Supervised seismic facies classification using Probabilistic Neural Networks: Which attributes should the interpreter use?

David Lubo-Robles, Thang Ha, S. Lakshmivarahan, and Kurt J. Marfurt, The University of Oklahoma*

Summary

In the last decade, machine learning algorithms such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Self-organizing Maps (SOM) have been adopted by geoscientists both to extract more detailed information and to accelerate the interpretation of their data. In this study, we present a novel technique called Exhaustive PNN which uses Probabilistic Neural Networks to determine the best suite of seismic attributes to perform a supervised seismic facies classification to differentiate salt from the background seismic response in a Eugene Island seismic survey, offshore Louisiana.

Introduction

Working with seismic attributes and well log data, Hampson et al. (2001) showed stepwise linear regression to be an effective means of picking the best number and collection of seismic attributes to predict a desire well log property. There are two limitations to this workflow, (1) it does not test all the possible combinations of seismic attributes, and (2) by searching for linear relationships between the attributes, it can miss non-linear relationships existing between them. However, for our problem, the limitation is slightly different; rather than predict a specific petrophysical property by *correlating* attributes to that property, we wish to *differentiat*e one or more seismic facies from each other and the background pattern.

Because of artificial neural network's capacity in exploring non-linear relationships, we evaluate whether a novel technique called Exhaustive PNN which, using Probabilistic Neural Networks coupled with an exhaustive search algorithm, are capable of determining the best combination of seismic attributes to distinguish between salt vs. non-salt seismic facies in a Gulf of Mexico 3D seismic survey.

Probabilistic Neural Networks (PNNs)

Probabilistic Neural Networks (PNNs) are feedforward neural networks that using Bayes's criteria and Parzen windows estimate the probability density function from random samples, and then classify an unknown variable into a certain class (Specht, 1995; Masters, 1995; Hajmeer and Basheer, 2002). According to Masters (1995), the most common kernel function used in the Parzen method is the Gaussian function. For this reason, given a set of training attributes **a**, the average estimated density function g(**a**) is given by:

$$\boldsymbol{g}(\mathbf{a}) = \frac{\mathbf{1}}{\boldsymbol{N}}\sum_{\boldsymbol{n=1}}^{\boldsymbol{N}} \boldsymbol{e}^{-\sum_{m=1}^{M}\frac{(\boldsymbol{b_m}-\boldsymbol{a_{nm}})^2}{\boldsymbol{2\sigma^2}}} \tag{1}$$

where, M is the number of training attributes, N is the number of training samples, **b** are the validation attributes in which we want to determine to what class they belong, and σ is a smoothing parameter that requires optimization.

Following Masters (1995), in order to classify an unknown sample (Figure 1), the PNN starts by computing the distance between the validation attributes and the training attributes, then it inputs that distance into the Gaussian activation function. In the summation layer, it computes the average estimated density function g(**a**) for each class. Finally, in the output layer, the PNN decides to what class the unknown sample belongs based on what g(**a**) is maximum. Also, PNN provides confidence estimates of the classification (Masters, 1995) given by:

$$\mathrm{P[A|B]} = \frac{\mathrm{g_A}(B)}{\sum_J \mathrm{g}_j(B)} \tag{2}$$

where, P is the probability of an observation B be the product of class A, g_A(**a**) is the estimated density function for class A, and J represent the total number of classes.

Finally, in order to optimize σ, we use an exhaustive search algorithm in which we test a range of values for σ, and select the one associated with the minimum error E given by:

$$E = \sum_{\mathrm{r=1}}^{\mathrm{R}} \mathrm{e(r)} \Big/ R \tag{3}$$

where, R is the number of validation samples, and e is 1 if the validation sample was not classified correctly, or 0 if the validation sample was classified correctly.

Exhaustive PNN workflow for attribute selection and supervised seismic facies classification

First, we use our geological insight to choose a suite of candidate seismic attributes. Next, we apply 3D Kuwahara median filter to smooth and block the attributes (Qi et al., 2016), preconditioning them for subsequent classification. We also define a group of polygons for each facies which represent the training and validation datasets (Figure 2). We compute the means $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$ from the training data which are used to Z-score normalize the and validation sets in order to avoid any bias related to different

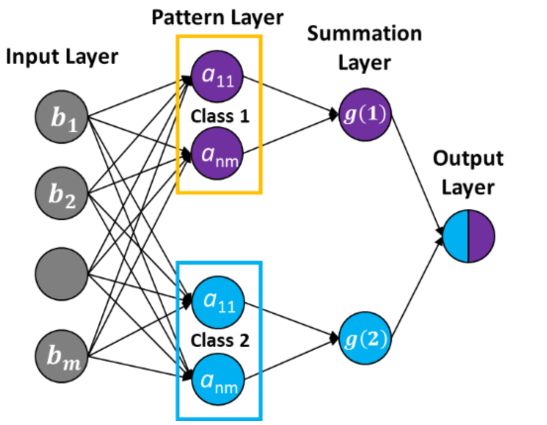

Figure 1. Basic framework for the Probabilistic Neural Network (PNN).

units between the seismic attributes. In order to initialize the PNN Exhaustive algorithm, we define our first seismic attribute combination and an initial smoothing parameter σ. We test a suite of values ranging from 0.1 to 15 through the exhaustive search algorithm, and we compute the error E in each iteration. We store the iteration associated with the minimum E and define another seismic attribute combination. After testing all possible combination, we select the combination of seismic attributes and smoothing parameter σ that provides the minimum E in order to perform our supervised seismic facies classification, and compute the probability of each classes measuring the confidence in the classification.

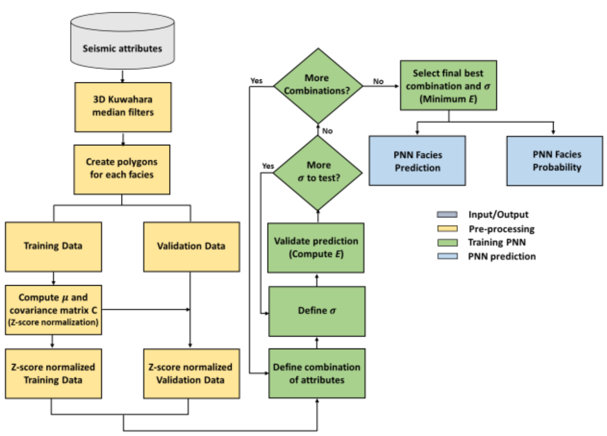

Figure 2. The proposed exhaustive PNN-based workflow for attribute selection and supervised seismic facies classification.

Geologic setting

The Eugene Island mini-basin is a giant Plio-Pleistocene oil and gas field located offshore Louisiana in the outer continental shelf of the Gulf of Mexico (Alexander and Flemings, 1995; Joshi and Appold, 2016). Following, Alexander and Flemings (1995), the mini-basin had three phases of evolution.

A prodelta phase characterized by deposition of bathyal and prodelta shales, distal deltaic sands, and turbidites on top of a salt sheet. This sediment loading caused the salt sheet to migrate outward laterally, creating new accommodation space in the area (Alexander and Flemings 1995; Joshi and Appold, 2016). Second, a proximal deltaic phase associated with lowstand deltas characterized by mud and sand sequences (Joshi and Appold, 2016). Also, a high sediment accumulation rate is seen during this phase due to salt withdrawal (Alexander and Flemings, 1995). Finally, the last phase is fluvial in which salt withdrawal waned, so little further accommodation space was created (Alexander and Flemings, 1995; Joshi and Appold, 2016).

Data description

The Eugene Island seismic survey is located in the Gulf of Mexico, offshore Louisiana. For this study, the seismic volume was cropped consisting of 700 inlines and 700 crosslines with a bin size of 82.5 ft by 82.5 ft, record length of 3 s, and an area of approximately 306 km^2. Figure 3 shows a representative time slice at t=2 s through the seismic amplitude volume, where we observe two salt bodies (Salt #1 and Salt #2; red arrows) associated with salt withdrawal during the prodelta and proximal deltaic phases of deposition.

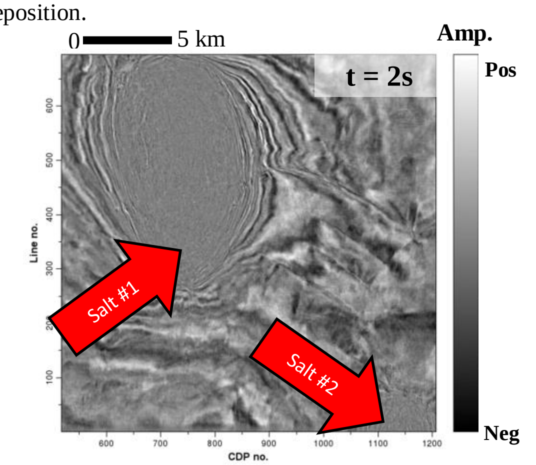

Figure 3. Time slice at t=2 s through the Eugene Island seismic amplitude volume, offshore Louisiana. Note the presence of two salt bodies (red arrows) and intervening minibasins associated with salt withdrawal in the study area.

Seismic attribute selection to discriminate salt from the background seismic pattern using Exhaustive PNN

To perform our supervised seismic facies classification, we

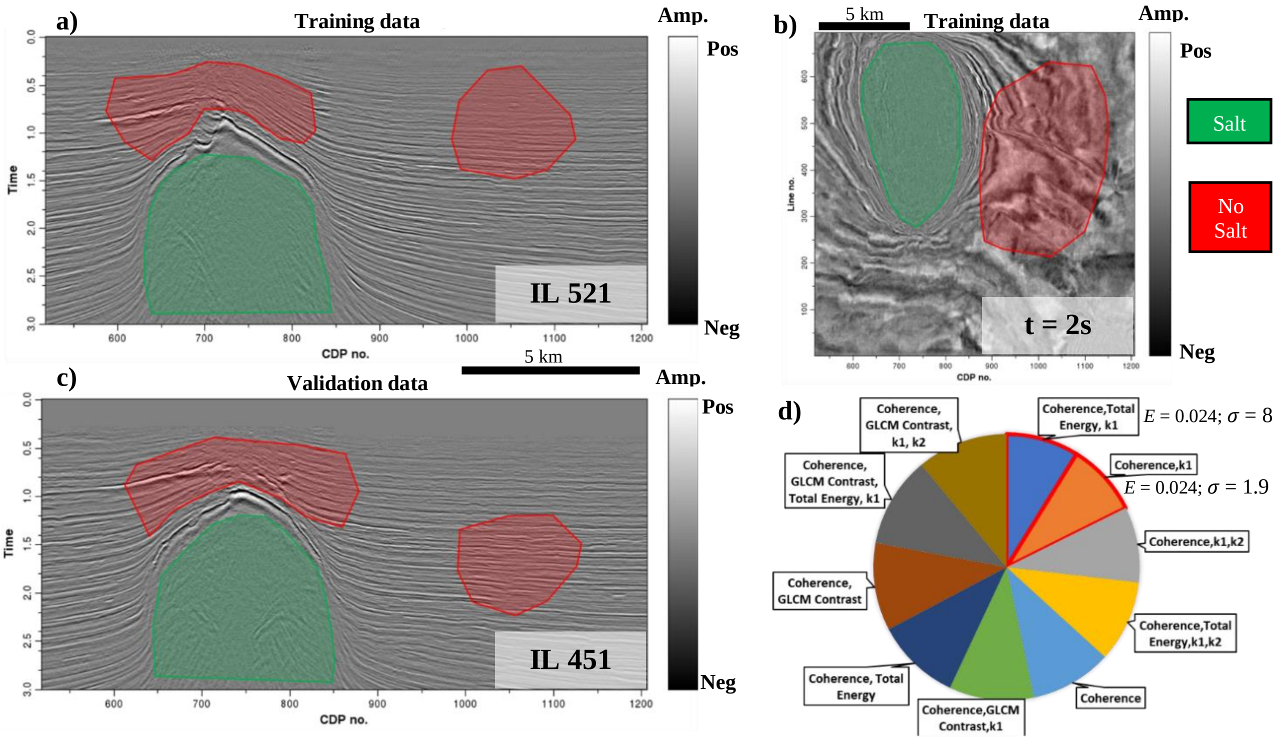

Figure 4. Training and validation datasets. (a) Vertical slice though the seismic amplitude along inline 521. We extract the training voxels from the salt (green polygons) and non-salt (red polygon) from the six seismic attributes. (b) Time slice at t=2 s through the seismic amplitude volume. This time slice is also used to generate training voxels used for the training set. (c) Vertical slice through the seismic amplitude volume along inline 451. We also extract voxels from the six seismic attributes for each class. However, these voxels are used as validation set in order to train our neural network by minimizing the prediction error E. (d) For six attributes, there are 63 possible combinations. We show the ten best combinations of seismic attributes computed from the Exhaustive PNN selected based on their associated E. The best combination is given by coherence and most-positive (k_1) curvature because it provides the minimum error using less attributes, thus is more computationally efficient.

evaluate six candidate seismic attributes based on our geologic insight and past experience: coherence, GLCM contrast, GLCM dissimilarity, total energy, most-positive curvature (k_1), and most-negative curvature (k_2). These six candidate attributes serve as input to our Exhaustive PNN algorithm with the goal of finding the best combinations of seismic attributes, and the ideal parameter σ to differentiate between salt vs. non-salt facies in the Eugene Island seismic volume.

To generate our training and validation sets, we pick a suite of polygons enclosing the salt and non-salt facies. We use inline 521 (Figure 4a) and time slice at t=2 s (Figure 4b) to extract the training voxels of the salt (green polygon) and non-salt (red polygons) seismic facies from the six seismic attributes used as input in the Exhaustive PNN workflow. We perform a similar analysis in inline 451 (Figure 4c), but in this case the extracted voxels from the seismic attributes enclosing the two target seismic facies are used as validation samples for training the neural network by minimizing the error E. Also, we only extract our training and validation sets from the Salt #1 diapir in order to leave the Salt #2 diapir as a blind test to further evaluate the performance of the PNN.

We maintain a similar number of voxels for each class in order to avoid any type of bias towards one of the facies.

When running the Exhaustive PNN algorithm, we test 63 different combinations between the training attributes, then we store the ideal smoothing parameter σ associated with the minimum E for each combination. In Figure 4d, we show the best 10 combinations of seismic attributes obtained from our proposed workflow selected based on their associated E. We observe that the best performance is obtained when classifying between salt and non-salt seismic facies using the coherence, total energy, most-positive (k_1) curvature, or coherence and most-positive (k_1) curvature, both with an E of 0.024 and smoothing parameter σ of 8 and 1.9 respectively. To perform our seismic facies classification, we use the coherence, and most-positive (k_1) curvature combination because it provides the minimum error using less seismic attributes, so it will be more computationally efficient.

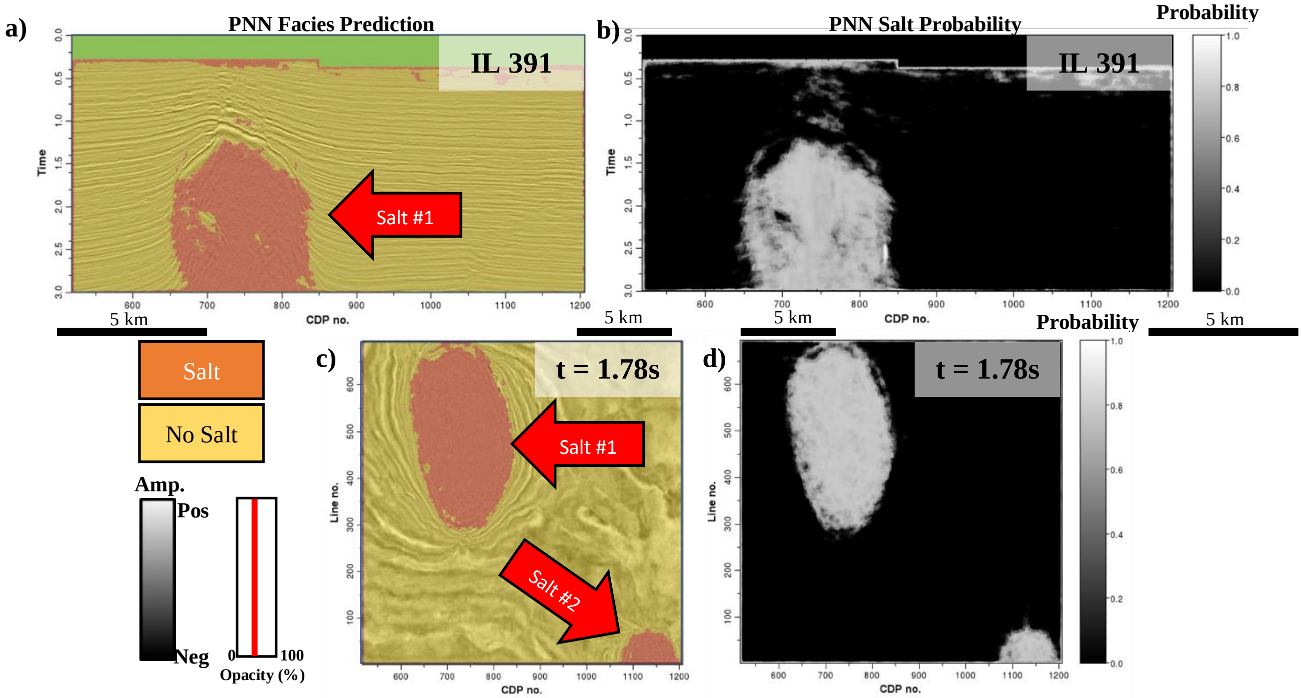

Figure 5. (a) PNN Facies Prediction co-rendered with the seismic amplitudes along inline 391. We note that the PNN correctly distinguishes between salt (red arrows) and non-salt seismic facies. (b) PNN Salt Probability along inline 391. We observe that the probability of the extracted orange facies of being salt ranges from 75 to 80%. (c) PNN Facies Prediction co-rendered with seismic amplitudes at time slice 1.78s. We note that the PNN still classifies the Salt#1 diapir correctly. Moreover, the Salt #2 diapir, that was left as blind test, is also classified correctly by the algorithm. (d) PNN Salt Probability at time slice 1.78 s. Salt #1 and Salt #2 diapirs show high probabilities ranging from 75 to 80%.

In Figure 5, we show the results obtained after applying the Exhaustive PNN in all the Eugene Island seismic survey using the coherence and most-positive (k_1) curvature attributes and σ of 1.9. Analyzing the PNN facies prediction co-rendered with the seismic amplitude along inline 391 (Figure 5a), we note that the neural network classifies correctly between the salt (red arrow) and non-salt seismic facies. Also, we are able to compute the PNN Salt probability volume which provide the confidence of the classification (Figure 5b). We observe that the algorithm classifies the extracted orange facies as salt with very high probabilities ranging from 75 to 80%. It is important to highlight that some orange facies visible on the top of the seismic volume are associated with missing or noisy data in the edges of the survey with little interpretational value.

In Figure 5c, we show the PNN Facies prediction co-rendered with the seismic amplitude volume at time slice 1.78 s. We observe that the Salt #1 (red arrow) diapir is still correctly classified by the Exhaustive PNN algorithm. Moreover, the Salt #2 diapir used as a blind test during the training of our neural network, is also correctly classified by our algorithm as a salt facies (red arrow). Finally, Figure 5d shows that both Salt #1 and Salt #2 diapirs (red arrows) show a high probability of being salt ranging from 75 to 80%.

Conclusions and future work

In this application Exhaustive PNN proved to be a powerful tool in determining the best combination of seismic attributes in order to perform a supervised seismic facies classification to differentiate between salt vs. non-salt seismic facies in the Eugene Island dataset. We determine that the best combination is given by using coherence and most-positive (k_1) curvature as training seismic attributes with a smoothing parameter σ of 1.9. For future work, we will improve the performance of our model by implementing a Gradient Descent algorithm in order to compute a different σ for each training attribute. Also, we will perform a geobody extraction to obtain a 3D distribution of the salt seismic facies along the Eugene Island seismic volume.

Acknowledgments

We thank the U.S. National Archive of Marine Seismic Surveys (NAMSS), and the U.S. Geological Survey (USGS) for providing the Eugene Island dataset. Also, we would like to thank the sponsors of the Attribute Assisted Seismic Processing and Interpretation (AASPI) consortium for their support. Finally, we thank Heather Bedle for her valuable comments.

Supervised seismic facies classification using Probabilistic Neural Networks: Which attributes should the interpreter use?

References

Alexander, F., and P. Flemings, 1995, Geologic Evolution of a Pliocene–Pleistocene Salt-Withdrawal Minibasin: Eugene Island Block 330, Offshore Louisiana: AAPG Bulletin, **79**, no. 12, 1737–1756.

Hajmeer M., and I. Basheer, 2002, A probabilistic neural network approach for modeling and classification of bacterial growth/no-growth data: Journal of Microbiological Methods, **51**, 217-226.

Hampson, D.P., Schuelke, J.S., Quirein, J.A., 2001, Use of multiattribute transforms to predict log properties from seismic data. Geophysics, **66**, no. 1, 220- 236.

Joshi, A., and M. Appold, 2016, Potential of porosity waves for methane transport in the Eugene Island field of the Gulf of Mexico basin: Marine and Petroleum Geology, **75**, 1-13.

Masters, T., 1995, Advanced Algorithms for Neural Networks.

Qi, J., T. Lin, T. Zhao, F. Li, and K. J. Marfurt, 2016, Semisupervised multiattribute seismic facies analysis: Interpretation, **4**, no. 1, SB91-SB106, doi: 10.1190/INT-2015-0098.1

Specht, D. F., 1995, Probabilistic Neural Networks and General Regression Neural Networks: Fuzzy Logic and Neural Network Handbook edited by C.H. Chen.