks had been used for prediction of lithology and hydrocarbon saturation in sandstone reservoirs of Lower Pontian (1st sandstone series) and Upper Pannonian (2nd sandstone series) age.

Neural analyses were performed using StatSoft STATISTICA 7.1. Lithology prediction was performed with data collected by manual determination (Cvetković et al. 2009) of lithological components selected in layers described as marl (0) and sandstone (1) from well logs (Bassiouni 1994). The neural network was trained on data from the 1st sandstone series, and prediction made on the intervals in the 2nd sandstone series, and vice versa. Thus it was possible to easily

recognize the success of the neural network as well as training and prediction error. Also, the training set was divided into true training (learning) and validation datasets. Stopping criteria in network training were the desired number of iterations, or amount of error. Moreover, for both cases and wells two types of networks were used: a multi-layer perceptron (abbr. MLP) network (type of backprop) and a radial basis function type analysis (abbr. RBF). Unfortunately the results of the neural analyses could not be spatially correlated, due to different relative shifts of SP curves in the selected wells. A comparison between lithology determined classically from logs (blue line) and using neural network (red dotted line) is given in Fig. 9.

Hydrocarbon saturation prediction uses cross-prediction, i.e. the neural network is trained on the Klo-A well, and the prediction was performed on the Klo-B well. Hydrocarbon saturation was manually determined from resistivity log R64, and defined as a categorical value (1 for hydrocarbon accumulation and 0 for none). Only the MLP neural network was used because an RBF network was characterized by a high selection and training error (Cvetković et al. 2009); the results are shown in Fig. 10.

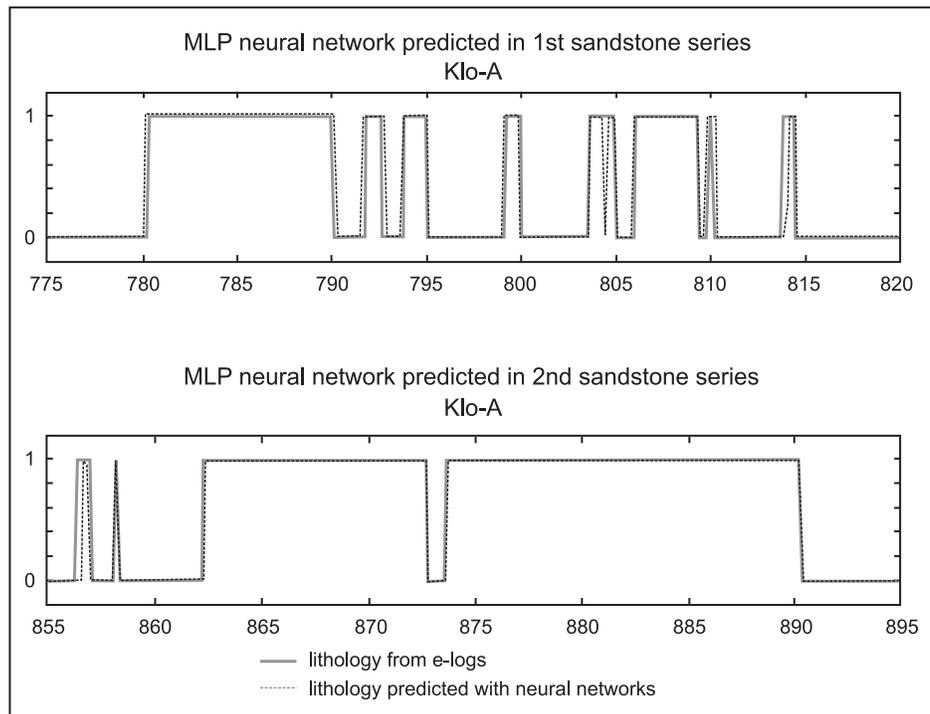


Fig. 9
Comparison of lithology predicted manually and by MLP network. Abscissa represents vertical depth and ordinate lithology (marl=0, sandstone=1; from: Cvetković et al. 2009)

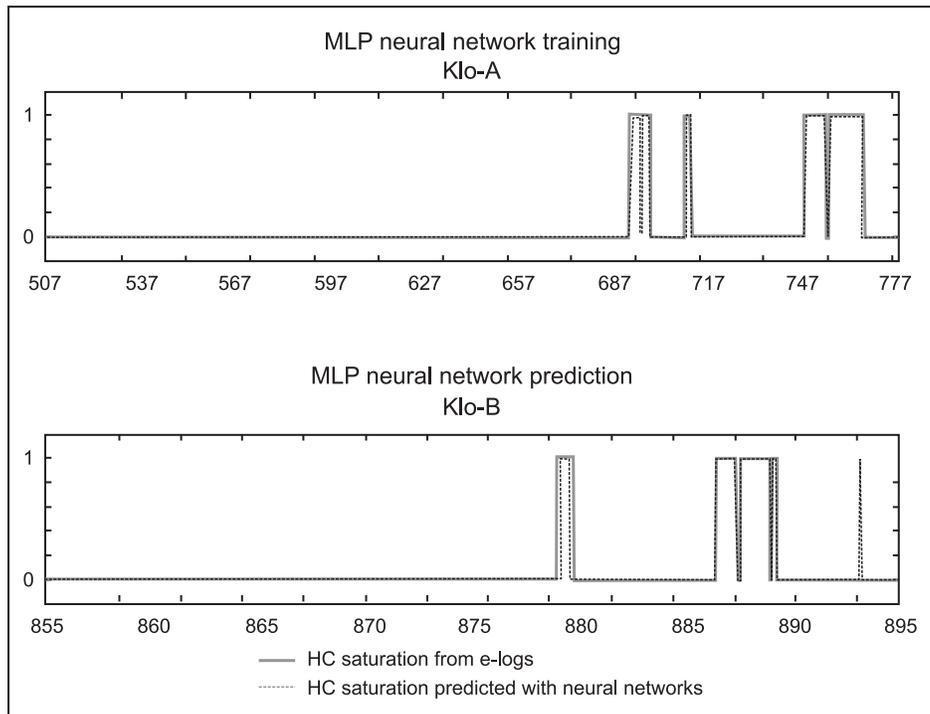


Fig. 10 Comparison of saturation predicted manually and by MLP network. Abscissa represents vertical depth and ordinate saturation (hydrocarbons=1, water=0; from: Cvetković et al. 2009)

Discussion

Three examples of using of neural networks in analyses of geologic data from hydrocarbon reservoir have been presented. All networks were trained with data originating from clastic reservoirs of Neogene age located in the Croatian part of the Pannonian Basin. Training always included similar reservoir variables, i.e. electric logs (resistivity, spontaneous potential), lithology determined from cores or logs and described as sandstone or marl with categorical values in intervals. Other variables were hydrocarbon saturation, also represented by a categorical variable, and average reservoir porosity calculated from interpreted well logs and seismic attributes. Some of mentioned inputs were used in all three neural models.

In this section the results are discussed in light of what is already known about the subject of the investigation, and explaining our new understanding of the problem after taking results into consideration.

The results obtained in the facies prediction in the Okoli Field for a monotonous alteration of sandstone and marl can be extended to the entire depression, because the analyzed lithology and relationships are very characteristic for the entire area. It is clear that such a facies analysis does not need such a large iteration set. A larger positive impact could be obtained by increasing the number of relevant input log curves. The present network is over-trained and does not contain enough criteria for recognizing marlstone sequences. In general, however, a well-trained backpropagation network was able to be used for lithology prediction in the well intervals where e-logs are of poor quality or partially absent. Such intervals need to be selected in the same reservoir where the network has successfully undergone learning and validation.

The Beničanci Field porosity map showed a tendency to group neural results in a narrow range. This is also often characteristic in geostatistics and regression. Not one of the selected seismic attributes shows high linear relationship with porosity (low correlation), and only one of their syntheses summarized in one particular value was able to show correlation with porosity. The neural porosity map was compared to previously interpolated Collocated Co-kriging porosity maps (Malvić and Đureković 2003), the best porosity estimation prior to using the neural approach, obtained with reflection strength as the secondary seismic source of information. A relatively smooth map, and one rarely reaching measured porosity minimum and maximum, strongly points to the conclusion that neural estimation is more precise than previous interpolations.

It turned out that for the Kloštar Field the neural results were the most successful, with very low network error. As Cvetković et al. (2009) concluded: "...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 errors." Prediction of lithology and hydrocarbon saturation in sandstone reservoirs was excellent and showed that the software used was the best one for analysis of geologic data collected for this analysis. Also, neural results confirmed the multi-layer perceptron algorithm as the one that can be applied most often as the appropriate one for modeling and prediction of hydrocarbon reservoir properties.

Conclusions

It is well known that neural networks favor numerous inputs, because in this way they can be trained with lower prediction error. However, this can result in a counter-effect because such networks also can be over-trained and provide prediction for almost every input item, which is useless. Such behavior can be compared to polynomial regression of a higher degree (polynomial curve of 4th, 5th, 6th, etc. degree) that can perfectly fit almost any geologic variable in a hydrocarbon reservoir, but cannot describe any geologic process, due to (a) our limited knowledge and (b) the high stochastic component of geologic processes.

This problem of an over-trained network and the selection of many inputs arose in the Benicanci Field network. The best training and prediction had been achieved using all three seismic attributes simultaneously in the prediction of lateral reservoir porosity. However, this solution was accepted based on geologic knowledge that connection between these three attributes could be described as "geophysically possible and even meaningful".

Using classical electric logs in the prediction of lithology, and even of saturation, indicated the importance of applied neural programs for analysis. For now, the famous Statistica package (StatSoft) showed very good performance for well log analyses, and not only with a neural algorithm. However, based on selected well logs and interpretation intervals inside the reservoir, training can also reach the critical point of an over-trained network, especially when (a) resistivity and especially spontaneous potential curves are not characterized with significant shifts and (b) there are too many lithological intervals selected for prediction inside the reservoir.

In the end all presented networks were of backpropagation type. Most of them have a multi-layer perceptron architecture and apply the sigmoid or log-sigmoid activation function. One network, however, used only three layers and a radial basis function as an activation function, emphasizing that different types of backpropagation architecture and activation can be used. Without regard to over-training in the case of numerous inputs, additional well logs that describe reservoir lithology or saturation, such as compensated neutron or density, would enhance prediction of every neural network applied to petroleum geology.

Acknowledgements

This work is part of a multidisciplinary geostatistical research study that was performed until 2010 within the project entitled "Stratigraphical and geomathematical research of petroleum geologic systems in Croatia" (project no. 195-1951293-0237), financed by the Ministry of Science, Education and Sports of the Republic of Croatia. The first results of the resulting analysis were presented at the 12th Congress of Hungarian Geomathematics and the 1st Congress of Croatian and Hungarian Geomathematics in Mórahalom, Hungary (29–31 May 2008).

References

- Anderson, J.A., E. Rosenfeld 1988: Neurocomputing: Foundations of Research. – MIT Press, Cambridge, 729 p.
- 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.

- Chen, M., M. Bai, J.C. Roegiers 1999: Permeability tensors of anisotropic fracture networks. – *Mathematical Geology*, 31/4, pp. 355–373.
- Cvetković, M., J. Velić, T. Malvić 2009: Application of neural networks in petroleum reservoir lithology and saturation prediction. – *Geologia Croatica*, 62/2, pp. 115–121, Zagreb.
- Fahlman, S.E. 1988: Faster-Learning Variations on Back-Propagation: An Empirical Study. – In *Proceedings, 1988 Connectionist Models Summer School*, Morgan-Kaufmann, Los Altos CA.
- Fruhworth, R.K. 2005: *cVision Users Guide Version-2*. Neuro Genetic Solutions GmbH. – Centre for Applied Technology, 31 p, Leoben.
- Gorse, D., A. Shepherd, J.G. Taylor 1997: The new ERA in supervised learning. – *Neural Networks*, 10/2, pp. 343–352.
- Horváth, J. 2010: Potential application of neural networks in identification of depositional environments. – 14th Annual Conference of the International Association for Mathematical Geosciences-IAMG 2010, Budapest, Earth and Environmental Science, Abstract book
- Huang, Z., M.A. Williamson 1996: Artificial neural network modeling as an aid to source rock characterization. – *Marine and Petroleum Geology*, 13/2, pp. 277–290.
- Jacobs, R.A. 1988: Increased rates of convergence through learning rate adaptation. – *Neural Networks*, 1, pp. 295–307.
- Jensen, J.L., L.W. Lake, P.W.M., Corbett, D.J. Goggin 2000: *Statistics for Petroleum Engineers and Geoscientists*. – Prentice Hall PTR, New Jersey, 390 p.
- Levenberg, K. 1944: A method for the solution of certain problems in least squares. – *Quart. Appl. Math.*, 2, pp. 164–168.
- Luthi, S.M., I.D. Bryant 1997: Well-log correlation using a back-propagation neural network. – *Math. Geol.*, 29/3, pp. 413–425.
- Malvić, T. 2006: Clastic facies prediction using neural network (case study from Okoli field) [Predviđanje klastičnih facijesa neuronskom mrežom (primjer polja Okoli)]. – *Nafta*, 57/10, pp. 415–424, Zagreb.
- Malvić, T., S. Prskalo 2007: Some benefits of the neural approach in porosity prediction (Case study from Beničanci field) [Koristi upotrebe neuronske mreže u procjeni poroznosti (na primjeru polja Beničanci)]. – *Nafta*, 58/9, pp. 455–467, Zagreb.
- Malvić, T., M. Đureković 2003: Application of methods: Inverse distance weighting, ordinary kriging and collocated cokriging in porosity evaluation, and comparison of results on the Beničanci and Stari Gradac fields in Croatia. – *Nafta*, 54/9, pp. 331–340, Zagreb.
- Marquardt, D.W. 1963: An algorithm for least-squares estimation of nonlinear parameters. – *J. Soc. Indust. Appl. Math.*, 11/2, pp. 431–441.
- McCormack, M.D. 1991: Neural Computing in Geophysics. – *The Leading Edge*, 10/1, Society of Exploration Geophysicists. pp. 11–15.
- McCulloch, W., W. Pitts 1943: A logical calculus of the ideas immanent in nervous activity. – *Bulletin of Mathematical Biophysics*, 5, pp. 115–133.
- Minsky, M., S. Papert 1969: *Perceptrons: An introduction to computational geometry*. – MIT Press, Cambridge MA. 308 p.
- Ouenes, A. 2000: Practical application of fuzzy logic and neural networks to fractured reservoir characterization. – *Computers and Geosciences*, 26, pp. 953–962.
- Park, J., I.W. Sandberg 1993: Approximation and radial-basis function networks. – *Neural computation*, 5, pp. 305–316.
- Riedmiller, M., H. Braun 1993: A direct adaptive method for faster backpropagation learning: The RProp algorithm. *Proc. of the IEEE Intl. Conf. – On Neural Networks*, San Francisco, pp. 586–591.
- Rojas, R. 1996: *Neural Networks: A systematic introduction*. – Springer, Berlin, 502 p.
- Rosenblatt, F. 1957: The perceptron: A perceiving and recognizing automaton. – Technical report 85 - 460-1, Project PARA, Cornell Aeronautical Lab.
- Rosenblatt, F. 1958: The perceptron: A probabilistic model for information storage and organization in the brain. – *Psychological Review*, 65, pp. 386–408.

- Rumelhart, D.E., G.E. Hinton, R.J. Williams 1986: Learning internal representations by error propagation. – In: Rumelhart, D.E., J.L. McClelland (Eds): *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press, vol. 1, Chapter 8.
- Sandberg, I.W., J.T. Lo, C.L. Fancourt, J.C. Principe, S. Katagiri, S. Haykin 2001: *Nonlinear Dynamical Systems: Feedforward Neural Network Perspectives*. – Wiley, New York, 256 p.
- Werbos, P.J. 1974: *Beyond regression: New tools for prediction and analysis in the behavioral sciences*. – Ph.D. thesis, Harvard University.