The Pennsylvania State University

The Graduate School

College of Earth and Mineral Sciences

POROSITY PREDICTION FROM SEISMIC DATA USING MULTIATTRIBUTE TRANSFORMATIONS, N SAND, AUGER FIELD, GULF OF MEXICO

A Thesis in
Geosciences

by

Joseph Christian Adam Frank Valenti

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

December 2009

The thesis of Joseph Christian Adam Frank Valenti was reviewed and approved* by the following:

Charles J. Ammon
Professor of Geosciences
Thesis Advisor

Turgay Ertekin
Professor of Petroleum and Natural Gas Engineering and George E. Trimble Chair in Earth and Mineral Sciences

Andy Nyblade
Professor
Department of Geosciences

Kate Freeman
Professor of Geosciences
Graduate Program Chair

*Signatures on file in the Graduate School

ABSTRACT

We compare two methods of predicting well-log porosity from seismic data. The data consist of a suite of well logs and a full stack 3D seismic survey over Auger Field in the deepwater Gulf of Mexico. The 3D seismic is transformed into a number of attribute volumes. These attributes are combined in a nonlinear manner, via an Artificial Neural Network (ANN), or in a linear manner, via multilinear regression analysis, in order to predict the target porosity logs from the available suite of field data.

A feed-forward back propagation ANN is trained using the seismic attributes as an input set and with the porosity logs as the output set. The linear mode uses the same training data, but derives a series of weights which when applied to the input set minimize the differences in a least-squares sense between the target and predicted outputs.

In order to measure the accuracy of the attribute to porosity transformation, cross-well validation was performed. In this procedure one well is removed from the training set and the transformation is re-derived. The accuracy of the transformation in predicting the log from the removed well is then measured. This is done to every well in the training set so that we may determine a reasonable expectation for the performance of the transformation.

We see a marked improvement of the performance of the ANN over that of the multilinear regression. These results are evident not only in the training data but more importantly also in the testing data.

TABLE OF CONTENTS

LIST OF FIGURES

value is below a threshold value, in this case 63 GAPI, the lithology is interpreted to be sandstone. This is shown by shading yellow. Above 63 GAPI, the rock is interpreted to be shale. The fluid type, either water or gas, saturating the rock is determined from the resistivity log **(b)**. In this well, if the resistivity is above 2 OHMM, shaded red, the fluid is interpreted to be gas, with a density of 0.42 g/cc. If the resistivity is below the threshold, the fluid saturant is interpreted to be water, with density of 1.1 g/cc. For each sample, this information is combined with the density log **(c)** to create the porosity log **(d)**. We determined that the predictive power of the methods being studied is greatly enhanced if a constant porosity value of 0.17 **(e)** is used wherever it a shale has been interpreted to be the rock type. The smoothed log in each case is shown in red. The porosity logs were calculated using the unfiltered logs, however.

Figure 11, Page 33. Architecture of the Feed Forward Network. Seven normalized input attributes are fed into twenty-two neurons in the hidden layer. The one output represents a porosity sample.

Figure 12, Page 34. These attributes **(a-g)** are combined in order to create a synthetic porosity log that resembles as closely as possible the target log at right **(h)**.

Figure 13, Page 35. Flow Chart Describing Training Method.

Figure 14, Page 37. Hinton plot that shows the strength of the weights between the input layer and the hidden layer. The size of the box corresponds to the amplitude of the weight, while the color indicates polarity. A red weight is negative, and a green weight is positive.

Figure 15, Pages 39-41. Test wells from the training sequence of the Neural Network with 1000 Epochs of Training. The CC value indicates the normalized cross-correlation coefficient between the neural network output and the target values.

Figure 16, Pages 41-42. Performance of ANN as a function of the number of wells used in the training set. **a)** Maximum Error Percentage in the test case. Note the asymptotic appearance of the curve. Error does decrease with an increasing number of wells, but the returns are diminishing. **b)** The mean testing error shows the same trend as in part **a)**. **c)** The cross correlation coefficient of the test set increases substantially with the number of wells in the training set. As the network is exposed to a greater variety of input-output pairs, it has more experience on which to call in order to make a predictive decision.

Figure 17, Page 43-44. Results of ANN porosity prediction from a Seismic Line passing through well A04BP1 at Trace 2206. a) InLine showing the Structure of the N_m^1. b) InLine flattened on the N_m^1 horizon. Time "0" corresponds to the peak trough amplitude picked in the seismic volume which was interpreted to represent the N sand.

Figure 18, Pages 46-53. In a 100 ms interval around the N sand, 25 porosity values were predicted by the ANN that correspond to the 25 seismic samples in the interval. **a)** The maximum predicted porosity at each trace location is mapped and overlain onto the

structure contours of the N_m^1. **b)** Porosity map at twelve ms, or 3 seismic samples, above the trough corresponding to the N_m^1. **c)** Porosity map at eight ms, or 2 seismic samples, above the trough corresponding to the N_m^1. **d)** Porosity map at four ms, or one seismic sample, above the trough corresponding to the N_m^1. **e)** Porosity map at the trough corresponding to the N_m^1. **f)** Porosity map at four ms, or one seismic sample, below the trough corresponding to the N_m^1. **g)** Porosity map at eight ms, or 2 seismic samples, below the trough corresponding to the N_m^1. **h)** Porosity map at twelve ms, or 3 seismic samples, below the trough corresponding to the N_m^1.

Figure 19, Pages 56-58. Test wells from the training sequence of the Multivariate Linear Regression. The CC value indicates the normalized cross-correlation coefficient between the neural network output and the target values.

Figure 20, Page 59. Results of the Regression porosity prediction from a Seismic Line passing through well A04BP1 at Trace 2206. a) InLine showing the Structure of the N_m^1. b) InLine flattened on the N_m^1 horizon. Time "0" corresponds to the peak trough amplitude picked in the seismic volume which was interpreted to represent the N sand.

Figure 21, Pages 60 -67. In a 100 ms interval around the N sand, 25 porosity values were predicted by multivariate linear regression that correspond to the 25 seismic samples in the interval. **a)** The maximum predicted porosity at each trace location is mapped and overlain onto the structure contours of the N_m^1. **b)** Porosity map at twelve ms, or 3 seismic samples, above the trough corresponding to the N_m^1. **c)** Porosity map at eight ms, or 2 seismic samples, above the trough corresponding to the N_m^1. **d)** Porosity map at four ms, or one seismic sample, above the trough corresponding to the N_m^1. **e)** Porosity map at the trough corresponding to the N_m^1. **f)** Porosity map at four ms, or one seismic sample, below the trough corresponding to the N_m^1. **g)** Porosity map at eight ms, or 2 seismic samples, below the trough corresponding to the N_m^1. **h)** Porosity map at twelve ms, or 3 seismic samples, below the trough corresponding to the N_m^1.

Figure 22, Pages 68-69. a) Histogram showing the distribution of porosity values in the training set. Note the range of values extends to the upper limit of 0.33. **b)** Histograms showing the distribution of predicted porosity values in maps shown above. The results given by the neural network demonstrate roughly the same distribution as the porosity training set. The results from regression analysis, however, lack the high-porosity information.

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to acknowledge Professor Charles J. Ammon and Professor Turgay Ertekin for their unwavering support during this degree, as well as Professor Andy Nyblade for agreeing to act on my committee. I would also like to thank: Eleanor for putting up with me; Matt for helping me move; Al and Dad for rescuing me from the winter more than once; Mom for the home cooked meals; and the dudes.

INTRODUCTION

Predicting subsurface properties, such as porosity, has always been a fundamental problem for geologists and geophysicists. Seismic data are often used to delineate the structure of reservoir bodies, but are not often used to estimate the spatial distribution of reservoir and rock properties. In this study, we compare two methods of predicting porosity from several 3D seismic attributes.

Schultz et. al., in 1994, were the first to propose the idea of using multiple seismic attributes to predict log properties away from well control [1]. Well-to-seismic ties are considered to be the training points for the method; both the seismic response and the well log property are known at the select well locations. A statistical relationship, either linear or nonlinear, is developed at the well sites to relate the seismic response to the well log response. We describe the use of an Artificial Neural Network (ANN) and Multivariate Linear Regression in determining these statistical relationships.

Geoscientists have been among the pioneers in developing uses for Artificial Neural Networks. In many ways, neural networks are perfectly suited for geostatistical analysis. Many geophysical measurements are related to rock properties in nonlinear ways and with no analytical relationships known as of yet. As computers have increased in computational speed and power, large scale data processing with neural networks has become feasible.

Artificial Neural Networks (ANN) have been used to predict core properties from well logs [2], well log to well log transformations, and have used seismic properties to predict lithology [3-6], sonic logs and shale content [7], shale stringers in a heavy oil

reservoir [8], density and gamma ray logs [8, 9], spontaneous potential [10, 11], permeability [2, 12], and porosity [3, 5, 6, 9, 12-17].

Table 1. Results of Past Studies Using ANN to Predict Porosity from Seismic Attributes				
Author	**CC Test**	**CC Training**	**Mean Testing Error (V/V)**	**Type of Network Used**
Pramanik [13]	0.86	0.95	0.037	PNN
Hampson [18]	0.62	0.95	NA	PNN
Leiphardt [16]	0.62	0.82	0.026	PNN
Dorrington [19]	0.57	0.86	NA	MLFN

Table 1. Results from previous porosity prediction studies using ANN. CC Test is the average cross correlation coefficient between the prediction and target sets in the test case. CC Training is the cross correlation coefficient between the prediction and target sets in the training case. The mean testing error documents the mean error between the prediction and target sets in the test case. Two types of ANN have been used in these previous studies. PNN refers to a Probabilistic Neural Network, while MLFN refers to a Multi-Layer Feed Forward Network.

Multivariate Linear Regression is a simple extension of the well-known univariate case. In these circumstances, log properties are estimated from a linearly weighted sum of a number of seismic attributes. This was first demonstrated to yield accurate results by Russell et. al., in 1997 [17].

Table 2. Results of Past Studies Using Multilinear Regression to Predict Porosity from Seismic Attributes			
Author	**CC Test**	**CC Training**	**Mean Testing Error (V/V)**
Pramanik [13]	0.56	0.66	0.056
Hampson [18]	0.60	0.69	NA
Leiphardt [16]	0.63	0.74	0.026

Table 2. Results from previous porosity prediction studies using a Multilinear Regression of Seismic Attributes. CC Test is the average cross correlation coefficient between the prediction and target sets in the test case. CC Training is the cross correlation coefficient between the prediction and target sets in the training case. The mean testing error documents the mean error between the prediction and target sets in the test case.

There are several advantages of these new algorithms over the conventional inversion methods: they predict logs other than acoustic impedance; they may use seismic attributes other than the conventional post-stack volume; they do not rely on any particular forward model; knowledge of the seismic wavelet is not required; they may enhance resolution; and they use testing and cross-validation as a measure of success [1, 16-18].

Theoretical Background

Feed Forward Back-Propagation Neural Network

An artificial neural network (ANN) creates a nonlinear mapping between a set of input data and target outputs. The nature of this nonlinear mapping depends upon the type of artificial neural network used. This thesis focuses on feed forward back propagation artificial neural networks.

A feed forward back-propagation neural network has an input layer, an output layer and a variable number of hidden layers (Figure 1) [20]. These layers are each composed of a number of neurons, which are each connected to every other neuron of the prior and next layers.

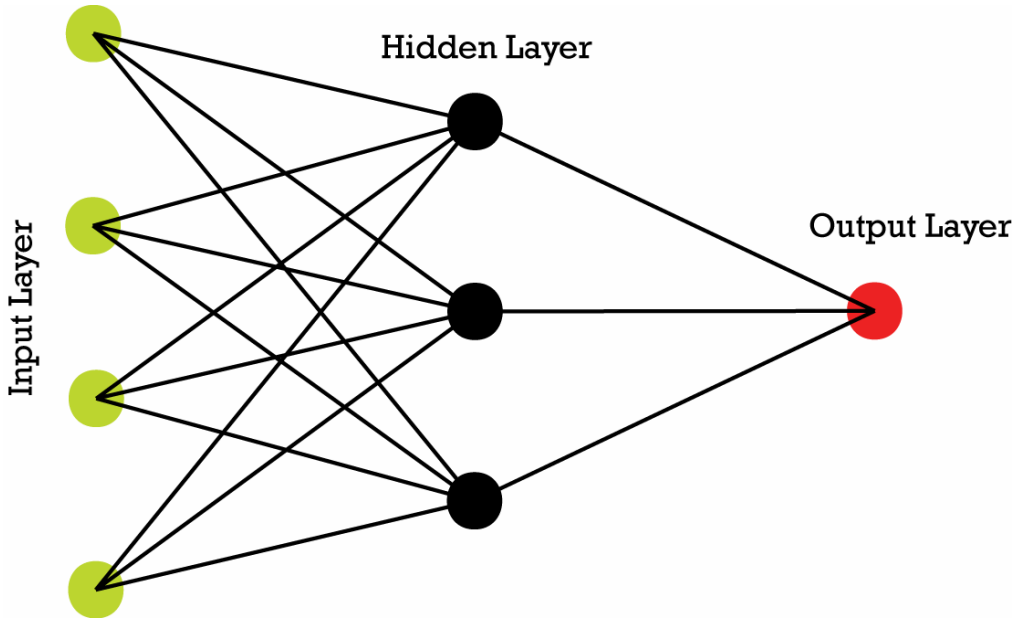


Figure 1. Architecture of a simple Neural Network with one hidden layer. In this scenario, the input layer is O_1, the hidden layer is O_2, and the output layer is O_3. Each circle represents a neuron. The input layer has four neurons, the hidden layer three, and the output one. This particular network would map four inputs to one output.

The neurons in the input layer do not perform any computation, as they are simply the input gate. The neurons in the hidden and output layers have weights and biases connecting them to the neurons in the previous layer [18, 20]. Each neuron sums the weighted and biased input from each neuron in the previous layer and then filters the sum with a transfer function. We assume that the relationship between the inputs and target is nonlinear. We therefore make use of a commonly employed nonlinear transfer function, the hyperbolic tangent sigmoid (Figure 2). The hyperbolic tangent sigmoid transfer function is:

$$\tanh(n) = \frac{e^{n} - e^{-n}}{e^{n} + e^{-n}} \qquad (1)$$

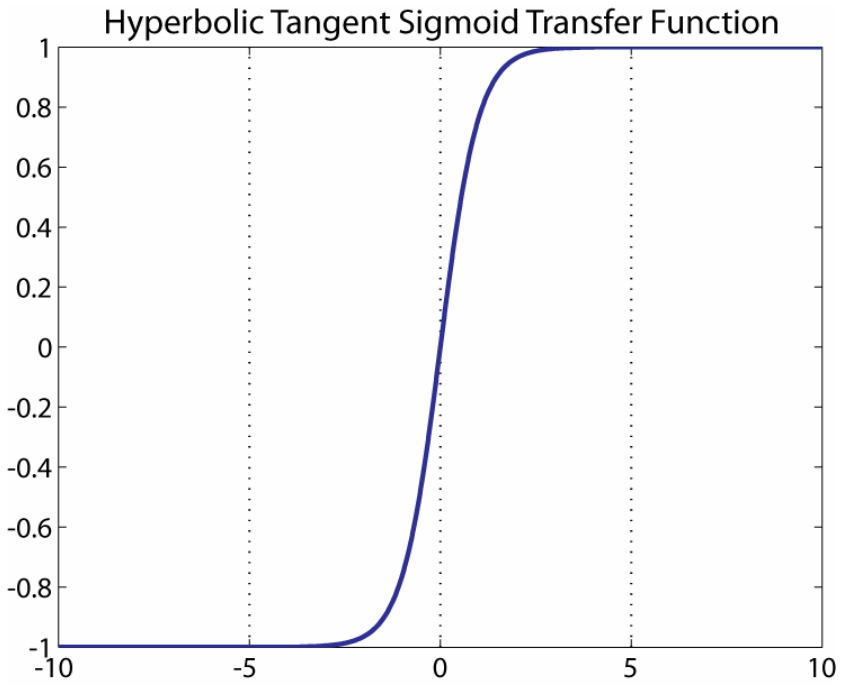


Figure 2. Graphic display of the hyperbolic tangent sigmoid transfer function. Note how quickly the function saturates for values greater than five. To make full use of the shape of this transfer function, the inputs are normalized to the range -1 to 1.

The output layer is the weighted and biased sum from the output of the neurons in the last hidden layer. This sum is not filtered, so the transfer function is described as linear with a slope of one.

If the output of layer m is denoted as O_m, then the output of the input layer is just the input vector, **I**, of length k.

$$O_1 = I \tag{2}$$

For simplicity, and since this applies directly to the network that is to be used in this study, assume that there is only one hidden layer. The output of the j^{th} neuron in this hidden layer, with transfer function f_2 is:

$$O_{2j} = f_2\left(\sum_{i=1}^{k} I_i w_{ij} + b_{ij}\right) \tag{3}$$

Where w_{ij} and b_{ij} are the weights and biases, respectively, that connect the i^{th} neuron in the input layer to the j^{th} neuron in the hidden layer [20].

The final output layer has only one neuron, corresponding to one property prediction. If there are n neurons in the hidden layer, then the final output, which is not filtered by a transfer function, is:

$$O_3 = \sum_{j=1}^{n} O_{2j} w_j + b_j \tag{4}$$

Where, once again, w_j and b_j are the weights and biases which connect the j^{th} neuron in the hidden layer to the single neuron in the output layer.

In order to accurately map the inputs to the output, the network will continuously update the weights and biases which connect each of the neurons until some performance criterion is achieved. This process is called training. Many different training algorithms exist, but error backpropagation is the most popular learning algorithm for multi-layered neural networks [20, 21].

An ANN minimizes the difference between the prediction and targets in any number of specified ways. In backpropagation it is usually the mean squared error between the network output and the targets that is used as the objective function, and is the quantity that is to be minimized. In this case, however, it was demonstrated that using the mean absolute error as the performance criterion yielded much more accurate results [21]. We will now describe backpropagation using mean absolute error, and the gradient descent method with momentum and a variable learning rate is used to train the neural network.

For prediction-target pairs, (p_1, L_1), (p_2, L_2), … (p_N, L_N), we want to minimize the mean absolute error:

$$E = \frac{1}{N} \sum_{j=1}^{N} \left| L_j - p_j \right| \tag{5}$$

This can be accomplished by the gradient descent algorithm. To descend towards the local minimum in a function, one can take steps in the direction of the negative gradient of the function at the current point.

For example, if γ is sufficiently small, then x^{n+1} will be closer to the local minimum in F then was x^n. Note that the superscripts do not represent exponents.

$$x^{n+1} = x^n - \gamma \nabla F(x^n) \tag{6}$$

This is an iterative process, which in many circumstances will converge towards the local minimum.

The objective, therefore, is to move towards the minimum in the mean absolute error surface, E. Weights and biases, *w* and *b*, are iteratively updated until this minimum is reached. This is done for all weights and biases in the network for a number of iterations until some stopping criterion is reached. The gradient descent algorithm in this case becomes:

$$w_{ij}^{n+1} = w_{ij}^{n} - \gamma \frac{\partial E(w_{ij}^{n})}{\partial w_{ij}^{n}} \tag{7}$$

$$b_{ij}^{n+1} = b_{ij}^{n} - \gamma \frac{\partial E(b_{ij}^{n})}{\partial b_{ij}^{n}} \tag{8}$$

The value of γ can change between successive iterations; this is called an adaptive learning rate. Convergence can be sped up if the learning rate is increased on flat parts of the error surface, and decreased where the slope is steep. To implement this simply, the learning rate is increased if the error decreases, and is decreased if the error increases.

A momentum operator can also be used to stabilize the trajectory of the convergence. This will essentially act as a low pass filter to smooth out any oscillations in the convergence trajectory. To illustrate momentum learning, recall that the weight update at iteration n is:

$$\Delta w_{ij}^{n} = -\gamma \frac{\partial E(w_{ij}^{n})}{\partial w_{ij}^{n}} \tag{9}$$

With momentum learning, this becomes:

$$\Delta w_{ij}^{n} = \alpha \Delta w_{ij}^{n-1} - (1-\alpha)\gamma \frac{\partial E(w_{ij}^{n})}{\partial w_{ij}^{n}} \qquad (10)$$

For some α that satisfies

$$0<\alpha<1 \qquad (11)$$

The value of α used in this study is 0.9.

Weights and biases will now converge quickly and stably towards the minimum in the mean absolute error surface [20]. In this way, neural networks can create a transformation that minimizes the error between the output of the network and the target mapping.

Multivariate Linear Regression

This method aims to combine many inputs, in this case many seismic attributes, in a linear manner to create a mapping to the target output [16-18]. It is a simple extension of conventional linear analysis.

Assume that there are m inputs each of length n, with the nth sample of the mth input denoted as A_{mn}. These inputs are to be mapped to an output log of length n, with the nth sample of the log denoted by L_n.

The log prediction is represented by the linear equation:

$$\begin{aligned} L_1 &= w_0 + w_1 A_{11} + w_2 A_{21} + \ldots + w_m A_{m1} \\ L_2 &= w_0 + w_1 A_{12} + w_2 A_{22} + \ldots + w_m A_{m2} \\ &\vdots \\ L_n &= w_0 + w_1 A_{1n} + w_2 A_{2n} + \ldots + w_m A_{mn} \end{aligned} \tag{12}$$

This set of equations can be written in matrix form as:

$$\begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{bmatrix} = \begin{bmatrix} 1 & A_{11} & A_{21} & A_{31} & \cdots & A_{m1} \\ 1 & A_{12} & A_{22} & A_{32} & \cdots & A_{m2} \\ 1 & A_{13} & A_{23} & A_{33} & \cdots & A_{m3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & A_{1n} & A_{2n} & A_{3n} & \cdots & A_{mn} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} \tag{13}$$

or

$$L = AW \tag{14}$$

The weights, w, in this equation may be derived by minimizing the difference between the prediction and the target log, in a least squares sense. The mean squared prediction error is:

$$E^2 = \frac{1}{n}\sum_{i=1}^{n}(L_i - w_0 - w_1 A_{1i} - w_2 A_{2i} - \ldots - w_m A_{mi})^2 \quad (15)$$

The minimum of this summation will occur where the first derivative with respect to the weights equals zero.

$$\begin{aligned} \frac{\partial E^2}{\partial w_0} &= 0 \\ &\vdots \\ \frac{\partial E^2}{\partial w_m} &= 0 \end{aligned} \quad (16)$$

Finding these partial derivatives and rewriting gives:

$$\begin{aligned} n w_0 + \sum_{i=1}^{n} w_1 A_{1i} + \sum_{i=1}^{n} w_2 A_{2i} + \ldots + \sum_{i=1}^{n} w_m A_{mi} &= \sum_{i=1}^{n} L_i \\ \sum_{i=1}^{n} w_0 A_{1i} + \sum_{i=1}^{n} w_1 A_{1i}^2 + \sum_{i=1}^{n} w_2 A_{2i} A_{1i} + \ldots + \sum_{i=1}^{n} w_m A_{mi} A_{1i} &= \sum_{i=1}^{n} L_i A_{1i} \\ &\vdots \\ \sum_{i=1}^{n} w_0 A_{mi} + \sum_{i=1}^{n} w_1 A_{1i} A_{mi} + \sum_{i=1}^{n} w_2 A_{2i} A_{mi} + \ldots + \sum_{i=1}^{n} w_m A_{mi}^2 &= \sum_{i=1}^{n} L_i A_{mi} \end{aligned} \quad (17)$$

Once again, in matrix notation this becomes:

$$\begin{bmatrix} n & \sum_{i=1}^{n} A_{1i} & \sum_{i=1}^{n} A_{2i} & \sum_{i=1}^{n} A_{3i} & \cdots & \sum_{i=1}^{n} A_{mi} \\ \sum_{i=1}^{n} A_{1i} & \sum_{i=1}^{n} A_{1i}^2 & \sum_{i=1}^{n} A_{2i}A_{1i} & \sum_{i=1}^{n} A_{3i}A_{1i} & \cdots & \sum_{i=1}^{n} A_{mi}A_{1i} \\ \sum_{i=1}^{n} A_{2i} & \sum_{i=1}^{n} A_{1i}A_{2i} & \sum_{i=1}^{n} A_{2i}^2 & \sum_{i=1}^{n} A_{3i}A_{2i} & \cdots & \sum_{i=1}^{n} A_{mi}A_{2i} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} A_{mi} & \sum_{i=1}^{n} A_{1i}A_{mi} & \sum_{i=1}^{n} A_{2i}A_{mi} & \sum_{i=1}^{n} A_{3i}A_{mi} & \cdots & \sum_{i=1}^{n} A_{mi}^2 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} L_i \\ \sum_{i=1}^{n} L_i A_{1i} \\ \vdots \\ \sum_{i=1}^{n} L_i A_{mi} \end{bmatrix} \quad (18)$$

In order to solve this system for the weights, we must invert the coefficient matrix on the left. The equation then becomes:

$$\begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} A_{1i} & \sum_{i=1}^{n} A_{2i} & \sum_{i=1}^{n} A_{3i} & \cdots & \sum_{i=1}^{n} A_{mi} \\ \sum_{i=1}^{n} A_{1i} & \sum_{i=1}^{n} A_{1i}^2 & \sum_{i=1}^{n} A_{2i}A_{1i} & \sum_{i=1}^{n} A_{3i}A_{1i} & \cdots & \sum_{i=1}^{n} A_{mi}A_{1i} \\ \sum_{i=1}^{n} A_{2i} & \sum_{i=1}^{n} A_{1i}A_{2i} & \sum_{i=1}^{n} A_{2i}^2 & \sum_{i=1}^{n} A_{3i}A_{2i} & \cdots & \sum_{i=1}^{n} A_{mi}A_{2i} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} A_{mi} & \sum_{i=1}^{n} A_{1i}A_{mi} & \sum_{i=1}^{n} A_{2i}A_{mi} & \sum_{i=1}^{n} A_{3i}A_{mi} & \cdots & \sum_{i=1}^{n} A_{mi}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n} L_i \\ \sum_{i=1}^{n} L_i A_{1i} \\ \vdots \\ \sum_{i=1}^{n} L_i A_{mi} \end{bmatrix} \quad (19)$$

The coefficient matrix is square and is invertible numerically. In this manner the weights which minimize the least squares prediction error can be determined. While this computation is realizable, there is a short cut which is much less computationally expensive [16-18].

Note that if

$$A = \begin{bmatrix} 1 & A_{11} & A_{21} & A_{31} & \cdots & A_{m1} \\ 1 & A_{12} & A_{22} & A_{32} & \cdots & A_{m2} \\ 1 & A_{13} & A_{23} & A_{33} & \cdots & A_{m3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & A_{1n} & A_{2n} & A_{3n} & \cdots & A_{mn} \end{bmatrix} \tag{20}$$

Then

$$A^T A = \begin{bmatrix} n & \sum_{i=1}^{n} A_{1i} & \sum_{i=1}^{n} A_{2i} & \sum_{i=1}^{n} A_{3i} & \cdots & \sum_{i=1}^{n} A_{mi} \\ \sum_{i=1}^{n} A_{1i} & \sum_{i=1}^{n} A_{1i}^2 & \sum_{i=1}^{n} A_{2i}A_{1i} & \sum_{i=1}^{n} A_{3i}A_{1i} & \cdots & \sum_{i=1}^{n} A_{mi}A_{1i} \\ \sum_{i=1}^{n} A_{2i} & \sum_{i=1}^{n} A_{1i}A_{2i} & \sum_{i=1}^{n} A_{2i}^2 & \sum_{i=1}^{n} A_{3i}A_{2i} & \cdots & \sum_{i=1}^{n} A_{mi}A_{2i} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} A_{mi} & \sum_{i=1}^{n} A_{1i}A_{mi} & \sum_{i=1}^{n} A_{2i}A_{mi} & \sum_{i=1}^{n} A_{3i}A_{mi} & \cdots & \sum_{i=1}^{n} A_{mi}^2 \end{bmatrix} \tag{21}$$

This matrix is equal to the coefficient matrix which must be inverted in equation 19. This matrix is also a very powerful tool used in data analysis, and is known as the covariance matrix.

Also. note that the right hand side of Equation 19 is equal to:

$$A^T L = \begin{bmatrix} \sum_{i=1}^{n} L_i \\ \sum_{i=1}^{n} L_i A_{1i} \\ \vdots \\ \sum_{i=1}^{n} L_i A_{mi} \end{bmatrix} \tag{22}$$

So that starting from equation 13,

$$AW = L \tag{23}$$

this leads to

$$A^T AW = A^T L \tag{24}$$

which is equivalent to equation 18.

A^TA is a square symmetric matrix and so in many cases may be invertible. To solve for W, we rearrange equation 24 to:

$$W = \left(A^T A\right)^{-1} A^T L \tag{25}$$

These weights in vector W provide a mapping between the attributes in matrix A, and the target log, L. The difference between the outputs of this mapping and the target log are minimized in a least squares sense.

Seismic Attributes

Seismic attributes are measures of dynamic, kinetic and statistical features of seismic data. A seismic trace is the result of complicated interrelationships between bed thickness, porosity, fluid saturation, lithological boundaries, and other rock properties [15, 22, 23]. As such, a seismic trace can be transformed into an attribute which amplifies one, or many, of these properties.

Hundreds of seismic attributes have been developed; only some of these are well enough understood to be quantitative, and many are redundant. We have chosen to use *Instantaneous Attributes*, so called because they are calculated at every time sample of the seismic trace. A detailed list of the attributes used, and their potential geologic significance is shown in Table 3 [15, 24, 25].

Table 3. List of seismic attributes and their significance.

Attribute	**Significance**
Amplitude	Acoustic Impedance Contrast
Instantaneous Phase	Indicative of Lateral Continutity
Instantaneous Frequency	Bed Thickness Indicator; DHI
Amplitude Envelope	Reflection Strength
First Derivative of the Amplitude	Absorption Effects
Second Derivative of the Amplitude	Bed Thickness; Reflection Strength
Integrated Absolute Amplitude	Low Frequency Trends
Sample Number	Vertical Continuity

The instantaneous phase, frequency, and envelope of the seismic traces is computed through complex trace analysis [26]. In complex trace analysis, a seismic trace, s(t), is viewed as the real part of an *analytical trace*, S(t), that is composed of both a real and imaginary part.

$$S(t) = s(t) + js^*(t) \tag{25}$$

Where j equals the square root of negative one, s*(t) is the quadrature component of the trace, and is uniquely determined by the in-phase component, if there is adherence to a liberal set of assumptions. If S(t) is described as the combination of a time varying amplitude, known as the amplitude envelope, A(t), and a time-varying phase, θ(t), then:

$$S(t) = s(t) + js^*(t) = A(t)e^{i\theta(t)} \tag{26}$$

Which leads to:

$$s(t) = \mathrm{Re}\left\{A(t)e^{i\theta(t)}\right\} = A(t)\cos(\theta(t)) \tag{27}$$

$$s^*(t) = \mathrm{Im}\left\{A(t)e^{i\theta(t)}\right\} = A(t)\sin(\theta(t)) \tag{28}$$

Solving for A(t):

$$A(t) = \sqrt{s^2(t) + s^{*2}(t)} \tag{29}$$

And θ(t), the instantaneous phase:

$$\theta(t) = \tan^{-1}\left(\frac{s^*(t)}{s(t)}\right) \tag{30}$$

The instantaneous frequency, *f*(t), is simply the time derivative of the instantaneous phase, divided by 2π:

$$f(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt} \tag{31}$$

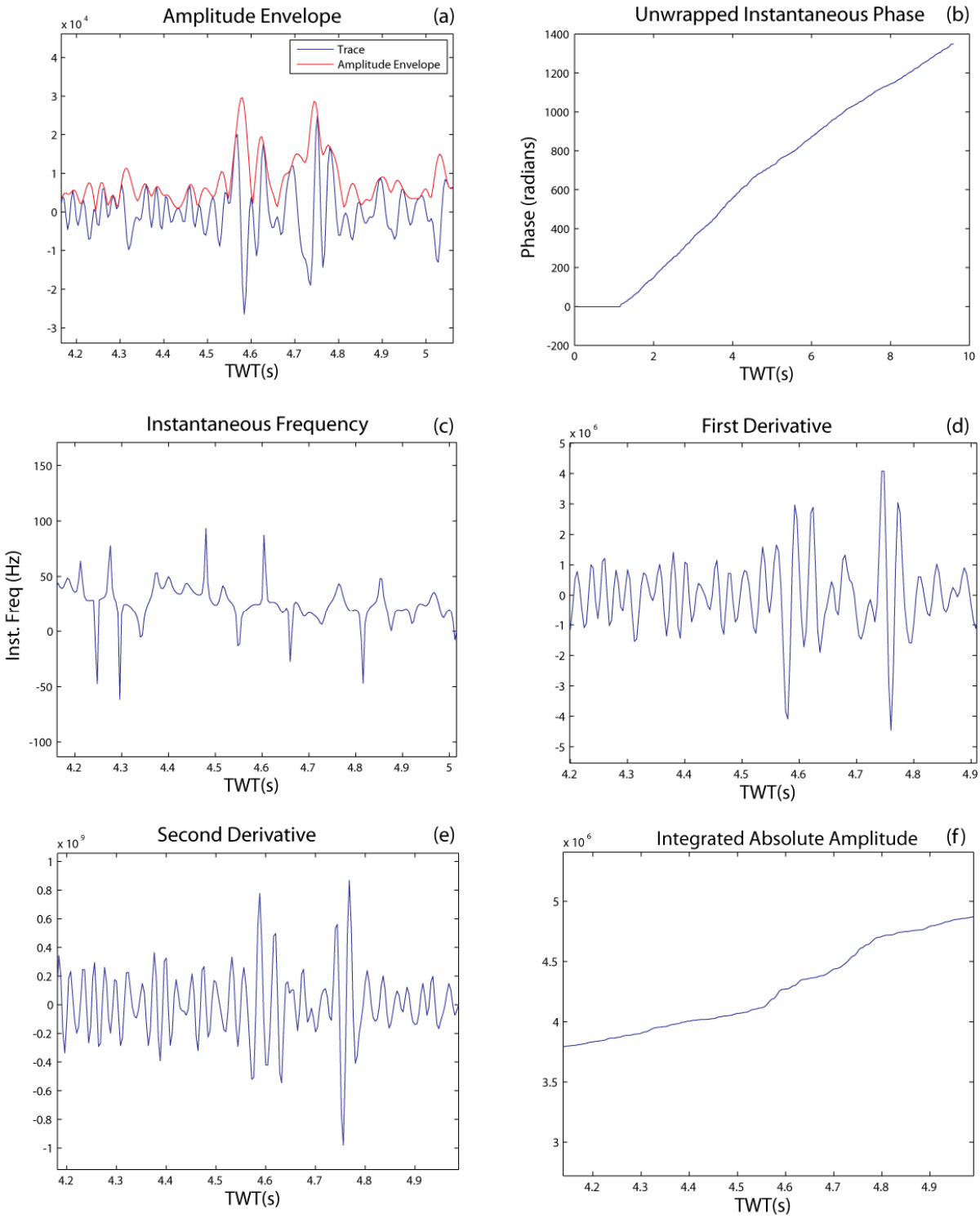


Figure 3. a) A seismic trace in blue with its amplitude envelope in red. For every trace sample there is an attribute sample. **b)** The unwrapped instantaneous phase of the trace in a). **c)** The

instantaneous frequency. **d)** First Derivative of the original signal with time. **e)** Second Derivative of the original signal with time. **f)** The integrated Absolute Amplitude is also used as an attribute. It may highlight low frequency vertical trends in the data.

Normalization and Principal Component Analysis

As discussed previously, the hyperbolic tangent sigmoid transfer function used in the neural network has a limited range. The domain is effectively limited as well since the function saturates quickly for absolute input values greater than around five. To overcome this, both the input and output data sets must be processed in order to take full advantage of the shape of the transfer function [27].

This is accomplished by processing the training set of the network to give both the input and output sets a mean of zero and a standard deviation of one. For example the normalized input set would be calculated as:

$$Normalized\_Input = \frac{Input - mean(Input)}{STD(Input)} \tag{32}$$

The same transformation is applied to the output set. This means that the network will be trained to yield outputs that have zero mean and unity standard deviation. These normalized results are then transformed back by multiplying by the standard deviation and adding the mean of the training output set.

In addition to normalizing the input set, we have also chosen to subsequently transform it using principal component analysis. While the input set may be large, many of the attributes may be highly correlated or redundant. Principal Component Analysis has three effects: it orthogonalizes the input vectors so that they are uncorrelated, it orders the resulting orthogonalized components such that those with the largest variation come first, and it eliminates the components that add the least amount of variation to the

data set [28]. In this case, we have chosen to eliminate data that contributes less than five percent of the variation.

In using principal component analysis we calculate the covariance matrix of the normalized input set, find its unit eigenvectors, sort the eigenvectors in order from largest to smallest eigenvalue, and then eliminate the undesired components. These remaining eigenvectors are now used as a transformation matrix, and essentially act as the axes in a new coordinate system.

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_7 \end{bmatrix} = \begin{bmatrix} -0.2547 & -0.0047 & 0.5835 & 0.6075 & -0.2103 & -0.0464 & 0.3314 & 0.1652 \\ 0.0298 & 0.0292 & 0.1552 & 0.1314 & 0.6962 & 0.6834 & 0.0528 & -0.0488 \\ 0.8177 & 0.1778 & 0.2053 & 0.1183 & 0.2726 & -0.4002 & 0.0818 & -0.0493 \\ -0.074 & 0.6729 & -0.2827 & -0.3194 & 0.0227 & 0.0362 & 0.5938 & 0.0722 \\ 0.0909 & -0.1316 & -0.173 & 0.0357 & 0.1057 & -0.0027 & -0.0241 & 0.9651 \\ -0.0686 & 0.6658 & 0.121 & -0.0085 & -0.0249 & 0.0374 & -0.7241 & 0.104 \\ -0.0047 & -0.1063 & 0.6838 & -0.7029 & -0.0517 & 0.0436 & 0.0495 & 0.1416 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8 \end{bmatrix}$$

The above equation shows the actual transformation matrix used in determining the principal components, p_i, of the eight normalized seismic attributes, s_i, used in the study.

Since all of the eigenvectors are orthogonal, this is the most efficient way to represent the data. This coordinate system now effectively communicates where along the trend line an individual data point sits.

CASE STUDY AND METHODOLOGY

Auger Field – N Interval

This study makes use of data over Auger Field in the deepwater Gulf of Mexico (GB 426, 427, 470, 471). Several studies of the geologic character of the field have been published [29-32]. Included in the Appendix of this thesis is "Characterization of the O and N Sands in Auger Field, Gulf of Mexico", which lays the stratigraphic framework for this more detailed rock property analysis [33].

This study focuses on the shallowest reservoir in Auger Field, the N Interval. Hydrocarbons in this interval are trapped by a three-way closure against the Auger Salt Ridge and the East Auger Fault. In the majority of the reservoir, the N Interval is composed of two massive sands, the N_M^1, and the N_M^2, termed the Lobe Facies by Bohn (2008). Each of these is capped by a laminated member, the N_L^1 and the N_L^2, respectively, which are termed the Levee Facies [33]. Close to the salt dome, however, the N Interval is comprised mainly of the N_M^1 and N_L^1 members, as the N_M^2 and N_L^2 members have onlapped the Auger Salt Ridge and pinched out.

Throughout deposition of the N Interval, incision was occurring from a northern source. This incision and bypass to the neighboring Andros Basin created a valley which thickens and widens to the south towards the depocenter. The evidence of this incision is seen downdip and to the west, where a third facies, the Incised Fill Facies, is present (Figure 14). The seismic signature of the interval becomes more chaotic; this represents the character of the Incised Fill Facies, which is composed of a lower net-to-gross sand than is present in the Lobe Facies. The N_L^1 and N_L^2 members are also seen to thicken at the valley margin. This is most likely due to overbank deposition by the incision.

There is log evidence that there is overall a lower average porosity in the incised valley fill than there is in the Lobe Facies.

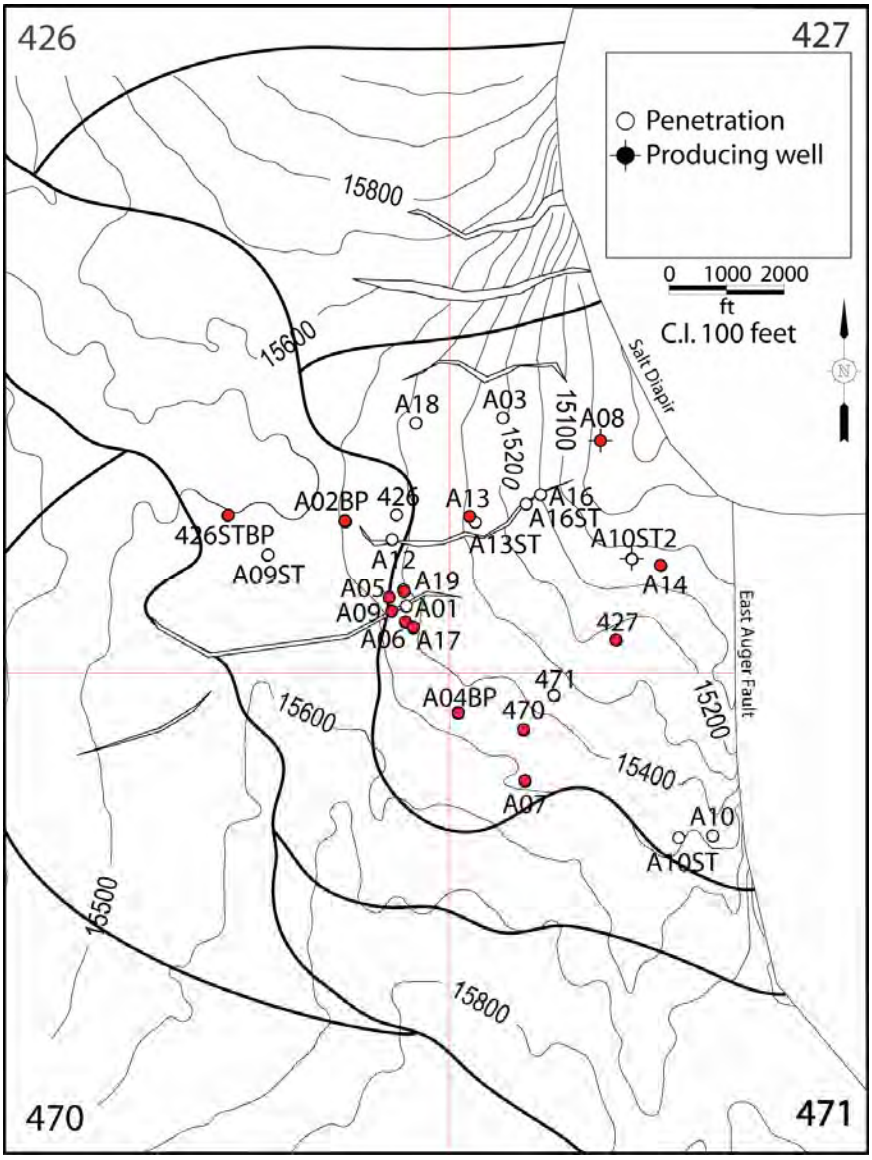


Figure 4. Structure Map on top of the N interval, outlining the Lobe Facies, Incised Fill, and well penetrations. Wells used in the training set are filled with red. The wells penetrating the Incised Fill that are used in this study include the A02BP1, the A05, and 426STBP1. The A01, A06, A07, and A09 penetrate the overbank deposition.

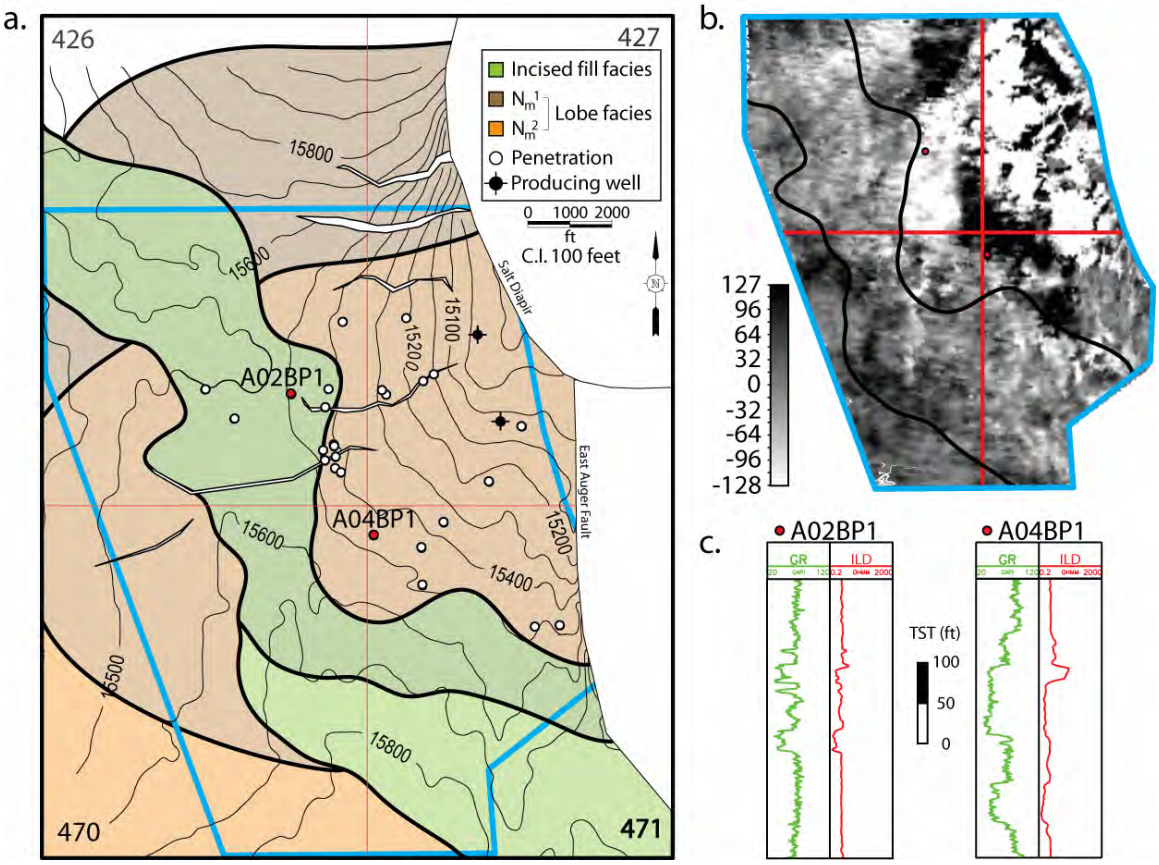


Figure 5. **a)** Structure map of the N_M^1 with a facies interpretation overlain. The Incised Fill Facies is shown in green; it widens and thickens to the south. **b)** Time slice showing lobate geometry and the basin axis channel. **c)** Type well logs showing considerable change in facies character going from the Lobe Facies (A04BP1) to the Incised Fill Facies (A02BP1).

There is extensive 3D seismic coverage over the field. The N Interval has 47 well penetrations, but only 14 wells were drilled before production began and have a suitable suite of well logs for analysis.

The seismic survey used was completed in 1990 before the field came on-line, and before the Tension Leg Platform was present. The survey was shot in a North-South orientation with three 1545 cu. in. airguns, at 2000 PSI. The far offset was 20000 feet,

with a CMP bin size of 100 ft. X 100 ft. giving a nominal fold of 30. The data underwent pre-stack time migration before undergoing integration and stacking.

Well Log Pre-Processing

In order for the well logs to be directly compared to the seismic data, a considerable amount of pre-processing was necessary. The primary processing steps are despiking and depth-to-time conversion.

The despiking and quality control of the well logs was accomplished by filtering the logs with a 35 point running mean filter. Applying the running mean filter has striking results that are visually evident in the log character (Figure 6).

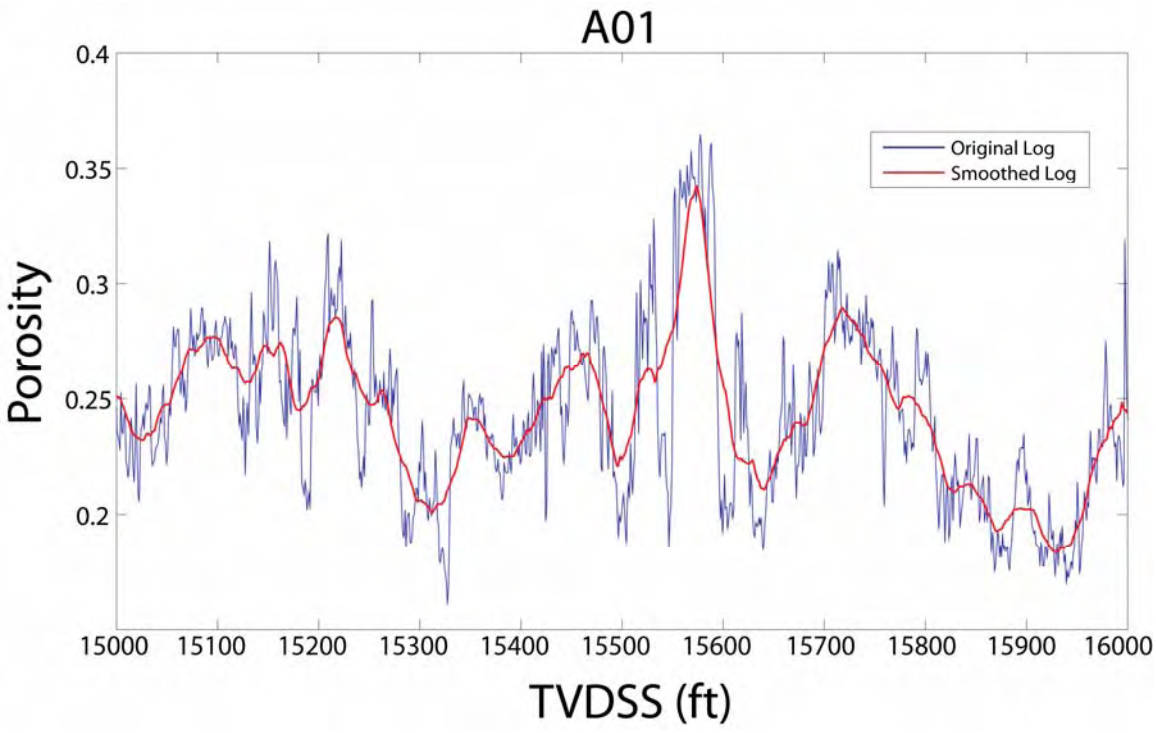


Figure 6. Comparison of the original, unedited porosity log from well A01 (in blue), and the log once it has been smoothed by a running mean filter (red).

The next pre-processing step is that of depth to time conversion. While the seismic traces are sampled in units of time, the well logs are sampled in units of depth.

Since the goal is to create synthetic well logs based on the seismic data, it is prudent to reference the well logs in the time domain.

Generally, as is the case for the wells being studied here, the seismic transit time to certain depths along the wellbore is measured directly in what is known as a checkshot survey. These data provides the most accurate transformation from depth to time. The checkshot coverage along the wellbore is sparse, and must be upsampled. The seismic survey being used has a sampling interval of 4 ms. For twenty samples around the N_m^1, giving a 100 ms window, depths are interpolated to every 4 ms sample through spline interpolation from the checkshot (Figure 7).

Now that the depth at which each seismic sample occurs is estimated, we must now be certain that a well log sample also occurs at that depth. For the most part, this is not the case, and we must interpolate the well log to give us a sample at that depth (Figure 8). The extent of these points on the well log corresponds to a 100 ms seismic window around the interval of interest, in this case the N_m^1.

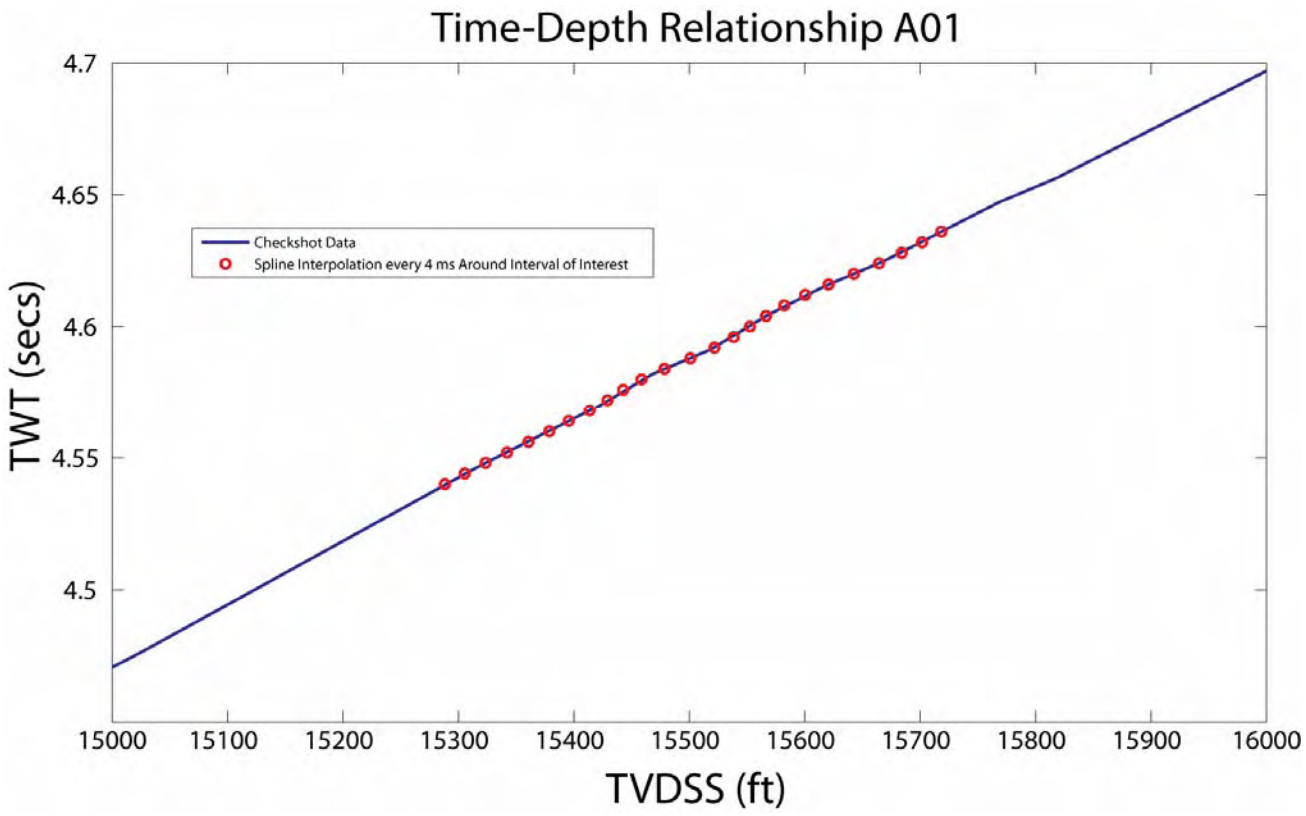


Figure 7. Example of the interpolation used to calculate the depth at which each seismic sample of interest occurs. This is data from the checkshot from well A01. At this well, the N_m^1 occurs at 4.580 seconds two-way-time. The depths are calculated for every 4 ms in a 100 ms window around the sample of interest.

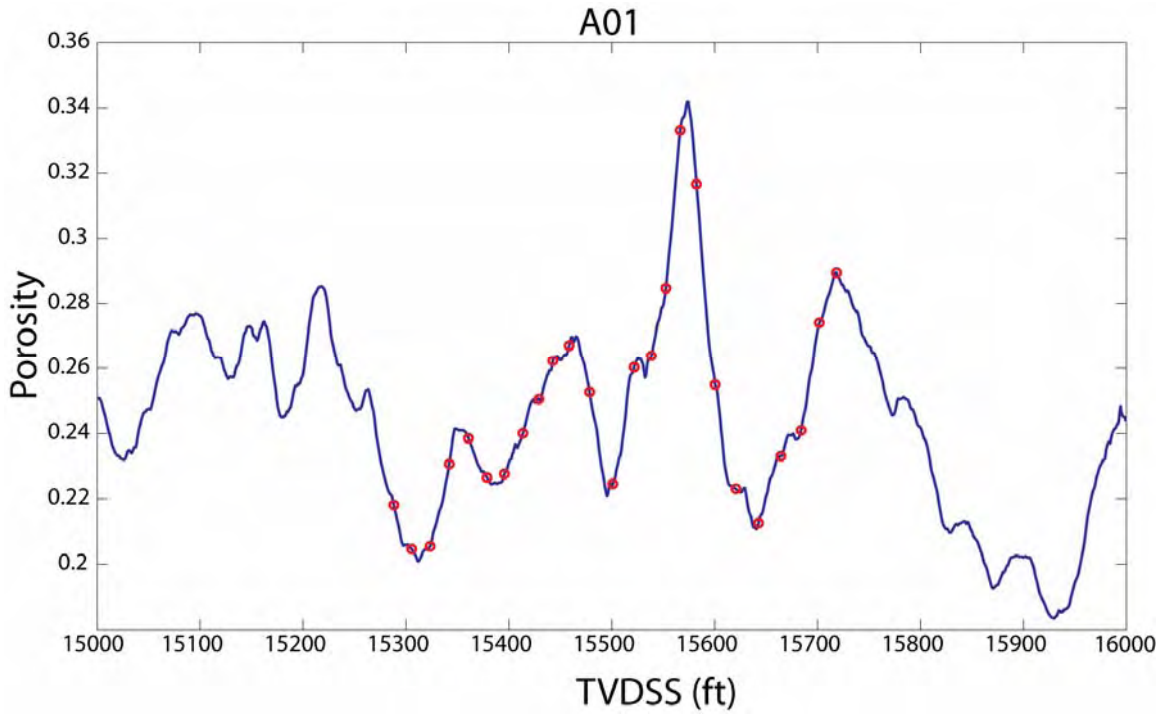


Figure 8. Porosity values that have been calculated via a spline interpolation for each of the depth points corresponding to Figure 7.

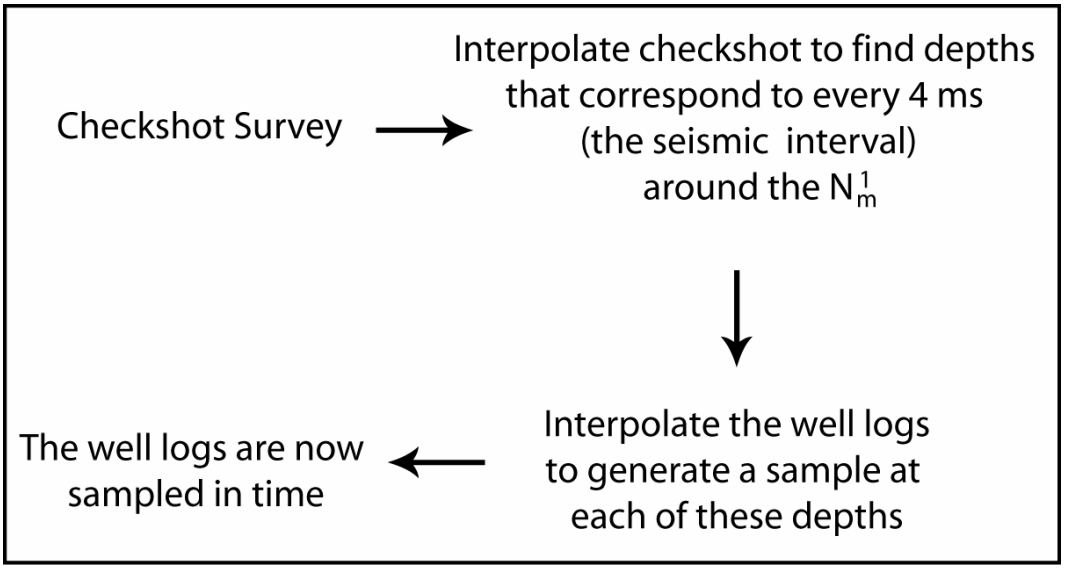


Figure 9. Flow Chart describing the methodology used to convert the well logs from being sampled in depth to being sampled in time.

Porosity Calculation from Density Logs

Although neutron porosity logs were available for analysis, it was uncertain whether each logging tool had been calibrated to the same fluid and rock type, and no information was available that described the overall calibration technique. Besides, it is generally accepted that porosity calculation from bulk density logs is more accurate[34].

To calculate porosity, φ, we use the rock matrix density, ρ_m, the fluid density, ρ_f, and the bulk density well log, ρ_b. The average rock density in the sandstones from core reports from four exploration wells is 2.66 g/cm^3. The average rock density in the shales is assumed to be 2.65 g/cm^3 [34].

The fluid density depends on whether the well encountered water or hydrocarbons. This was determined by the electrical resistivity log. If the resistivity was above an interpreted threshold value, than the hydrocarbon value was used; if below the threshold, the water value was used. The hydrocarbon density calculated from composition and phase considerations is 0.42 g/cc at 10,500 PSI, which is roughly the in situ reservoir pressure. The water density used was 1.1 g/cc. The calculation is then:

$$\phi = \frac{\rho_m - \rho_b}{\rho_m - \rho_f} \tag{23}$$

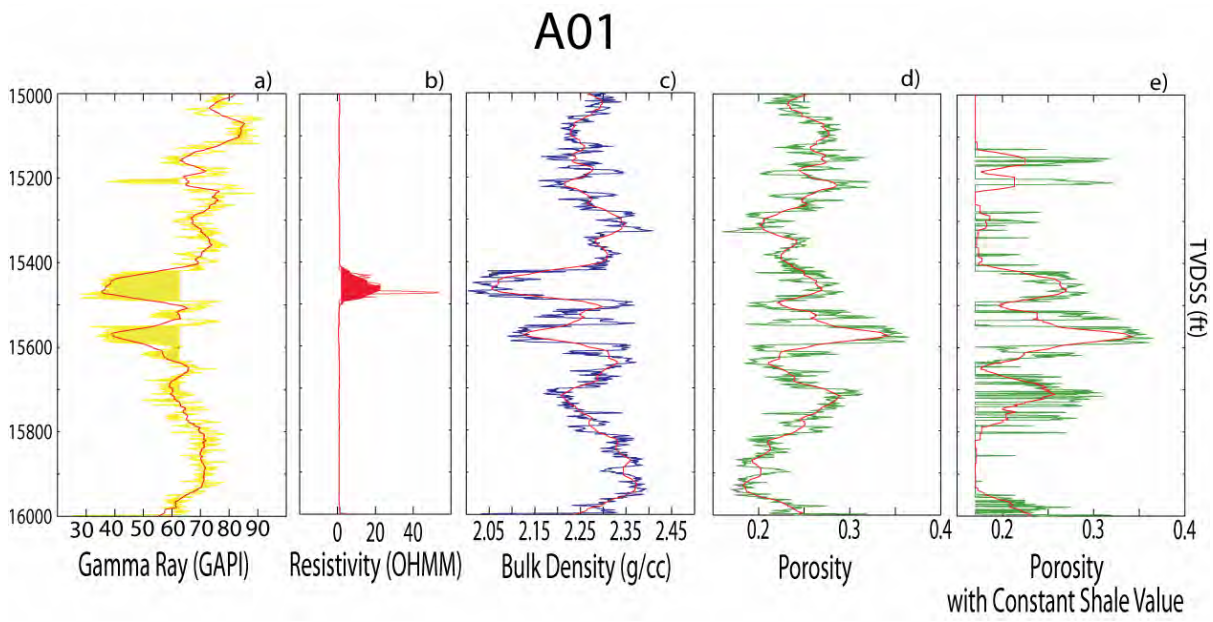


Figure 10. An example of the calculation of the Porosity Logs used in the training of the neural network and multilinear regression analysis. If the Gamma Ray **(a)** value is below a threshold value, in this case 63 GAPI, the lithology is interpreted to be sandstone. This is shown by shading yellow. Above 63 GAPI, the rock is interpreted to be shale. The fluid type, either water or gas, saturating the rock is determined from the resistivity log **(b)**. In this well, if the resistivity is above 2 OHMM, shaded red, the fluid is interpreted to be gas, with a density of 0.42 g/cc. If the resistivity is below the threshold, the fluid saturant is interpreted to be water, with density of 1.1 g/cc. For each sample, this information is combined with the density log **(c)** to create the porosity log **(d)**. We determined that the predictive power of the methods being studied is greatly enhanced if a constant porosity value of 0.17 **(e)** is used wherever it a shale has been interpreted to be the rock type. The smoothed log in each case is shown in red. The porosity logs were calculated using the unfiltered logs, however.

Porosity Prediction with an Artificial Neural Network

Once the well logs and seismic attributes had been processed as described in Chapter 2, an ANN could be suitably trained to generate a suite synthetic well logs from volumes of seismic attributes.

The selection of an appropriate neural network is heuristic in nature. Before an appropriate architecture was found, dozens of networks were designed, tested, and ultimately rejected. These networks varied in the number of hidden layers, the number of neurons in each of those hidden layers, training algorithms, transfer functions, and optimization criteria. Among the many networks tested, the most accurate and well-generalized network encountered was a feed-forward back-propagation network consisting of one hidden layer of 22 neurons with tangent sigmoid transfer functions that used a gradient-descent with momentum and adaptive learning rate training algorithm. The performance criterion was set to minimize the mean absolute error, in contrast to the more standard mean square error. This choice was once again heuristic; the network was simply more accurate and generalized better.

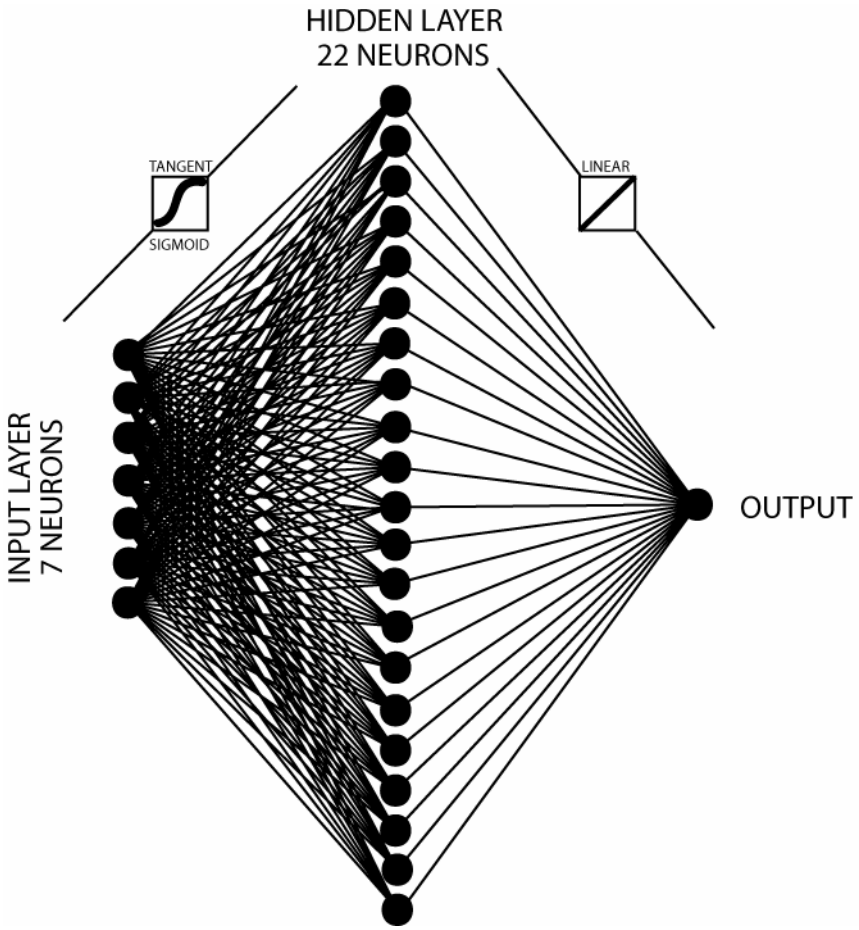


Figure 11. Architecture of the Feed Forward Network. Seven normalized input attributes are fed into twenty-two neurons in the hidden layer. The one output represents a porosity sample.

All of the logs and the corresponding seismic data were compiled into one set. The seismic attributes were normalized and underwent principal component analysis which reduced the number of inputs from eight to seven. The porosity targets were also normalized to a mean of zero. The training set consisted of the well log and seismic data from all but one well, which was separated as the test set.

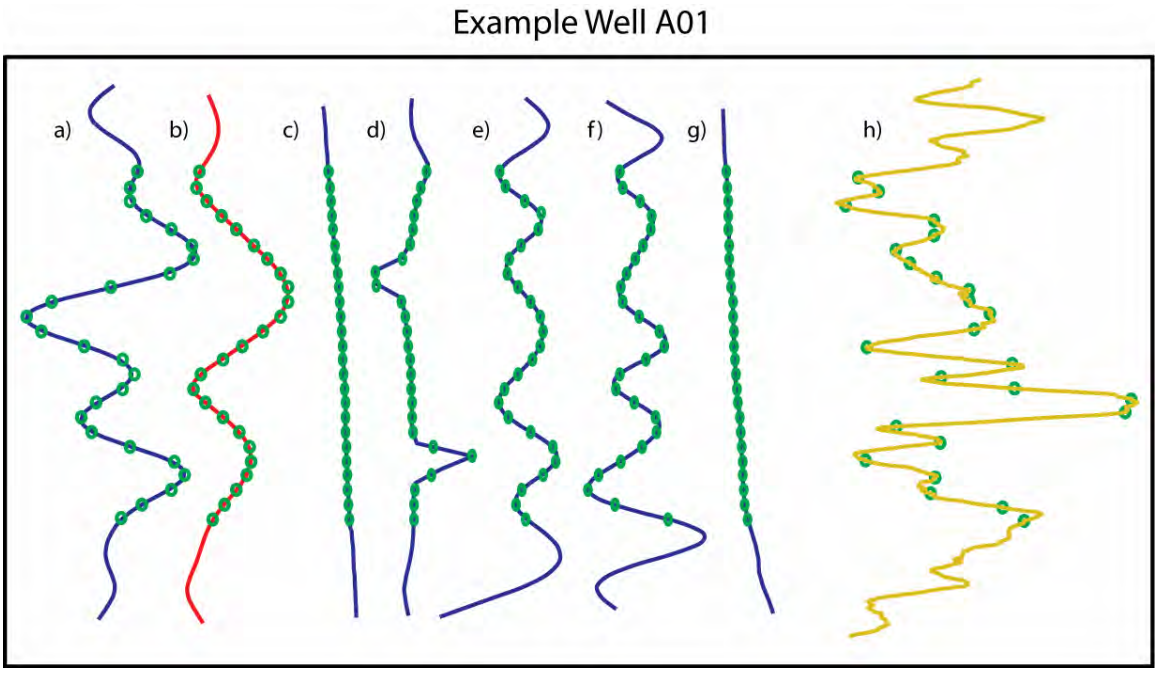


Figure 12. These attributes **(a-g),** also listed on page 18, are combined in order to create a synthetic porosity log that resembles as closely as possible the target log at right **(h)**.

We rotated through the wells, successively taking one aside for the test set and training a new network with the same network architecture in order to appropriately gauge the predictive power of this network architecture. Training commenced with both validation stopping and with stopping after 1000 epochs. When validation stopping was used, every sixth sample of the training set was set aside for validation. Results for each training cycle are shown in Table 3, where the well indicated is the well used in the test set.

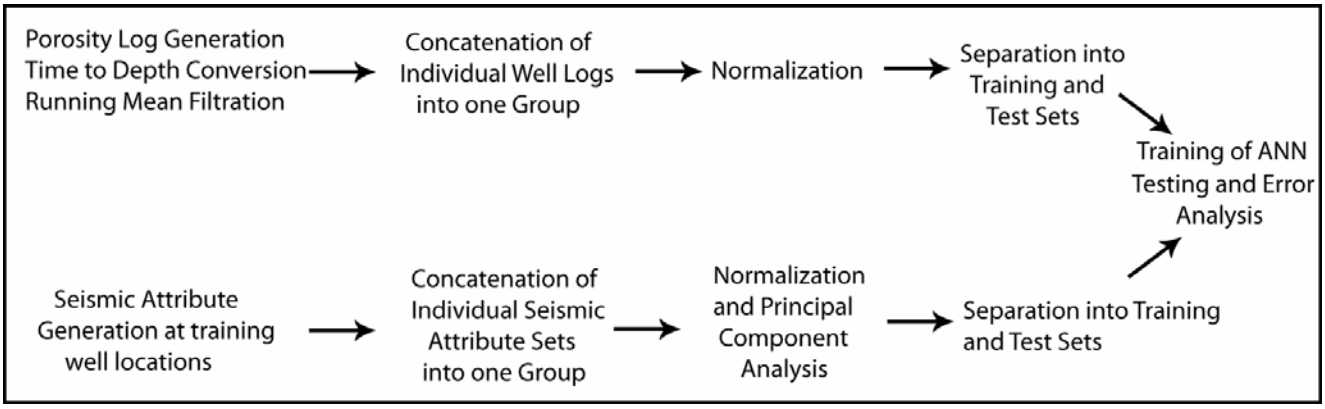


Figure 13. Flow Chart Describing Training Method.

RESULTS

ANN Predictions Using a Constant Shale Porosity

Well Name	Max Error	Mean Error	CC Test	CC Training
A01	0.107	0.039	0.537	0.871
A02BP1	0.095	0.034	0.437	0.879
A04BP1	0.101	0.020	0.825	0.872
A05	0.125	0.047	0.618	0.895
A06	0.057	0.026	0.952	0.899
A07	0.102	0.034	0.835	0.895
A08	0.141	0.070	0.450	0.892
A09	0.086	0.025	0.886	0.889
A13	0.071	0.021	0.865	0.880
A14	0.098	0.030	0.565	0.872
A19	0.106	0.030	0.658	0.890
426STBP1	0.082	0.016	0.896	0.852
427	0.057	0.026	0.833	0.872
470	0.116	0.033	0.825	0.874
AVERAGES	**0.096**	**0.032**	**0.728**	**0.881**

Table 4. Results of porosity prediction after 1000 epochs of training. All wells except the one indicated were used in the training set. For example, the second row of the table contains the data for the A01. All of the wells except for the A01 were used to train the network; A01 was left out of the training as the test case. The maximum and average errors are in units of porosity (V/V). CC Test and CC Training are the normalized cross correlation coefficients between the prediction and the targets for the test and training sets, respectively.

The performance of the ANN was analyzed by keeping one well aside as a test case, and training the network on the remaining thirteen wells. This procedure was performed fourteen times, each time testing the network on a new well.

The testing results demonstrate that the network generalizes well, and has impressive predictive powers. With an average cross correlation coefficient of 0.72 for

the test case, this network outperforms many of those from previous studies (see Intro) [16, 18, 19]. The mean error is on the order of 3% porosity which is much lower than the range predicted which was between 16% and 33%.

This was accomplished without the use of an inverted acoustic impedance volume as input, which has been demonstrated to have the highest correlation with porosity of any single attribute [19]. We attribute this success to the quality control of the well logs and well-ties, the preprocessing of the input seismic attributes and output porosity logs which includes normalization and principal component analysis, and the choice of training algorithm.

The Hinton Plot below shows that the normalized inputs are given roughly equal weighting when estimating the porosity log. This suggests that all of the attributes used in the input set have a functional relationship with porosity, and that that relationship may be nonlinear in nature.

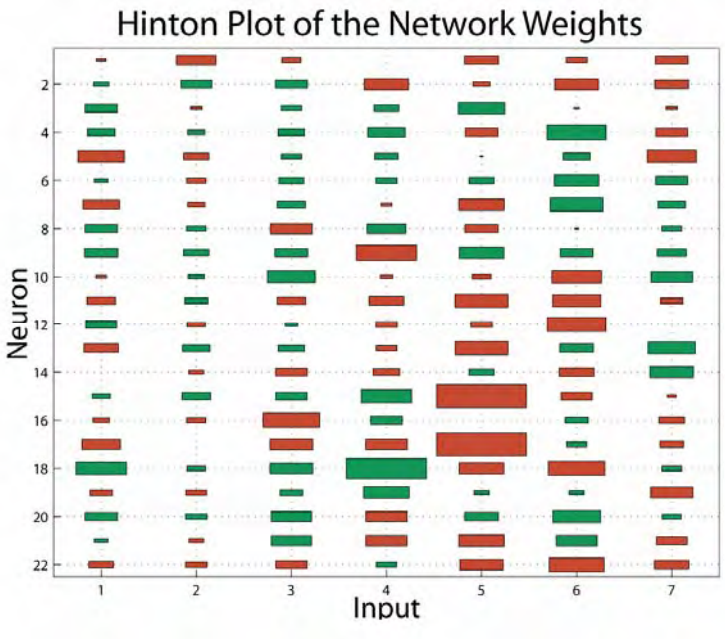
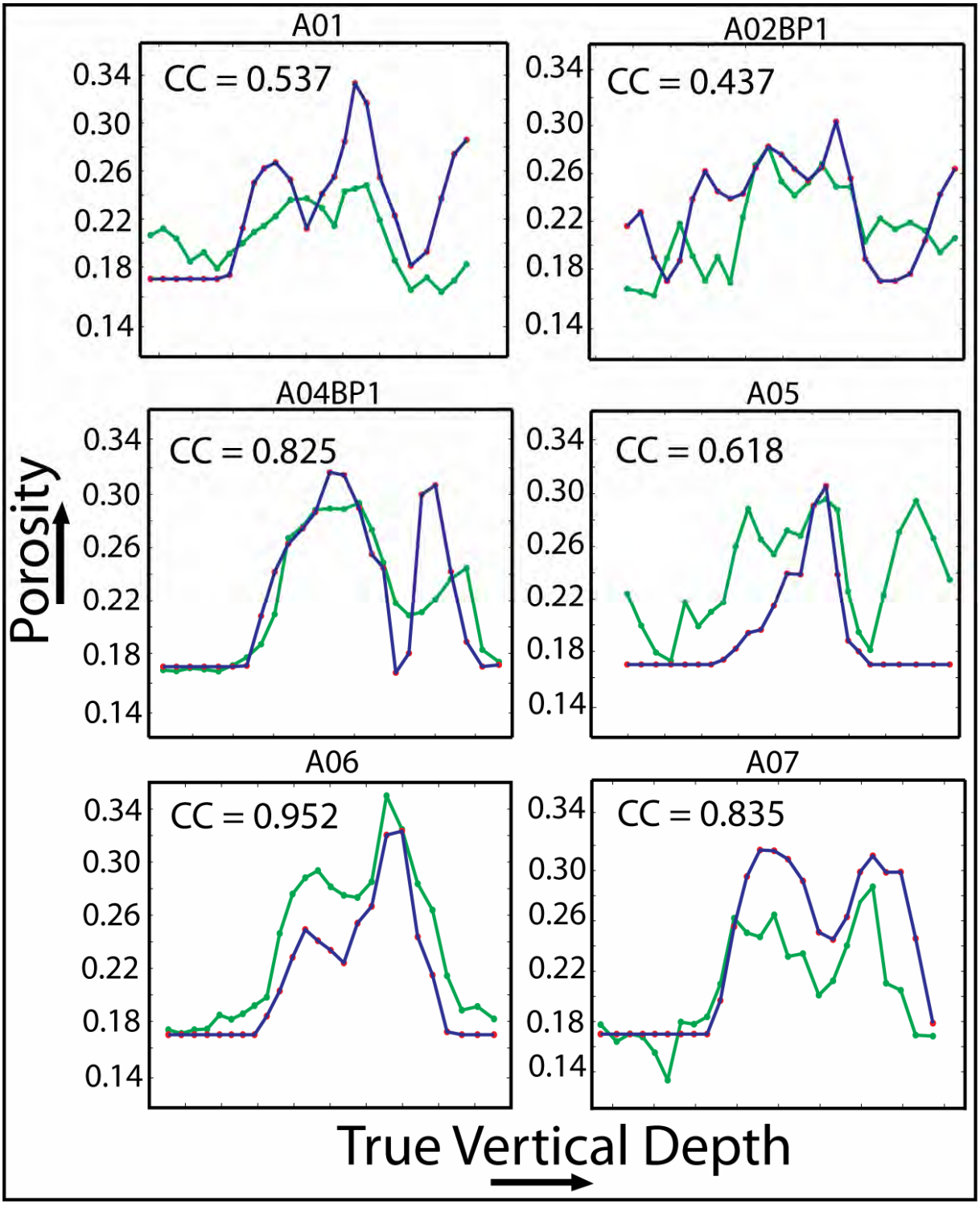
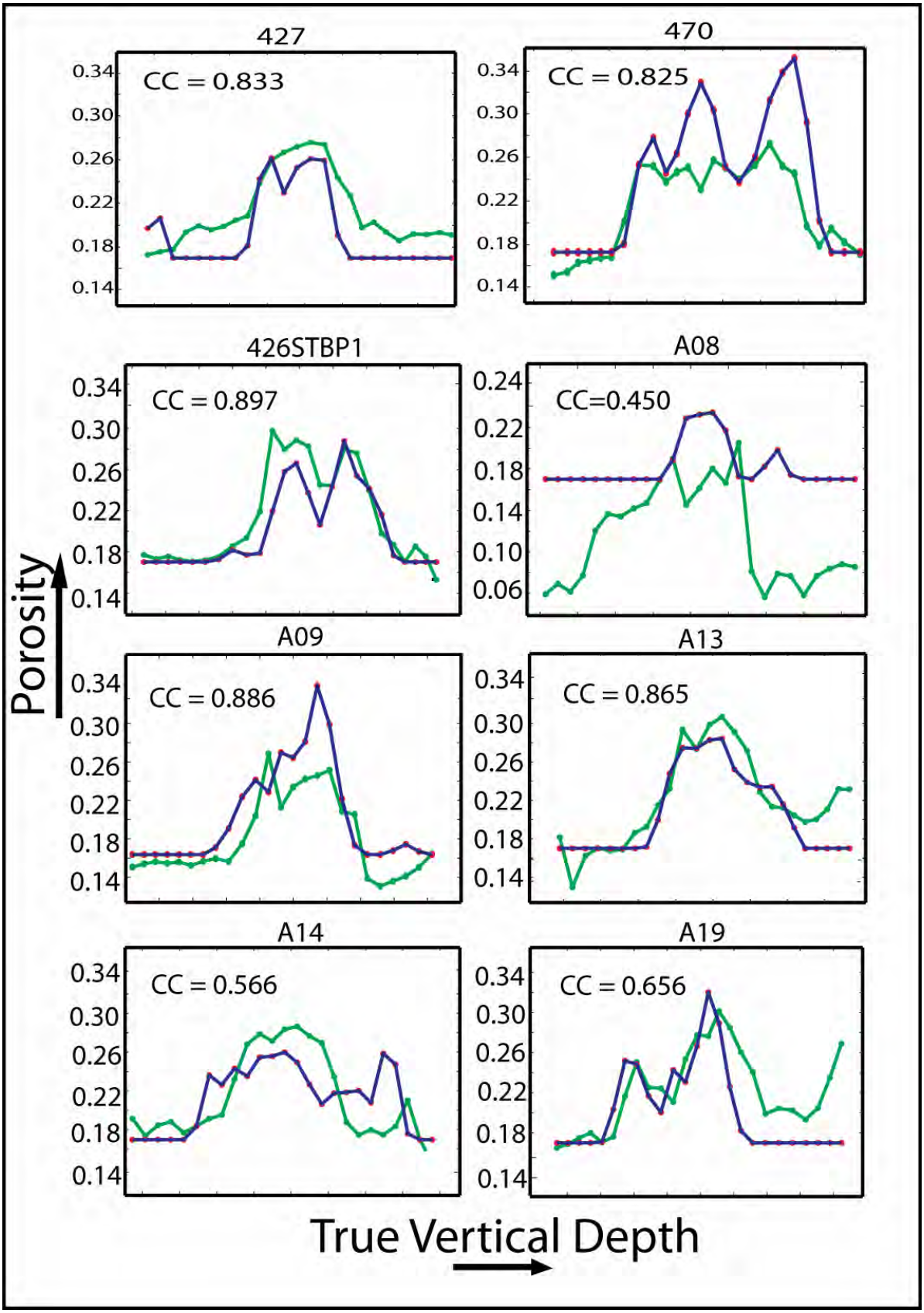


Figure 14. Hinton plot that shows the strength of the weights between the input layer and the hidden layer. The area of the box corresponds to the amplitude of the weight, while the color indicates polarity. A red weight is negative, and a green weight is positive. The seven inputs being weighted are the principal components of the original, normalized seismic attributes.


A01
CC = 0.537
A02BP1
CC = 0.437
A04BP1
CC = 0.825
A05
CC = 0.618
A06
CC = 0.952
A07
CC = 0.835
0.34
0.30
0.26
0.22
0.18
0.14
Porosity
True Vertical Depth


427
CC = 0.833
470
CC = 0.825
426STBP1
CC = 0.897
A08
CC=0.450
A09
CC = 0.886
A13
CC = 0.865
A14
CC = 0.566
A19
CC = 0.656
Porosity
True Vertical Depth

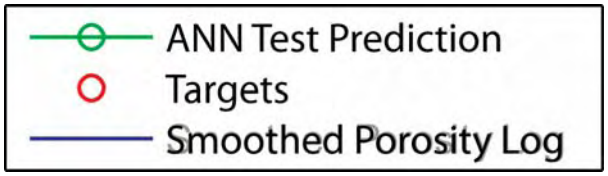
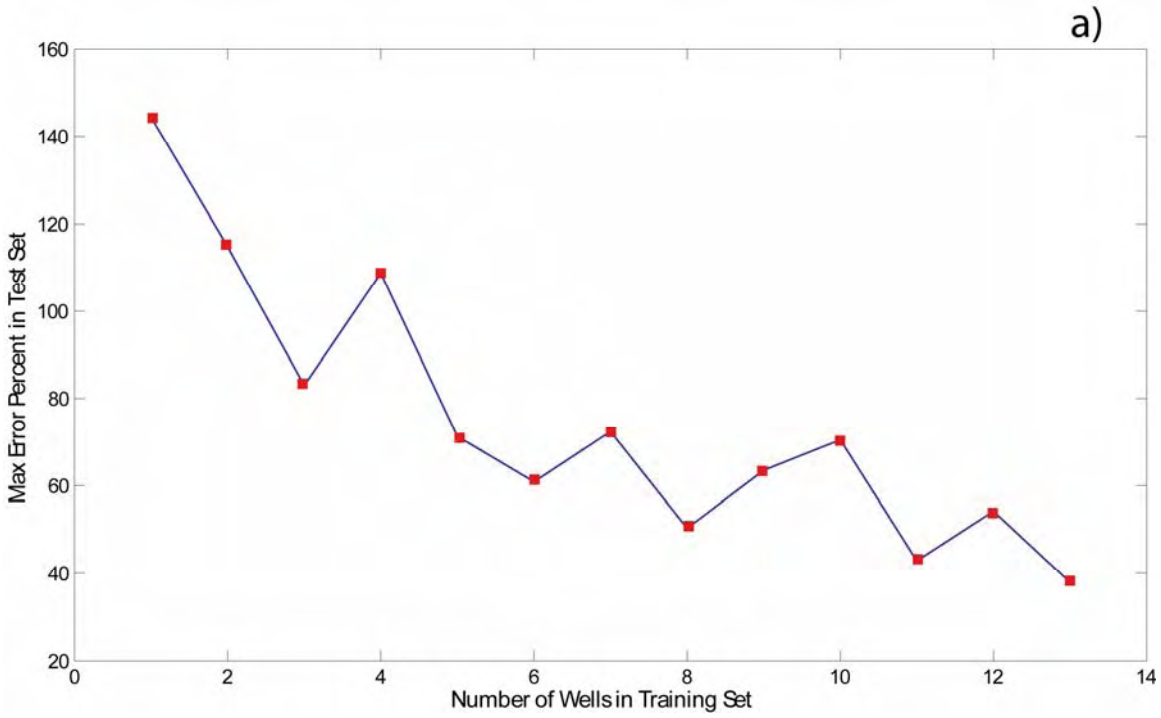


Figure 15. Test wells from the training sequence of the Neural Network with 1000 Epochs of Training. The CC value indicates the normalized cross-correlation coefficient between the neural network output and the target values.

We also performed an additional test of the predictive capabilities of the ANN. Instead of removing only one well as the test case, we successively removed additional wells until only one well served as the training case. As expected, the predictive power of the ANN grew as the number of training wells increased. Both absolute error, as well as error percent decreased, while the testing cross correlation coefficient increased. In Figure 16 a) the decrease in error percent appears to be asymptotic. The benefits of adding more training wells diminish rapidly.



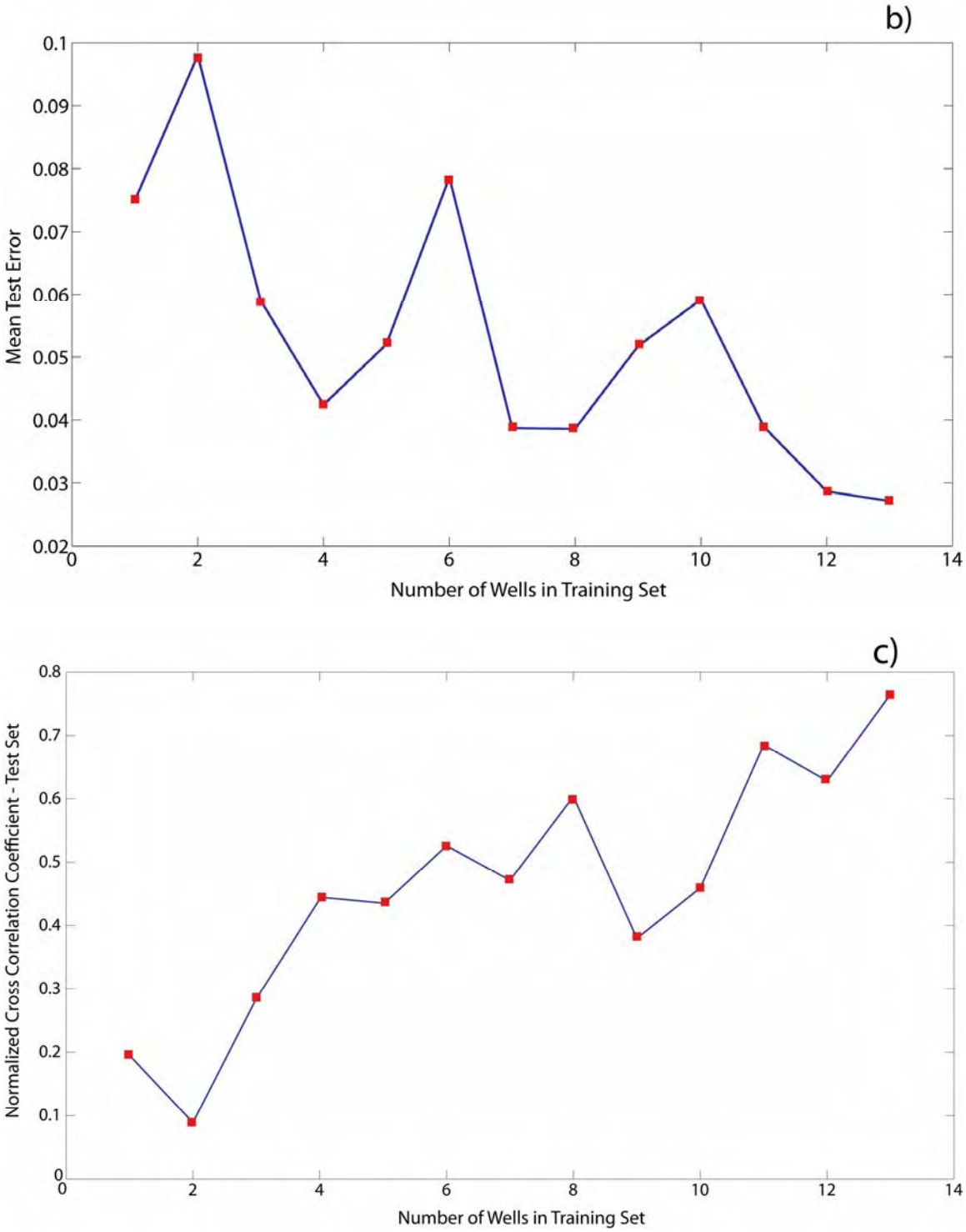
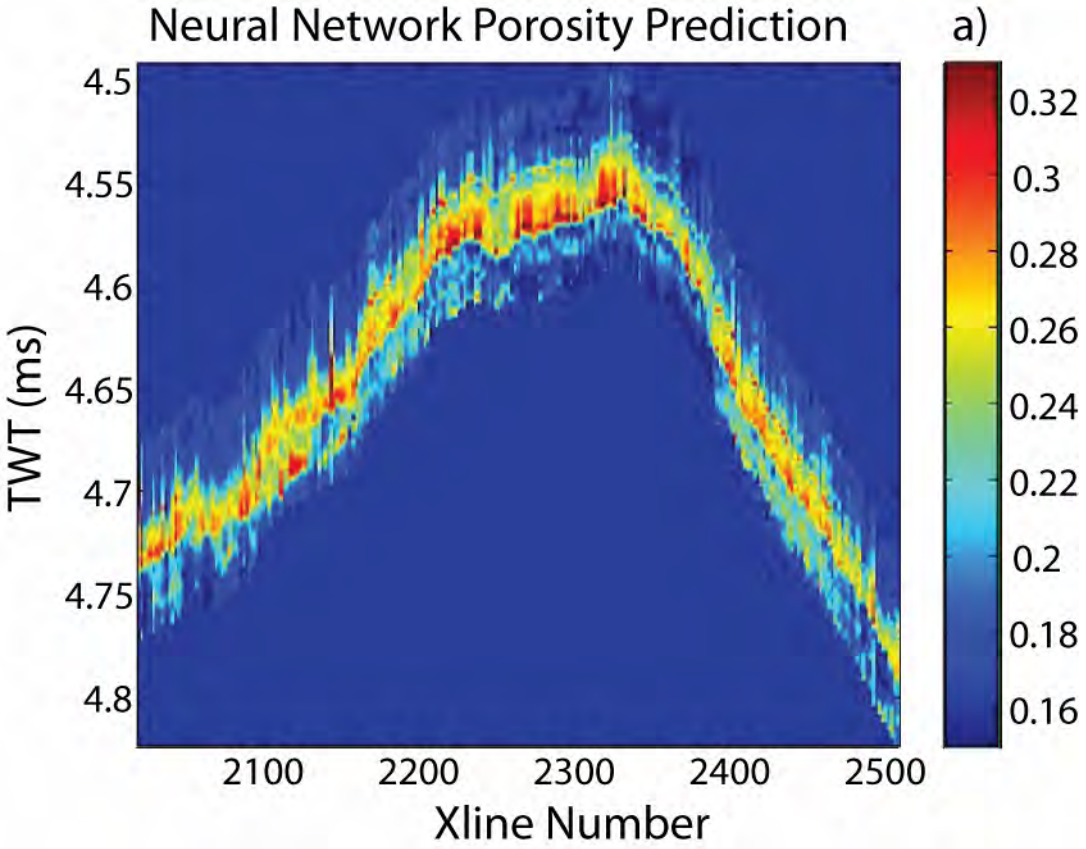


Figure 16. Performance of ANN as a function of the number of wells used in the training set. **a)** Maximum error percentage in the test case. Note the asymptotic appearance of the curve. Error

does decrease with an increasing number of wells, but the returns are diminishing. **b)** The mean testing error shows the same trend as in part **a)**. **c)** The cross correlation coefficient of the test set increases substantially with the number of wells in the training set. As the network is exposed to a greater variety of input-output pairs, it has more experience on which to call in order to make a predictive decision.



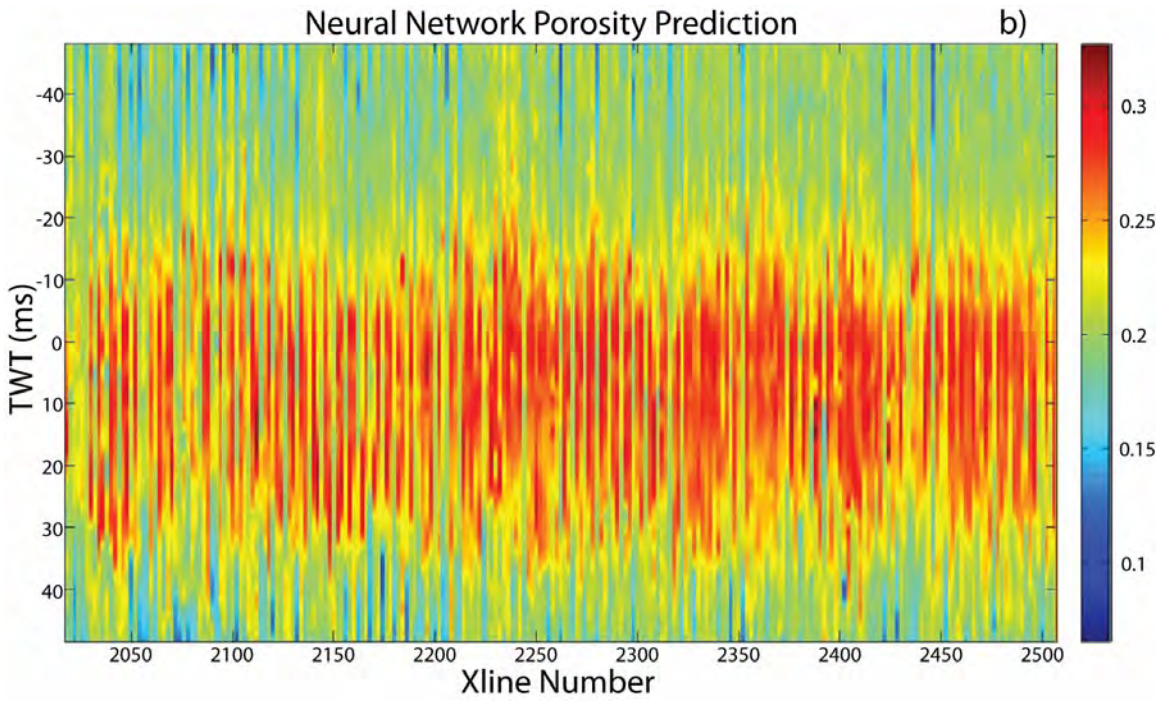


Figure 17. Results of ANN porosity prediction from a Seismic Line passing through well A04BP1 at Trace 2206. a) InLine showing the Structure of the N_m^1. b) InLine flattened on the N_m^1 horizon. Time "0" corresponds to the peak trough amplitude picked in the seismic volume which was interpreted to represent the N sand.

Mapped Results

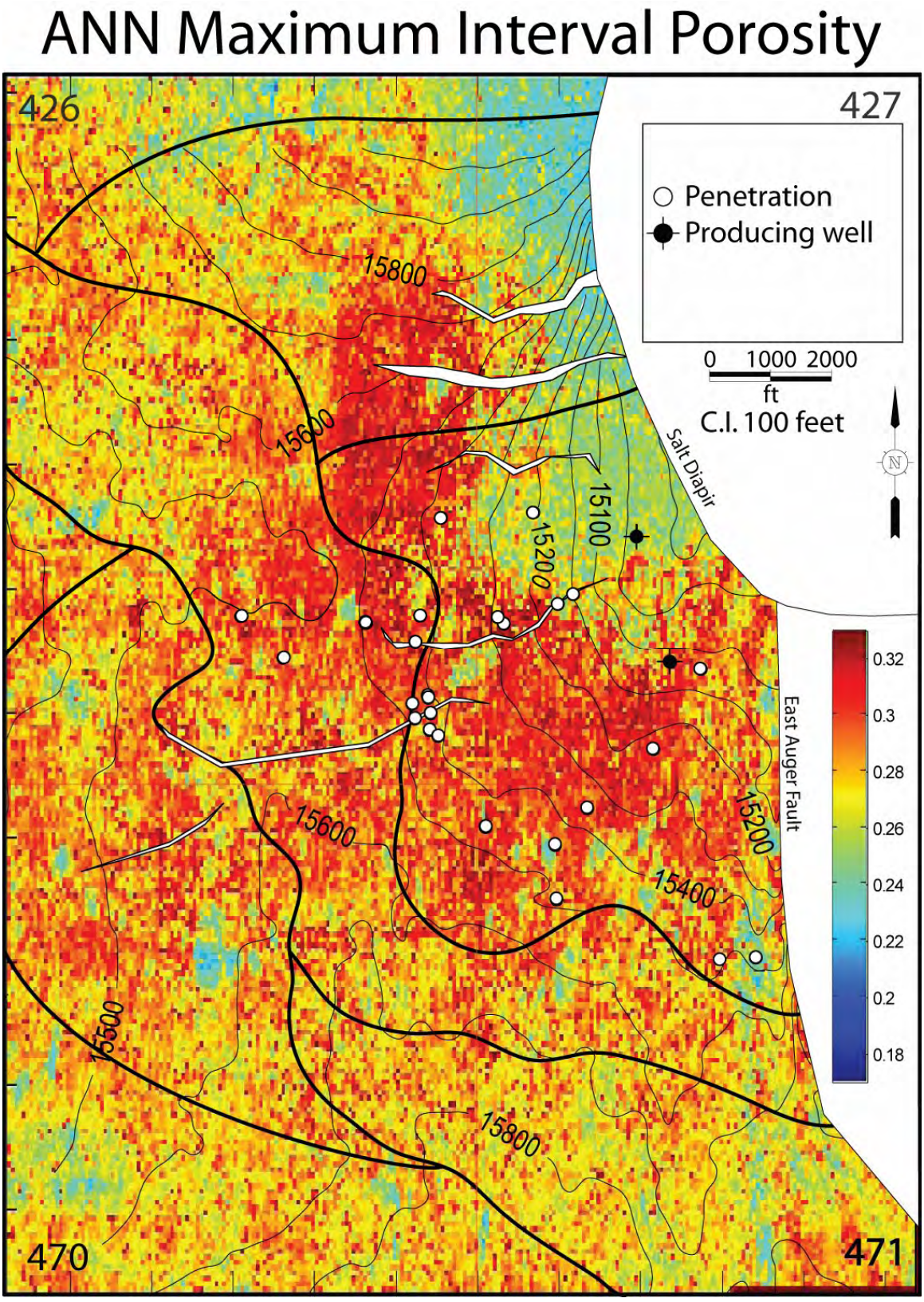
Porosity was predicted in a 100 ms interval surrounding the N_m^1. Since the sampling rate is 4 ms, there are 25 porosity samples generated for each trace. One can then construct 25 maps connecting corresponding porosity samples. Those results are shown below for predictions from both the ANN and the multilinear regression. The network and weights that were used came from the training sequence that used the constant value of porosity of 0.17 for shales, since this method was deemed to be more accurate by the sensitivity tests.

Note that labels of negative times show a map above the interval of maximum seismic amplitude corresponding to the N_m^1, while positive numbers refer to maps showing porosity values below the time of maximum amplitude in the seismic trace.

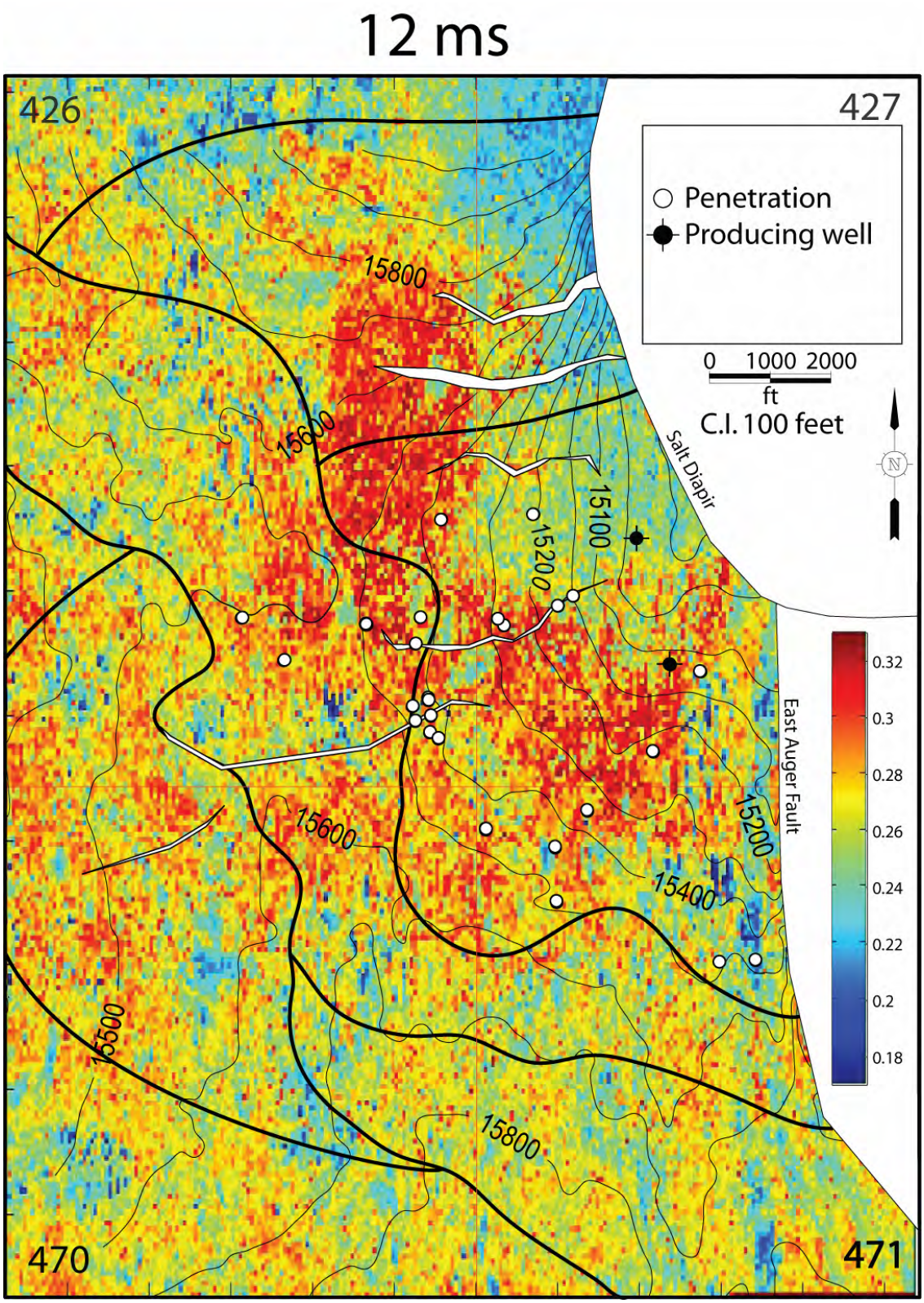
Both methods show clear indications of fan lobes. (See the two figures immediately below) The results may even suggest a fining upward facies, as the coarsest and most porous part of the sand package occurs at the bottom of the N_m^1 (+8-12 ms). This fining upward characteristic is noticed in the well logs, and is indicative of turbiditic origins.

The maps generally show an area of high porosity to the southwest of the salt diapir. This corresponds to the thickness area of the sand, as indicated by both an isochron map and well penetrations. Directly to the northeast of that, the sand thins and onlaps the salt, and the porosity decreases. We believe that these results are geologically feasible and accurate.

Figure 18. In a 100 ms interval around the N sand, 25 porosity values were predicted by multivariate linear regression that correspond to the 25 seismic samples in the interval. **a)** The maximum predicted porosity at each trace location is mapped and overlain onto the structure contours of the N_m^1. **b)** Porosity map at twelve ms, or 3 seismic samples, above the trough corresponding to the N_m^1. **c)** Porosity map at eight ms, or 2 seismic samples, above the trough corresponding to the N_m^1. **d)** Porosity map at four ms, or one seismic sample, above the trough corresponding to the N_m^1. **e)** Porosity map at the trough corresponding to the N_m^1. **f)** Porosity map at four ms, or one seismic sample, below the trough corresponding to the N_m^1. **g)** Porosity map at eight ms, or 2 seismic samples, below the trough corresponding to the N_m^1. **h)** Porosity map at twelve ms, or 3 seismic samples, below the trough corresponding to the N_m^1.

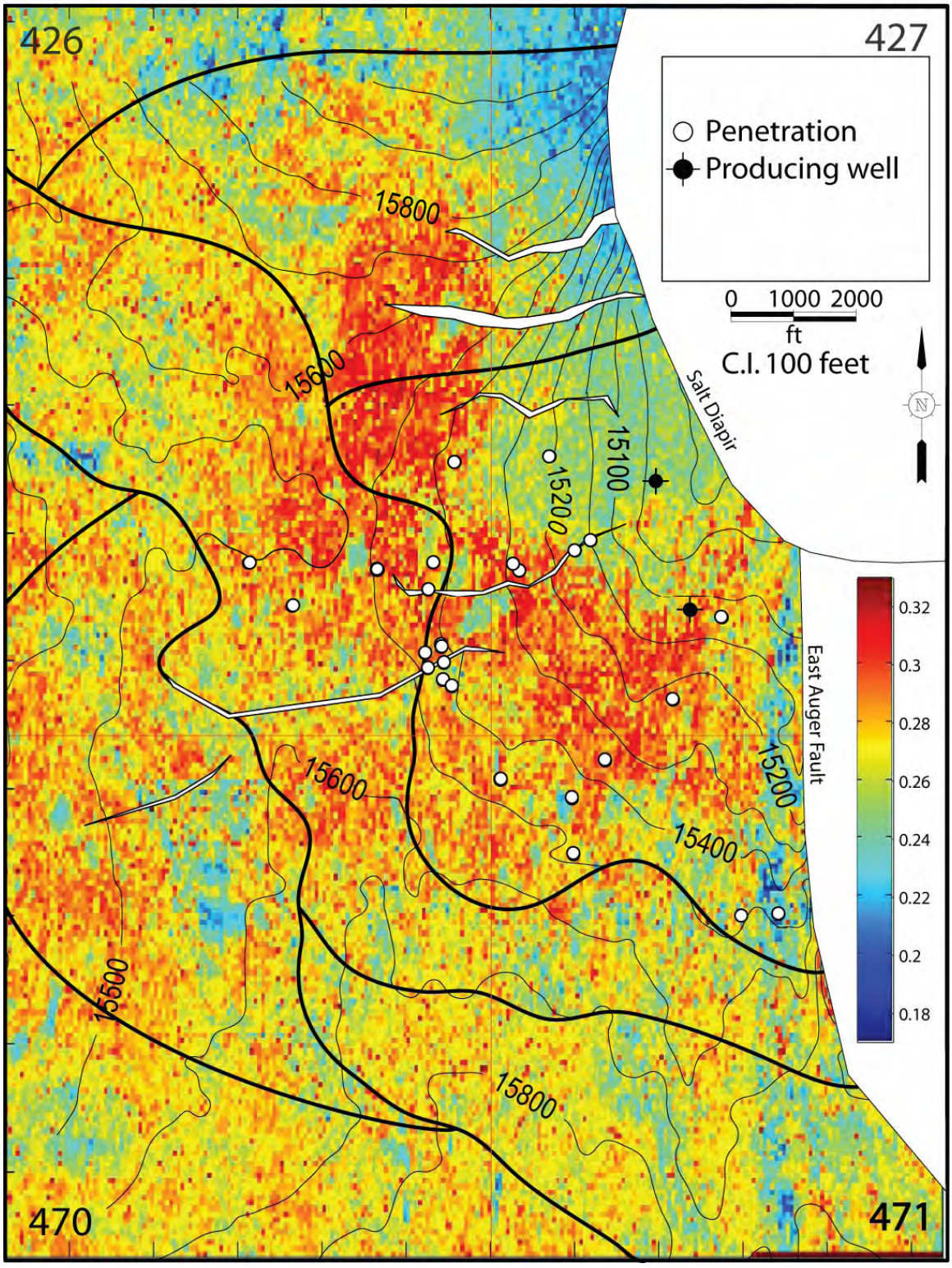

ANN Maximum Interval Porosity
426
427
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
Salt Diapir
East Auger Fault
15800
15600
15100
15200
15600
15400
15200
15500
15800
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18
470
471

18a

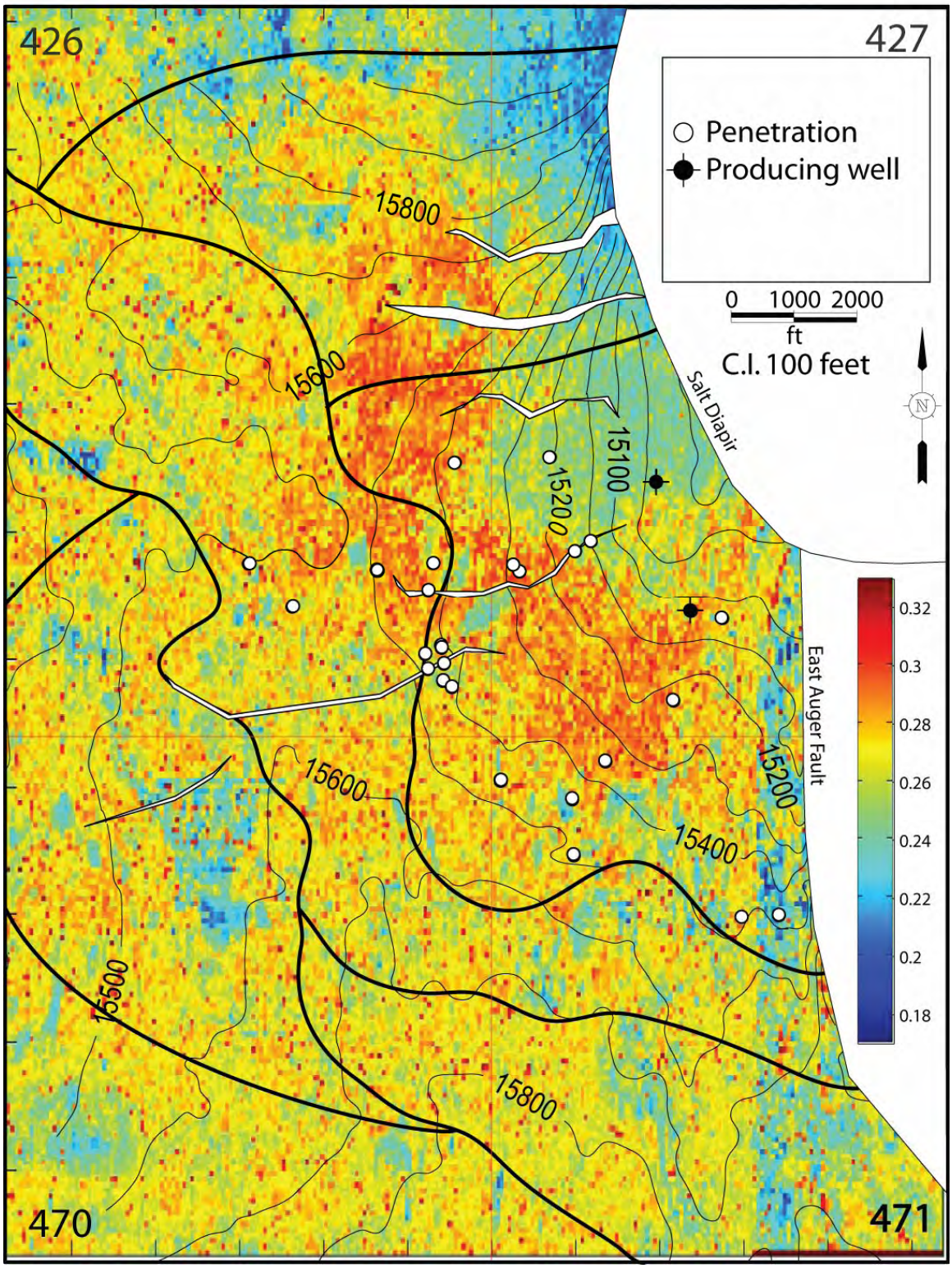

12 ms
426
427
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
Salt Diapir
East Auger Fault
15800
15600
15100
15200
15400
15500
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18
470
471

18b

8 ms

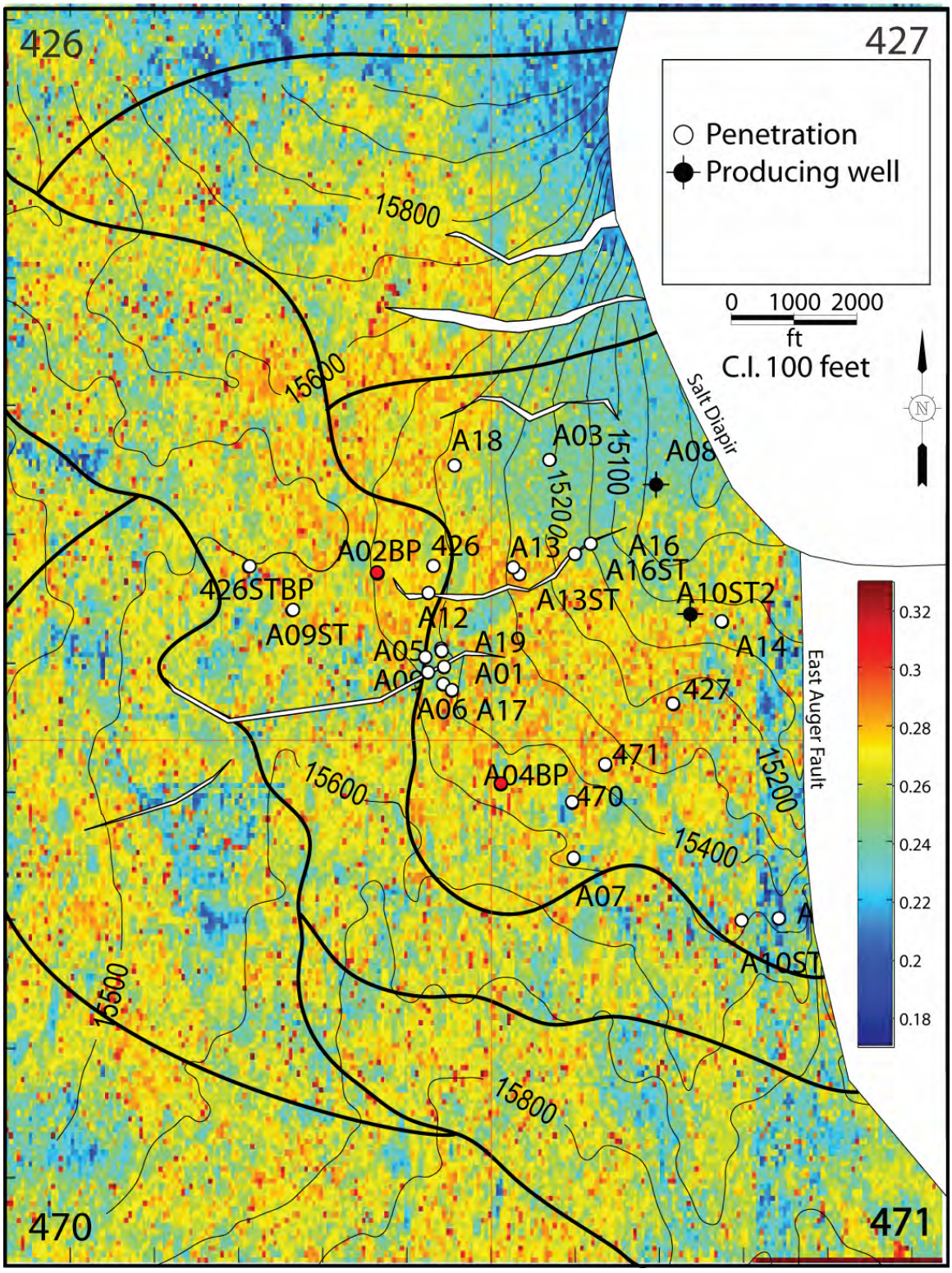

426
427
470
471
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
Salt Diapir
East Auger Fault
15800
15600
15100
15200
15400
15500
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18

18c

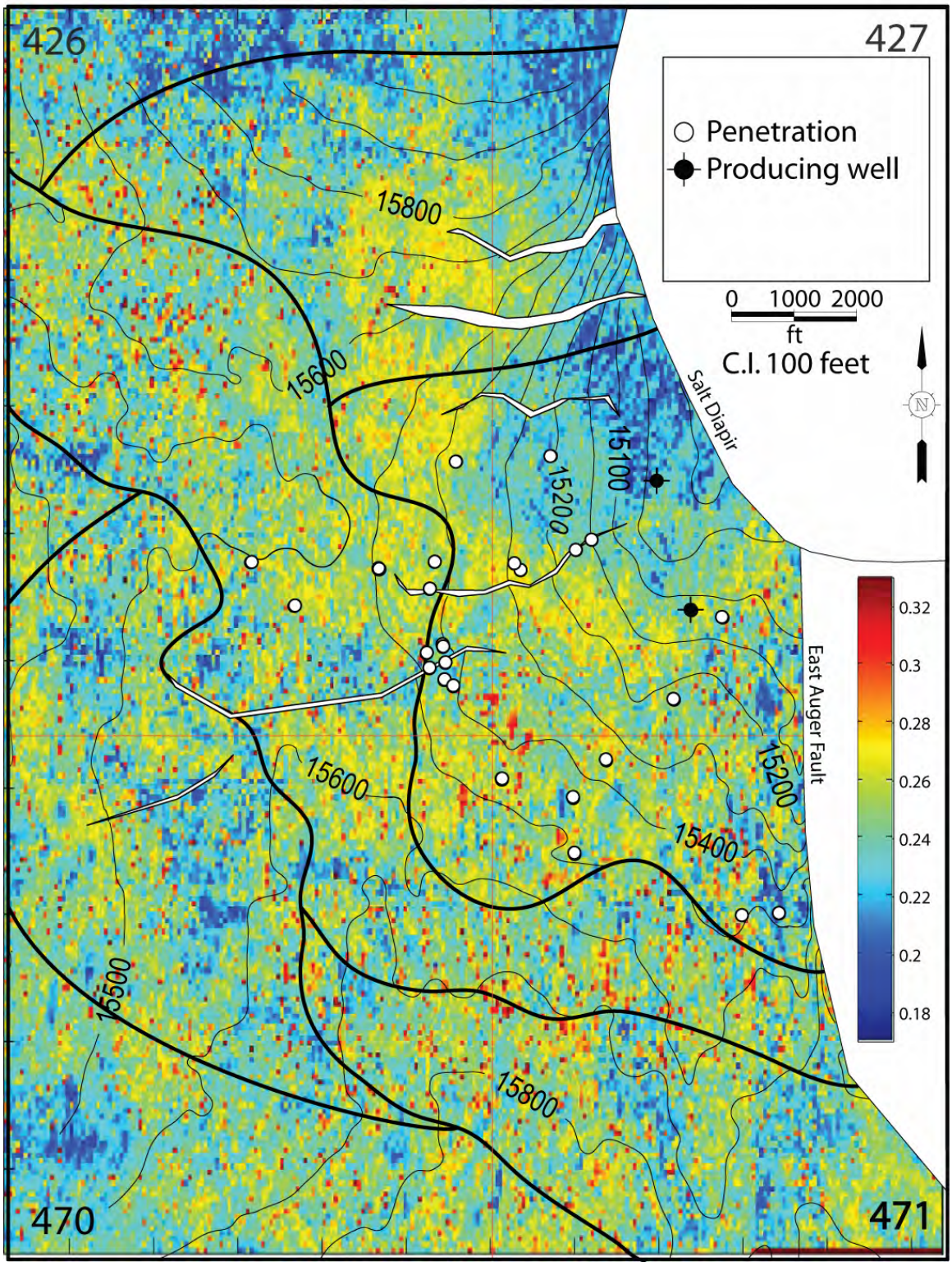

4 ms
426
427
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
15800
15600
15100
15200
Salt Diapir
East Auger Fault
15600
15200
15400
15500
15800
470
471
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18

18d

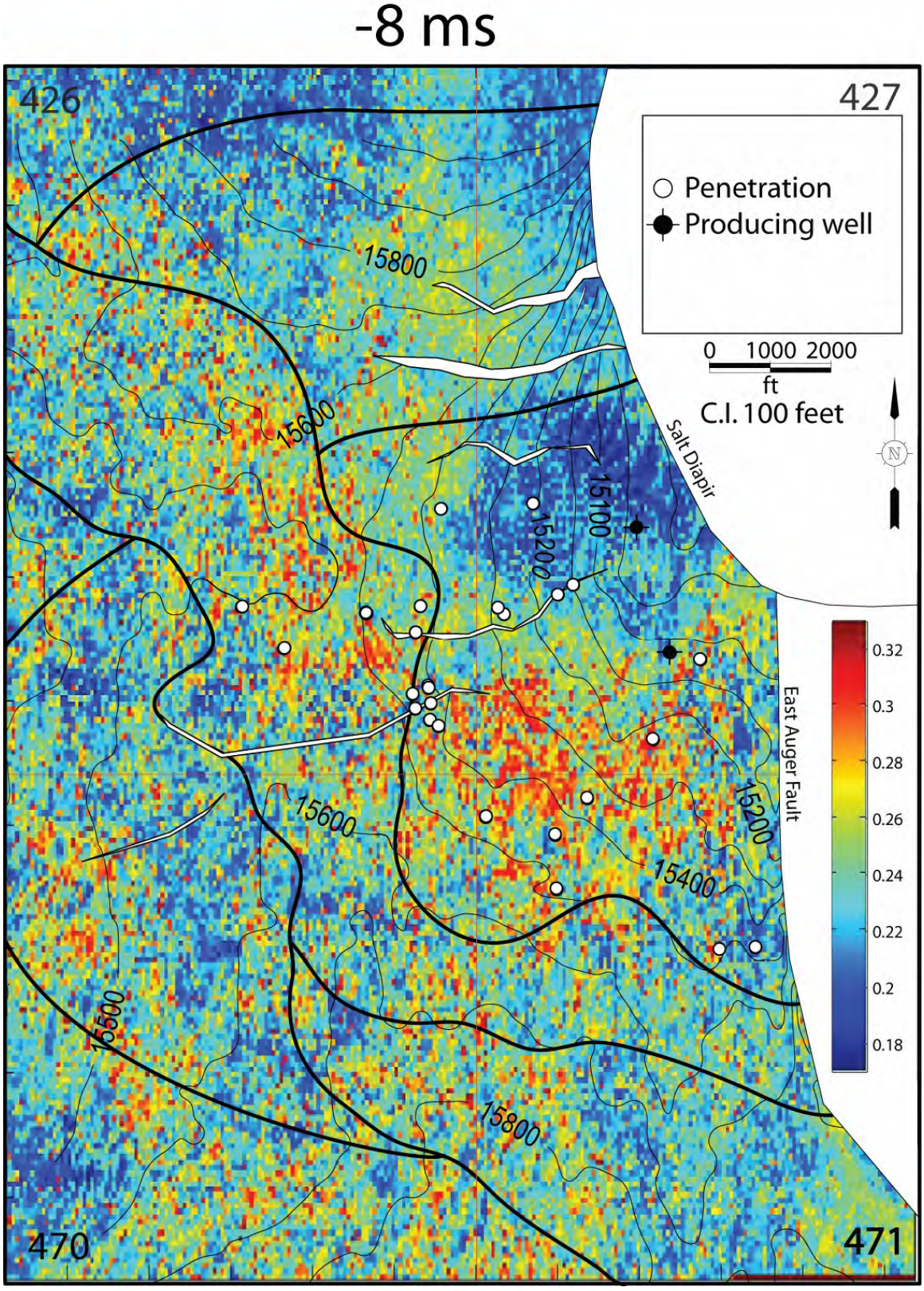
0 ms


426
427
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
15800
15600
15100
15200
Salt Diapir
A18
A03
A08
A02BP
426
A13
A16
A16ST
426STBP
A13ST
A10ST2
A09ST
A12
A19
A14
A05
A01
A17
427
471
A04BP
470
15400
A07
A10ST
15500
East Auger Fault
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18
470
471

18e

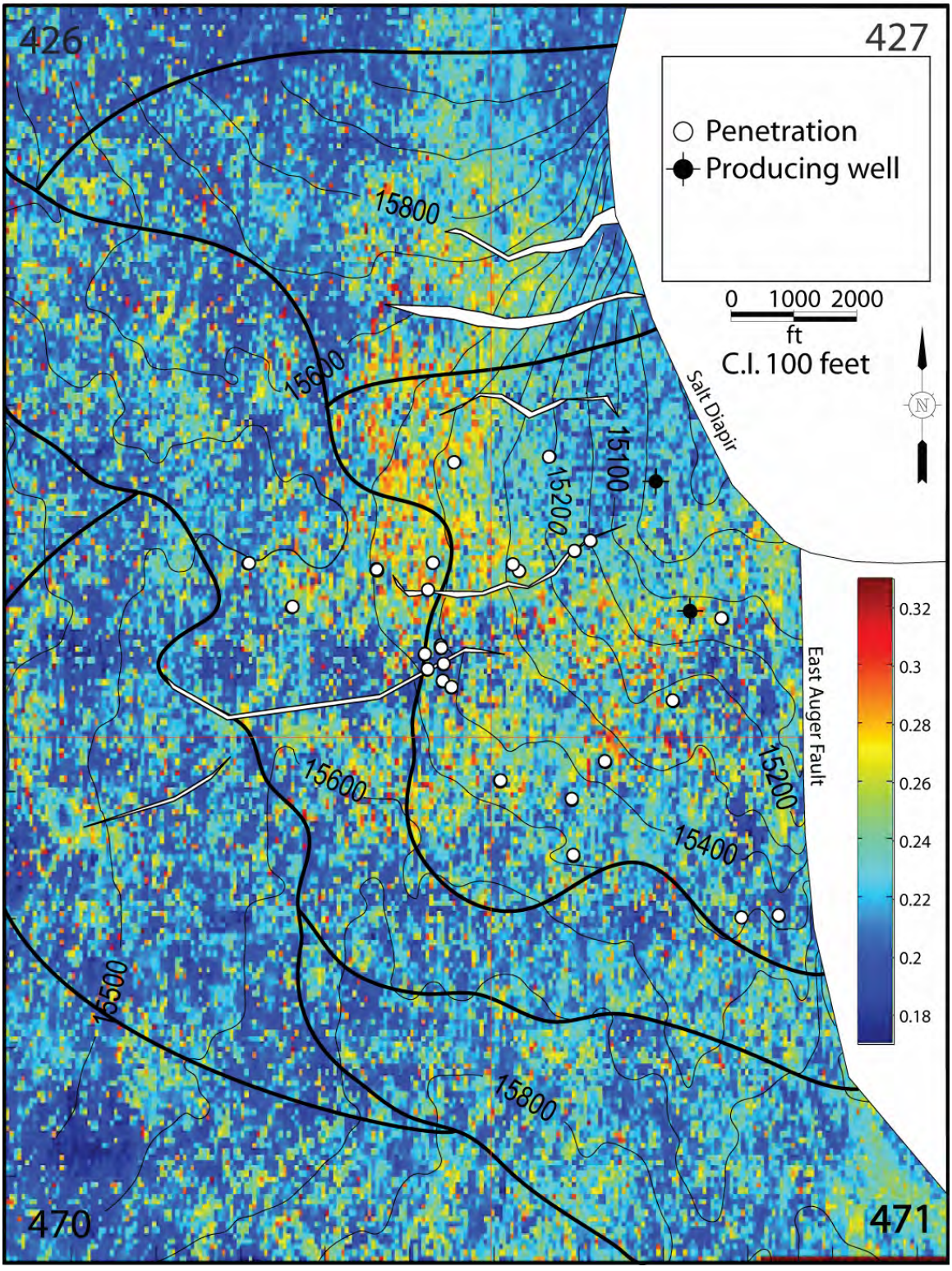

-4 ms
426
427
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
15800
15600
15100
15200
Salt Diapir
East Auger Fault
15600
15400
15200
15500
15800
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18
470
471

18f


-8 ms
426
427
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
15800
15600
15100
15200
Salt Diapir
East Auger Fault
15400
15500
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18
470
471

18g

-12 ms


426
427
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
15800
15600
15100
15200
Salt Diapir
East Auger Fault
15600
15200
15400
15600
15800
470
471
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18

18h

Multivariate Linear Regression Results with Constant Shale Porosity

Well Name	Max Error	Mean Error	CC Test	CC Training
A01	0.091	0.028	0.800	0.601
A02BP1	0.105	0.034	0.388	0.622
A04BP1	0.108	0.033	0.697	0.602
A05	0.099	0.049	0.405	0.642
A06	0.095	0.026	0.766	0.599
A07	0.098	0.041	0.734	0.595
A08	0.085	0.038	0.746	0.606
A09	0.107	0.031	0.682	0.605
A13	0.087	0.029	0.564	0.611
A14	0.076	0.029	0.373	0.622
A19	0.112	0.038	0.411	0.630
426STBP1	0.056	0.019	0.864	0.595
427	0.087	0.028	0.540	0.612
470	0.132	0.041	0.757	0.605
Averages	**0.096**	**0.033**	**0.623**	**0.610**

Table 5. Results of porosity prediction from multivariate linear regression. All wells except the one indicated were used in the training set. For example, the second row of the table contains the data for the A01. All of the wells except for the A01 were used to train the network; A01 was left out of the training as the test case. The maximum and average errors are in units of porosity (V/V). CC Test and CC Training are the normalized cross correlation coefficients between the prediction and the targets for the test and training sets, respectively.

The results of this analysis are in inline with past studies [16-18]. The test sets had on average a cross correlation with the target sets of around 0.62. This is considerably lower than the results from the ANN. Although the number is low, there is still good visual agreement between the estimates and targets. The average error is also quite low at 0.033 porosity units.

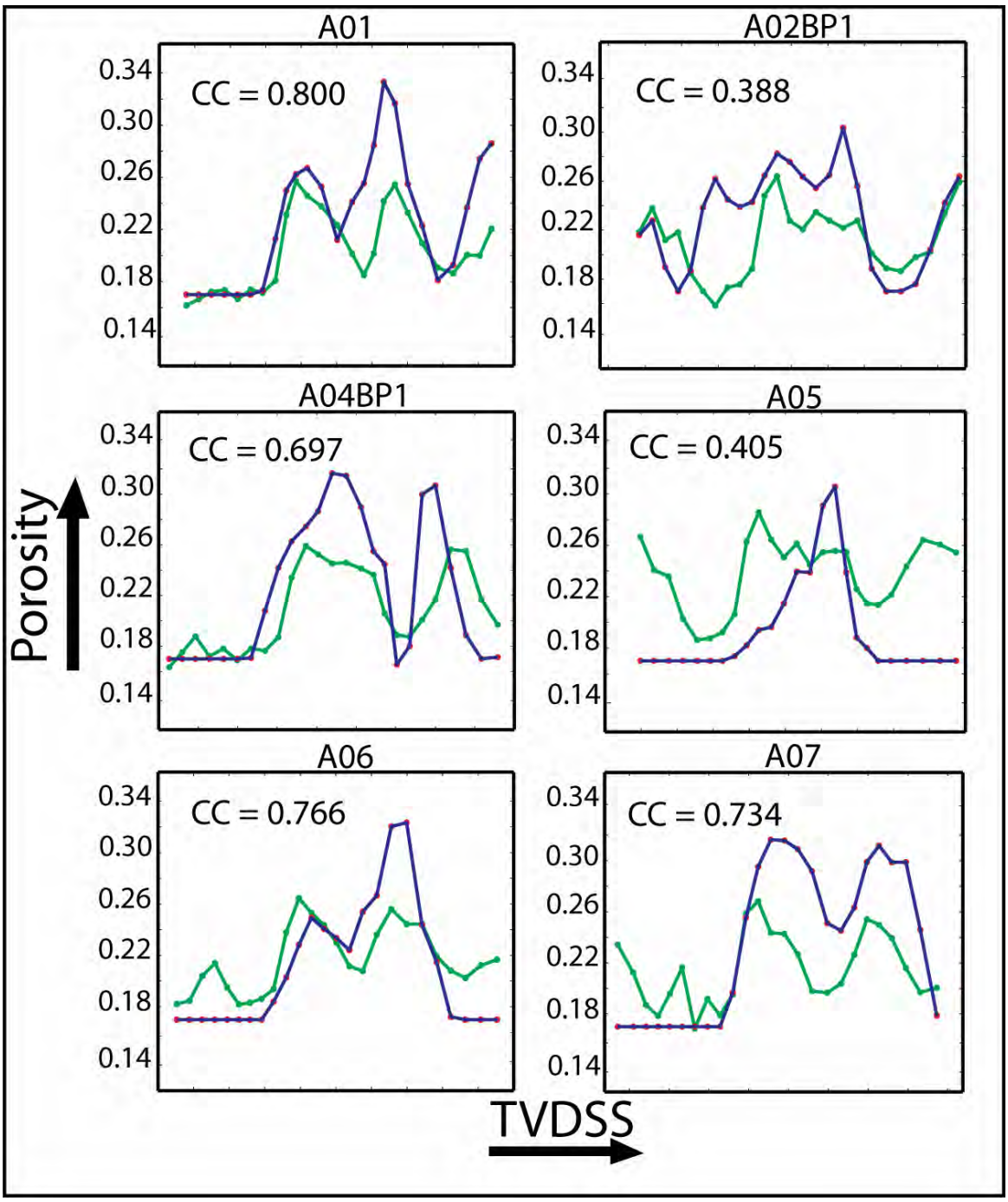
The weights derived through analysis are:

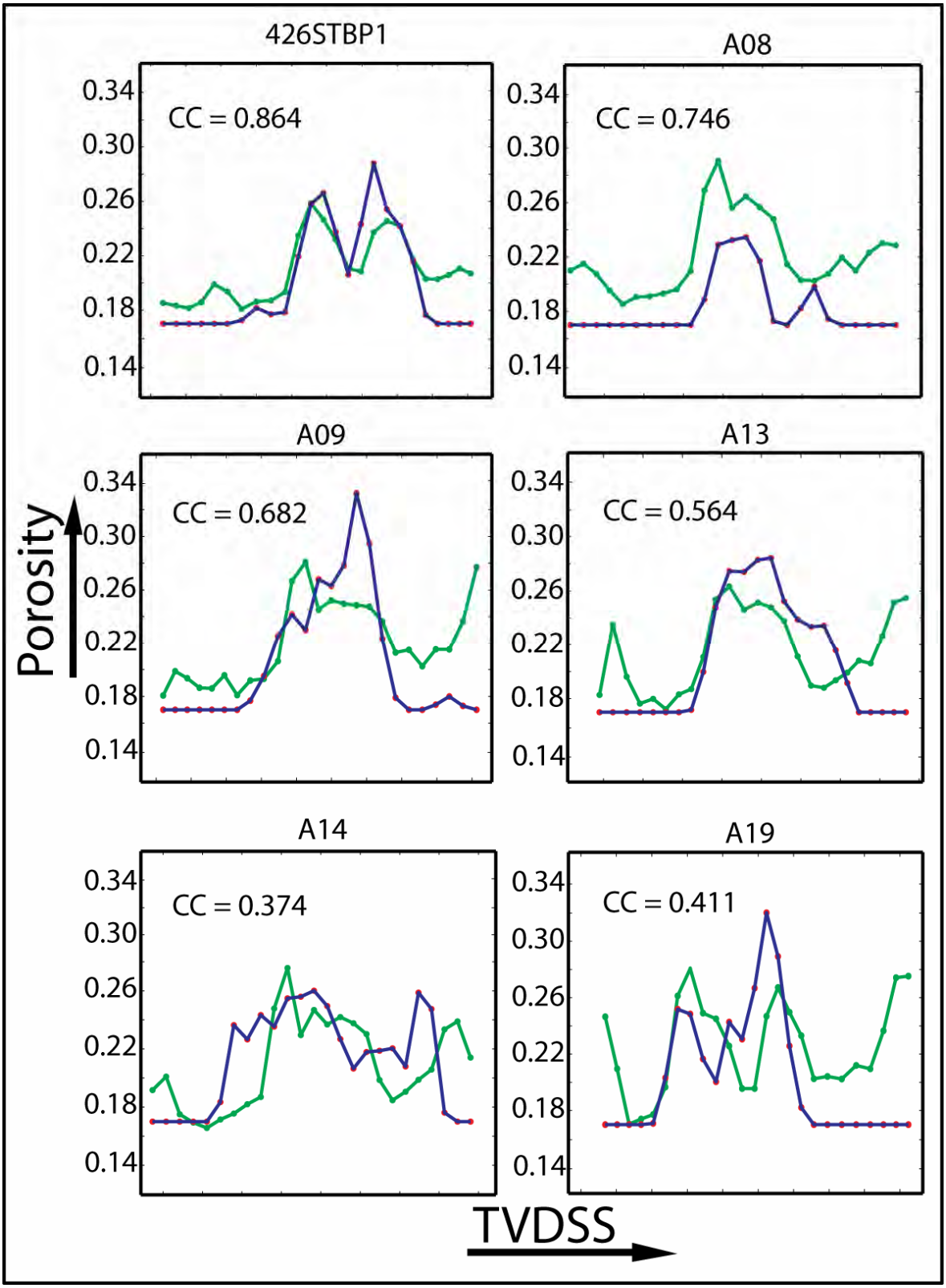
$$\begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.156 \\ -0.183 \\ 0.426 \\ 0.123 \\ 0.133 \\ 0.0142 \\ -0.172 \end{bmatrix}$$

Keep in mind that these weights operate on normalized and transformed inputs.

With these weights having been derived, they were used to transform the entire seismic volume into a porosity volume 100 ms, or 25 samples, in length around the $N_m{}^1$ horizon.

The first map shows the maximum porosity value at each trace position. Although it predicts lower porosity values than the ANN, both methods show the same trends. Namely, that porosity is highest in a northwest-southeast trending belt just to the southwest of the salt diapir.


A01
CC = 0.800
A02BP1
CC = 0.388
A04BP1
CC = 0.697
A05
CC = 0.405
A06
CC = 0.766
A07
CC = 0.734
0.34
0.30
0.26
0.22
0.18
0.14
Porosity
TVDSS


426STBP1
CC = 0.864
A08
CC = 0.746
A09
CC = 0.682
A13
CC = 0.564
A14
CC = 0.374
A19
CC = 0.411
0.34
0.30
0.26
0.22
0.18
0.14
Porosity
TVDSS

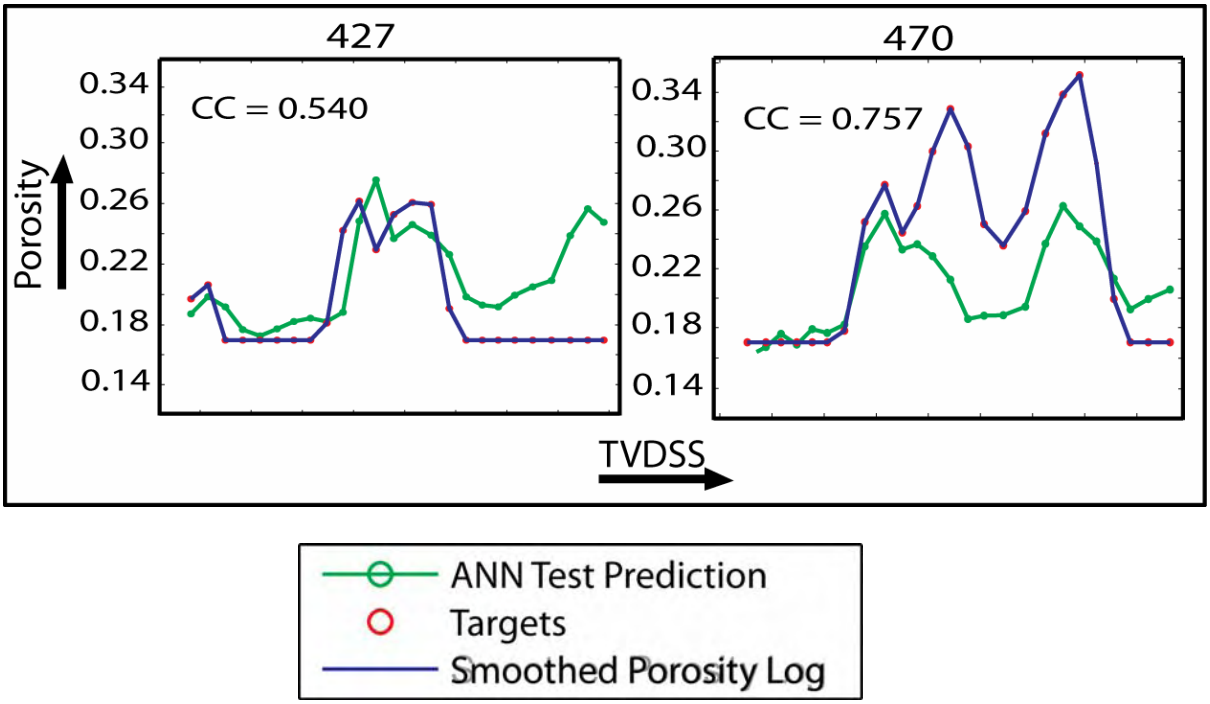


Figure 19. Test wells from the training sequence of the Multivariate Linear Regression. The CC value indicates the normalized cross-correlation coefficient between the neural network output and the target values.

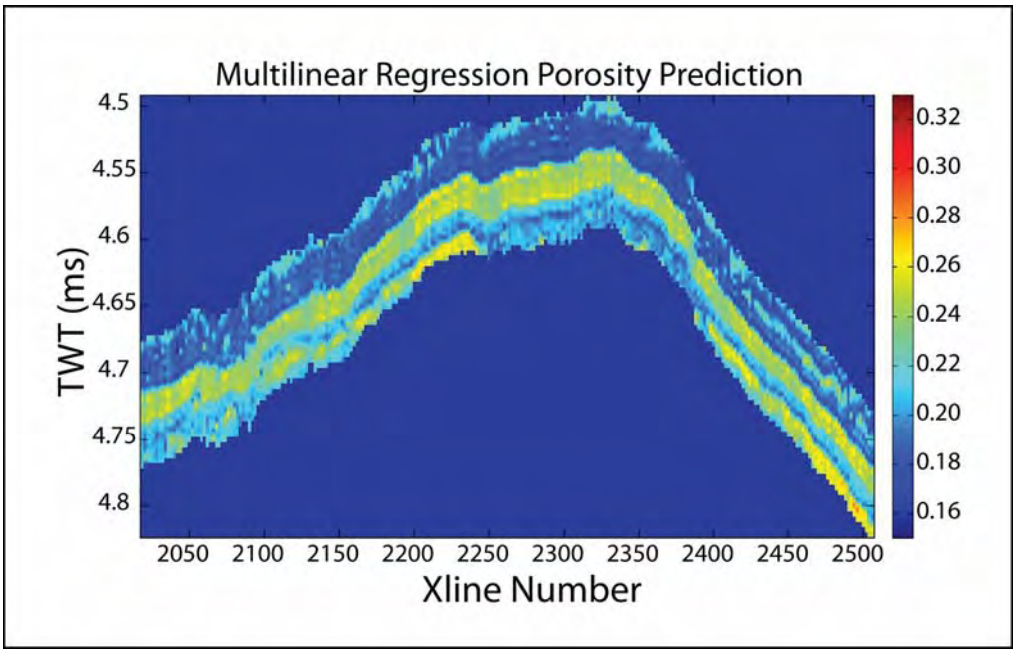
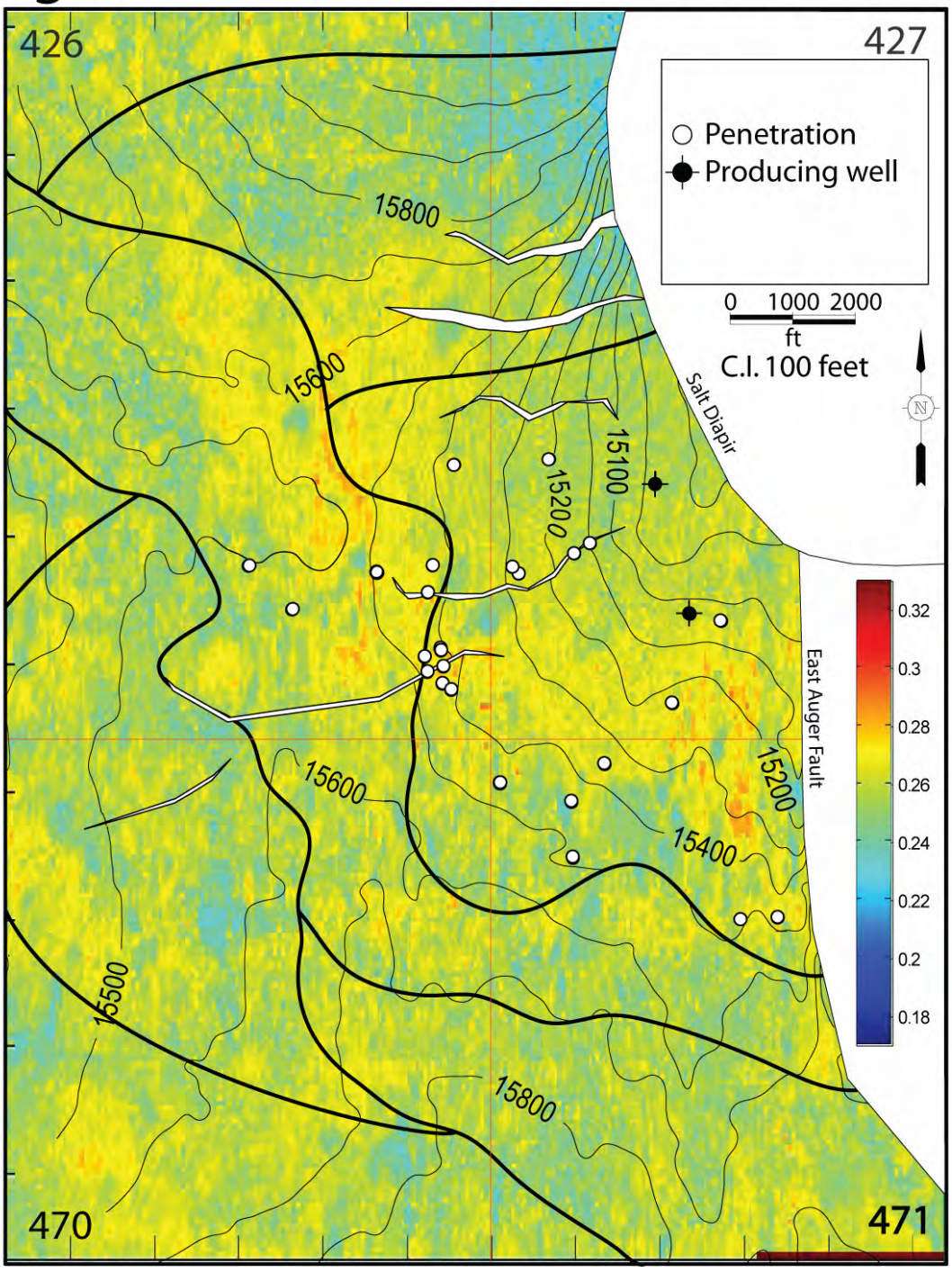


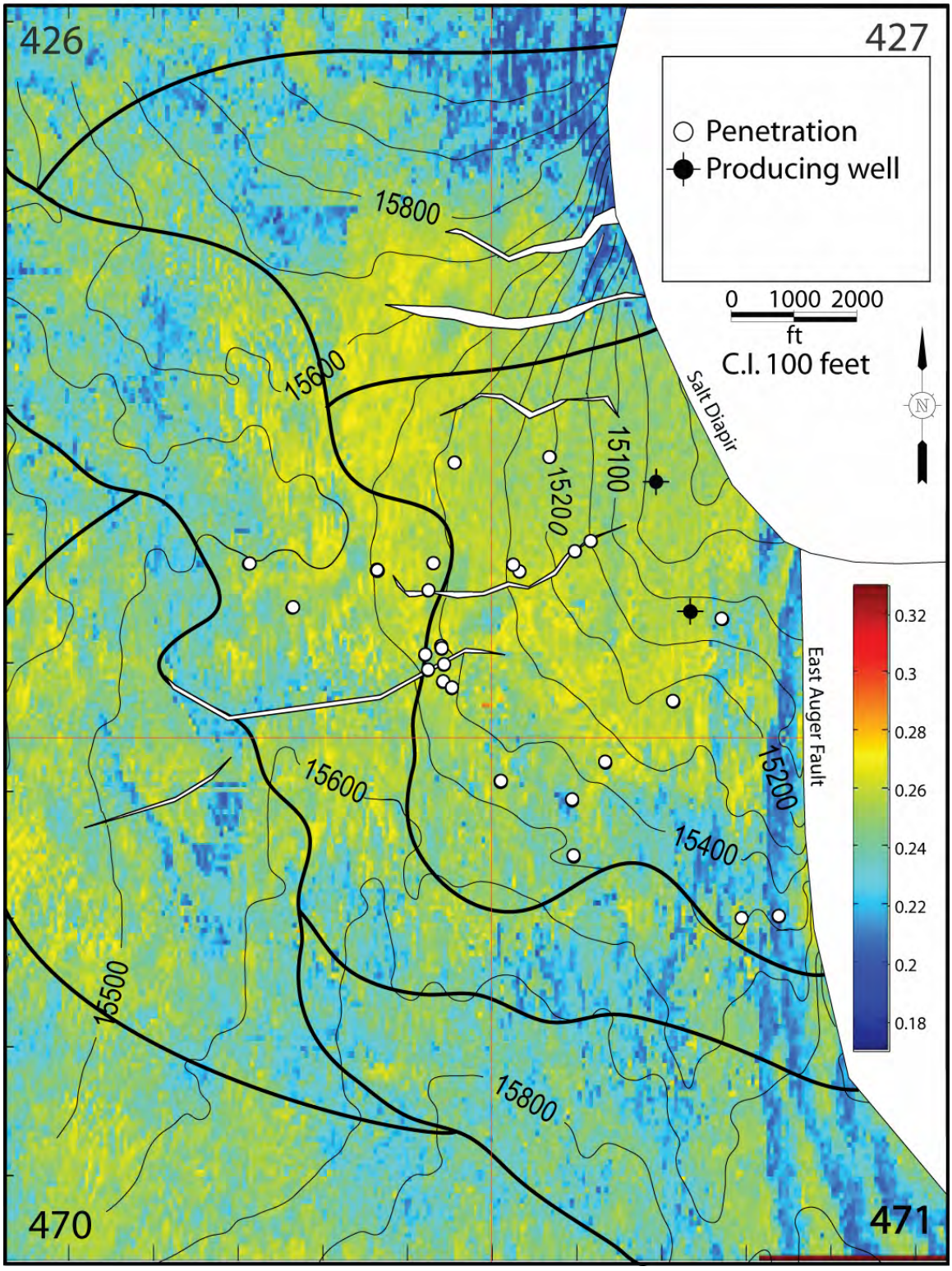
Figure 20. Results of the Regression porosity prediction from a Seismic Line passing through well A04BP1 at Trace 2206. a) InLine showing the Structure of the N_m^1.

Figure 21, Pages 60 -67. In a 100 ms interval around the N sand, 25 porosity values were predicted by multivariate linear regression that correspond to the 25 seismic samples in the interval. **a)** The maximum predicted porosity at each trace location is mapped and overlain onto the structure contours of the N_m^1. **b)** Porosity map at twelve ms, or 3 seismic samples, above the trough corresponding to the N_m^1. **c)** Porosity map at eight ms, or 2 seismic samples, above the trough corresponding to the N_m^1. **d)** Porosity map at four ms, or one seismic sample, above the trough corresponding to the N_m^1. **e)** Porosity map at the trough corresponding to the N_m^1. **f)** Porosity map at four ms, or one seismic sample, below the trough corresponding to the N_m^1. **g)** Porosity map at eight ms, or 2 seismic samples, below the trough corresponding to the N_m^1. **h)** Porosity map at twelve ms, or 3 seismic samples, below the trough corresponding to the N_m^1.

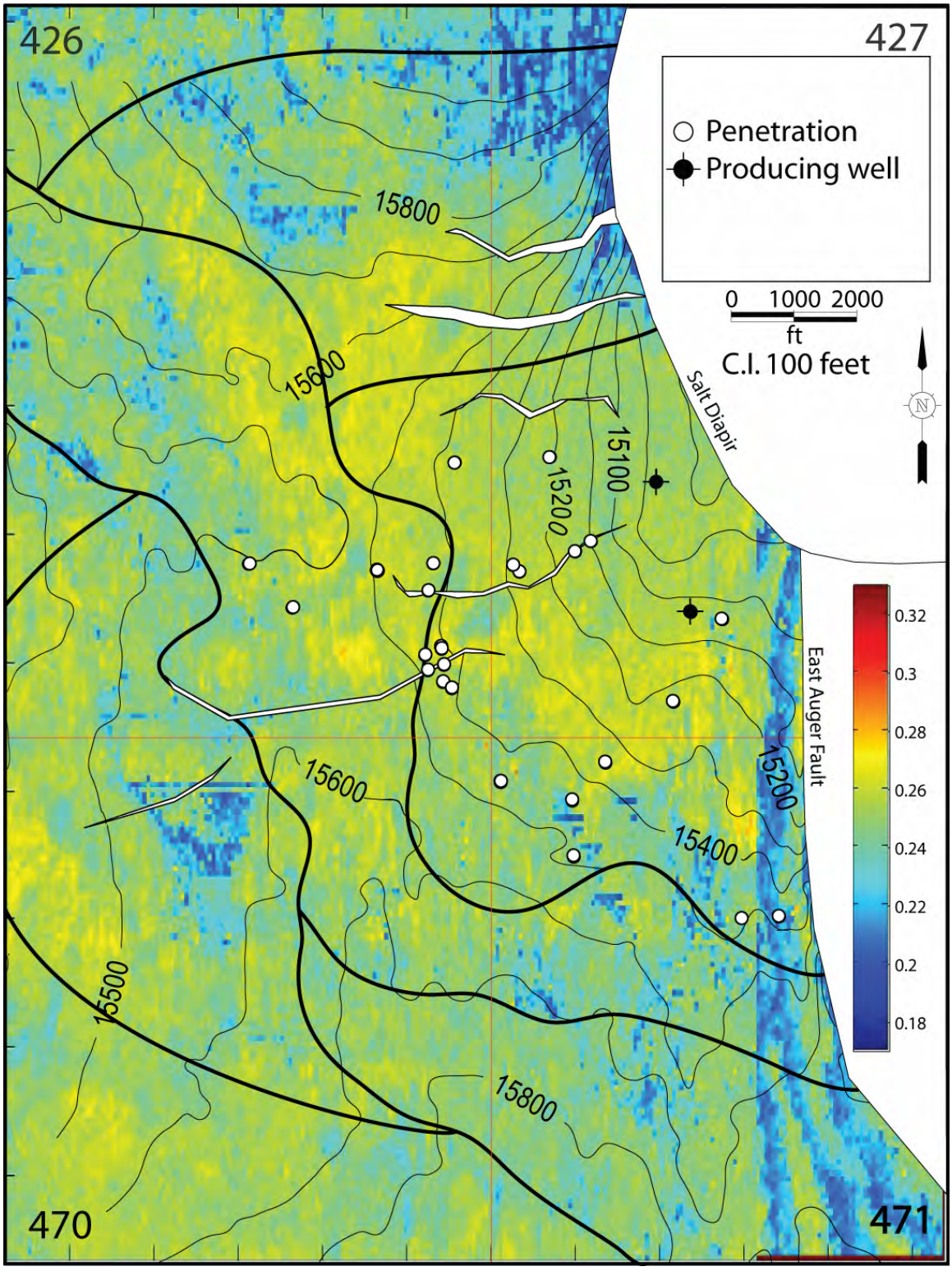
Regression Maximum Interval Porosity



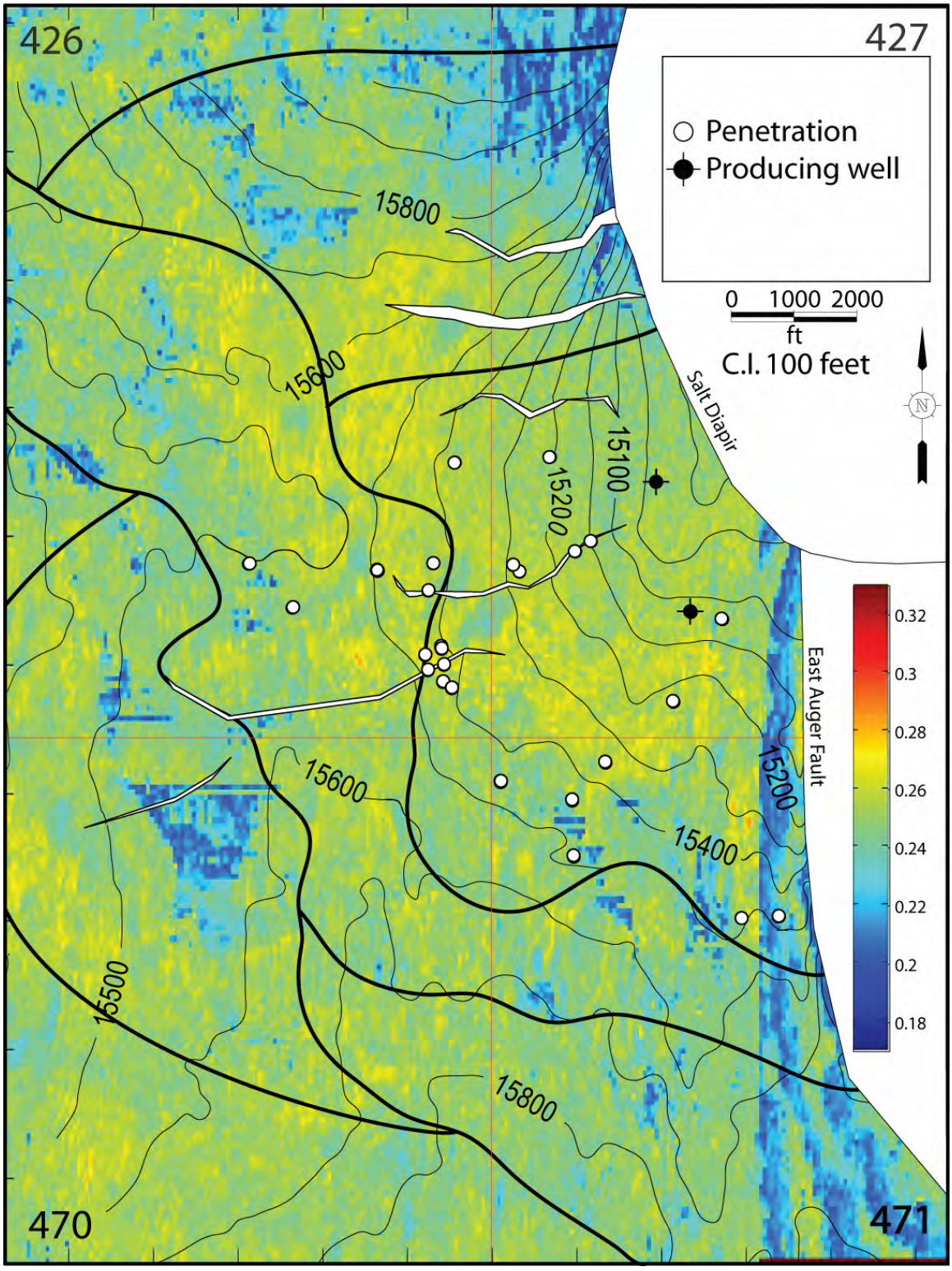
12 ms


426
427
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
Salt Diapir
East Auger Fault
15800
15600
15100
15200
15600
15400
15200
15600
15800
470
471
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18

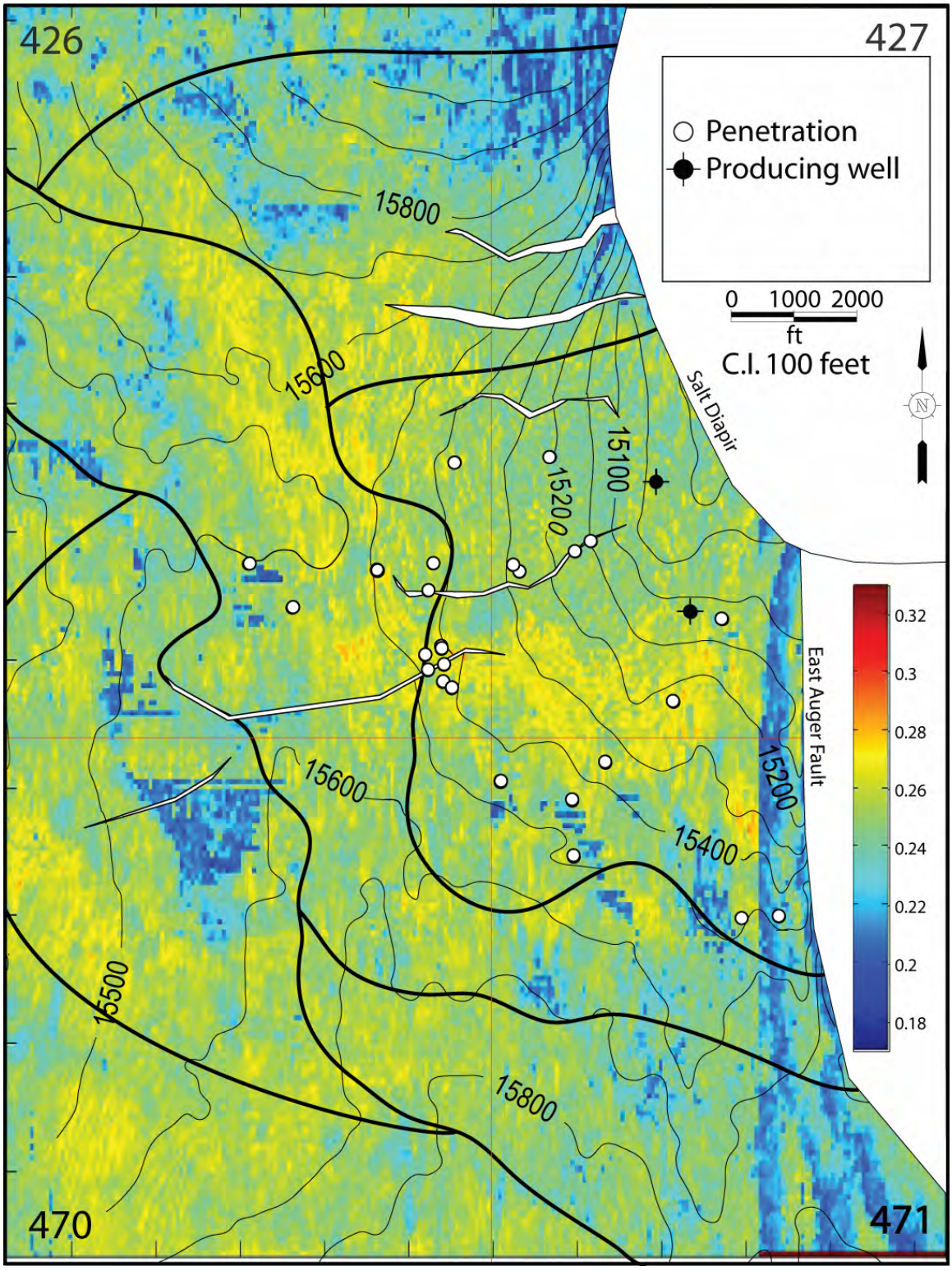
8 ms


426
427
470
471
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
Salt Diapir
East Auger Fault
15800
15600
15100
15200
15600
15400
15500
15800
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18

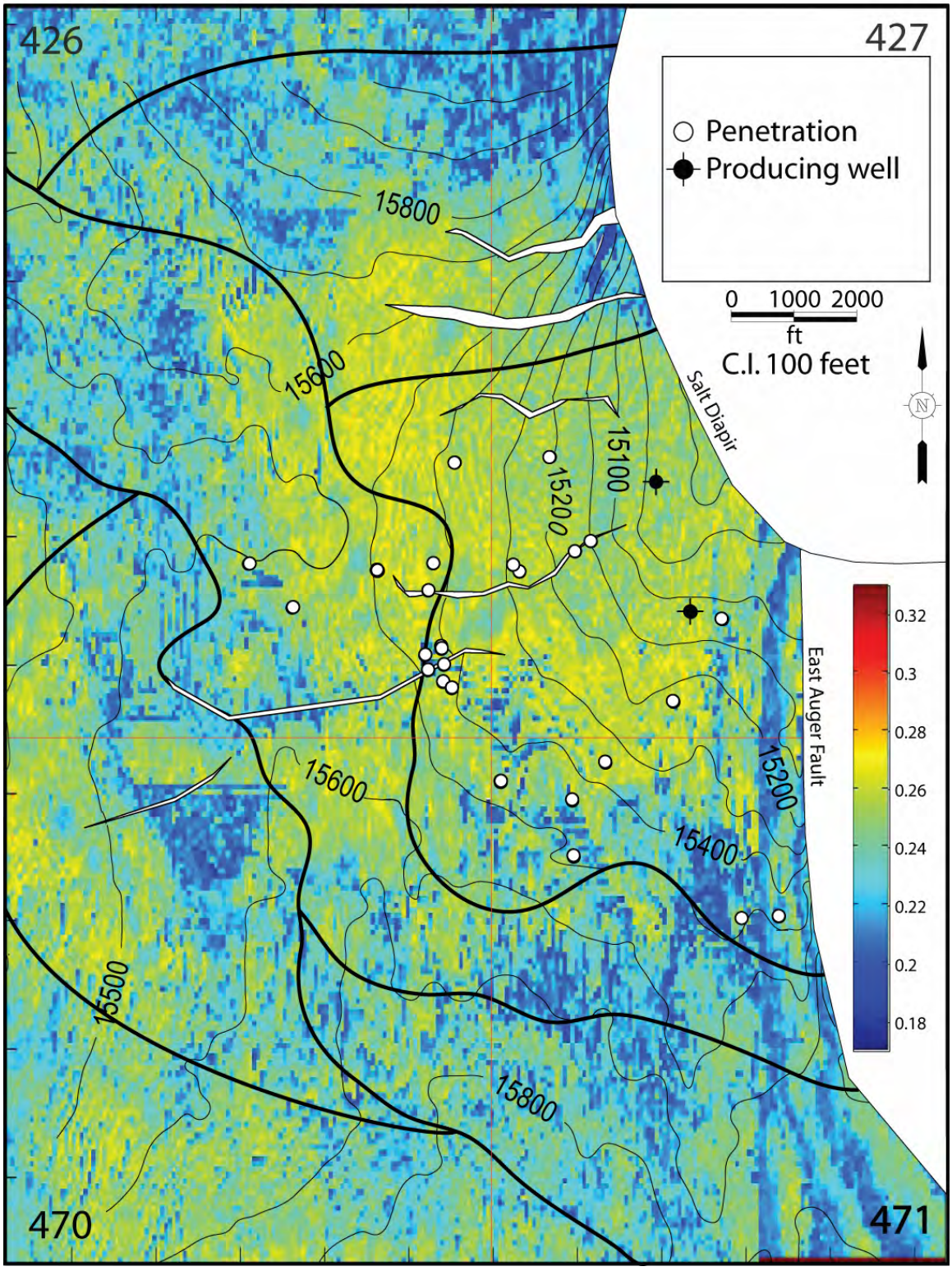
4 ms


426
427
470
471
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
Salt Diapir
East Auger Fault
15800
15600
15100
15200
15600
15400
15200
15600
15800
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18

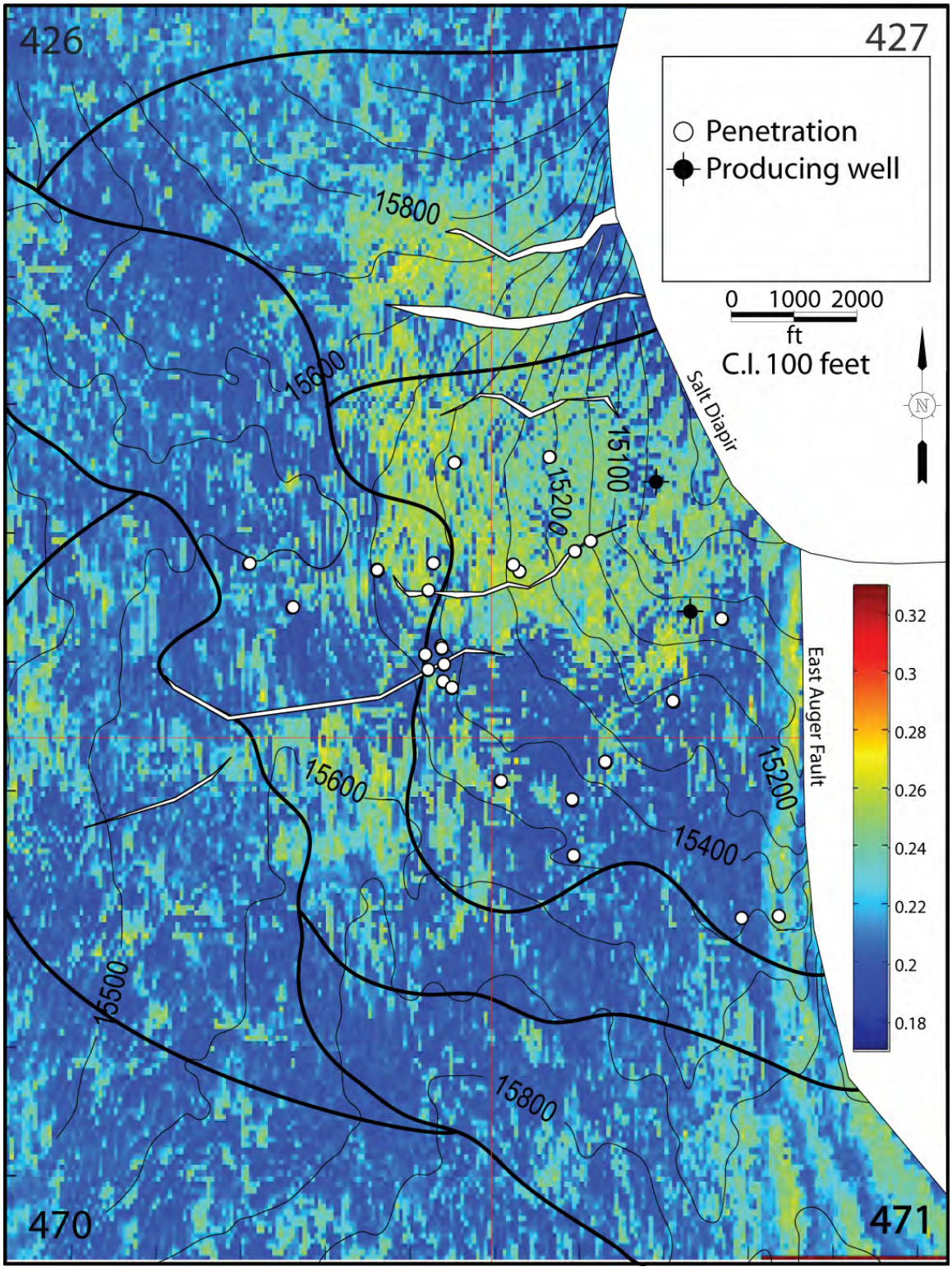
0 ms


426
427
470
471
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
Salt Diapir
East Auger Fault
15800
15600
15100
15200
15600
15400
15500
15800
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18

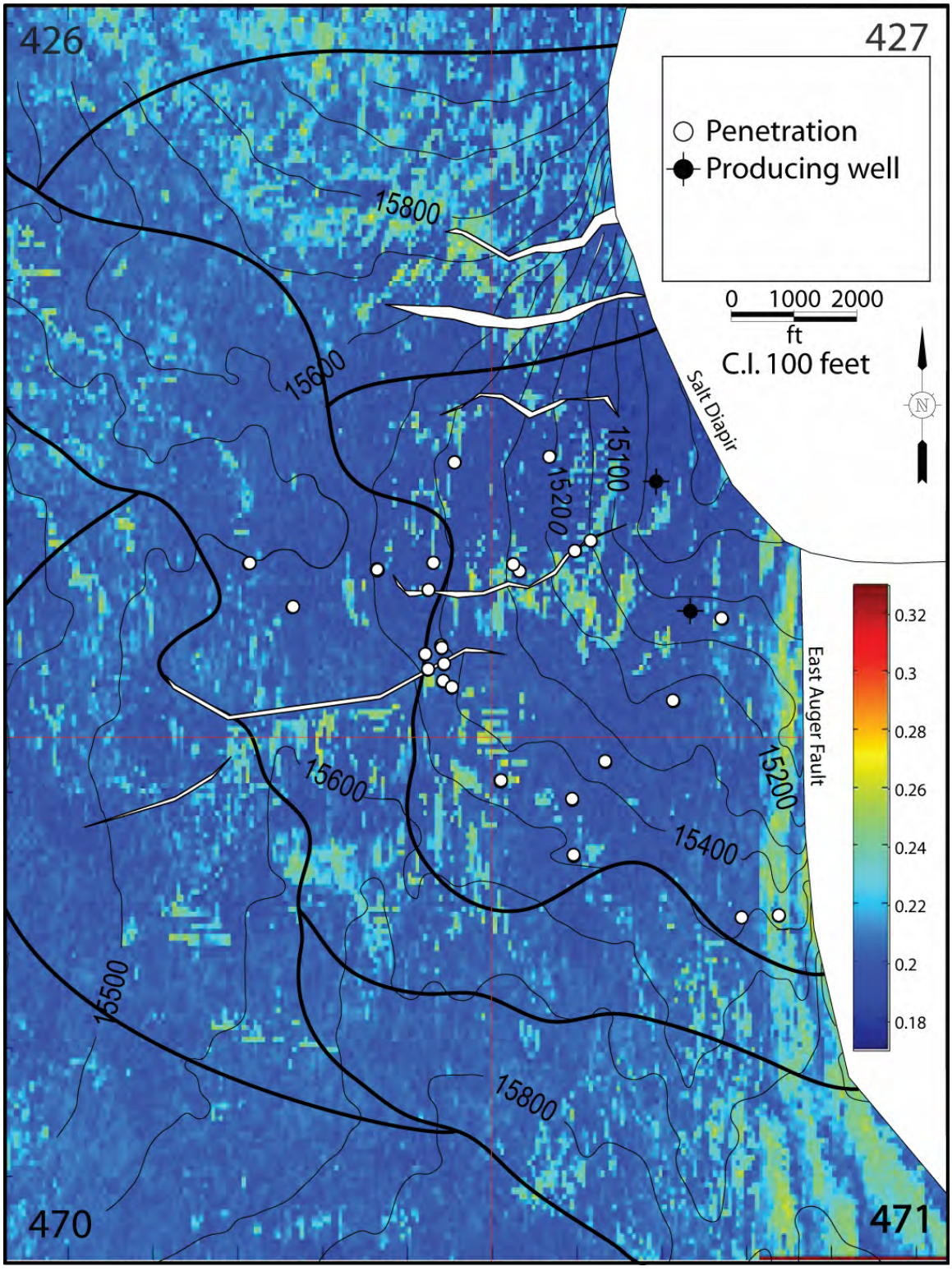
-4 ms


426
427
470
471
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
Salt Diapir
East Auger Fault
15800
15600
15100
15200
15400
15600
15600
15800
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18

-8 ms


426
427
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
Salt Diapir
East Auger Fault
15800
15600
15500
15400
15200
15100
15200
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18
470
471

-12 ms

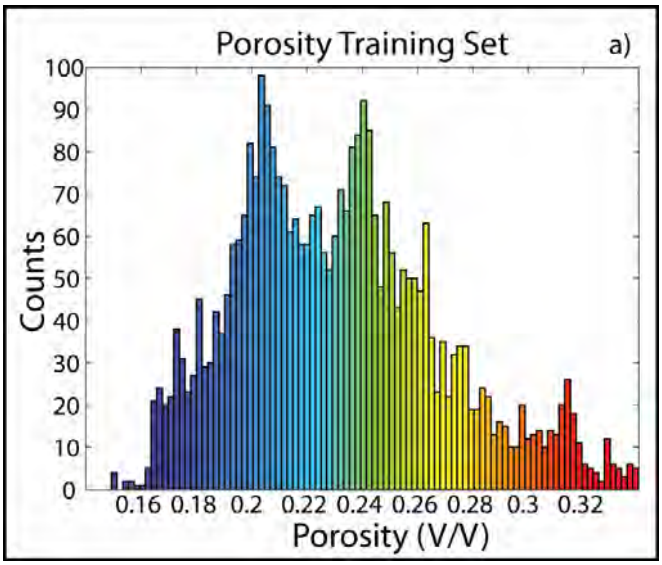

426
427
470
471
Penetration
Producing well
0 1000 2000
ft
C.I. 100 feet
N
Salt Diapir
East Auger Fault
15800
15600
15100
15200
15600
15200
15400
15500
15800
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18

CONCLUSIONS

Prediction of seismic porosity from seismic data improves reservoir characterization by estimating the rock property away from well control. This improved image of the reservoir allows for more advanced and accurate reservoir models, performance prediction, and eventually to more advantageous placement of future production wells.

We have investigated two methods for the prediction of porosity from multiple 3D seismic attributes: Artificial Neural Network estimation, and Multivariate Linear Regression. The valor of these two methods has been demonstrated on a data set from an amalgamated turbidite gas reservoir in the deepwater of the Gulf of Mexico.

Overall, the ANN predicted a greater range of porosity values than did multilinear regression (Figure 22). The ANN also outperformed the regression analysis considerably, both in average error, but also in the test cross correlation coefficient values. Although more computationally expensive, the ANN prediction was on average 10.4% more similar to the target than was the regression prediction.



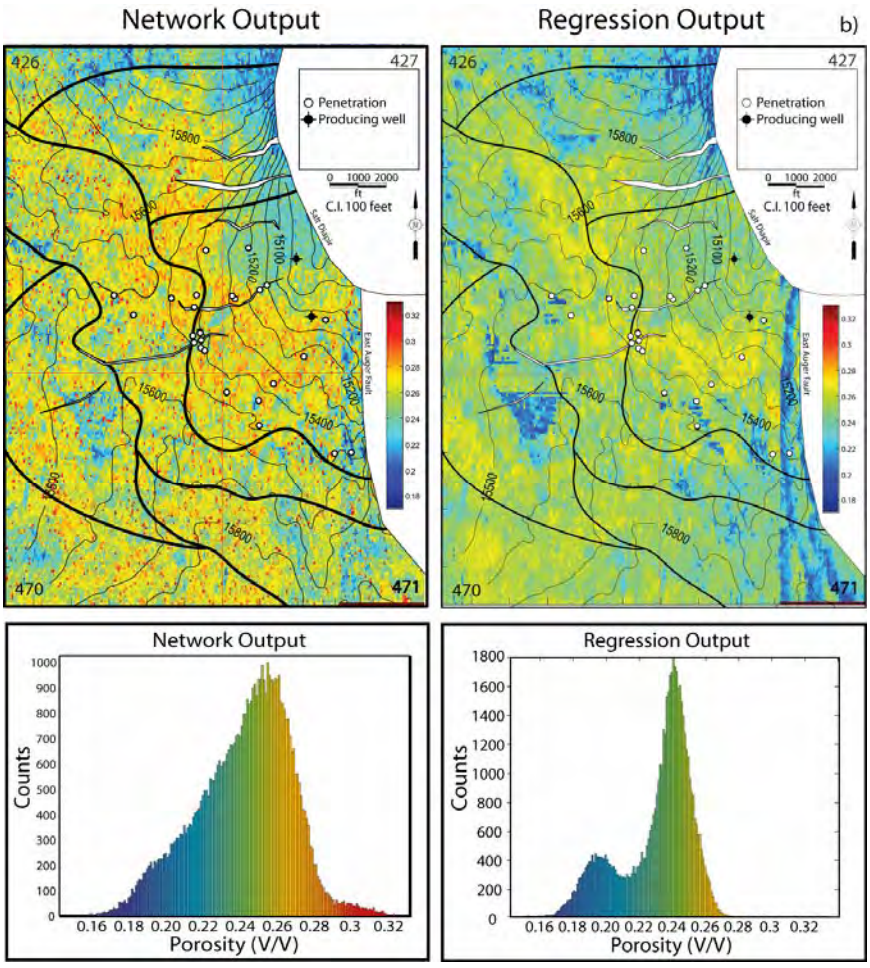


Figure 22. **a)** Histogram showing the distribution of porosity values in the training set. Note the range of values extends to the upper limit of 0.33. **b)** Histograms showing the distribution of predicted porosity values in maps shown above, which are correspond to Figures 18e and 22e. The results given by the neural network demonstrate roughly the same distribution as the porosity training set. The results from regression analysis, however, lack the high-porosity information.

The several advantages of these new algorithms over the conventional inversion methods are that: they predict logs other than acoustic impedance; they may use seismic attributes other than the conventional post-stack volume; they do not rely on any

particular forward model; knowledge of the seismic wavelet is not required; they may enhance resolution; and they use testing and cross-validation as a measure of success [1, 16-18].

Despite these advantages, these techniques, like any statistical rock property estimation, may not always work well. Like any data driven techniques, the quality, uniformity, and processing of the data set will determine the effectiveness of both ANN prediction and Multivariate Linear Regression.

Future Directions

While the accuracy of both the neural network and the regression analysis has been demonstrated through error testing and correlation studies, it is important to exercise prudence if allowing these results to influence capital expenditure. As noted in the introduction, other methods of predicting reservoir porosity exist that are independent of seismic data. Well testing and reservoir simulation may serve as verification for a seismic-based porosity model.

Spatial interpolation of the well log porosity can be very accurate in close proximity to the well data. Further from well control, however, the accuracy can quickly deteriorate. In future studies, it may be worthwhile to devise a scheme whereby spatially based prediction techniques could be married with seismic based prediction techniques. Spatial data were not used in this study because of the strong correlation between lateral position and porosity. When included in the ANN analysis, spatial data were over emphasized by the network, ignoring needed information in the seismic data. Since this

study was designed to examine the possibility of using seismic data in porosity prediction, any spatial data were ignored.

The number of seismic attributes used in porosity prediction need not be limited to the eight used here. Hundreds of seismic attributes are known, and new ones could be invented solely for rock property prediction. A neural network may very well prove to be the quickest way to determine which of the hundreds of seismic attributes correlate most strongly with rock properties, including porosity.

As shown in Figure 22, the two types of analysis yielded different porosity "spectra". The two methods attempt to minimize two different types of statistical dispersion, namely mean squared error in the regression case, and mean absolute error in the network case. This may account for some of the dissimilarity. It should be possible to find appropriate linear weights which minimize mean absolute error, and therefore may provide for a more direct comparison between the results given by the neural network and the regression. This however, may not account for the difference in the shape of the porosity spectra. When a completely random array of data is fed into the neural network, for example, the output is strongly centered around the peak of 25% (see Figure 22), but there is virtually no spread in the results whatsoever. This is an intriguing area of study, which could be a focus of future efforts.

Bibliography

1. Schultz, P.S., et al., *Seismic-guided estimation of log properties (Part 1: A data-driven interpretation methodology)*. 1994, SEG. p. 305-310.
2. Lim, J.-S., *Reservoir properties determination using fuzzy logic and neural networks from well data in offshore Korea.* Journal of Petroleum Science and Engineering, 2005. **49**(3-4): p. 182-192.
3. Singh, V., et al., *Neural networks and their applications in lithostratigraphic interpretation of seismic data for reservoir characterization*. 2007. p. 1244-1260.
4. Walls, J.D., et al., *Seismic reservoir characterization of a U.S. Midcontinent fluvial system using rock physics, poststack seismic attributes, and neural networks*. 2002. p. 428-436.
5. Calderon, J.E. and J. Castagna, *Porosity and lithologic estimation using rock physics and multi-attribute transforms in Balcon Field, Colombia*. 2007. p. 142-150.
6. Joel, D.W., et al., *Interpreter's Corner---Seismic reservoir characterization of a U.S. Midcontinent fluvial system using rock physics, poststack seismic attributes, and neural networks*. 2002, SEG. p. 428-436.
7. Zhengping, L. and L. Jiaqi, *Seismic-controlled nonlinear extrapolation of well parameters using neural networks*. 1998, SEG. p. 2035-2041.
8. Tonn, R., *Neural network seismic reservoir characterization in a heavy oil reservoir*. 2002. p. 309-312.
9. Victor Mercado, H., R. Brian, and F. Adilce, *Neural networks in reservoir characterization*. 2006, SEG. p. 402-411.
10. Banchs, R.E. and R.J. Michelena, *From 3D seismic atrributes to pseudo-well-log volumes using neural networks: Practical considerations*. 2002. p. 996-1001.
11. Rafael, E.B. and J.M. Reinaldo, *From 3D seismic atrributes to pseudo-well-log volumes using neural networks: Practical considerations*. 2002, SEG. p. 996-1001.
12. Helle, H.B., A. Bhatt, and B. Ursin, *Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study.* Geophysical Prospecting, 2001. **49**: p. 431-444.
13. Pramanik, A.G., et al., *Estimation of effective porosity using geostatistics and multiattribute transforms: A case study*. 2004, SEG. p. 352-372.
14. Daniel, P.H., S.S. James, and A.Q. John, *Use of multiattribute transforms to predict log properties from seismic data*. 2001, SEG. p. 220-236.
15. Kevin, P.D. and A.L. Curtis, *Genetic-algorithm/neural-network approach to seismic attribute selection for well-log prediction*. 2004, SEG. p. 212-221.
16. Leiphart, D.J. and B.S. Hart, *Comparison of linear regression and a probabilistic neural network to predict porosity from 3-D seismic attributes in Lower Brushy Canyon channeled sandstones, southeast New Mexico*. 2001, SEG. p. 1349-1358.
17. Russell, B., et al., *Multiattribute seismic analysis*. 1997, SEG. p. 1439-1444.
18. Hampson, D.P., J.S. Schuelke, and J.A. Quirein, *Use of multiattribute transforms to predict log properties from seismic data*. 2001, SEG. p. 220-236.
19. Dorrington, K.P. and C.A. Link, *Genetic-algorithm/neural-network approach to seismic attribute selection for well-log prediction*. 2004, SEG. p. 212-221.

20. Martin T. Hagan, H.B.D., Mark Eagle, *Neural Network Design*. First ed. 1996, Boston, MA: PWS Publishing Company.
21. Taji, K., Miyake, T., Tamura, H., *On Error Backpropagation Algorithm Using Absolute Error Function.* 1999 IEEE International Conference on Systems, Man, and Cybernetics, 1999. IEEE SMC '99 Conference Proceedings. , 1999. **5**: p. 401-406.
22. Satinder, C. and J.M. Kurt, *Emerging and future trends in seismic attributes*. 2008, SEG. p. 298-318.
23. Satinder, C. and J.M. Kurt, *Introduction to this special section---Seismic Attributes*. 2008, SEG. p. 296-297.
24. Taner, M.T., *Seismic Attributes.* CSEG Recorder, 2001. **26**(7): p. 48-56.
25. Bruce, H., *Stratigraphically significant attributes*. 2008, SEG. p. 320-324.
26. Taner, M.T., F. Koehler, and R.E. Sheriff, *Complex seismic trace analysis*. 1979, SEG. p. 1041-1063.
27. Ecoublet, P., and Symes, W., *Application of Probabilistic Neural Network to Porosity Prediction*, in *The Rice Inversion Project, Annual Report*. 1998, Rice University: Houston, TX.
28. Dunteman, G.H., *Principal Components Analysis*. Quantitative Applications in the Social Sciences. 1989, Thousand Oaks, CA: Sage Publications Inc. 96.
29. Booth, J.R., et al., *Paleo-bathymetric controls on the stratigraphic architecture and reservoir development of confined fans in the Auger Basin: central Gulf of Mexico slope.* Marine and Petroleum Geology, 2003. **20**(6-8): p. 563-586.
30. Booth, J.R., DuVernay III, A.E., Pfieffer, D.S., Styzen, M.J., *Sequence Stratigraphic Framework, Depositional Models, and Stacking Patterns of Ponded and Slope Fan Systems in the Auger Basin: Central Gulf of Mexico Slope.* GCSSEPM Foundation 20th Annual Research Conference - Deep Water Reservoirs of the World, 2000: p. 82-103.
31. McGee. D.T., B., P.W., Gary, P.S., Pfeiffer, D.S., Sheiman, J.L., *Geologic Models and Reservoir Geometries of Auger Field, Deepwater Gulf of Mexico.* GCSSEPM Foundation 15th Annual Research Conference - Submarine Fans and Turbidite Systems, 1994: p. 245-256.
32. Kendrick, J.W., *Turbidite Reservoir Architecture in the Northern Gulf of Mexico Deepwater: Insights from the Development of Auger, Tahoe, and Ram/Powell Fields.* GCSSEPM Foundation 20th Annual Research Conference - Deep Water Reservoirs of the World, 2000: p. 450-468.
33. Bohn, C., Reilly, M., Seren, D., and Valenti, J., *Accommodation history, reservoir architecture, and production behavior in N and O sands of Auger Field, Deepwater Gulf of Mexico.* . AAPG Bulletin, 2008.
34. Issler, D.R., *A new approach to shale compaction and stratigraphic restoration, Beaufort-Mackenzie Basin and Mackenzie Corridor, northern Canada*. 1992. p. 1170-1189.

Table 6. Variable Definitions	
Variable	**Description**
tanh	Hyperbolic Tangent Sigmoid Transfer Function
I_i	*i*th Sample of the input vector, I
O_1	Output of the Input Layer of ANN
O_{2j}	Output of the *j*th neuron in the 2nd, or hidden, layer of the ANN
f_2	Transfer function associated with the 2nd, or hidden, layer of ANN
w_{ij}	Weight connecting the *i*th neuron in the input layer to the *j*th neuron in the hidden layer
b_{ij}	Bias connecting the *i*th neuron in the input layer to the *j*th neuron in the hidden layer
E	Mean Absolute Error between target and prediction
p_j	*j*th sample of the predicted log
L_j	*j*th sample of target Log
γ	Adaptive Learning Rate constant
α	Constant used in Momentum Learning
w_i	Weight multiplied by the *i*th attribute input in regression analysis
W	Vector composed of weights w_i
A	Matrix of Input Attributes
E^2	Mean Squared Prediction Error
A^T	Transpose of matrix A
$S(t)$	Analytical Trace
$s(t)$	Seismic Trace, the real part of S(t)
$s^*(t)$	Quadrature component of the Analytical Trace
A(t)	Amplitude Envelope of Analytical Trace
$\theta(t)$	Instantaneous Phase
$f(t)$	Instantaneous Frequency
z_i	Depth to log sample *i*
v_i	Wave Velocity in interval *i*
t_i	Cumulative Transit Time to depth, z_i
ϕ	Porosity
ρ_m	Matrix Density
ρ_f	Fluid Density
ρ_b	Bulk Density

APPENDIX

Results from Validation Training with ANN

In order to demonstrate that using a fixed number of training epochs was justified, included below are the testing results from a neural network with the same architecture used in Chapter 3, but trained using validation training.

In all measures of success, the validation trained network performed more poorly than the network that was simply trained for 1000 epochs. The results below, in Table 7, can be compared to the results in Table 5.

Well Name	Max Error	Mean Error	Max Error Percent	CC Test	CC Training
A02BP1	0.119183655	45.45776074	0.035007468	0.489084205	0.571225
A04BP1	0.126220945	43.74880299	0.047002472	0.263560985	0.638067575
A05	0.145012221	85.30130652	0.03960731	0.457423974	0.669981543
A06	0.106158558	60.23677415	0.037137763	0.631343879	0.587573059
A07	0.076808279	45.176815	0.030712686	0.876160938	0.576227064
A08	0.117827823	69.31048433	0.034462862	0.763788719	0.614099391
A09	0.115866659	68.15685804	0.030843336	0.60888349	0.628635092
A13	0.101556141	59.7388582	0.028509361	0.60487601	0.612899496
A14	0.100966756	59.39220951	0.032074924	0.437389553	0.592034589
A19	0.137616819	80.95106995	0.038444913	0.344678982	0.573031905
426STBP1	0.126965017	74.68530394	0.04363576	0.8539127	0.49864843
427	0.092790164	54.58244966	0.030825303	0.462827272	0.617828737
470	0.130205815	76.59165575	0.03879009	0.654684162	0.616431442
AVERAGES	**0.115167604**	**63.33310375**	**0.03592725**	**0.502597455**	**0.599744871**

Table 7. Prediction results of an ANN trained with validation training. In all cases, it performed more poorly than a network with the same architecture but trained for 1000 epochs.