ESTIMATION OF LITHOLOGIES AND DEPOSITIONAL FACIES FROM WIRELINE LOGS

Muhammed M. Saggaf

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

Ed L. Nebrija

Reservoir Characterization Department
Saudi Arabian Oil Company
Dhahran, Saudi Arabia

ABSTRACT

We approach the problem of identifying facies from well logs though the use of neural networks that perform vector quantization of input data by competitive learning. The method can be used in either an unsupervised or supervised manner. Unsupervised analysis is used to segregate a well into distinct facies classes based on the log behavior. Supervised analysis is used to identify the facies types present in a certain well by making use of the facies identified from cores in a nearby well. The method is suitable for analyzing lithologies and depositional facies of horizontal wells, which are almost never cored, especially if core data is available for nearby vertical wells. Both types of modes are implemented and used for the automatic facies analysis of horizontal wells in Saudi Arabia. In addition to the identification of facies, the method is also able to calculate confidence measures for each analysis that is indicative of how well the analysis procedure can identify those facies given uncertainties in the data. Moreover, constraints derived from human experience and geologic principles can be applied to guide the inference process.

INTRODUCTION

Since horizontal wells are rarely cored, a method is needed to indirectly derive the facies distributions within the formations penetrated by those wells. There is a material correlation between the behavior of well logs and the lithologic and depositional facies of penetrated formations, since modern logs are sensitive to factors that vary with the makeup of those formations.

Figure 1 shows histograms of log values for each of the facies of well SVC, a vertical oil well in the marine carbonate environment of the Lower Cretaceous Shu'aiba formation in eastern Saudi Arabia. This formation is the result of deposition on a carbonate ramp along the irregular margins of an intrashelf basin (Aktas *et al.*, 1999).

The open algal platform facies (OAP) is typically thin (15–20 ft) and consists of wackestones and packstones characterized by the consistent presence of the calcareous alga, *lithocodium aggregatum*. The deep open platform facies (DOP) is also usually thin, and is composed almost entirely of planktonic "chalky" lime mudstones and wackestones deposited under low-energy platform conditions. The lithocodium-coral shelf facies (LICO) contains encrusting calcareous alga (*lithocodium aggregatum*) together with simple corals, and represents inner ramp deposits partially barred from the open marine basin by rudist shoals. The lagoonal facies (LAG) is relatively high in detrital mud content, and consistently contains skeletal fragments of the elevator rudist *agriopleura spp.* The back barrier facies (BB) consists of wackestones and packstones with abundant large skeletal fragments of *glossomyphorous* and *pachytraga* elevator rudists. The rudist barrier facies (RB) were deposited within the high-energy barrier banks and shoals of a ramp crest complex separating the lagoon and the open marine basin. They typically consist of cross-bedded, mud-starved packstones and grainstones that contain abundant, large reworked fragments of *offneria murgensis* rudists, arranged into upward-shoaling cycles.

The left panel of Figure 1 shows how gamma ray values are distributed for each facies. The facies exhibit different characteristic log behavior, though it could be difficult to differentiate them based on a single log type. It can be concluded, for example, that gamma rays in this case can be effective in differentiating the back barrier facies from the lagoonal facies, but not from the lithocodium-coral facies. The addition of more log types (adding extra dimensions), however, can remedy this problem. The right panel of Figure 1 shows histograms for the density log. Here, the back barrier facies can be distinguished from the lithocodium-coral facies. Adding other extra dimensions (more log types) helps differentiate the other facies.

Figure 2 shows a cross-plot of the log values of well SHA, a horizontal well in the same environment described above, for three logs: neutron porosity, density, and deep resistivity. The log values show distinct characteristics that can be utilized to classify the facies for this well based on log character. These characteristics are manifested as pronounced clusters on the cross-plot, each cluster corresponding to a particular facies or sub-facies. In general, more than three logs are used, and thus the entire problem

cannot be visualized in a cross-plot.

Manual interpretation of facies from well logs is a labor-intensive process that involves the expenditure of a considerable amount of time by an experienced well log analyst, even with the aid of graphical techniques like cross-plotting. The problem becomes even more difficult as the number of simultaneous logs to be analyzed is increased. Automatic methods that rely on hard-coded thresholds and implicit cross-relations are sensitive to inconclusive data and are often unreliable in unclean formations or atypical environments. 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 of the input data representation by projecting the input space to a new vector space of lesser dimension.

Wolff and Pelissier-Combescure (1982) used principal component analysis to cluster the log values into separate facies that can be regarded as indicative of lithology. A principal component analysis projects the original log space into a target space of lesser dimensions such that the distances between the projected points are closest to the distances in the original space. In other words, the emphasis is to minimize the distortion inflicted by the projection. Delfiner *et al.* (1987) and Busch *et al.* (1987) used discriminant factor analysis to correlate the log values to lithological facies. Discriminant factor analysis projects the original log space into a target space of lesser dimensions such that the projected cluster centers are as far apart as possible and so that the projected points of the same cluster are closest to each other. In other words, the emphasis is on the selection of the parameters that provide the most discrimination. Anderberg (1973) and Rao (1973) give a more complete treatment of these multidimensional statistical analysis methods.

Recently, methods relying on self-organizing maps (Baldwin *et al.*, 1990) and back-propagation feed-forward neural networks (Rogers *et al.*, 1992) were introduced for the estimation of lithologies from logs. For data of several dimensions (multiple logs), self-organizing maps based on fully connected hyper-spheres require a very large number of nodes and are thus too compute-intensive. Therefore, the data range had to be heavily restricted in order for the problem to be tractable with existing computer resources. Also, the lack of an unambiguous metric hindered the estimation of well-defined quantitative confidence measures. Although back-propagation feed-forward networks for the most part give satisfactory results, their convergence is often unreliable, they have a nonmonotonous generalization behavior (overfitting, large networks often perform worse than smaller ones in test data), have no smoothness constraint, require considerable parameter tweaking, and lack biological plausibility (since the neurons have to transmit the signal in two different directions forward and back through the network to back-propagate the error) (Hrycej, 1992).

In this paper, we approach the problem of identifying both lithologic and depositional facies from well logs through the use of neural networks that perform vector quantization of input data by competitive learning. These are uncomplicated one-

layer or two-layer networks that are small, compute-efficient, inherently well suited to classification and pattern identification, and avoid the difficulties associated with back-propagation feed-forward neural networks mentioned above. Not only can the method perform quantitative facies analysis, classification, and identification for uncored horizontal wells, it is also 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.

METHOD AND EXAMPLES

Facies analysis can be carried out in one of three modes: unsupervised, supervised, and semi-supervised. These three modes are discussed in the following sections.

Unsupervised Analysis

The objective of the unsupervised mode of analysis is to segregate the well into distinct facies classes based on the log behavior. This process emulates of what a petrophysicist would do by performing a careful examination of the log characteristics at various intervals in the well. Figure 2 shows a cross-plot of three of the six logs used to perform the analysis on well SHA.

This mode of analysis is implemented by a single-layer competitive neural network that takes the various well logs as input and classifies each depth interval into its corresponding facies category. This mode is often called feature-discovery or a "let-the-data-talk" scheme. It performs clustering, or quantization of the input space, but the resulting classes are unlabeled. 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. The neuron that most resembles the input vector 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 progressively migrates 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.

Prior to the analysis, the logs may be preprocessed by applying a running average or a low-pass filter to remove the high-frequency jitter in the data. Each log is then mapped to the interval [0,1] through a linear transform. This process is preferable to scaling, since it ensures equal contribution to the solution from each log type. After the mapping, individual log types can also be emphasized by scaling them differently. This optional step may be useful were there is more confidence in the quality of particular log types over others, and is a means of assigning fuzzy *a priori* importance to particular log types.

The bottom two panels of Figure 3 shows the result of the automatic unsupervised analysis on well SHA compared with that performed manually. The manual analysis was carried out by an experienced geologist and took substantial time and effort to produce. The two maps show considerable agreement. In fact, they are almost identical except for the facies transition zones, where the log behavior is inconclusive. This is evident in the top two panels of the figure, which display the analysis confidence measures. Confidence is relatively high throughout the well, except at the transition zones, where it decreases significantly.

Figure 4 shows a cross-plot similar to that of Figure 2, but each facies class identified by the analysis is colored differently. Also shown are the cluster centers derived by the technique for these facies classes. These class representatives can be regarded as the most typical log values for each facies. It can be seen that the method was able to zone-in and identify distinct facies categories. Note that only three dimensions are shown in the plot. All dimensions (log types) contribute to the solution, though, and the method is influenced by every one of them.

The confidence measures for the interval from 6700 to 7000 ft in Figure 3 (indicated by the horizontal line) and the proximity of the class representatives of classes 2 and 3 in Figure 4 (the middle cluster centers) suggest that these two classes may represent a single facies. The analysis is repeated, restricting the facies classes to 3. Figure 5 shows the result of that analysis, where it is evident that classification confidence has increased for that interval. The figure also shows the well trajectory on the top panel. The plausibility of the result of the analysis is evident here, as the facies classes correlate well with the different depths penetrated by the well, suggesting that the well most likely penetrates almost flat, or very gently dipping, beds. One such interpretation of the results is shown on the top panel of Figure 5.

Supervised Analysis

The objective of supervised analysis is to identify the facies types present in a certain well by making use of facies identified from cores in a nearby well. This mode, often called guided or directed classification, is implemented by a two-layer neural network. The first layer is a competitive network like the one mentioned previously that preclassifies the input into several distinct subclasses. The second layer is a linear network that combines the subclasses and associates them with the final target classes. 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 class 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 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. The technique is described in more detail in the appendix.

As a test, the logs and core-derived facies of well SVB, a vertical well in the same

environment described above, were used to train the technique, which was then applied to the logs of well SVC, a few kilometers away. Figure 6 compares the result of the analysis with the actual facies derived from the cores of well SVC. (Facies 3, 4, and 7 are not present in either well, although they appear in some other wells in the reservoir.) The two show considerable agreement, signifying that the method can be an effective means of identifying the facies of the horizontal wells by utilizing their logs and the logs and core-derived facies descriptions of nearby cored vertical wells. In general, if the log types used during training are different than those of the application well, the solution is restricted to only the common log types of the two wells by discarding the extra dimensions from the trained network.

Semi-Supervised Analysis

Semi-supervised analysis is a hybrid between the two modes mentioned above. Instead of using the core facies and logs of specific wells, regional log response averages for each facies (calculated or estimated qualitatively from experience) may be used. Unsupervised analysis is run, and then the resulting classes are automatically labeled with the facies names by comparing their class representatives with the regional averages.

All modes of analysis can be applied to classify and identify not only depositional facies, but also lithologic facies. In fact, identification of lithologies is most often more facile, since the relation between lithology and log response is stronger. A simple test is shown on Figure 7, which displays the result of running the semi-supervised analysis on well UVD, a deep vertical gas well in eastern Saudi Arabia, to identify the dolomite and limestone intervals within the well. The bottom two panels of the figure compare the lithological facies identified automatically with the ones derived from the cuttings. The two show considerable agreement. The top panel shows the individual similarity measures for each lithology, which will be described below.

Confidence Measures

One of the most useful pieces of information generated by facies analysis is a measure that describes the degree of confidence in the analysis. Humans describe confidence qualitatively. For example, a petrophysicist looking at a suite of logs may say that he has a lot of confidence that a certain interval of the well consists of facies A, some confidence that some other interval is facies B, but very little confidence in a third interval where the logs are inconclusive. We want to convey the same type of information, only quantitatively. The method actually generates three types of confidence measures:

Distinction. This is a measure of how distinct the log behavior of the identified facies is from the rest of the log characteristics. A high distinction means that the logs at that location have behavior that is quite distinct and far away from the log behavior of all facies but the one to which it belongs. Such a point plots on the log cross-plot far away from all other cluster centers but its own. A low distinction means that the logs are

inconclusive. In fact, zero distinction indicates that the location could be ascribed to two different facies equally well. This measure depends on the relative distances in the vector space from a particular point representing the log values to the two nearest cluster centers. Figure 8 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.

Overall similarity. This is a measure of how similar the log behavior of the identified facies is to the class representative. A high overall similarity means that the logs at that location have behavior that is quite similar to the behavior typically exemplified by the identified facies. Such a point plots on the log cross-plot quite close to its class representative. This measure depends on the distance in the vector space from a particular log point to the nearest cluster center. Figure 9 is a schematic that shows that two points may have similar overall similarity, but different distinction. The top panels in Figure 3 show the distinction and overall similarity measures for that analysis, respectively.

Individual similarity. This is akin to the above, except a measure is generated for each facies type, not just the one identified. At each point in the well, the identified facies will always have the highest individual similarity measure. In fact, the facies was identified precisely because the log behavior at that location has the highest similarity to the class representative of that facies. Individual similarity measures can be regarded as indicators of the relative composition of the facies at each interval in the well. The top panel of Figure 7 shows the individual similarity measures for each lithology.

Noisy Data Sets

One of the advantages of neural networks over conventional artificial intelligence techniques is its ability to generalize and its graceful degradation (Hrycej, 1992). Figure 10 shows the result of the analysis performed on the same well as in Figure 3, except this time one of the logs was perturbed by adding random noise to an interval within it (the bottom panel shows the perturbed log). Comparing the result in Figure 10 with that in Figure 3, we see that the method was able to see through the noise and was not adversely affected by it. Note, however, that the analysis confidence for that zone has slightly decreased as expected (compare the top panels in Figures 3 and 10).

Constrained Inference

A petrophysicist never makes log interpretations in a vacuum; he incorporates all available sources of knowledge to make his inference. The same thing is accomplished in facies analysis by making constrained inferences. Constraints derived from human experience and geological principles can be formulated as rules that guide the analysis

process. These constraints are especially useful where different facies have similar log characteristics. Two types of constraints can be applied: static and dynamic constraints.

Static constraints. These are formulated and incorporated *a priori*. They have fixed depth locations, and the rules do not change within the inference process. For example, we already know in some fields that certain facies do not occur before given depths, while others do not occur after certain depths. These constraints can be applied to guide the automatic facies identification process. The information needed for these constraints can be readily obtained from the field-wide facies model derived by interpolating between the wells of known facies.

Dynamic constraints. These constraints do not refer to specific depth locations, but rather express general geologic notions. The application of these rules is dynamically updated within the inference process. For example, in most carbonate environments of the Arabian Peninsula, the lithocodium-coral facies does not occur deeper than the algal platform facies. This notion can be formulated into a dynamic constraint that precludes the method from ascribing the lithocodium-coral facies to a depth location after the algal facies has been identified. The supervised analysis of Figure 6 was repeated, except here the constraint, *lithocodium-coral facies do not occur deeper than the algal platform facies*, was added. The result is shown in Figure 11, where it can be seen that the application of the dynamic constraints resulted in a noticeable improvement in the accuracy of prediction in the deeper part of the well.

Both types of constraints are implemented by applying a mask to the individual similarity measures prior to identification. Thus, to constrain a particular facies to not be identified below a certain depth, its individual similarity measure is multiplied by zero below that depth. The constraints can also be implemented in a fuzzy sense. In other words, instead of precluding certain facies from being identified below a certain depth, their identification below that depth may be hindered but not completely prevented. This is done by multiplying the individual similarity measure by a nonzero number that represents the fuzzy possibility of that facies occurring below that depth (for example, 20%). Therefore, the likelihood of identifying that facies below that depth is decreased by that amount, but not completely precluded.

CONCLUSIONS

The proposed method of automatic log analysis based on competitive neural networks can be an effective means for identifying lithologies and depositional facies from wireline logs. Utilizing this method, facies classification yields closely similar results to those obtained manually by an experienced geologist. In addition, facies identification tracks the results derived from the cores quite well. The analysis may be applied in several modes: unsupervised, supervised, or semi-supervised, depending on the availability of data and knowledge of the field. Constraints derived from human experience and geologic principles can be applied to guide the inference process. The analysis may be assessed through various confidence measures: distinction, overall similarity, and individual similarity, and iterated depending on the results of these measures. The technique is compute-efficient, easy to apply, can handle data of large size and multiple log types, and does not suffer from input space distortion or nonmonotonous generalization. The method is also automatic, fast, and robust in the presence of noise.

ACKNOWLEDGMENTS

We would like to thank Saudi Arabian Oil Company (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

Aktas, G., G.W. Hughes, A. Abu-Bshait, and S. Al-Garni, 1999, Sequence stratigraphic framework of Shu'aiba formation, Shaybah field, Saudi Arabia: A basis for reservoir development of an Aptian carbonate ramp complex, AAPG 1999 Annual Convention, Abstract, San Antonio, Texas.

Anderberg, M.R., 1973, *Cluster Analysis for Applications,* Academic Press.

Baldwin, J.L., R.M. Bateman, and C.L. Wheatley, 1990, Application of a neural network to the problem of mineral identification from well logs, *The Log Analyst, 3,* 279–293.

Busch, J.M., W.G. Fortney, and L.N. Berry, 1987, Determination of lithology from well logs by statistical analysis, *SPE Formation Evaluation, 2,* 412–418.

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

Delfiner, P., O. Peyret, and O. Serra, 1987, Automatic determination of lithology from well logs, *SPE Formation Evaluation, 2,* 303–310.

Hrycej, T., 1992, *Modular Learning in Neural Networks: A Modularized Approach to Neural Network Classification*, John Wiley and Sons, New York.

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

Rao, C.R., 1973, *Linear Statistical Inference and its Application,* John Wiley and Sons, New York.

Rogers, S.J., J.H. Fang, C.L. Karr, and D.A. Stanley, 1992, Determination of lithology from well logs using a neural network, *AAPG Bulletin, 76,* 731–739.

Wolff, M. and J. Pelissier-Combescure, 1982, FACIOLOG: Automatic electrofacies determination, *SPLWA Annual Logging Symposium, Paper FF,* 6–9.

APPENDIX

We describe here the implementation of each mode of analysis.

Unsupervised Analysis

Unsupervised 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 ith 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 correlation between the neuron and input, or the semblance (in the latter two cases, the distance is taken as the negative of the norm in order to maximize 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 closet to the input vector according to the distance $\mathbf{d}$, and zero for all other neurons:

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

$\mathbf{f}(\mathbf{C})$ 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:

$$\begin{aligned} \Delta C_j &= r(\mathbf{x}^T - C_j) f(C_j) \\ C_j &= C_j + \Delta C_j \end{aligned}$$

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 further 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 progressively migrates 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

Supervised 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:

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Each neuron is compared to the input vector like above. 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 ith row. This is transformed by

$$\mathbf{t}_c = \mathbf{L}^T \mathbf{t}_l$$

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

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

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 which target class the input vector belongs.

Confidence Measures

Let $\mathbf{x}$ be the input vector, $\mathbf{n}_1$ be the winning neuron, and $\mathbf{n}_2$ be 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 log values to the two nearest cluster centers. We define it as:

$$D = 1 - \sqrt{\frac{\|\mathbf{x} - \mathbf{n}_1\|^2}{\|\mathbf{x} - \mathbf{n}_2\|^2}}$$

where the norm is taken to be the same one used throughout the analysis (the l_1, $l_2, \ldots$, etc.). When the distances between the point and the two nearest cluster centers are the same, this measure vanishes.

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

$$S = 1 - \sqrt{\frac{\|\mathbf{x} - \mathbf{n}_1\|^2}{\|\mathbf{x}\|^2 + \|\mathbf{n}_1\|^2}}.$$

When the distance between the point and the winning neuron vanishes, the measure equals unity. In this case, the point is most similar to the identified facies. The individual similarity measure is defined in the same way, except that a measure is generated for each neuron, rather than just the winning one.

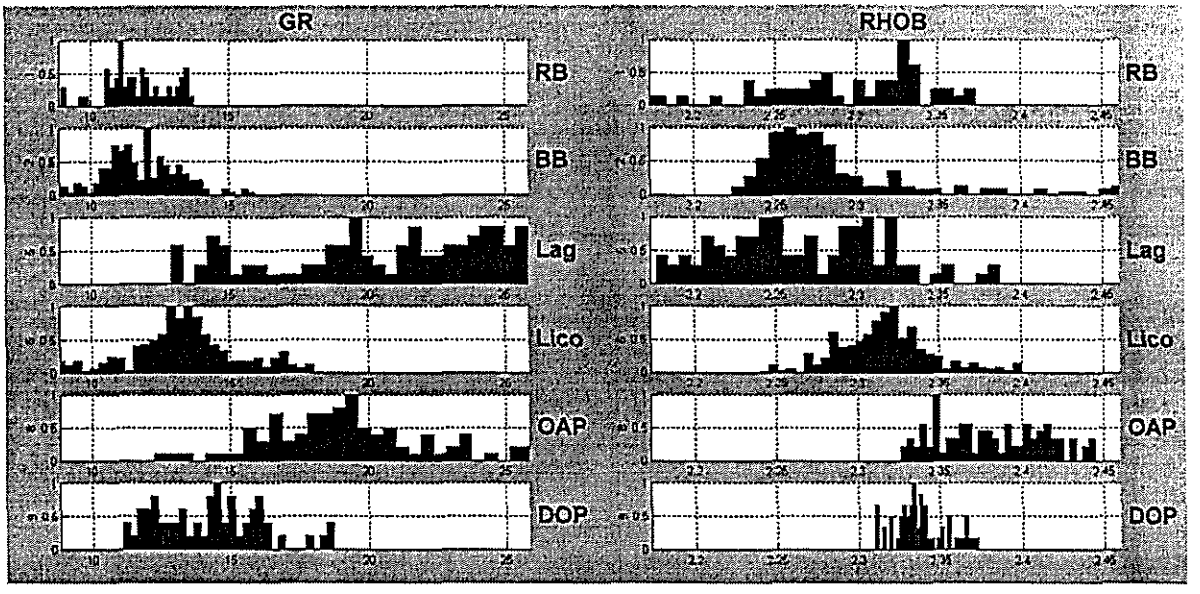

Figure 1: Histograms of log values for each facies of well SVC. The left panel shows how gamma ray values are distributed for each facies. Gamma rays in this case can be effective in differentiating the back barrier (BB) facies from the lagoonal (LAG) facies, but not from the lithocodium-coral (LICO) facies. The right panel shows the histograms for the density log. Here, the back barrier facies can be distinguished from the lithocodium-coral facies. Adding extra logs helps differentiate the other facies.

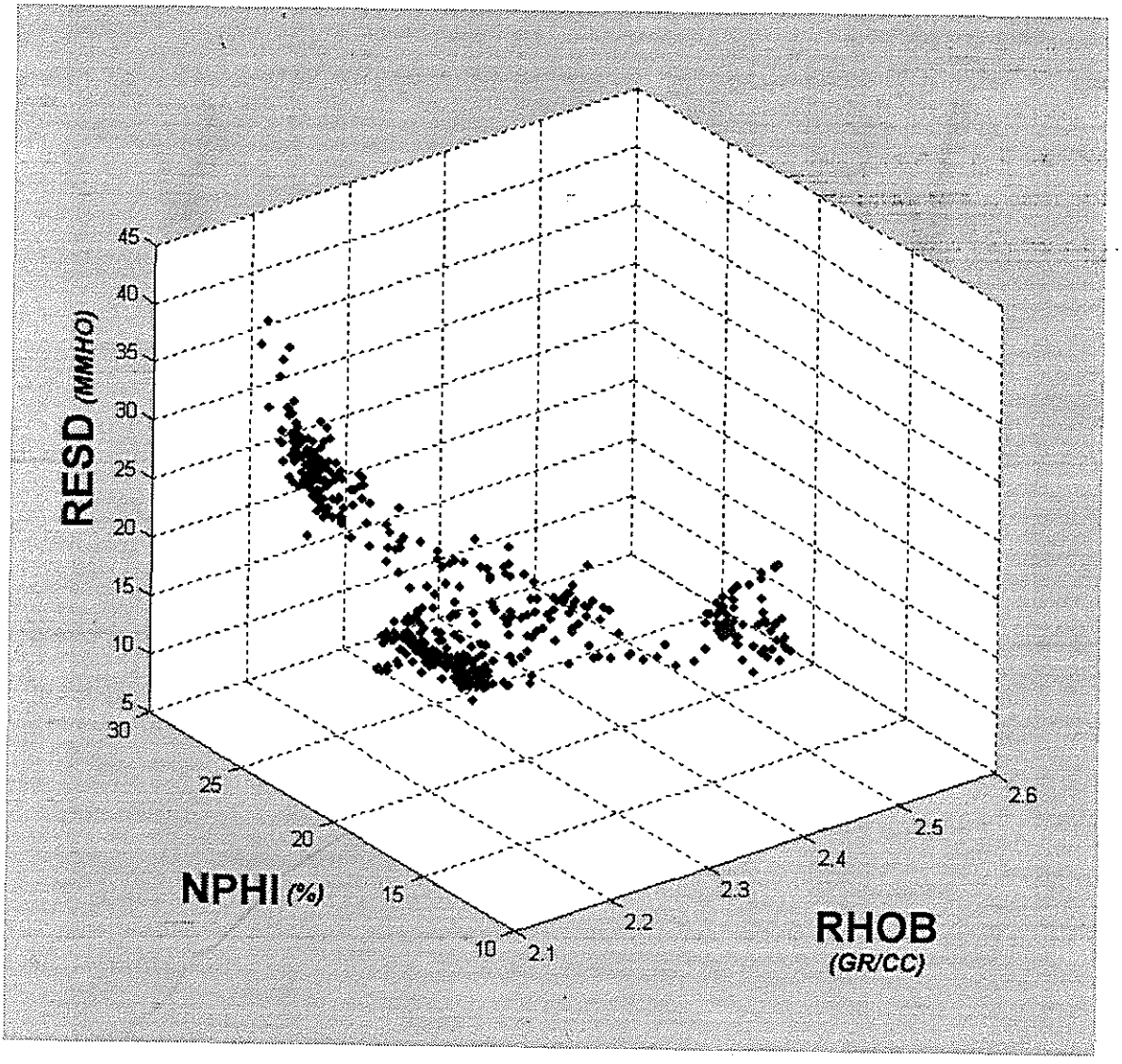

Figure 2: Cross-plot of the log values of well SHA for neutron porosity, density, and deep resistivity. The log values show distinct characteristics that can be utilized to classify the facies based on log character.

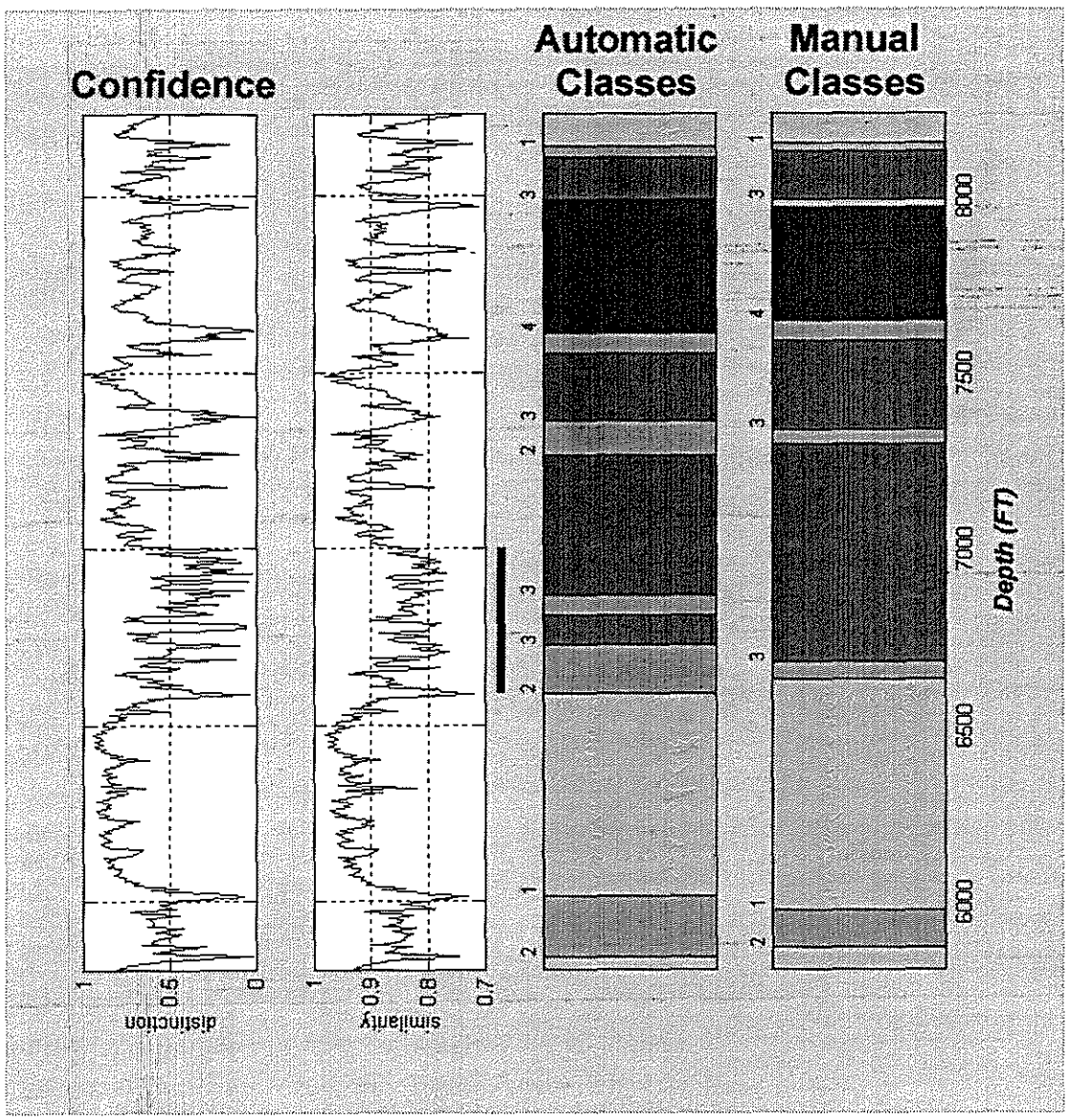

Figure 3: Unsupervised analysis applied to well SHA. The bottom two panels compare the results of the manual and automatic facies classification. The top panels show the classification confidence: distinction and overall similarity, respectively, from the top.

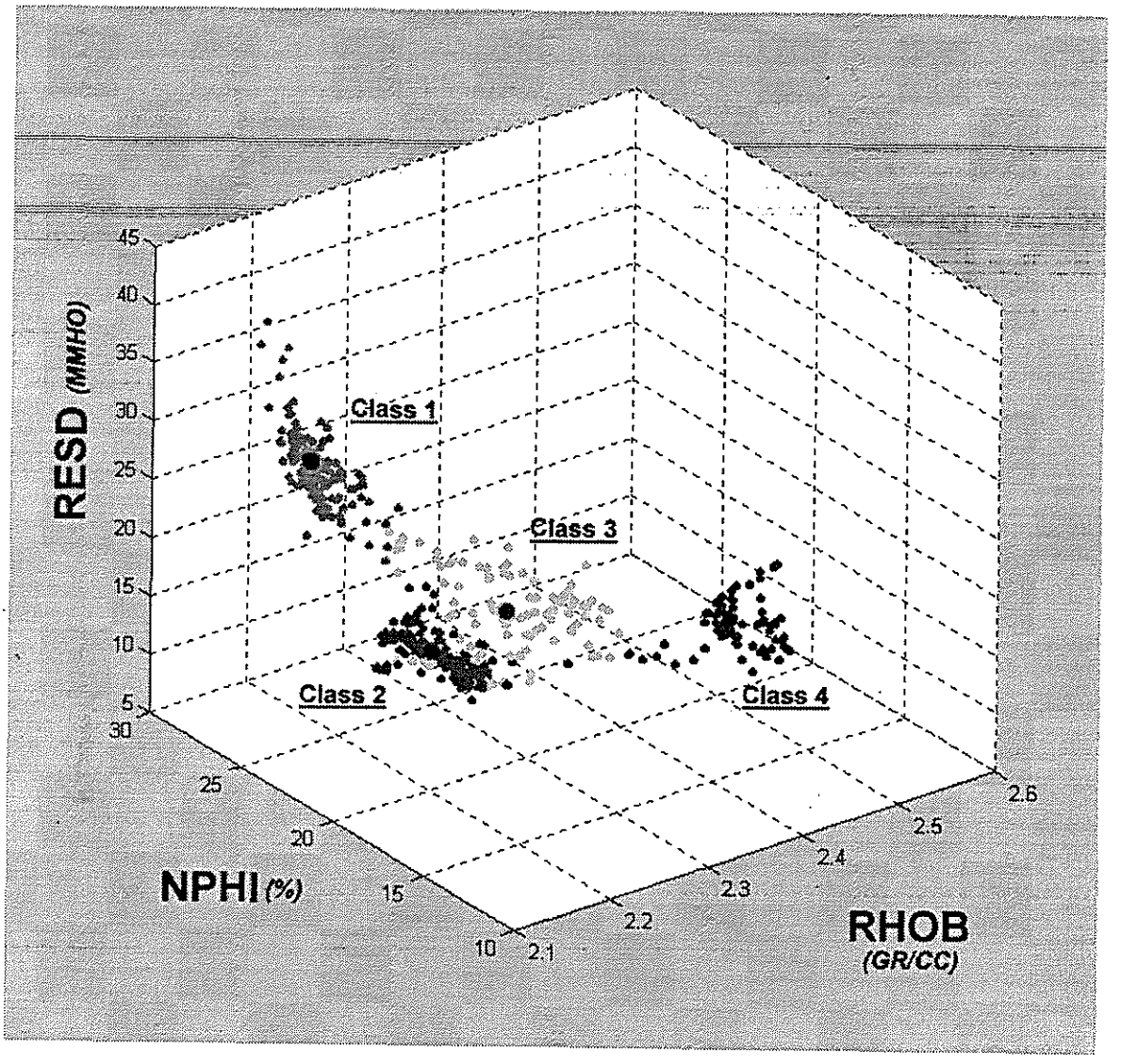

Figure 4: Similar to Figure 2, but the class representatives generated by the automatic analysis are superimposed on the plot (dark), and the points that belong to each class are colored differently.

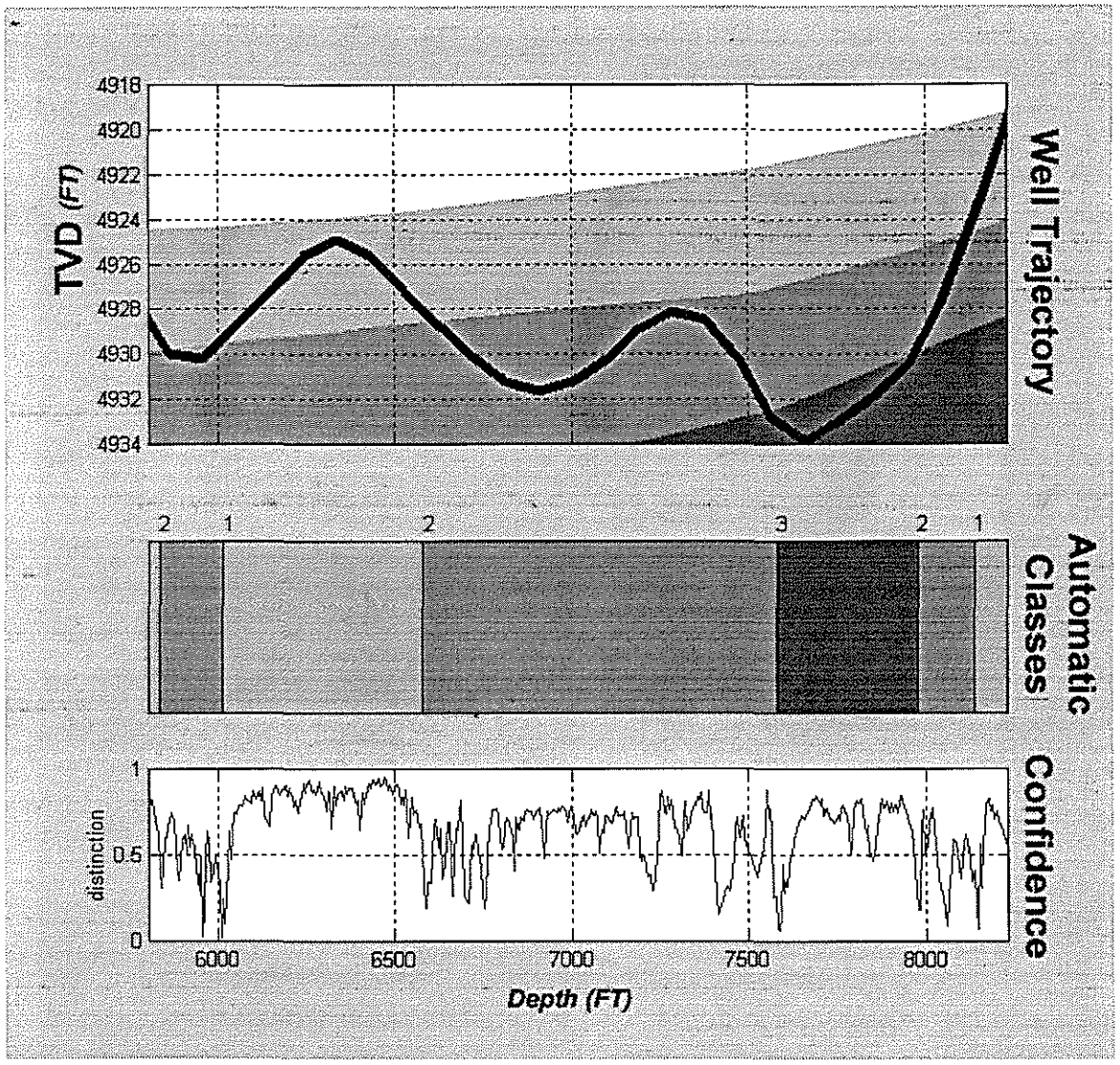

Figure 5: The analysis of Figure 3 was repeated, but the number of classes was restricted to three. The two most similar classes thus joined to represent a single facies. The top panel shows the well trajectory (TVD is the well's true vertical depth) and a possible formation interpretation of the resulting classification.

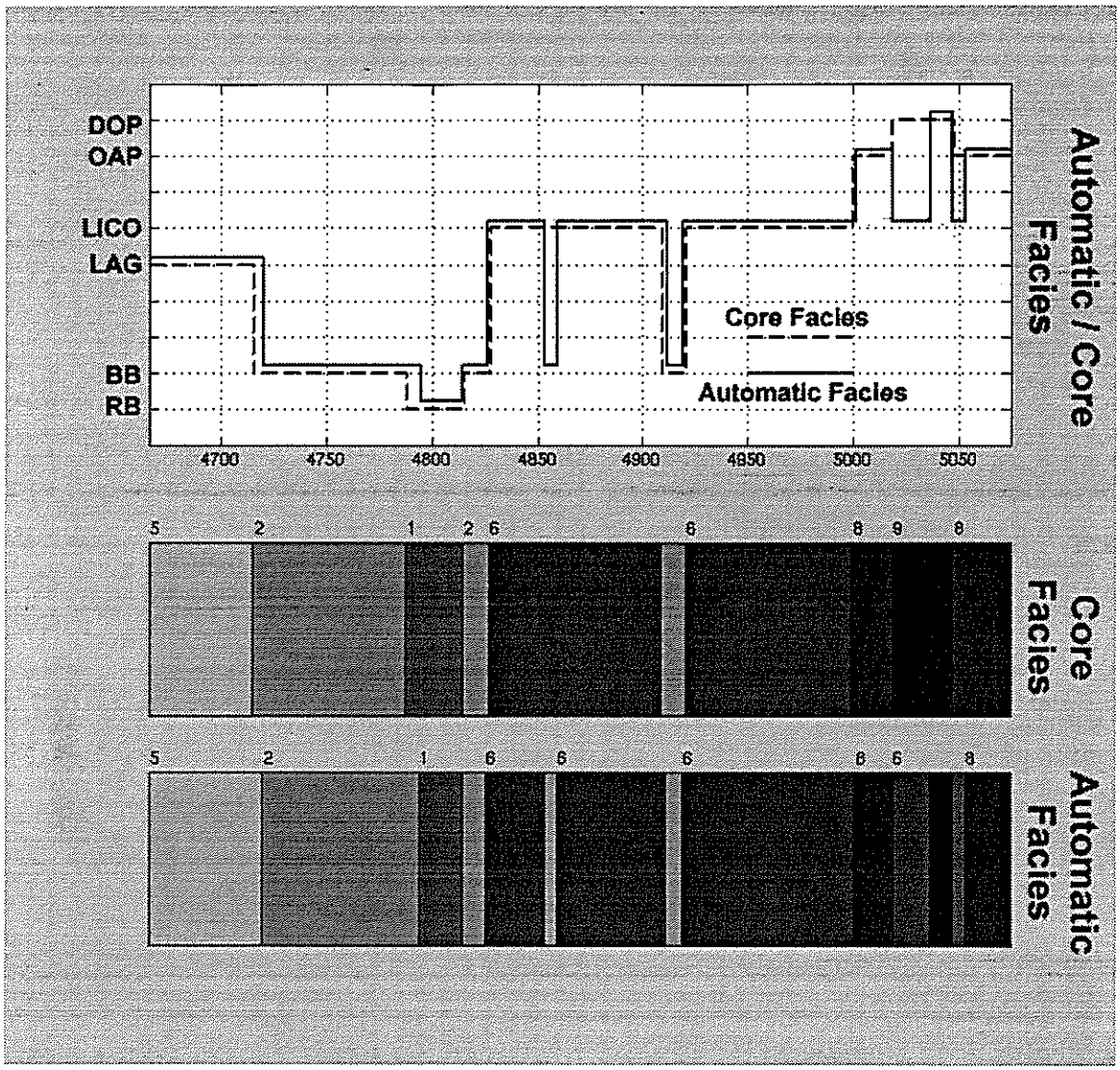

Figure 6: Supervised analysis applied to well SVC after training on the log and core data of well SVB. The bottom two panels compare the results of the automatic facies identification with the facies map derived from the cores of well SVC. The top panel shows the same comparison, but in a graphical plot (facies 3, 4, and 7 are not present in either well, although they appear in some other wells in the reservoir).

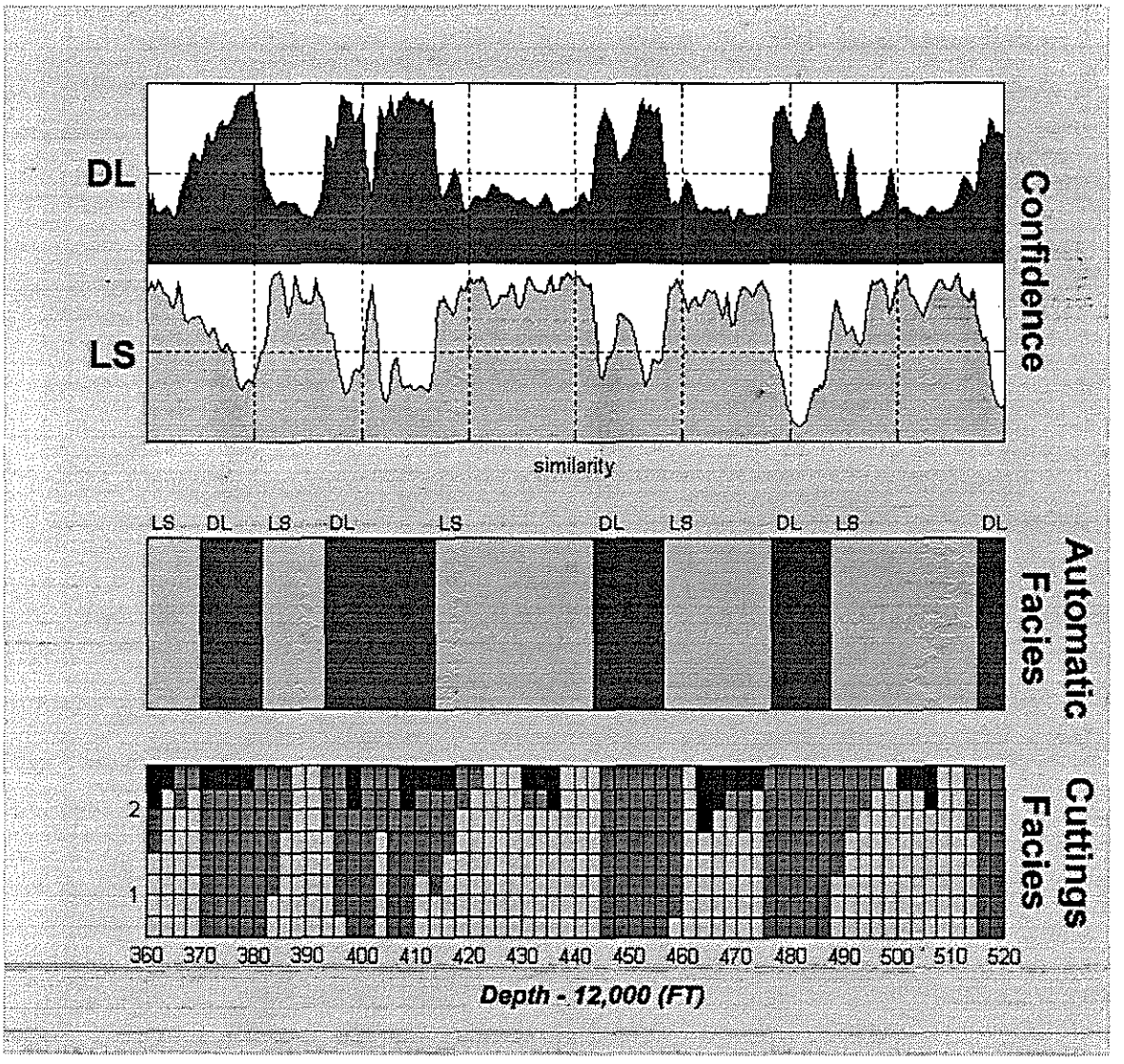

Figure 7: Semi-supervised analysis applied to well UVD for the identification of its lithological facies. Labeling was performed afterwards via regional averages of log behavior for each lithology. The bottom two panels compare the lithologies identified from the logs (dark = dolomite, light = limestone) with those derived from the cuttings (small anhydrite streaks, barely seen here, were not considered for the analysis). The top panels show the individual similarity confidence measure for each lithology.

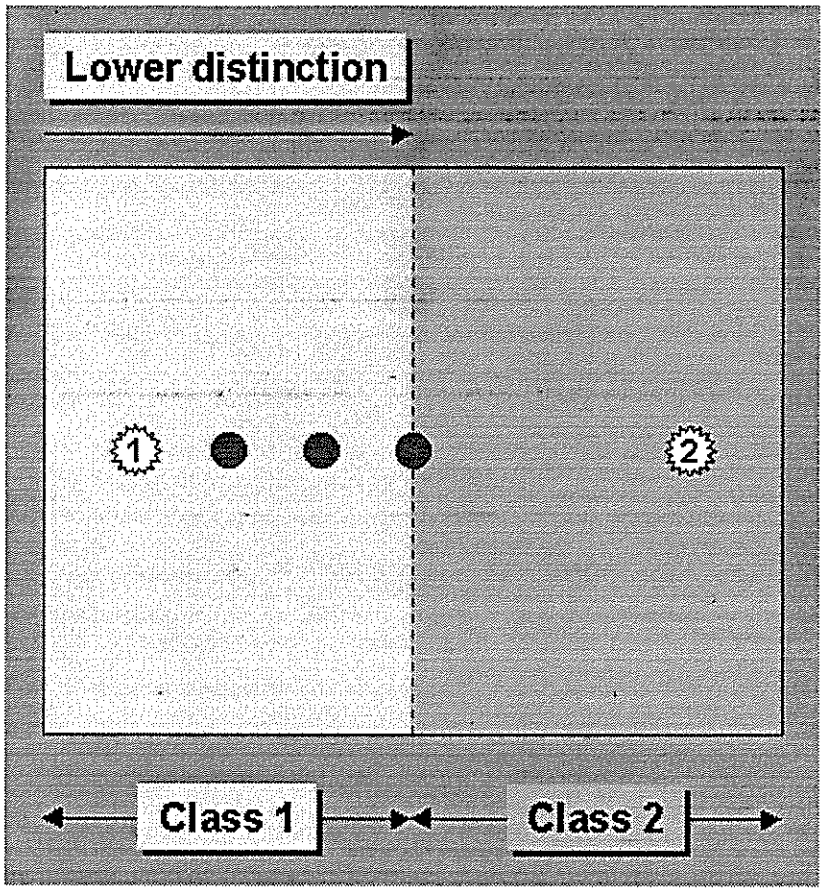

Figure 8: Schematic showing the vector space divided into two classes defined by the class representatives (denoted by 1 and 2). The direction of decreasing distinction is marked. As the distances from the point to the two nearest cluster centers become equal, the distinction measure vanishes.

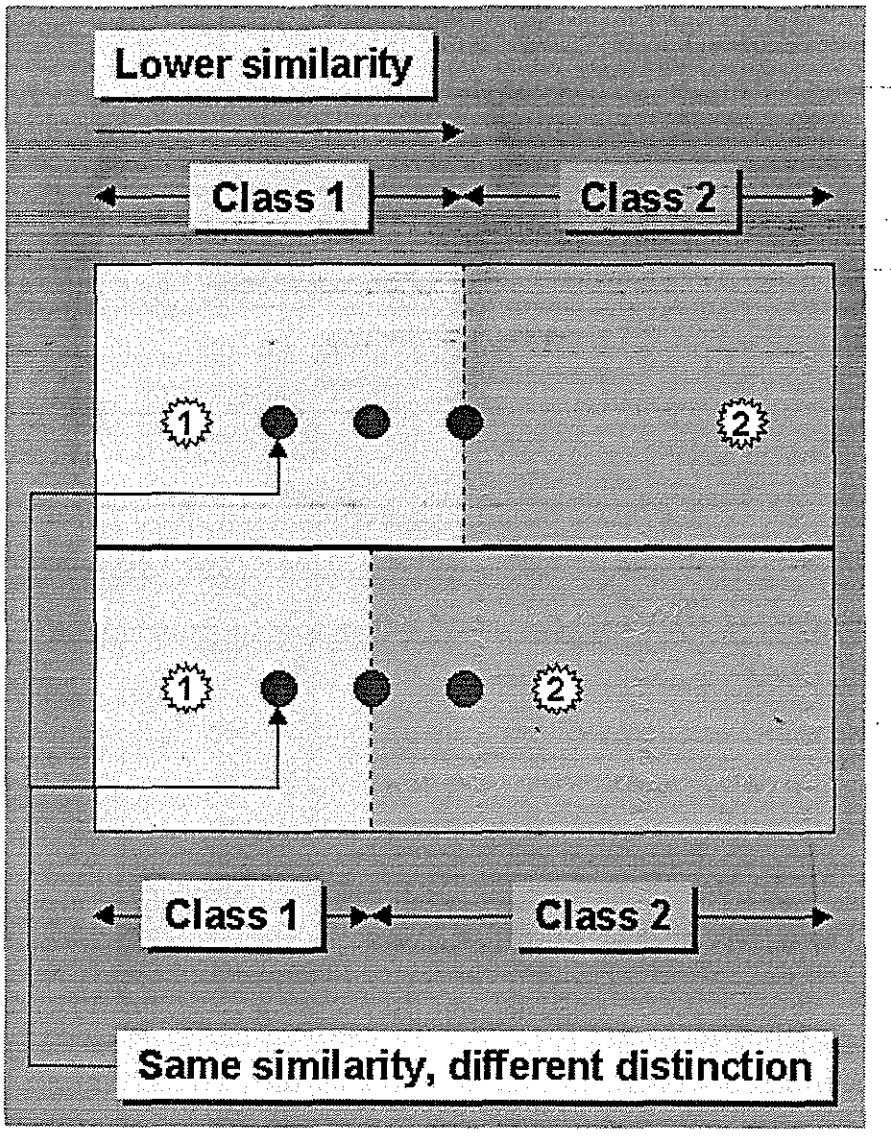

Figure 9: Schematic showing two different classifications of the vector space. The direction of decreasing overall similarity is marked. As the class representative of class 2 draws closer to a point without affecting its classification, its overall similarity remains the same, but its distinction decreases.

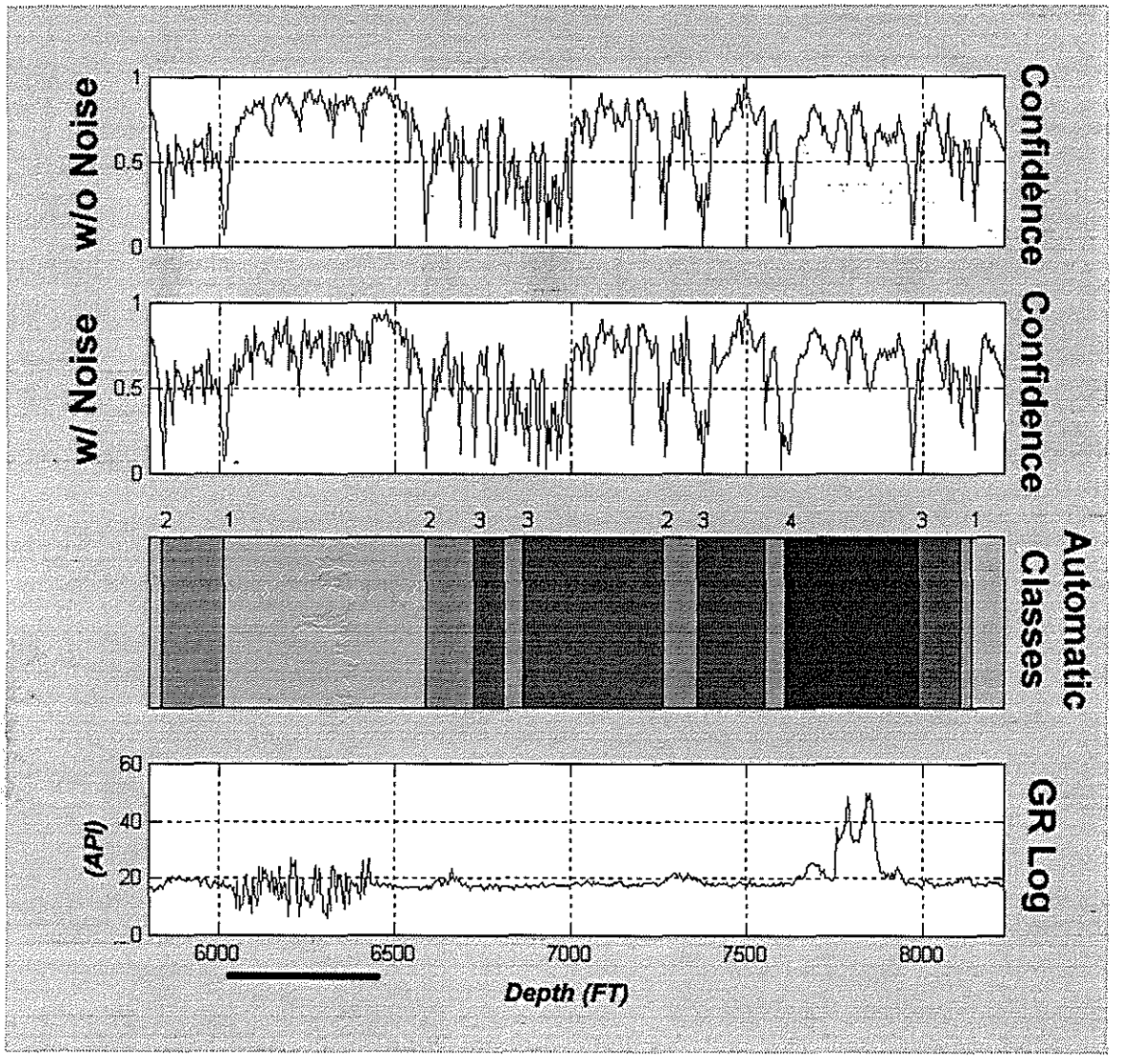

Figure 10: The analysis of Figure 3 was repeated, but one of the logs was perturbed by adding random noise to an interval within it (denoted by the line). Due to the graceful degradation of neural networks, the analysis was not adversely affected by the noise. However, confidence decreased slightly due to the noise, as expected.

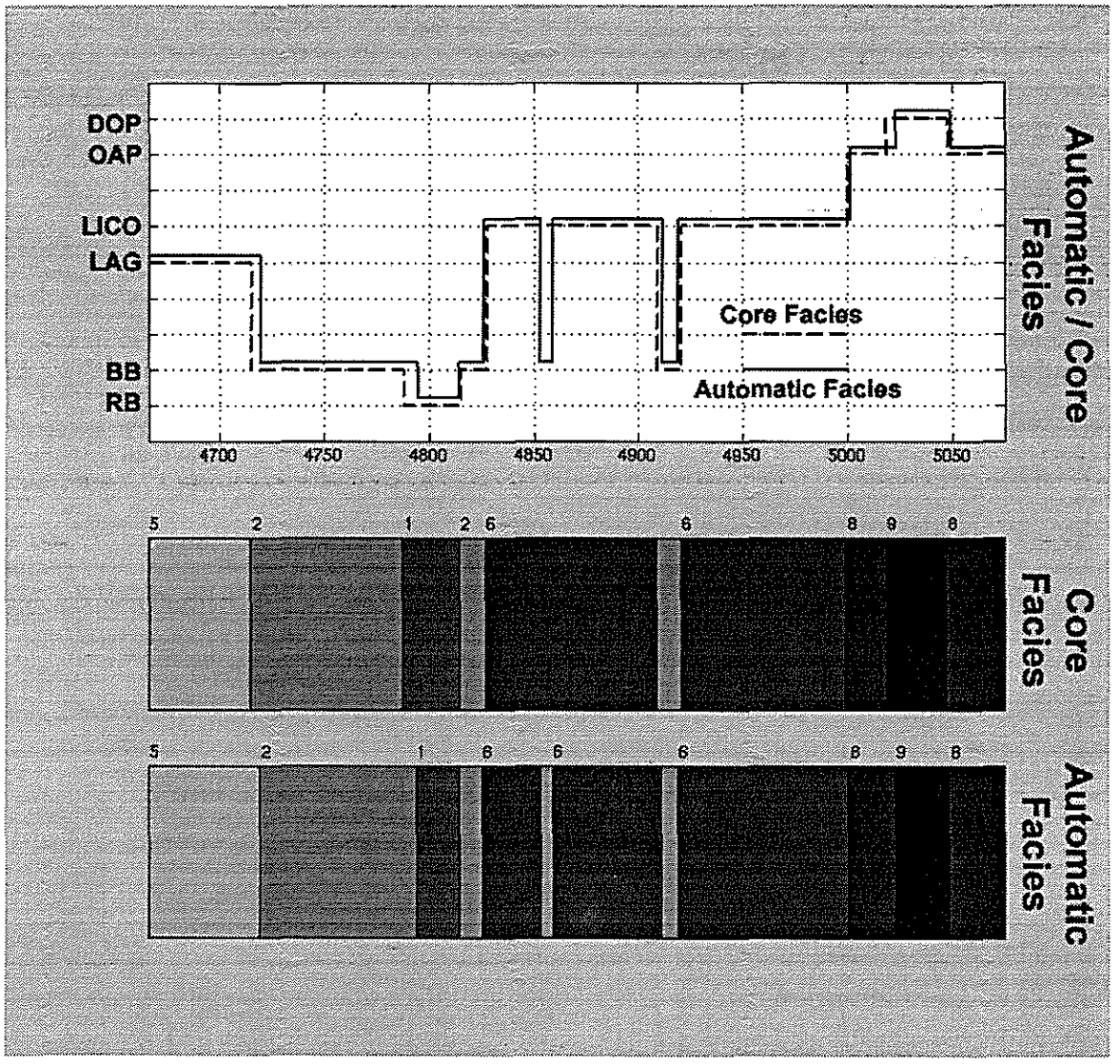

Figure 11: The supervised analysis was repeated on well SVC, but this time the dynamic constraint, *no lithocodium-coral facies deeper than algal facies*, was used. This is an easy constraint to make here, as it is observed to hold all over the field. Application of that constraint further enhanced the accuracy of identification process (compare with Figure 6).