AAPG Search and Discovery Article #90023©2002 AAPG Southwest Section Meeting, Ruidoso, New Mexico, June 6-8, 2002 (2003)

Using Artificial Intelligence to Predict Drilling Success Using Regional Data, Brushy Canyon Formation, Delaware Basin, New Mexico

R. S. Balch, New Mexico Petroleum Recovery Research Center, D. M. Hart, Sandia National Lab, W. W. Weiss, New Mexico Petroleum Recovery Research Center, and R. F. Broadhead, New Mexico Bureau of Geology and Mineral Resources.

Abstract

Incomplete and sparse information can introduce a high level of risk for oil exploration and development projects. "Expert" systems developed and used in several disciplines and industries have demonstrated beneficial results in modeling the decision making process of knowledgeable experts. A state-of-the-art exploration "expert" tool using a computerized data base and computer maps generated by neural networks, is being developed using fuzzy logic, a relatively new mathematical treatment of imprecise or non-explicit parameters.

Analysis to date includes generation of regional scale maps of aeromagnetic, gravity, structure, thickness, and production data for the target Brushy Canyon Formation in the Delaware Basin, New Mexico. For each regional scale map, data attributes were also computed to look for more subtle trends. These attributes include directional first and second derivatives, dip azimuth and magnitude, and azimuth and magnitude of curvature. These data were mapped and gridded to a 40 acre spacing, the current well spacing for Delaware pools in New Mexico, and compared to average monthly production in the first year for Delaware Brushy Canyon wells. The geophysical and geologic data covers 60478 bins (3780 square miles), of which 2434 of these bins have oil, gas and water production data. Using a new fuzzy ranking tool each data attribute was ranked for its ability to predict production potential at these well locations. The highest ranked attributes were gravity dip-azimuth, second latitude derivative of thickness, longitude derivative of gravity, and longitude derivative of structure. These attributes were used to generate a production potential map for the Delaware basin, using neural networks and expert systems, at the scale of 40 acres.

Expert systems operate by developing rule sets that can be used to answer questions related to the problem at hand, in this case prospect evaluation. Since prospect generation data often contains non-crisp data, such as "low porosity" or "high on structure", the expert system will necessarily allow fuzzy inputs. The approach taken is to accumulate all available public domain data and incorporate them into online databases, which can be accessed by the expert system. The primary goal after development of the production potential map is to teach the expert system to add or detract to each prospect's estimate of risk, in a fashion similar to that employed by human explorationists. Such a map would be a useful tool for evaluating the potential of infill, step out, and wildcat wells in the Delaware basin, both at reservoir and regional scales. This paper discusses the development of this production potential map and initial rules development for the expert system.

Introduction

Expert systems are computer programs that are designed to make decisions similar to the manner in which a human expert would. In the past expert systems have been primarily restricted to medical and industrial applications, but with DOE support an expert system to prospect for oil is now being developed to automate and accelerate prospect development for the Brushy Canyon formation in the Delaware basin. Expert systems operate by developing rule sets that can be used to answer questions related to the problem at hand, in this case prospect evaluation. Since prospect generation data often contains non-crisp data, such as "low porosity" or "high on structure", the expert system will necessarily allow fuzzy inputs. The approach taken is to accumulate all available public domain data and incorporate them into online databases, which can be accessed by the expert system. A primary goal was to develop a map of production potential based on available regional data from which the expert system could add or detract to each prospects estimate of risk.

Regional Data

A key component to the success of this study is the analysis of the regional data to provide baseline data to correlate with production potential; this also provides a source of heuristic rules for the expert system. Four major categories of regional data were selected and compiled. Regional gravity surveys cover the entire area of the Delaware basin and have been compiled with an accuracy of a few milligals. The survey measurements are on the order of a few thousand feet apart, but sample point locations are highly variable as gravity is measured in easily benchmarked locations, such as along roadways. Gravity measures variations in density and tends to highlight large-scale regional structures at basement depths and if structure has an impact on maturation, migration or trapping of hydrocarbons in the basin useful information can be obtained. Regional Aeromagnetic data, primarily collected via over-flights with 1 mile spacing re-gridded to 0.296 miles longitude and 0.346 miles latitude, also exists for the region. Aeromagnetic data highlights contrasts in the magnetic susceptibility between rocks and can help indicate basement blocks, large-scale faults, and possible large-scale alluvial deposits. The structure of the lower Brushy Canyon was picked on 700 wells in the basin covering a geographically large area[1]. Large-scale maps of these attributes covering the region were constructed with a kriging algorithm.

Structure can play more than one role in trapping and migration of hydrocarbons and two potentially helpful attributes for this study are structural highs, and flexures, which while commonly induce fracturing along the flanks of structures may also help locate subsurface faults which can compartmentalize fields, or allow preferential migration paths for hydrocarbons and water. Finally, the wells used to compute structure were used to generate an isopach map for the Brushy Canyon in the region(*Broadhead and Justman, 2000*). Thickness may indicate areas of greater potential production and also can indicate pinch-outs and other nonstructural features that may form hydrocarbon migration pathways or traps.

A number of attributes were calculated from these 4 principal data types. These attributes are 1st and 2nd derivatives along latitude and longitude; dip azimuth and magnitudes; and curvature azimuths and magnitudes. These values were computed to expose finer scale features in the basic data types that might be useful for correlating back to a production indicator. A total of 36 maps were generated using the Zmap tool of Landmark Graphics Release 98 plus interpretation package **(Table 1)**.

Each of these maps was gridded at a scale of 1320-ft (quarter section) because that is the regulatory spacing for wells in the Brushy Canyon in New Mexico. The gridded data was exported and loaded into the project production database. Our current production database is a subset of the Onguard database (courtesy of the SW PTTC) containing production information on all New Mexico wells. In this database we have also identified Brushy Canyon wells and using grid locations from the Zmap maps we were able to correlate producing wells with grid numbers. This essentially allows regressions to be formed using the production data as control points (training and testing) and the attribute data as variables. Any regression formed in this manner could then be used to predict production in all other 60,478 40-ac bins in the basin.

Fuzzy Ranking

There are two primary considerations when trying to form regressions: the first involves the quality of the data you are attempting to predict with the generated regression model; the second deals with the choice of attributes or variables that will be used in forming the regression model. An optional consideration is the application of linear models (least squares regression) or more complicated non-linear solutions such as polynomial regressions or neural networks. An average of the first 12 producing months oil production at each well was chosen as the data to be modeled. PredictOnline (*http://ford.nmt.edu/PredictOnline6/index.html*) is an in-house devloped neural network package that is available online and is based on the fast-converging, feed-forward, back-propagation conjugate gradient algorithm(*Moller, 1993*). A back-propagation feed-forward algorithm such as the conjugate gradient algorithm used here is "trained" using known inputs and outputs in an iterative fashion, with weights being sequentially adjusted until the desired fit (if possible) is achieved.

There are a number of ways to determine which of a set of inputs (attributes) would best be used to form a regression for a particular output. Simply crossplotting each input against the output can give an indication of the quality of linear or multiple linear regression models that could be formed. For more complicated relationships found in many oil field problems such simple tools often do not provide adequate solutions.

In previous studies we have used a single stage fuzzy-ranking algorithm to select inputs best suited for predicting the desired output(*Balch et al., 1999; Balch et al., 2000; Hart et al., 2000; and Weiss, 2000-1*). The algorithm statistically determines how well a particular input could resolve a particular output with respect to any number of other inputs using fuzzy curve analysis.

To illustrate the technique a simple example is given. Consider a set of random numbers in the range {0,1} using $x=\{x_i\}$, $i=1,2,\ldots,99$, and $x_i=0.01*i$, and plot each value ($y_I=$ Random(x_i)). Next add a simple trend to the random data ($y_i=(x_i)$^0.5+Random (x_i)) and plot those values. For each data (x_i, y_i) a "fuzzy" membership function is defined using the following relationship:

$$F_i(x) = \exp\left(-\left(\frac{x_i - x}{b}\right)^2\right) * y_i$$

Sample fuzzy membership functions are shown in **Figures. 1 and 2**. Here, b=0.1, since b is typically taken as about 10% of the length of the input interval of x_i. A fuzzy curve was constructed using a summation of all individual fuzzy membership functions in (x_i, y_i), and this final curve can prioritize a set of inputs for linear or non-linear regressions. The fuzzy curve function is defined below:

$$FC(x) = \frac{\sum_{i=1}^{N} F_i(x) * y_i}{\sum_{i=1}^{N} F_i(x)}$$

Where N is the size of the data set or the total number of fuzzy membership functions. **Figure 3** shows the curves for the data sets shown in **Figs. 1 and 2**. This simple example illustrates the ability of the fuzzy ranking approach to screen apparently random data for obscure trends such as the correlation between seismic attributes and reservoir properties[4].

More information is needed however to advance this analysis from the art of reading these fuzzy curves to a more robust and systematic elimination of less useful inputs, not only allowing selection of optimal inputs but also to allow an estimate of data quality and uniqueness. As such, we developed software based on a two-stage fuzzy ranking (*Weiss, 2000-2*). The two-stage fuzzy ranking (TSFR) has two improvements: 1) Reduction of input variable space through random characterization and 2) setting hard rules for selection of best-input variables. TSFR introduces second stage fuzzy curves, with first and second stage fuzzy surfaces to select the most important and independent input variables for modeling, while removing the input variables that show random characteristics.

TSFR uses first and second stage fuzzy curves to generate the fuzzy curve performance index (P_c):

$$P_c = \frac{P_{stage1}}{1 + P_{stage2}}$$

With the addition of a known random variable into the input space the FCPI is normalized by the random $P_{c,R}$ to produce the normalized fuzzy curve performance index ($P_{c,N}$):

$$P_{c,N} = \frac{P_c}{P_{c,R}}$$

The input variable with the smallest $P_{c,N}$ value is the most important variable. Input variables with $P_{c,N}$ greater than 1.0 are eliminated from the selection process. Once the most important variable is determined fuzzy surface analysis is performed.

Analogously, for fuzzy surfaces there exists a performance index using the first and second stage fuzzy surfaces (P_s):

$$P_s = \frac{P_{stage1}}{1 + P_{stage2}}$$

A similar normalization procedure produces the normalized fuzzy surface performance index ($P_{s,N}$):

$$P_{s,N} = \frac{P_s}{P_{s,R}}$$

The input variable with the smallest $P_{s,N}$ is the next most important and independent. In an iterative process, the input variables with $P_{s,N}$'s above 1.0 are eliminated from selection process. The fuzzy surface analysis continues until no input variables remain. Therefore, Two-stage fuzzy ranking can be used to automatically and quickly identify the important and independent inputs needed to model the system of interest.

For this study each of the 36 data and data attributes calculated and loaded into the database were analyzed using the second stage fuzzy ranking algorithm. Each data attribute was ranked for its ability to predict production potential at these well locations. The four best attributes selected were dip azimuth of gravity, second latitude derivative of thickness, longitude derivative of gravity, and longitude derivative of structure (**Table 2**).

Multivariate Regressions

Using PredictOnline, our in-house web driven neural network, a regression relationship was formed between these four inputs and the average first month's production. It is best when forming regressions to hold out a randomly selected sample of the data for testing. This data is used for testing the ability of the regression to accurately predict data not used in forming the regression. For this study a 520 well subset of the available 2434 wells in the basin was used to train the neural network. These 520 wells were selected as they were verified to have produced only from lower Brushy Canyon, and because they included dryholes in which a completion effort was made to generate production, as well as being distributed fairly evenly across the basin. Of these 520 wells 466 were used to form the regression while 54 were held out for blind testing. A 4-10-10-10-1 Neural network with 250 weights provided an excellent solution with a 2 to 1 ratio of data to weights and CC=0.90 for the training data and CC=0.81 for the blind test data. BOPM (average barrels of oil per month expected in first year) at all 60478-40ac bins in the basin were predicted using this model, including nearly 2000 other wells with Brushy Canyon production. Training and testing crossplots can be found in **Figures 4 and 5**, respectively.

Results

The calculated BOPM for each 40-ac bin in the New Mexico portion of the Delaware basin, Brushy Canyon formation was used to generate a map (**Figure 6**) to highlight potential areas of exploration. Neural network analysis necessarily results in a non-crisp solution, and examination of the cross-plots in figures 4 and 5 demonstrate that there is some possible error in the maps, though in general the high cross correlation means the overall fit is good. Therefore it would be inappropriate to expect that any give drilling locations would produce exactly as mapped, there are simply too many variables and the algorithm is designed to form generalized solution (**Figure** 7). Our goal is to use this generated map as the basis for an expert system that will quantify the risk associated with each prospect by answering questions often posed by human experts exploring in the Brushy Canyon, as well as questions posed by statistical analyses of the data itself.

Conclusions

Applying the technologies of fuzzy ranking and neural network analysis has allowed the generation of a "look here" map for the Brushy Canyon formation of the Delaware basin. It has been estimated, by us, which between 350 and 850 million barrels of oil remain to be recovered from just the New Mexico portion of the Delaware basin in the lower Brushy Canyon formation[10]. However, high water cuts and thin interbedded layers, which make log analysis less reliable, make production expensive. At this stage of the project we have already generated a map, which will allow explorationists to focus on underexplored productive regions, and allow testing through recompletions in many areas. As the expert system is developed, the presented "look here" map will be refined and tested under a barrage of questions which mimic those used by human explorationists and more reliable and consistent risk estimates should result for each of 60,478 potential 40-ac drilling sites in the New Mexico portion of the Delaware basin at an accelerated pace.

Acknowledgements

This project is supported through DOE project DE-AC26-99BC15218, Project Manager Jim Barnes.

References

Balch, R.S., et al. (1999). "Using Artificial Intelligence to correlate Multiple Seismic Attributes to Reservoir Properties," Paper SPE 56733 presented at the 1999 Annual Conference and Exhibition, Houston, TX, October 3-6.

Balch, R.S., et al. (2000)."Predicting Core Porosity using Wire-Line Logs at Dagger Draw Field, SE New Mexico", Paper SPE 59554 presented at the 2000 Permian Basin Oil and Gas Recovery Conference, Midland TX, March 21-23.

Broadhead, R. F., and Justman, H. A. (2000)."Regional controls on oil accumulations, Lower Brushy Canyon Formation, southeast New Mexico", in DeMis, W.D., Nelis, M.K., and Trentham, R.C., The Permian Basin: proving ground for tomorrows technologies: West Texas Geological Society, Publication00-109, pp. 9-18. 2000.

Hart, D.M., et al. (2000)." Time-to-Depth Conversion of Nash Draw "L" Seismic Horizon", Paper SPE 59555 presented at the 2000 Permian Basin Oil and Gas Recovery Conference, Midland TX, March 21-23

Weiss, W.W.:" Data Mining at a Regulatory Agency to Forecast Waterflood Recovery", Paper SPE 71057 presented at the 2000 SPE Rocky Mountain Petroleum Technology Conference, Keystone, Colorado, May 21–23.

Moller, M. F. (1983). "A Scaled Conjugate Gradient Method for Fast Supervised Learning," *Neural Networks*, (v6 1993), 525-533.

Weiss, W.W. (2000-1). " Data Mining at a Regulatory Agency to Forecast Waterflood Recovery", Paper SPE 71057 presented at the 2000 SPE Rocky Mountain Petroleum Technology Conference, Keystone, Colorado, May 21–23.

Weiss, W.W. (2000-2). "Risk Reduction with a Fuzzy Expert Exploration Tool," 3[d] Semi-Annual Technical Report, Contract No. DE-AC-26-99BC15218, U.S. DOE, November 2001.

Table 1. Regional Datatypes

VARIABLE	*SOURCE*	*MAP*	*EXPLANATION OR USE*
1	Gravity	Bouger Anomalies	Variations in regional densities incl. basement
2	Gravity	X-Derivative	X direction rate of change in gravity
3	Gravity	Y-Derivative	Y direction rate of change in gravity
4	Gravity	2nd X-Derivative	Removes basement features–sedimentary gravity
5	Gravity	2nd Y-Derivative	Removes basement features–sedimentary gravity
6	Gravity	Dip Azimuth	Data trends
7	Gravity	Dip Magnitude	Data trends
8	Gravity	Curvature Azimuth	Data trends
9	Gravity	Curvature Magnitude	Data trends
10	Aeromag	Aeromagnetic anomalies	Variations in magnetic susceptibility
11	Aeromag	X-Derivative	Data Trends
12	Aeromag	Y-Derivative	Data Trends
13	Aeromag	Second X-Derivative	Sedimentary section aeromagnetics
14	Aeromag	Second Y-Derivative	Sedimentary section aeromagnetics
15	Aeromag	Dip Azimuth	Differentiation of basement blocks?
16	Aeromag	Dip Magnitude	Data trends – scale of susceptibilities
17	Aeromag	Curvature Azimuth	Data trends
18	Aeromag	Curvature Magnitude	Data Trends
19	Structure	Brushy Canyon Subsea	Structural traps in Brushy Canyon
20	Structure	X-Derivative	X directional derivative - data trends of slope
21	Structure	Y-Derivative	Y directional derivative - data trends of slope
22	Structure	2nd X-Derivative	X directional derivative - rate of slope change
23	Structure	2nd Y-Derivative	Y directional derivative - rate of slope change
24	Structure	Dip Azimuth	data trends – faults - anticlines
25	Structure	Dip Magnitude	Steepness or scale of data trends
26	Structure	Curvature Azimuth	Rate of change in dip – flexure
27	Structure	Curvature Magnitude	Scale of curvature changes – fracture indicator
28	Thickness	Thickness of Sand >15% porosity	Potential net pay thickness
29	Thickness	X-Derivative	X directional derivative - data trends of slope
30	Thickness	Y-Derivative	Y directional derivative - data trends of slope
31	Thickness	2nd X-Derivative	X directional derivative - rate of slope change
32	Thickness	2nd Y-Derivative	Y directional derivative - rate of slope change
33	Thickness	Dip Azimuth	data trends – pinchouts
34	Thickness	Dip Magnitude	Steepness or scale of data trends
35	Thickness	Curvature Azimuth	Reservoir scaling
36	Thickness	Curvature Magnitude	Pinchouts

TABLE 2. VARIABLES SELECTED AS OPTIMAL

RANKING	*VARIABLE*	*PC/PCR*
1	6	0.8584718
2	32	0.8490006
3	2	0.8779765
4	20	0.88185640

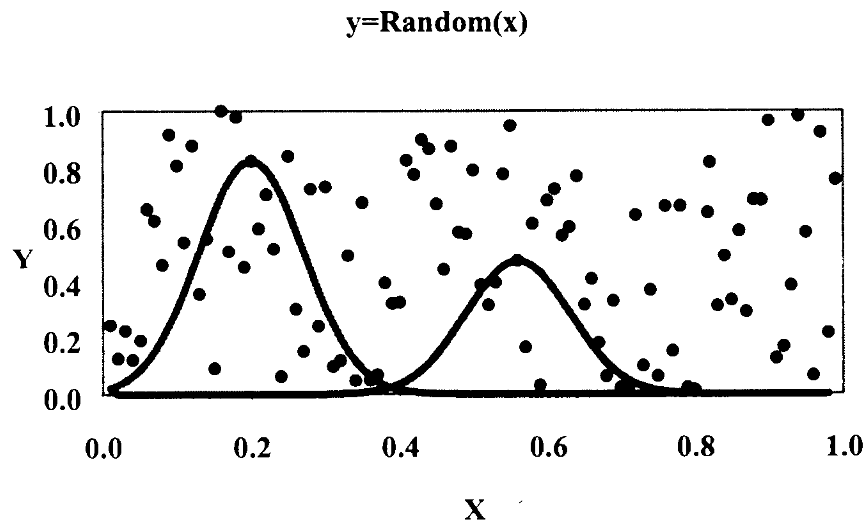


Figure 1. Conventional cross plot of a random data set (0-1). No correlation between X and Y. The trend is 0.5 (average between 0 and 1) is not evident. For each point a fuzzy membership function is defined, two example functions are shown on this plot.

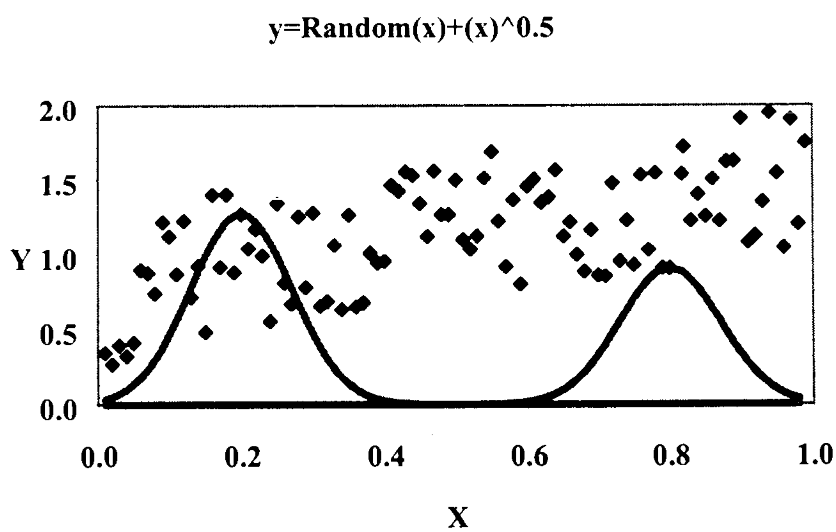


Figure 2. A conventional cross plot of a rando m data set (0-1) plus a square root trend. Again two sample fuzzy membership functions are illustrated

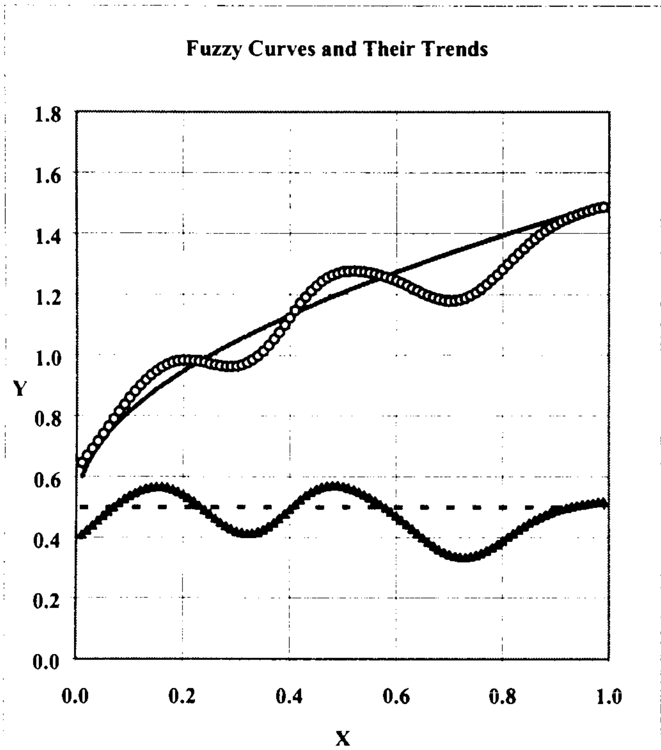
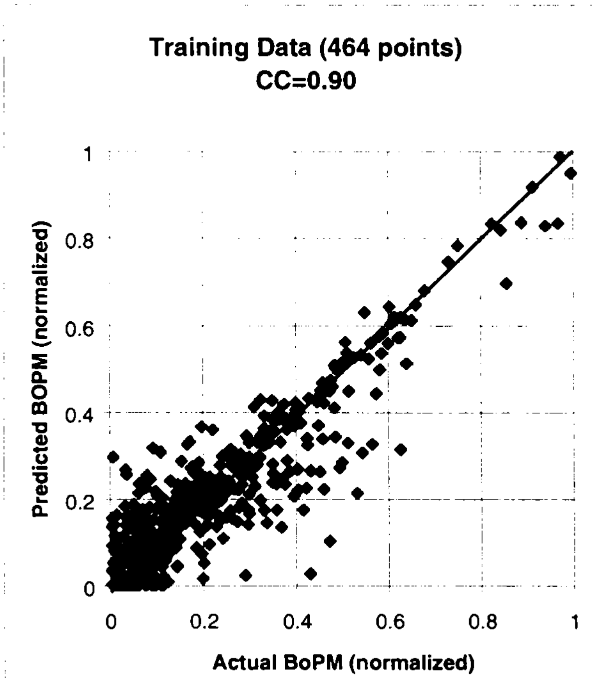


Figure 3. Fuzzy ranking curves. The trends are clearly evident.



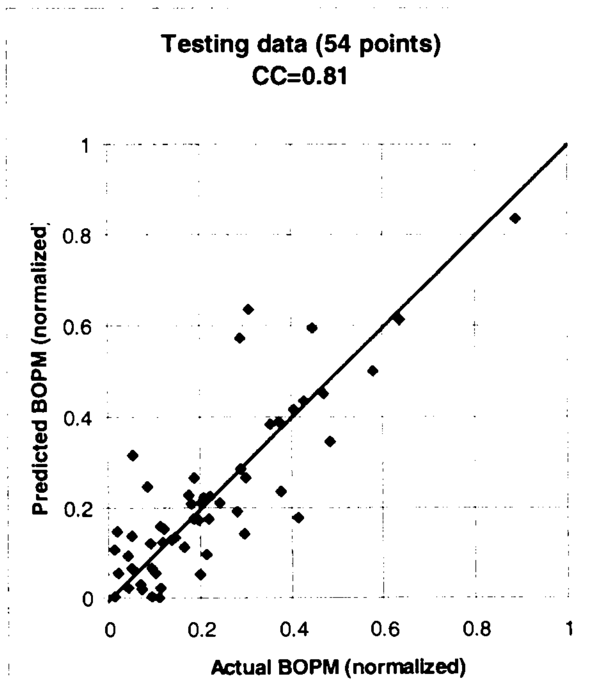


Figure 4. Crossplot showing training results for the 4-10-10-1 neural network used to form the regression relationship between 4 regional data attributes anmd ave BOPM for 1st year production.

Figure 5. Test points withheld from regression analysis.

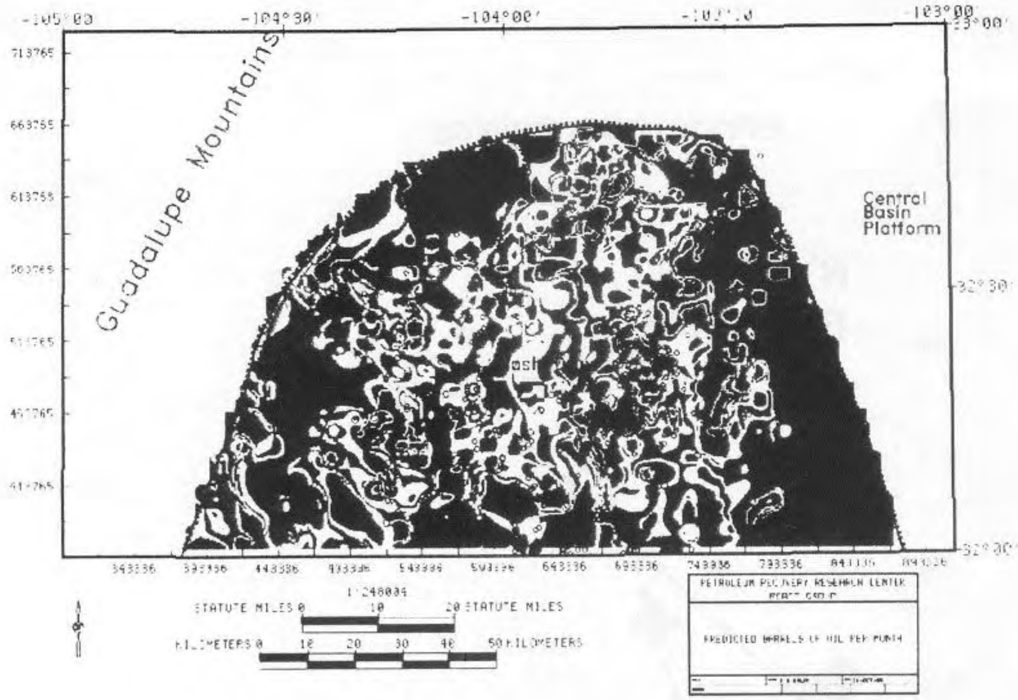


Figure 6. Predicted BOPM over the first year. This three-shade map can be loosely interpreted as poor (blue), average (yellow) and good (green). Small black circles are training points, larger blue circles are testing points.

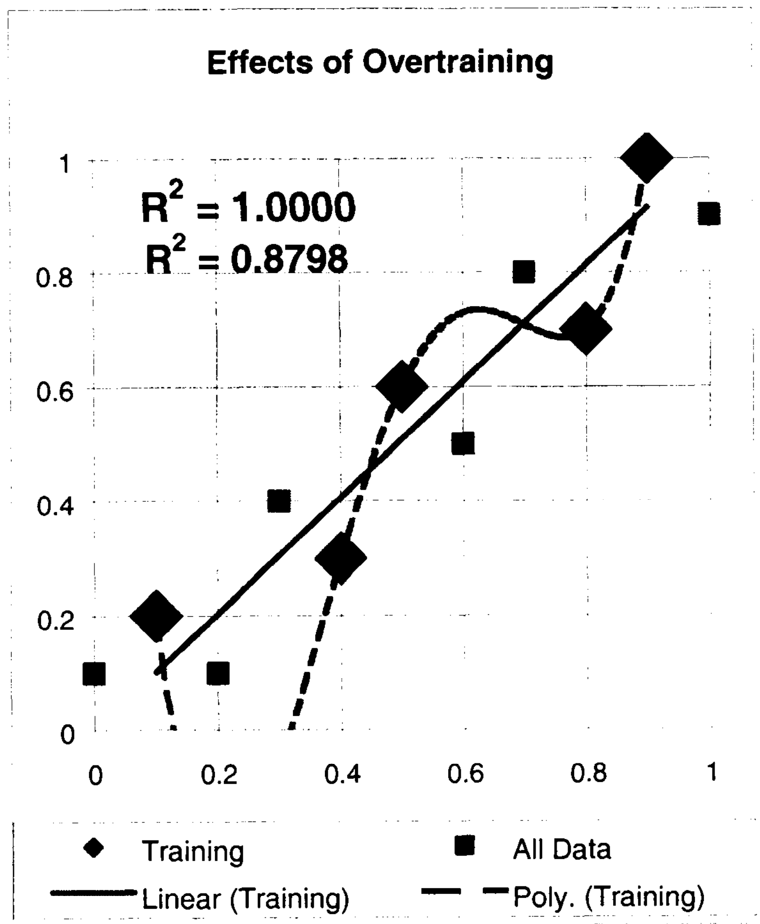


Figure 7. Neural Network Analogy. Square points are the complete data set in this simple example. Assume that only the points with diamonds are known and you want to find a solution. If you fit a polynomial to those four points you can easily find a solution in which the training data is exactly fit by the regression equation. However, as you acquire more data it becomes obvious that the solution is not good at all, and cannot predict accurately other points in the over-all data set. A straight line fit to the originally known data does not fit exactly, but does allow a more accurate prediction of the other points in this simple distribution. This shows by analogy the power of neural networks to find general solutions without knowing the complete data distribution