Carbonate reservoir characterization with lithofacies clustering and porosity prediction

Abdulrahman Al Moqbel and Yanghua Wang

Centre for Reservoir Geophysics, Department of Earth Science and Engineering, Imperial College London, South Kensington, SW7 2BP, UK

Abstract

One of the objectives in reservoir characterization is to quantitatively or semi-quantitatively map the spatial distribution of its heterogeneity and related properties. With the availability of 3D seismic data, artificial neural networks are capable of discovering the nonlinear relationship between seismic attributes and reservoir parameters. For a target carbonate reservoir, we adopt a two-stage approach to conduct characterization. First, we use an unsupervised neural network, the self-organizing map method, to classify the reservoir lithofacies. Then we apply a supervised neural network, the back-propagation algorithm, to quantitatively predict the porosity of the carbonate reservoir. Based on porosity maps at different time levels, we interpret the target reservoir vertically related to three depositional phases corresponding to, respectively, a lowstand system tract before sea water immersion, a highstand system tract when water covers organic deposits and a transition zone for the sea level falling. The highstand system is the most prospective zone, given the organic content deposited during this stage. The transition zone is also another prospective feature in the carbonate depositional system due to local build-ups.

Keywords: carbonate reservoirs, reservoir characterization, lithofacies clustering, porosity estimation, self-organizing map, back propagation algorithm

1. Introduction

How to accurately mapping the reservoir properties including heterogeneity is a key element in reservoir characterization. The artificial neural networking approach has some advantages of handling the reservoir nonlinearity over conventional methods. The latter sometimes fails to accurately classify the reservoir or to estimate its properties due to the complexity of reservoirs, especially for carbonate reservoirs, which pose real challenges in petroleum exploration. For a carbonate reservoir, the heterogeneity nature of the reservoir adds a complexity that can fail the reservoir prediction.

We adopt neural networking approaches to resolve the reservoir heterogeneity. We first explore the clustering capacity of the self-organizing map (SOM) algorithm, to classify the reservoir lithofacies. SOM with non-supervised learning performs a feature projection nonlinearly from the high-dimensional (input) space into a low-dimensional (output) space (feature space) consisting of 2D array of neurons in an orderly fashion (Kohonen 1989, Aras et al. 1999). In addition, the method also has the ability to cluster, classify and visualize the data which makes it a useful approach to characterize the reservoir heterogeneity.
We then use a back-propagation (BP) neural networking algorithm to estimate the porosity from seismic attributes. BP is a feed-forward neural network with supervised learning and is one of the most popular approaches for modelling and computer learning (Looney 1997). It is an iterative process to estimate neural connection strengths (the weights) by minimizing an error function based on a gradient-descent technique (Bishop 1995, Kosko 1996, Haykin 1999). Both SOM and BP neural networking methods have been used successfully in reservoir studies (Huang et al., 1996; Huang and Williamson, 1997; Zhang et al., 2000; Helle et al., 2001). However, we here focus on a carbonate reservoir, for its lithofacies and porosities.

The main input to the neural networks is a 3D seismic cube which covers an area of approximately $40 \times 40 \, \text{km}^2$ (figure 1a). We select 401 inlines (in the West–to–East direction) and 401 crosslines (in the South–to–North direction) from the data cube, with line interval of 100 m in each direction. We focus on the time window between 700 and 1500 ms, with a sample rate of 4 ms. The study area covers eight wells, with different production classifications: high, moderate and low. Figures 1b and 1c display the inline and crossline seismic profiles through well A15.

\[a\] A basemap of the study area coverage and the wells production classification. \[b\] The inline section through well A15. \[c\] The crossline section through well A15.
2. Self-organizing map (SOM) for lithofacies clustering

The SOM architecturally has two layers, the input layer and the output layer, which form a feed-forward structure from the (input) data space to the (output) feature space. The input layer presents different input patterns to the network, from which the network learns and then generalizes. The output layer consists of neurons (or nodes) that are regularly spaced and arranged in a rectangular grid. Both layers are fully connected and each connection is given an adjustable weight. The weight adjustment is determined only by the difference between the input pattern and the neuron which is closest to it (the winner).

When an input vector \( x \) is introduced to the network, the competition among neurons located on a discrete lattice is enforced. The winner of competition is the unit whose reference vector is the nearest to the input \( x \), as

\[
m_c = \min_i \{|| x - m_i ||^2 \},
\]

where \( m_i \) is the \( i \)th weight vector. The competitive learning algorithm allows not only the winning neuron \( m_c \) but also its neighbours on the lattice to learn. That is, the weight vectors \( m_i \) in the neighbourhood of the winner \( m_c \) are also modified, getting closer to the input vector \( x \). This is an adaptive feature of the algorithm. The weight vectors are changed iteratively according to an adaptation rule (Luttrell 1990) in the form,

\[
m_i(k+1) = m_i(k) + h_i(k)(x(k) - m_i(k)),
\]

where \( x(k) \) is the input at the \( k \)th iteration, and \( h \) is a neighbourhood kernel, a decreasing function of the distance of the units from the winning unit \( c \) on the map lattice (Kaski and Kohonen 1994). A widely applied kernel function is expressed in terms of a Gaussian function,

\[
h_{ic}(k) = \alpha_0(k) \exp\left(-\frac{|| r_i - r_c ||^2}{\sigma(k)^2}\right),
\]

where \( r_i \) and \( r_c \) are the two-dimensional vectors representing the locations of units \( i \) and \( c \) on the map grid, \( \alpha_0(k) \) is the learning rate parameter, and \( \sigma(k) \) is the width of the kernel. Both \( \alpha_0(k) \) and \( \sigma(k) \) decrease monotonically as the iteration runs.

As the mapping from the input (signal) space to the SOM (surface) space is based only on the internal relations in the structure of input signal, the ordering takes place automatically without external supervision. Therefore, the algorithm creates an internal representation of the incoming signal structure which ensures that similar patterns in the input space are correspondingly located near each other in the SOM space.

SOM consists of two phases, training and mapping. In the training phase, a square grid of neurons is initialized to train the network. Each node is represented by a weight vector of the same dimension as the input vector and the weights are initially set to random values. The network is fed with random input patterns in an iterative process from the high-dimensional input space to the SOM low-dimensional feature space.

Beginning with an initial state that correspond to a completely disordered map (figure 2a), the goal of the SOM training is to arrive to a state that corresponds to an ordered, “topology-conserving map” of the input space, in which some relevant lithofacies features of input patterns are two-dimensionally represented.

When the training starts, the neighbourhood is broad and the self organizing takes place on the global scale. However, the neighbourhood radius shrinks as the training continues where the weight update converges to local scale. Most of the nodes get the chance to be the best
matching unit to the introduced patterns. By the end of training phase, the resulting map is a topographically self-organized indicating that the network is trained and that the network is ready to map the whole input space data into the SOM feature space (figure 2b).

Figure 2. Training in self-organized map (SOM). (a) The initial weight vectors before training. (b) The weight vectors after training.

Figure 3. The lithofacies clustering map, generated by SOM. (a) The result based on 3D seismic data. (b) The result based on multiple attributes.

In the mapping phase, the goal is to estimate the main lithofacies of the 3D seismic data. As each input point corresponds to a point in the lattice in the feature space such that the neighbourhood relation between points is preserved. After running the algorithm on the seismic data, the resulting lithofacies in the area are shown in figure 3a with nine main types.

We also develop a version of SOM which can run multiple attributes as the input. Figure 3b is the lithofacies map generated when using the instantaneous amplitude, the instantaneous phase, the instantaneous frequency, and the acoustic impedance as the input. The results from the single input (seismic amplitude) and from the multiple attributes show similar general
trends in the area.

The results indicate that the low-producing wells (A01, A03, A09, A18 and A20) fall in a group of similar clusters on both figures (note different colours even for the same clusters in two figures due to different input to the network). These particular facies are expected to be non-hydrocarbon bearing types of facies such as mudstone. On the other hand, the intermediate-producing and high-producing wells (A15, A17, and A21) fall in another group of similar clusters. Lithofacies such as grainstone, wackestone and packstone are expected to be dominant in the area due to their hydrocarbon-bearing properties. This map is useful in choosing future planned drilling wells.


The BP neural network is a multi-layered feed-forward neural network consists of three layers: input, hidden, and output (figure 4). Neurons in the hidden and output layers perform main tasks to mimic the processes of the biological neuron: summation of the incoming signals, and their activation through a transfer function.

Mathematically, neuron $j$ has two equations for summation and activation, respectively. They can be written as

$$NET_{py} = \sum_{i=1}^{N} w_{pi} x_{pi},$$  \hspace{1cm} (4)

and

$$y_{pj} = f(NET_{pj} - \phi_{pj}),$$  \hspace{1cm} (5)

where $x_{p1}, x_{p2}, \ldots, x_{pN}$ are the input signals, $w_{ij}, w_{j2}, \ldots, w_{jk}$ are the synaptic weights of neuron $j$ within a total of $N$ samples that connect neurons in a layer to the neurons in the next layer, $NET_{pj}$ is the linear combiner (weighted sum) output, $\phi_{pj}$ is the threshold, $f$ is the activation function, and $y_{pj}$ is the output signal from the neuron.

The basic procedure for training the network consists of four steps. (1) Apply an input (data) vector to the network and calculate the corresponding output values; (2) Measure the
error by comparing the actual outputs with the outputs; (3) In order to reduce the error, determine in which direction (positive or negative) to change each weight, and the amount by which to change each weight; and (4) Apply the correction to the weights. The errors of the nodes of the output layer are back-propagated; this process gives the error BP neural network its name (Rumelhart and McClelland 1986).

This entire process is continued in an iterative manner. The network converges when its output is within acceptable proximity of the desired output. The computational capabilities of the network are sufficient for learning input-output mapping between patterns (Omlin and Giles 2000).

For predicting the porosity of the reservoir, we train the network using the input/output pairs at the eight well locations. We feed four input attributes (the instantaneous amplitude, the instantaneous phase, the instantaneous frequency, and the acoustic impedance, after normalization) into the network to estimate the reservoir porosity. The correlation coefficient between the original porosity values at the well locations and porosity values estimated by BP ranged between 67% and 97%. The overall correlation coefficient is approximately 85%.

However, a validation of the result is needed before generating these 2-D porosity maps to ensure that the network is generalizing and not over-fitting. Over-fitting, is also referred to as memorization, could be a common problem in the prediction of a neural network. Once a network memorizes a data set, it is incapable of generalization even though it fits the training data set very accurately. Over-fitting means getting acceptable prediction by the network for the data used in training, but unacceptable prediction for inputs that have never been used in the training stage (Hagan et al. 1996, Hart 1999). Over-fitting leads to statistically significant, but physically meaningless results. One good way for validation is the take-one-out approach. One well is excluded at a time from the training stage. The network is trained with the data from seven wells to predict the measurements from the eighth well (the well that has been excluded from that particular training). This practice is repeated eight times, each time well logs from a test well are not used in the training. Therefore, the accuracy of the network can be evaluated each time by comparing the predicted result (the testing well is not included in the training) with the actual well data at the excluded well (figure 5).

Figure 5. A comparison between the network prediction (gray curves) and the target porosity logs (black curves). This is a validation result using take-one-out method when well A09 is excluded.
Figure 6. Carbonate depositional system associated with sea-level fall (lowstand) and rise (highstand). The deep sea deposits are dominated by mud, while grainstone dominates the shallow deposits. One of the most important carbonate reservoir features is the reef, forming as a sheet dominated by high organic contents (corals/shoals), which makes its facies a good target in hydrocarbon exploration.

Once the system is fully trained and validated, we apply it to field seismic data, for estimating the porosity of the target reservoir. As carbonate rocks lack bedding and have higher seismic velocities than clastic sediments, as well as having lower acoustic impedance contrasts within a platform succession (Read 1985), carbonate facies are more difficult to predict seismically than clastic facies which usually are well-behaved and laterally predictable. The carbonate platform geometry forms complex structures with steep dips and usually contains evaporites (Schuelke and Rick 2003). The porosity of a carbonate reservoir is affected by the stratigraphic setting and the depositional facies (deposits).

In general, the main elements of a depositional carbonate system are ramp, shelf, reef and lagoon (figure 6). One of the key processes affecting the carbonate porosity is the transgression and regression of sea level (Jacquin et al. 1991). Transgression and regression are defined as "long-term sea rise" and the "long-term sea fall", respectively.

In the lowstand case, as the sea-level falls, the land is exposed, which allows the high-energy erosional processes to dominate. This environment is characterized by the deposition of grain-supported facies (Gadallah and Fisher 2005). On the other hand, in the highstand case, the land is immersed by water, and the energy is low on the ramp and medium at the shoreface. This environment is characterised by the deposition of mud facies. One of the most common hydrocarbon exploration targets is the reef depositional setting, which occurs in shallow marine environments with organic-dominated facies, and provides the best environment for high-porosity facies because of the good accommodation space.

We generate a series of 2D porosity maps at the different reservoir levels. As shown by figure 7, we divide the zone of interest (12 time slices) into three separate sub-zones, based on our cluster analysis using seismic data (not shown here). Zone 1 (upper) is comprised of time slices from 924 to 932 ms, zone 2 (the middle) is comprised of slices from 936 to 948 ms, and zone 3 (lower) is comprised of time slices from 952 to 968 ms. We interpret these three subzones as follows in a depositional sequence:

Zone 3 shows a relatively low porosity which could correspond to a lowstand system tract (before the sea water immersing).
Figure 7. The regional estimation of porosity for the different reservoir levels. Zone 3 with a relatively low porosity corresponding to a lowstand system tract (before the sea water immersing). Zone 2 has the highest porosity corresponding to a highstand system tract. Zone 1 represents a transition zone where local build-ups could make another prospective feature of the carbonate depositional system.

Zone 2 has the highest porosity which corresponds to a highstand system tract (when water covers organic deposits). Therefore, zone 2 is the most prospective zone, given the organic contents deposited during this stage.

Finally, zone 1 corresponds to a major drop of the sea level. It represents a transition zone
when the sea level starts dropping from the highest level. Local build-ups could make another prospective feature of the carbonate depositional system within this zone.

4. Conclusions

Carbonate reservoir characterization based on 3D seismic data is implemented in a two-stage approach, using respectively unsupervised and supervised neural network algorithms.

The unsupervised self-organizing map (SOM) is developed for the purpose of clustering the reservoir dominant hidden lithofacies. Results indicate that the reservoir is clustered to nine different heterogeneity-quality lithofacies. In general, the intermediate-producing and high-producing wells fall within different clusters that those of the low-producing wells. SOM conducts mapping based on the internal structure of the seismic data and attributes, and reveals an important piece of information that contributes in raising the confidence for future drilling locations, as well as, reducing the drilling risk.

Then back-propagation neural network has successfully integrated the seismic information with the geological knowledge, to predict carbonate reservoir porosity quantitatively based on seismic attributes, and helped to define the main trends at the different reservoir levels.

Acknowledgements

We are grateful to the sponsors of the Centre for Reservoir Geophysics, Imperial College London, for supporting this research.

References


Networks 1, 229–32.