first break volume 22, March 2004

Soft computing for qualitative and quantitative seismic object and reservoir property prediction

Part 1: Neural network applications

Fred Aminzadeh[1] and Paul de Groot[2] of dGB Earth Sciences begin a major series of three articles on the increasing use of soft computing techniques for E&P geoscience applications, focusing first on how neural networks can enhance seismic object detection

Soft computing has been used in many areas of petroleum exploration and development. With the recent publication of three books on the subject, it appears that soft computing is gaining popularity among geoscientists. In this paper we focus on one aspect of soft computing: neural networks, in qualitative and quantitative seismic object detection. In subsequent papers we will review other aspects of soft computing in exploration. Highlighted here will be the role neural networks play in combining different seismic attributes and effectively bringing together data with the interpreter's knowledge to decrease exploration risk in four categories (geometry, reservoir, charge and seal).

Introduction

Three new books in the general area of soft computing applications in exploration and development, Wong *et al* (2002), Nikravesh *et al* (2003) and Sandham *et al* (2003) represent a comprehensive body of literature on recent applications of soft computing in exploration. Soft computing is comprised of neural networks, fuzzy logic, genetic computing, perception-based logic and recognition technology. Soft computing offers an excellent opportunity to address the following issues:

- Integrating information from various sources with varying degrees of uncertainty
- Establishing relationships between measurements and reservoir properties
- Assigning risk factors or error bars to predictions.

Deterministic model building and interpretation are increasingly replaced by stochastic and soft computing-based methods. The diversity of soft computing applications in oil field problems and the prevalence of their acceptance can be judged by the increasing interest among earth scientists and engineers. Given the broad scope of the topic, we will limit the discussion in this paper to neural network applications. In subsequent papers we will review other aspects of soft computing, such as fuzzy logic in exploration.

Neural networks have been used extensively in the oil industry. Approximately 10 years after McCormack's review (1991) of neural network applications in geophysics, much work has been done to bring such applications to the main stream of geophysical interpretation. Some of these efforts are documented in Wong *et al* (2002), Nikravesh *et al* (2003) and Sandham *et al* (2003) which include many papers and extensive references on neural network applications. Most of these applications have been in reservoir characterization, seismic object detection, creating pseudo logs, and log editing. In the next section, we will focus on two general areas of applications of neural networks. This will include qualitative methods with the main aim of examining seismic attributes to highlight certain seismic anomalies without having access to very much well information. In this case neural networks are primarily used for classification purposes. The second category involves quantitative methods where specific reservoir properties are quantified using both seismic data and well data, and neural networks serve as an integrator of the information.

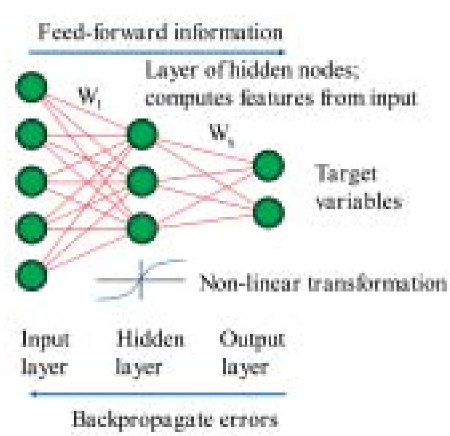

Figure 1 *Fully-connected Multi-Layer-Perceptron (MLP) type of Artificial Neural Network (ANN).*

[1]*dGB-USA, 1 Sugar Creek Center Blvd., Suite 935, Sugar Land TX 77478 , USA*
[2] *dGB Earth Sciences, Boulevard 1945 nr. 24, 7511 AE Enschede, The Netherland*

The most widely-used ANN is the Multi-Layer-Perceptron (MLP). This network is built up of many small processing units that emulate the working of biological neurons. Each neuron has one or more inputs, x_1, …, x_n, with associated weights, w_1, ..., x_n, and one output (called activation level), y. The activation level is computed from the node's inputs in two simple steps:

1. The node's net input x_s is computed as the sum of the inputs multiplied by the corresponding weights:
   $x_s = w_1{}^*x_1{+}w_2{}^*x_2{+}...{+}w_n{}^*x_n$
2. The activation level is computed by passing the net input through an activation function. Most often this is a non-linear function to ensure that the ANN can capture non-linear relationships. The most common function is the sigmoid function, given by:
   $y{=}1/(1{+}exp({-}x_s)$

The nodes in an MLP are organized in three layers: input, hidden and output (Figure 1). The information in Figure 1 flows from left to right. Nodes in the input layer simply pass the information on to the hidden layer, where the net input is passed through the activation function. The final output is computed by summation over the outputs of the nodes in the hidden layer multiplied by the corresponding weights. If all nodes are connected via their weights to all nodes in the next layer, the network is called a fully-connected MLP.

The functionality of an ANN is determined by its design and is controlled by the user, who defines input, output, the number of nodes in the hidden layer and optionally, which activation function to use. The knowledge of an ANN is incorporated in the weights. One of the main characteristics of MLP networks is that they can learn from examples. This is done in a training phase where the network is given a training set with known combinations of inputs and outputs. In a supervised learning process the network uses the provided information to adjust the connection weights between the nodes. There are different ways to adjust the network's weights, the most popular method being the back-propagation algorithm (Rumelhart et.al., 1986). It is good practice to test the performance of the network during the training phase by employing a second, independent test set. Examples from the test set are passed through the network to monitor the error but the error is not used to update the weights, as is done with examples from the training set. The test set helps to determine the optimal point for stopping the training phase, which is the point where the error on the test set is minimal. With prolonged training the error on the training set may continue to decrease while the error on the test set increases. The network loses its generalization capabilities and starts to recognize individual examples from the training set, a process called overfitting.

Let us discuss in a bit more detail what is happening inside an MLP network to demystify the common misconception of ANN being considered a 'black box' approach. Figure 2 shows an MLP network with two inputs, three nodes in the hidden layer and one output. Let us assume this network is designed to predict the average reservoir porosity (y) from two seismic attributes (x1 and x2). The network is trained on a data set with known examples. These are the porosities taken from real wells and the attributes extracted from seismic data at the well locations. The aim is now to find the optimal (possibly non-linear) relationship between seismic attributes and porosity. This corresponds to finding the surface that fits the data points best. In the training phase the network starts with a random set of weights. An example from the training set is passed through the network and the error is used to update the weights. This corresponds to stretching and squeezing the sigmoid shape surfaces. The summation of the three sigmoid surfaces yields the desired output surface. The process of passing examples through the network and adjusting the weights is repeated until the error is minimal, hence the best fitting surface is found. We now have a trained network. At every trace location we extract the two input attributes and pass these through the network to obtain a predicted porosity value. Thus the final product in this example is a porosity map.

Another way of looking at the internal functioning of an MLP is by comparing it with conventional statistical methods. If the sigmoid shape function is replaced by a linear function the network equals multi-dimensional linear regression. With the sigmoid shape activation function the network performs a multi-dimensional non-linear regression. It is common to design networks with a decreasing number of nodes from input to hidden layer and from hidden to output layer. This implies that the network tries to progressively reduce the dimensionality of the problem. From the input layer to the hidden layer the input attributes are replaced by a new set of attributes that are linear combinations of all input attributes multiplied by their corresponding weights (step 1, described above). This dimensionality reduction is comparable to Principal Component Analysis in conventional statistics, where we also aim to replace the input attributes by a new, smaller set of attributes.

Another popular supervised neural network is Radial Basis Functions (RBF). In this network the sigmoid shape activation function is replaced by a Gaussian function or other type of RBF. The training algorithm locates the functions at the positions where the data are located in multi-dimensional input space. This implies that the data cannot be uniformly distributed for RBF networks to perform well. RBF networks may perform better than MLP networks in low-dimensional problems (say less than 20 inputs) and where a high correlation exists between the input variables. The latter is often the case with seismic attributes.

Instead of supervised neural networks we can also employ unsupervised networks. The main difference between supervised and unsupervised approaches lies in the amount of a priori information that is supplied. In unsupervised (or competitive learning) approaches, the aim is to find structure within the data. This is a form of clustering, or segmentation. The data is clustered, for example, in groups of similar waveforms along a mapped horizon. The result in this case is a seismic pattern map revealing areas in the seismic data with similar seismic response. What these patterns mean in terms of geological or a petrophysical variation is subject to further study and interpretation. Popular unsupervised neural networks are Self Organizing Maps (e.g. Kohonen) and Unsupervised Vector Quantizers.

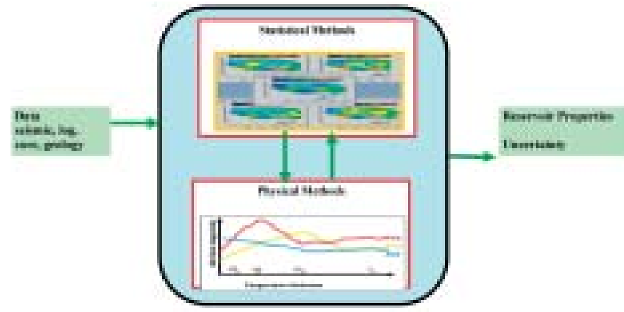

Figure 3 *Hybrid reservoir property prediction*

Physics-based vs. statistics-based methods

Physics or model-based methods attempt to exploit the changes in seismic character or seismic attribute to a given reservoir property based on physical phenomena. Many geophysical analysis methods and consequently seismic attributes are based on physical phenomena. In other words, based on certain theoretical physics (wave propagation, Biot-Gassman Equation, Zoeppritz Equation, tuning thickness, shear wave splitting, etc.), certain attributes may be more sensitive to changes in certain reservoir properties.

In the absence of a theory, using experimental physics (for example, rock property measurements in a laboratory environment) and/or numerical modelling, one can identify or validate suspected relationships. Although physics-based methods and direct measurements (the ground truth) are the ideal and reliable way to establish such correlations, for various reasons it is not always practical. Those reasons range from lack of known theories, difference between the laboratory environment and field environment (noise, scale, etc.) and the cost of conducting elaborate physical experiments.

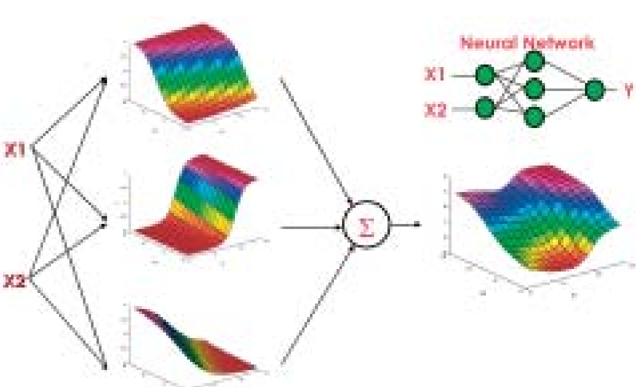

Figure 2 *MLP network with two inputs, three nodes in the hidden layer and one output. The training algorithm updates the weights, which corresponds to stretching and squeezing the sigmoid shape surfaces, such that the summation of the individual surfaces yields the optimal (non-linear) surface through the training examples.*

Statistics based methods aim at deriving an explicit or implicit heuristic relationship between measured values and properties to be predicted. Different statistical methods such as clustering (Aminzadeh and Chatterjee, 1984), regression analysis, cross-plotting, principal component analysis, cross correlation, geostatistics and neural networks (de Groot, 1995 and Aminzadeh *et al*, 2000), and fuzzy logic (Nikravesh and Aminzadeh, 2001) are used to establish a relationship between different seismic attributes, petrophysical data, laboratory measurements and different reservoir properties.

When performing clustering or regressions, we should examine the impact of noise, data population used for statistical analysis, scale, geologic environment, and the correlation between different attributes. Thus conclusions have to be re-examined and their physical significance explored. Moreover, the statistical correlations between the seismic attributes and reservoir properties from regression analysis or clustering should take into account the geologic environments (clastics vs. carbonates, formation age, type of sedimentation, pressure regime, trap type, etc.). Some perfectly strong correlations in certain areas may not apply to a different area. Also, mixing data from different geologic regimes may have an adverse effect on the results.

It can be argued that a hybrid method, combining the strength of the statistics and physics based method would be most effective. Furthermore, an examination of the discrepancies between the two methods can provide us with some insight into the uncertainty of the estimates. Figure 3 shows a conceptual view of such a hybrid approach. As will be discussed later, a good 'hybrid' method iterates and goes back and forth between the physics and statistics based methods to ensure cross validation.

Qualitative method

A distinction is made between the conventional attribute-based methods, the 'black box' approach of clustering or combining the two and 'meta attributes'. The mechanism to combine different attributes, such as clustering or neural net-

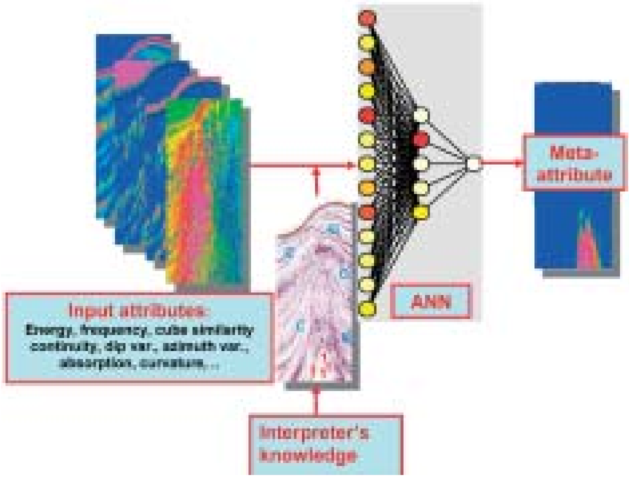

Figure 4 *Meta-attribute combining machine based multi-attribute and human knowledge*

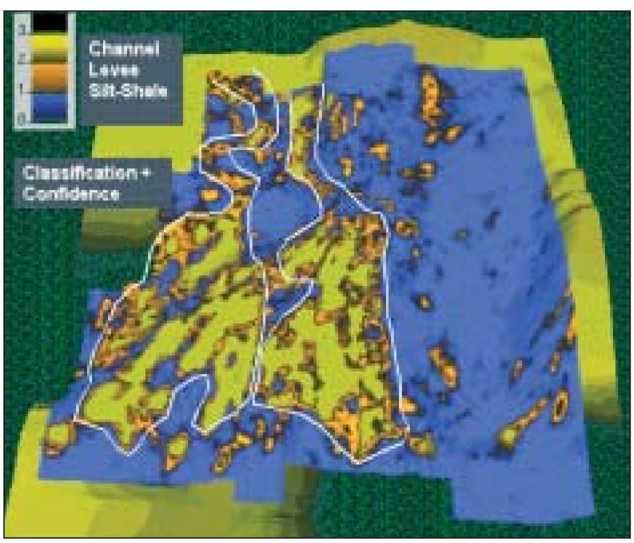

Figure 5 *Qualitative method of lithology classification using meta attributes*

works as shown in Figure 3, assist in the task of evaluating and visualizing the impact of different attributes on the output. However, these methods on their own can be considered a black box. Usually there is no possibility of incorporating the knowledge and insight of the interpreter in conventional clustering or neural network approaches

Meldahl *et al* (2001) introduced a method that forms the basis for the meta-attribute approach. Meta-attributes are versatile in the training process. The following are the main features of meta-attributes:

- Pre-processing through steering and filtering
- Combining attributes
- Non-linear mapping of attributes via use of neural networks
- Ability to use interpreter's knowledge to zero in on a specific object

Figure 4 shows the procedure, which is similar to a conventional way of using neural networks with the important addition of the 'Interpreter's knowledge' box. For example, let us assume the focus of the analysis is to determine three classes of lithologies: channel, levee and silt-shale in a 3D volume of seismic data. The first step is to examine the data set and identify areas with known or suspected lithologies from well entries, geologist interpretation from visual inspection of the data.

After attribute calculations and going through the training, testing and application phase, we can then create an implicit non-linear transformation of all the attributes that we can call 'lithology attribute'. In an ideal situation, a lithology attribute should highlight only those areas within the 3D volume that correspond to one of the three lithologies we intend to classify. Practically, we create a three class 'lithology probability' volume with large probability values associated with those areas that have closer overall 'likeness' to the given lithologies from the training set. Figure 5 shows such a lithology classification, with the colour intensity in each class representing the associated uncertainty in the classification.

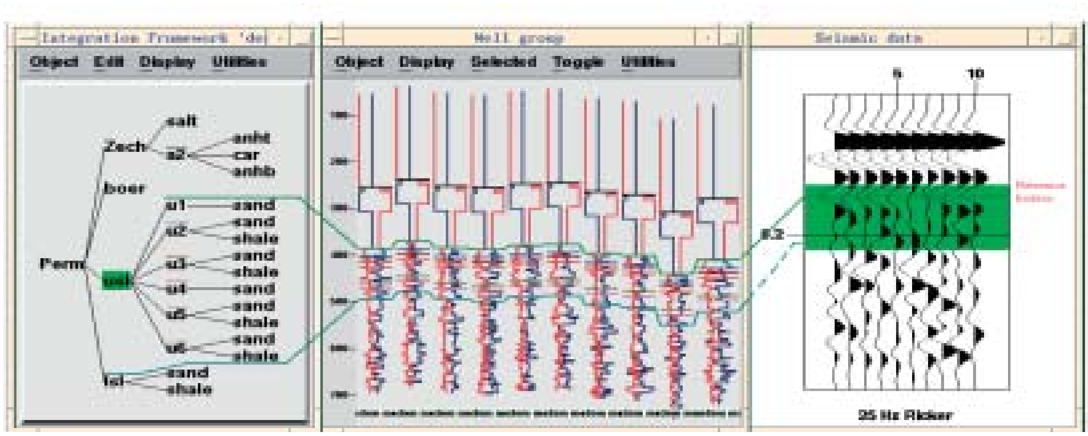

Figure 6 *Pseudo well generation to determine sensitivities of seismic response to changes in well properties*

This method can be useful for qualitative analysis of seismic data and highlighting anomalies of interest that can help reduce exploration risk. This would include highlighting stratigraphic features, structures (such as salt or faults) or vertical gas migration (chimneys). Combining different volumes of 'seismic object' data can help predict hydrocarbon phase and charge efficiency, distinguish between sealing and leaking faults and determine dynamic changes in the reservoir or accumulation of hydrocarbon next to the flank of a salt dome. For example, chimney cubes in conjunction with fault cubes can reveal where hydrocarbons originated, how they migrated into a prospect, and how they spilled or leaked from the prospect to create shallow gas, mud volcanoes or pockmarks at the sea bottom.

Quantitative method

de Groot (1995) developed a neural networks based method to test different hypotheses regarding geology of a reservoir by generating pseudo-wells. Figure 5 shows a number of stochastically simulated pseudo wells to determine sensitivities of seismic response to different reservoir properties and to validate statistics based results. The synthetic seismograms in Figure 6 (to the right) are generated for the corresponding pseudo wells. Seismic character within the time gate, or the reservoir interval where perturbations are made, can then be linked to different geologies, reservoir properties, or lithologies.

By combining qualitative and quantitative methods, one can use all the available data and the known physical relationship (through modelling) as well as the geoscientist's knowledge and insight. Specifically, using the statistical information from clustering, regression, kriging, and neural networks, in conjunction with geological interpretation, we establish the initial relationship between seismic attributes or seismic characters and reservoir properties (qualitative method). We then use real wells, pseudo-wells, acoustic and elastic inversion volumes as well as other geological and production data to confirm and/or modify such relationships. The procedure can be iterated to further establish the match between the statistics and physics based methods. We also evaluate contributions of different data components to the final results and study their impact on the uncertainty ranges.

Figure 7a demonstrates how information from seismic characters combined with other data (such as acoustic and elastic impedance volume and information from different wells and pseudowells can be combined to create an output highlighting certain reservoir properties. For example, Figure 7b shows segmentation of different classes of seismic characters or patterns and distribution of those patterns for different lithologies derived from a set of pseudo wells similar to those described in Figure 6.

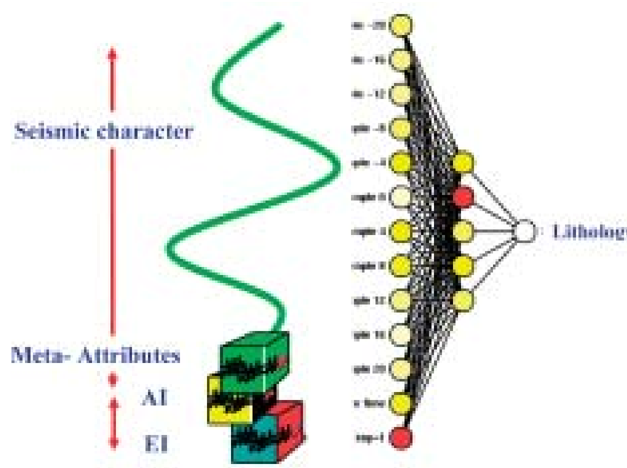

Figure 7a *Neural network as an integrator of data to predict a reservoir property*

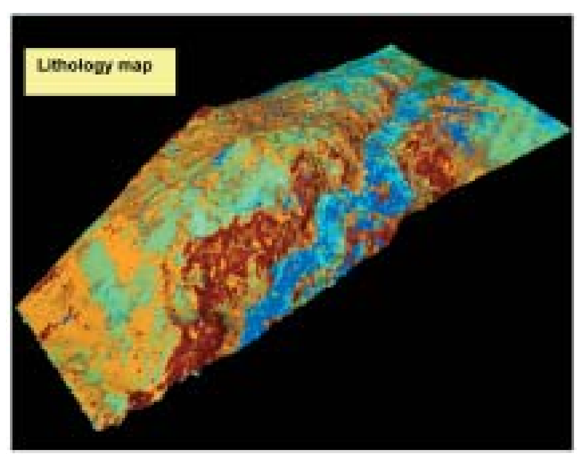

Figure 7b *Distribution of different seismic character classes for different lithologies.*

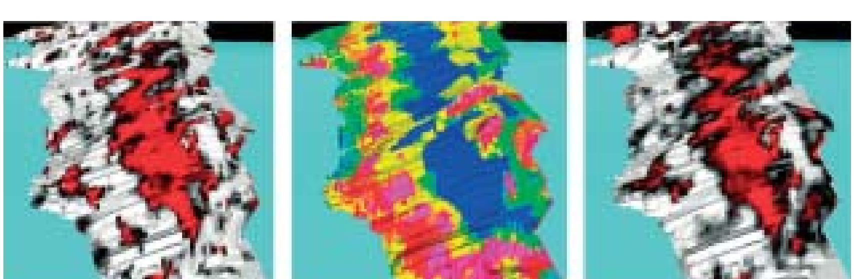

Figure 8 *Comparison of 4D analysis using different multi-attribute techniques. All views are top-reservoir displays taken from 3D volumes. Left: multi-attribute match (a vector differencing technique), centre: unsupervised neural network segmentation, and right: supervised neural network 4D anomaly 'probability' (the meta attribute approach described above).*

Dynamic changes in reservoir properties

Another important issue in reservoir characterization is to detect and monitor changes in reservoir properties with time. We should recognize the inherently static nature of some reservoir parameters (e.g. porosity) versus the dynamic nature of others (e.g. fluid saturation) and use the right set of data for appropriate predictions. For example, Oldenziel *et al* (2003) use a combination of two different vintage data sets to predict porosities while using data from different temporal measurements (time lapse data) to predict changes in the fluid fronts. Another recent improvement in detecting changes in reservoir properties was reported by Meldahl *et al* (2001), where multi-attributes were used to highlight changes in the reservoir, instead of the conventional amplitude differencing. Use of the meta attribute concept further enhances results where known changes in the reservoir are also employed. Figure 8 shows a comparison of results of the different techniques.

Conclusions

We discussed the existing challenges in going from collected seismic and log data to description of reservoirs. We suggested use of a hybrid method (physics+statistics) for relating seismic character to reservoir properties. Quantification of uncertainties of predictions and associated confidence levels are of paramount importance. Reduction of uncertainties is possible by a more comprehensive use of the available data as well as utilizing vector wavelets, which involves classification of an ensemble of angle gathers. For example, in the case of near, mid and far offset gathers, we are dealing with a vector wavelet with three elements rather than a single element in the case of full-stack data. Proper treatment of dynamic change in reservoir properties was also discussed and a few suggestions were made.

References

Aminzadeh, F. and Chatterjee, S. L. [1984] Application of Clustering in Exploration Seismology, *Geoexploration,* **23**, 147-159.

Aminzadeh, F. *et al*, [2000] Reservoir Parameter Estimation Using a Hybrid Neural Network, *Computers and Geoscience*, **26**, 860-875.

de Groot, P. F. M. [1995] *Seismic Reservoir Characterization Employing Factual and Simulated Wells,* PhD Thesis, Delft University Press.

Oldenziel, T. [2003] *Time-lapse seismic within reservoir engineering,* PhD Dissertation, Delft University.

McCormack, M. D. [1991] Neural computing in geophysics, *The Leading Edge,* **10**,1, 11-15.

Meldahl, P.,Heggland, R., Bril, B., and de Groot, P, [2001] An iterative method for identifying seismic objects by their texture, orientation and size, *Extended Abstracts, 71st Annual SEG Meeting,* San Antonio.

Nikravesh, M., Aminzadeh, F., Zadeh, L. A. [2003] *Soft Computing and Intelligent Data Analysis,* Elsevier.

Nikravesh, M. and Aminzadeh, F. [2001] Mining and Fusion of Petroleum Data with Fuzzy Logic and Neural Network Agents, *Journal of Petroleum Science and Engineering*, **29**, 221-238.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J. [1986] *Learning internal representations by error propagation, Parallel Distributed Processing.* Editors: Rumelhart, D.E., McClelland, J.L. and the PDP Research group, 318-362, Cambridge, MA, MIT Press.

Sandham, W., Leggett, L and, Aminzadeh, F. [2003] *Applications of Artificial Neural Networks and Fuzzy Logic,* Kluwer Academic Publisher.

Wong, P.M., Aminzadeh, F., and Nikravesh, M. [2002] *Soft Computing for Reservoir Characterisation and Modeling,* Studies in Fuzziness and Soft Computing, ed. Physica-Verlag, Springer-Verlag,