FUSION BASED CLASSIFICATION METHOD AND ITS APPLICATION

LONG JIN, MRINAL K. SEN and PAUL L. STOFFA

Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, U.S.A.

(Received February 13, 2008; revised version accepted January 16, 2009)

ABSTRACT

Jin, L., Sen, M.K. and Stoffa, P.L., 2009. Fusion based classification method and its application. *Journal of Seismic Exploration*, 18: 103-117.

Classification algorithms have many applications both in exploration and production seismology. Many classification algorithms have been reported in the literature. However, for facies identification, lithology/fluid prediction etc, improper choice of an algorithm and parameters for a specific problem will create incorrect classification results. Here, we elaborate on some of these issues and propose a new method based on combining multiple classifiers with Dempster-Shafer theory (DS) that increases the accuracy of classification. The philosophy of our approach is that different classifiers offer complementary information about the patterns to be classified. Thus combining classifiers in an efficient way can achieve better classification results than a single classifier alone can. The effectiveness of this method is demonstrated with a real well log data from North Sea.

KEY WORDS: facies classification, lithology prediction, Dempster-Shafer theory, fusion, neural network, back-propapagation.

INTRODUCTION

Several methods of classification or pattern recognition exist that can automatically extract information from seismic data or well log data. They have been used in seismic facies, lithology/fluid prediction, and time-lapse seismic anomaly classification and identification, etc. Over the years, researchers have

0963-0651/09/$5.00

made rigorous efforts to apply advanced classification methods in exploration and production seismology, including cluster, discriminant analysis, Bayesian classification, neural network, and support vector machine etc. There are numerous papers on this subject. For example, Saggaf et al. (2003) presented an approach based on competitive neural network for the classification and identification of reservoir facies from seismic data. Li and Castagna (2004) reported on the application of a support vector machine in AVO classification of gas and wet sand. Avseth and Mukerji (2002) used three different methods to classify six different facies based on well-log measurements of p-wave velocity and gamma ray. However, there are many uncertainties in the conventional classification methods. One of them is that the accuracy of classification relies on the choice of a classification algorithm and on the parameters related to a specific algorithm. The quality of data is another factor which affects the accuracy and stability of the classification. For supervised classification, the choice of learning data will also influence classification results. Given so many uncertainties, it is a challenging task to get accurate classification results. It appears that there is no single method that can solve all the problems. Here, we propose a fusion based classification method which attempts to address some of these issues.

ALGORITHM

Dempster-Shafer theory

The Dempster-Shafer (DS) theory is a mathematical theory of evidence based on belief functions and plausible reasoning, which is used to combine separate pieces of information (evidence) to calculate the probability of an event. The theory was developed by Arthur P. Dempster and Glenn Shafer (1976). In the following, we summarize the theory and the basic concepts.

There are three important functions in DS theory: the basic probability assignment function (bpa or m), the belief function (Bel), and the plausibility function (Pl).

Let χ be the universal set, the set of all states under consideration. The power set, $P(\chi)$, is the set of all possible sub-sets of χ, including the empty set, Φ.

The basic probability assignment bpa, represented by m, defines a mapping of the power set to the interval between 0 and 1, where the bpa of the null set is 0 and the summation of the bpa's of all the subsets of the power set is 1. The value of the bpa for a given set A [represented as m(A)], expresses the proportion of all relevant and available evidence that supports the claim that a particular element of χ belongs to the set A. The value of m(A) pertains only to the set A and makes no additional claims about any subsets of A. The description of m can be represented with the following three equations:

$$m: P(\chi) \rightarrow [0,1] \quad , \tag{1}$$

$$m(\Phi) = 0 \quad , \tag{2}$$

$$\sum_{A \in P(\chi)} m(A) = 1 \quad , \tag{3}$$

where $P(\chi)$ represents the power set of χ which is the set of all possible sub-sets of χ, Φ is the null set, and A is a set in the power set.

From the basic probability assignment, the upper and lower bounds of an interval can be defined. This interval contains the precise probability of a set of interest and is bounded by two non-additive continuous measures called *belief* and *plausibility*. The lower bound, belief, for a set A is defined as the sum of all the basic probability assignments of the proper subsets B of the set of interest $A(B \subseteq A)$. The upper bound, plausibility, is the sum of all the basic probability assignments of the sets B that intersect the set of interest $A(B \cap A \neq \Phi)$. The formulas are as follows,

$$\mathrm{Bel}(A) = \sum_{B \mid B \subseteq A} m(B) \quad , \tag{4}$$

$$\mathrm{Pl}(A) = \sum_{B \mid B \cap A \neq \Phi} m(B) \quad . \tag{5}$$

The Dempster rule of combination combines multiple belief functions through their basic probability assignments m. These belief functions are defined on the same frame of discernment, but are based on independent arguments or bodies of evidence. Specifically, the combination (called the joint m_{12}) is calculated from the aggregation of two bpa's m_1 and m_2 in the following manner:

$$m_{12}(A) = m_1 \oplus m_2 = [\sum_{B \cap C = A} m_1(B) m_2(C)]/(1 - K) \quad ,$$

$$m_{12}(\Phi) = 0 \quad ,$$

$$K = \sum_{B \cap C = A} m_1(B) m_2(C) \quad , \tag{6}$$

where $\oplus$ is the combination operator, K represents the basic probability mass associated with conflict. This is determined by summing the products of the bpa's of all the sets where the intersection is null. The combination rule can be easily extended to several belief functions by repeating the rule for new belief functions. Thus the pairwise orthogonal sum of n belief functions can be formed as,

$$[(m_1 \oplus m_2) \oplus m_3] \ldots \oplus m_n \quad . \tag{7}$$

This is the basis of our application.

Classification algorithm

A typical scenario of a classification problem is as follows (Avseth and Mukerji, 2005). There is a set of input variables or predictors that influence one or more "outcomes" or "response". The goal of classification is to predict the outcome based on the observed inputs. The classification technique can be carried out in either of the two modes: unsupervised or supervised. The former unsupervised classification tries to cluster the data into groups that are statistically different. For the latter, we have a training data set where we have observed both the inputs and the outcomes. Using the training data we can devise a classification rule or prediction model that will allow us to predict the outcomes for new data where the outcomes are unknown. In this study, we apply two supervised classification algorithms. One is radial basis function network (Buhmann and Ablowitz, 2003) and the other is backpropagation neural network (Dowla and Rogers, 1995). Both of these methods show good properties. The theories of these two methods are described in detail in the cited references and are not discussed here.

Fusion based classification

To account for the uncertainties of choosing different classification algorithms, different parameters, and different input attributes, a fusion based classification method is presented. The workflow of this method is described in Fig. 1.

Fig. 1 shows that multiple classifiers can be combined (fusion) using the DS combination rule (7). The final output from this fusion process is expected to be better than any of the individual outputs from different classifiers. In this way, the accuracy of classification can be improved.

In this paper, we choose radial basis function network (RBF) and backpropagation neural network (BP) as the classifiers; other classifiers can also be used. The outputs of the classifier are considered as the belief functions. The DS combination rule is applied to get a final result.

In a similar manner, we can reduce the uncertainty of attribute selection by combining outputs of multiple classifiers with different attributes.

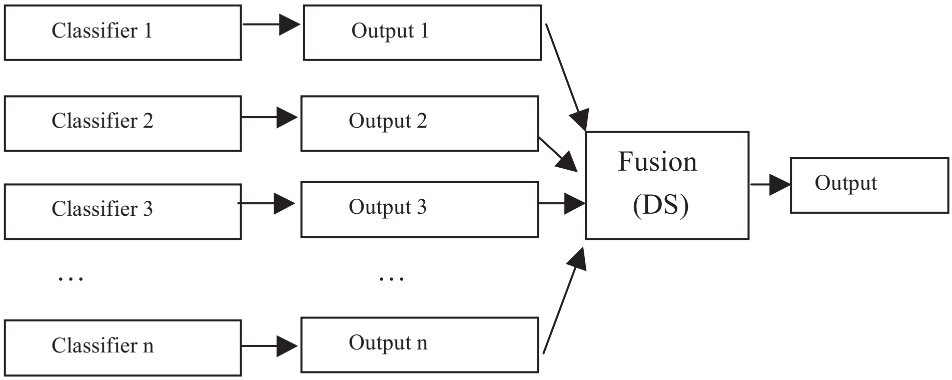

Fig. 1. Workflow of fusion based classification.

EXAMPLE

We apply this fusion based classification method on a real dataset. It is a type well in the oil field which is located in the South Viking Graben in the North Sea. This exercise uses a method similar to that shown in Avseth and Mukerji (2002) to classify different facies based on well measurements of V_P and gamma ray. From the crossplot of V_P and gamma ray, we see that shale can be classified easily from the sand facies (Fig. 2). A three step classification can be used, which may increase the accuracy of classification. First, we separate the sand and shale facies. Then, we discriminate between cemented sand, clean sand and silty-sand. Finally we separate the silty sand into silty-sand1 and silty-sand2. Here, we choose three sand facies including cemented sand, clean sand and silty-sand1 as classification targets (Fig. 3). It is similar to the step 2 of a three step classification. We can also group the silty-sand1 and silty-sand2 to be one class for the step 2 classification.

There are 303 samples for the cemented sand, 67 samples for the clean sand and 106 samples for the silty-sand1. Six samples are randomly chosen as test data for every facies, which are excluded from the training dataset. A three layers BP neural network and radial basis network are used as classifiers. Fig. 4 shows the probability of different classes for the test dataset using BP neural

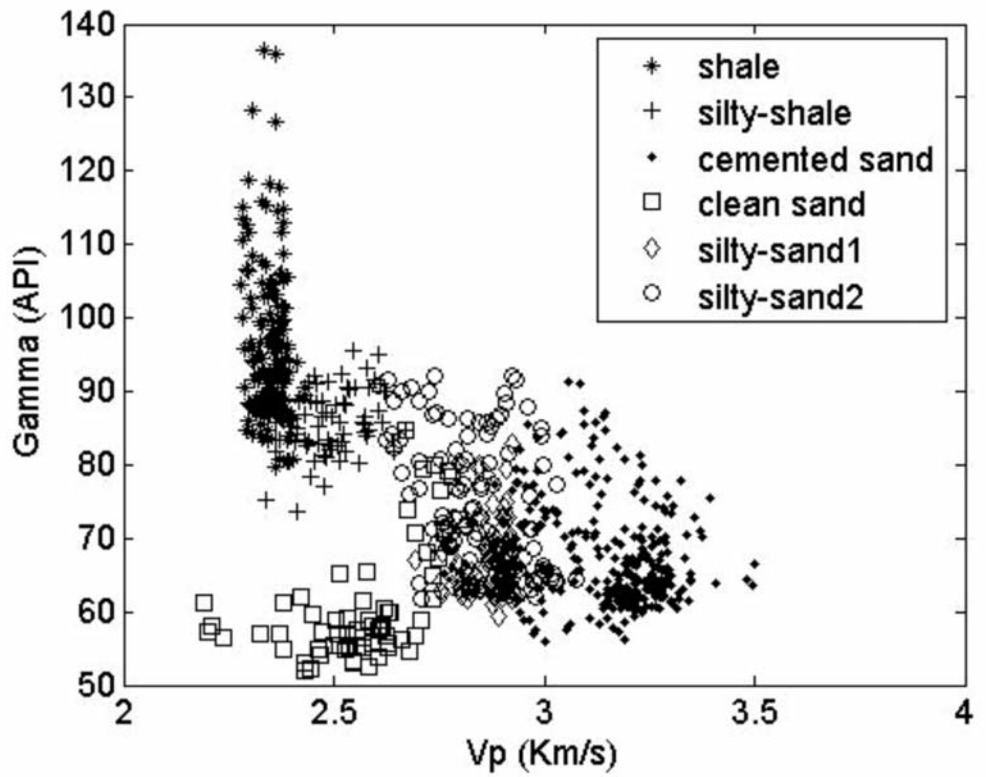

Fig. 2. The crossplot of a well logging data in the North Sea. Six facies are shown in this figure. The shale and sand facies have a good separation.

network. Fig. 5 is the classification result. One sample is misclassified using BP neural network. Fig. 6 shows the probability of different classes for the test dataset using RBF. Fig. 7 is the classification result. Two samples are misclassified with RBF. Then we combine the outputs of BP (Fig. 4) and RBF (Fig. 6) to get the fusion result which is shown as Fig. 8, which combines the two outputs of RBF and BP. The classification result is shown is Fig. 9. It demonstrates that all the samples are correctly classified.

Next, we perform another experiment. The BP and RBF are run 20 times by choosing different input parameters. For the BP neural network, we adjust the size of the second layer. For the RBF, we adjust the spread of the radial basis functions. Then, compute the mean probability of these 20 runs. The mean probabilities from the two methods are further combined to get a final result. Fig. 10 is the mean of the probability for different facies from 20 realizations of BP. Fig. 11 gives the classification result. All the samples are classified correctly. Fig. 12 is the mean of the probability for different facies from 20 realizations of RBF. Fig. 13 shows the classification result based on Fig. 12. All the samples are classified correctly. Thus we know that the accuracy of classification increased by the statistical analysis of multiple runs of classifiers. We further combine the probabilities shown in Figs. 10 and 12 using DS theory. The result is shown in Fig. 14 which shows better separation of different facies than the input probabilities. The classification result is shown is Fig. 15 which shows that all the samples are classified correctly. We define the classification variance as the difference between classification probability and output of the classifier for the corresponding facies. Fig. 16 gives the variances for different methods. It shows that the fusion result has the lowest variance among the three methods. We finally note that even though we apply RBF and BP neural network as classifiers, this fusion based classification method does not depend on the choice of a specific classifier.

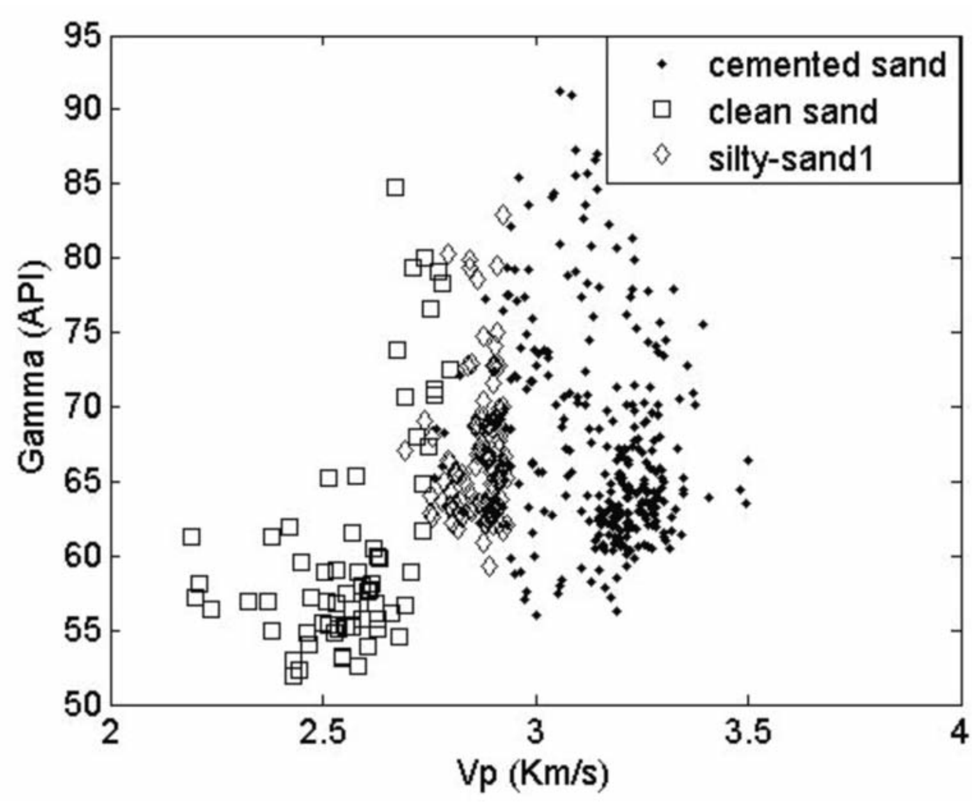

Fig. 3. The crossplot of our chosen data. The data includes three facies: cemented sand, clean sand and silty-sand1. The cemented is considered as class 1. Clean sand is treated as class 2. Silty-sand1 is represented as class 3.

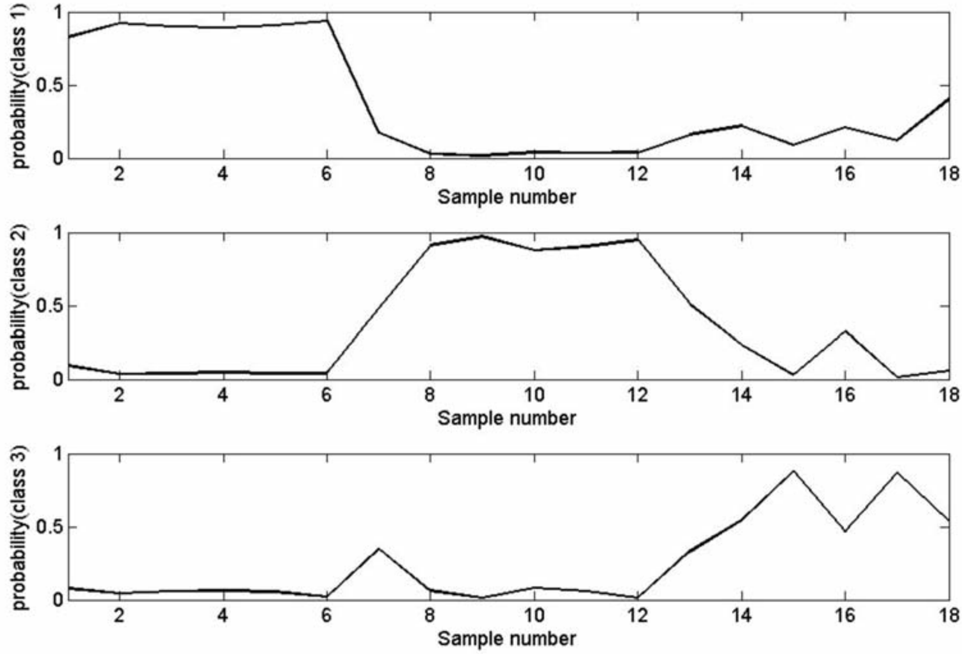

Fig. 4. Probability of different classes using BP.

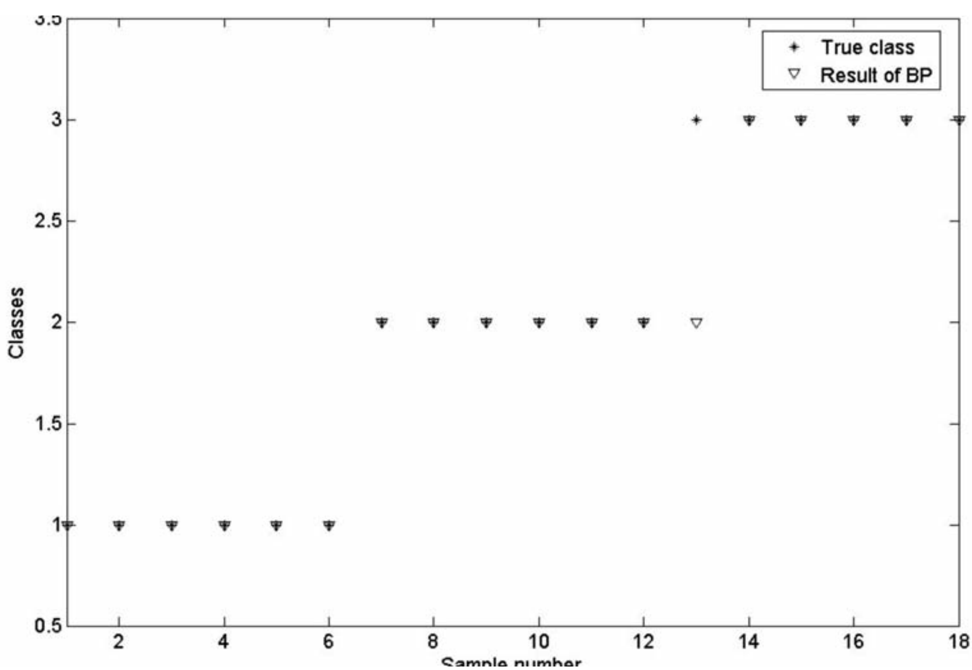

Fig. 5. Classification result of BP neural network based on the probability shown in Fig. 4. One sample is misclassified for the class 2.

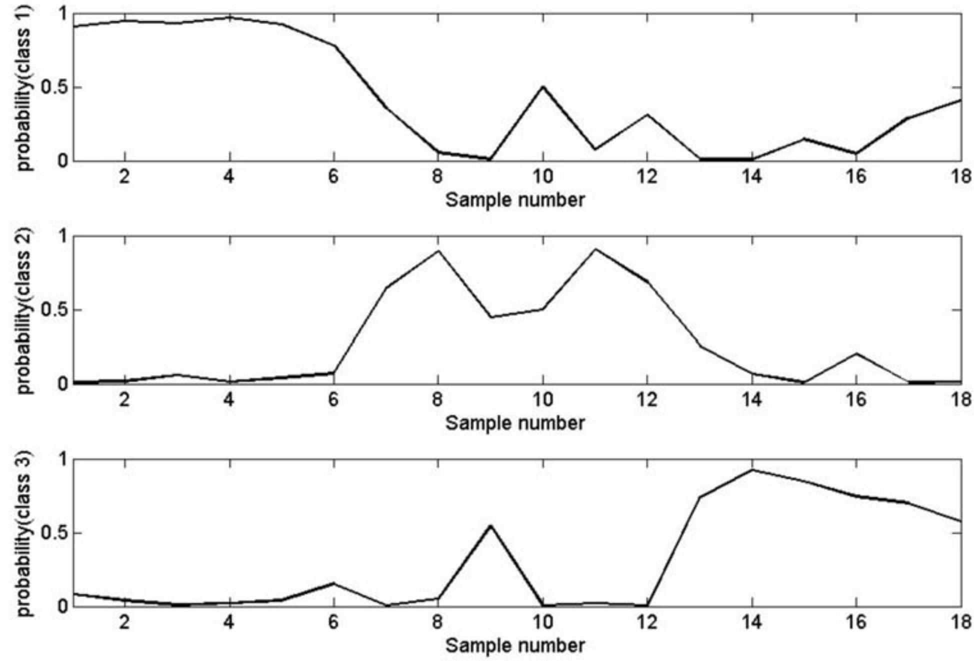

Fig. 6. Probability of different classes using RBF.

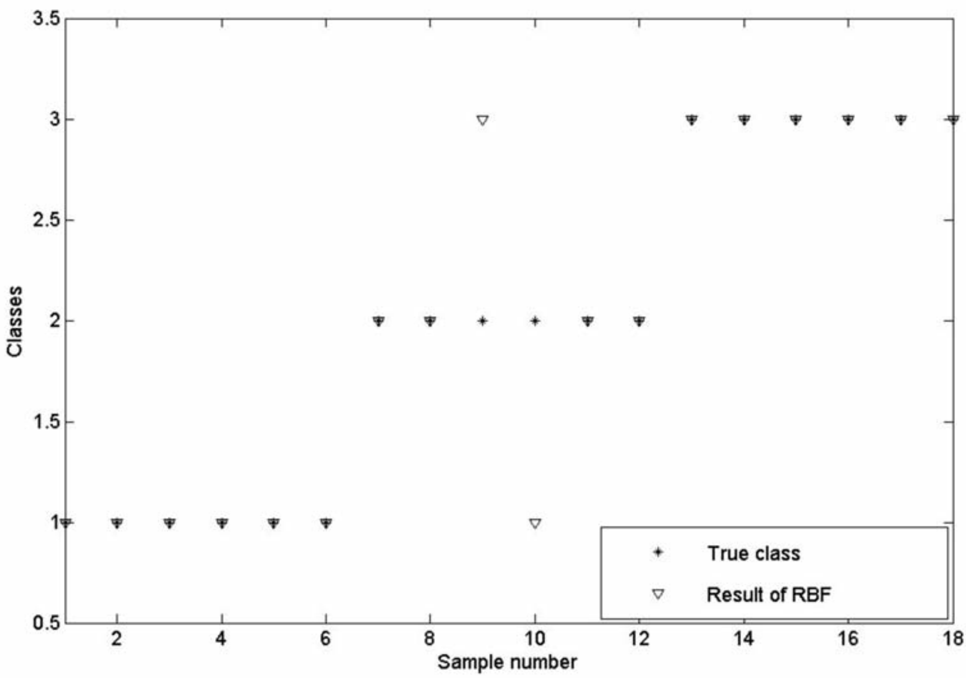

Fig. 7. Classification result of RBF based on the probability shown in figure 6. Two samples are misclassified for the class 2.

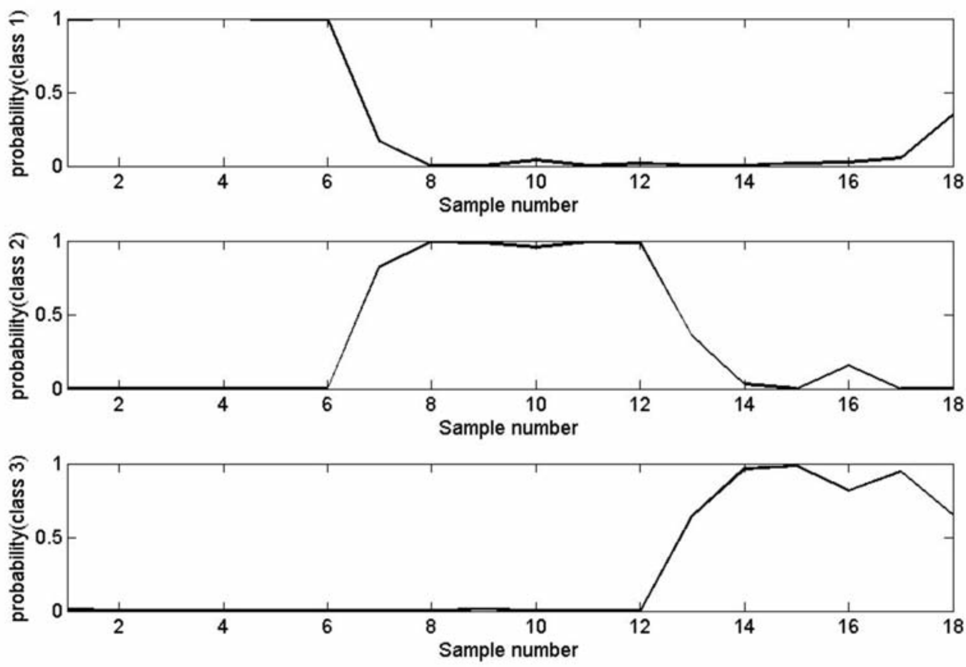

Fig. 8. The Fusion probability of BP and RBF.

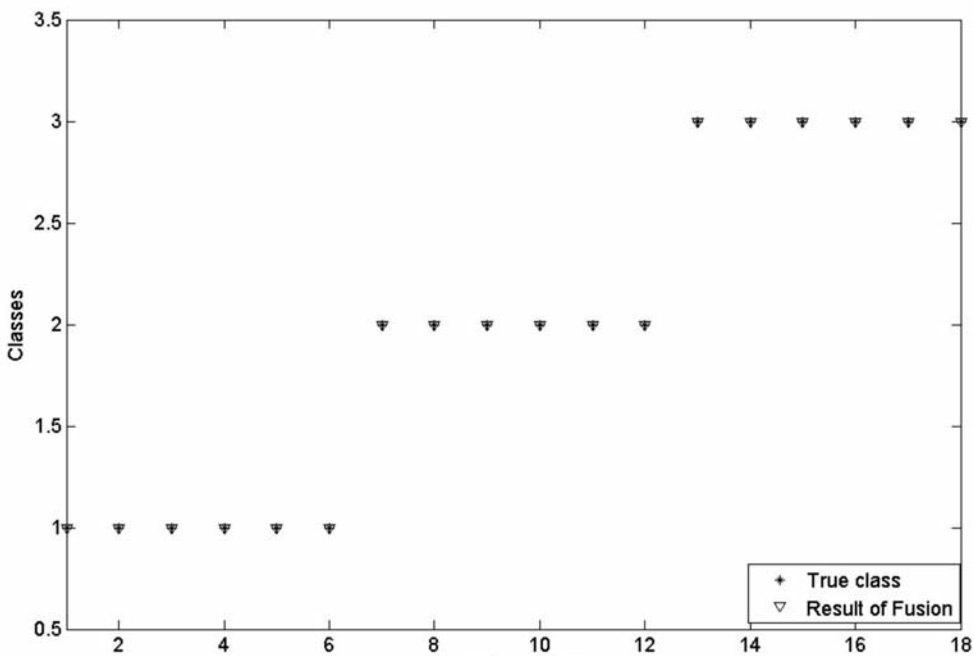

Fig. 9. Classification result of fusion probability shown in Fig. 8. All the samples are correctly classified.

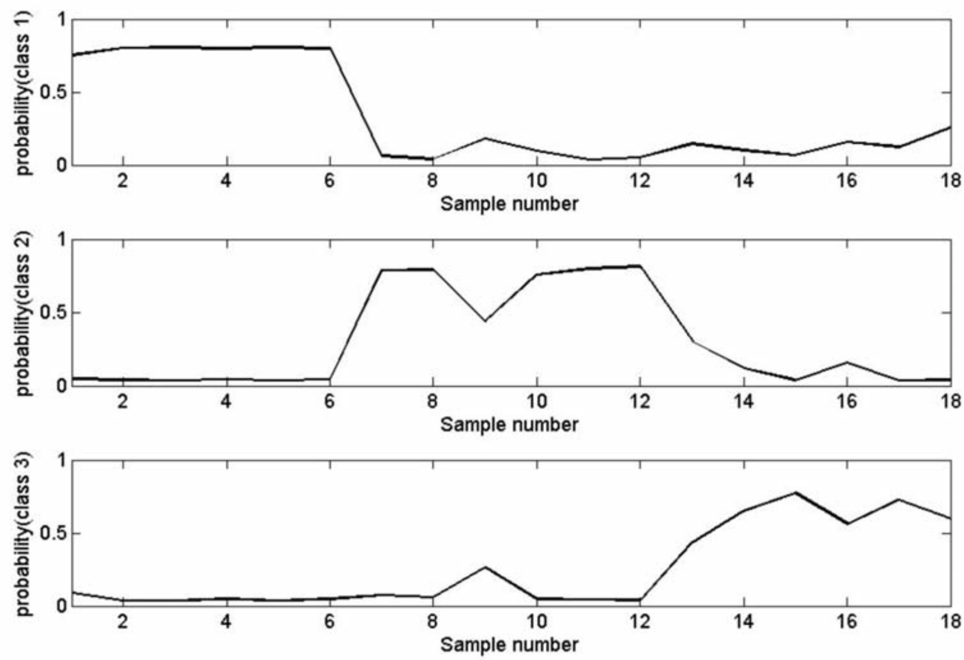

Fig. 10. The mean probability of twenty runs of BP with different parameters.

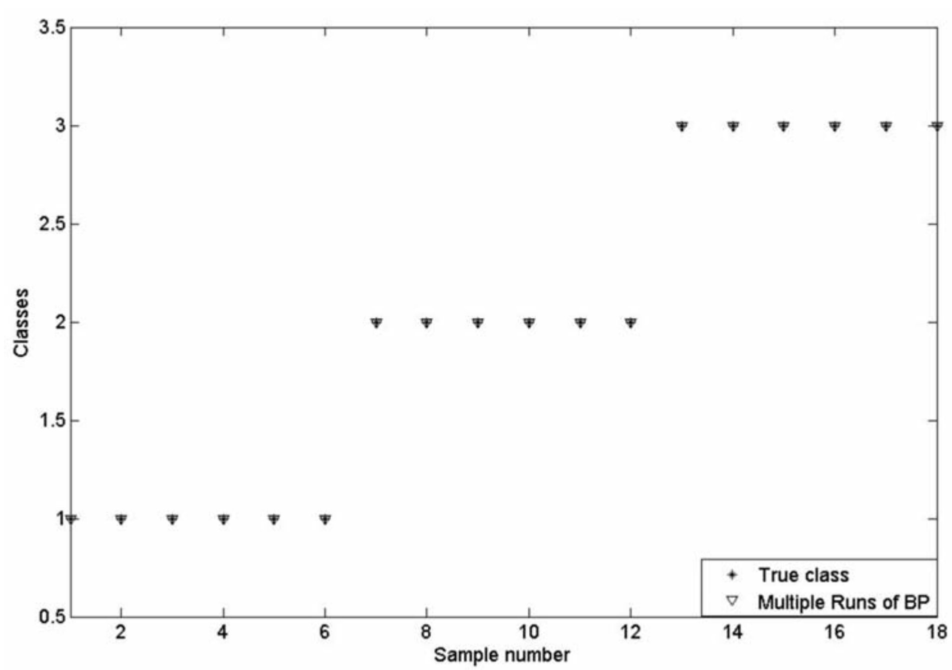

Fig. 11. Classification result based on the probability shown in Fig. 10. All the samples are correctly classified.

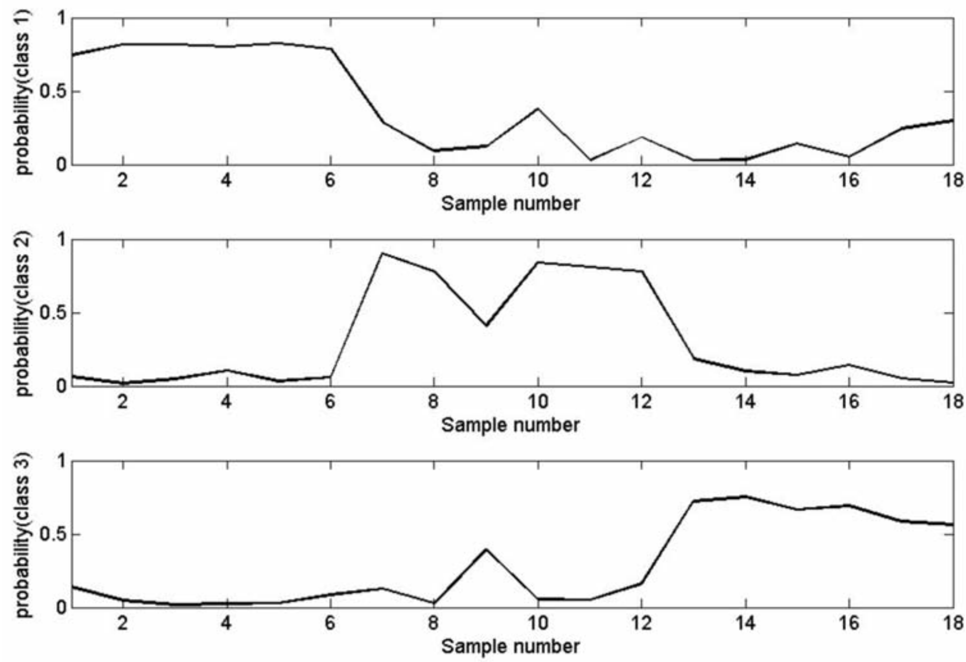

Fig. 12 The mean probability of twenty runs of RBF with different parameters.

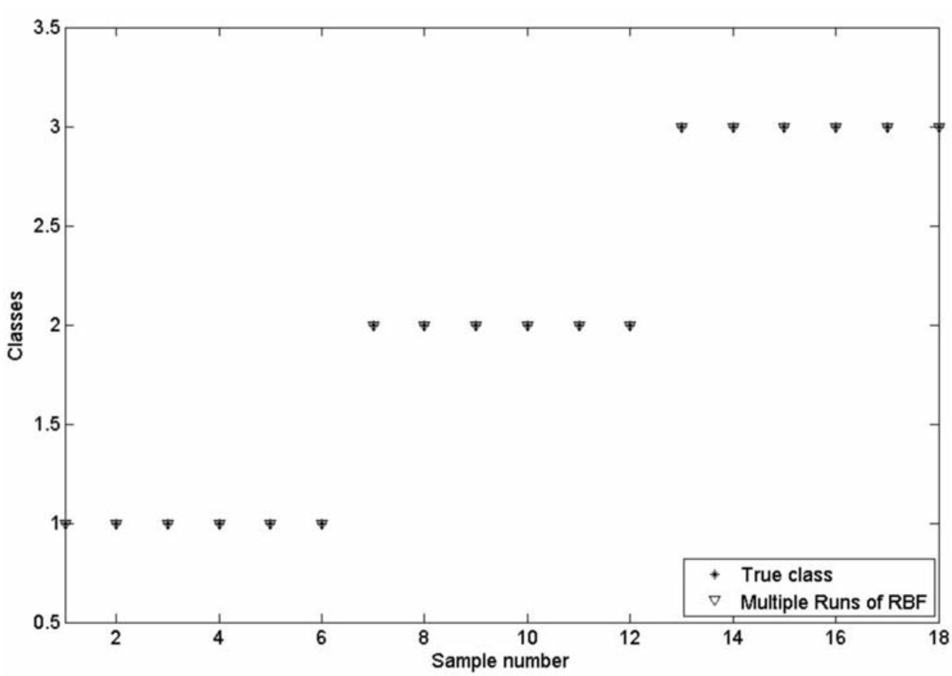

Fig. 13. Classification result based on the probability shown in Fig. 12. All the samples are correctly classified.

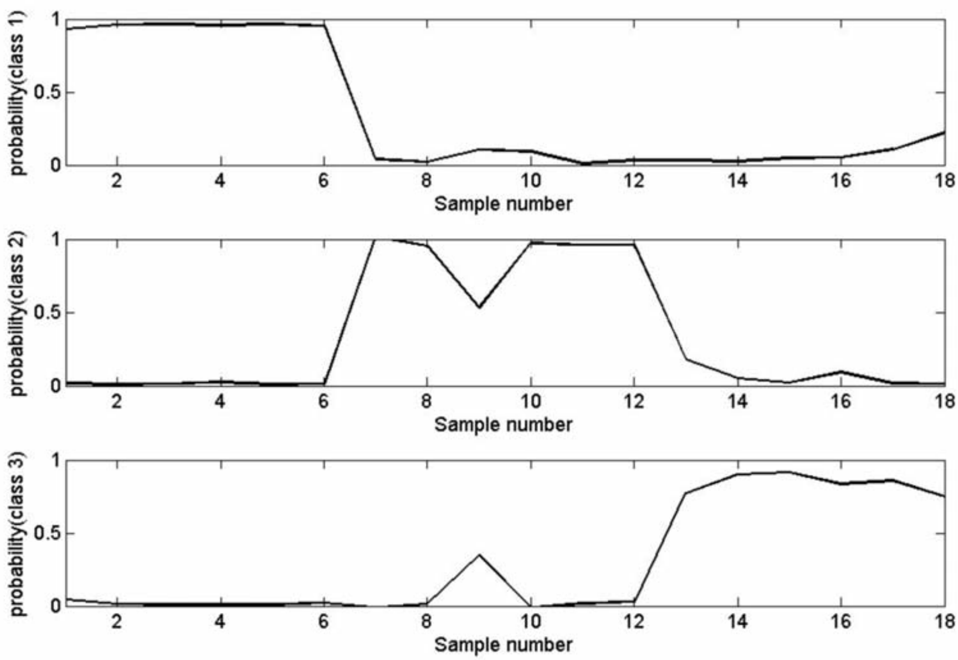

Fig. 14. Fusion result using the mean probabilities shown in Fig. 10 and 12.

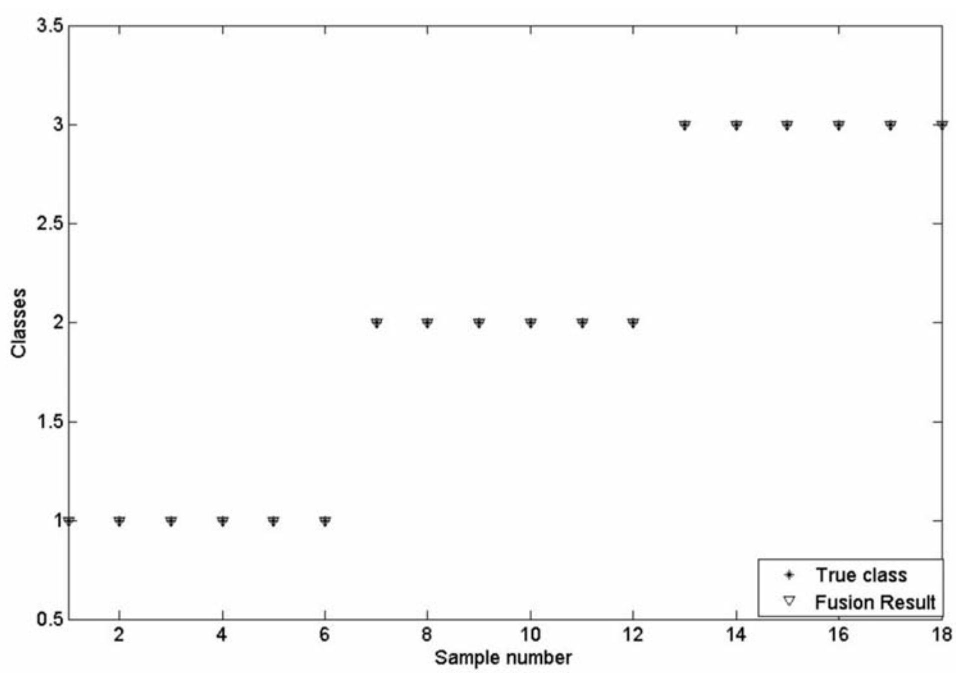

Fig. 15. Classification result based on the fusion probability shown in Fig. 14. All the samples are correctly classified.

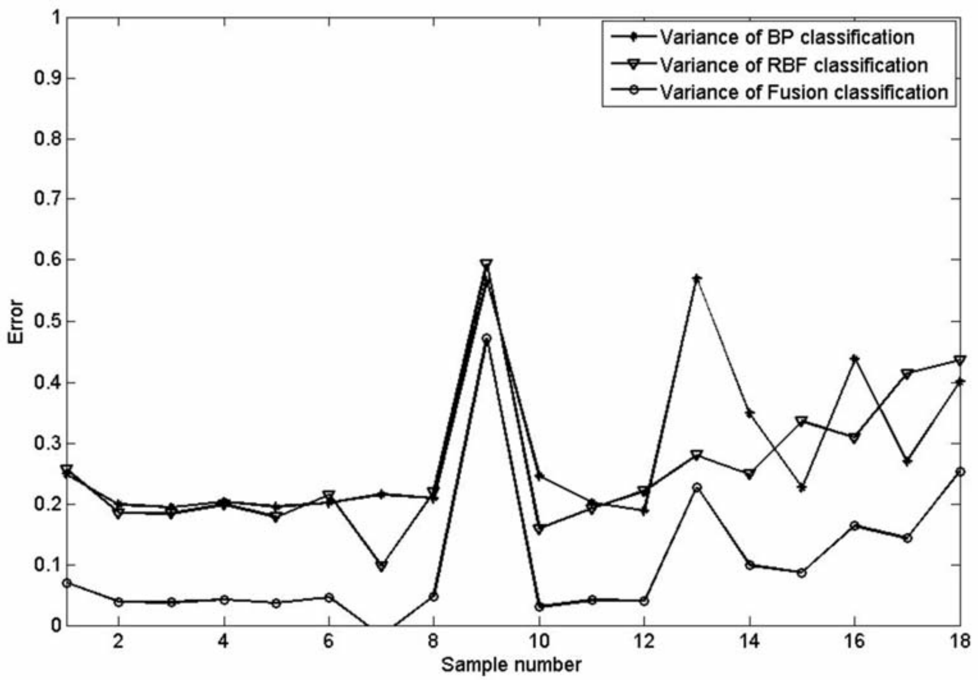

Fig. 16. The variance of different classification methods: multiple runs of BP, multiple runs of RBF and the fusion of multiple runs BP and RBF. The fusion based method can further reduce the variance of classification.

CONCLUSION

We presented a method that combines multiple classifiers based on DS combination rules to improve classification accuracy and reduce the uncertainties related to the choice of suitable classification algorithms and parameters. An example is presented which is based on the real well log data from the North Sea. It shows that the fusion based classification does improve the accuracy and stability of classification of shale and different sand facies. The statistical analysis of multiple runs of a specific classifier with different parameters is another way to reduce the uncertainty of the choice of parameters. The fusion of statistical classification results further reduces the variance of the classification. We demonstrated that this fusion based classification is a general method and does not depend on the specific classifiers and therefore, appears to be a promising tool.

ACKNOWLEDGMENTS

Long Jin was supported on a grant from Conoco-Phillips and Jackson School of Geosciences. We thank Dr. Xiaohong Chen and Dr. Shoudong Wang for many helpful discussions.

REFERENCES

Avseth, P. and Mukerji, T., 2002. Seismic lithofacies classification from well logs using statistical rock physics. Petrophysics, 43: 70-81.

Avseth, P. and Mukerji, T., 2005. Quantitative Seismic Interpretation. Cambridge University Press, Cambridge.

Dowla, F.U. and Rogers, L.L., 1995. Solving Problems in Environmental Engineering and Geosciences with Artificial Neural Networks. MIT Press, Cambridge, MA.

Holmes, C.C. and Mallick, B.K., 1998. Bayesian radial basis functions of variable dimension. Neural Computat., 10: 1217-1233.

Li, J. and Castagna, J., 2004. Support vector machine pattern recognition to AVO classification. Geophys. Res. Lett., 31: L02609.

Martin, D., Buhmann, M. and Ablowitz, J., 2003. Radial Basis Functions: Theory and Implementations. Cambridge University Press, Cambridge.

Saggaf, M.M., Toksöz, M,N. and Marhoon, M.I., 2003. Seismic facies classification and identification by competitive neural networks. Geophysics, 68: 1984-1999.

Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press, Princeton.

Shafer, G., 1990. Perspectives on the theory and practice of belief functions. Internat. J. Approxim. Reason., 3: 1-40.