Application of Probabilistic Neural Network to Porosity Prediction

Philippe Ecoublet and William W. Symes *

ABSTRACT

In a previous report (Ecoublet, Symes and Levin, 1997) we investigated, using a backpropagation neural network, how seismic inversion can improve the prediction of fluid properties and lithologies as compared to seismic attributes prediction. Although promising results were obtained, we had no systematic procedure to set the optimal parameters of the backpropagation neural network, in terms of number of layers and number of neurons per layer, according to the complexity of the mapping to be achieved. Also training the network is a lengthy process without the certitude of finding the global optimum.

One alternative to the backpropagation neural network is the probabilistic neural network which offers the advantage of not requiring any training. The probabilistic neural network uses non-parametric estimators of the density probability function to compute decision boundaries based on Bayes' inference theory and which are asymptotically Bayes-optimal. A probabilistic neural network is applied to porosity prediction from well logs. A performance analysis is conducted on predicting porosity from seismic inversion using either a backpropagation neural network or a probabilistic network.

Introduction

By substituting the sigmoid transfer function used in backpropagation neural network with an estimator of the probability density function, a probabilistic neural network (PNN) can be built to classify input patterns (Specht, 1990a). We propose to use such an approach to predict porosity intervals from well logs and seismic inversion, and compare with prediction capabilities previously achieved with a backpropagation neural network.

*The Rice Inversion Project, Department of Computational and Applied Mathematics, Rice University, Houston TX 77005-1892 USA, email `ecoublet@caam.rice.edu`

We previously observed (Ecoublet, Symes and Levin, 1997) that relative fluctuations of porosity ϕ are best predicted from relative fluctuations of density ρ, P-impedance Z_p, and S-impedance Z_s.

Let's assume a division of the porosity range into N_{class} intervals or *classes*. We want to determine in which class an input vector $\mathbf{X}$, with components (ρ, Z_p, Z_s), belongs. The probabilistic neural network estimates the probability density function of each class from the training vectors and computes the probability that the input vector $\mathbf{X}$ belongs in each of the classes.

The Bayes strategy for pattern classification

The goal of Bayes' strategy is to classify a new vector $\mathbf{X}$ between N_{class} classes in such a way as to minimize the probability of misclassification.

A set of training vectors and their corresponding porosity values constitute the measurements. The porosity is divided into intervals and each training vector is assigned to its proper class θ_k, $k \in \{1, ..., N_{class}\}$.

The prior probability of occurence of class θ_k, denoted $P(\theta_k)$, is then given by the ratio of the number of training vectors belonging to class θ_k, to the total number of training vectors.

The joint probability density of observing a new vector $\mathbf{X}$ which belongs to class θ_k, $p(\mathbf{X}, \theta_k)$, is given by the product of the class-conditional probability density of observing the vector $\mathbf{X}$ given that it belongs to class θ_k, $p(\mathbf{X}|\theta_k)$, with the prior probability of class θ_k. The joint probability density takes the form

$$p(\mathbf{X}, \theta_k) = p(\mathbf{X}|\theta_k)P(\theta_k).$$

Following a similar argument, the joint probability density can also be written in the form

$$p(\theta_k, \mathbf{X}) = P(\theta_k|\mathbf{X})p(\mathbf{X}),$$

where $P(\theta_k|\mathbf{X})$ is the probability that the class is θ_k given the vector $\mathbf{X}$, and $p(\mathbf{X})$ is the probability density of observing a vector $\mathbf{X}$ independently of the class membership.

The Bayes' decision theory, see for example (Bishop, 1996), gives the posterior probability $P(\theta_k|\mathbf{X})$ in terms of the prior probability $P(\theta_k)$ and class-conditional probability density $p(\mathbf{X}|\theta_k)$,

$$P(\theta_k|\mathbf{X}) = \frac{p(\mathbf{X}|\theta_k)P(\theta_k)}{p(\mathbf{X})}.$$

To ensure that the posterior probability sums to unity, i.e.

$$\sum_{k=1}^{N_{class}} P(\theta_k|\mathbf{X}) = 1,$$

the probability density of $\mathbf{X}$, $p(\mathbf{X})$ satisfies

$$p(\mathbf{X}) = \sum_{k=1}^{N_{class}} p(\mathbf{X}|\theta_k)P(\theta_k).$$

The posterior probability of the class θ_k given the vector $\mathbf{X}$ can then be expressed in the form

$$P(\theta_k|\mathbf{X}) = \frac{p(\mathbf{X}|\theta_k)P(\theta_k)}{\sum_{k=1}^{N_{class}} p(\mathbf{X}|\theta_k)P(\theta_k)}.$$

According to Bayes' theorem, the probability of misclassification of a new vector is minimized by assigning that vector to the class having the largest posterior probability, so that a new vector $\mathbf{X}$ is assigned to class θ_k if

$$P(\theta_k|\mathbf{X}) > P(\theta_j|\mathbf{X}), \quad \forall j \neq k.$$

A different decision strategy, which consists of minimizing a "cost" of misclassification, may also be applied. We define a loss matrix $\mathcal{L}$, with elements $\mathcal{L}_{kj}$, specifying the loss associated with assigning a vector $\mathbf{X}$ to class θ_j when it actually belongs to class θ_k. The overall risked loss is given by

$$\mathcal{R}_k = \sum_{j=1}^{N_{class}} \mathcal{L}_{kj} P(\theta_j|\mathbf{X}).$$

Devising an appropriate definition for the loss matrix is not trivial and different scenarios have been tested. In any case the diagonal terms of $\mathcal{L}$ are null, since there is no loss in classifying correctly. We observed a minimum classification root mean square error when defining the loss of misclassification between two consecutive classes to be a constant α. The loss matrix in this measure becomes

$$\mathcal{L}_{kj} = \alpha|j-k|.$$

Following such a decision strategy, the new vector $\mathbf{X}$ will be assigned to class θ_k for which the loss of misclassification is minimum, i.e.

$$\sum_{i=1}^{N_{class}} \mathcal{L}_{ki} P(\theta_i|\mathbf{X}) < \sum_{i=1}^{N_{class}} \mathcal{L}_{ji} P(\theta_i|\mathbf{X}), \quad \forall j \neq k.$$

Observing that the coefficient α cancels, we can take the loss matrix to be

$$\mathcal{L}_{kj} = |j - k|,$$

and a new vector $\mathbf{X}$ will be assigned to class θ_k if

$$\sum_{i=1}^{N_{class}} |i-k| P(\theta_i|\mathbf{X}) < \sum_{i=1}^{N_{class}} |i-j| P(\theta_i|\mathbf{X}), \quad \forall j \neq k.$$

Probability density estimation

The class-conditional probability density $p(\mathbf{X}|\theta_k)$ can be estimated by non-parametric functional forms for which no prior assumption is made about the shape of the distribution. The shape of the distribution is entirely determined by the data set drawn from the distribution, i.e. the training data clustered in each class θ_k. Among the non-parametric techniques of probability density estimation (Bishop, 1996), kernel-based methods model the probability density as a sum of kernel functions centered at each of the training data vectors. Such a modeling technique is asymptotically optimal, i.e. the decision surface defined by $p(\mathbf{X}|\theta_k)$ can approach Bayes' optimal criterion.

For a kernel function we chose a multivariate Gaussian distribution of the form

$$p(\mathbf{X}|\theta_k) = \frac{1}{N_{vect}^k} \sum_{n=1}^{N_{vect}^k} \exp\left(-\frac{||\mathbf{X} - \mathbf{X_n^k}||^2}{2\sigma^2}\right),$$

where N_{vect}^k is the number of training vectors clustered in class θ_k, $\mathbf{X_n^k}$ is the n^{th} training vector belonging to class θ_k, and σ is a smoothing parameter.

The shape of the decision surfaces can be made as complex as necessary by varying the value of the smoothing parameter σ (Specht, 1990b).

Probabilistic neural network for porosity prediction at log scale

As in the experiments described in (Ecoublet, Symes and Levin, 1997) and carried out on log data, wells 58c and 58cst are used as training wells, and the wells 55a and 56c used to test the network.

The relative fluctuations of porosity for the 3550 training vectors range between -0.89 and 0.905. We chose for this study to divide the porosity into 10 intervals, from [-1.0,-0.8] to [0.8,1.0]. The training vectors are then sorted according to the class of porosity they fall into. The number of training vectors belonging to each class is given in Table 1.

The prior probability of occurence of class θ_k, $P(\theta_k)$, should be calculated from the observations, i.e. the training vectors, summarized in Table 1. We observed that for a relatively large value of σ, typically $\sigma \simeq 1$, $p(\mathbf{X}|\theta_k)$ becomes negligible compared to $P(\theta_k)$ and any new vector $\mathbf{X}$ is classified into class θ_8 which has the highest prior probability. However, $p(\mathbf{X}|\theta_k)$ contains the most valuable information for classification.

For this reason we assume that the prior probability of occurence of each of the N_{class} classes is independent of the training data, and is given by a uniform distribution

$$P(\theta_k) = \frac{1}{N_{class}}.$$

Therefore, $P(\theta_k)$ can be dropped in the computation of the posterior probability $P(\theta_k|\mathbf{X})$.

Class θ_k	Relative fluctuations of ϕ	N_{vect}^k	Class θ_k	Relative fluctuations of ϕ	N_{vect}^k
1	[-1.0,-0.8]	131	6	[0.0, 0.2]	296
2	[-0.8,-0.6]	314	7	[0.2, 0.4]	539
3	[-0.6,-0.4]	186	8	[0.4, 0.6]	948
4	[-0.4,-0.2]	120	9	[0.6, 0.8]	637
5	[-0.2, 0.0]	222	10	[0.8, 1.0]	157

Table 1. Clustering of the training data into 10 porosity classes

Our PNN has only one parameter, the smoothing coefficient σ which controls the shape of the class-conditional probability density $p(\mathbf{X}|\theta_k)$. The smaller σ is, the sharper and more selective $p(\mathbf{X}|\theta_k)$ becomes.

The porosity logs of the training wells 58c and 58cst have been predicted for different values of σ and the prediction error reported in Table 2. The classification decision is either based on maximizing the posterior probability $P(\theta_k|\mathbf{X})$ for class θ_k or on minimizing a risked loss from misclassification $\mathcal{R}_k$ for class θ_k.

	$max\ P(\theta_k\|\mathbf{X})$		$min\ \mathcal{R}_k$	
σ	rms $\times 10^2$	$C_p\%$	rms $\times 10^2$	$C_p\%$
1.	08.61	95.42	09.32	94.70
0.1	08.21	95.39	06.27	97.13
0.01	04.64	98.45	04.15	98.74
0.001	01.27	99.88	01.27	99.88
$\leq$ 0.0005	0	100	0	100

Table 2. Rms error and prediction coefficient between targeted and predicted interval porosity from training wells 58c and 58cst, for different σ. The prediction error is given for classification either based on $P(\theta_k|\mathbf{X})$ maximization or on $\mathcal{R}_k$ minimization.

A prediction coefficient C_p defined by

$$C_p = \frac{T_\phi \cdot P_\phi}{||T_\phi|| \; ||P_\phi||},$$

is computed in addition to the rms error, and provides a measure of the match between predicted P_ϕ and targeted T_ϕ porosity.

As the following results will confirm, classification based on risk minimization enabled us to predict porosity with a smaller error than classification based on posterior probability maximization.

An illustration of the influence of the smoothing parameter σ on the probability density estimation is illustrated in Figure 1 where the posterior probability $P(\theta_k|\mathbf{X})$ is plotted for each class θ_k for the first hundred training vectors from wells 58c and 58cst.

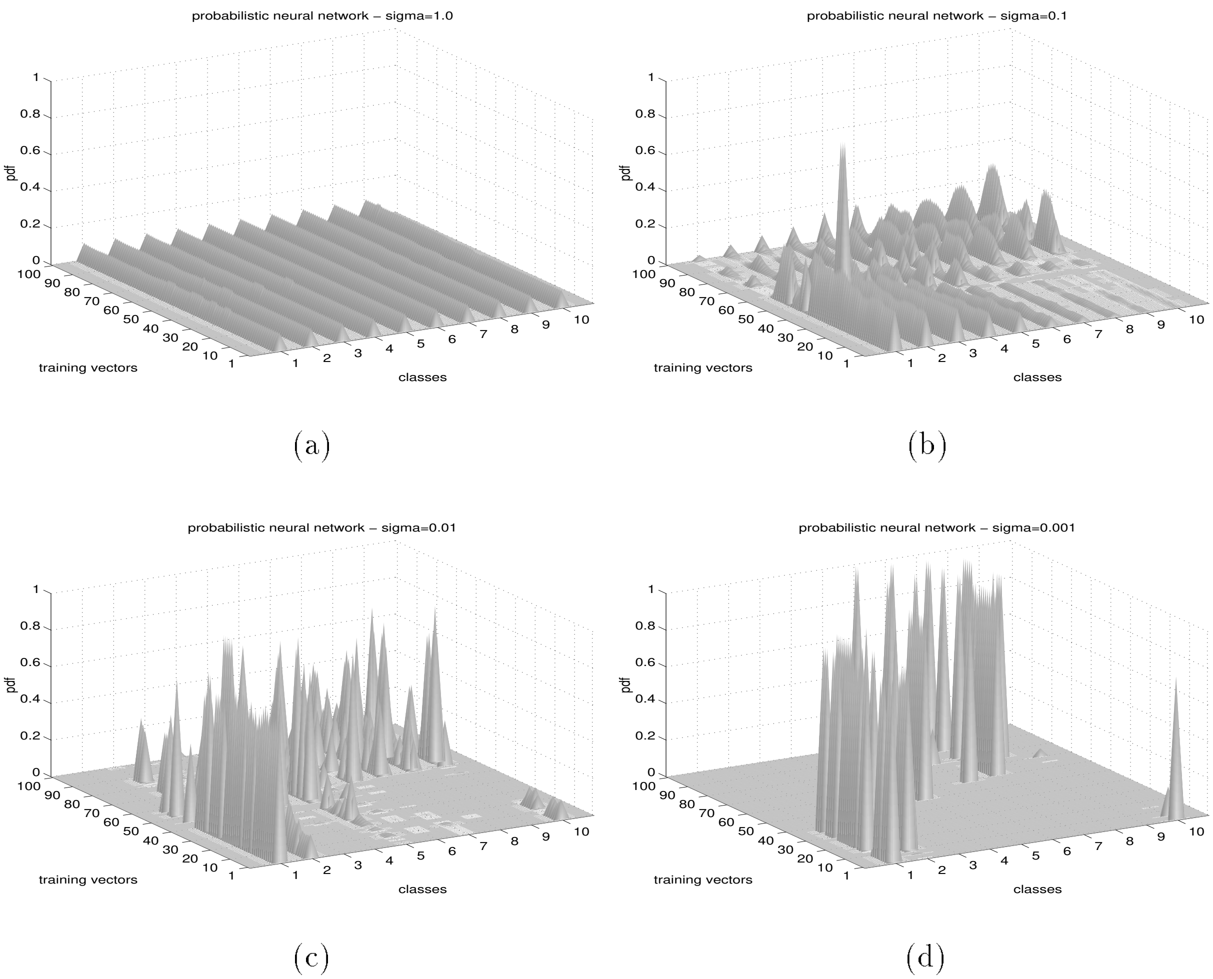

FIG. 1. The posterior class probability $P(\theta_k|\mathbf{X})$ is plotted for the first hundred training vectors $\mathbf{X}$ from training wells 58c and 58cst, with different smoothing parameter: (a) $\sigma = 1.0$, (b) $\sigma = 0.1$, (c) $\sigma = 0.01$, and (d) $\sigma = 0.001$.

For $\sigma = 1.0$ (Figure 1 (a)), the class-conditional probability density $p(\mathbf{X}|\theta_k)$ is smooth, and, although posterior class probability $P(\theta_k|\mathbf{X})$ looks the same for each class, porosity is predicted with an rms error of 0.0932 (Table 2) when minimizing $\mathcal{R}_k$. As σ decreases (Figure 1 (b), (c) and (d)), $p(\mathbf{X}|\theta_k)$ becomes sharper and the probability of misclassification tends to zero.

Too large a choice of σ produces a smooth PDF with an increased risked loss from misclassification. On the other hand, if σ is too small, the network will either classify a new vector with a high probability or will fail to classify it at all.

To illustrate these observations, we applied our PNN to predict the porosity log of the test well 56c by chosing different values for σ. The classification errors are summarized in Table 3, where the classification decision is either based on posterior probability maximization, $P(\theta_k|\mathbf{X})$, or based on minimizing the risked loss from misclassification $\mathcal{R}_k$.

	max $P(\theta_k\|\mathbf{X})$		*min* $\mathcal{R}_k$		
σ	rms $\times 10^2$	$C_p\%$	rms $\times 10^2$	$C_p\%$	no classification
0.25	08.10	95.57	05.22	98.87	0
0.2	08.06	95.59	04.73	98.94	0
0.15	08.01	95.60	04.53	98.81	0
0.1	07.80	95.68	04.65	98.45	0
0.05	07.44	95.90	05.08	98.06	0
0.01	06.20	97.15	05.36	97.85	0
0.005	06.39	97.03	05.90	97.47	50

Table 3. Rms error and prediction coefficient between targeted and predicted interval porosity from test well 56c, for different values of the smoothing parameter σ.

Again, porosity is better predicted when the minimization of a risked loss is chosen to be the classification rule. For σ smaller than 0.01, some input vectors are not classified, the unconditional probability density $p(\mathbf{X})$ equals zero. Note that the minimum prediction error does not correspond to the smallest σ for which all vectors are classified.

We investigated a way of reducing the rms error by chosing an appropriate value for the smoothing parameter σ. Table 4 gives the standard deviation computed for each input log of the training and test wells.

standard deviation				
Well	Number of Vectors	σ_ρ	σ_{Z_p}	σ_{Z_s}
58c+st	3550	0.064	0.113	0.118
56c	2574	0.048	0.092	0.110
55a	2088	0.065	0.138	0.133

Table 4. Standard deviation of the relative fluctuations of ρ, Z_p and Z_s well logs, for training wells (58c and 58cst) and for test wells (55a and 56c).

Instead of defining the class-conditional probability density $p(\mathbf{X}|\theta_k)$ with a constant parameter σ, we can take the deviation of each component of the input vectors into account by replacing σ with a diagonal covariance matrix, denoted $\mathbf{\Sigma}$, which has the squared standard deviations for diagonal terms,

$$\mathbf{\Sigma} = \begin{pmatrix} {\sigma_\rho}^2 & 0 & 0 \\ 0 & {\sigma_{Z_p}}^2 & 0 \\ 0 & 0 & {\sigma_{Z_s}}^2 \end{pmatrix}.$$

The expression of the class-conditional probability density $p(\mathbf{X}|\theta_k)$ becomes

$$p(\mathbf{X}|\theta_k) = \frac{1}{N^k_{vect}} \sum_{n=1}^{N^k_{vect}} \exp\left(-\frac{1}{2}(\mathbf{X} - \mathbf{X^k_n})^T \mathbf{\Sigma}^{-1} (\mathbf{X} - \mathbf{X^k_n})\right),$$

where ^{T} denotes the vector transpose.

The prediction results for test wells 56c and 55a are summarized in Table 5, and are to be compared with results obtained using a Backpropagation Neural Network (BPNN), (Table 10 in (Ecoublet, Symes and Levin, 1997)) for the prediction version of the BPNN, (Table 11 in (Ecoublet, Symes and Levin, 1997)) for the classification version of the BPNN.

Well	*max* $P(\theta_k|\mathbf{X})$		*min* $\mathcal{R}_k$	
	rms $\times 10^2$	$C_p\%$	rms $\times 10^2$	$C_p\%$
58c+st	08.05	95.69	06.20	97.18
56c	06.88	96.81	04.10	98.76
55a	11.25	92.88	08.55	94.05

Table 5. Rms error and prediction coefficient between targeted and predicted interval porosity from training on relative fluctuations of ρ, Z_p and Z_s at log scale. The smoothing coefficients $\mathbf{\Sigma}$ are computed from the logs.

The predicted porosity logs for wells 58c plus 58cst, 56c and 55a are plotted in Figures 2, 3 and 4 respectively. The comparison can be made with interval porosity prediction achieved using the BPNN in its classification version (Figures 29, 30 and 31 in (Ecoublet, Symes and Levin, 1997)).

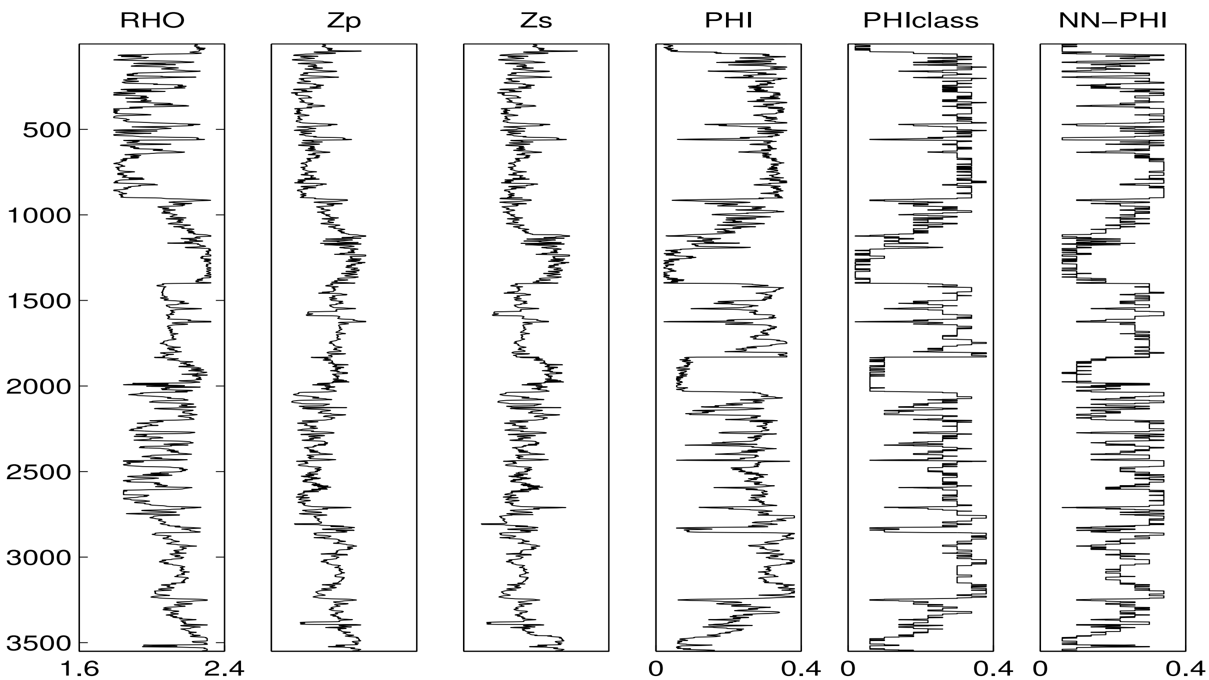

FIG. 2. Interval porosity prediction from relative fluctuations of density and impedance logs. Training data from wells 58c and 58cst. Porosity is classified into 10 intervals. *PHI* is the actual log, *class-PHI* the targeted interval porosity log represented by its mean value, and *NN-PHI* the predicted interval porosity. Vertical coordinates are in depth samples. P-impedance ranges within [0.01,0.03] and S-impedance within [0.005,0.015].

We investigated the possibility of reducing the residual error by adding synthetic training vectors to the actual training data set. The objective was to strengthen the discrimination of the network by artificially increasing the training data set. The shape of the probability density is then modified accordingly to the statistics of the synthetic data.

Up to a thousand of synthetic vectors were generated from the covariance matrix of the actual training data clustered in each class. By applying a factor to the covariance, the shape of the probability density can be flatten or made sharper. Two observations can be made from numerical results. First, increasing the number of synthetic vectors from a hundred to a thousand does not yield to a reduction of the residual error. Second, increasing the covariance of the synthetic data reduces the number of no classification by spreading the probability density at the cost of an higher misclassification rate.

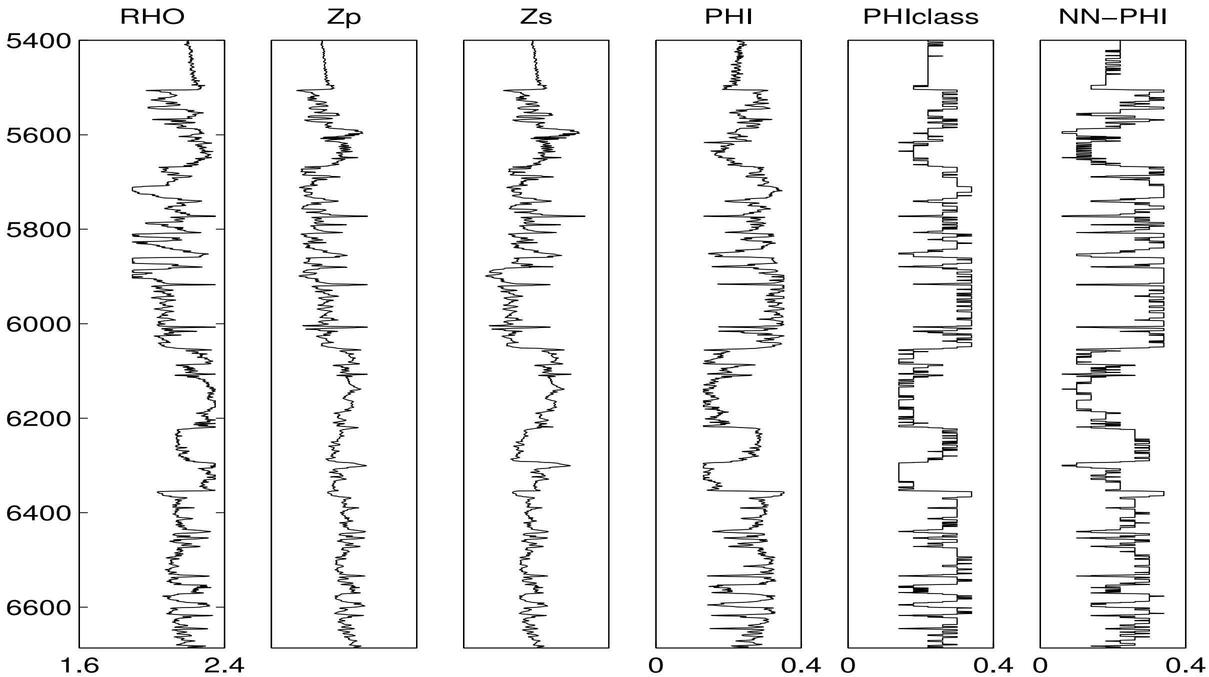

FIG. 3. Interval porosity prediction from relative fluctuations of density and impedances logs. Test data from well 56c. Vertical coordinates are in measured depth.

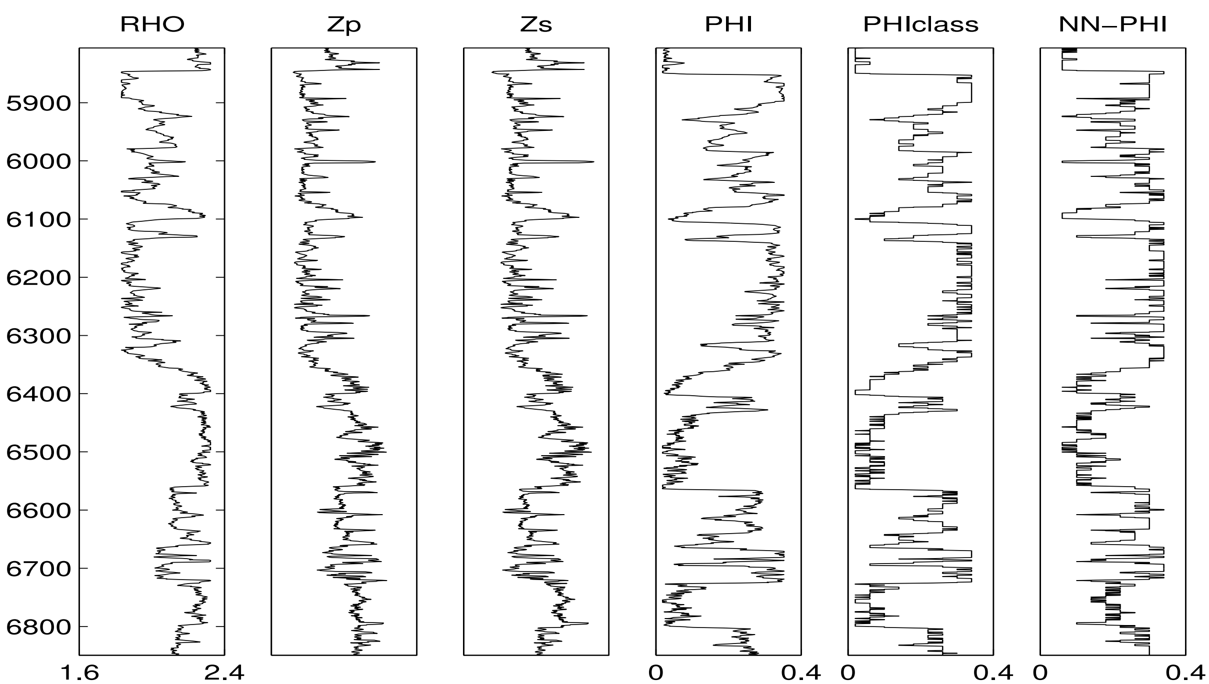

FIG. 4. Interval porosity prediction from relative fluctuations of density and impedances logs. Test data from well 55a. Vertical coordinates are in measured depth.

Porosity prediction from seismic inversion

Synthetic seismic traces were generated using an elastic 1D model built from the log data upscaled at seismic scale. The seismic data were then inverted for the relative fluctuations of ρ, V_p, and V_s. The modeling-inversion experiment is described in (Ecoublet, Symes and Levin, 1997) in the section entitled "Fluid properties predictions from seismic inversion."

The impedance logs are then constructed from the seismic inversion results and used to predict porosity at seismic scale. A number of tests have been conducted to estimate which value for σ produces the smallest porosity prediction error. Once again the best prediction was achieved from using a diagonal matrix $\mathbf{\Sigma}$ whose terms are calculated from the data. The standard deviations of inverted density, σ_ρ, P-impedance, σ_{Z_p}, and S-impedance, σ_{Z_s}, are given in Table 6.

standard deviation				
Well	Number of Vectors	σ_ρ	σ_{Z_p}	σ_{Z_s}
58c + 58cst	591	0.0505	0.0812	0.0771
56c	347	0.0448	0.0631	0.0757
55a	426	0.0586	0.0969	0.0876

Table 6. Standard deviation of the relative fluctuations of ρ, Z_p and Z_s resulting from seismic inversion.

The calculated matrix $\mathbf{\Sigma}$ is used to compute the class-conditional probability density $p(\mathbf{X}|\theta_k)$ for classifying the test data from wells 56c and 55a. Additionally, the PNN is also fed with the training data to assess how well porosity can be estimated from the training data using $\mathbf{\Sigma}$, knowing that a perfect fit can be achieved for a small constant σ. The porosity prediction error is summarized in Table 7 for each well.

Well	*max* $P(\theta_k|\mathbf{X})$		*min* $\mathcal{R}_k$	
	rms $\times 10^2$	C_p%	rms $\times 10^2$	C_p%
58c+st	08.74	94.76	05.94	97.48
56c	07.87	95.48	05.40	97.84
55a	09.10	94.85	06.79	95.85

Table 7. Probabilistic neural network. Rms error and prediction coefficient between targeted and predicted interval porosity from relative fluctuations of ρ, Z_p and Z_s resulting from seismic inversion. The smoothing coefficients $\mathbf{\Sigma}$ are computed from the logs.

The predicted porosity for the training data and the test wells 56c and 55a are plotted in Figures 5, 6, and 7 respectively.

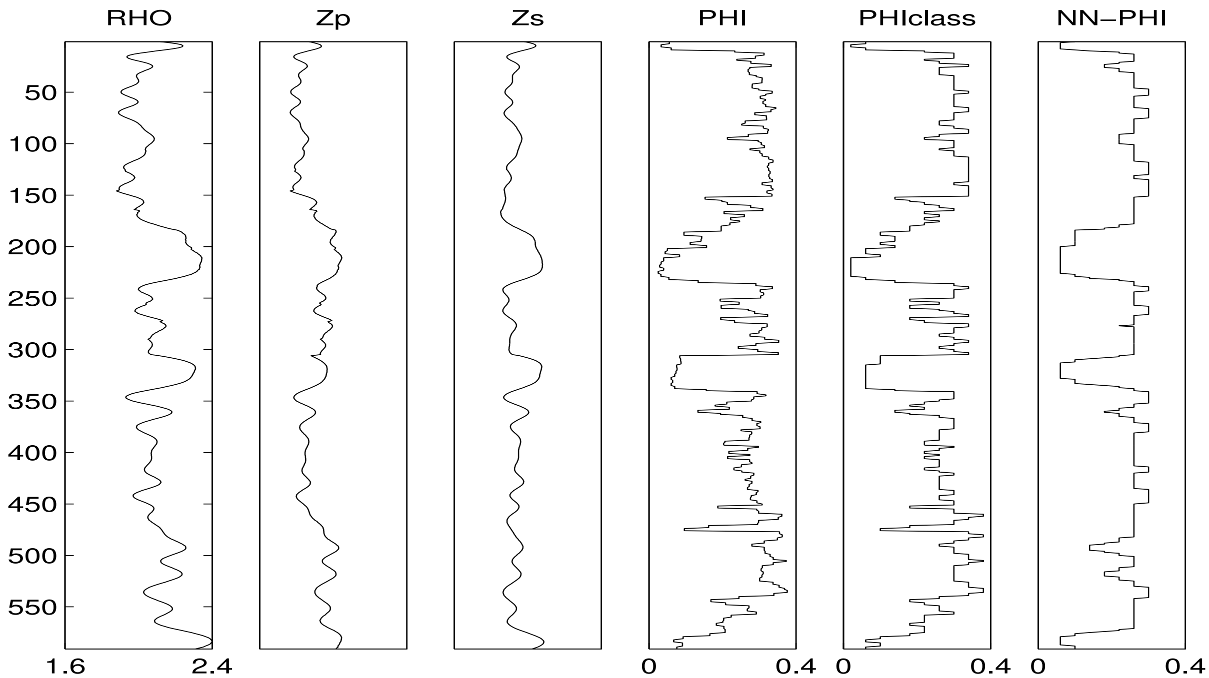

FIG. 5. Probabilistic neural network. Interval porosity prediction from relative fluctuations of density and impedances resulting from seismic inversion. Training data from wells 58c and 58cst. Porosity is classified into 10 intervals. *PHI* is the actual log, *class-PHI* the targeted interval porosity log represented by its mean value, and *NN-PHI* the predicted interval porosity. Vertical coordinates are in depth samples. P-impedance ranges within [0.01,0.03] and S-impedance within [0.005,0.015].

Performance of probabilistic neural network compared with backpropagation neural network

The backpropagation neural network (BPNN) described in (Ecoublet, Symes and Levin, 1997) was used to predict porosity from seismic inversion in order to compare the accuracy achieved from using the different networks. It has been demonstrated for various applications (Specht, 1990b) that classification accuracy can be reached equally well using either PNN or BPNN.

Our BPNN consists of three layers. The input layer has three neurons associated to relative fluctuations of ρ, Z_p and Z_s. The hidden layer is composed of five neurons. In its prediction version, for which porosity prediction error is given in Table 8, the output layer has one neuron producing the relative fluctuation of ϕ.

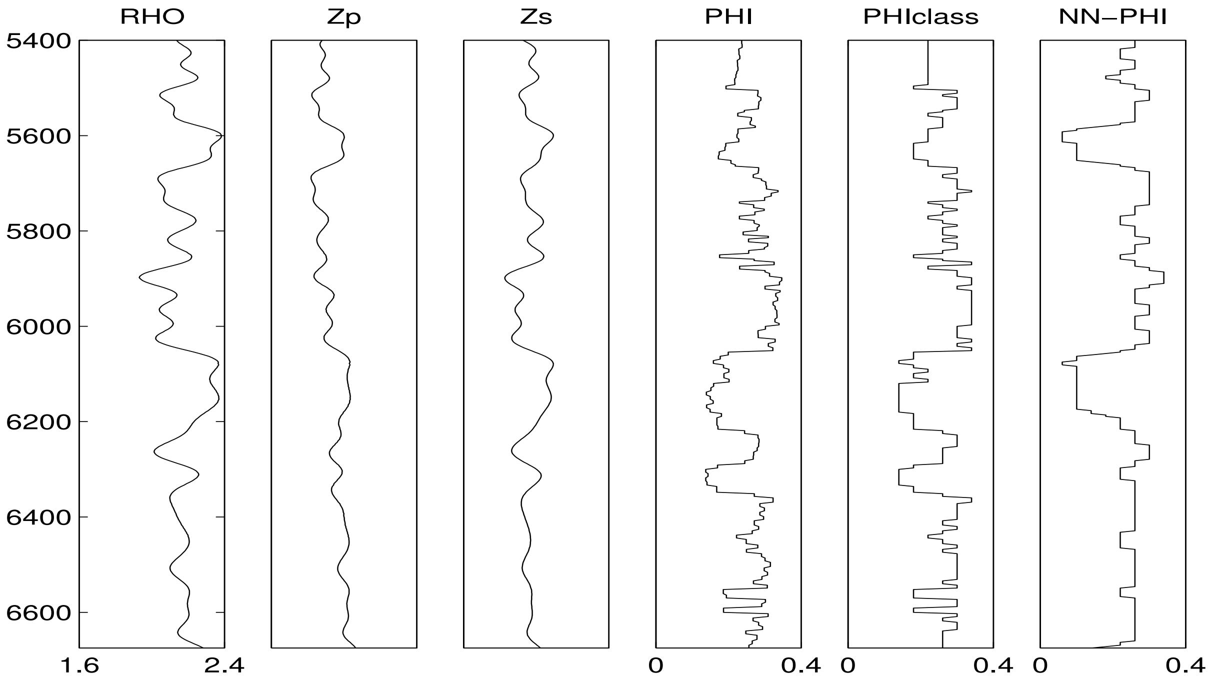

FIG. 6. Probabilistic neural network. Interval porosity prediction from relative fluctuations of density and impedances resulting from seismic inversion. Test data from well 56c. Vertical coordinates are in measured depth.

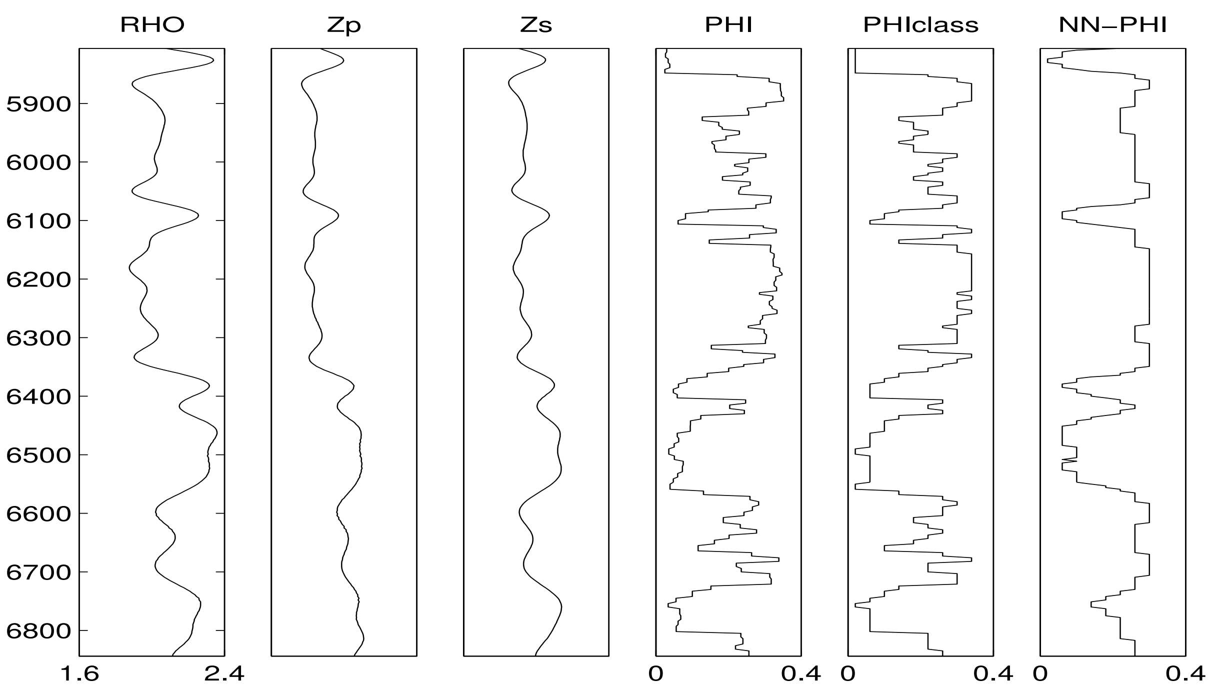

FIG. 7. Probabilistic neural network. Interval porosity prediction from relative fluctuations of density and impedances resulting from seismic inversion. Test data from well 55a. Vertical coordinates are in measured depth.

Well	No. of samples	ϕ	
		rms $\times 10^2$	C_p %
58c+st	591	05.56	97.69
56c	426	05.09	98.05
55a	347	08.22	95.00

Table 8. Backpropagation neural network for prediction. Rms error and prediction coefficient between targeted and predicted porosity relative fluctuations of ρ, Z_p and Z_s.

The BPNN for classification has ten neurons in its output layer, one for each of the porosity intervals as previously detailed for the PNN. The porosity prediction error obtained with the BPNN for classification is given in Table 9.

These results need to be compared with the PNN results given in Table 7. The performance of the three neural networks does not significantly differ in terms of residual error. Note that the PNN, compared to the classification version of the BPNN, performs a 15 to 20% reduction in the rms prediction error.

The porosity estimated using the prediction version of the BPNN is plotted in Figures 8, 10 and 12. The porosity predicted using the classification version of the BPNN is shown in Figures 9, 11 and 13. These results can be compared to those obtained using the PNN by refering to Figures 5, 6 and 7.

This comparison reveals that better resolution achieved with the PNN. Most of the porosity features resolved by the classification version of the BPNN are more accurately recovered by the PNN. The BPNN, in its prediction version, does not recover the amplitude as well as the classification version and PNN networks do. Note for example, that the prediction version of the BPNN produced a negative porosity value at the top of well 55a as seen in Figure 12.

Well	No. of samples	ϕ	
		rms $\times 10^2$	C_p %
58c+st	591	05.12	98.11
56c	426	06.27	97.09
55a	347	08.77	93.93

Table 9. Backpropagation neural network for classification. Rms error and prediction coefficient between targeted and predicted porosity from training on relative fluctuations of ρ, Z_p and Z_s.

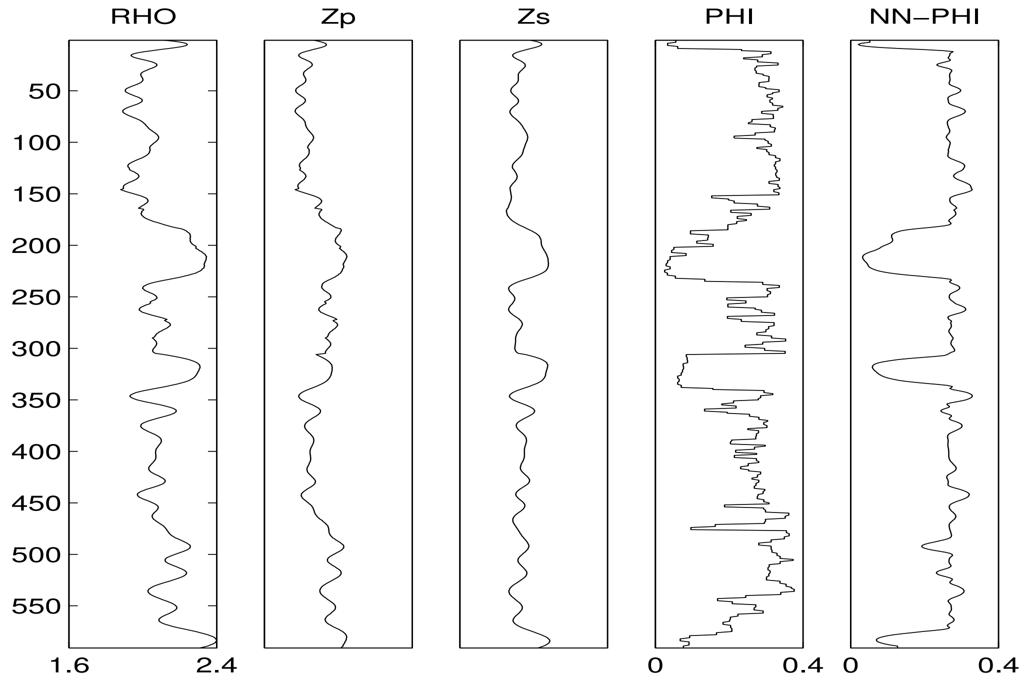

FIG. 8. Backpropagation neural network for prediction. Interval porosity prediction from relative fluctuations of density and impedances resulting from seismic inversion. Training data from wells 58c and 58cst. Vertical coordinates are in depth samples.

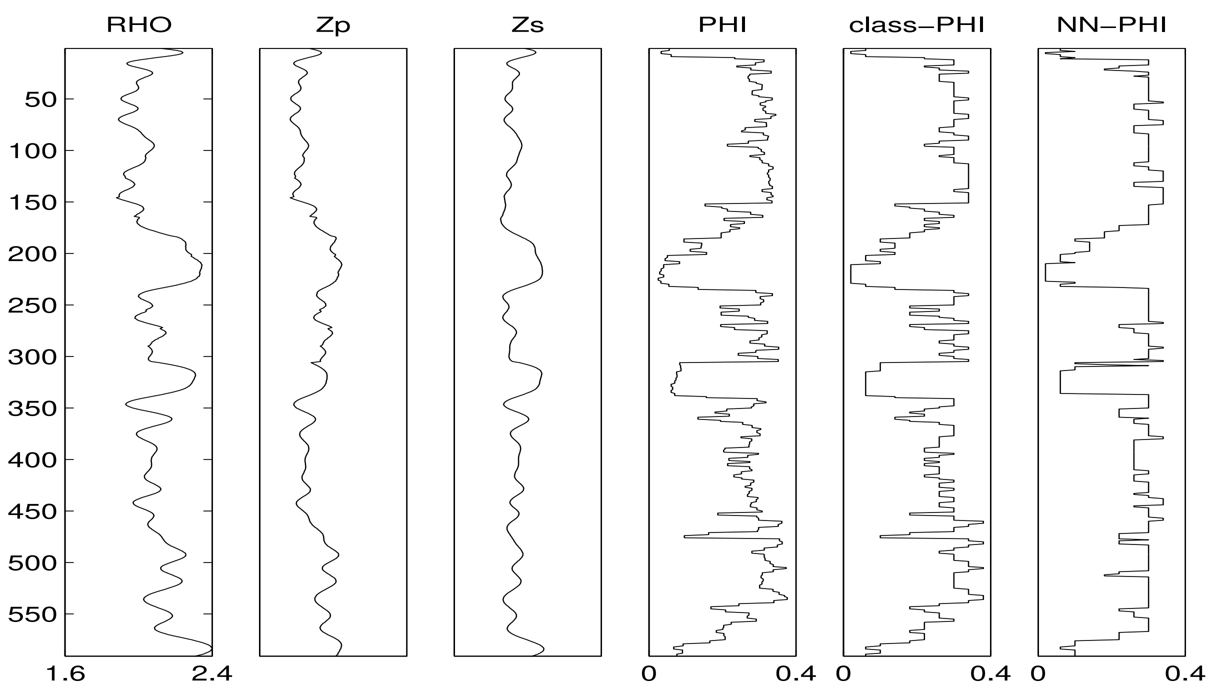

FIG. 9. Backpropagation neural network for classification. Interval porosity prediction from relative fluctuations of density and impedances resulting from seismic inversion. Training data from wells 58c and 58cst. Porosity is classified into 10 intervals.

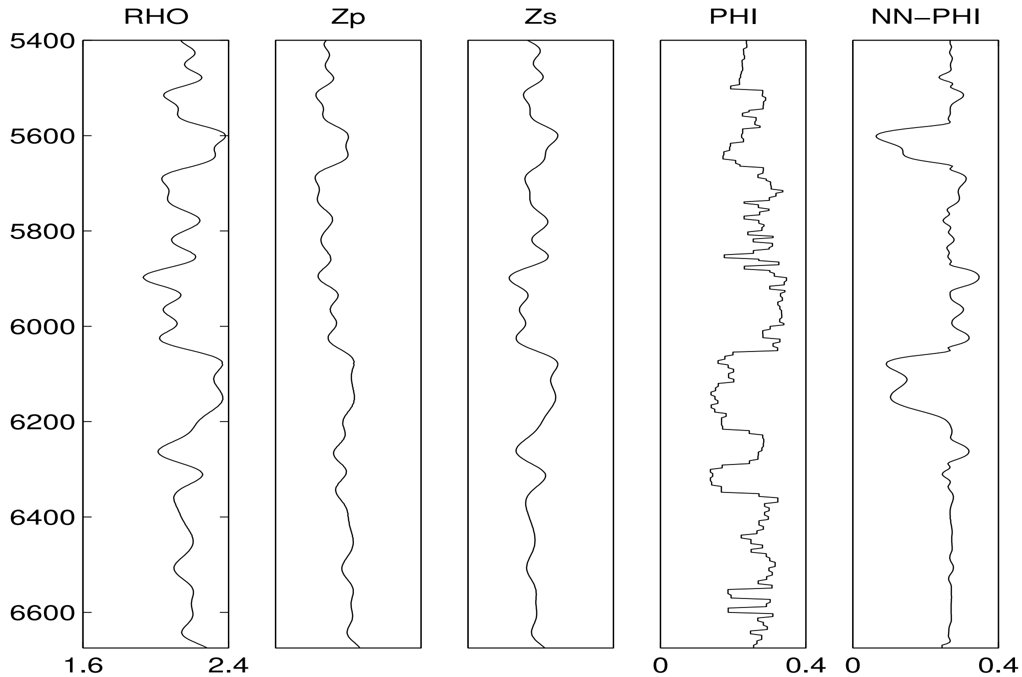

FIG. 10. Backpropagation neural network for prediction. Interval porosity prediction from relative fluctuations of density and impedances resulting from seismic inversion. Test data from well 56c. Vertical coordinates are in measured depth.

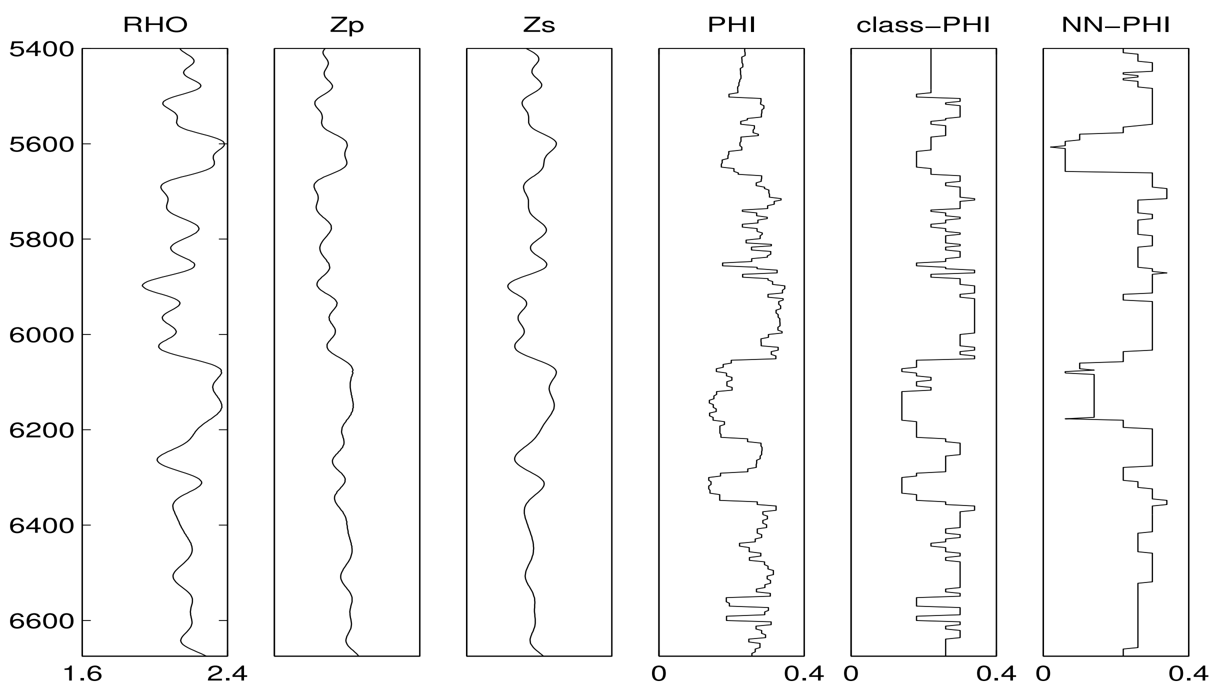

FIG. 11. Backpropagation neural network for classification. Interval porosity prediction from relative fluctuations of density and impedances resulting from seismic inversion. Test data from well 56c. Vertical coordinates are in measured depth.

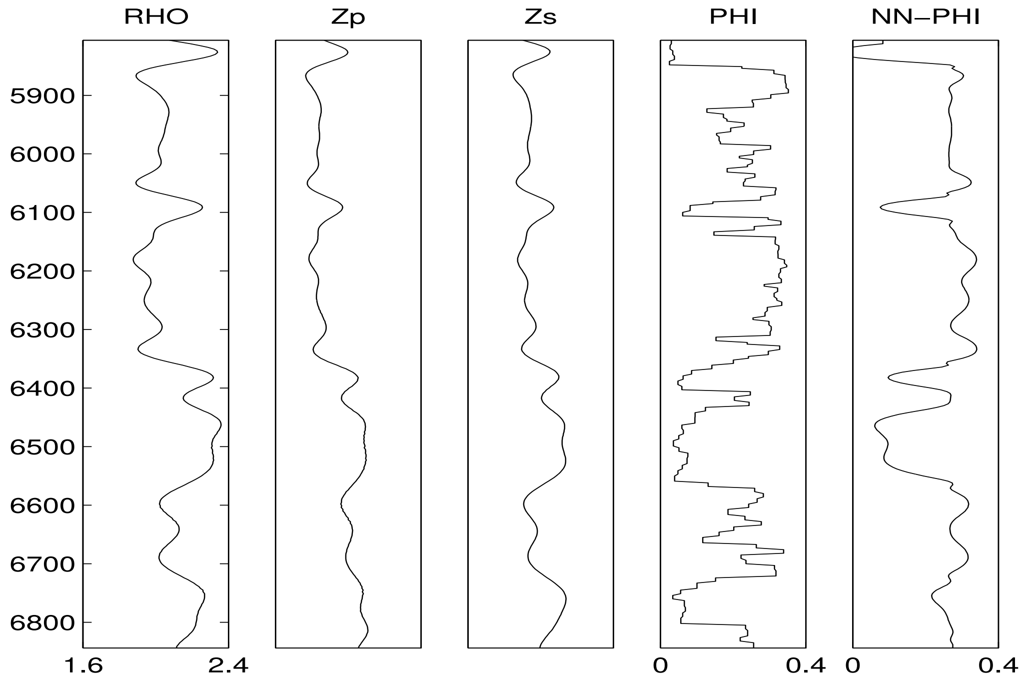

FIG. 12. Backpropagation neural network for prediction. Interval porosity prediction from relative fluctuations of density and impedances resulting from seismic inversion. Test data from well 55a. Vertical coordinates are in measured depth.

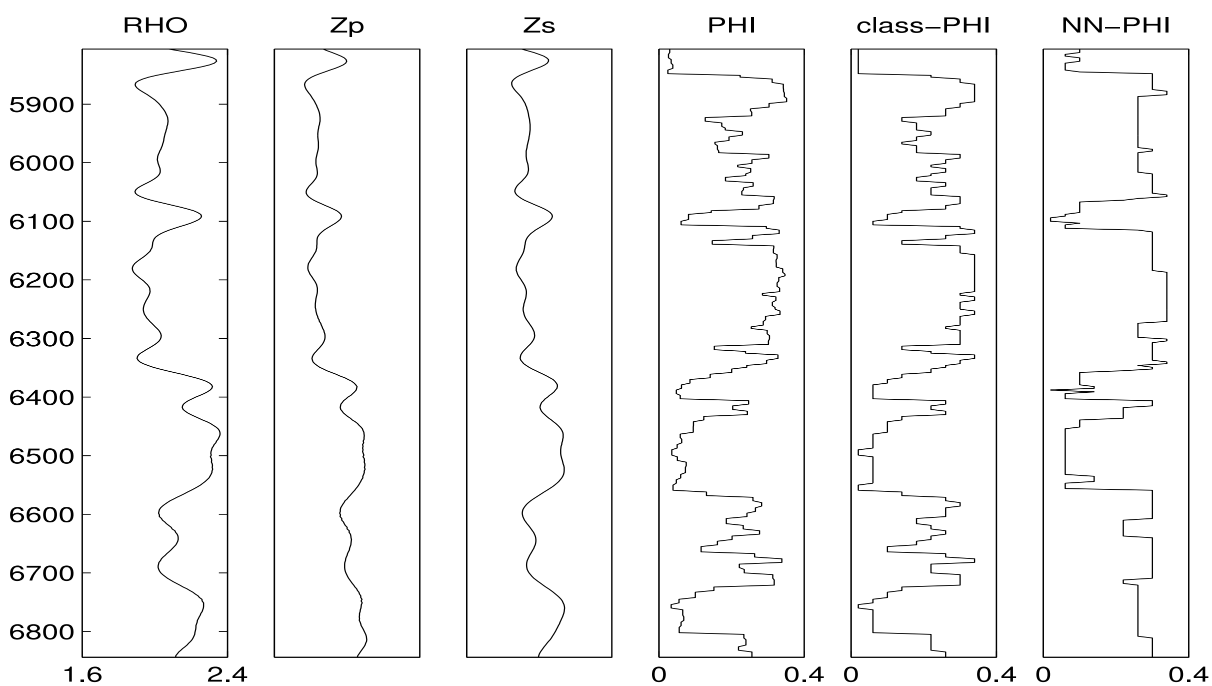

FIG. 13. Backpropagation neural network for classification. Interval porosity prediction from relative fluctuations of density and impedances resulting from seismic inversion. Test data from well 55a. Vertical coordinates are in measured depth.

Conclusion

The results presented in this report on porosity prediction from seismic inversion demonstrate that performance of the PNN is superior to those of the BPNN. Although residual errors varies from well to well, the porosity mapping is better resolved by the PNN.

The PNN has major advantages over the BPNN. Its nonparametric approach makes the decision entirely based on statistical description of the data itself, whereas the capability of the heuristic BPNN to train and generalize depends on how the network parameters are set. The decision boundary, i.e. the shape of the probability density, can be modified anytime through the addition of training vectors. Rock physics may be introduced by adding synthetic training vectors to the existing training data set.

Another major advantage of the PNN is that the PNN generalizes without having to go through the lengthy training process of the BPNN, which is susceptible to convergence toward local minima. On the other hand, generalization may become costly for a large training data set, since all training vectors must be stored and used to classify new vectors. One alternative is to compute an approximation of the exponential form or use a different kernel function to determine the underlying probability density. Different kernel functions, including the Parzen window (Parzen, 1962) have been tested, but the best results so far remain those achieved using a Gaussian kernel.

Acknowledgement and Disclaimer: This work was facilitated by a gift from the Strategic Research Center of Mobil Technology Company. The development of the Hilbert Class Library, used in this project, was partially supported by the National Science Foundation under grant numbers DMS 9404283 and DMS 9627355; the Office of Naval Research under grant number N00014-96-1-0156; the Air Force Office of Scientific Research, Air Force Materials Command, USAF, under grant number F49620-95-1-0339; and The Rice Inversion Project. TRIP Sponsors for 1997 are Advance Geophysical, Amoco Production Co., Conoco Inc., Cray Research Inc., Discovery Bay, Exxon Production Research Co., Interactive Network Technologies, Mobil Research and Development Corp., and Shell International Research. The US Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research, the Office of Naval Research, the National Science Foundation, or the US Government.

REFERENCES

Bishop, C. M., 1996, *Neural Networks for Pattern Recognition*: Oxford, Clarendon Press.

Ecoublet, P., Symes, W. W., and Levin, S., 1997, Seismic Inversion for Porosity using a Backpropagation Neural Network: *TRIP 97*, Rice University.

Parzen, E., 1962, On estimation of a probabilistic density function and mode: *Annals of Mathematical Statistics*, 33:1065-1076.

Specht, D. F., 1990a, Probabilistic Neural Networks: *Neural Networks*, 3:109-118.

Specht, D. F., 1990b, Applications of Probabilistic Neural Networks: *SPIE Applications of Artificial Neural Networks*, 1294:344-353.

Specht D. F., and Shapiro, P. D., 1991, Generalization accuracy of probabilistic neural networks compared with back-propagation networks: *Proc. IEEE International Conference on Neural Networks*, I:525-532.