# SEISMIC FACIES CLASSIFICATION AND IDENTIFICATION BY COMPETITIVE NEURAL NETWORKS

Muhammad M. Saggaf and M. Nafi Toksöz

Earth Resources Laboratory
Department of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

Maher I. Marhoon

Area Exploration Department
Saudi Aramco
Dhahran, Saudi Arabia

## ABSTRACT

We present an approach based on competitive networks for the classification and identification of reservoir facies from seismic data. This approach can be adapted to perform either classification of the seismic facies based entirely on the characteristics of the seismic response, without requiring the use of any well information, or automatic identification and labeling of the facies where well information is available. The former is of prime use for oil prospecting in new regions, where few or no wells have been drilled, whereas the latter is most useful in development fields, where the information gained at the wells can be conveniently extended to inter-well regions. Cross-validation tests on synthetic and real seismic data demonstrated that the method can be an effective means of mapping the reservoir heterogeneity. For synthetic data, the output of the method showed considerable agreement with the actual geologic model used to generate the seismic data, while for the real data application, the predicted facies accurately matched those observed at the wells. Moreover, the resulting map corroborates our existing understanding of the reservoir and shows substantial similarity to the low frequency geologic model constructed by interpolating the well information, while adding significant detail and enhanced resolution to that model.

## INTRODUCTION

Traditionally, seismic data have been used predominantly for determining the structure of petroleum reservoirs. However, there is also significant stratigraphic and depositional information in seismic data, and it is becoming increasingly essential to utilize this information to facilitate seismic facies mapping for delineating the extent of the reservoir and characterizing its heterogeneity, as most of the larger and simpler reservoirs have already been developed and many are already depleted, and ever smaller and more complex reservoir targets are therefore being prospected. This task of facies interpretation has been aided by considerable recent advances in seismic recording and processing technology, such as the more widespread acquisition of 3D surveys and the advent of processing methods that help preserve the integrity of relative amplitudes in the seismic data.

In order to utilize seismic data to characterize the heterogeneity of the reservoir, seismic facies are interpreted by inspecting the data in various regions of the reservoir and grouping them together based on the characteristics of the seismic response in these regions. This grouping, or classification, is performed either visually (by examining the raw trace itself or attributes derived from that trace), with the help of graphical aids like cross-plots and star diagrams, or by automatic techniques. Matlock *et al.* (1985) used a Bayesian linear decision function to identify the boundaries of rapidly varying sand facies; Hagan (1982) employed principal component analysis to study the lateral differences in porosity; Mathieu and Rice (1969) used discriminant factor analysis to determine the sand/shale ratio of various zones in the reservoir; Dumay and Fournier (1988) employed both principal component analysis and discriminant factor analysis to identify the seismic facies; Simaan (1991) developed a knowledge-based expert system to segment the seismic section based on its texture; and Yang and Huang (1991) used a back-propagation neural network for detecting anomalous facies in the data.

Manual interpretation of seismic facies is a labor-intensive process that involves the expenditure of a considerable amount of time and effort by an experienced stratigraphic interpreter, even with the aid of graphical techniques like cross-plotting. The problem becomes especially more difficult as the complexity of the seismic data increases or the number of simultaneous attributes to be analyzed is raised. Automatic methods that rely on hard-coded *a priori* prescribed patterns and implicit cross-relations are sensitive to inconclusive data and are often unreliable with noisy data or atypical environments. Linear decision functions cannot adequately handle the inherent nonlinearity of the seismic data, and methods that rely on multi-variate statistics are inflexible, require a large amount of statistical data, and often need complex dimension-reduction techniques like principal component and discriminant factor analyses that are themselves compute-intensive and inflict an unnecessary distortion on the input data representation by projecting the input space to a new vector space of a lesser dimension.

Neural network methods are in general superior to knowledge-based and rule-based expert systems, as they have better generalization and fault tolerance. Better gener-

alization means that for similar inputs, outputs are similar. The network is able to generalize from the training data so that its output is similar for input data that are almost, but not quite, the same. Conventional expert systems, on the other hand, have no concept of similarity, since they are based on logic, and the concept of similarity cannot be ascribed to logical rules. Fault tolerance signifies that a small distortion of the input data produces only a proportionally small performance deterioration. In other words, noisy or incomplete input data would not produce considerable deviation in the inference results, a problem often encountered in expert systems. Saggaf *et al.* (2000a, 2000b) utilized regularized back-propagation networks to estimate the point-values of porosity in the inter-well regions of the reservoir. Although back-propagation networks are effective in problems involving point-value estimations, they are not well suited to the task of discrete facies classification and identification, thus their performance for this task is often less than optimal. A further shortcoming of the above approaches is that they require *a priori* training examples (usually drawn from seismic traces gathered near well locations, where the stratigraphy is reasonably well-known) to perform their analysis, and the accuracy of the final outcome is significantly dependent on the proper choice of these training examples.

In this paper, we approach the problem of automatic classification and identification of seismic facies through the use of neural networks that perform vector quantization of the input data by competitive learning. These are uncomplicated one-layer or two-layer networks that are small, compute-efficient, naturally nonlinear, and inherently well suited to classification and pattern identification. Not only can the method perform quantitative seismic facies analysis, classification, and identification, but also it is able to calculate for each analysis confidence measures that are indicative of how well the analysis procedure can identify those facies given uncertainties in the data. Moreover, this approach can be conveniently adapted to problems where prior training examples are present, such as development fields where numerous wells have already been drilled, as well as cases with no training data, such as exploratory areas where few or no wells are available. We describe next the methodology and implementation of the approach in these two scenarios.

## METHOD DESCRIPTION

The technique described here can be carried out in either of two modes: unsupervised and supervised. In the former, no well information is used, and the seismic data is classified without labeling the resulting classes, whereas in the latter, well information is utilized within the technique to assign and label the resulting outcome, thus identifying (not just classifying) the seismic facies.

### Unsupervised Analysis

The objective of unsupervised analysis is to classify seismic data into regions of distinct characteristic seismic behavior without making use of any extra information like

well logs. The classification is based entirely on the internal structure of the seismic data; the resulting facies map is indicative of the reservoir heterogeneity (e.g., channel limits and orientation, or high and low porosity regions), but those classes are not labeled. Additional information may be used to identify each of the classes (for example, by inspecting the impedance contrast in the typical traces of each class and identifying the high porosity region with class whose typical seismic response indicates higher impedance). Such analysis is of prime use for oil prospecting in new regions, where no well information is available.

Unsupervised analysis is implemented by a single-layer competitive network that takes the entire seismic trace as input and maps the location at which this trace was gathered into its corresponding seismic facies category. The input/output architecture of the network is shown in Figure 1a. This mode is often called feature-discovery or a "let-the-data-talk" scheme. It performs clustering, or quantization of the input space. The competitive layer has a number of neurons equal to the desired number of clusters. Individual neurons are considered class representatives and they compete for each input vector during training. The neuron that resembles the input vector the most wins the competition and is rewarded by being moved closer to the input vector via the instar or Kohonen learning rules (Kohonen, 1989). Thus, each neuron migrates progressively closer to a group of input vectors, and after some iterations, the network stabilizes, with each neuron at the center of the cluster it represents. Biases are introduced during the training phase to ensure that each neuron is assigned to some cluster, and to distribute the neurons according to the density of the input space. The technique is described in more detail in the appendix.

Figure 2 shows a schematic of the above process, where two neurons are utilized to classify a group of input vectors in the plane. The neurons start at the same location (Figure 2a), then each one migrates gradually toward a group of input vectors (Figure 2b), until eventually it stabilizes at the center of a cluster of input vectors to which it is most similar (Figure 2c). The neuron is considered the class representative of that group, and it exhibits typical behavior of the input vectors in that cluster. The network convergence is continuously monitored, and the training is stopped once the neurons stabilize and no longer migrate appreciably, or when a prescribed maximum number of iterations has been exhausted.

To measure the resemblance between the input vectors and each neuron, either distance norms, such as the $l_1$ and $l_2$ norms, or correlation norms, such as the cross-correlation and semblance, may be used (note that the term "norm" is being used loosely). Table 1 shows the definition of each of these norms (the $l_2$ norm is the familiar Euclidian distance). The implementation differs slightly in the detail when correlation norms are used instead of distance norms, since the correlation is directly proportional to the resemblance between the vectors (a unity implies perfect resemblance), while the distance is inversely proportional (perfect resemblance is indicated by a distance of zero). Thus, for distance norms, the inverse or negative of the norm is used as a measure of the similarity between the vectors.

| Norm | Definition |
|---|---|
| $l_1$ | $\|x-y\| = \sum_{i=1}^{n} \|x_i - y_i\|$ |
| $l_2$ | $\|x-y\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$ |
| Cross-correlation | $\|x,y\| = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \cdot \sqrt{\sum_{i=1}^{n} y_i^2}}$ |
| Semblance | $\|x,y\| = \frac{1}{2} \frac{\sum_{i=1}^{n} (x_i + y_i)^2}{\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} y_i^2}$ |

Table 1: Definition of some of the distance and correlation norms that can be used to compute the resemblance between neuron and input vectors.

The selection of the particular norm to utilize depends on the properties of the data on which the interpreter wishes to base the analysis. For example, within the distance norms, the $l_1$ norm penalizes outliers to a greater extent than does the $l_2$ norm, and so the selection of the former would be more appropriate if numerous outliers are present in the data and suppressing them is imperative. Correlation norms are insensitive to the absolute amplitudes of the seismic data and depend only on the shape of the trace, whereas distance norms are sensitive to variations in the amplitude as well as the shape of the seismic traces. Thus, distance norms are more suitable to detecting properties manifested in amplitude changes, while correlation norms are more suited to shape-only characteristics. An example is shown in Figure 3, where two traces have the exact same shape but quite different amplitudes, and their cross-correlation and semblance are therefore unity, although their $l_1$ and $l_2$ norms are nonzero. Additionally, correlation norms are in general dependent on accurate picking of the interpreted horizon from which the analysis begins, and shift errors due to poor picks can adversely affect the analysis when such norms are utilized. This shortcoming can be alleviated somewhat, however, by applying a taper to the seismic traces prior to the analysis. This de-emphasizes the first few samples of the trace and reduces the sensitivity to mispicks of the horizon.

When instantaneous seismic attributes are to be utilized in the analysis, the input/output architecture we described above (Figure 1a) can be extended so that multiple attributes are fed to the network. A schematic of this extended architecture is shown in Figure 1b. Since the various attributes have in general different ranges in their values, the input in this case must be normalized such that all attributes have fairly similar ranges. Otherwise, the attribute with the highest absolute values will dominate the analysis when distance norms are utilized, as these norms would be disproportionately affected by that attribute relative to the other attributes. The normalization can be done by scanning the extracted attributes for their minimum and maximum values, and then applying a transformation based on these extrema that scale each attribute to be within the interval $-1$ to 1. Note that this concept may also be used to artificially emphasize one attribute to be more prominent in the analysis than the others, if desired. This can be achieved by simply adjusting the transformation such that the values of that attribute are mapped to a larger interval than that of the other attributes.

Finally, when interval attributes, rather than instantaneous attributes, are to be employed, a slightly different input/output architecture is needed (shown in Figure 1c, where the input to the network consists of a vector containing one attribute at each sample point). The same normalization and relative emphasis comments made above apply here as well. Additionally, the correlation norms cannot be used in this case, as each sample in the input vector comes from a different attribute. Therefore, the shape of the input reflects only the rather arbitrary ordering of the attributes in the input vector and conveys no useful information.

## Supervised Analysis

The objective of supervised analysis is to characterize the reservoir by identifying its facies based on well log and seismic information. The network is first trained on facies information available from the logs and the seismic traces close to the wells at which those logs were measured. Once this training is completed, the network is applied to the seismic data in the inter-well regions to predict the facies between the well locations. The boundaries of the reservoir (e.g., sand/shale border, sand pinch-out, or high porosity region) can thus be mapped. Mapping is of prime use especially where 3D seismic surveys are available, as an aerial map of the reservoir limits may be extracted from the seismic survey using this technique.

Supervised analysis, often called guided or directed classification, is implemented by a two-layer network. The first layer is a competitive network like the one mentioned previously that pre-classifies the input into several distinct subclasses. The second layer is a linear network that combines those subclasses and associates them with the final target facies. The neurons are fully connected between the two layers, and the objective of the linear layer is to assign (aggregate) several of the competitive neurons into each of the target facies at the output of the network. Numerous subclasses are usually used in the competitive layer since a single facies is allowed to encompass two or more distinct characteristics in the data. This behavior would not have been accommodated had the number of subclasses in the competitive layer been equal to the number of target facies.

The facies identification procedure is illustrated in Figure 4a. To identify the facies present at a particular location in the field, an input vector consisting of the seismic trace at that location is presented to the network. The vector is assigned to the competitive neuron to which it is most similar (this is indicated by the black neuron in the figure). The final target facies $T_2$ at the output of the network encompasses the subclass defined by this winning neuron as well as other subclasses (designated by the gray neurons in the figure). The linear layer maps all those subclasses to the final target facies $T_2$ at the output of the network. Thus, the location is classified in this case to be of facies $T_2$. In contrast to the unsupervised analysis, the final outcome here is automatically labeled with the facies names according to the behavior learned during training, in which the subclasses of the competitive layer were assigned to final target facies. Thus, in supervised analysis, the seismic facies are actually identified rather than just classified, and the class labeling is accomplished intrinsically within the technique. However,

well log information needs to be utilized during training, whereas it is not needed in unsupervised analysis.

The learning rule is modified such that the winning neuron is moved closer to the input vector only if the subclass defined by that neuron belongs to the target facies of the input vector. Otherwise, the neuron is moved away from the input vector. Thus, competitive neurons move closer to the input vectors that belong to the target facies of those neurons, and away from those that belong to other target facies. After some iterations the network stabilizes, with each neuron in the competitive layer at the center of a cluster, and each group of such neurons (subclasses) mapped to a certain target facies. The same comments made in the previous section regarding the input/output architectures, type of norm utilized, and data normalization apply here as well. The technique is described in more detail in the appendix.

## CONFIDENCE MEASURES

One of the useful pieces of information generated by our method is a measure that describes the degree of confidence in the analysis. Humans describe confidence qualitatively. For example, a stratigraphic interpreter inspecting a seismic survey may say that the data is inconclusive at a particular region of the field and so his analysis there is suspect. In other words, his confidence of the analysis is low. However, in another region of the field, the data has distinct characteristics, and so he has a lot of confidence in the analysis of that region. The same type of information can be conveyed in this method only quantitatively. The method actually generates three types of confidence measures: distinction, overall similarity, and individual similarity.

### Distinction

Distinction is a measure of how distinct the seismic behavior of the identified facies is from the rest of the seismic data characteristics. A high distinction means that the seismic response at that location has behavior that is quite distinct and far away from the seismic behavior of all facies except the one to which that location belongs. Such a point plots on the trace cross-plot far away from all other cluster centers but its own. A low distinction means that the data are inconclusive. In fact, zero distinction indicates that the location could be ascribed to two different facies equally well. For distance norms, this measure depends on the relative distances in the vector space from a particular point representing the seismic trace to the two nearest cluster centers. Figure 4b is a schematic that shows an example of the direction of decreasing distinction. As the distances from the point to the two nearest cluster centers become equal, the distinction measure vanishes. The interpretation is similar when correlation norms are utilized, but the measure is computed in a slightly different manner. The distinction confidence measure is described further and defined for both categories of norms in the appendix.

### Overall Similarity

Overall similarity is a measure of how similar the seismic behavior of the identified facies is to the class representative. A high overall similarity means that the seismic response at that location has behavior that is similar to the behavior typically exemplified by the identified facies. Such a point plots on the trace cross-plot quite close to its class representative. This measure depends on the distance in the vector space from a particular seismic trace to the nearest cluster center. Figure 4c is a schematic that shows that two points may have comparable similarity but different distinction. The overall similarity confidence measure is described further and defined for both categories of norms in the appendix.

### Individual Similarity

Individual similarity is akin to overall similarity, except a measure is generated for each facies type—not just the one identified. At each well location, the identified facies will always have the highest individual similarity measure. In fact, these facies were identified precisely because the seismic behavior at that location has the highest similarity to the class representative of those facies. Individual similarity measures can be regarded as indicators of the relative composition of the facies at each well location.

## SYNTHETIC DATA EXAMPLE

We illustrate the use of our method by applying both unsupervised and supervised analyses on a simple synthetic example. A geologic model was constructed to simulate a fluvial environment consisting of horizontally-layered bedding with a sand channel embedded in otherwise shaly sediments. This model is shown in Figure 5a. Synthetic seismic traces were then generated with zero-offset image rays, and the data were subsequently encumbered by 10 db random Gaussian noise to reflect typical recording circumstances. The generated seismic cross section is shown on Figure 5b. This data was then fed to an unsupervised competitive network to classify the section laterally according to the characteristics of the seismic response. The network consisted of a two-neuron layer, and the $l_2$ norm was utilized. The result of this unsupervised analysis is shown in Figure 6a.2, where it is compared to the actual model (Figure 6a.1). The two maps show considerable agreement. In fact, they are for the most part identical except at the channel edges, where the seismic data are naturally inconclusive due to the inherent limitations of the seismic resolution. Figure 6b shows the two class representatives generated by the method. As we mentioned previously, these can be considered the typical seismic responses of the two output classes.

The distinction and overall similarity confidence measures generated by the method are shown in Figures 6a.3 and 6a.4, respectively. We see that confidence in the analysis was reasonably high throughout the section except at the class boundaries, where confidence drops sharply since the data are inconclusive there. The confidence measures

convey vital information regarding the validity of the prediction at each location and they should always be considered an integral and indispensable part of the solution. For example, Figure 6a.3 clearly implies that the analysis output at the edges of the classes should not be considered to have the same validity as the rest of the output, and that the classification accuracy at these edges is rather suspect. Thus, one should be cautious in those two locations, and important decisions regarding drilling wells there must be taken into account.

Note that although the analysis was able to predict the channel location and extent quite accurately, the output (Figure 6a.2) does not specify which of the two classes is the channel. Further work needs to be done to ascertain this. For example, it is evident from Figure 6b that the typical seismic response of Class 2 exhibits more impedance contrast, and one would be able to conclude that Class 2 represents the channel. This manual step can, however, be avoided in supervised analysis, where the method is presented with prior examples of each facies. Two traces were selected for that purpose, one in the shaly deposits and another in the channel about halfway between its center and its left edge. The result of supervised analysis is shown in Figure 7a.2, where it is compared to the actual model (Figure 7a.1). The locations of the two training traces are also indicated in the figure. The actual and predicted maps again show considerable agreement (except at the edges of the channel, where the data are inconclusive), but this time the analysis output is automatically labeled with the actual facies names. There is no need here to inspect the class responses and manually infer class labels. Of course, the price paid is that example traces were needed to train the supervised network, whereas none were utilized for the unsupervised analysis.

Supervised analysis represents a significant improvement over direct methods that simply compute a norm between each seismic trace in the data and the training (so-called pilot) traces and classifies the data accordingly, thus partitioning the input space by spheroids each centered at one of the pilot vectors. There are two key advantages of the supervised technique we presented here over such direct methods. First, there may be several distinct characteristics in the data that belong to the same target facies. The supervised analysis handles this naturally by optionally utilizing numerous subclasses in the competitive layer. The relation between these distinct subclasses and the target facies may not be obvious, and so manually picking several pilot traces for each target facies in the direct method becomes unwieldy as the number of these subclasses (distinct characteristics) increases.

The second, and more important, advantage is that the selected pilot traces may define class representatives that are not at the optimal centers of the data clusters but are rather at the fringes. This does not present a problem for the supervised analysis described here (where the first layer pre-classifies the input and the second layer aggregates these subclasses), since the training traces are used only for labeling the classes and do not enter directly in the classification process. This makes the method much less sensitive to the choice of training traces employed in the analysis. For example, picking any training trace within the channel gives rise to the same eventual

class representative for the channel cluster (Figure 6b), since that class representative is computed by the first layer of the network and is not affected directly by the training trace.

On the other hand, the pilot traces are used firsthand in the direct methods and their choice defines the centers of the spheroids that partition the input space. In other words, the pilot traces are themselves the class representatives. Therefore, these pilot traces are involved directly in the actual classification process, and the final outcome is greatly influenced by their selection. When these pilot traces do not coincide with the optimal cluster centers of the data (and they seldom do, since their choice is subjective), the boundaries between the different predicted classes are erroneous, and incorrect classification ensues. This is illustrated in Figure 7b, which depicts an example where the optimal cluster centers of the data do not fully overlap the classes defined by the pilot traces, since those traces, although at the fringes of the optimal clusters, define the direct classes to be centered around them. The supervised method, on the other hand, always computes and utilizes the optimal cluster centers of the data.

In fact, the pilot trace for the channel area in Figure 7a was purposely selected to be less than optimal (near the edge of the channel rather than in the middle). The two pilot traces are shown in Figure 8b. The result of the supervised analysis (shown previously in Figure 7a.2 and again in Figure 8a.1) was not adversely affected by this rather inferior but typical choice of the channel training trace. However, when the direct method is applied, the substandard selection of the pilot trace has a considerable negative impact on the accuracy of the outcome. The result of the direct method is shown in Figure 8a.3, where it can be seen that significant portions of the section have been misidentified.

Figure 8a.4 shows the difference for each trace in the survey between the Euclidian distances from that trace to the two pilot traces (the class representatives for the direct method). Thus, the channel area is identified whenever that difference is positive and the shaly facies is identified otherwise (note that the peaks in the figure correspond to the locations of the two pilot traces). Figure 8a.2 shows a similar plot but for supervised analysis, where the distances are computed from each trace in the survey to the two class representatives generated by the first layer of the network. Comparison of Figures 8a.2 and 8a.4 shows that the separation difference for the supervised method has a regular and consistent behavior that does not oscillate around the border of the two facies, whereas the corresponding separation difference for the direct method behaves erratically and is close to the zero line (border between the two facies) almost throughout the shaly area. It is not surprising, then, that the facies prediction of the direct method was quite poor.

We presented the above synthetic example to demonstrate the methodology and illustrate the difference between the two modes of analysis, although this case is simple enough that the seismic facies can actually be deduced by visual inspection of the data (Figure 5b). In general, though, the application of this method remains simple and straightforward as the complexity of the geology increases, while the accuracy of manu-

ally analyzing the data deteriorates quickly with environments that are more complex. We investigate a real data example in the next section.

## APPLICATION TO FIELD DATA

We applied the method presented above to classify and identify high porosity and low porosity facies at the top of the Unaizah Formation in the Ghinah Field of central Saudi Arabia. The Unaizah Formation (Permian) is a result of a complex succession of continental clastics, and is bounded by the Pre-Khuff and Pre-Unyazh unconformities. The rock composition consists of interbedded red sandstone and siltstone, and the reservoir facies change between braided streams, meandering streams, and eolian dunes. During the late upper Permian time a regional erosional surface with incised valleys developed on the exposed paleoshelf due to a significant fall in sea level. At that time, uplift occurred in the source area and produced large amounts of fine to very coarse sands. These sands were transported by incised valleys across the shelf, and were deposited as low stand deltaic wedges. During relative sea level rise, good quality reservoir sands were deposited by aggradations within the lower and middle parts of the channels. The very fine-grained carbonaceous sandstones and shales that were deposited in a marginal marine environment cap the good quality reservoir sands. Deposition within the valleys terminated at the same time, and shallow marine shales, argillaceous limestones, dolomites, and anhydrites of the Khuff formation during maximum sea level high stand covered the entire area. More detail about the depositional environment of the Unaizah Formation is given by Alsharhan and Kendall (1986), Al-Laboun (1987), McGillivray and Husseini (1992), and Evans *et al.* (1997).

The 3D seismic data consist of 181 sublines by 261 crosslines, for a total of nearly 50,000 traces, each sampled at 2 ms. The subline spacing is 50 m, while the CDP spacing is 25 m, and the total survey area is approximately 9 x 6.5 km. The data had been processed with special attention to preserving the true amplitudes, and was time-migrated after stacking. Additionally, the data were also filtered post-stack by the spectral compensation filter based on fractionally integrated noise described by Saggaf and Toksöz (1999) and Saggaf and Robinson (2000). This filtering helped to improve the event continuity, wavelet compression, and signal resolution of the data, and to attenuate persistent low frequency noise that was otherwise difficult to reduce. The top of the Unaizah Formation was interpreted and the horizon was picked manually. Twenty-eight milliseconds of seismic data beneath this horizon were utilized in the study a layer that encompasses the main region of interest in this zone. Thus, each seismic trace in the survey consisted of 15 time samples.

### Application of Unsupervised Analysis

We first apply the unsupervised analysis. A competitive network of four neurons was used, and the seismic data were segregated into four categories. The raw seismic trace was employed as the input (Figure 1a), and the norm utilized was the l2 norm (Euclidian

distance). The result of the analysis is shown in Figure 9a, where the reservoir is classified into four distinct classes (note that the selection of colors in this map is arbitrary). The output map is quite plausible, and it corroborates our existing understanding of the reservoir heterogeneity in this region. Figure 9b shows a geologic model constructed prior to the development of the unsupervised analysis method in which the average porosity measured at several wells in this region at the top of the Unaizah Formation was manually interpolated between the wells. The two maps (Figures 9a and 9b) show considerable agreement, which is remarkable considering that no well information was used to generate the classification map of the unsupervised analysis.

Not only does the classification map corroborate the existing geologic model constructed by interpolating the well data, it also adds significant detail to the broad picture conveyed by the low frequency interpolated model, with considerably enhanced resolution and more well-defined boundaries between the various facies. Note that the interpolated model is strictly valid only at the well locations, since the interpolation has no objective source of information to guide it in the inter-well regions. On the other hand, the advantage of seismic facies mapping via unsupervised analysis is that the inter-well facies estimates are much more objective and consistent, since the method utilizes an extra piece of information (seismic data) that directly samples the reservoir between the well locations. At the same time, the result obtained via this method remains in agreement with the broad picture gained from the interpolated map.

It should be noted that the classes in Figure 9a are not labeled with the facies names. Comparing the maps in Figures 9a and 9b, one can easily recognize the high porosity classes in the output of the analysis. However, in general, the interpolated model is unknown since it relies on well information, and thus the class labeling must be done manually by inspecting the seismic characteristics of each class. Figure 10 shows the four class representatives generated by the analysis, which, as we mentioned previously, exemplify the typical behavior of each class. Inspecting these traces, one can readily draw the conclusion that Class 1 represents the high porosity facies, as its representative trace exhibits larger impedance contrast than the other traces, with Class 2 representing a high but slightly lower porosity facies (for the same reason), while the other two classes are the low porosity facies. As can be seen, this inference is strongly corroborated by the interpolated geologic model of Figure 9b, although it was made without reference to that model.

In fact, to take this a step further, we compute the individual similarity measure for each trace in the seismic survey relative to the representative of Class 1, which we have just identified as the highest porosity class. The resulting map is thus a measure of how similar each trace in the survey is to the typical seismic response of the high porosity class. This map is shown in Figure 11a, where, except for the scale, it looks strikingly close to the interpolated geologic model of Figure 9b—a testament to the accuracy of the unsupervised analysis technique and how effective it can be as indicative of the porosity heterogeneity of the reservoir.

The distinction and overall similarity confidence measures generated by the analysis

are shown in Figures 12a and 12b, respectively. Inspecting these maps indicates that confidence in the analysis is reasonably high in most of the field, except at the southeast corner, where the overall similarity drops to some extent. Thus, the classification result produced in this region should be considered with caution. Even though the similarity measure is low in this region, the more important distinction confidence measure is fairly high. It is interesting to note that this region is precisely where the analysis outcome (Figure 9a or Figure 11a) and the interpolated geologic model (Figure 9b) show the least resemblance.

Finally, Figure 11b shows the result of an unsupervised analysis similar to the one discussed thus far, except semblance was used instead of the $l_2$ norm. The maps generated using the two norms are mostly similar to each other, as expected (compare Figures 9a and 11b). However, since semblance is more sensitive to the trace shape characteristics, its map is in general more responsive to edge anomalies, and thus edge discontinuities in this map are sharper.

One of the advantages of the unsupervised method presented here over other techniques, such as those based on multi-variate statistics and back-propagation neural networks, is that it is not dependent on the well coverage of the field, and is thus not adversely affected by the difficulties encountered in other methods when that coverage is limited. Statistical and back-propagation network methods cannot extrapolate correctly to data not seen in their training set. For example, a back-propagation network was trained on a subset of the interval [0,5] to approximate the nonlinear function $T(z) = 1 - 2e^{-z}$ over that interval. The network was subsequently applied to a larger set inside the training range, and it was quite successful at predicting the values of the function within that range (Figure 13a). However, when the network is applied to values outside the training interval, it performs quite poorly. Figure 13b compares the original function to the output of the network in the interval $[-2,10]$. It is evident that the network was able to approximate the function only within the training range, and the applicability of the network is not valid outside that range.

In short, the back-propagation network is adept at interpolating between the training points, but is a poor extrapolator outside its training interval. This problem is not really limited to back-propagation networks, but is suffered also by other methods such as the statistical techniques we mentioned previously. What this means is that regions at the fringes of the field, which usually have little or no well coverage, are poorly modeled by those methods. On the other hand, no such problem is encountered by the unsupervised technique, as it does not depend on the use of any well data; it is thus not affected by the poor well coverage of the field. Note that even though, in the interpolated geologic model (Figure 9b) the southeast region shows high porosity, this region was never drilled, since the interpreters did not believe the model to be accurate because it is far away from well control. This decision is validated by the result of the unsupervised analysis, where that region was not assigned the same class as the known high porosity areas of the reservoir.

Furthermore, although supervised analysis utilizes well information, and is therefore

invariably affected to some degree by well coverage of the field, it is much more resistant to extrapolation difficulties than the statistical and back-propagation methods, since, as we mentioned previously, the well information is used only to label the resulting classes and is not involved directly in the classification process. In fact, the first stage of supervised analysis is identical to unsupervised analysis, and thus the two approaches have much in common when it comes to their resistance to the problems associated with extrapolating beyond the range of the training data set.

## Application of Supervised Analysis

We next apply supervised analysis. The well data utilized is comprised of averaged values of the porosity measured at the top of the Unaizah Formation. Twenty-six wells in the field were used in the analysis. These wells have a somewhat good coverage of the seismic survey area, especially at the center and northern regions. The southern region and the fringes of the area have rather sparse well coverage.

Only a single facies was assigned to each well, and the facies for all wells were divided into high porosity and low porosity facies according to the average porosity of the well. The division used the arbitrary threshold of 18% porosity. Wells whose average porosity is above that threshold were considered to have high porosity facies, whereas the other wells were deemed to have low porosity facies. Figure 14a.1 shows the average porosity in each well and threshold that was used to divide the wells into high and low porosity facies. The training data set thus consisted of 26 wells, each assigned a high or low porosity facies, and the seismic traces close to those wells. The actual average porosity values were not utilized in the analysis, as the objective here is to identify and map the facies, not estimate the point-values of the porosity (Saggaf *et al.*, 2000a,b).

To gauge the accuracy of the supervised analysis, systematic cross-validation tests of the data were carried out. In each of these tests, one well is removed from the training data set, the network is trained on the remaining wells and applied on a seismic trace at that hidden well location, and the output of the network is compared to the known facies present at that well. The test is then repeated for each well in the training data set. This cross-validation offers an excellent measure of the accuracy of the method and provides an objective assessment of its effectiveness.

Figure 14a.2 shows the result of the supervised analysis, where the predicted facies are compared to the actual facies known to be present in each well. Figures 14a.3c and 14a.4d show the distinction and overall similarity confidence measures generated by the method. Inspecting Figure 14a.2b, we see that supervised analysis correctly identified the facies present in 24 out of the 26 wells in the study (an accuracy of 92%). Although this is not a perfect outcome, it is quite good nonetheless. This is especially true since the confidence measures are quite low for the two wells whose facies were misidentified (Figures 14a.3 and 14a.4d). Thus, an erroneous decision to drill these wells would not have been made based on this analysis, as the low confidence measures clearly indicate that the data are inconclusive there and thus the validity of the facies identification at these two locations is suspect. The arbitrary division of the training data set into high

and low porosity facies based on a rather subjective and rigid choice of the threshold was no doubt also detrimental.

To obtain the spatial distribution of the reservoir facies, a final supervised network was built using all 26 wells in the training data set (i.e., utilizing the entire training data set). This network was then applied on the whole 3D seismic survey to produce a map representing the facies distribution of the reservoir. This final map, shown in Figure 14b, is quite similar to that produced by the unsupervised analysis (Figure 9a), and is therefore also quite similar to the interpolated geologic model we discussed previously (Figure 9b). However, in contrast to the outcome of the unsupervised analysis, the classes have been automatically assigned here to the target reservoir facies and labeled accordingly. The result of the supervised analysis is thus quite plausible and it corroborates the results of the unsupervised analysis (predictably) as well as the low frequency map produced by interpolating between the well locations, yet this method adds significant detail and enhanced resolution to the map. Note that we could have utilized more than two target facies in this supervised analysis, hence segregating the reservoir into high, intermediate, and low porosity facies, for example. However, we opted to keep the analysis simple in order to more clearly illustrate the method. Figure 14b shows that the two high porosity classes we identified earlier manually in Figure 9a have now been aggregated and identified as the high porosity facies, as they should, whereas the low porosity classes we identified manually in the unsupervised analysis have now been combined and identified as the low porosity facies. The supervised analysis thus represents a consistent and convenient method that relieves the interpreter from performing the manual class assignment and labeling step required by the unsupervised analysis (at the expense of requiring well information).

## CONCLUSIONS

We presented a method for quantitative mapping of the reservoir facies distribution by classifying and identifying the seismic facies. The method is based on competitive neural networks, and it can be applied in either an unsupervised or a supervised mode. The unsupervised analysis classifies the seismic data based entirely on the characteristics of the seismic response, without requiring the use of well information. It is thus of prime use for oil prospecting in new regions, where few or no wells have been drilled. The supervised analysis automatically assigns the output classes to the target reservoir facies and labels them accordingly. Therefore, the seismic facies are actually identified here rather than just classified, and the class labeling is accomplished intrinsically within the technique. This mode of analysis is therefore most useful in development fields, where the information gained at the wells can be conveniently extended via this method to the inter-well regions. It is especially valuable where 3D seismic surveys are available, as an aerial map of the reservoir limits may be extracted from the seismic survey using this technique. In addition to classifying and identifying the facies distribution, the method, in both of its modes, also ascribes to the output confidence measures that are

indicative of the validity of the result given the uncertainties and inconclusiveness of the data. These confidence measures represent an integral part of the solution, as they convey vital information about the range of applicability of the prediction output.

Tests conducted on synthetic and real data demonstrated that the method can be an effective means of mapping the reservoir heterogeneity. For synthetic data, the output of the method showed considerable agreement with the actual geologic model used to generate the seismic data. And for the real data application, the systematic cross-validation results indicate that the method can predict the facies distribution of the reservoir quite accurately. Furthermore, the resulting map corroborates our existing understanding of the reservoir and shows substantial similarity to the low frequency geological model constructed by interpolating the well information, while adding significant detail and enhanced resolution to that model. Additionally, the method is, compute-efficient, simple to apply, and does not suffer from input space distortion or nonmonotonous generalization. Not only that, it also does not depend on accurate selection of the training traces, and it is fairly resistive to the extrapolation shortcomings of the traditional statistical and back-propagation methods. Additionally, the norm utilized in the analysis can be selected to tailor the sensitivity of the method to the desired properties of the data, and several input/output architectures can be used depending on whether the raw seismic trace, multiple instantaneous attributes, or multiple interval attitudes are employed as the input. This method thus provides a convenient and robust approach to characterize the facies heterogeneity of the reservoir.

## ACKNOWLEDGMENTS

We would like to thank Saudi Aramco for supporting this research and for granting us permission for its publication. This work was also supported by the Borehole Acoustics and Logging/Reservoir Delineation Consortia at the Massachusetts Institute of Technology.

## REFERENCES


Al-Laboun, A.A, 1987, Unaizah Formation—A new Permian-Carboniferous unit in Saudi-Arabia, *AAPG Bull., 71,* 29-38.

Alsharhan, A.S. and Kendall, C.G., 1986, Precambrian to Jurassic rocks of Arabian Gulf and adjacent areas: Their facies, depositional setting, and hydrocarbon habitat, *AAPG Bull., 70,* 977-1002.

Chiu, S., 1994, Fuzzy model identification based on cluster estimation, *J. Intelligent & Fuzzy Systems, 2,* 267-278.

Dumay, J. and Fournier, F., 1988, Multivariate statistical analyses applied to seismic facies recognition, *Geophysics, 53,* 1151-1159.

Evans, D.S., Bahabri, B.H., and Al-Otaibi, A. M., 1997, Stratigraphic trap in the Permian Unaizah Formation, Central Saudi Arabia, *GeoArabia, 2,* 259-278.

Hagan, D.C., 1982, The applications of principal component analysis to seismic data sets, *Geoexploration, 20,* 93-111.

Kohonen, T., 1989, *Self-Organization and Associative Memory,* Springer-Verlag, Berlin.

Mathieu, P.G. and Rice, G.W., 1969, Multivariate analysis used in the detection of stratigraphic anomalies from seismic data, *Geophysics, 34,* 507-515.

Matlock, R.J., McGowen, R.S., and Asimakopoulos, G., 1985, Can seismic stratigraphy problems be solved using automated pattern analysis and recognition?, *Expanded Abstracts, 55th Ann. Internat. Mtg., Soc. Expl. Geophys.,* Session S17.7.

McGillivray, J.G. and Husseini, M.I., 1992, The Paleozoic petroleum geology of Central Arabia, *AAPG Bull., 76,* 1473-1490.

Saggaf, M.M. and Robinson, E.A., 2000, A unified framework for the deconvolution of traces of nonwhite reflectivity, *Geophysics,* in publication.

Saggaf, M.M. and Toksöz, M.N., 1999, An analysis of deconvolution: modeling reflectivity by fractionally integrated noise, *Geophysics, 64,* 1093-1107.

Saggaf, M.M., Toksöz, M. N., and Mustafa, H.M., 2000a, Estimation of reservoir properties from seismic data by smooth neural networks, this report, 2-1–2-36.

Saggaf, M.M., Toksöz, M. N., and Mustafa, H. M., 2000b, Application of smooth neural networks for inter-well estimation of porosity from seismic data, this report, 3-1–3-26.

Simaan, M.A., 1991, A knowledge-based computer system for segmentation of seismic sections based on texture, *Expanded Abstracts, 61st Ann. Internat. Mtg., Soc. Expl. Geophys.,* 289-292.

Yang, F.M. and Huang, K.Y., 1991, Multi-layer perception for the detection of seismic anomalies, *Expanded Abstracts, 61st Ann. Internat. Mtg., Soc. Expl. Geophys.,* 309-312.

# APPENDIX

We describe the implementation of the training of the unsupervised and supervised competitive networks and discuss the computation of their confidence measures.

## Unsupervised Analysis

This mode of analysis is implemented by a network that consists of a single competitive layer. This layer is represented by the matrix $\mathbf{C}$, where each row vector $C_i$ is a neuron that will ultimately be a cluster center describing one of the resulting classes. The size of the network is thus dictated by the desired number of classes. The network is initialized by simply setting each neuron to be in the middle of the interval spanned by the input. More sophisticated algorithms like subtractive clustering (Chiu, 1994) can also be used to initialize the network and suggest the optimal number of classes.

Each input vector is compared to the neuron vectors by computing the distance $d_i$ between the input vector and the $i$th neuron. This distance is usually computed as the $l_2$ norm, but other metrics can also be used, such as the $l_1$ norm, the cross-correlation between the neuron and input, or the semblance (the first two are examples of distance norms, whereas the last two are examples of correlation norms; for the latter two, the distance is taken as the negative of the norm in order to reverse the sense of the norm). The $l_1$ norm, for example, would penalize outliers more so than the $l_2$ norm. The transfer function $f$ is computed such that its output is one for the neuron that is closest to the input vector according to the distance $\mathbf{d}$, and zero for all other neurons:

$$f(C_i) = \begin{cases} 1 & \text{if } d_i = \min(\mathbf{d}) \\ 0 & \text{otherwise.} \end{cases} \tag{A-1}$$

$\mathbf{f}(\mathbf{C})$ would thus be a column vector with a single nonzero value that corresponds to the winning neuron. Such a neuron is said to have won the competition, and is rewarded by being moved closer to the input vector. This is done by updating the network matrix according to the rule:

$$\Delta C_j = r(\mathbf{x}^T - C_j) f(C_j), \tag{A-2}$$

$$C_j = C_j + \Delta C_j, \tag{A-3}$$

where $\mathbf{x}^T$ is the transpose of the input vector and $r$ is a prescribed learning rate, which is set to 0.1 here.

In the above formulation, some neurons may start far away from any input vector and never win any competition. To make the competition more equitable, biases are introduced to penalize neurons that win quite often, in order to ensure that all neurons win some of the competitions and are thus assigned to some clusters. Such a bias can be done by adding a decaying running average of each neuron transfer function to the

distance between the neuron and input vector. Therefore, neurons that have recently won look farther away from the input vector than they would have without the bias, and other neurons are thus given a better chance to win the competition.

Each neuron thus migrates progressively closer to a group of input vectors, and after some iterations, the network stabilizes, with each neuron at the center of the cluster it represents. To perform the prediction, the input vector is compared to each neuron, and it is assigned to the class represented by the neuron that is closest to that input vector.

## Supervised Analysis

This mode of analysis is implemented by a network that consists of a competitive layer followed by a linear layer. The transfer function of the entire network becomes $\mathbf{Lf(C)}$, where $\mathbf{L}$ is the matrix representing the linear layer. This layer maps the subclasses produced by the competitive layer into the final target classes. For example, if subclasses 1 and 3 belong to target class 1, and subclasses 2, 4, and 5 belong to target class 2, then $\mathbf{L}$ would be:

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}. \tag{A-4}$$

Each neuron is compared to the input vector as before. However, in this case, the winning neuron is moved closer to the input vector only if the subclass defined by that neuron belongs to the target class of the input vector. Otherwise, the neuron is moved away from the input vector. If the input vector belongs to target class $i$, then the target vector $\mathbf{t}_l$ would be a column vector that has a nonzero value in the $i$th row. This is transformed by:

$$\mathbf{t}_c = \mathbf{L}^T \mathbf{t}_l \tag{A-5}$$

to yield the competitive vector $\mathbf{t}_c$ that has nonzero values in the positions for all neurons that belong to the target class $i$. This can be used to determine whether a particular neuron belongs to that target class.

The learning rule thus becomes:

$$\Delta C_j = \begin{cases} r(\mathbf{x}^T - C_j) f(C_j) & \text{if } \mathbf{t}_c(j) \neq 0 \\ -r(\mathbf{x}^T - C_j) f(C_j) & \text{if } \mathbf{t}_c(j) = 0 \end{cases}. \tag{A-6}$$

Therefore, competitive neurons move closer to the input vectors that belong to the target classes of those neurons, and away from those that belong to other target classes. After some iterations, the network stabilizes, with each neuron in the competitive layer at the center of a cluster, and each group of such neurons (subclasses) mapped to a certain target class. To perform the prediction, the output of the prediction of the competitive layer (which is performed like before) is multiplied by $\mathbf{L}$ to determine to which target class the input vector belongs.

## Confidence Measures

Let $\mathbf{x}$ be the input vector, $\mathbf{n}_1$ the winning neuron, and $\mathbf{n}_2$ the next closest neuron to that input vector. The distinction measure depends on the relative distances in the vector space from a particular point representing the input vector to the two nearest cluster centers. For distance norms, we define it as:

$$D = 1 - \frac{\|\mathbf{x} - \mathbf{n}_1\|_d}{\|\mathbf{x} - \mathbf{n}_2\|_d}, \tag{A-7}$$

where $\|\cdot\|_d$ denotes a distance norm, and the norm is taken to be the same one used throughout the analysis ($l_1$, $l_2 \ldots$, etc.). When the distances between the point and the two nearest cluster centers are the same, this measure vanishes. For correlation norms, this confidence measure conveys the same information but is computed slightly differently since correlation norms are directly, rather than inversely, related to the resemblance between the neurons and the input vectors. In this case, we define the distinction measure as:

$$D = 1 - \left[\frac{\|\mathbf{x}, \mathbf{n}_1\|_c + 1}{\|\mathbf{x}, \mathbf{n}_2\|_c + 1}\right], \tag{A-8}$$

where $\mathbf{x}, \mathbf{n}_1$, and $\mathbf{n}_2$ are as defined above, and $\|x, y\|_c$ denotes a correlation norm between $x$ and $y$.

The overall similarity measure depends on the distance in the vector space from a particular point to the nearest cluster center. For distance norms, we define it as:

$$S = 1 - \frac{\|\mathbf{x} - \mathbf{n}_1\|_d}{\|\mathbf{x}\|_d + \|\mathbf{n}_1\|_d}, \tag{A-9}$$

where all symbols are as defined above. When the distance between the point and the winning neuron vanishes, the measure equals unity. In this case, that point is most similar to the identified facies. The individual similarity measure is defined the same way, except that a measure is generated for each neuron, rather than just the winning one. For correlation norms, the overall similarity measure is almost the norm itself, since correlation norms are directly proportional to similarity. Thus, for distance norms, we define the overall similarity measure as:

$$S = \frac{1}{2}[\|\mathbf{x}, \mathbf{n}_1\|_c + 1], \tag{A-10}$$

where all symbols are as defined above. The individual similarity measure is defined in a similar manner, except that a measure is generated for each neuron, rather than just the winning one.

Note that all the confidence measures above are defined in such a way that their value is always in the range 0 to 1, since for distance norms, $\|y\|_d \geq 0$ for all $y$; and for correlation norms, $-1 \leq \|x, y\|_c \leq 1$ for all $x$ and $y$.

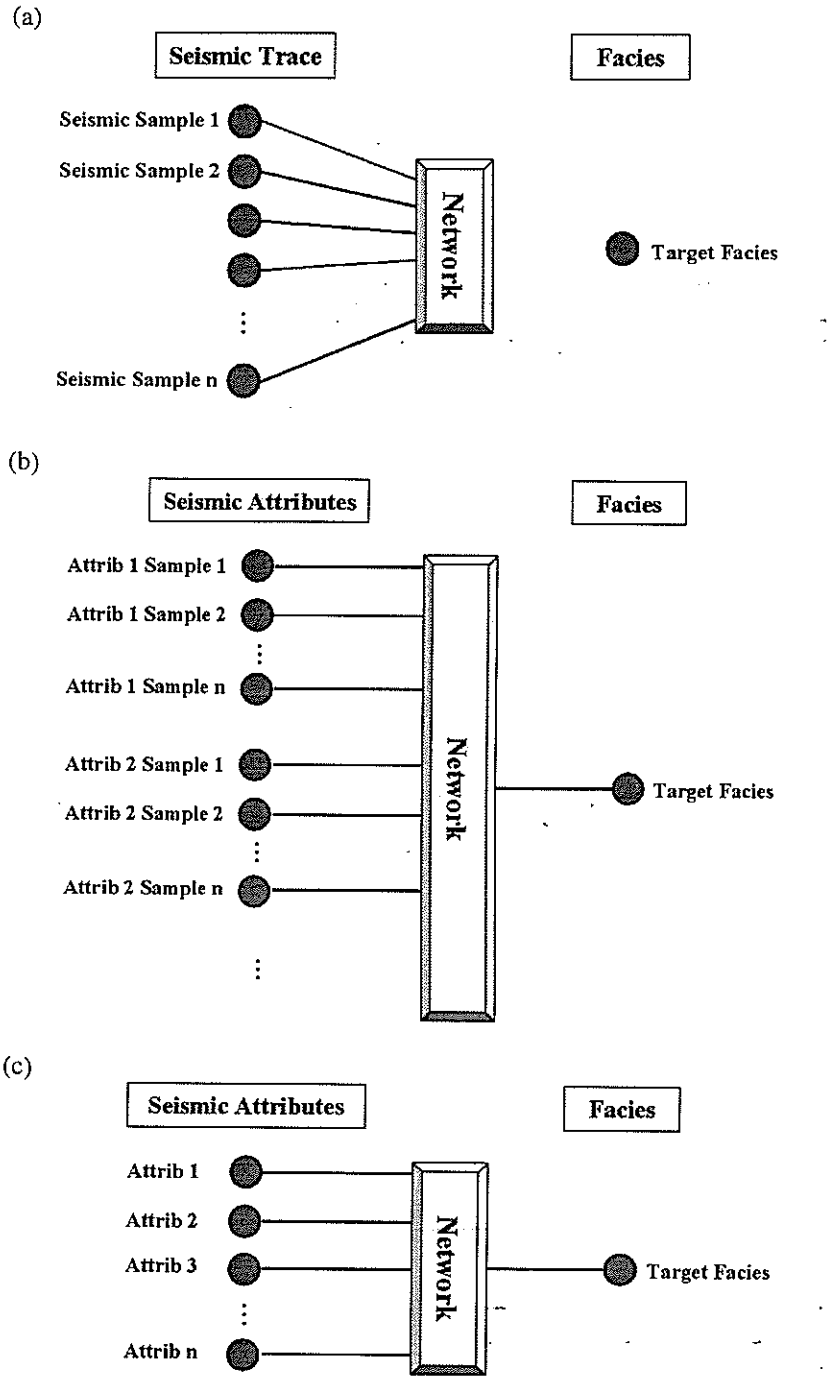


Figure 1: Schematic of the input/output architecture of the supervised and unsupervised networks when the input consists of: (a) the entire raw seismic trace; (b) multiple instantaneous seismic attributes; and (c) multiple interval seismic attributes.

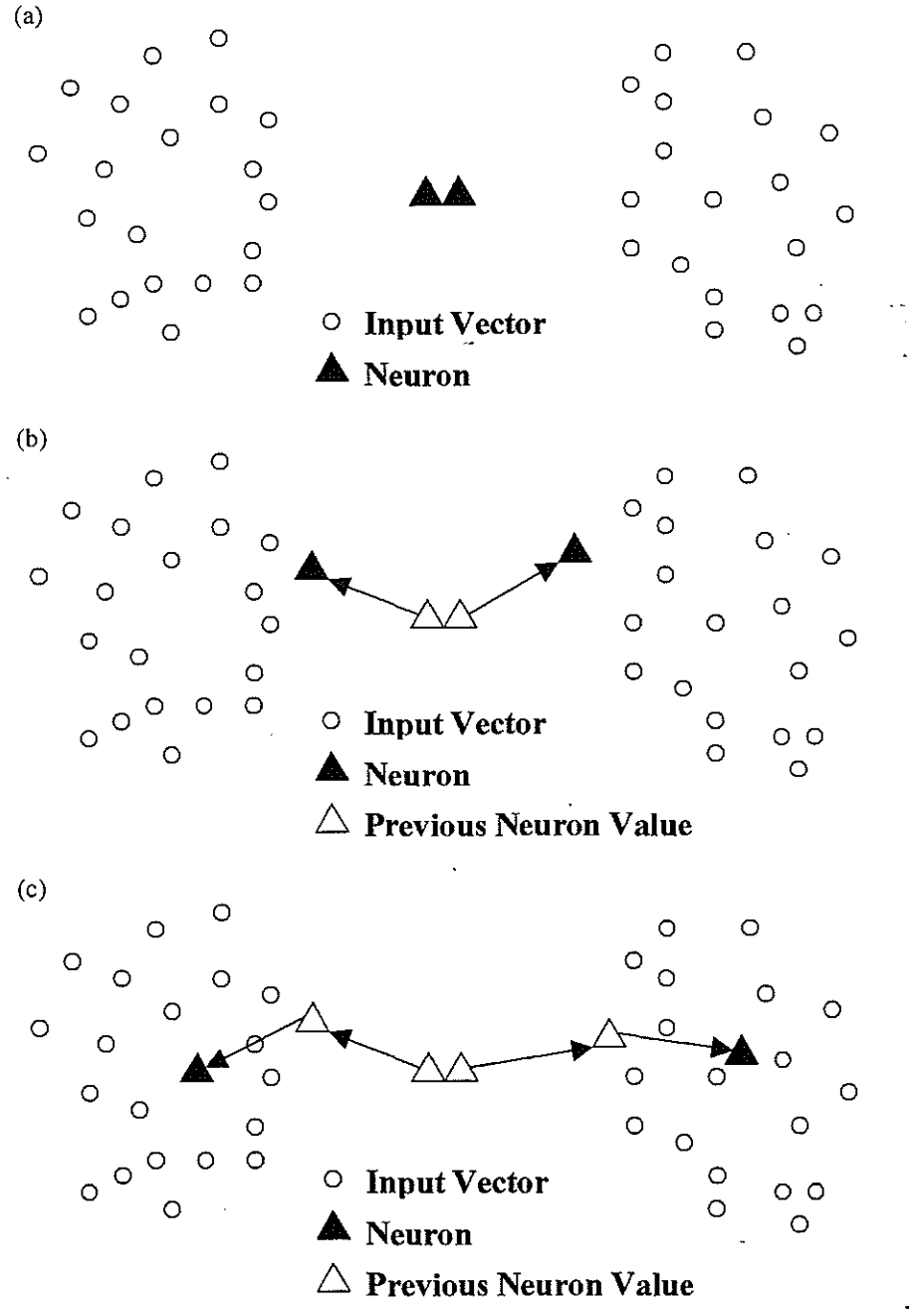


Figure 2: Schematic depicting the training of the competitive network. The neurons start at the same location (a), then each one migrates gradually towards a group of input vectors (b), until eventually it stabilizes at the center of a cluster of input vectors to which it is most similar (c).

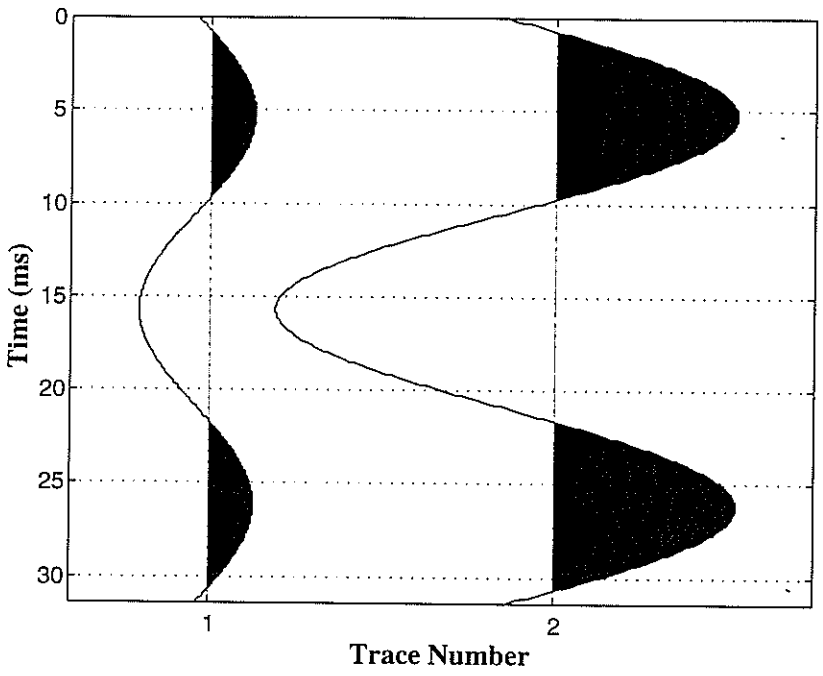


Figure 3: Example illustrating that correlation norms are insensitive to the absolute amplitudes of the seismic data and depend only on the shape of the trace. The two traces shown have the same shape but quite different amplitudes. Thus, their cross-correlation and semblance norms are unity, indicating that the two traces are similar, but their distance norms ($l_1$and $l_2$) are nonzero, indicating that the traces are dissimilar.

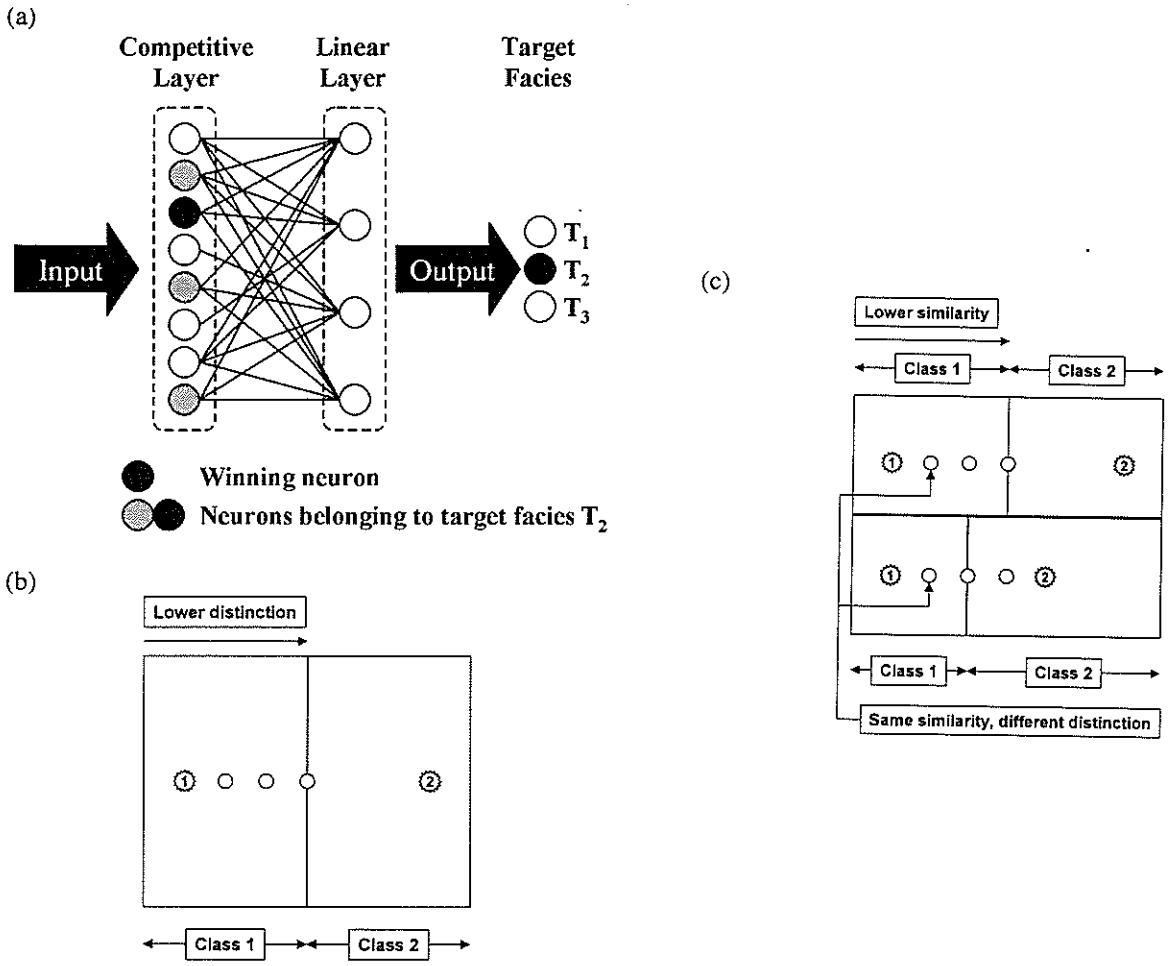


Figure 4: Schematics showing: (a) the basic structure of a supervised network, the linear layer assigns several of the competitive neurons to each of the target facies (all neurons are fully connected between the two layers, but for clarity, not all connections are shown); (b) vector space divided into two classes defined by the class representatives (denoted by 1 and 2), as the distances from the point to the two nearest cluster centers become equal, the distinction measure vanishes; and (c) two different classifications of the vector space, as the class representative of Class 2 draws closer to a point without affecting its classification, its overall similarity remains the same, but its distinction decreases.

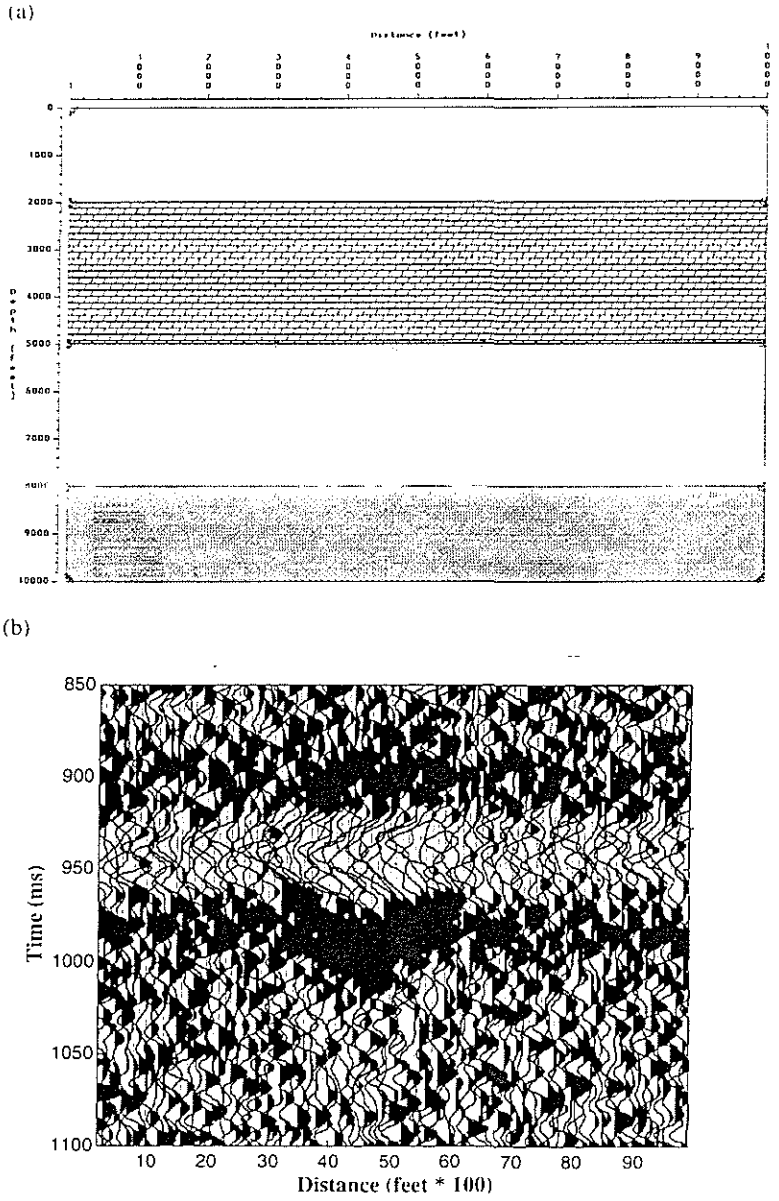


Figure 5: For the test conducted on synthetic data: (a) geologic model constructed to simulate a fluvial environment; and (b) seismic cross section constructed from the geologic model by generating synthetic seismic traces with zero-offset image rays and adding random Gaussian noise (this seismic section corresponds only to the relevant portion of the geologic model).

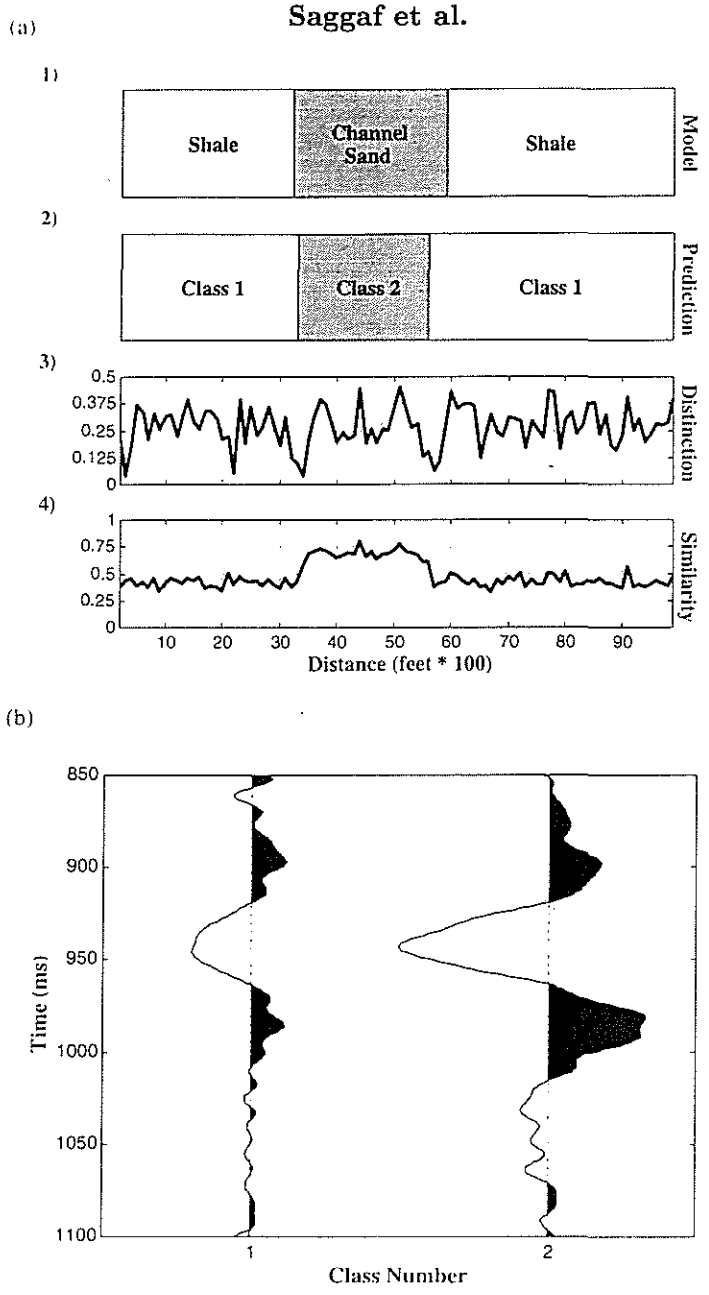


Figure 6: Application of unsupervised analysis in the synthetic data example: (a1) actual model used to generate the seismic data; (a2) facies classification predicted by the unsupervised network; (a3) distinction confidence measure; (a4) similarity confidence measure; and (b) the two class representatives, the seismic response of the trace belonging to Class 2 exhibits more impedance contrast, which indicates that Class 2 represents the channel area. Note that the predicted classes in (a2) are not labeled with the facies names. Also note how the distinction measure drops sharply at the edges of the channel, indicating that the data is inconclusive there.

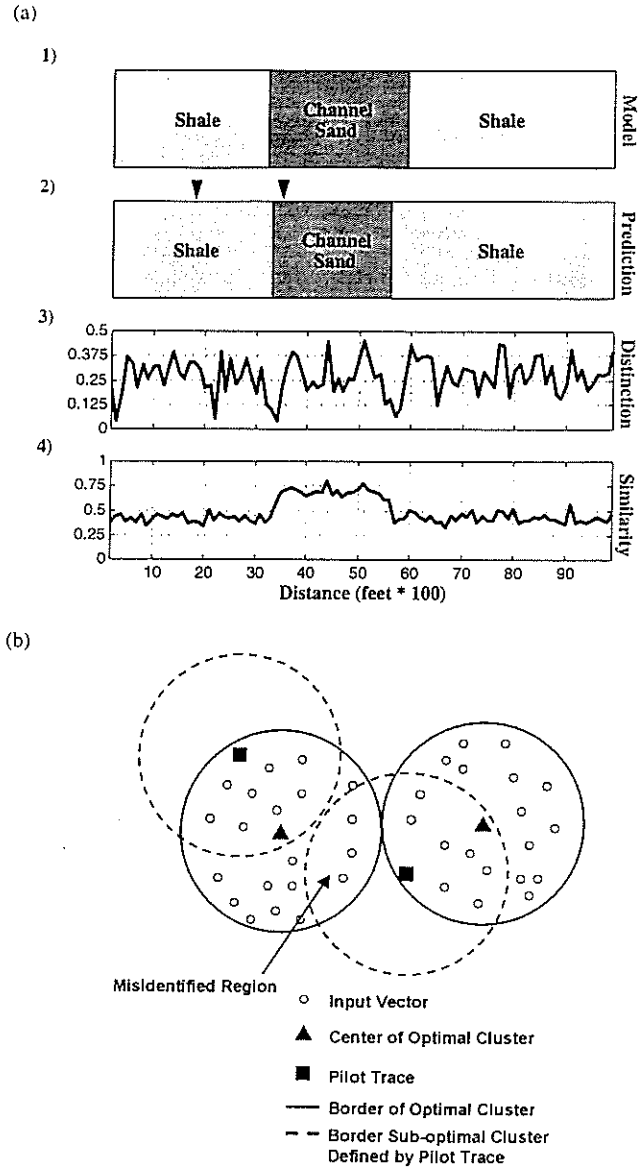


Figure 7: Application of supervised analysis in the synthetic data example: (a1) actual model used to generate the seismic data; (a2) facies identified by the supervised network; (a3) distinction confidence measure; and (a4) similarity confidence measure. The locations of the two training traces are indicated by the arrows. Note that the predicted classes in (a2) are labeled with the facies names. (b) Schematic depicting an example where the optimal cluster centers of the data do not fully overlap the classes defined by the pilot traces, since those traces, although at the fringes of the optimal clusters, define the direct classes to be centered around them. The region indicated is thus erroneously assigned to Class 1, although it optimally belongs to Class 2.

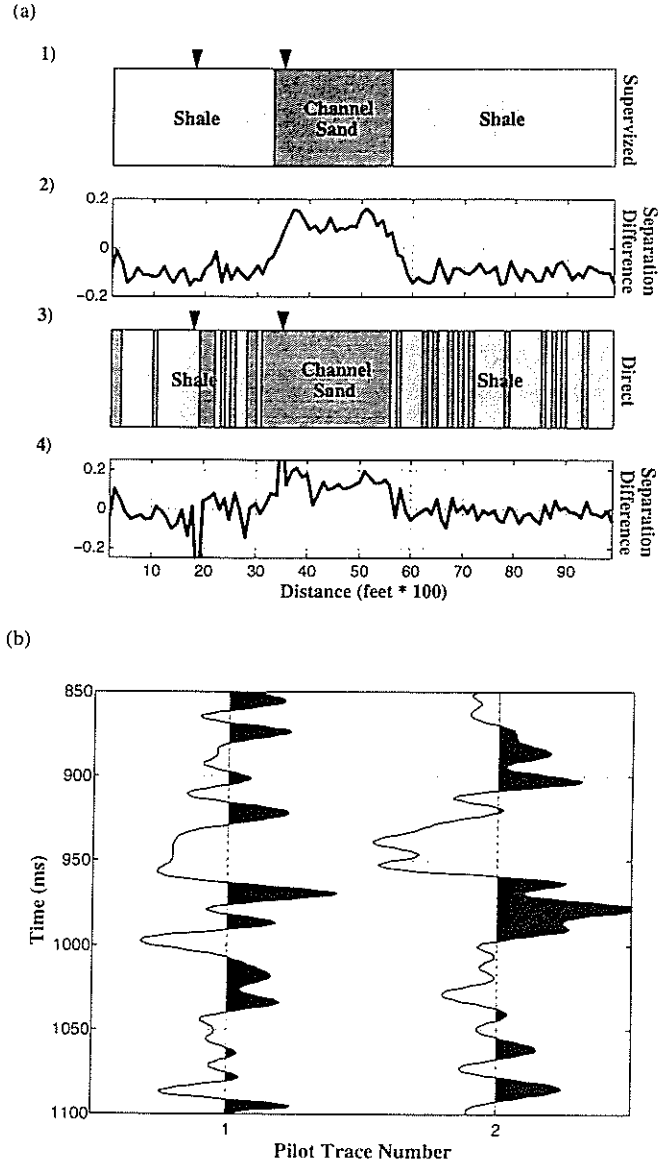


Figure 8: Application of supervised analysis and the direct method in the synthetic data example: (a1) facies identified by the supervised network; (a2) difference in distances from each trace in the survey to the two class representatives generated by the first layer of the network; (a3) facies identified by the direct method; (a4) difference in distances from each trace in the survey to the two pilot traces; and (b) the two pilot (training) traces used in the supervised analysis and direct inference method. The locations of the two training traces are indicated by the arrows. Note how the separation distance in (a4) fluctuates erratically around the baseline.

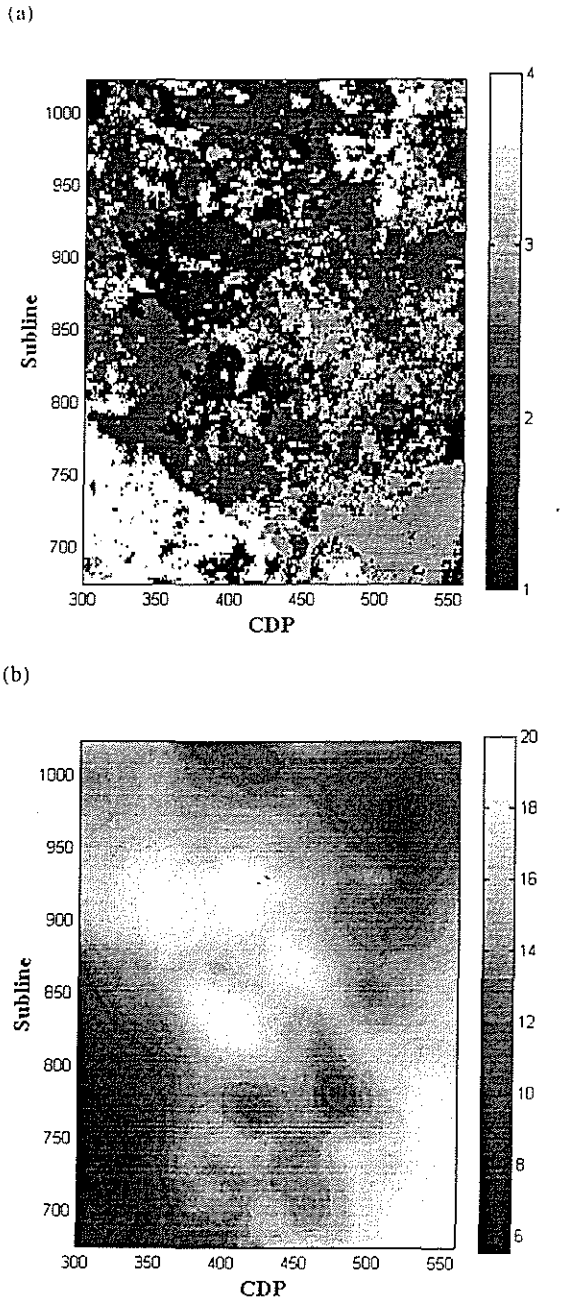


Figure 9: For the test conducted on field data: (a) facies classification predicted by the unsupervised network utilizing the l2 norm, the reservoir is segregated into four distinct classes representing the facies distribution of the reservoir (note that the selection of colors in this map is arbitrary); and (b) geologic model constructed by interpolating the average porosity measured at several wells in the region, the porosity scale is shown on the right.

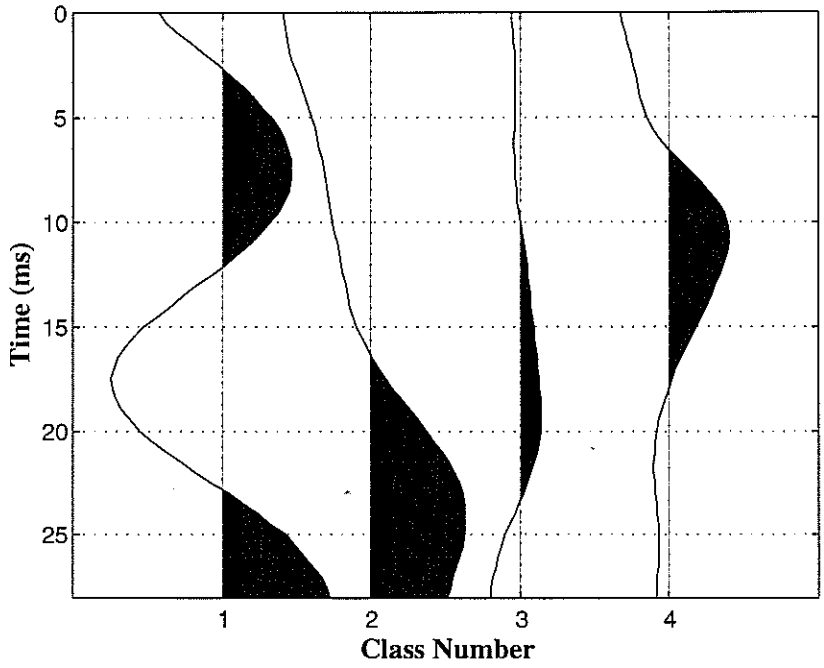


Figure 10: Four class representatives generated by the unsupervised analysis in the real data example. These traces can be considered the typical seismic responses of the output classes. The seismic responses of the first two traces exhibit more impedance contrast, which indicates that these two classes represent the high porosity facies.

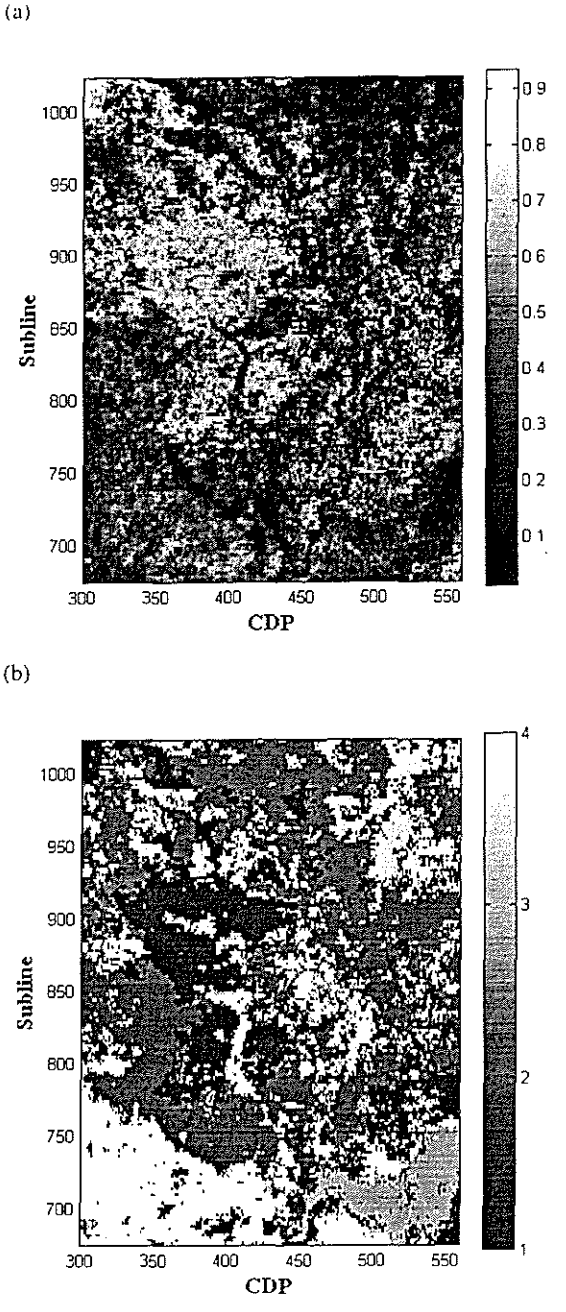


Figure 11: For the test conducted on field data: (a) individual similarity measure for each trace in the seismic survey relative to the representative of Class 1, which we identified as the high porosity class, the resulting map is thus a measure of how similar each trace in the survey is to the typical seismic response of the high porosity class (note how similar it looks to the interpolated geologic model); and (b) facies classification predicted by the unsupervised network utilizing the semblance norm.

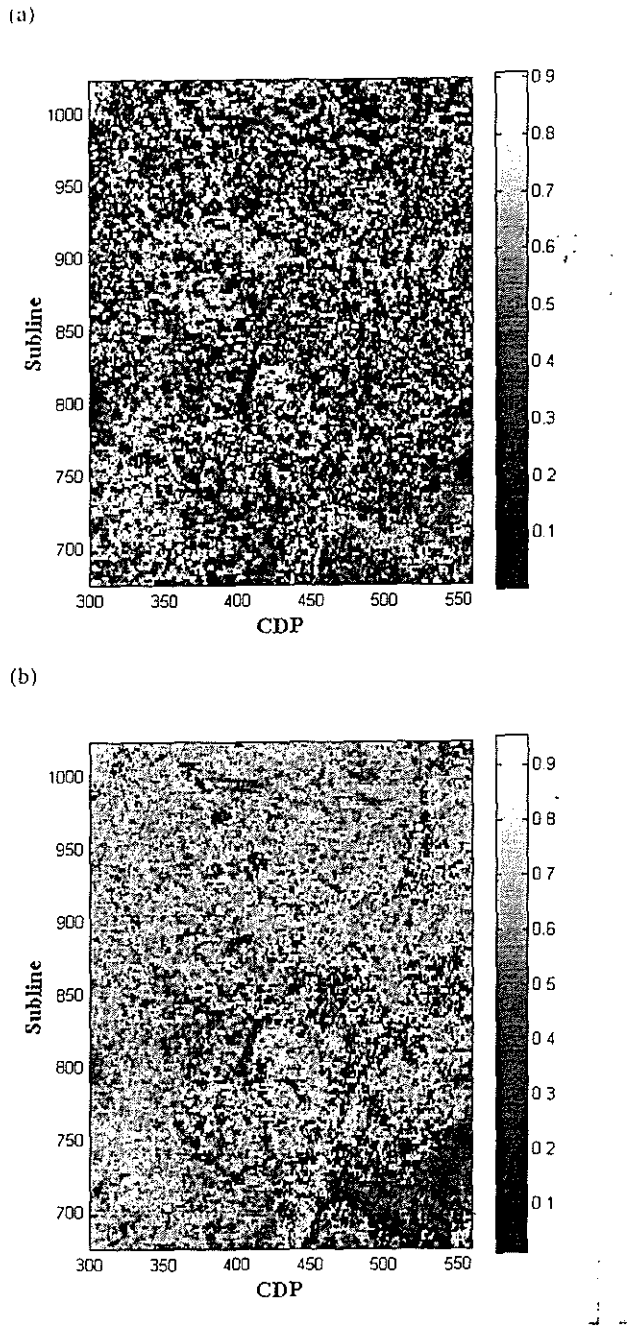


Figure 12: For the test conducted on field data: (a) the distinction, and (b) overall similarity confidence measures generated by the unsupervised network.

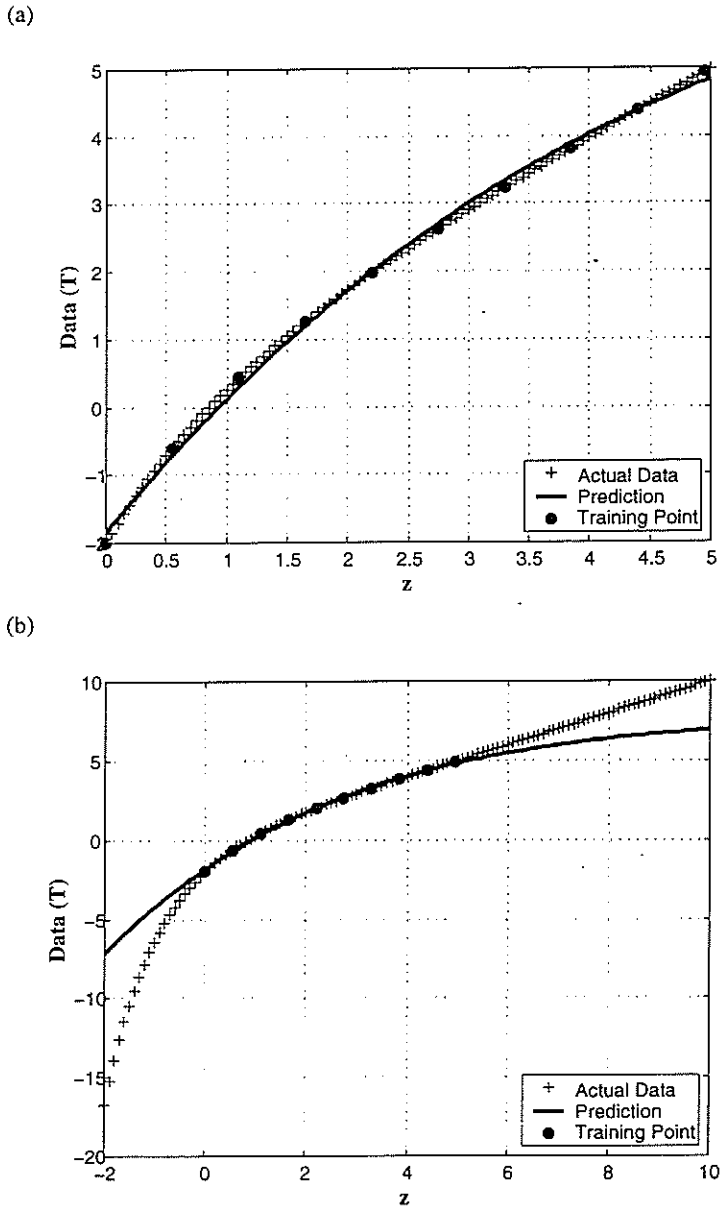


Figure 13: Illustration of the difficulties encountered by a back-propagation network in extrapolating beyond the training range: (a) the network was able to reproduce the data very well within the training range; but (b) the network prediction outside the training range was quite poor. The network is a good interpolator of the data but a poor extrapolator. The training data points are indicated on the plots.

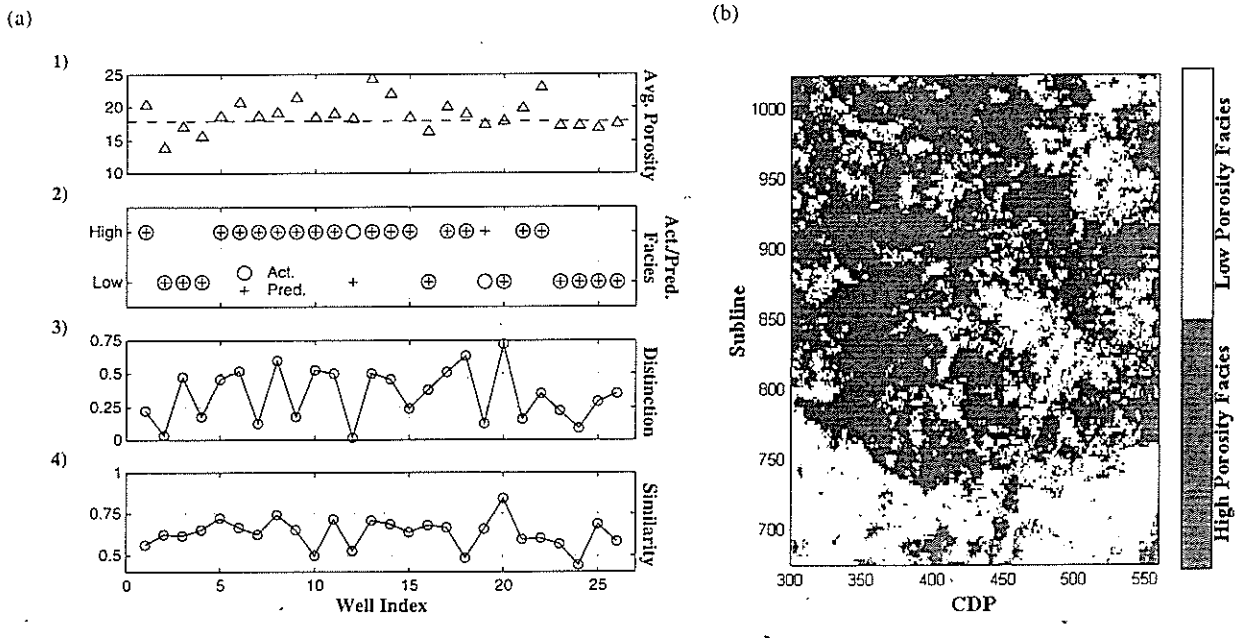


Figure 14: Results of systematic cross-validation tests conducted for the application of the supervised analysis in the field data example: (a1) actual average porosities measured at each well, the dashed line is the threshold utilized to divide the wells into low porosity and high porosity facies; (a2) facies identified by the supervised network (plusses) compared to the actual faces in each well (circles); (a3) distinction confidence measure; and (a4) similarity confidence measure for each well location. Note how the misidentified facies have very low confidence measures. Part (b) shows the low porosity and high porosity facies identified by the supervised network for the real data example. Note how the two high porosity classes encountered previously have now been aggregated and identified as the high porosity facies, as they should, whereas the low porosity classes encountered previously have now been combined and identified as the low porosity facies.