PRRC 00-13

Risk Reduction with a Fuzzy Expert Exploration Tool

(First Annual Technical Progress Report)

DOE Contract No. DE-AC-26-99BC15218

New Mexico Petroleum Recovery Research Center
New Mexico Institute of Mining and Technology
Socorro, NM 87801
(505) 835-5142

Date of Report:	June 30, 2000
Contract Date:	March 15, 1999
Anticipated Completion Date:	March 15, 2004
Award Amount for Current Fiscal Year:	$420,000
Project Manager:	James Barnes, NPTO
Principal Investigator:	William W. Weiss
Contributors:	Ron Broadhead & Andrew Sung
Contracting Officer's Representative:	William R. Mundorf, FETC
Reporting Period:	March 15, 1999 through May 31, 2000

Table of Contents

List of Tables

List of Figures

Overview

Objective

Incomplete or sparse information on geologic or formation characteristics introduces a high level of risk for oil exploration and development projects. Expert systems have been developed and used in several disciplines and industries, including medical diagnostics, with favorable results. A state-of-the-art exploration "expert" tool, relying on a computerized data base and computer maps generated by neural networks, is proposed through the use of "fuzzy" logic, a relatively new mathematical treatment of imprecise or non-explicit parameters and values. This project will develop an Artificial Intelligence system that will draw upon a wide variety of information to provide realistic estimates of risk. "Fuzzy logic," a system of integrating large amounts of inexact, incomplete information with modern computational methods to derive usable conclusions, has been demonstrated as a cost-effective computational technology in many industrial applications.

Accomplishments

During project year 1, 90% of geologic, geophysical, production and price data were assimilated for installation into the database. Logs provided geologic data consisting of formation tops of the Brushy Canyon, Lower Brushy Canyon, and Bone Springs zones of 700 wells used to construct regional cross sections. Regional structure and isopach maps were constructed using kriging to interpolate between the measured points. One of the structure derivative maps (azimuth of curvature) visually correlates with Brushy Canyon fields on the maximum change contours.

Derivatives of the regional geophysical data also visually correlate with the location of the fields. The azimuth of maximum dip approximately locates fields on the maximum change contours. In a similar manner the second derivative in the x-direction of the gravity map visually correlates with the alignment of the known fields. The visual correlations strongly suggest that neural network architectures will be found to correlate regional attributes with individual well production.

On a local scale, given open-hole log information, a neural network was trained to predict the product of porosity and oil saturation as reported in whole core analysis. Thus a direct indicator of an oil show is available from log information. This is important in the thin-bedded Delaware sand reservoirs.

Fuzzy ranking was used to prioritize 3D seismic attributes that were then correlated to formation depth with a neural network. The results were superior to those obtained using linear interpolation or low order polynomial interpolation as time-to-depth conversion tools.

A radial basis function neural network was developed and used as a log evaluation tool. This new technology gives an additional tool to the more commonly used multilayer perceptron (MLP) neural network. An interactive web based MLP, PredictOnline, was coded in Java and made available to consortium members for beta testing. PredictOnline demonstrates the power of Java programming language for web-based applications. A draft design of the Fuzzy Expert Exploration (FEE) Tool system based on readily available software was completed. The recent development of a Java Expert System Shell, JESS, facilitates expert rule development.

Discussion

This first year report is divided into six sections covering: data acquisiton and assimilation, geology, geophysics, engineering, computational intelligence, and technology transfer. The section titles are descriptive, except for the engineering section, which includes production and log analysis. Each section is self-contained.

Section 1—Data Acquisition and Assimilation

Most data used in this project are in the public domain. The exceptions are core and seismic information that were provided by oil and gas companies. The aeromagnetic survey is from Researchers at the NASA-funded Pan American Center for Earth and Environmental Studies (University of Texas at El Paso), headed by Dr. Randy Keller. They recently compiled and merged data sets from west Texas and southeastern New Mexico into an aeromagnetic database for the region. The data was obtained from several aeromagnetic surveys flown over these areas during the 1950s and 60s. Gravity data was acquired from the National Geophysical Data Center in Denver Colorado.

The visual information conveyed in conventional maps of the data may be increased using several trend-identifying data attributes. Attributes of the mapped properties that we used for this purpose are the slope, the change in slope (curvature), and directional dependent combinations of slope and curvature. The following equations generate the attributes and were applied to all surface mapped data.

$$\text{Slope} - \frac{dz}{dy}, \frac{dz}{dx} \qquad \text{Curvature} - \frac{d^2z}{dy^2}, \frac{d^2z}{dx^2} \qquad \text{Azimuth} - \arctan\left(\frac{d^2z/dy^2}{d^2z/dx^2}\right)$$

Production history data is from the New Mexico Oil & Gas Engineering Committee, Inc. A digital history is available from http://octane.nmt.edu/dataserv/DataMain.asp. Recently, Bob Emery, of the Petroleum Technology Transfer Council SW District, linked the pre- and post-1993 ONGARD data formats into a common digital format from the Go-Tech web (address above).

Oil price information is a record of the spot price at the time API stocks are published in the *Oil & Gas Journal*. A digital record has been maintained by the Petroleum Recovery Research Center since 1985. Additionally, commodity price history is available at http://quotes.ino.com/exchanges/nymex/ courtesy of Barchart.com.

Log data is from the New Mexico Bureau of Mines and the Roswell Geological Society Log Libraries.

Public domain computational intelligence tools used in the project are available at http://www-ra.informatik.uni-tuebingen.de/SNNS (Institute for Parallel and Distributed High Performance Systems (IPVR) at the University of Stuttgart) and http://herzberg.ca.sandia.gov/jess (Sandia National Laboratory). Additionally, http://ai.iit.nrc.ca/IR_public/fuzzy/fuzzyJavaToolkit.html is a site maintained by the National Research Council of Canada's Institute for Information.

Fuzzy ranking is a term new to petroleum engineering. Figures 1-3 illustrate this concept. The bell shaped lines in Figures 1-2 are fuzzy membership functions used to identify trends in what can appear to be random data.

$$\text{Fuzzy Membership Function}, \quad F_i(x) = \exp\left(-\left(\frac{x_i - x}{b}\right)^2\right) \bullet y_i$$

The fuzzy curves, a type of average of the points encompassed by the curve, are illustrated in Fig. 3 and clearly identify trends that may not be easily seen in conventional cross plots. The range, change in Y value, of the random data is about 0.26, indicating no meaningful correlation. The range of the random data plus the square root trend is ~0.9, indicating a strong corre-

lation between X and Y. The following equation is the fuzzy curve function used to define trends in Figure 3.

$$\text{Fuzzy Curve Function,} \quad FC(x) = \frac{\sum_{i=1}^{N} F_i(x)}{\sum_{i=1}^{N} F_i(x) / y_i}$$

The use of fuzzy curve analysis dramatically increases the speed and accuracy of data assimulation.

Section 2—Geology

Data generation and analysis in this section was done by Ron Broadhead , co-Pi, and student workers Heidi Justman and Jeff Anderson.

Geology

A regional geologic picture of the Brushy Canyon is a project requirement. Therefore co-PI Ron Broadhead focused his efforts on regional geologic interpretations. His findings are summarized below.

Geologic data

1. An extensive network of 16 well log cross sections was correlated for data quality control and assurance. Correlated formation tops of 697 wells were tied to the network of regional cross-sections (Figs. 4, 5). Top of Brushy Canyon Formation, top of Lower Brushy Canyon Formation, and top of Bone Spring Formation were correlated (Fig. 6). The geologically significant markers have been used by others[1] and are relatively easy to recognize and correlate.
2. Porosity data were obtained for Lower Brushy Canyon Formation from 530 wells using density porosity logs. We acquired:

- Net thickness of sandstones in Lower Brushy Canyon with porosity > 10%
- Net thickness of sandstones in Lower Brushy Canyon with porosity > 15%

Source rock geochemical data

1. Rock-Eval pyrolysis measurements were made on 70 samples of cuttings and sidewall cores from 26 wells in order to determine total organic carbon content, thermal maturity, kerogen type, and source-rock generative capability.
2. Visual kerogen analyses were made on 16 samples of cuttings and sidewall core from five wells in order to determine kerogen type and thermal maturity.
3. Thin sections were cut on 20 samples for evaluation of source rock type.

Production data

1. Initial production data were compiled from 508 productive wells from the Lower Brushy Canyon Formation. These data include initial production daily rates of oil, gas, and water (Fig. 7).

General well data

1. General well information was entered into computer databases for all wells, including operator, well number, lease name, and legal location.
2. Latitude and longitude of each well were calculated with digital land grid.

The acquired geologic and source-rock geochemical data were used to map structural, stratigraphic, depositional, and source-rock geochemical variables thought to exert control, to varying extents, on Brushy Canyon production. These data were compared to Lower Brushy Canyon production as a means of determining visual spatial relationships between the data and oil and gas productivity.

Geology

1. *Structure maps* on the top of the Bone Spring Formation, the top of the Lower Brushy Canyon Formation, and the top of the Upper Brushy Canyon Formation (Figs. 8, 9, 10) show similar basinal and local structure. Structural relief on the Upper Brushy Canyon is more subdued, indicating sedimentary infilling of the Delaware Basin by sediments of the Brushy Canyon Formation. Superposition of areas with Lower Brushy Canyon production on structure (Fig. 11)

indicates that structural culminations ("highs") have not localized oil accumulations. If anything, the reverse is true and areas of Lower Brushy Canyon production coincide with local structural lows.

2. *Isopach maps* of the Lower and Upper Brushy Canyon formations (Figs. 12, 13) generally show thicker areas in the basin and thinner areas around the basin margins. Local elongate to lobate areas coincide with structural lows in the underlying Bone Spring Formation and are indicative of localization of Brushy Canyon turbidite deposits in underlying paleovalleys carved into the Bone Spring.

3. *Porosity isolith maps* that indicate net thickness of Lower Brushy Canyon sandstones with porosity greater than 10% and net thickness of Lower Brushy Canyon sandstones with porosity greater than 15% show patterns similar to the isopach maps, but with better-defined depositional trends (Figs. 14, 15). When superimposed on wireframe relief maps of the structure of the upper Bone Spring surface (Figs. 16, 17) there is a clear indication that submarine fan complexes within the Lower Brushy Canyon coincide with incised paleovalleys at the shelf edge. The most porous, reservoir-quality sandstones were transported into the basin along axes of paleovalleys and submarine canyons. Reservoir-quality sandstones originated along the shelf margins of the northern and western flanks of the basin. The shelf margin formed by the Central Basin Platform along the eastern side of the basin did not act as a source of reservoir-quality sandstones in the Lower Brushy Canyon Formation. Superposition of Lower Brushy Canyon producing areas on the porosity isolith maps shows a strong positive correlation (Figs. 17, 18). Almost all oil pools that produce from the Lower Brushy Canyon are localized in areas that have at least 25 ft of sandstone with a minimum of 15% porosity (Fig. 18). Established production is almost always confined to areas with at least 100 ft of sandstone with at least 10% porosity (Fig. 19). However, there are barren areas between productive areas within these reservoir fairways.

Source-rock geochemistry

Collection of source-rock geochemical data is 95+% complete and analysis of data commenced at the end of our first project year. Previous investigations by other workers[2] indicate that oils produced from Brushy Canyon sandstones are, at least in part, from organic-rich strata interbedded with the Brushy Canyon sandstone reservoirs. Because of this self-sourced nature of the Brushy Canyon, distribution of hydrocarbon source-rocks and source-rock productivity should exert a fundamental control on oil and gas accumulations within the Brushy Canyon. To that end, we have analyzed samples of cuttings and sidewall core for hydrocarbon source rock parameters that indicate organic richness, thermal maturity, and kerogen type (oil prone vs. gas prone vs. nongenerative).

Our analyses to date have consisted of mapping raw source-rock parameters: **TMAX** (a measure of thermal maturity; and **TOC** (total organic carbon, a measure of organic richness). Some preliminary analyses have also been made on basinal distribution of kerogen type. The contour map of TMAX for Lower Brushy Canyon source facies (Fig. 20) indicates that although the Lower Brushy Canyon is thermally mature throughout, maturity is far from uniform. Lower Brushy canyon source facies have been matured to peak oil generation in the southern and western parts of the basin but are in earlier, less prolific stages of oil generation in the northern and

eastern parts of the basin. Similar but somewhat more complex patterns of thermal maturity can be discerned in the Upper Brushy Canyon (Fig. 21).

TOC measurements indicate that that dark, organic-rich Lower Brushy Canyon strata contain sufficient organic matter for oil and gas generation throughout the basin but organic richness of source facies is not uniformly distributed. Percentages of TOC exceed 1.4% in the northwestern and southeastern parts of the basin (Fig. 22). More marginal levels of TOC (<1.2%) occur in a swath that cuts from the southwestern to the east-central parts of the basin.

TOC is also unevenly distributed in source facies of the Upper Brushy Canyon (Fig. 23). TOC levels are marginal (<1%) for hydrocarbon generation in the eastern part of the basin but are very good (>1.5%) in the south-central and western parts of the basin.

Kerogen type governs whether the primary hydrocarbons generated from a source rock will be oil or gas. Preliminary analysis indicates that oil-prone kerogens are prevalent throughout large parts of the basin but that areas exist where gas-prone kerogens are significant, especially in the northern and northwestern parts of the basin. We will investigate whether this distribution, in conjunction with thermal maturity, has enhanced gas charge in reservoirs in these areas and has therefore affected oil recovery.

Fuzzy Rules

During the first year, we have constructed a set of geologically based rules that will be fuzzified in the second year and integrated into the expert system.

Section 3—Geophysics

This section reflects the work performed by Robert Balch, research associate and Darren Hart, graduate student. Geophysics includes the regional aeromagnetic and gravity data plus their respective attributes. The analysis of the attributes generated from the regional geologic maps, structure & isopach, are included in this section.

Geophysics

Gravity and magnetic data was used to verify the extent of the Delaware Basin and to perform a preliminary analysis of regional structure in the study area. Both gravity (Fig. 24) and aeromagnetics (Fig. 25) clearly show the location of the basin in the regional data.

Fine-scale studies will use much more detailed local maps to compare production attributes to gravity and magnetic data. Figure 24 shows the regional gravity structure for the New Mexico portion of the Delaware basin. A strong negative Bouger anomaly clearly defines the deeper parts of the basin, while the Central Basin Platform and Guadalupe Mountains exhibit the smallest Bouger anomalies in the region. Black dots represent the 1395 wells that were operating in the beginning of 1999, or plugged and abandoned after 1992. Additionally, 700 wells plugged and abandoned in the 1970–1980 time period are included in the production database, and accurate positions are being computed for display on maps. The figure supports the need for the FEE Tool, as large areas of the Basin remain unexploited.

Second derivative maps in both the X and Y directions were studied to search for regional scale anomalies such as faults in the Brushy Canyon by removing the effects of basement features from the gravity map. No such features were apparent on this scale, but future analysis at finer resolutions may reveal field-sized anomalies in the upper crust. Rigorous and systematic studies of the data using Fuzzy Ranking and Neural Networks may help reveal subtle features that are difficult to interpret visually.

The magnetic data included in the FEE Tool database was collected through an airborne survey flown at a constant elevation of 1000 feet with flight lines spaced one mile apart. This profiled line data was then gridded into equally spaced data points of 0.296 miles longitude and 0.346 miles latitude. An airborne magnetic survey provides data that is “smoother” than data collected on the ground. The airborne survey is similar to the upward continuation modeling problem, which acts like a low pass filter taking out the high frequency components of the signal, i.e. surface effects. Because of the great depths to basement, known from drilling, airborne magnetic data is the preferred data type to use for this investigation. The airborne magnetometer used to collect this data is known to have collected data to an accuracy of ±2.0 gamma. The applied diurnal correction, loop-based method allowed for a reliability of the reading within the same order of magnitude as the accuracy.[2]

Aeromagnetic data is generally used to determine the depth to and structure of the basement. "Basement" is used here to define the local igneous intrusive structure responsible for measured signal. Our goal is to process the aeromagnetic data further to isolate the effects associated with the Delaware Mountain group. This processing will include band-pass filtering to target the Delaware Mountain group source depth and calculation of directional derivatives to indicate possible trends related to fault sets within the Delaware Mountain group.

Attributes were generated from the regional gravity and areomagnetic maps. Dip azimuth, magnitude, and first and second order derivative maps were also completed for gravity and aeromagnetic data for the Delaware Basin and the surrounding region. Second derivative maps of gravity are widely interpreted[2] to represent removal of basement features, leaving a regional residual for rocks above the basement. Figure 26 shows a second derivative gravity, residual map in the X direction. Notice that alignment of fields (clusters of producing wells) seems to occur in the Y direction.

Figure 27 is a dip azimuth map for the aeromagnetic data in Figure 25. The amplitude of the whole-field aeromagnetic anomaly is very dependent on the strength of the magnetic susceptibility contrasts in the basin. The largest contrasts in susceptibility represent offset basement

blocks or lateral differences in basement intrusions. These features have many closely spaced lines in the dip azimuth map and may represent boundaries of individual basement blocks. Production seems to be related to margins of these "blocks" and may be due to faults propagating up-section into the Brushy Canyon, or local and regional structural highs built on uplifted blocks down-section of the Brushy Canyon. The boundary of the potash reserve follows the basement feature in a manner similar to the oilfields.

A total of 27 maps were generated for the regional structure, gravity, and aeromagnetic data. Table 1 lists the existing geophysical maps.

A new technique[3] for time-to-depth conversion of seismic horizons was completed during this quarter and tested at the Nash Draw field for the Brushy Canyon interval. The Brushy Canyon has been interpreted as being deposited through slope to basin turbidity or saline density currents.[4] This complex depositional process limits an interpretation of depth structure constructed strictly using the commonly-used method of geostatistical extrapolation of well control points. The geologic model is fairly complex, so a sophisticated pattern recognition tool was used to determine an empirical relationship between well and seismic data. Neural networks are such tools, and for this study they were used to predict a velocity model. Thus a robust computed depth map using seismic attributes to control velocity calculations in bins away from well control was obtained when the velocity and time pick maps are combined.

Several aspects of reservoir development require an accurate understanding of depth structure to adequately evaluate a play. Reservoir volume calculations, trapping mechanism analysis, and drilling are all depth-dependent, so an accurate depth structure map is an invaluable resource. The accuracy of the traditional techniques (TDQ and Zmap by Landmark Graphics) in predicting depth away from the central part of the Nash Draw field (Table 2) needed to be improved. An artificial intelligence approach, a multilayer perceptron neural network, (MLP) was used to evaluate the dataset. Training and prediction cross-correlations were well above 95%. Figure 28 shows the TDQ map, which poorly predicted the three offset points, and Figure 29 shows the MLP depth map, which accurately and blindly predicted the same points.

Section 4—Engineering

The engineering section encompasses analyses by all members of REACT group, Bill Weiss, Shaochang Wo, and Robert Balch, plus students David Welsh, Najib Benslimane, Weimin Ma, and Benedict Ekal. Production and log analyses are included in this section.

Engineering—Production

A purely analytical estimate of oil remaining to be produced from the Brushy Canyon zone of the Delaware sands in New Mexico was made by accepting that M. King Hubbert's famous bell-shaped curve applies to all basins. A logistics curve derivative, P = 2Pm/(1+COSH(-b(t-tm))), was fit to his lower 48 states curve first published in 1956 where Hubbert accurately forecast that peak US production would occur in 1970. CP is cumulative production, U is an asymptote representing ultimate recovery (EUR), tm is the time of peak oil production, and b=4Pm/U. Two EUR estimates are included in Fig. 30. Ultimate recovery is expected to be between 0.5 and 1.0 billion barrels and serves as a target for the Fuzzy Expert Exploration Tool.

Examination of the real data in Fig. 31 shows several anomalies in the rate of annual production. These anomalies are due to the effect of oil price, which more strongly affects production rates in a single basin, than for all US basins as a whole. The spike in annual production rate occurred in 1993, when the Brushy Canyon was a major play, and the flattening of the production rate during the rest of the 1990's is due to a decrease in oil prices. This prompted the two forecasts of Delaware production seen in Fig. 30. Note that peak production, assuming that price supports aggressive exploration and development, will occur between 2009 and 2013 at a rate of-20–40 million barrels per year.

The EUR estimates seen in Fig. 31 are highly dependent on price; accordingly price forecast technology is being developed. The goal is to see if a correlation exists between the daily closing price of west Texas intermediate grade crude oil and the future closing prices of a large variety of commodities. Some of the commodities are the U.S. Dollar Index, the Deutsche Mark, the Japanese Yen, crude oil, gold, wheat, soybeans, etc. The necessary information has been acquired and is being prepared for neural network processing.

Engineering—Log Analysis

Recently several concepts based on forward modeling have been applied to log interpretation. Traditional log analysis is considered an example of forward modeling. New adaptations to the forward modeling method include tuning conventional wireline logs with core information to estimate thin bed resistivity[5] and resistivity modeling.[6]

This project includes an inverse method based on training a neural network with core information to predict porosity, water saturation and a direct estimate of oil saturation. A new computational intelligence method based on fuzzy analysis is being investigated. This innovative approach focuses the initial scoping effort on lithology identification.

Oil shows are perhaps the most important factor when prospecting for Delaware zone oil. Historically, mud logs have provided the oil show information used to complete intervals in the Delaware sands. Mud log information is notoriously ambiguous, resulting in numerous non-commercial Delaware completions in good porosity zones. Or conversely, commercial oil zones are overlooked because the fine-grained, thin beds are determined to be wet by conventional log analysis. A new log analysis technique[5] incorporating information from core analyses has greatly reduced the risk of depending solely on mud logs. The new interpretation method relies on porosity and oil saturation information from core analysis. Public information concerning the majority of the wells drilled through the Delaware sands does not include core analyses. There is a need for a direct estimate of oil saturation from the public domain log information.

Neural Network Log Interpretation

Work by Schlumberger[7] thirty years ago demonstrated the relationship between neutron and formation density logs and oil saturation. The neutron log porosity, Φ_N, is related to hydrocarbon saturation in the flushed zone, S_{rh}, through

$$S_{rh} = \frac{\Phi - \Phi_N}{\Phi\left(1 - \frac{\alpha}{\beta}\right)}$$

where α is the hydrogen index of hydrocarbons, β is the hydrogen index of the mud filtrate, ϕ is the real porosity, and S_{rh} is the residual hydrocarbon saturation. The hydrogen index is dependent on the oil gravity.

The residual hydrocarbon saturation is related to the density log porosity, ϕ_D, through

$$S_{rh} = \frac{\Phi_D\left(\rho_{ma} - \rho_{mf}\right) \Big/ \left(1.07\Phi\right) + C_{mf}\rho_{mf} - C_{ma}\rho_{ma}}{C_{mf}\rho_{mf} - C_h\rho_h}$$

where ρ is the density of the matrix, *ma*, or mud filtrate, *mf*. The Compton effect is *C*, which is about 1.0 except in the case of hydrogen, *h*, where it is about 2.0 and ϕ is the real porosity. Thus, there is a physical basis to determine oil shows from log-only information.

A dataset consisting of a full suite of logs and whole core data through a 200 ft. interval of the Brushy Canyon zone (Nash Draw Well 23) provides the information required to correlate measured core data with log measurements. The following procedure details the correlating method.

A suite of 17 wells from the Nash Draw field in the Delaware Basin was obtained for correlation of wireline log data with core data. One well, Nash Draw #23, had full core data including permeability (K_{air}), porosity (Φ), grain density (GD), water saturation (S_w), oil saturation (S_o) and fluorescence while the other wells within the field had core plug measurements with these same parameters. All of the wells had wireline logs that included caliper, DPHI, DRHO, GR, LLD, LLS, MSFL, NPHI, NPOR, PEF, RHOB, SP and TNPH.

The purpose of this study is to use the neural network program (*Predict*) to correlate the wireline log data in Nash Draw #23 with the corresponding full core measurements. Once this is accomplished the neural network will be used to make predictions of core parameters such as Φ, S_w and S_o using the wireline log data from the other 16 wells. Each of these wells also has core plug measurements that will provide some measure of the neural network's ability to make reasonable predictions. If this procedure is successful it will be used to make predictions on approximately 500 wells in the initial expert system database.

A fuzzy ranking program was used on the wireline log data before attempting any correlations with the neural network. This procedure quickly ranks each of the log variables according to their likelihood of correlating with core parameters such as Φ, S_o, S_W, etc. The fuzzy ranking output is analogous to a correlation coefficient; the higher the ranking is (0.5, 0.6), the greater is the probability that a significant correlation will be found between the two variables.

Before running the fuzzy ranking program, each of the input and output variables is normalized between 0 and 1 using the relationship

$$\overline{X} = \frac{X - X\,\min}{X\,\max - X\,\min}$$

where the maximum and minimum values are obtained for each variable. Any number of input variables may be used, in this case wireline log parameters, to rank against one output variable such as core Φ, S_w, S_o etc. The end result is a list of variables with coefficients that are a measure of the likelihood of a correlation between a specific input variable and the output variable, as shown in Table 3.

Input and output data for the neural network program are normalized between 0 and 1, in the same fashion as for the fuzzy ranking. The maximum and minimum values for all the wells in the field are used in each case, instead of the local maximum and minimum for a single well. These were determined by finding the maximum and minimum values for each parameter for all the wells in the field, then adding 1σ to the minimum or subtracting one standard deviation (1σ) from the maximum.

The performance of the neural network was evaluated by exclusion testing in which a subset of data points was extracted from the training set to use for testing. For the Nash Draw set one out of every five data points was extracted for testing, resulting in 172 samples for training and 42 for testing. Once the neural network is trained to the desired correlation coefficient and is validated by testing, it will be used to make predictions of core parameters on wells for which only wireline logs exist.

With the data sets properly formatted the neural network was then trained using the training inputs (DPHI, LLD and LLS) to correlate with the desired training output (core Φ, S_w). The linear correlation coefficient is used as the measure of the closeness of correlation between the actual values and neural network determined training values. When the training is successful the ability of the neural network to make predictions is examined using the testing data set. The measure of how well the neural network can correlate the same input testing variables against the output testing variables is also determined by the linear correlation coefficient.

The training and testing process involves trying many different neural network architectures to determine which produces the best results. The example shown in Fig. 32 shows a sample 4-4-2-1 architecture in which there are four input variables (A1-4) and one output variable (core Φ). The number of nodes in the first layer is equal to the number of input variables (DPHI, LLD, LLD and PEF) and there is only one output variable (Φ). The number of nodes in the hidden layers (there can be one, two, or three hidden layers) can be altered to achieve the best correlations. This results in several hundred or thousand architectures that may be tested.

Initially input variables were selected from fuzzy ranking results to predict core porosity. Other combinations of input variables were later tried to make predictions of core Φ, S_w, S_o and $\Phi*S_o$ (Table 4). Because the neural network can overtrain, it is trained to several different correlation coefficients to determine which produces the highest testing correlation coefficient.

Results of the fuzzy ranking show an overall weak correlation between wireline log data and core data, with relatively low ranking coefficients. Among the inputs tested, DPHI, LLD, LLS, PEF and TNPH ranked highest, with coefficients above 0.3. Coefficients were identical for core Φ, S_o, S_w, K_{air} and $\Phi*S_o$, indicating a poor correlation between log and core data. These results were used as a guide for selecting input variables for initial neural network training.

Before beginning the neural network analysis, conventional linear correlation cross-plots were generated in an Excel spreadsheet. Core porosity showed a good correlation with core per-

meability (0.84 cc), DPHI (0.74 cc), LLD (0.65 cc) and SQRT((NPHI^2+DPHI^2)/2) (0.60 cc); poorer correlations were found with core S_w (0.45 cc) and NPHI (0.35 cc) as summarized in Table 4. A good correlation was found between S_o and S_w (0.69 cc) and between Sw and LLD (0.53); correlations between S_o and SQRT ((NPHI^2+DPHI^2)/2) (0.0 cc), and Sw with DPHI (0.39 cc) were not as good, as seen in Table II. Reasonably good correlations also exist between LLD and core permeability (0.51 cc) and DPHI (0.68 cc), Table 4.

Results of the neural network training are compiled in Table 5. Good correlations between log measurements and core porosity were consistently obtained by several different methods. The best correlations (training to 0.70, testing to 0.78) were found using adjusted data sets where resistivity values >100 ohm were excluded from training and testing (145 training points, 36 test points). Most of the excluded values exceeded 1000 ohm, whereas the retained values averaged less than 20 ohm. However, due to the large number of wells that will be tested in this study, this method of selective exclusion of data points is not reasonable.

Another variation was obtained by deleting every fourth point for exclusion testing (161 training points and 53 test points) instead of every fifth point (172 training points and 42 test). This procedure did not change the results noticeably and higher correlations were obtained by parsing every fifth point.

The best correlations with core porosity that included all data points were obtained using 11 inputs including Cal, DPHI, GR, LLD, LLS, MSFL, NPHI, NPOR, PEF, RHOB and SP (trained to 0.84 and tested to 0.77). The inputs DPHI and RHOB are basically the same with RHOB being calculated from DPHI. NPHI and NPOR are virtually the same as well, although they have slightly different fuzzy ranking values. Therefore, complications could occur with these results due to the extra weight placed on the two duplicated inputs. Other input combinations that had no duplication of input variables showed slightly lower results with the best training cc's at about 0.78-0.8 and testing cc's of 0.72-0.74. The neural network correlations are summarized in Table 5.

Although the neural network provides potentially fair correlations between log measurements and core S_w (training and testing to 0.56 cc), visual examination of the output data shows otherwise, as seen in Fig. 33.

Outlier values cause an apparent correlation. An attempt to correct this by using resistivity values that excluded outliers did not improve the correlation.

Poor correlations were also obtained between log measurements and core S_o alone (training to 0.56 and testing to 0.27 cc), Table 5. Changing the input variables did not significantly change these results and in some cases no correlations were found. The correlation was improved by multiplying core Φ by core S_o (training to 0.62 and testing 0.52 cc), Table 5.

Application of PredictOnline

Training and testing

During the course of the first year additional neural network software was developed. The latest neural network was applied to the same dataset described above and the results follow. The development of *PredictOnline* is detailed later in this report.

The objective of using a neural network is predicting core data from conventional openhole logs. It is desirable to have a neural network that can be applied to a broad range of areas in the Basin other than Nash Draw.

Nash Draw #23 in the Nash Draw field, had a full core data including permeability (K_{air}), porosity (Phi), grain density (GD), water saturation (S_w), oil saturation (S_o) and fluorescence and wire line logs that included caliper, DPHI, DRHO, GR, LLD, LLS, MSFL, NPHI, NPOR, PEF, RHOB, SP, and TNPH.

Neural network correlation

Optimum neural network training and testing is dependent on the neural network architecture. *PredictOnline* requires that the user specify the number of hidden layers. For example a 4-2-2-1 neural network architecture contains four input variables (NPHI, DPHI, Log LLD and Log LLS), two hidden layers and one output. The data file is split into two portions, training and testing determined by the value of perc for_testing, which in this case is 20% of the training file on all the training operations.

The neural network initially was trained and optimized for porosity prediction using full core data and the well logs from the Nash Draw #23 well. Training and optimization of the neural network for S_w, S_o and $S_o*\Phi$ from core measurements was performed to make predictions (validation tests) of other Brushy Canyon wells within the Delaware basin. The other wells include data supplied by Read Steven Inc., Roswell, NM, who provided 14 logs with core plugs associated with the Lea Penn, Northeast Lea Delaware, and Quail Ridge Delaware fields. Other locations with the requisite log and corresponding core data are in the East Livingston Ridge (Delaware) field and a Delaware wildcat

Table 6 compiles the neural network results; good correlations were obtained by using the core data as the output data. However a better correlation was obtained by averaging the core data over 2, 2.5 and 3 ft intervals (approximate interval of logging measurements), The best correlation with core porosity were obtained using 2 ft interval with 10 inputs variables (CAL, DPHI, GR, LOG LLD, Log LLS, Log MSFL, NPHI, TNPH, PEF, SP) trained to 0.91 and tested to 0.92.

The neural network provides good correlation between log measurements and core S_w using 2.5-ft interval with 10 input variables trained to 0.79 and tested to 0.82. A poor correlation was obtained between log measurements and core S_o (oil show). The net trained to 0.53 and tested to 0.49 with 3-ft interval used to average the core measurements. A good correlation was obtained by multiplying core porosity by core oil saturation, S_o. $\Phi*S_o$ averaged over a 2.5 ft interval trained to 0.83 (Figs 34 and 35) and tested to 0.73 as seen in Figs. 36 and 37. The predictions are presented in conventional vertical log format as well as a measurement versus predicted graph to illustrate limitations of a numerical correlation coefficient.

Since oil shows are a key exploration parameter, this new approach to a direct estimate of a production indicator independent of mud logs or core is significant.

Blind Predictions

Using the neural network trained with the four wireline logs generally available throughout the Basin (NPHI, DPHI, Log LLD and Log LLS) predictions of core plug $\Phi*S_o$ were partially successful. Logs from the 14 Read and Stevens wells, the East Livingston Ridge wells including a dry hole and a dry hole wildcat served as blind test data. The initial predictions were not acceptable. However, after accounting for wireline extension during the logging/sidewall coring operations the density porosity log correlated reasonably well with the core plugs in several wells. (Line tension continued to pull the coring tool up hole after the hoist was stopped.)

In addition to estimating core plug porosity with *PredictOnline*, predictions of Φ^*S_o were made for each of the Read and Stevens wells using neural network trained with the Nash Draw 23 dataset. The Hudson #6 and #2 correlation coefficients were 0.59 and 0.44 respectively. The Hudson #1 and Truman # 4 datasets did not correlate. The East Livingston Ridge wells produced only poor results (Cercion #8 cc=0.48 and Cercion #1 cc=0.16).

The potential usefulness of a direct oil indicator from wireline logs is illustrated in Fig. 38. Using the neural network predicted average Φ^*S_o for each of 19 Delaware Basin wells and plotting the summation versus Peak Oil production, it appears that average Φ^*S_o must exceed 100 porosity-oil saturation percent squared to be commercial.

Section 5—Computational Intelligence

The information presented in this section is primarily the work of Shaochang Wo and Mateusz Pietrzyk, student. The focus of this section is the development of new computer tools used in the project: fuzzy logic, neural networks and the FEE Tool system.

Computational Intelligence

Work on the algorithms required to utilize the database progressed at a rapid pace. A first generation of *Predict*, a multilayer perceptron (MLP) neural network, was developed and used to estimate core plug porosity values at a Brushy Canyon field. A second generation MLP, web-based neural network was also developed.

A Radial Basis Function, RBF, neural network was developed to improve the analysis of localized nonlinear features where a great deal of detailed information is available, such as logs.

Development of Radial Basis Function Neural Network

Neural network methods are increasingly receiving interest in reservoir characterization studies[8,9] where correlations among diverse data sources are often not straightforward when more than two variables are involved. Applied as a multivariable nonlinear regression tool, the radial-basis function (RBF) network provides a universal approximator similar to another type of neural network, the multilayer perceptron (MLP). However, RBF networks focus more on localized nonlinear features, which makes them a promising method to deal with clustering problems such as correlating different formation measurements. A cluster is an area where the data concentrate.

Typically, an RBF network, shown in Fig. 39, has one single hidden layer and the weighted links only between the hidden and output layers. In fact, Fig. 39 is a graphical representation of a function of the form

$$f(\vec{x}) = w_0 + \sum_{i=1}^{N} w_i G(\|\vec{x} - \vec{t}_i\|) \quad \text{.....(1)}$$

where $\vec{x}$ is the input vector of *p*-dimensional space; $\{G_1, \ldots, G_N\}$ are radial-basis functions, most often the Gaussian function. Moreover, each radial-basis function *G* is associated with a center $\vec{t}$. For Gaussian functions,

$$G_i = G(\|\vec{x} - \vec{t}_i\|) = exp(-\frac{\|\vec{x} - \vec{t}_i\|^2}{2\sigma^2}) \quad \text{.....(2)}$$

where σ is the standard deviation. The Euclidean distance between the input vector $\vec{x}$ and the center $\vec{t}$ determines the output of *G* that gives 1.0 at the center and approximates to zero for larger distance.

The design of an RBF network, i.e. selecting the links, deviations, and locations of the centers, can be viewed as a curve-fitting problem.[10] Given *m* pairs of inputs, $\{\vec{x}_1, \ldots, \vec{x}_m\}$, and desired response $\{d_1, \ldots, d_m\}$, we are looking for a function *f* of Eq. 1 that minimizes the following error functional *H*:

$$H[f] = \sum_{j=1}^{m} (d_j - f(\vec{x}_j))^2 + \lambda \|Pf\|^2 \quad \text{.....(3)}$$

where λ is called the regularization parameter and *P* is the stabilizer, usually a differential operator. Detailed mathematical establishment of the RBF network can be found in the references.[10,11,12,13]

Various learning strategies have been developed for optimally selecting centers, links, and weighted norms (standard deviations in this case). In this work, a hybrid learning process was used where the k-means clustering algorithm[14,15] moves the locations of centers in a self-organized fashion; the links are learned with supervising by the least-mean-square (LMS) algorithm; and the deviation of all the Gaussian functions is valued at

$$\sigma = \frac{D_{max}}{\sqrt{2N}} \quad \text{.....(4)}$$

where D_{max} is the maximum distance between current centers.

The Nash Draw well 23 dataset was used to evaluate the new algorithm. The density, deep resistivity, and shallow resistivity logs were used with the radial basis function neural network to predict porosity measured by whole core analysis. The correlation coefficient of the actual versus predicted porosity is 93% as seen in Fig. 40. The cross-plot in Fig. 41 provides a direct comparison of density log porosity versus the core porosity. Notice the correlation coefficient is 73%. The additional information included in the resistivity logs significantly increases the ability to predict core porosity.

Development of PredictOnline

Web implementation of a second generation MLP-neural network for correlating attributes with measurements (*PredictOnline)* by Mateusz Pietrzyk was initiated. The software was successfully tested in-house and generated correlations equal to or superior to the first generation MLP developed several years ago.

The Java-sourced program was made available online to a consortium member for further testing. *PredictOnline* can be found at http://ford.nmt.edu and is representative of the eventual interface that will be used for the Fuzzy Expert Exploration Tool. The following documentation describes the utility of the software.

1. *Overview*

PredictOnline is a web-based neural network application. More precisely, it is a web-based interface to SNNS—Stuttgart Neural Network Simulator developed at the University of Stuttgart, Germany and the University of Tübingen, Germany.[16] SNNS is a comprehensive neural network package featuring a full GUI. It can also be used without GUI, using a special SNNS program called batchman. This program is the engine of *PredictOnline.*

"Web-based" means that end users can run the program through their web browsers on local computers connected to the Internet. The computation will be performed remotely on the server machine. Thus, the users are not limited by the speed of their local computers, as neural network training can be computationally demanding.

The second goal of *PredictOnline* is to create a simple user interface with the minimum number of user-settable parameters. The current version (v6) uses Swing for its user interface -- a state-of-the-art Java GUI package.

2. *Neural networks*

Neural networks are useful for correlating more than two variables. They can be viewed as an inverse problem solver where a system of non-linear equations is developed to yield a known value. The variables in the equations are the input parameters. The constants in the equations are called weights. During the training of a neural network, the weights are adjusted to yield the known output value.

A non-linear, feed forward, back propagation, neural network with a fast matrix solver is used to correlate the input parameters with the known independent variable. Fully connected neural network architectures are evaluated for speed and efficiency. As an example, two different neural network architectures (4-4-2-1 and 4-3-2-1) are seen in Fig. 42. A1, A2, A3, and A4

are the input variables; any number of inputs can be used. The tie lines are the weights and circles are the non-linear functions. Maintaining a satisfactory ratio of training data to weights (coefficients of the regression equation) is desired. A good overview of a back propagation neural network is available at http://www.npto.doe.gov/Software/miscindx.html (see the Risk Analysis manual).

Neural networks are trained to yield an output value. The goodness of the training is expressed as a correlation coefficient with 100% being perfect. In reality networks trained to about 80% are suitable for forecasting. Once a network is trained it can be used to predict unknown output values given the input parameters. The accuracy of the predictions can be judged with exclusion testing.

3. *Input files*

PredictOnline accepts files in ASCII text format. There are two types of input files:

- *data files*—A single data file is used by *PredictOnline* for network training and testing. This is the file with all known patterns. Each line in that file defines a single pattern. For example a simple data file may look like this:

  10.0 0.7 5.81 3.1
  11.0 1.5 4.21 2.19
  12.0 0.9 3.1 2.1

  For instance, the first pattern (i.e., the first line) is: ``for input vector [10.0 , 0.7 , 5.81] the output value is 3.1".

  Each line in the data file must contain the same number of values (in the example each line has 4 values).

  The network will be trained and tested using all patterns (i.e., all lines in the data set). The data file will be split into two portions, training and testing, according to the value of perc for testing. Each line (i.e., each pattern) from the data file has perc for testing probability of ending up in the testing file. Therefore, if one wants to train the network using the entire data file, one should set perc for testing to 0, etc.
- *predict files*—When the network is trained, one can use it to make predictions. A predict file contains input vectors with no output. Output will be predicted by the trained network. So, for the network trained using the above data file, a simple predict file may look like this:

  10.5 0.8 5.3
  11.0 1.0 4.9

  In this file the last column (output) is missing. Each line in the predict file must contain the same number of values (which is the number of columns in the data file minus 1).

4. *Output*

- *Training/testing output*

This is the trained network with its performance, i.e., how precisely it can predict the output. The trained network is not shown on the screen. The performance is measured with the following two parameters:

$$r^2 = 1 - \frac{\sum_i (x_i - y_i)^2}{\sum_i (x_i - \overline{Y})^2} \quad \text{and,} \quad \text{if } r^2 \geq 0, \; r = \sqrt{r^2},$$

where n is the number of patterns, x_i is the desired output for pattern i, y_i is the neural network output for pattern i, and $\bar{Y} = \sum_i y_i / n$.

These two parameters (r^2 and r) are computed for both training and testing portion of the data file.

- *Predict output*

  The predict output is simply the vector of neural network outputs $Y = [y_1 , ... , y_m]$, where m is the number of patterns in predict file. When predicting, we do not have the desired output and, of course, we cannot measure the performance of the prediction. The predict output file looks just like the input file, except that it has an additional column—the output obtained from the trained neural network.[17]

5. *Usage*

PredictOnline is just a network-based interface to SNNS's batchman. Therefore, all computation is performed on the server machine. Figure 43 shows the architecture of *PredictOnline* (see Figure 44 for additional detail).

The most important consequence of the web-based nature of the system is that the *PredictOnline* engine (i.e., batchman) has no direct access to the input file(s), which reside at the client side. So, in some way, those input files have to be transferred to the server side and the output transferred back to the client (see Figure 44). In previous versions FTP was used for file transmission. In the current release (v6) *PredictOnline* takes care of file transmission as well. However, since *PredictOnline* is running as an applet inside a web browser, it has no access to the local disk.[18] Such access is necessary to perform file operations (i.e., open and save). Therefore, *PredictOnline* users need to grant certain privileges to the *PredictOnline* applet. This easy one-time procedure is explained at the start point of *PredictOnline*:
http://ford.nmt.edu/PredictOnline6/.

PredictOnline supports multiple users. Each user will have a *PredictOnline* account on the server machine. Let us use an example account name, John, to illustrate the usage of *PredictOnline*. Let us also assume that the *PredictOnline* server is located at ford.nmt.edu.

- The user starts a web browser at http://ford.nmt.edu/PredictOnline6/, the login point of *PredictOnline*. After a successful login, the user should see a window as shown in Figure 45. The PROJECT NAME list contains all of the user's projects. By clicking on a project name, particular project can be selected. Figure 46 shows project p01 selected. One can see that this project contains previously uploaded data files and results obtained by training and predicting. Figure 47 shows project ddd selected. This project is empty.

  The user has the option of creating new projects and deleting old ones. Further, the user can upload data files from the local computer to the server and download the results to their own computer. Let us now assume that the user is working on p01 project. In the DATA FILE list, the user selects the file RSonline.txt and in the PREDICT FILE list, the file RSonline-pred.txt.

  Next, the user selects the architecture of the network. In the ARCH field, the user enters 6 5 so the actual architecture of the network will be 2-6-5-1 (because the RSonline.txt file has 2 inputs and 1 output).

  Next, the user specifies the percentage of the data file that will be used for testing the network (see Section 3). In this case it is 10.

  The network is ready to be trained. The user selects one of two training options: regular training or training with cross-validation. When training with cross-validation, during the train-

ing process *PredictOnline* will periodically check the network's performance on the testing portion of the data file. If it starts decreasing, the process is stopped due to over-training. When training without cross-validation (regular training) it is a good idea to do testing after the network has been trained. In this case the user chooses "Training with cross-validation." The screen is shown in Figure 48.

- The progress of the training process will be displayed on the screen. Typically, this will look like this:

  connecting...
  request sent - please wait...
  process started on server...
  cycles = 200 trainMSE = 0.0157214 testMSE = 0.0163658
  cycles = 400 trainMSE = 0.0152363 testMSE = 0.0161779
  cycles = 600 trainMSE = 0.0150482 testMSE = 0.0156563
  cycles = 800 trainMSE = 0.0149772 testMSE = 0.0155811
  cycles = 1000 trainMSE = 0.0149772 testMSE = 0.0155811
  Cycles trained: 1000
  Training stopped at train MSE = 0.0149772
  Training stopped at test MSE = 0.0155811
  train (r2 r) = (0.71258 0.84414)
  test (r2 r) = (0.66466 0.81526)
  end

  This means that the training has been completed in 1,000 cycles (or epochs) and the performance measures for training and testing are r2train = 0.71258, r2test = 0.66466, rtrain = 0.84414, and rtest = 0.81526. Also, during training, *MSE* for training and testing set is displayed. *MSE* is defined as

$$MSE = \frac{SSE}{m},$$

  where *SSE* is the sum of the squared differences between all actual outputs and desired outputs, and m is the number of patterns.
- When training is finished one can view the training/testing plots. The x-axis represents desired output, y axis represents actual output from the neural network. Figure 49 shows the screen with the training plot and the contents of the main data file used for training (RSonline.txt). Also, the above-explained statistics of training (i.e., r2 and r) are shown on the screen.

- Next, the network can be used for prediction. When the user clicks the Predict button, *PredictOnline* will make predictions for all selected predict files.

- The result files (training, testing, and predicted) can be downloaded to the local computer. Also, both data files and result files can be viewed in the main text area of the *PredictOnline* applet. However only files smaller than 32 KB can be viewed this way. This is to avoid over-

loading the network with such large files. On the other hand, the upload and download features of *PredictOnline* can handle any file sizes.

Expert System Shell

Expert system shells have been developed to emulate human expertise. The C Language Integrated Production System (CLIPS) from NASA is briefly compared to three other system shells (shown in Table 7). The initial choices were to develop an in-house system based on CLIPS such as other vendors have created, or to begin with one of the other systems.

The expert system shells have been superceded with the recent advent of Jess,[19] an expert system shell and scripting language created at Sandia National Laboratories. Jess, when coupled with the NRC Fuzzy Java Toolkit[20] now available, greatly simplifies the development of the FEE Tool. A design draft follows.

A Draft Design of The Fuzzy Expert System Tool

The design of such a complex system is state-of-the-art work. There is no predetermined blueprint to follow. However, the knowledge, either successful or unsuccessful, gained from other AI[21] projects provides some useful guidelines.

In terms of concepts, a top-down design is outlined to define the goal, elicit the knowledge domain, identify the major difficulties, and select methods and tools. As the single goal, the expert system will provide the risk of an oil/gas prospect at any location of the Brushy Canyon formation in the Delaware Basin.

The knowledge domain will cover diverse data types and expert resources, which is in many ways derived from wireline log, production history, seismic, aeromagnetic, and gravity data. Conventional geology/engineering techniques and expertise are able to provide an accurate estimate of prospect risk if all of the required information is available. In most cases (locations), the difficulty comes from the lack of a complete data set, which is essential for performing a conventional analysis. Therefore, how to estimate unavailable data from available data is one of the key issues.

Neural network and fuzzy logic techniques can resolve this sparse data issue. The REACT group has demonstrated that neural networks can be applied to a number of topics requiring multivariable nonlinear regressions. A web-based neural network package, *PredictOnline*, was recently completed and is now available to consortium members.

<u>Guidelines for Designing a Complex System</u>

In terms of AI applications, the proposed system consists of a spectrum of sub-applications: risk analysis, basin appraisal, rock analysis, log interpretation, log correlation, geophysical attribute correlation, etc. In the last two decades, the petroleum industry has increasingly funded AI projects and knowledge-based systems are now an ordinary part of information systems in many industries. However, unsuccessful projects are not unusual. The lessons learned from a project's failure are as valuable as the knowledge gained from a successful one. There are four potential pitfalls that AI developers should be aware of:

- Using AI when conventional techniques will solve the problem;
- Selecting an inappropriate AI technology or combinations of technologies to solve a problem;
- Attempting to capture expertise in a knowledge domain that is too large;

- Matching user expectations with the capabilities of the AI system.

It is obvious that our proposed knowledge domain is very large and perhaps too diverse to be organized into a single reasoning system. This represents a pitfall to be avoided. The following guidelines present the basic philosophy of the design:

- Dividing the whole system into several self-organized reasoning modules, each with its own well-defined outputs and precise knowledge domain;
- Avoiding a fuzzy rule if a crisp rule is sufficient;
- Using industry-proven products (software and hardware) if they fit our requirements;
- Creating the system as an assistant tool that involves the user in the decision-making procedure.

The current trend is that expert systems, rather than being an independent technology, are becoming part of an integrated strategy. In this project, we will integrate the expert system, neural network, fuzzy logic, Internet communication, and visualization together as a web-based exploration assistance tool.

FEE Tool Architecture Overview

Internally, the FEE Tool is a multi-tier web-based system. The top tier (or presentation tier) is running on a client (in a web browser) as a Java *applet*. The middle tier (or control tier) is a collection of Java *servlets* running on the server. This tier communicates directly with the top tier on a request-reply basis. The user applet sends a request to the appropriate servlet, which, after performing some task(s), sends a reply. Requests to the middle tier are initiated through the graphical interface of the applet running on the client side. Finally, the bottom tier (or base tier) is a collection of tools that perform all required work. The base tier is also installed on the server side and performs services for the middle tier. However, it does not have to be located on a single physical machine. The base tier itself can be distributed, and in some cases, even divided further into sub-tiers.

Using this multi-tier approach, FEE Tool users will be able to access it through the Internet using a web browser. All input parameters can be entered via graphical interface. In some cases, input parameters are too big to be entered manually. For instance, if a parameter is a file, one does not expect the user to enter the contents of a file by hand. In such cases, the applet will access the local file and transfer it to the server, where it can be reached by FEE Tool components.

The output of the FEE Tool will be displayed on the user screen inside a web browser. Similarly, if the output is too big to be displayed, it can be transferred from the server and saved locally to a file by the applet.

The above two features (*uploading* inputs and *downloading* results) require that the applet running on the user machine has access to the local disk. Such access, for security reasons, is by default denied. Fortunately, Java features excellent security control and users can grant certain *privileges* to selected applets. The *uploading* and *downloading* features are demonstrated in *PredictOnline*.

Inference Work Flow

The kernel system consists of several task-oriented modules. Figure 50 shows the process flow and the connections among modules. Note that four self-organized reasoning modules, Regional Indication, Trap Assessment, Formation Assessment, and Enhanced Recovery Scheme are

processed in parallel. In fact, each of them can be considered as an independent expert system. Aiming at web-based applications, Jess and NRC Fuzzy Java Toolkit (detailed in next two sections), both written in Java will be used as the fuzzy expert system shell. For a seamless integration, all modules will be developed in Java except the database. Currently, data are stored in a Microsoft Access database. In the future, we plan to transfer data to a Microsoft SQL Server that is more suitable for internet access. To switch data from Microsoft Access to Microsoft SQL Server requires no change to the data format, but only the data relations. Starting with an introduction to Jess and NRC Fuzzy Java Toolkit, descriptions of the basic functions of each module follows.

Jess – A Java Expert System Shell

Jess is an expert system shell and scripting language created by Ernest J. Friedman-Hill at Sandia National Laboratories. The license agreement allows us to use Jess for non-commercial and academic use without charge. Jess was entirely written in Sun Microsystem's Java language, therefore it can be easily linked to code written in the portable Java language. Jess, being just a set of Java classes, is easily incorporated as a library into graphical applications written in Java. It is also possible to write graphical applications in the Jess language itself. Jess's rule engine uses an improved form of a well-known algorithm called Rete[22] to match rules against the knowledge base. Jess is actually faster than some popular expert system shells written in C, especially on large problems, where performance is dominated by algorithm quality. Jess supports both forward and backward chaining inferencing mechanisms. In Jess, each individual jess.Rete object represents an independent reasoning engine. A single program can then include several independent engines. For our system, there are four independent reasoning modules. Jess can be used in a multithreaded environment, e.g. more than one user can access the expert system at the same time. The jess.Rete class internally synchronizes itself using several synchronization locks. The most important lock is a global lock on IF statements of all rules: only one assert or retract may be processing in a given jess.Rete object at a time. This restriction is likely to be relaxed in a future release. Jess and NRC Fuzzy Java Toolkit will be controlled entirely by the graphical user interface of the user-input module.

NRC Fuzzy Java Toolkit

The NRC Fuzzy Java Toolkit is a Java API for representing and manipulating fuzzy information. The toolkit was developed by the Integrated Reasoning Institute for Information Technology at National Research Council Canada. The access to the NRC Fuzzy Java Toolkit is free for use in a research or educational purposes. The toolkit consists of a set of classes that allow a user to build fuzzy systems in Java as well as a class of Jess UserFunctions that provide integration with Jess. Therefore, fuzzy rules can be created using only Jess code. As an object-oriented style, fuzzy concepts are implemented by a spectrum of Java classes that includes fuzzy variable, fuzzy set, fuzzy value, fuzzy modifier, fuzzy rule, and certainty factor. Since the Jess (or FuzzyJess) package is too big to be an applet we will use Jess/FuzzyJess only on the server side as a servlet.

Rete Algorithm

The Rete algorithm is the basis for a whole generation of fast expert system shells: OPS5, ART, CLIPS, and Jess. Without the Rete algorithm, a finding-fact approach will keep a list of the rules and continuously cycle through the list, checking each one's left-hand-side (LHS), the IF

statement, against the knowledge base and executing the right-hand-side (RHS), the THEN statement, of any rules that apply. This is inefficient because most of the tests made on each cycle will have the same results as on the previous iteration and its computational complexity is of the order of O(RF^P), where R is the number of rules, P is the average number of patterns per rule LHS, and F is the number of facts on the knowledge base. This escalates dramatically as the number of patterns per rule increases.

In the Rete algorithm, the inefficiency described above is alleviated (conceptually) by remembering past test results across iterations of the rule loop. Additionally, new facts are tested against only the rule LHSs to which they are most likely to be relevant. As a result, the computational complexity per iteration drops to something more like O(RFP), or linear in the size of the fact base.

The Rete algorithm is implemented by building a network of nodes, each of which represents one or more tests found on a rule LHS. Facts that are being added to or removed from the knowledge base are processed by this network of nodes. At the bottom of the network are the nodes that represent individual rules. Where a set of facts filters all the way down to the bottom of the network, it has passed all the tests on the LHS of a particular rule and this set becomes activated. The associated rule may have its RHS executed (fired) if the activation is not invalidated first by removal of one or more facts from its activation set.

User Inputs via Web Interface

The top tier (or presentation tier) running on the client via the web-based user interface can be accessed by any Java capable browser. The main interface will consist of four areas: control area, view area, parameter-input area, and information area. The control area will include the account controls, such as user name and password, project controls, file operations, and downloading of reports. The view area is designed to be interactive; selecting, zooming, or clicking on a map will invoke a view of a resized map and updated location information in the information area. All parameters and data related to queries will be grouped in the parameter-input area in a style of property pages in which one page includes all location-related parameters (possibly a map with 40 acre grids) and each of the other four pages contains parameters related to a specific reasoning module. Any progress or error information will be dynamically output from the information area.

Formulating Queries and Retrieving Data

After receiving a user request with contributed parameters, the server (the control tier) interprets parameters into queries that can be recognized by reasoning modules and retrieves related data from the database. Database driver(s) of Access or SQL Server will be created by Open Database Connectives (ODBC) to allow direct access by Java scripts. Besides the four reasoning modules, the control tier may also trigger other appropriate tool(s) in the base tier. The results (and possibly the execution progress) from the final report and other base tier are passed back to the control tier, which sends the final reply, properly formatted, to the top tier (the client side), where the user can see it from the view area and information window.

Regional Indication

Attributes from geophysical surveys, such as seismic, aeromagnetic, and gravity, plus the regional structure/isopach maps, are the only global parameters that can be obtained in the public domain. Measured data from wells, including wireline logs, core measurements, and production

data, are usually available only from established oil fields. It is expedient to find correlations between geophysical attributes and measured well data since any intrinsic correlation can become a global indicator especially for explored regions.

The current focus of applications regarding regional attributes is on correlation with production data and distributions of producing and non-producing wells. Examining data attributes such as slope and curvature can further optimize regional map data. Visual correlations between production and regional attributes have already been made, strongly suggesting that linear or neural network relationships exist and that production trends are mappable, and with neural networks, predictible away from well control.

In this reasoning module, fuzzy ranking and neural network techniques will be embedded into the inference engine. For example, a simple IF-THEN rule could be replaced by an IF-CALL A NEURAL NETWORK-THEN procedure. Predicted regional maps will be included in the overall evaluation report.

Trap Assessment

In the Brushy Canyon formation, reservoir sandstones can be divided into a series of major productive trends related to proximal/slope and more distal/basin-floor depositional settings. Oil entrapment is observed to be stratigraphic, mainly related to pinch-out of reservoir-quality sandstone. Structural highs have not localized oil accumulations. Identifying connections between a selected location (area) and any established production trend is the main focus of this reasoning module. Connection examples are if the selected location is in trend or on the strike of established production, downdip or updip of an established production, and a close distance to any established production. Additionally, structure pinch-out, sandstone thickness, average porosity, and dry hole indication will be all referenced as important facts.

Formation Assessment

Formation evaluation, well-log interpretation and correlation are among the earliest AI applications used in the petroleum industry. The reason is that there are no precisely defined "best" methods of solving problems for any given set of circumstances. Human experts in the field apply heuristics acquired from years of learning successful techniques for particular circumstances. Their decisions are often based on intangibles, such as patterns or shapes observed in log curves.

In the Brushy Canyon formation, significant clay content, lamination, and close interbedding between oil- and water- units make wireline log analysis and reserve estimates problematic. As a result, a success in formation assessment requires extra inputs from the expert knowledge and other indicators, such as cores, mud logs, and source-rock analysis. Neural network correlation and fuzzy inference will play an important role in this module. The module will estimate the local properties, porosity, permeability, organic richness, fluid (oil, gas, and water) saturations, and pay intervals at any given location of the Brushy Canyon formation. As part of module inputs, wireline logs, at and around the selected location, will be retrieved from the database at run time. The user inputs used by this module, to replace the default settings, are oil indication from the mud log, net thickness of sandstone, kerogen type, and etc.

Enhanced Recovery Scheme

As observed in the production data, well productivity is variable in different regions and trends, but primary recovery rarely exceeds 10% from the Brushy Canyon formation. Low permeability, typically between one to five md, is often the cause of a low primary recovery. Options for enhanced recovery will have to be factored into any risk analysis. Waterflooding and carbon dioxide flooding are schemes that have been tested. Some successful examples suggest that carbon dioxide flooding may be most appropriate in these low-permeability, clay-bearing formations. The expert knowledge of this module will be mainly based on case studies. For a selected location (area), the expert system will search for the cases that have the similar circumstance, region, deposition, composition, and productive trend. The retrieved cases will be summarized into the final summary report. With more and more cases being collected, this module will become the most active knowledge base in the future.

Overall Evaluation Report

The judging criteria for the quality of a selected location (area) will evolve from the fuzzy values with estimates of their uncertainty from the four reasoning modules, Regional Indication, Trap Assessment, Formation Assessment, and Enhanced Recovery Scheme. The first step for giving an overall evaluation is to defuzzify the fuzzy values into crisp values, which involves mostly human expertise and is subject to tuning and optimizing during the development phase. The second step, consisting of how to weight factors from different aspects into the overall assessment, varies by case and often depends on the operator's prospect. Users options, such as the weighting method, will be incorporated into related calculations. Moreover, oil price history and projected oil price will be provided in comparison with the operating cost estimated by the user. The final evaluation results will be provided in a report, which includes the assessments from the four reasoning modules with their uncertainty indexes, the methods used for the overall evaluation, and the overall risk indication. Additionally, major parameters, including some user inputs, will be tabulated by categories. Maps will be generated if the user checks the map option. This module will be coded in Java because it involves more complex computations than inference rules, and requires extra Java classes. In fact, maps and formatted data will be prepared in the control tier as part of the reply message.

Section 6—First Year Technology Transfer

This section relates the technology transfer activities of the group as a whole, through their presence at professional meetings as well as their papers and presentations. Additionally, future technology transfer activities are discussed.

Section 6—First Year Technology Transfer

The Fuzzy Expert Exploration Tool project was well represented at the 2000 SPE Permian Basin Oil and Gas Recovery Conference. A session devoted to computing gave the project members an opportunity to showcase advances in computational intelligence and some specific case studies where unique applications of those technologies were made. The three paper presented in the Computing Session were given by project workers. Shaochang Wo presented **SPE 59553,** "New Technique To Determine Porosity And Deep Resistivity From Old Gamma Ray And Neutron Count Logs," Robert Balch presented **SPE 59554**, "Predicting Core Porosity Using Wire-Line Logs At Dagger Draw Field, SE New Mexico," and graduate student Darren Hart presented **SPE 59555,** "Time-To-Depth Conversion Of Nash Draw "L" Seismic Horizon Using Seismic Attributes And Neural Networks."

Attendees showed a great deal of interest, resulting in numerous questions on how these techniques could be used in other projects. The most common questions revolved around software availability and utility. Interested attendees were referred to poster presentation **SPE 62810,** "Risk Reduction with a Fuzzy Expert Exploration Tool," presented by Bill Weiss, which outlined the FEE Tool Project in the Delaware Basin of New Mexico. Several prospective consortium members resulted from the poster and new sources of seismic and core data appear to have resulted from the conference.

During the first year the following papers were presented. All papers resulted from application of concepts developed under **Risk Reduction with a Fuzzy Expert Exploration Tool** research.

1. Weiss, W.W., Wo, S. and Balch, R.S.: "Integrating Core Porosity and Sw Measurements with Log Values," paper SPE 55642 presented at the 1999 SPE Rocky Mountain Regional Meeting, Gillette, May 15–18.
2. Weiss, W. W.: "Advanced Oil Recovery Technologies for Improved Recovery from Slope Basin Clastic Reservoirs," paper presented at the 1999 Oil & Gas Conference, DOE Office of Fossil Energy, Dallas, June 28–30.
3. Wo, S., Weiss, W.W., Balch, R.S., Scott, L., Roe, J., and Kendall, R.: "Producing GOR Used to Predict Permeability Distribution in a Tight Heterogeneous Reservoir," paper SPE 56505 presented at the 1999 SPE Annual Technical Conference, Houston, October 3–6.
4. Balch, R.S., Stubbs, B.S., Weiss, W.W., and Wo, S.: "Using Artifical Intelligence to Correlate Multiple Seismic Attributes to Reservoir Properties," paper SPE 56733 presented at the 1999 SPE Annual Technical Conference, Houston, October 3–6.
5. Weiss, W.W., Wo, S., Balch, R.S., Scott, L., and Kendall, R.P.: "Assessing the Potential Redevelopment of a 1960's Vintage Oil Field," paper SPE 59297 presented at the 2000 SPE/DOE Improved Oil Recovery Symposium, Tulsa, April 3–5.
6. Wo, S., Weiss, W.W., Balch, R.S., Scott, L., and Kendall, R.P.: "New Technique to Determine Porosity and Deep Resistivity from Old Gamma Ray and Neutron Count Logs," paper SPE 59553 presented at the 2000 SPE Permian Basin Oil and Gas Recovery Conference, Midland, March 21–23.
7. Balch, R.S., Weiss, W.W., and Wo, S. and Welch, D.M.: "Predicting Core Porosity Using Wireline Logs at Dagger Draw Field, Southeast New Mexico," paper SPE 59554 presented at the 2000 SPE Permian Basin Oil & Gas Recovery Conference, Midland, March 21–23.
8. Hart, D.M., Balch, R.S., Tobin, H.J., and Weiss, W.W.: "Time-to-Depth Conversion of Nash Draw "L" Seismic Horizon Using Seismic Attributes and Neural Networks," paper SPE 59555 presented at the 2000 SPE Permian Basin Oil and Gas Recovery Conference, Midland, March

21–23.
9. Weiss, W.W., Sung, A.H., and Broadhead, R.: "Risk Reduction with a Fuzzy Expert Exploration Tool," poster SPE 62810 presented at the 2000 SPE Permian Basin Oil and Gas Recovery Conference, Midland, March 21–23.

Future

An available web-based shell, FuzzyJess, will be the core of the FEE Tool. Rules will be developed with the FuzzyJess language to estimate the likelihood of drilling a successful Brushy Canyon well on any of the 40-acre records in the database. A Java code will be developed to extract a record from the database. Following development, the FEE Tool will be evaluated and the rules adjusted via validation testing in a manner similar to neural network blind testing.

A database in Microsoft Access format consisting of approximately 50,000, 40-acre records with the associated fields (database terminology for entries) will be created. Each record will have perhaps 100 fields. Each record will contain the regional entries gravity, aeromagnetic, and geologic interpolations, plus the respective attributes. Sparse well information such as porosity, permeability, saturations, production, and oil shows, will be included as a field where available.

A trial and error approach is envisioned for generating the first generation Brushy Canyon interval rules. A second generation FEE Tool will evolve during the Permian Basin Devonian formation phase of the project.

Plans for Upcoming Geologic Project Year

During the second project year the geologic analysis and data acquisition of the petroleum source rocks as well as the reservoir analysis of the upper Brushy Canyon Formation will be concluded. Derivative source rock parameters will be mapped that assess areal and stratigraphic variations in generative potential and productive capacity of source rocks as well as distribution of source facies in both the lower Brushy Canyon and the upper Brushy Canyon. These will be compared with mappable production parameters such as presence or absence of established production and initial production rates of oil, gas, and water. Net thickness maps of porous sandstones (at the 10% and the 15%) levels will be made for the upper Brushy Canyon in order to determine controls of reservoir distribution on known production. A link between structure and reservoir distribution that is similar to that for the lower Brushy Canyon (established during the first project year) will be evaluated. Areas of enhanced reservoir distribution will be investigated to determine fuzzy logic approaches to identifying productive and non-productive areas within the reservoir fairways. Finally, all geologic and geochemical information will be input into the fuzzy tool in the development of the expert system.

Geologic Tech transfer efforts will be aggressively pursued within the next year. Presentations will be made on our efforts at the 2000 AAPG Rocky Mountain Section Meeting (September 2000, Albuquerque NM) and at the Annual West Texas Geological Society Symposium (November 2000, Midland TX). Written papers will be submitted for publication in conjunction with the latter meeting.

Conclusions

All goals for the first project year were met. Primary goals were data acquisition and assimilation, and construction of needed computational intelligence tools. Ninety percent of the needed data has been acquired and is in the database or in the process of being input. Geological and geophysical data, and their derived attributes cover wide areas of the basin. Tools have developed and tested which allow correlating the production histories of 2200 wells in the Delaware basin with the regional attribute maps. An online neural network program (*PredictOnline*) was developed and is being used as a prototype for development of the online Fuzzy Expert System interface.

With data and correlating tools prepared, progress toward the development of the rules and online interface for the Fuzzy system will move forward. Involvement of consortium members will increase during the development and testing phases. The first year research has resulted in nine papers presented at meetings.

References

1. Broadhead, R.F., Luo, F., and Speer, S.W.: "Oil And Gas Resources at the Waste Isolation Pilot Plant (WIPP) site, Eddy County, New Mexico," New Mexico Bureau of Mines and Mineral Resources Circular 206 (1998)
2. Telford, W.M., L.P. Geldart and R.E. Sheriff. *Applied Geophysics*. Second Edition, Cambridge University Press (1990).
3. Hart, D.M., Balch, R.S., Weiss, W.W., and Wo, S.: "Time-to-Depth Conversion of Nash Draw "L" Seismic Horizon Using Seismic Attributes and Neural Networks," paper SPE 59555 presented at the Permian Basin Oil and Gas Recovery Conference, Midland, March 21–23.
4. Montgomery, S.L., Worrall, J., and Hamilton, D.: "Delaware Mountain Group, West Texas and Southeastern New Mexico, A Case of Refound Opportunity: Part 1-Brushy Canyon," *AAPG Bulletin* (December 1999) **83**(12) 1901–1926.
5. Martin, F.D., Murphy, M.B., Stubbs, B.A., Uszynski, B.J., Hardage, B.A., Kendall, R.P., Whitney, E.M., and Weiss, W.W.: "Reservoir Characterization as a Risk Reduction Tool at the Nash Draw Pool," *SPE Reservoir Evaluation and Engineering* (April 1999) **2** (2).
6. Arbogast, J.S. and Frankline, M.H.: "Artificial Neural Networks And High Speed Resistivity Modeling Software Speeds Reservoir Characterization," two-part series, *Petroleum Engineer International* (May 1999).
7. Gaymard, R. and Poupon, A.: "Response of Neutron and Formation Density Logs in Hydrocarbon Bearing Formations," *The Log Analyst* (Sept.–Oct. 1968).
8. Chawathe, A., Ouenes, A. and Weiss, W. W.: "Predicting Unconventional Well Logs from Conventional Logs Using Neural Networks," *In Situ*, (1997) **21**(2) 145–159.
9. Arbogast, J. S. and Franklin, M. H.: "Artificial Neural Networks and High Speed Resistivity Modeling Software Speeds Reservoir Characterization," Petroleum Engineer International (June 1999).
10. Haykin, S.: *Neural Networks*, Macmillan College Publishing Company, Inc. (1994).
11. Powell, M.J.D.: "Radial Basis Functions for Multivariable Interpolation: A Review," in: J.C. Mason and M.G. Cox, editors, *Algorithms for Approximation*. Clarendon Press, Oxford (1987).
12. Poggio, T. and Girosi, F.: "Regularization Algorithms for Learing That Are Equivalent to Multilayer Networks," *Science*, (February 1990) **247**, 978-982.
13. Poggio, T. and Girosi, F.: "Networks for Approximation and Learning," *Proceedings of the IEEE*, (September 1990) **78** (9) 1481–1497.
14. Lloyd, S. P.: "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory*, (March 1982) **28**(2).
15. Moody, J. and Darken, C. J.: "Fast Learning in Networks of Locally-Tuned Processing Units," *Neural Computation* (1989) **1**, 281-294.
16. http://www.informatik.uni-stuttgart.de/ipvr/bv/projekte/snns/snns.html and http://www-ra.informatik.uni-tuebingen.de/SNNS/ ...network.
17. http://ford.nmt.edu/PredictOnline6/.
18. Deitel, H.M. and Deitel, P.J.: *Java How to Program.* Prentice Hall (1997).
19. Ernest J. Friedman-Hill: *Jess, The Java Expert System Shell*, Version 5.0 released on January 28, 2000, Sandia National Laboratories, Livermore, CA.

20. R. A. Orchard: "*NRC Fuzzy Java Toolkit, User's Guide*," Version 1.0 Beta 1, Integrated Reasoning Institute for Information Technology of National Research Council Canada, (November 1999).
21. *Expert Systems in Engineering Applications*, SPE Reprint Series No. 41, Society of Petroleum Engineers, Richardson, TX, 1996.
22. Forgy, Charles L.: "*Rete: A Fast Algorithm for the Many Pattern/Many Object Pattern Match Problem*," *Artificial Intelligence* (1982) **19**, 17–37.

Table 1. Regional Maps		
SOURCE	*MAP*	*EXPLANATION OR USE*
Structure	Brushy Canyon Subsea	Structural traps in Brushy Canyon
Structure	Dip Azimuth	data trends – faults - anticlines
Structure	Dip Magnitude	Steepness or scale of data trends
Structure	Curvature Azimuth	Rate of change in dip – flexure
Structure	Curvature Magnitude	Scale of curvature changes – fracture indicator
Structure	X-Derivative	X directional derivative - data trends of slope
Structure	Y-Derivative	Y directional derivative - data trends of slope
Structure	2nd X-Derivative	X directional derivative - rate of slope change
Structure	2nd Y-Derivative	Y directional derivative - rate of slope change
Gravity	Bouger Anomalies	Variations in regional densities incl. basement
Gravity	X-Derivative	X direction rate of change in gravity
Gravity	Y-Derivative	Y direction rate of change in gravity
Gravity	2nd X-Derivative	Removes basement features–sedimentary gravity
Gravity	2nd Y-Derivative	Removes basement features–sedimentary gravity
Gravity	Dip Azimuth	Data trends
Gravity	Dip Magnitude	Data trends
Gravity	Curvature Azimuth	Data trends
Gravity	Curvature Magnitude	Data trends
Aeromag	Aeromagnetic anomalies	Variations in magnetic susceptibility
Aeromag	Dip Azimuth	Differentiation of basement blocks?
Aeromag	Dip Magnitude	Data trends – scale of susceptibilities
Aeromag	Curvature Azimuth	Data trends
Aeromag	Curvature Magnitude	Data Trends
Aeromag	X-Derivative	Data Trends
Aeromag	Y-Derivative	Data Trends
Aeromag	Second X-Derivative	Sedimentary section aeromagnetics
Aeromag	Second Y-Derivative	Sedimentary section aeromagnetics

Table 2. Relative Errors at Each Control Point for Three Time-to-Depth Conversion Tools							
Well	Depth	MLP Dep	MLP-Err	TDQ Dep	TDQ Err	Zmap Dep	Z,map Err
T-Fed	7081.5	7073.9	7.62	6852.11	*229.39*	6976.84	*104.66*
T-Fee	6994.0	6950.3	43.66	6864.6	*129.4*	6946.35	*47.65*
38	6800.0	6816.5	-16.53	6787.65	*12.35*	6795.89	*4.11*
25	6764.0	6763.9	0.06	6764.0	0	6762.86	1.14
29	6803.0	6807.0	-4.02	6803.0	0	6803.48	-0.48
5	6882.0	6876.9	5.1	6882.0	0	6880.50	1.50
1	6862.0	6866.5	-4.47	6862.0	0	6860.96	1.04
6	6891.5	6891.4	0.12	6891.5	0	6889.74	1.76
10	6824.0	6807.1	16.9	6824.0	0	6823.66	0.34
24	6746.0	6747.3	-1.31	6746.0	0	6746.51	-0.51
23	6753.5	6749.5	4.02	6753.5	0	6754.01	-0.51
20	6912.0	6913.6	-1.61	6912.0	0	6912.20	-0.20
9	6807.5	6822.7	-15.18	6807.5	0	6808.40	-0.90
14	6850.0	6846.2	3.78	6850.0	0	6850.20	-0.20
15	6771.5	6767.9	3.62	6771.5	0	6770.92	0.58
11	6773.0	6775.7	-2.74	6773.0	0	6773.32	-0.32
12	6794.0	6799.4	-5.44	6794.0	0	6793.72	0.28
19	6787.5	6787.7	-0.22	6787.5	0	6787.51	-0.01

Table 3. Results of Fuzzy Ranking with Dta Set from Nash Draw #23 Well with Full Core					
Input Wire-line	Output Core Φ	Output Core S_o	Output Core S_W	Output Core A_{ir}	Output Core $\Phi * S_o$
Caliper	.25	.25	.25	.25	.25
DPHI	.33	.33	.33	.33	.33
DRHO	.17	.17	.17	.17	.17
GR	.20	.20	.20	.20	.20
LLD	.36	.36	.36	.36	.36
LLD	.36	.36	.36	.36	.36
MSFL	.22	.22	.22	.22	.22
Log LLD	.27	.27	.27	.27	.27
Log LLS	.28	.28	.28	.28	.28
Log MSFL	.21	.21	.21	.21	.21
NPHI	.26	.26	.26	.26	.26
NPOR	.25	.25	.25	.25	.25
PEF	.33	.33	.33	.33	.33
RHOB	.33	.33	.33	.33	.33
SP	.41	.41	.41	.41	.41
TNPH	.32	.32	.32	.32	.32

Table 4. Table Of Conventional Correlation Coefficients From Cross Plots. Correlation Coefficients Listed Represent Best-Fit Line	
Variables	Correlation Coefficients, %
Core Φ vs. Core Perm.	0.84
Core Φ vs. DPHI	0.74
LLD vs. Core Φ	0.65
SQRT(Φ) vs. Core Φ	0.60
S_w vs. Core Φ	0.45
Core Φ vs. NPHI	0.35
S_w vs. S_o	0.69
S_w vs LLD	0.53
S_o vs. SQRT(Φ)	No Correlation
S_w vs. DPHI	0.39
LLD vs. Core Perm.	0.51
LLD vs. DPHI	0.68

SQRT(Φ) is the same as SQRT((NPHI^2)+(DPHI^2))

Table 5. Nash Draw Neural Network Correlations. The Neural Net Program Calculates Correlation Coefficient as the Best-fit Line			
Input variables	Output variable	Best Architecture	Trn/tst cc, %
*DPHI, PEF, SP	Core Φ	3-4-2-1	0.78/0.73
*Log LLD, Log LLS, PEF, SQRT(Φ)	Core Φ	4-6-2-1	0.78/0.68
*Log LLD, Log LLS, SQRT(Φ)	Core Φ	3-6-6-1	0.71/0.66
*DPHI, Log LLD, Log LLS, NPHI	Core Φ	4-4-3-1	0.78/0.74
*DPHI, Log LLD, Log LLS, NPOR, PEF	Core Φ	5-5-4-1	0.80/0.74
*DPHI, Log LLD, Log LLS, NPOR, PEF, GR	Core Φ	6-7-2-1	0.80/0.72
*DPHI, LLD, LLS, PEF, SP	Core Φ	5-6-5-1	0.78/0.73
*DPHI, Log LLD, Log LLS, PEF	Core Φ	4-4-3-1	0.78/0.73
*DPHI, Log LLD, Log LLS	Core Φ	3-5-2-1	0.78/0.74
*Cal, DPHI, GR, LLD, LLS, MSFL, NPHI, NPOR, PEF, RHOB, SP	Core Φ	11-6-5-1	0.84/0.76
*Cal, DPHI, GR, LLD, LLS, MSFL, NPOR, PEF, SP	Core Φ	9-3-3-1	0.75/0.74
*Cal, DPHI, GR, LLD, LLS, MSFL, NPOR, NPHI, PEF, SP	Core Φ	10-4-3-1	0.81/0.74
**DPHI, LLD, LLS, PEF, SP	Core Φ	5-5-5-1	0.76/0.71
**DPHI, LLD, LLS	Core Φ	3-6-5-1	0.78/0.69
***DPHI, Log LLD, Log LLS	Core Φ	3-6-3-1	0.67/0.71
***DPHI, Log LLD, Log LLS, NPHI, PEF	Core Φ	5-5-5-5-1	0.70/0.78
***DPHI, Log LLD, Log LLS, NPHI, PEF, GR	Core Φ	6-5-2-1	0.80/0.77
***DPHI, Log LLD, Log LLS, NPOR, PEF	Core Φ	5-6-2-1	0.70/0.76
***DPHI, NPOR, PEF	Core Φ	3-4-3-1	0.68/0.78
***DPHI, NPHI, PEF	Core Φ	3-7-4-1	0.69/0.74
****Cal, DPHI, GR, Log LLD, Log LLS, Log MSFL, NPHI, NPOR, PEF, SP	Core Φ	10-5-5-1	0.70/0.74
*DPHI, PEF, SP	Core S_w	3-7-7-1	0.54/0.3
*DPHI, LLD, LLS, PEF, SP	Core S_w	5-6-3-1	0.60/0.27
*GR, LLD, PEF, SP	Core S_w	4-5-4-4-1	0.57/0.55
*GR, LLS, PEF, SP	Core S_w	4-4-4-1	0.57/0.54
*LLD, PEF, SP	Core S_w	3-5-3-1	0.57/0.54
*GR, Log LLD, PEF, SP	Core S_w	4-5-3-1	0.57/0.53
*GR, LLD, PEF, SP	Core S_w	No Correlation	No Correlation
*Cal, DPHI, GR, LLD, LLS, MSFL, NPOR, NPHI, PEF, SP	Core S_w	10-6-2-1	0.57/0.52
*Log LLD, Log LLS, PEF, SQRT(Φ)	Core S_o	4-4-3-1	0.47/0.51
*DPHI, LLD, LLS, PEF, SP	Core S_o	No Correlation	No Correlation
*DPHI, LLD, PEF	Core S_o	No Correlation	No Correlation
*GR, LLD, PEF, SP	Core S_o	No Correlation	No Correlation
*Cal, DPHI, GR, LLD, LLS, MSFL, NPOR, NPHI, PEF, SP	Core S_o	No Correlation	No Correlation
*Log LLD, Log LLS, PEF, SQRT(Φ)	Core S_o*Φ	4-6-3-1	0.65/0.42
*Cal, DPHI, GR, LLD, LLS, MSFL, NPOR, NPHI, PEF, SP	Core S_o*Φ	10-3-2-1	0.61/0.51

*Used 172 training points 42 test points (pulled every 5 for testing).

**Used 161 training points, 53 test points (pulled every 4 for testing).

***Used 145 training points, 36 test points (pulled every 5 for testing, excluded all LLD, LLS and MSFL <100 ohm).

SQRT (Φ) = SQRT (NPHI^2+DPHI^2)/2)

Table 6. Nash Draw #23 Neural Network Training/Test Correlation Coefficients				
Input Variables	Output Var.	Train/Test 2 ft Inter-val	Train/Test 3 ft Inter-val	Train/Test 2.5 ft In-terval
DPHI, PEF, SP	CoreΦ	085/084	0.83/0.81	0.83/0.89
Log LLD, Log LLS, PEF, SQRT(PHI)	CoreΦ	0.90/0.85	0.90/0.85	0.91/0.81
Log LLD, Log LLS, SQRT(PHI)	CoreΦ	0.89/0.85	0.90/0.81	0.88/0.91
DPHI, Log, LLD, Log LLS, NPHI	CoreΦ	0.90/0.83	0.89/0.88	0.88/0.94
DPHI, Log LLD, Log LLS, TNPH, PEF	CoreΦ	0.90/0.88	0.90/0.88	0.90/0.82
DPHI, Log LLD, Log LLS, TNPH, PEF, GR	CoreΦ	0.90/0.90	0.92/0.89	0.90/0.88
DPHI, LLD, LLS, PEF, SP	CoreΦ	0.870.87	0.82/0.92	0.84/0.86
DPHI, Log LLD, Log LLS, PEF	CoreΦ	0.89/0.91	0.89/0.89	0.89/0.85
DPHI, Log LLD, Log LLS	CoreΦ	0.90/0.82	0.88/0.91	0.88/0.90
CAL, DPHI, GR, LLD, LLS, MSFL, NPHI, TNPH, PEF, RHOB, SP	CoreΦ	0.88/0.86	0.88/0.85	0.89/0.85
CAL, DPHI, GR, LLD, LLS, MSFL,TNPH, PEF, SP	CoreΦ	0.88/0.86	0.87/0.89	0.87/0.89
CAL, DPHI, GR, LLD, LLS, MSFL, TNPH, NPHI, PEF, SP	CoreΦ	0.88/0.86	0.87/0.86	0.87/0.90
DPHI, LLD, LLS	CoreΦ	0.85/0.85	0.84/0.84	0.84/0.85
DPHI, Log LLD, Log LLS, NPHI, PEF	CoreΦ	0.89/0.88	0.89/0.87	0.90/0.84
DPHI, Log LLD, Log LLS, NPHI, PEF, GR	CoreΦ	0.88/0.89	0.90/0.84	0.90/0.86
DPHI, TNPH, PEF	CoreΦ	0.86/0.83	0..84/0.84	0.84/0.84
DPHI, NPHI, PEF	CoreΦ	0.85/0.85	0.84/0.83	0.84/0.83
CAL, DPHI, GR, Log LLD, Log LLS, Log MSFL, NPHI, TNPH, PEF, SP	CoreΦ	0.91/0.92	0.90/0.87	0.91/0.84
DPHI, PEF, SP	Core S_w	0.73/0.77	0.75/0.74	0.73/0.73
DPHI, LLD, LLS, PEF, SP	Core S_w	0.74/0.65	0.76/0.72	0.76/0.70
GR, LLD, PEF, SP	Core S_w	0.71/0.74	0.74/0.77	0.75/0.77
GR, LLS, PEF, SP	Core S_w	0.72/0.65	0.75/0.77	0.74/0.80
LLD, PEF, SP	Core S_w	0.72/0.73	0.74/0.74	0.74/0.77
GR, Log LLD, PEF, SP	Core S_w	0.71/0.73	0.73/0.79	0.74/0.78
CAL, DPHI, GR, LLD, LLS, MSFL, TNPH, NPHI, PEF, SP	Core S_w	0.75/0.73	0.79/0.71	0.79/0.82
Log LLD, Log LLS, PEF, SQRT(PHI)	Core S_o	0.52/0.50	0.53/0.49	0.52/0.50
DPHI, LLD, LLS, PEF, SP	Core S_o	0.49/0.46	0.52/0.57	0.53/0.59
DPHI, LLD, PEF	Core S_o	0.45/0.47	0.53/0.47	0.53/0.30
GR, LLD, PEF, SP	Core S_o	0.52/0.56	0.54/0.51	0.51/0.51
CAL, DPHI, GR, LLD, LLS, MSFL, TNPH, NPHI, PEF, SP	Core S_o	0.67/0.58	0.68/0.58	0.67/0.45
Log LLD, Log LLS, PEF, SQRT(PHI)	Core S_o*Φ	0.80/0.80	0.82/0.79	0.82/0.76
CAL, DPHI, GR, LLD, LLS, MSFL, TNPH, NPHI, PEF, SP	Core S_o*Φ	0.82/0.63	0.83/0.73	0.81/0.74

Table 7. Expert System Shells			
System	Vendor	Pro	Con
CLIPS	NASA	Free code, well documented, numerous examples.	No fuzzy logic or internet interface.
Fuzzy CLIPS	IIT, National Research Council of Canada	Incorporates fuzzy logic	Not commercial, incomplete documentation, no internet interface.
Clips/R2	Production Systems Technology	Fastest algorithm	No fuzzy logic or internet interface
ECLIPS	Haley Enterprise, Inc.	Fully commercial application with internet interface.	No fuzzy logic

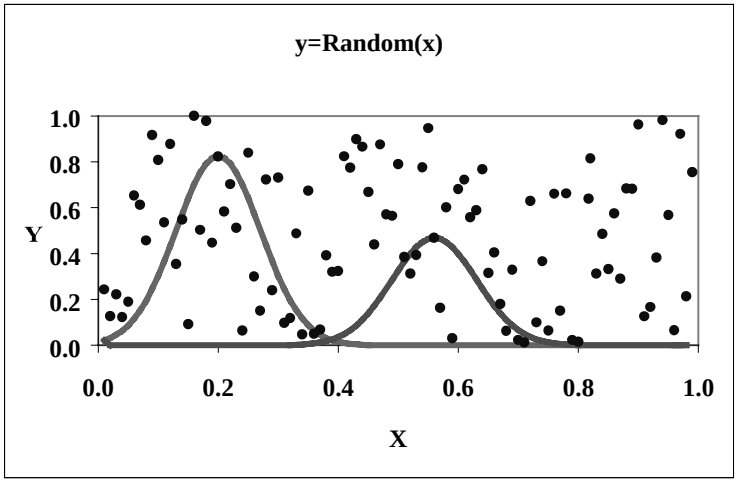

Figure 1. Conventional cross plot of a random data set (0-1). No correlation between X and Y. The trend is 0.5 (average between 0 and 1) is not evident. For each point a fuzzy membership function is defined, two example functions are shown on this plot..

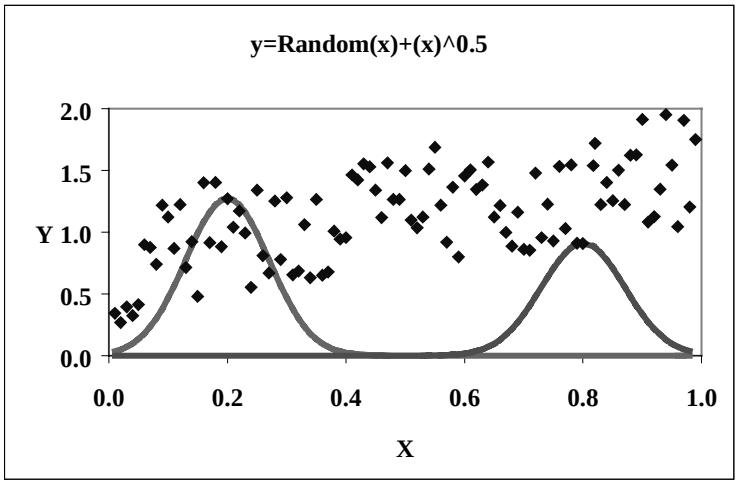

Figure 2. A conventional cross plot of a random data set (0-1) plus a square root trend. Note the apparent correlation. Again two sample fuzzy membership functions are illustrated.

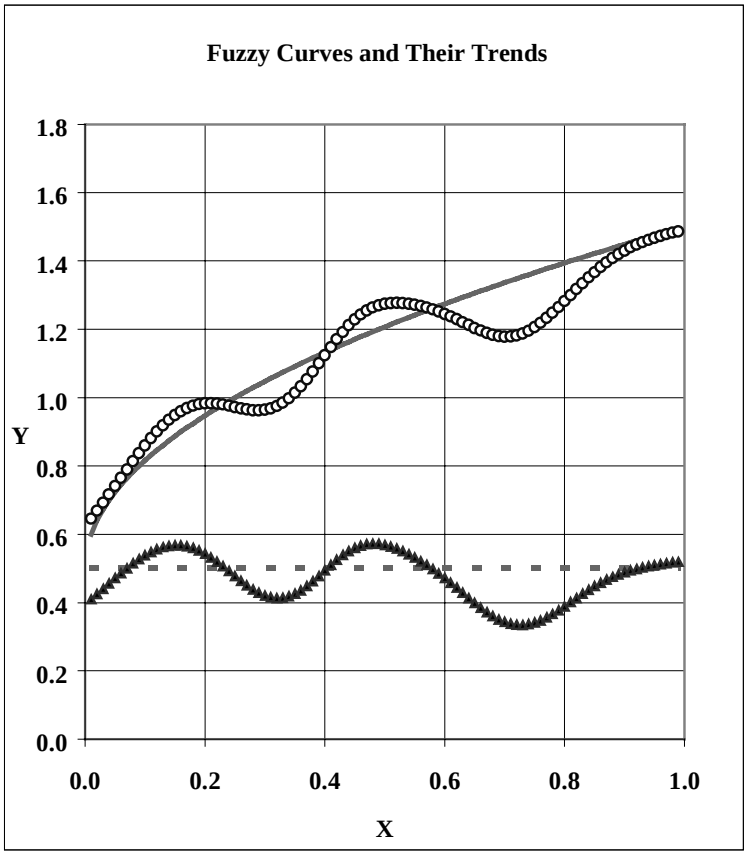

Figure 3. Fuzzy ranking curves. The trends are clearly evident.

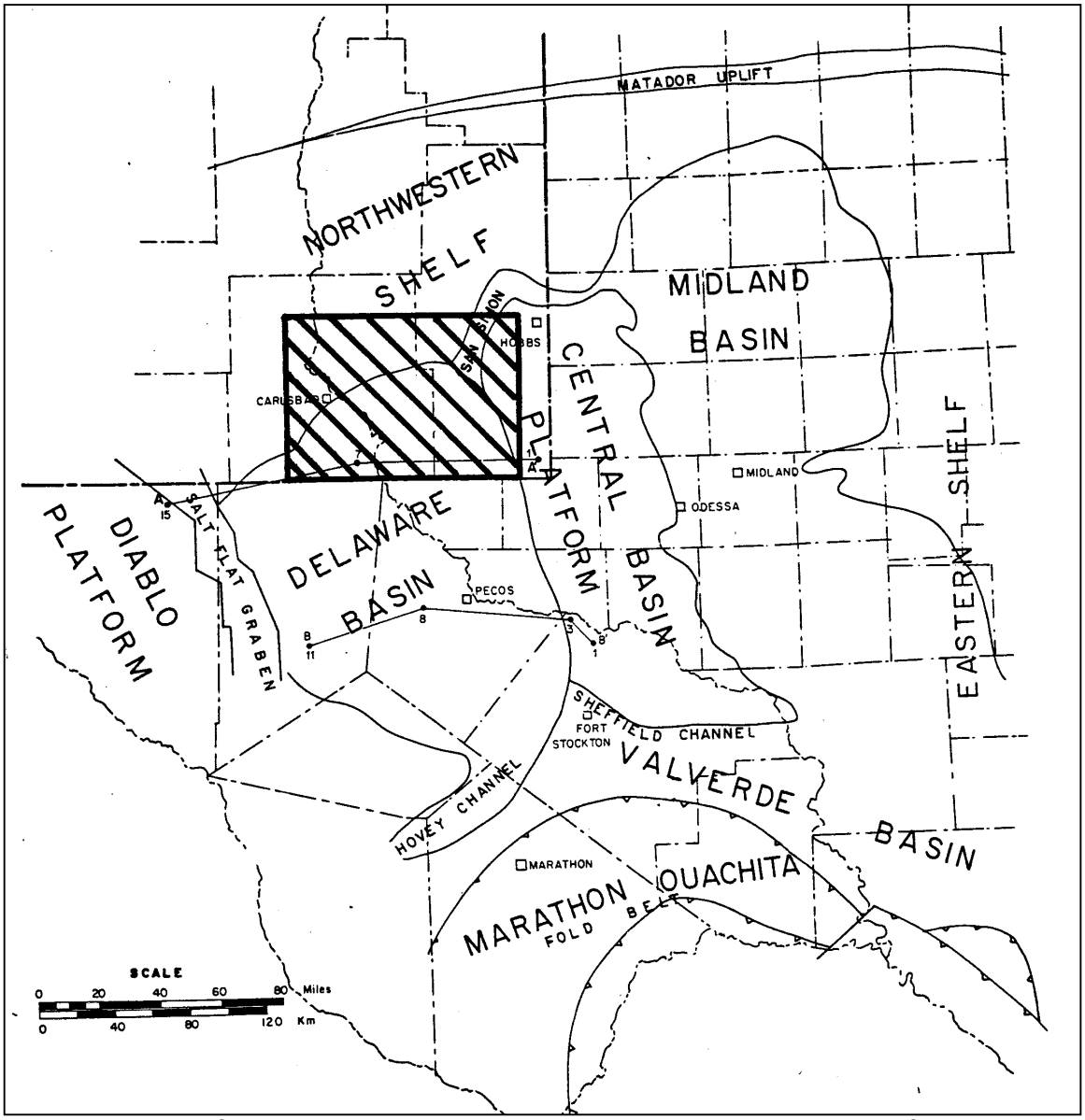

Figure 4. Map of Permian Basin showing data area used in development of FEE Tool.

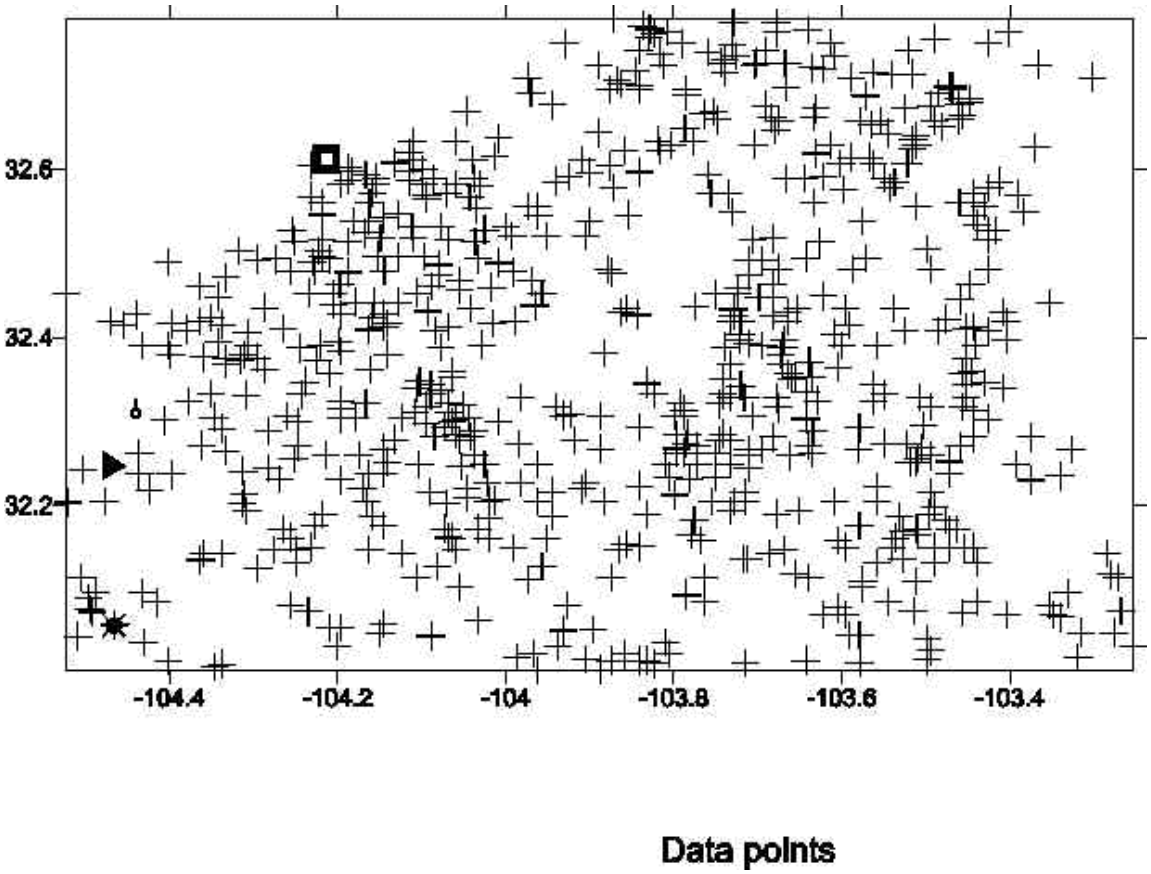

Figure 5. Map of Brushy Canyon project area showing locations of wells used for geologic data for FEE Tool.

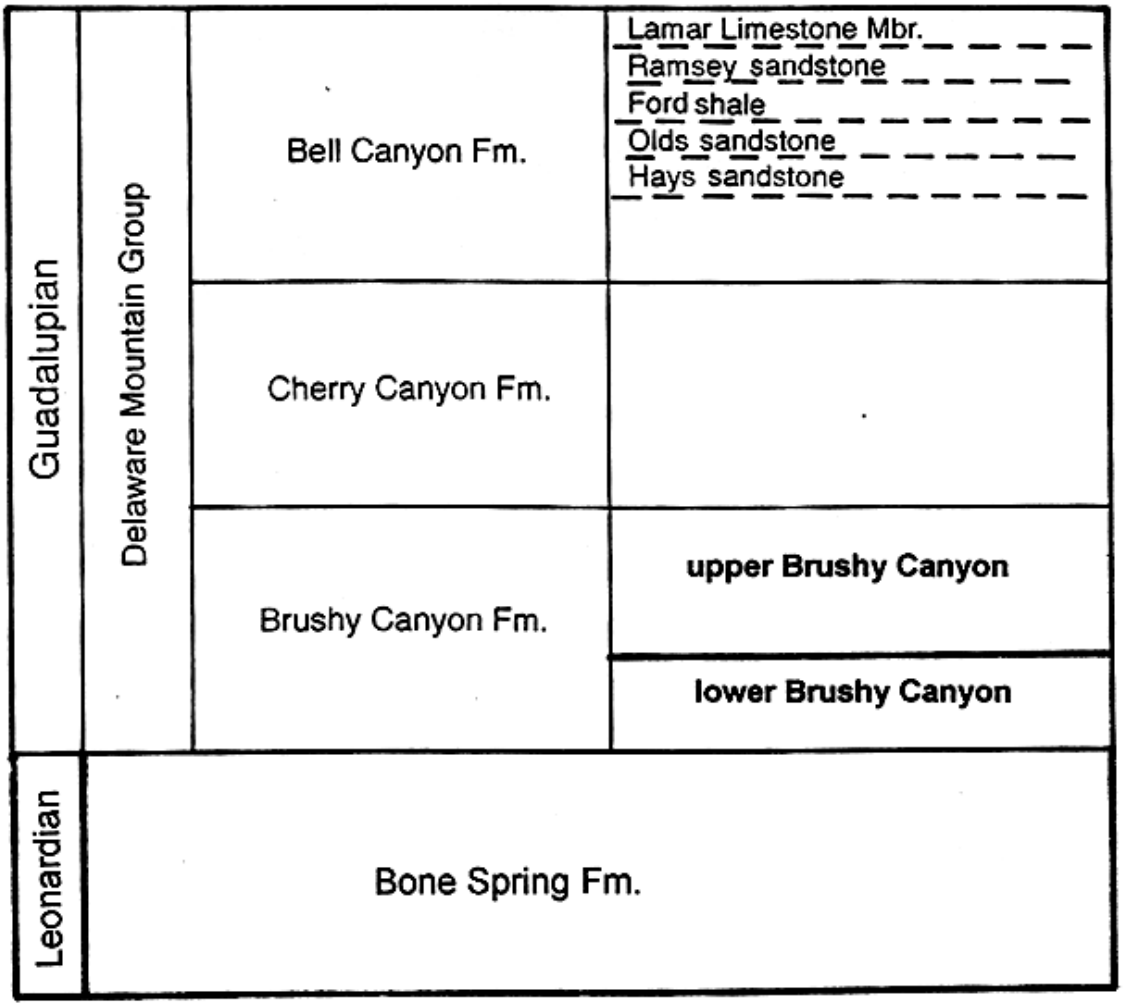

Figure 6. Stratigraphic column of Delaware Mountain Group in Delaware Basin.

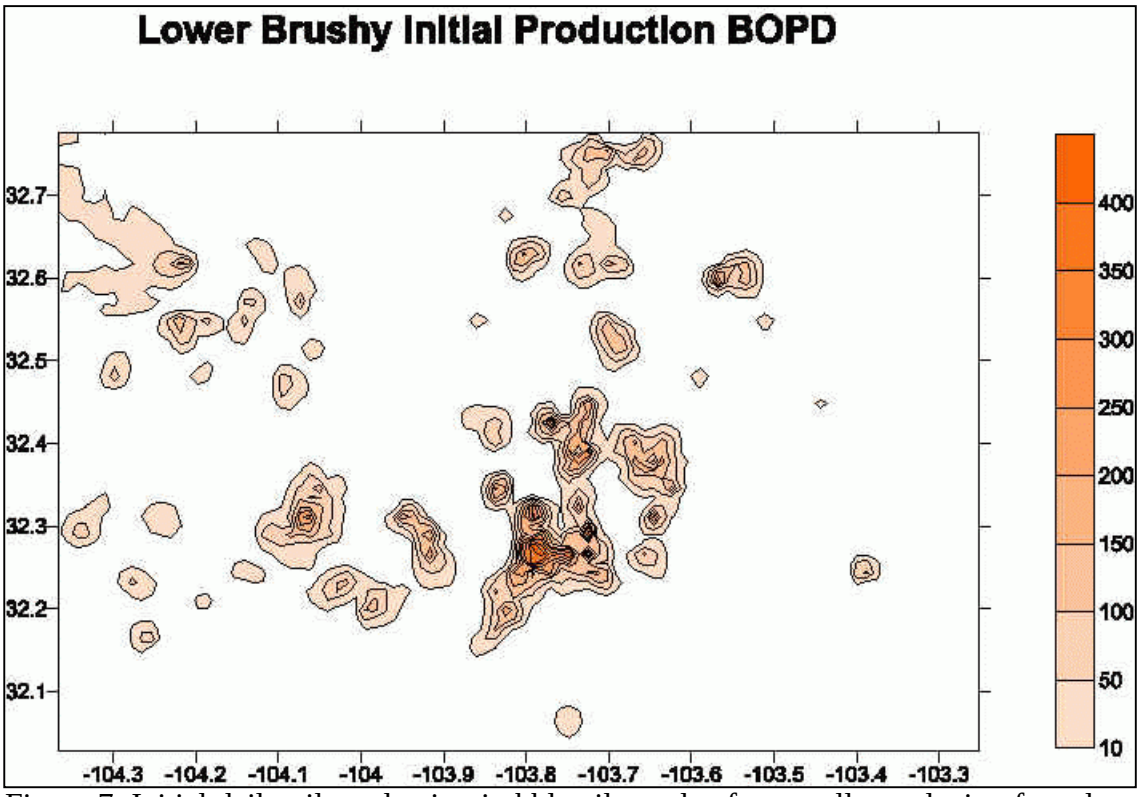

Figure 7. Initial daily oil production in bbls oil per day from wells producing from lower Brushy Canyon Formation.

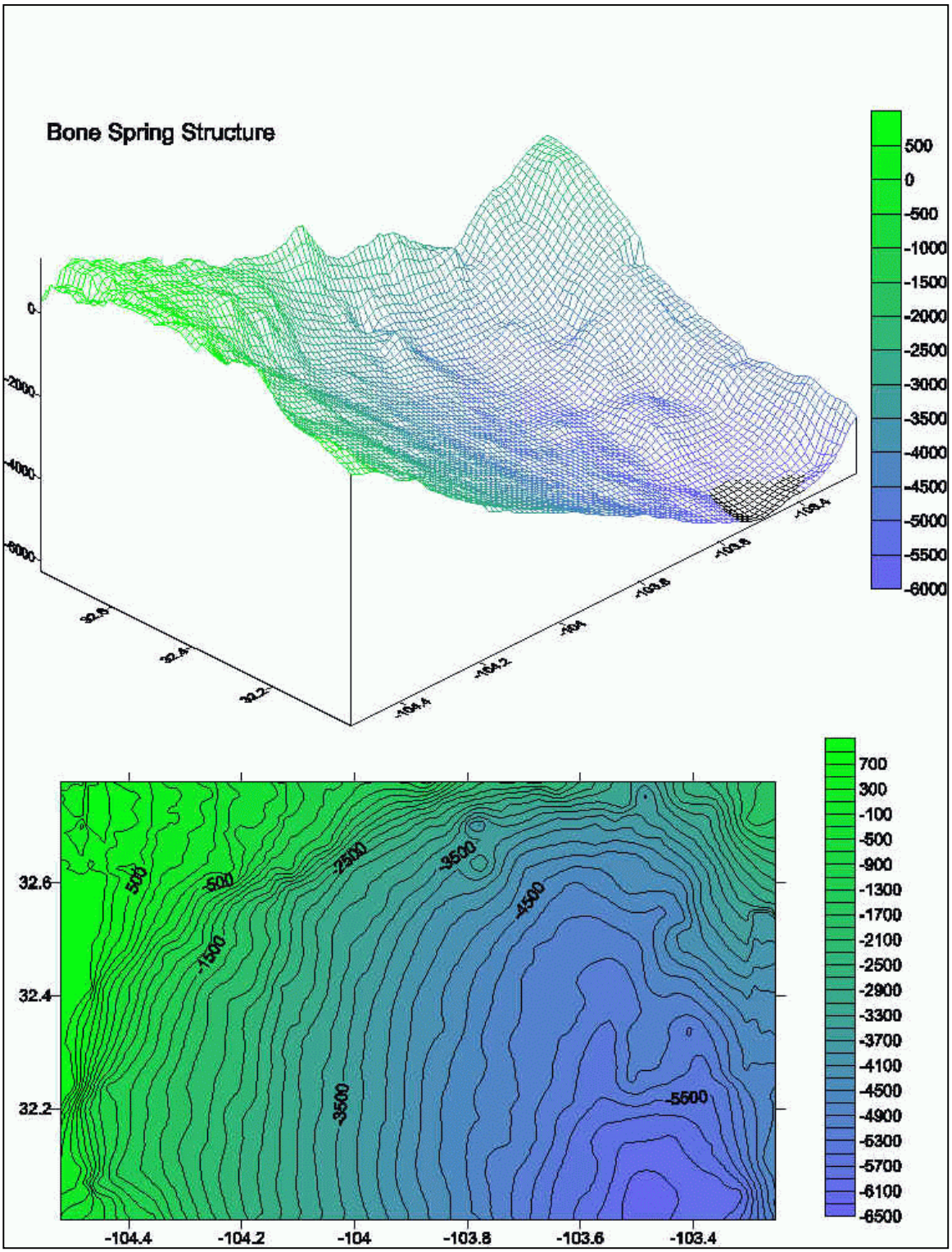

Figure 8. Structure contour map and wireframe relief map showing structure on top of Bone Spring FM.

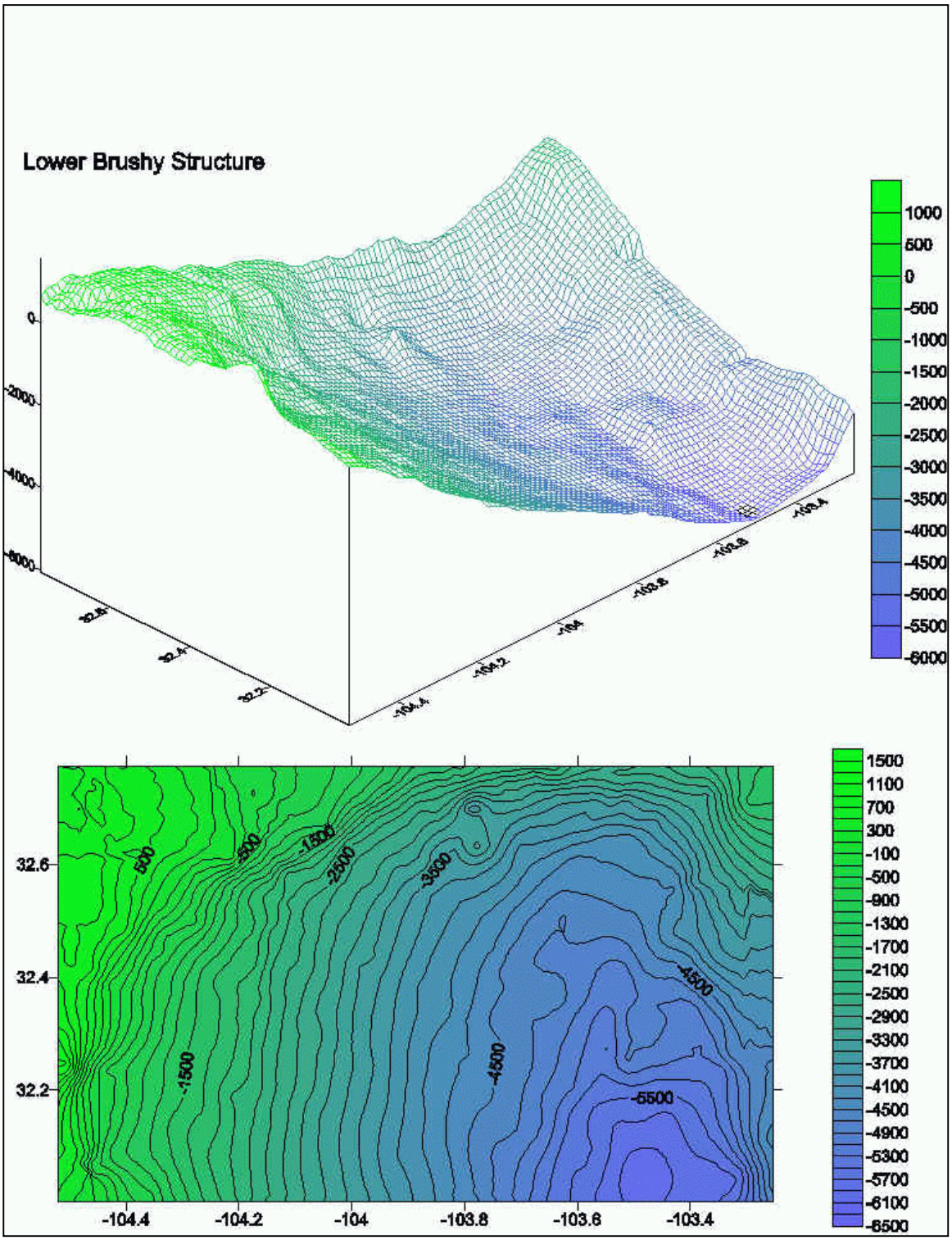

Figure 9. Structure contour map and wireframe relief map showing structure on top of lower Brushy Canyon formation.

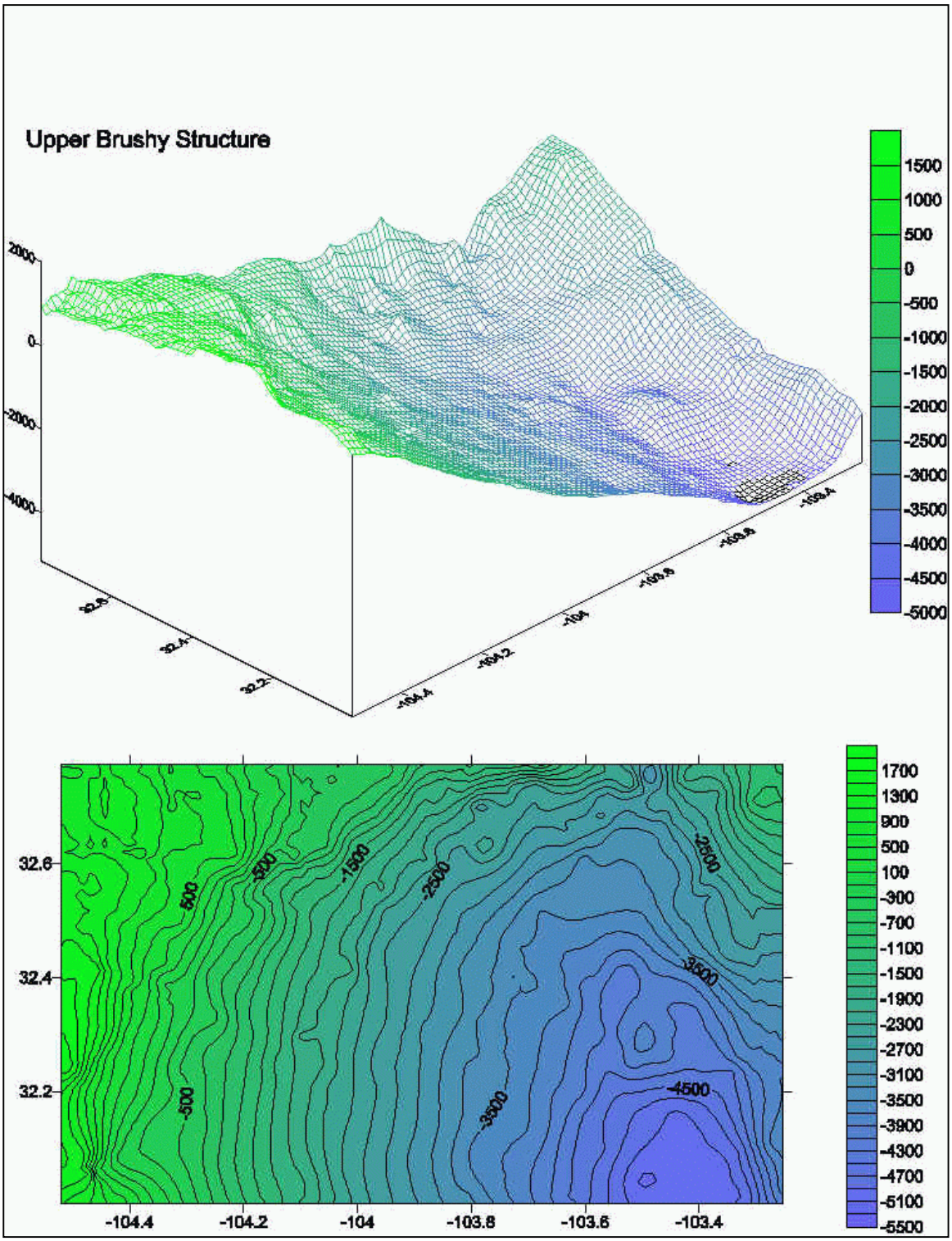

Figure 10. Structure contour map and wireframe relief map showing structure on top of upper Brushy Canyon formation.

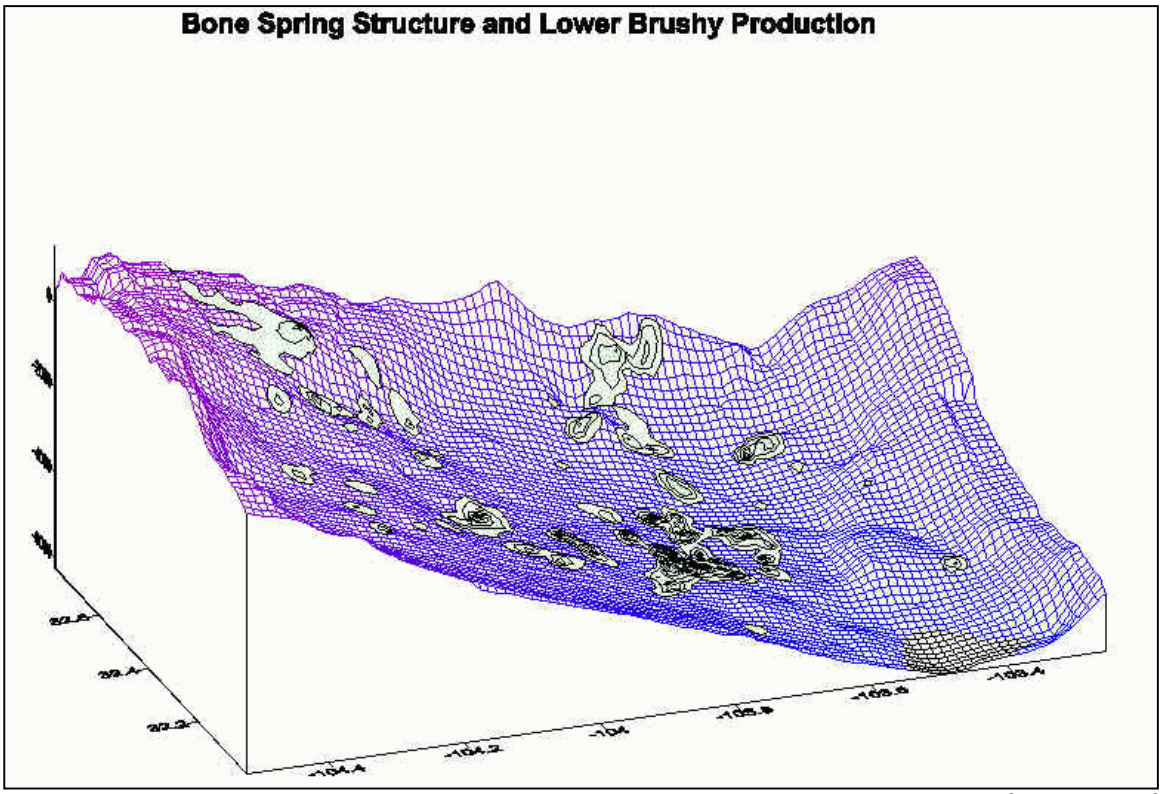

Figure 11. Lower Brushy Canyon productive areas superimposed on wireframe relief map of Bone Spring structure.

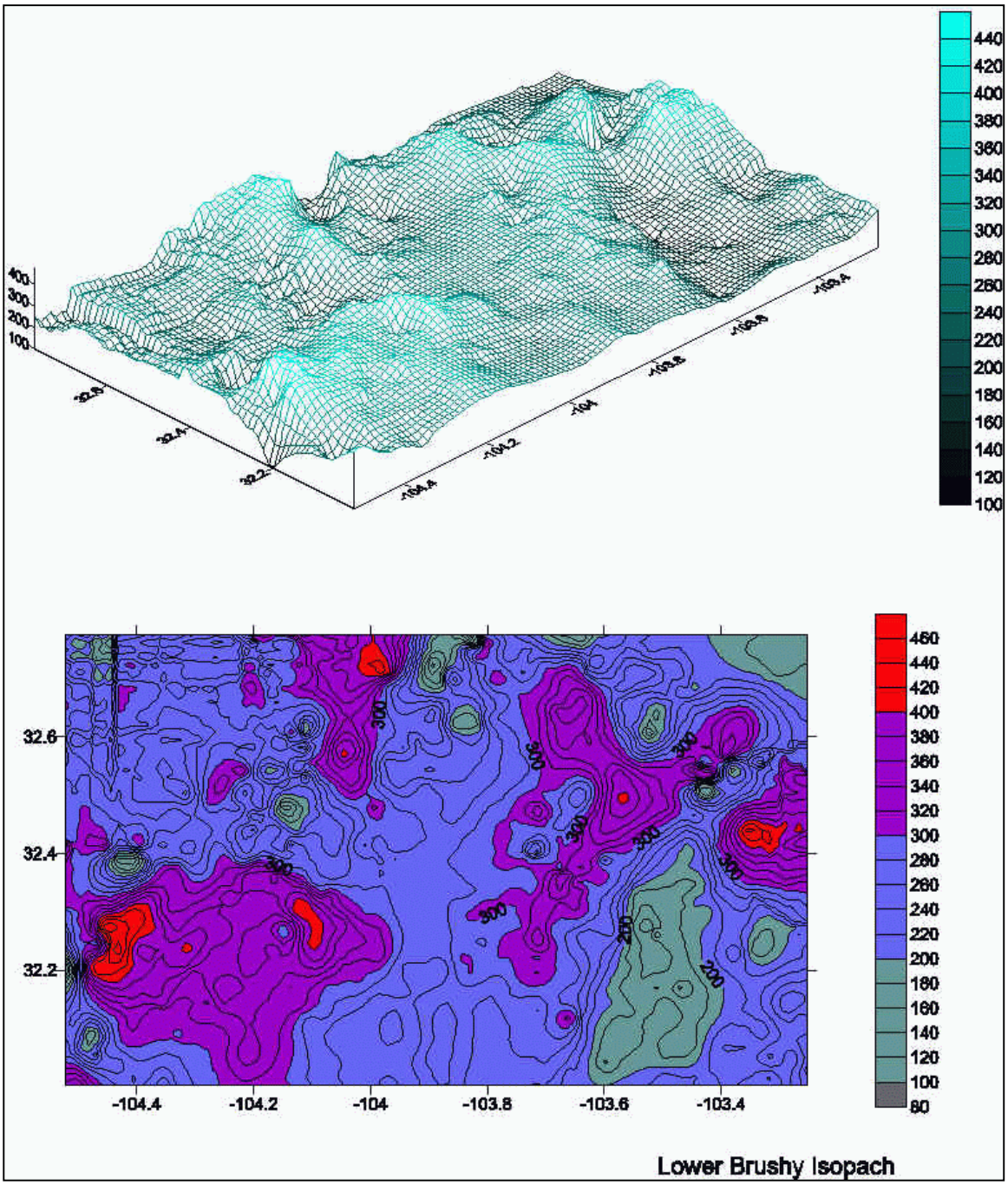

Figure 12. Thickness of lower Brushy Canyon formation, portrayed as a contour map and a wireframe relief map.

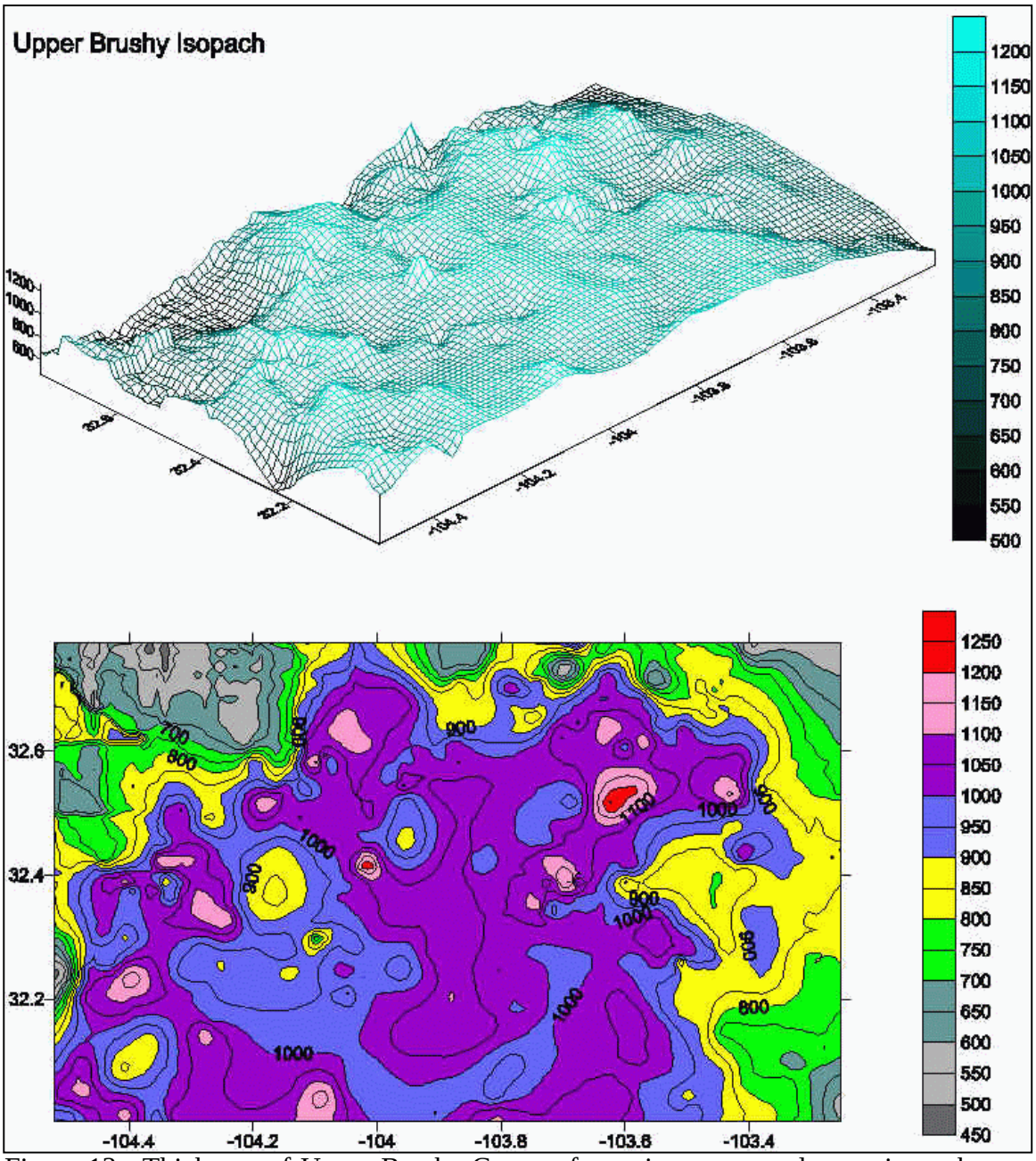

Figure 13. Thickness of Upper Brushy Canyon formation, portrayed as an isopach contour map and as a wireframe relief map.

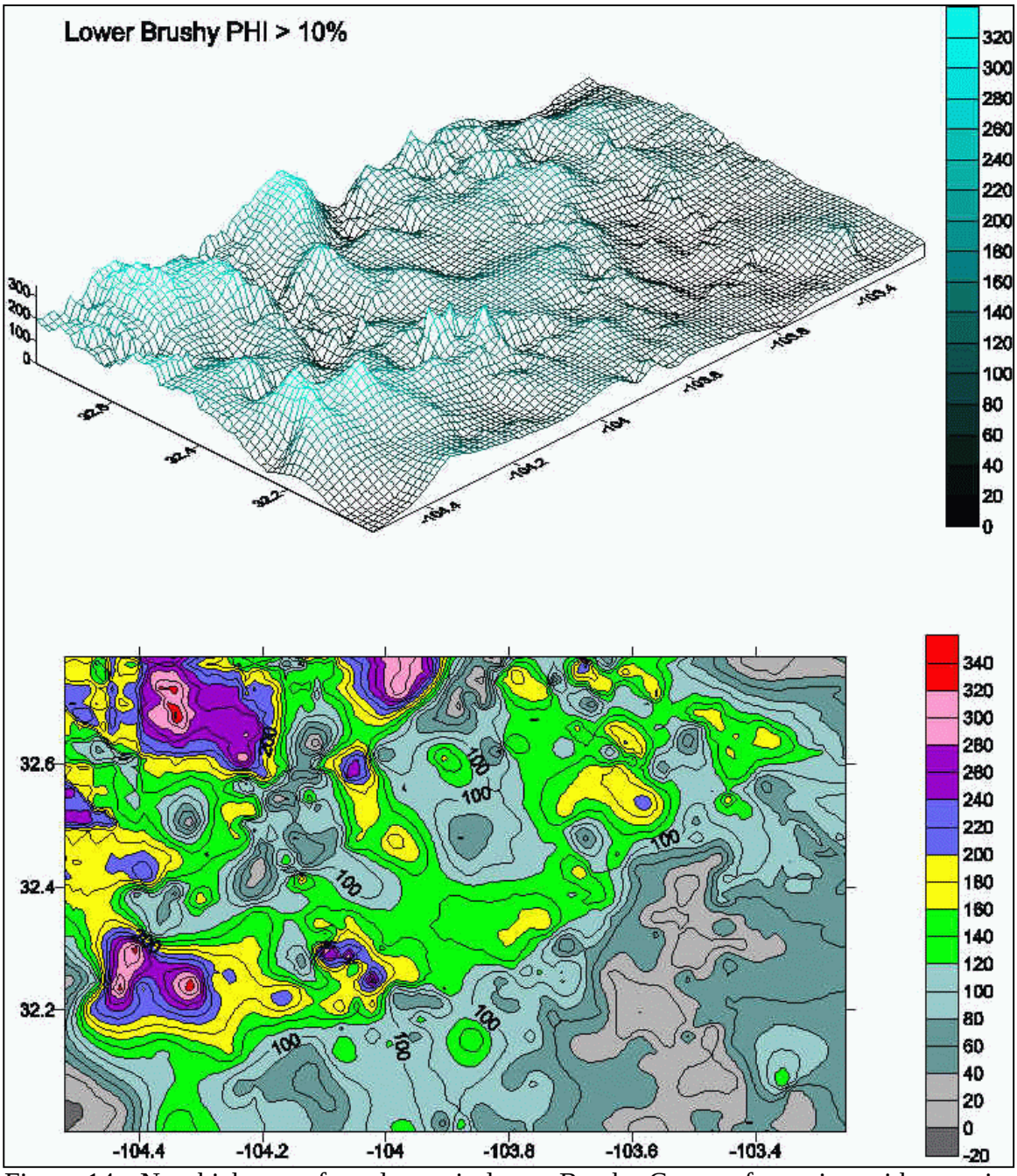

Figure 14. Net thickness of sandstone in lower Brushy Canyon formation with porosity greater than 10%, portrayed as contour map and as a wireframe relief map.

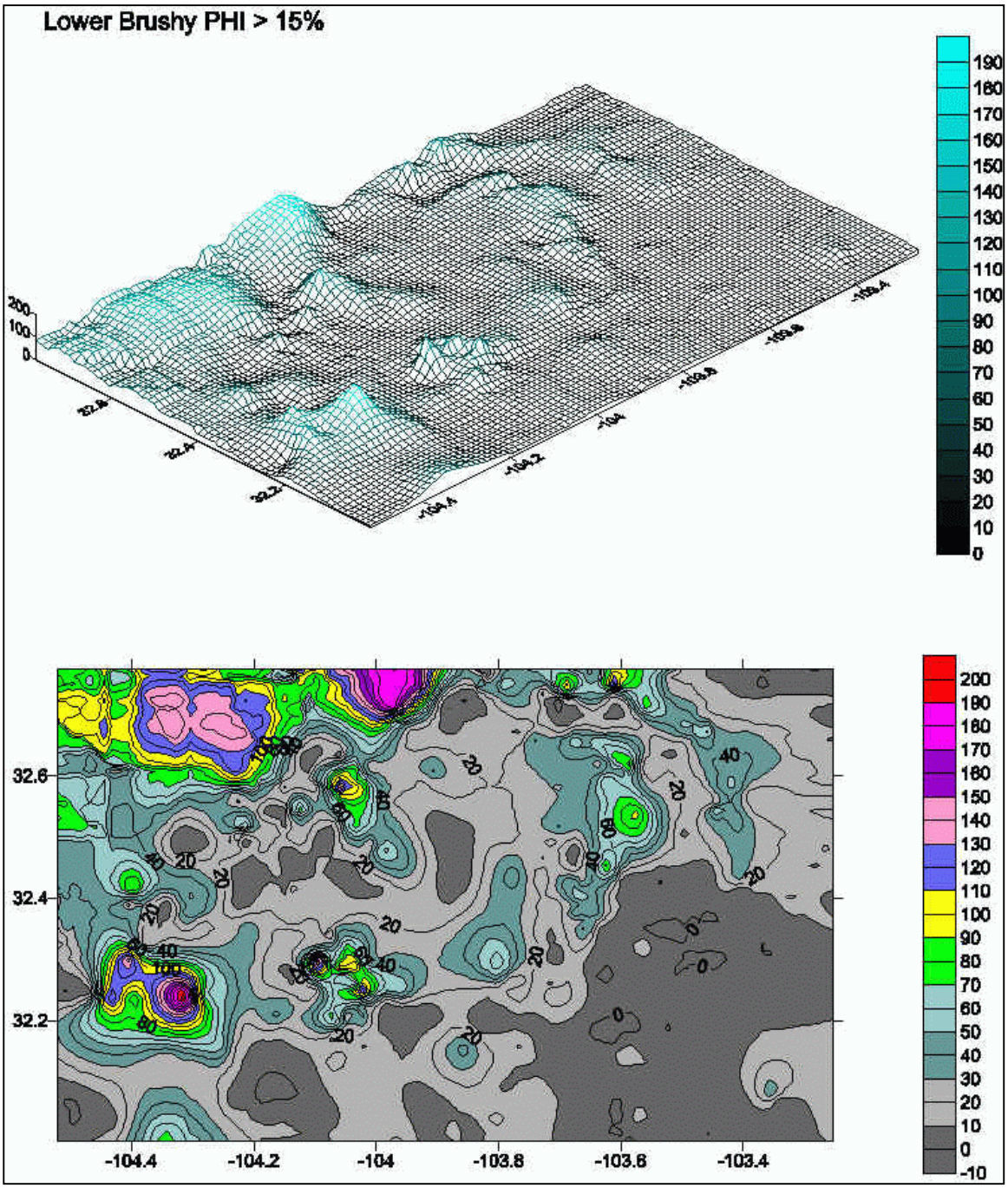

Figure 15. Net thickness of sandstone in lower Brushy Canyon formation with porosity greater than 15%, portrayed as a contour map and as a wireframe relief map.

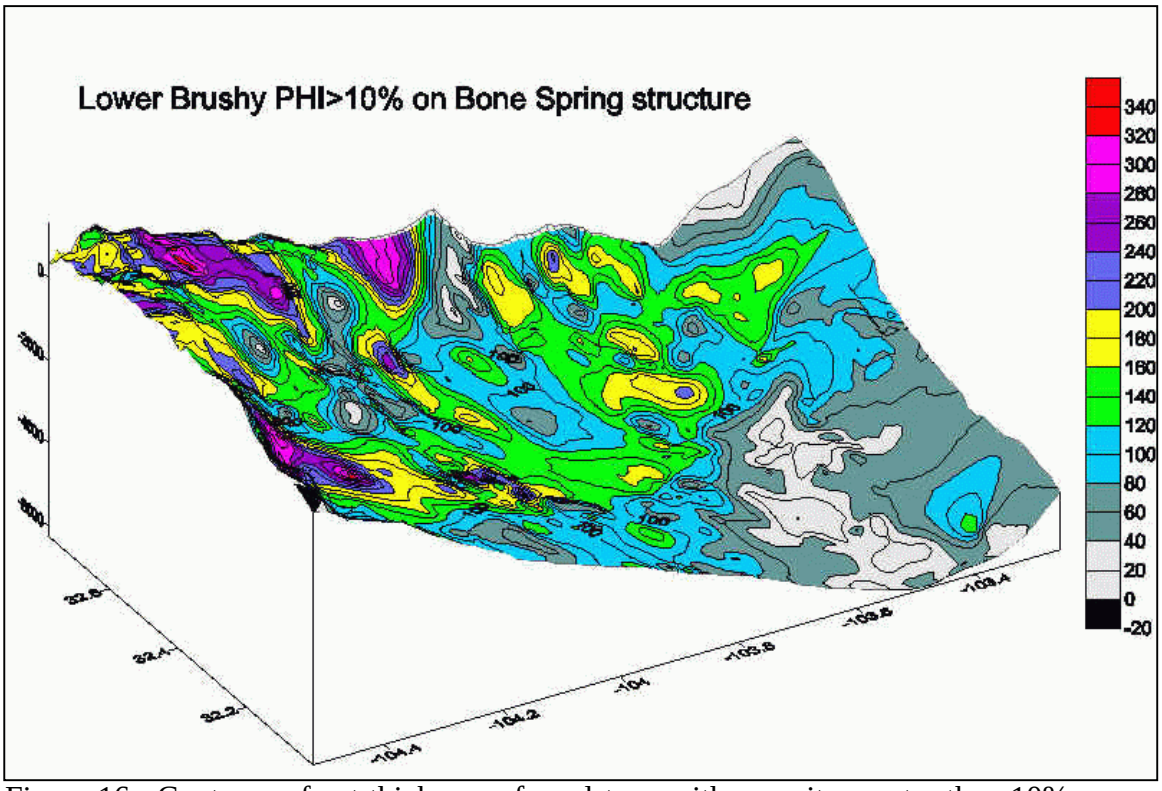

Figure 16. Contours of net thickness of sandstone with porosity greater than 10% superimposed on wireframe relief map showing structure on top of Bone Spring formation.

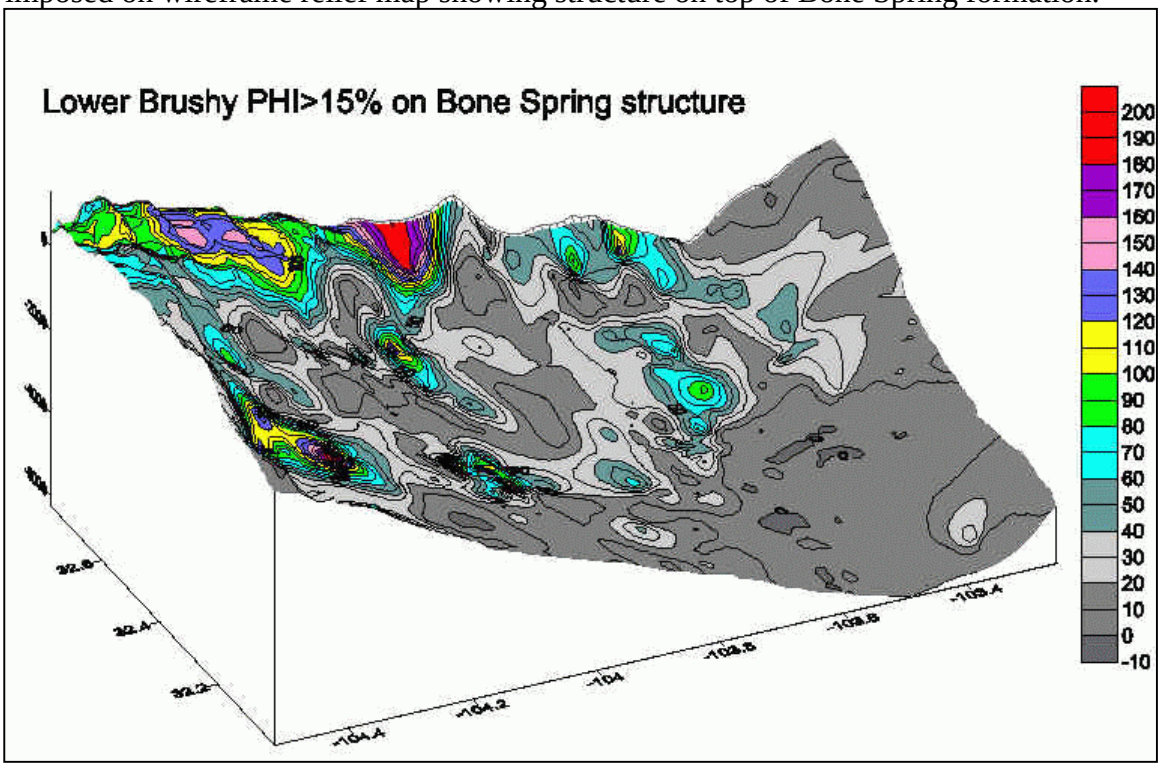

Figure 17. Contours of net thickness of sandstone with porosity greater than 15% superimposed on wireframe relief map showing structure on top of Bone Spring formation.

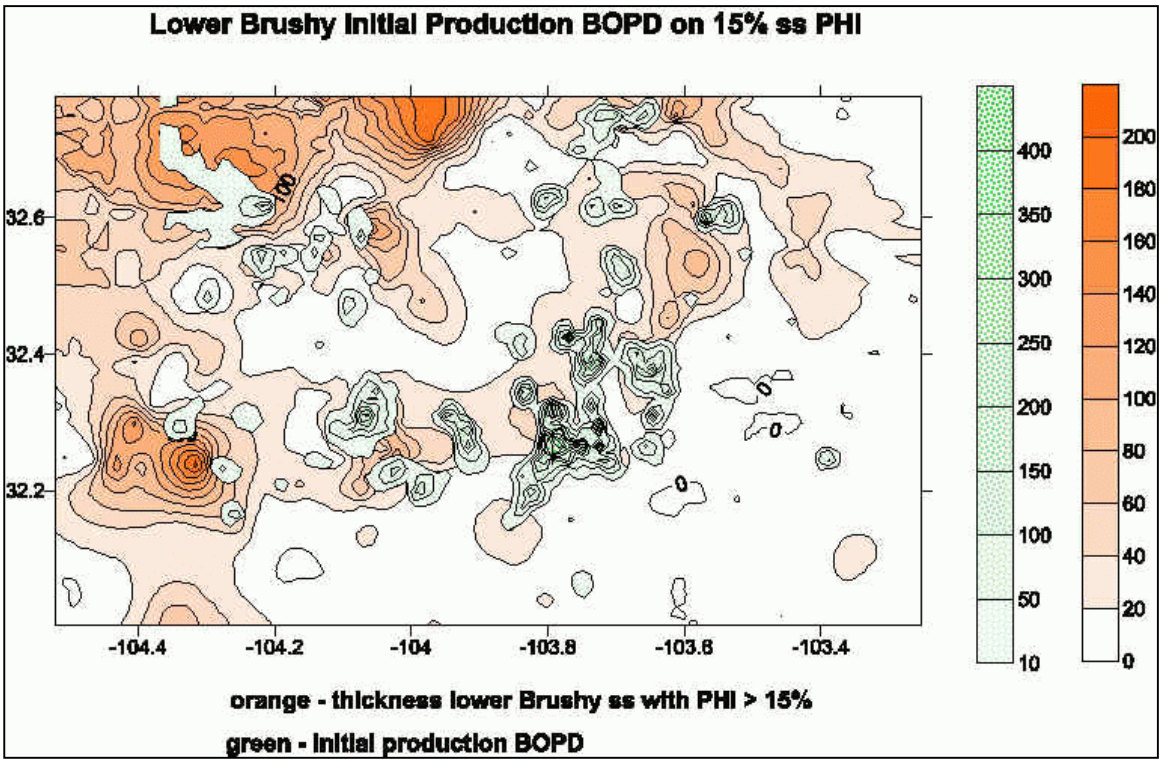

Figure 18. Initial oil production from lower Brushy Canyon formation superimposed on contour map of net thickness of sandstones with porosity greater than 15%.

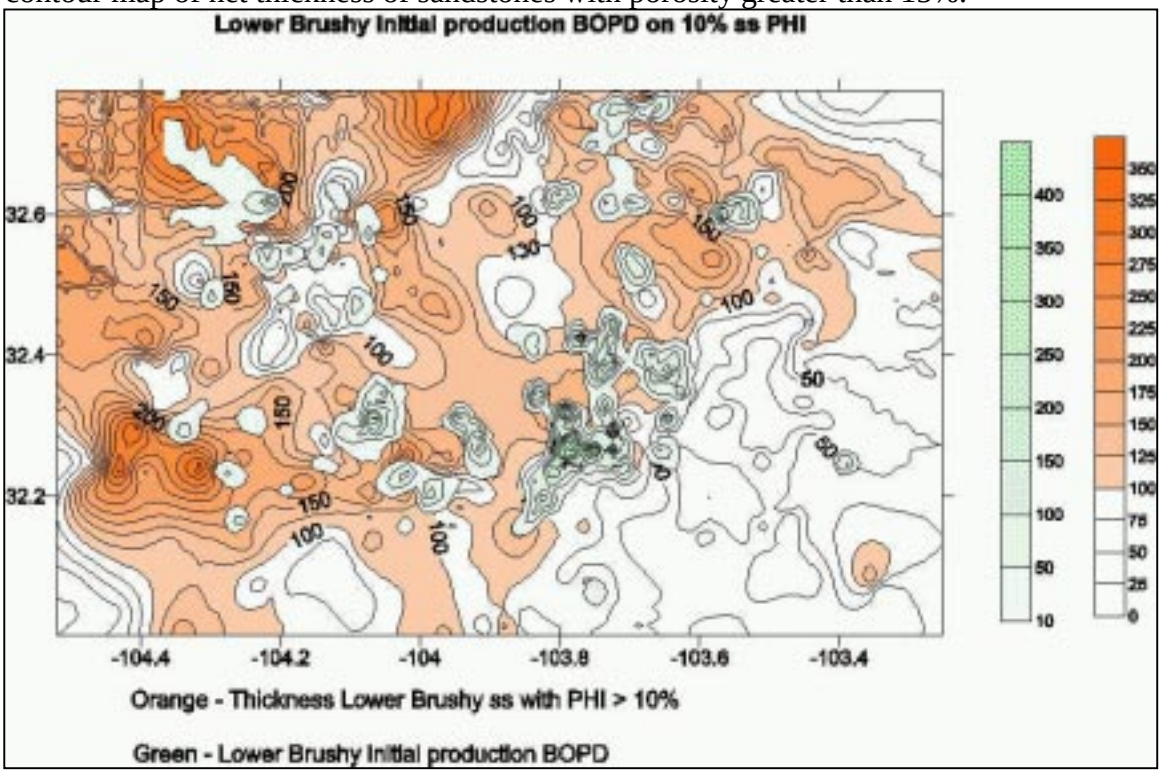

Figure 19. Initial oil production from lower Brushy Canyon formation superimposed on contour map of net thickness of sandstones with porosity greater than 10%.

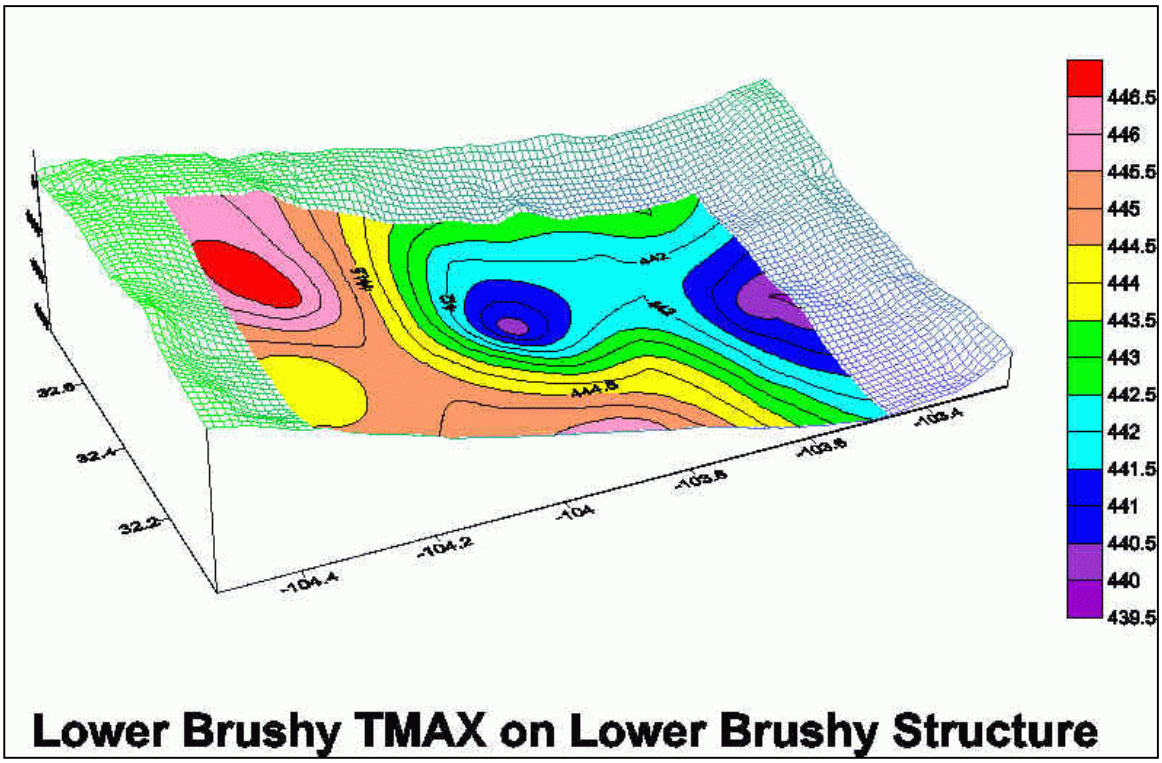

Figure 20. Values of Rock_Eval TMAX, lower brushy Canyon formation, superimposed on wireframe relief map of structure on top of lower Brushy Canyon formation.

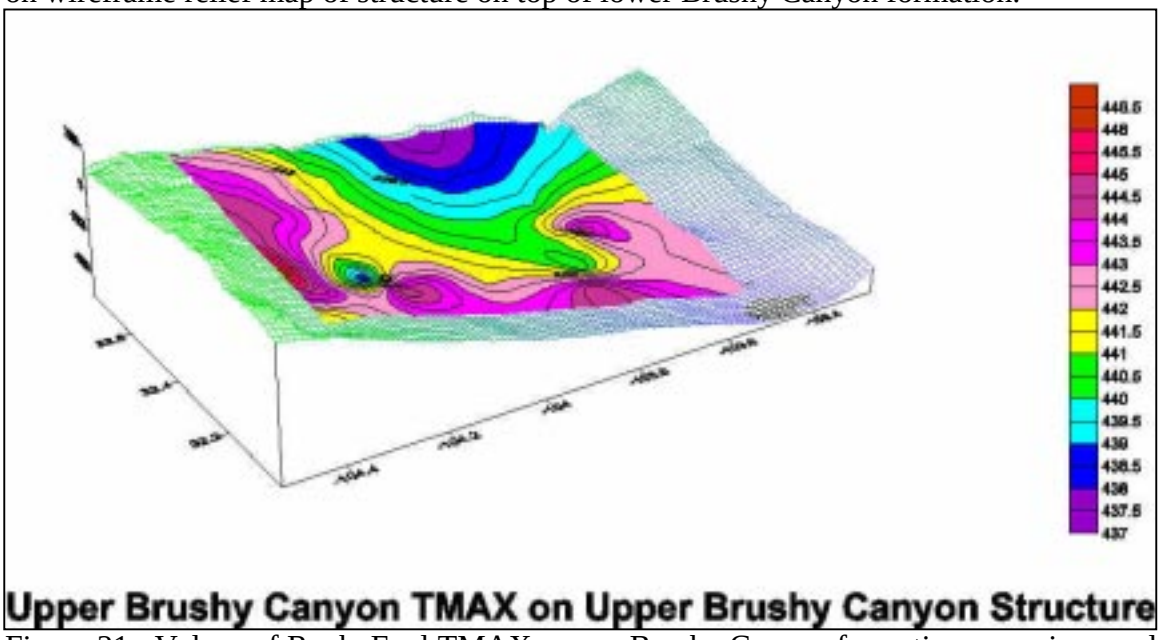

Figure 21. Values of Rock_Eval TMAX, upper Brushy Canyon formation, superimposed on wireframe relief map of structure on top of upper Brushy Canyon formation.

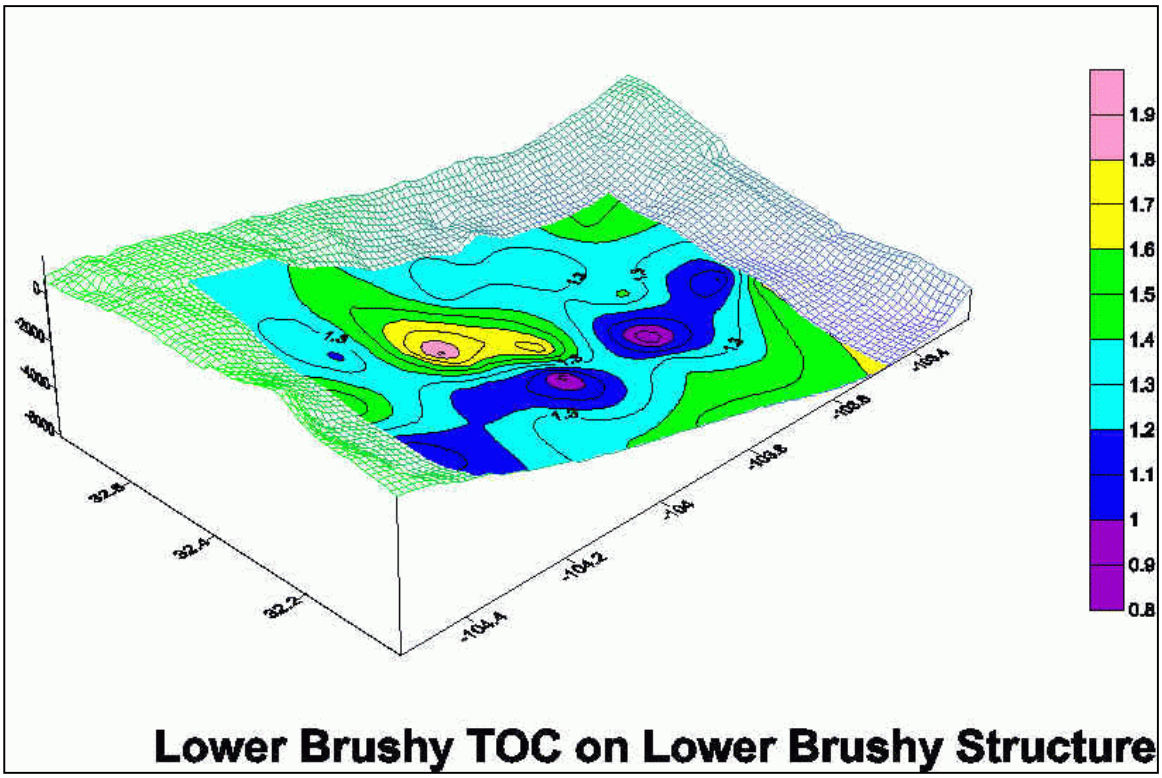

Figure 22. Total organic carbon contours of Lower Brushy Canyon formation superimposed on wireframe relief map of structure on top of lower Brushy Canyon.

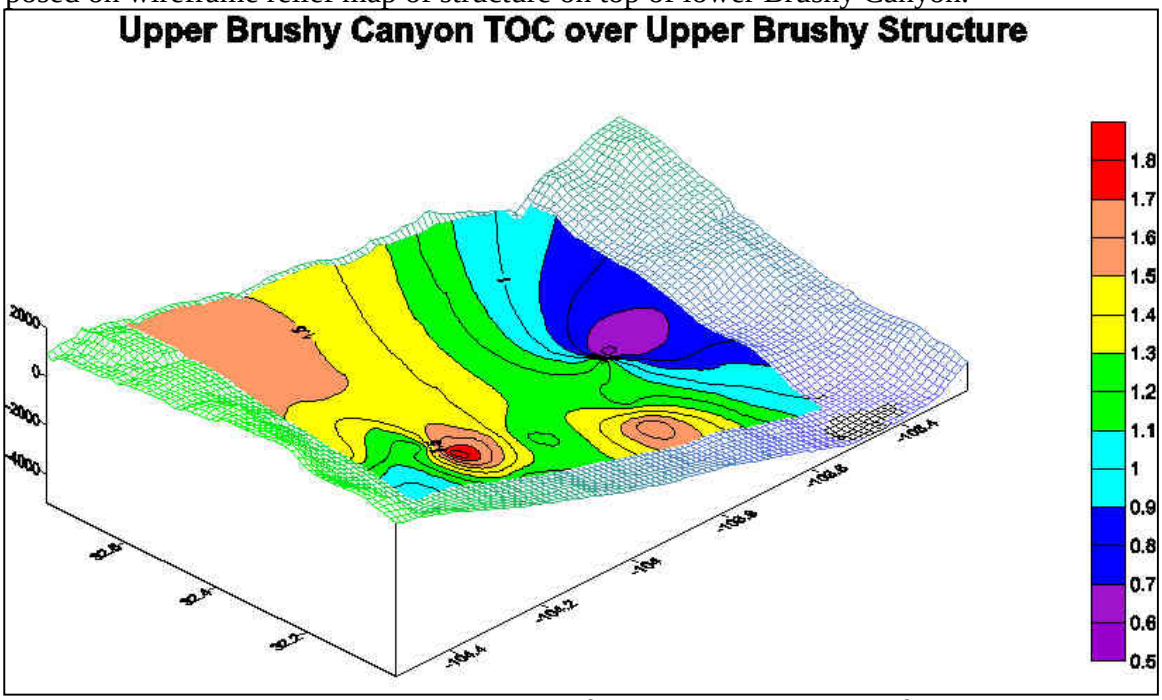

Figure 23. Total organic carbon contours of upper Brushy Canyon formation, superimposed on wireframe relief map of structure on top of upper Brushy Canyon formation.

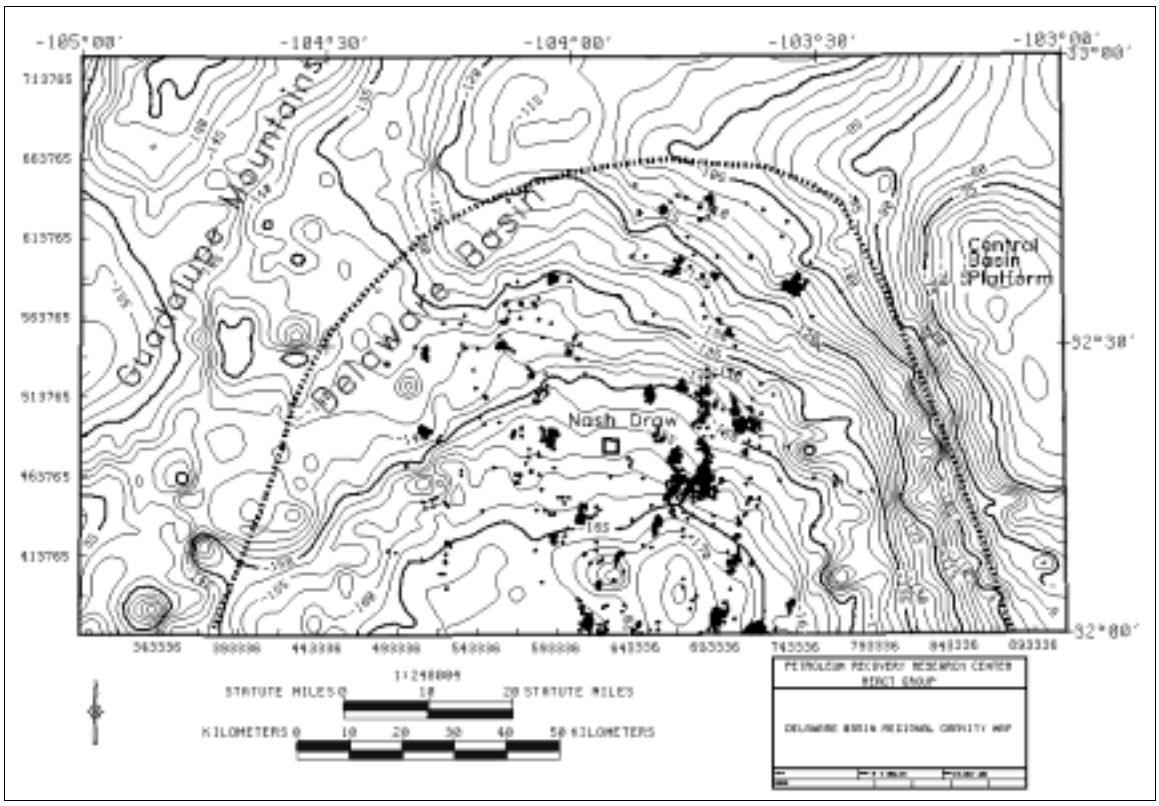

Figure 24. Residual gravity anomaly for the Delaware basin and surrounding regions and Delaware producing wells (dots). The gravity signature is quite strong.

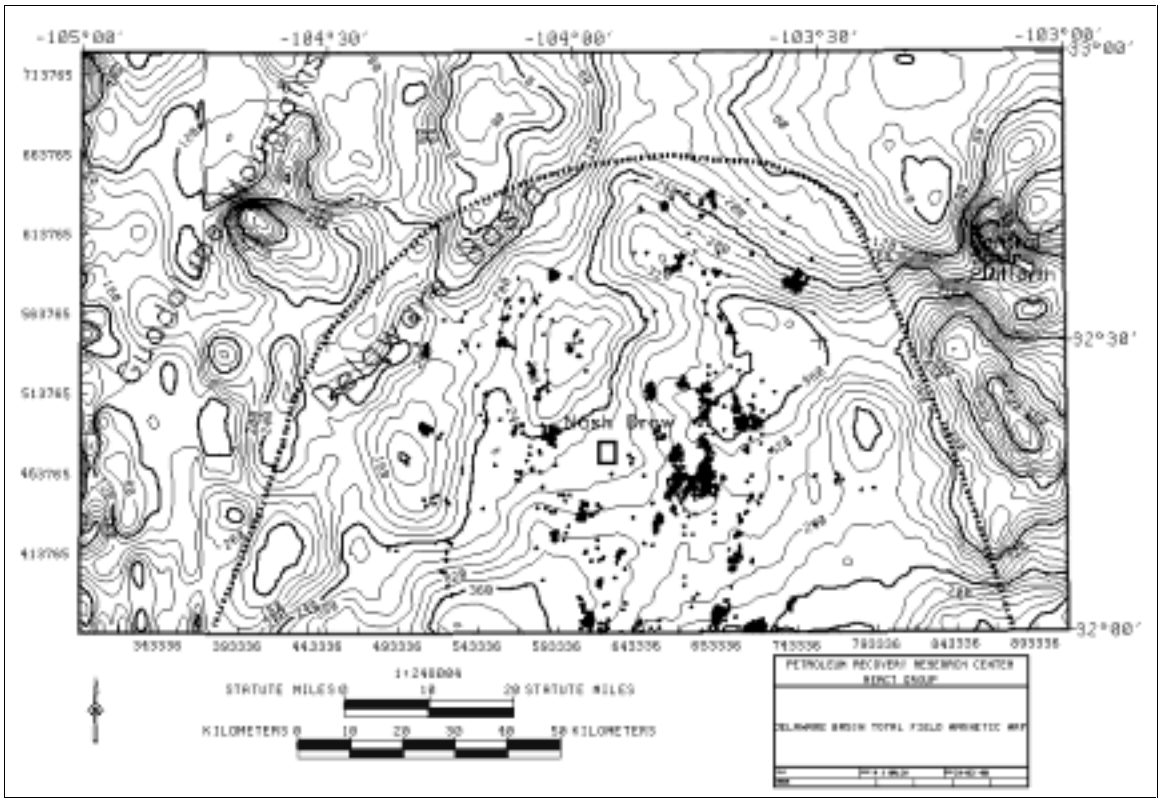

Figure 25. Total field magnetic map for the Delaware Basin and surrounding regions. The basin can be seen encompassed by low magnetic values. Delaware producing wells (dots) are concentrated along the flanks of magnetic highs.

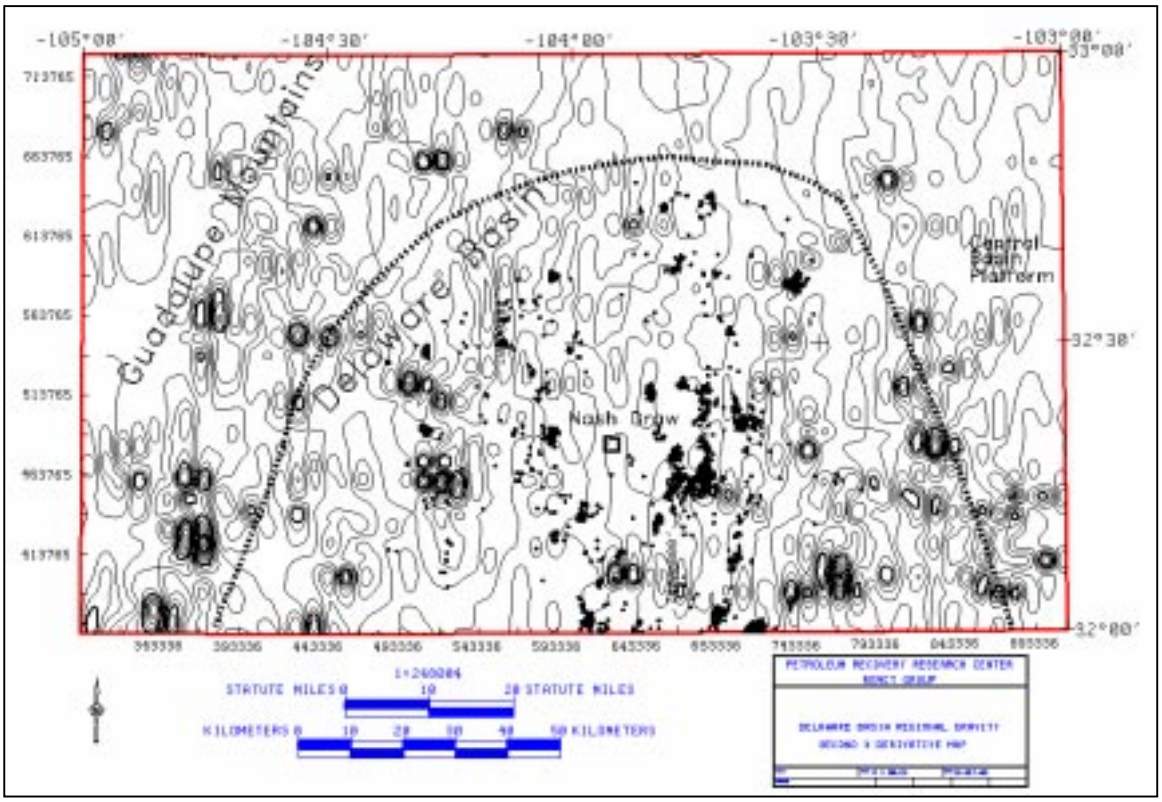

Figure 26. Second derivative gravity map (x direction). The second derivative map acts as a high pass filter and tends to remove larger scale features such as those caused by basement features.

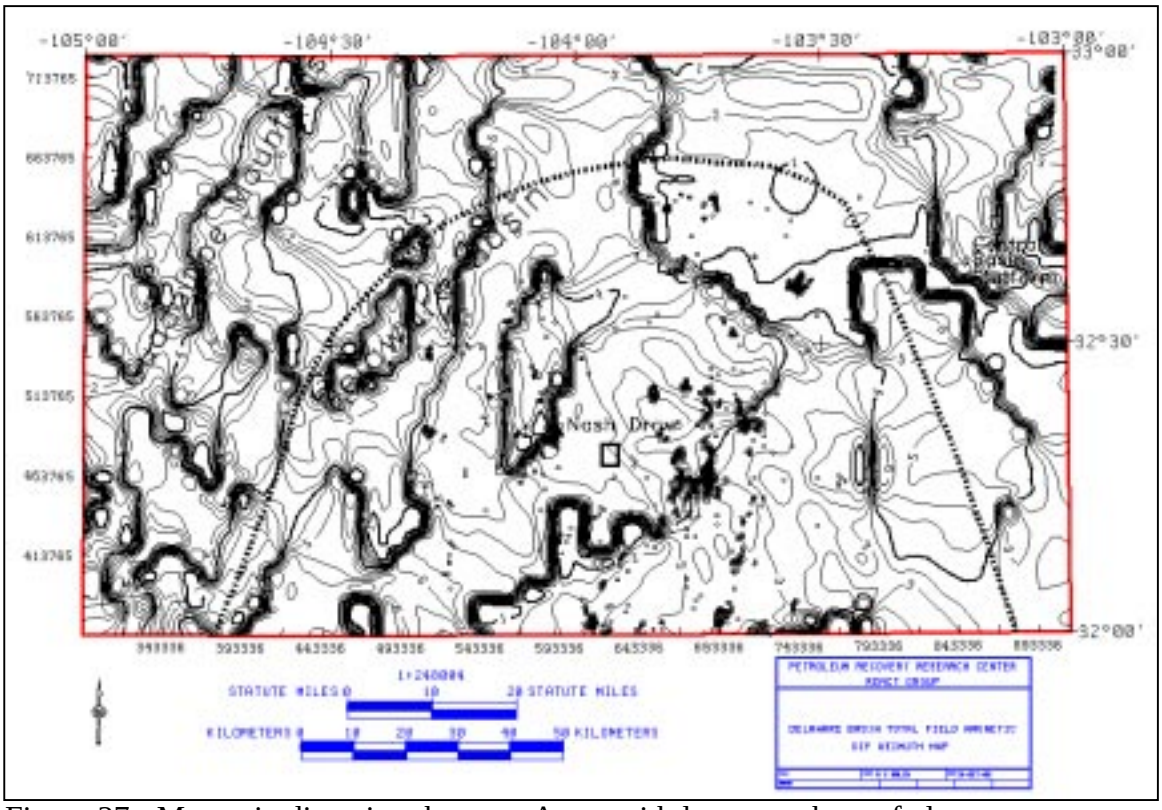

Figure 27. Magnetic dip azimuth map. Areas with large numbers of close contours represent areas of large-scale magnetic susceptibility, which can really only occur in basement blocks or intrusions.

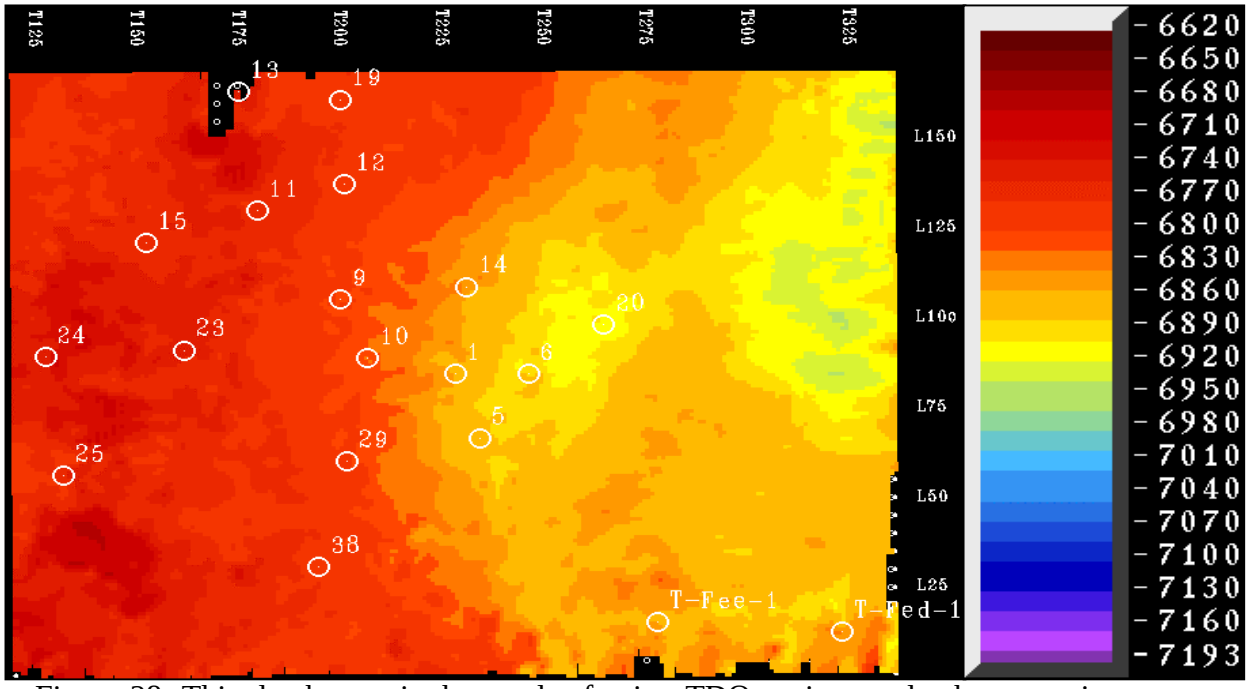

Figure 28. This depth map is the result of using TDQ, a time-to-depth conversion program, with a simple one-layer velocity model based on well control points. This program honors the data points exactly, and uses a weighted linear interpolation to convert the time horizon to depth. Test points are poorly predicted (T-Fee-1 and T-Fed-1). Scale is in feet.

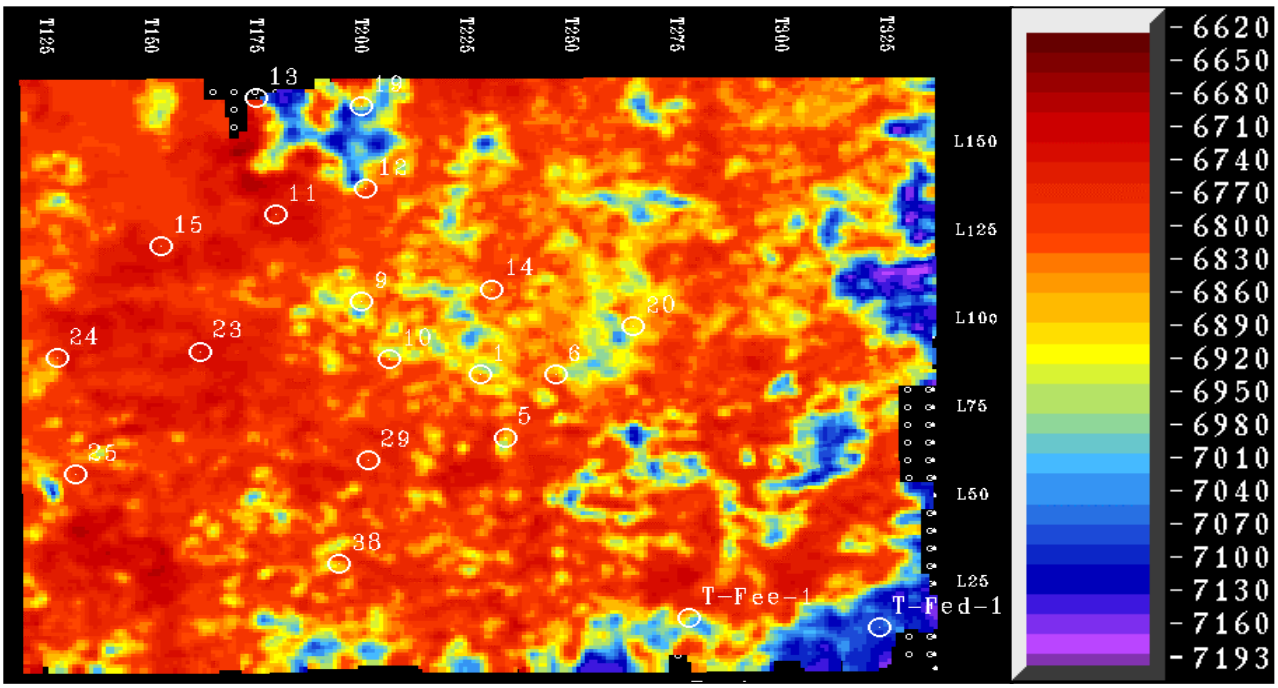

Figure 29. This is the depth map predicted using the MLP with a network architecture of 3-3-3-2-1, and training to CC=0.9895. This map was filtered lightly to highlight trends. The filter was based on 3x3-bin size equally weighted low pass filter. Test wells T-Fee-1 and T-Fed-1 were accurately predicted. Scale in feet.

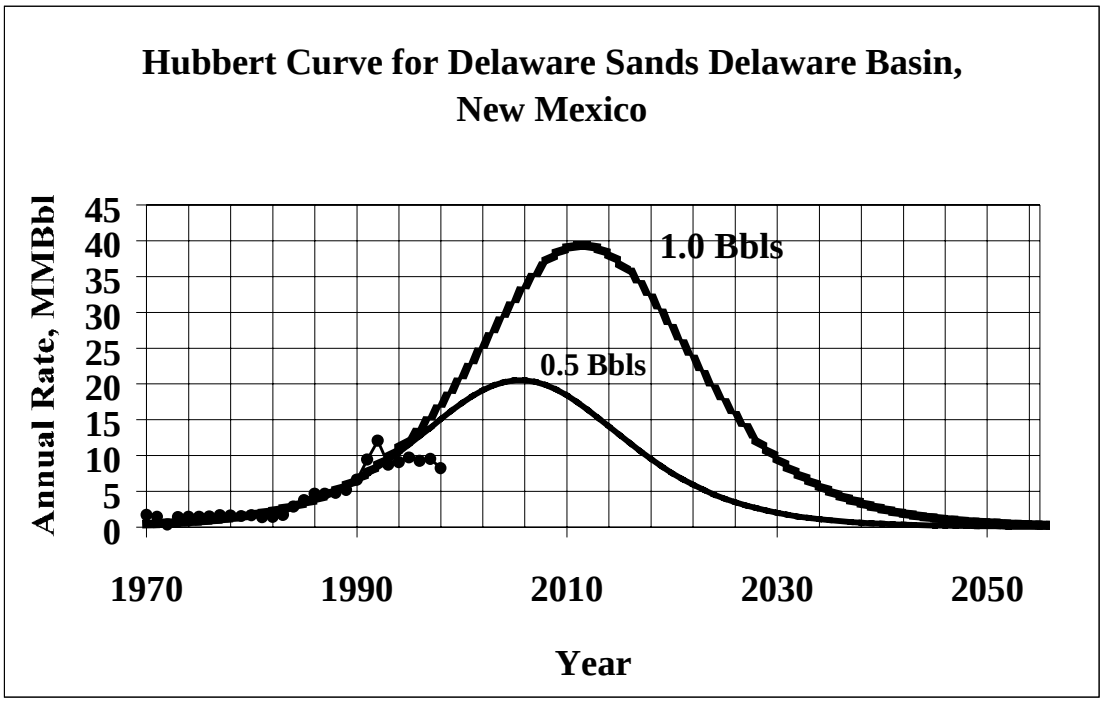

Figure 30. Curves constructed with the derivative of the logistic function.

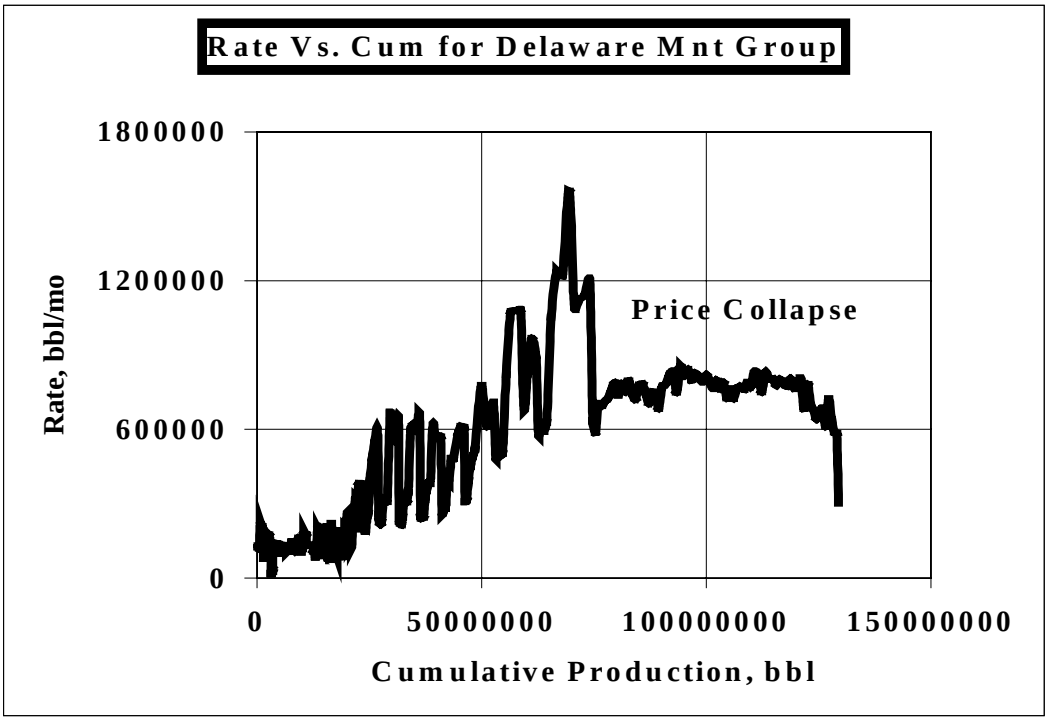

Figure 31. Price collapses of 1985 and 1998.

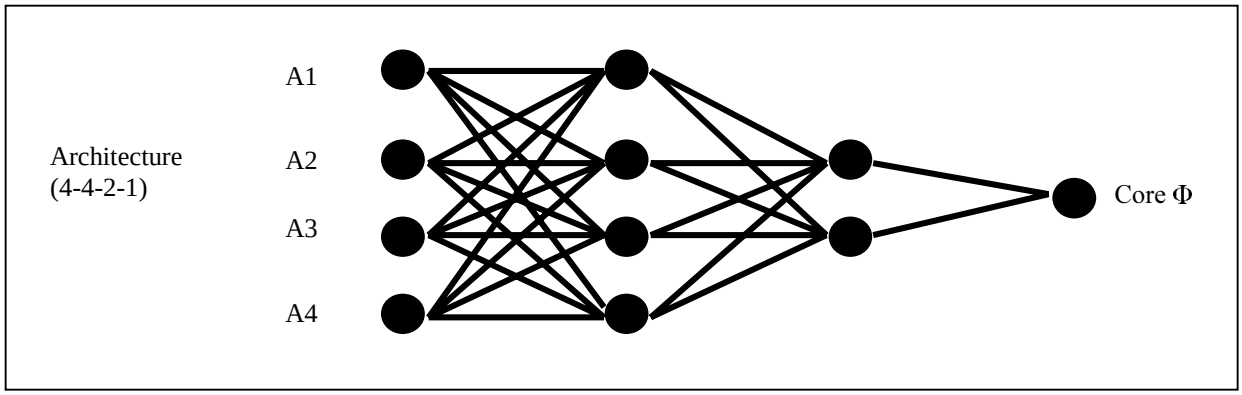

Figure 32. A 4-4-2-1 neural network architecture. This architecture is four-layered with one input layer, two hidden layers and one output layer.

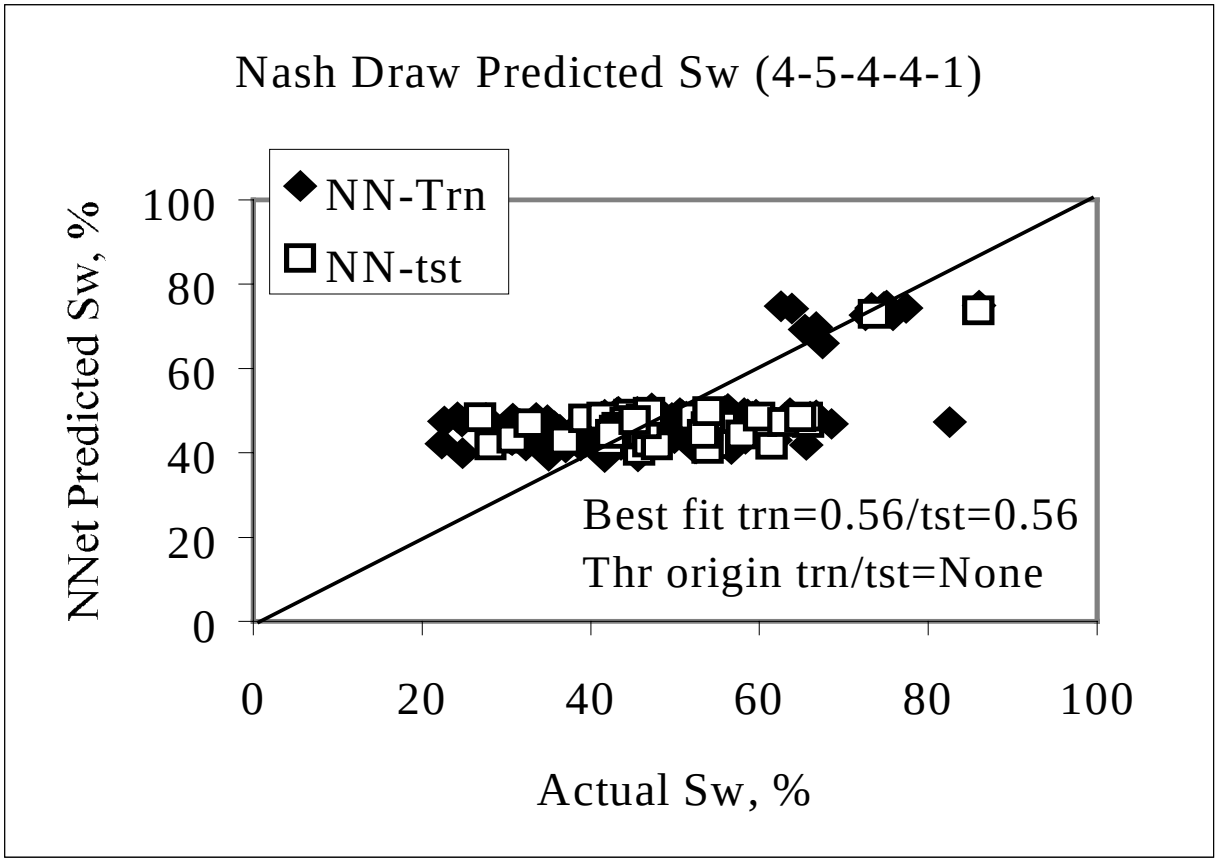

Figure 33. No visual correlation.

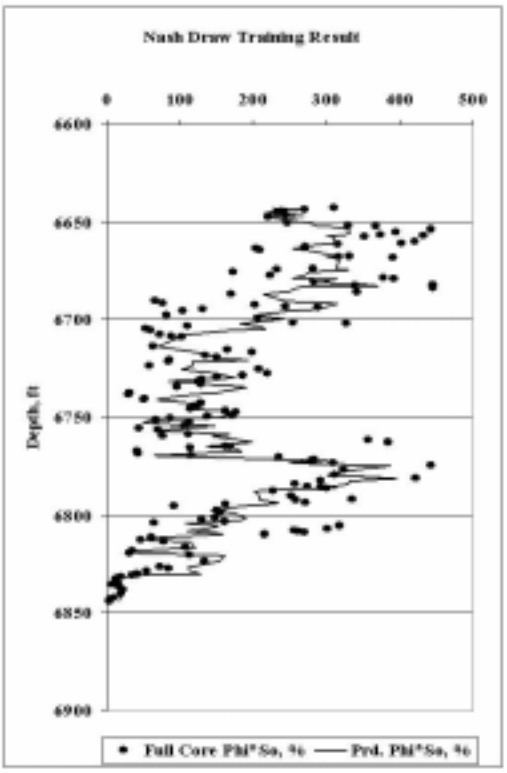

Figure 34. Nash Draw #23 training result using NPHI, DPHI, Log LLD and Log LLS.

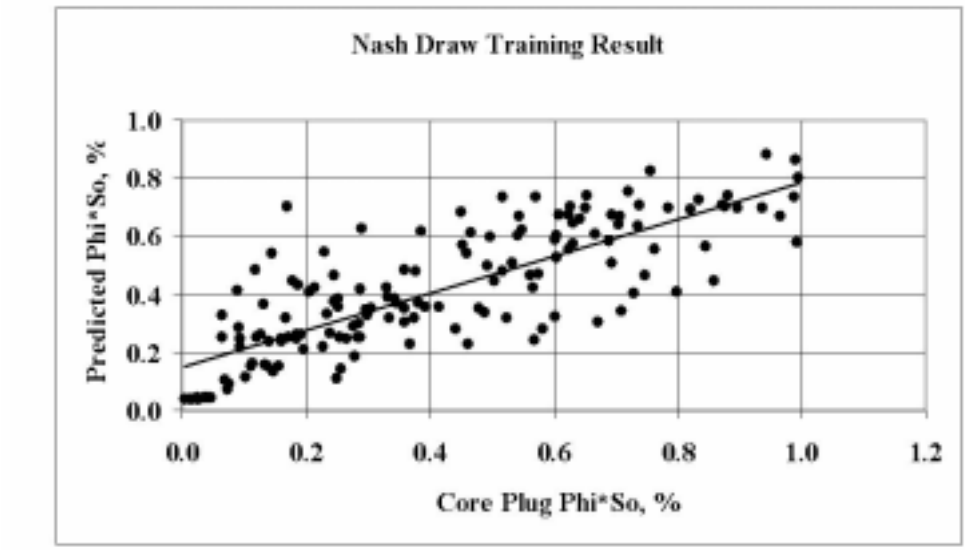

Figure 35. Core plug Φ^*S_o vs. predicted Φ^*S_o (training result).

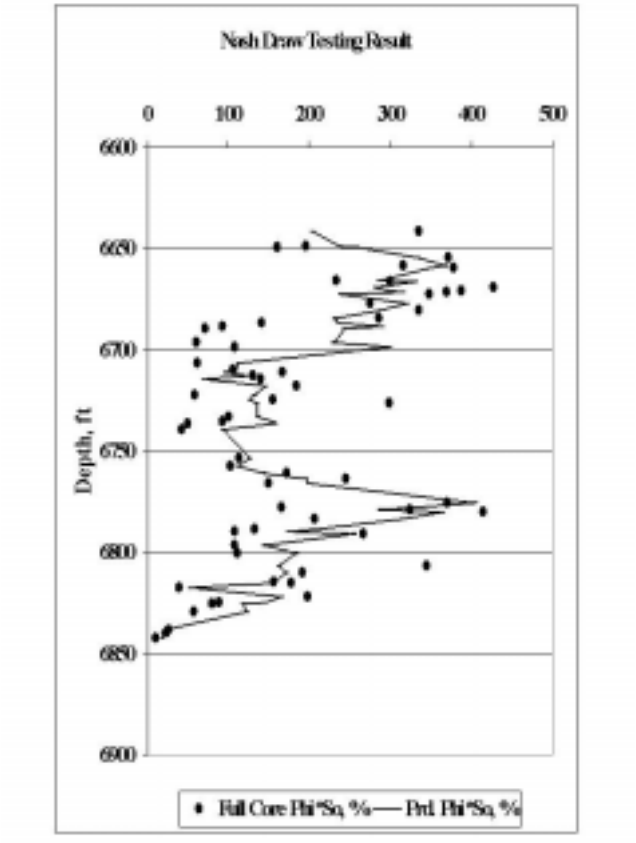

Figure 36. Nash Draw #23 testing result using NPHI, DPHI, Log LLD and Log LLS.

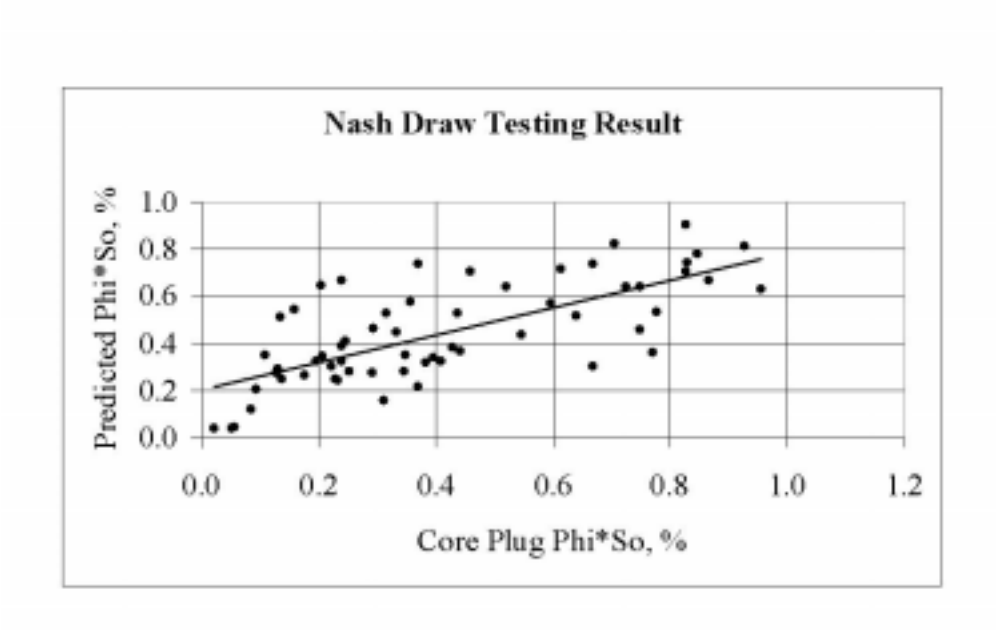

Figure 37. Core plug $\Phi*S_o$ vs. predicted $\Phi*S_o$ (testing result).

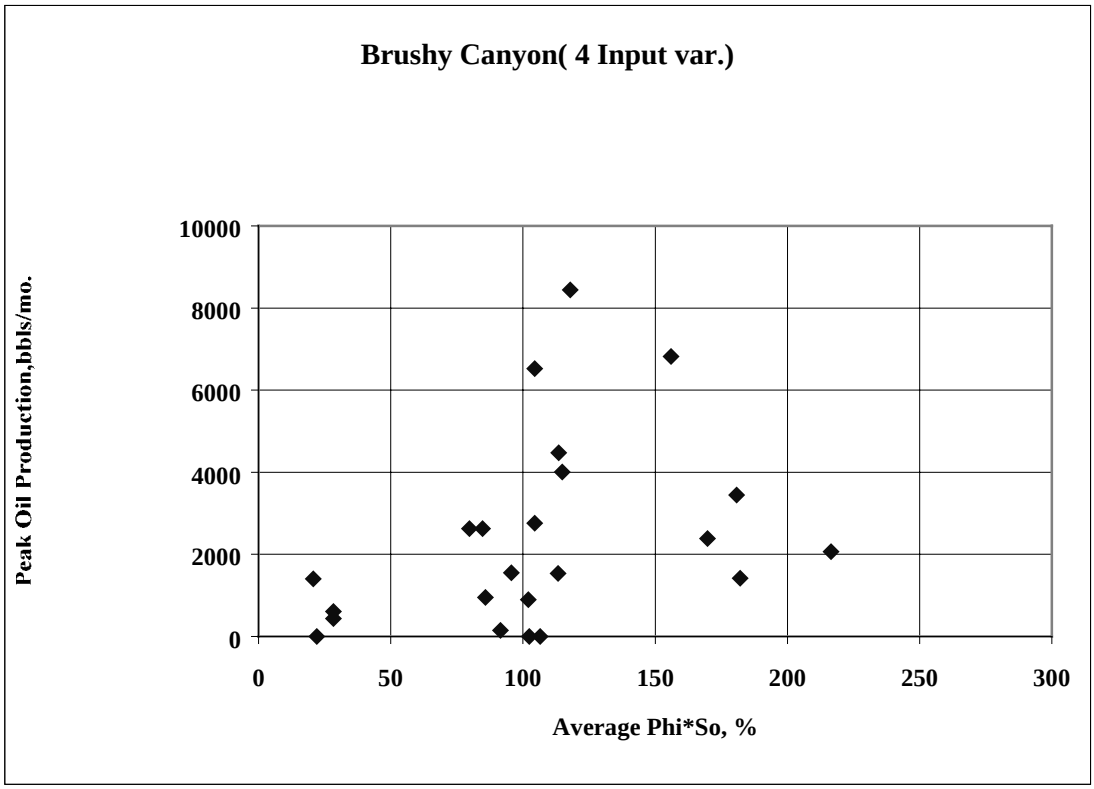

Figure 38. Cum. Φ*So vs. peak oil production.

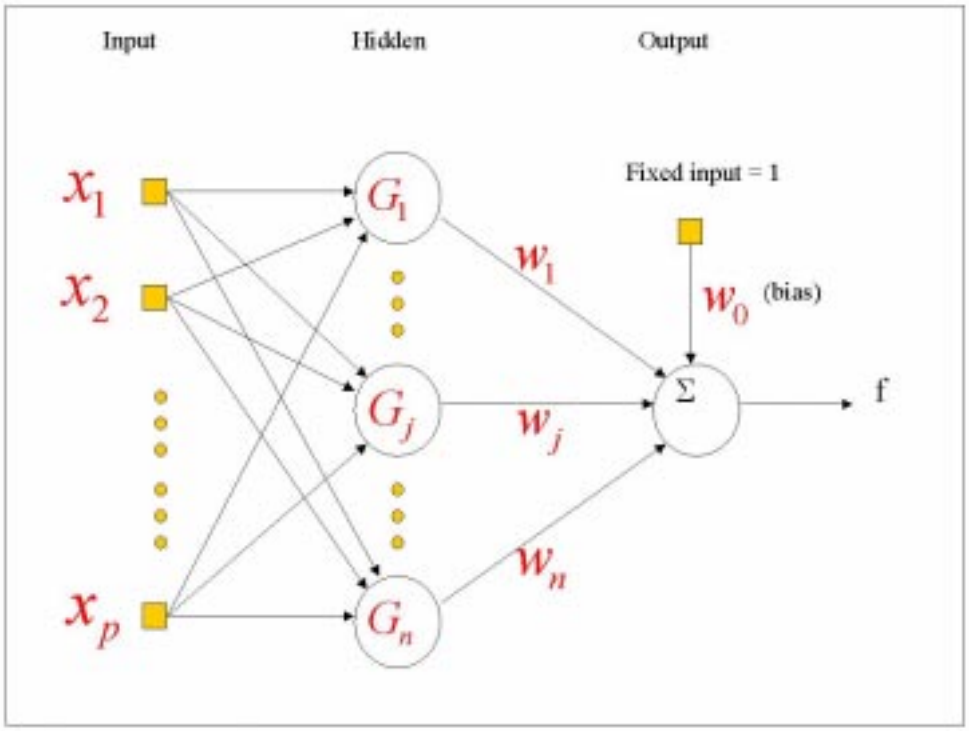

Figure 39. RBF neural network architecture.

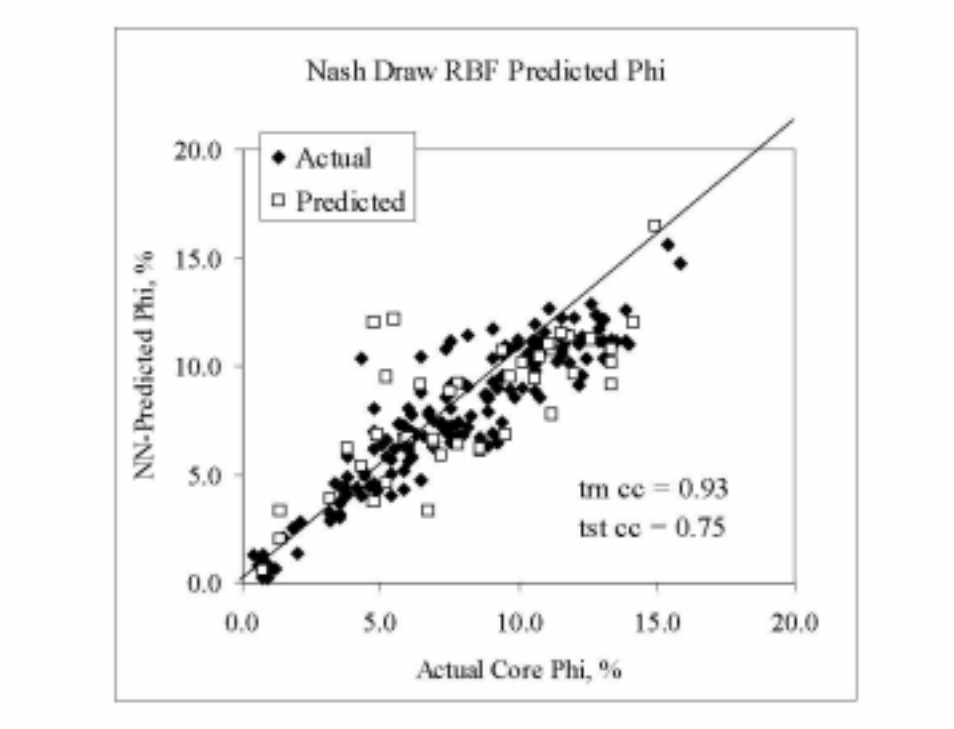

Figure 40. RBF Correlation coefficient of core porosity with density plus deep and shallow resisitivity logs is 93 %.

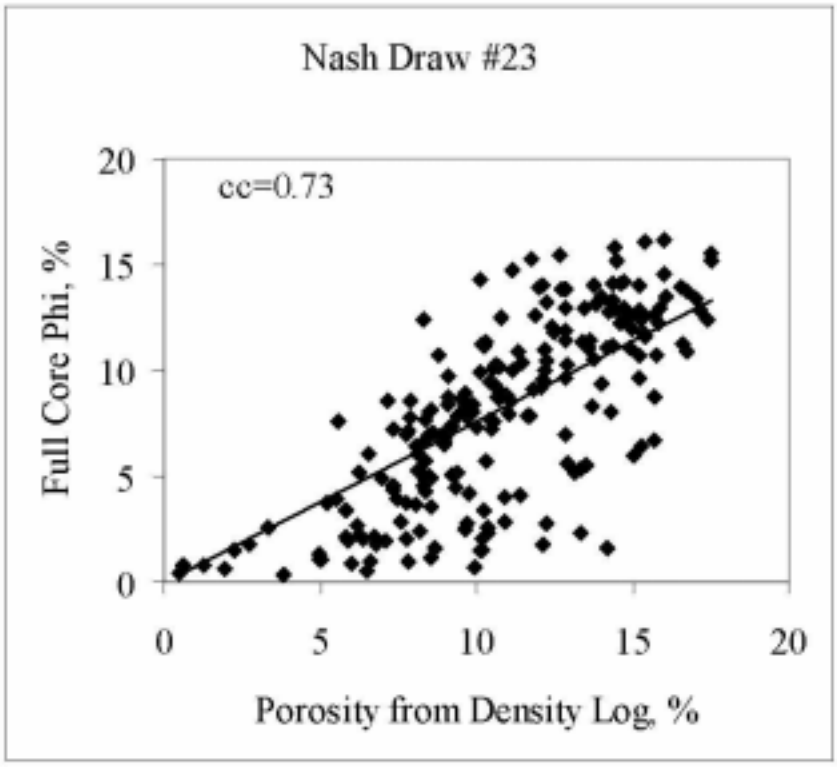

Figure 41. Correlation coefficient of core porosity with density log porosity is 73%.

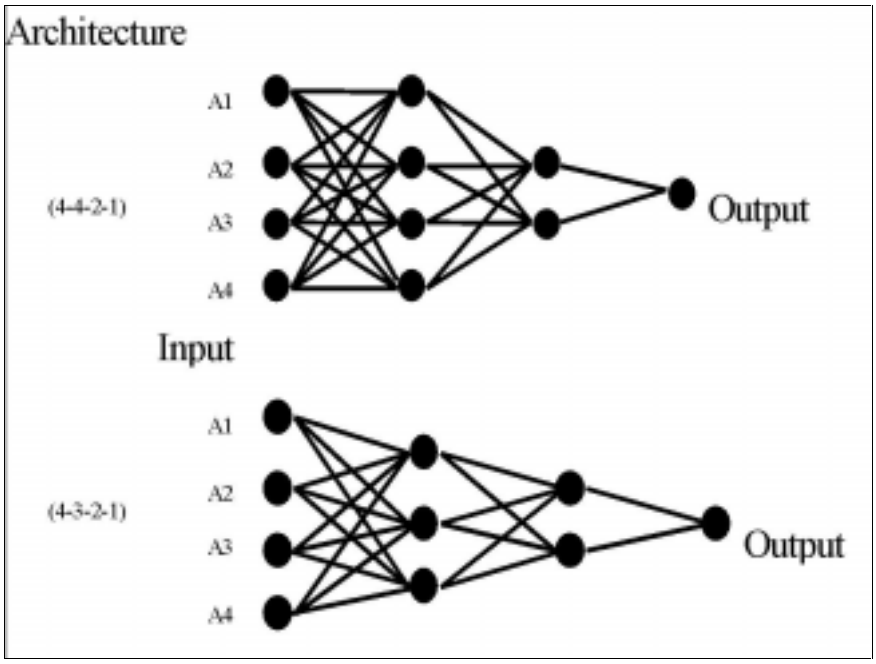

Figure 42 . Example of neural network architectures.

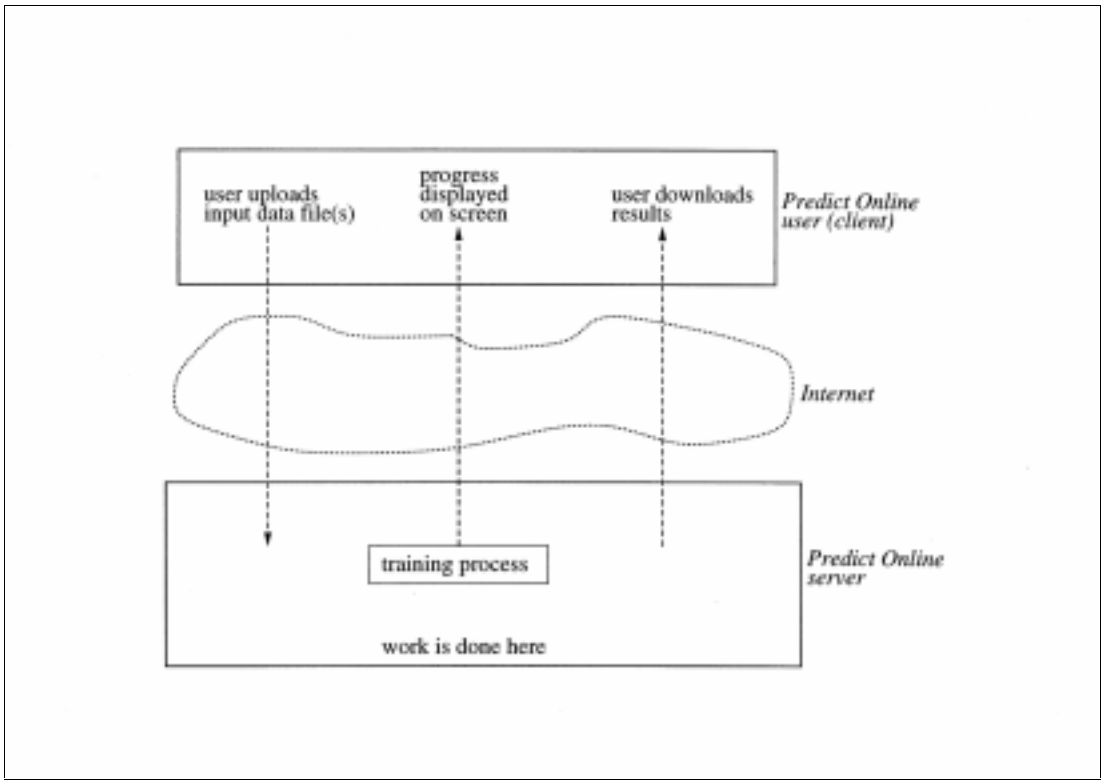

Figure 43. General architecture of *PredictOnline*.

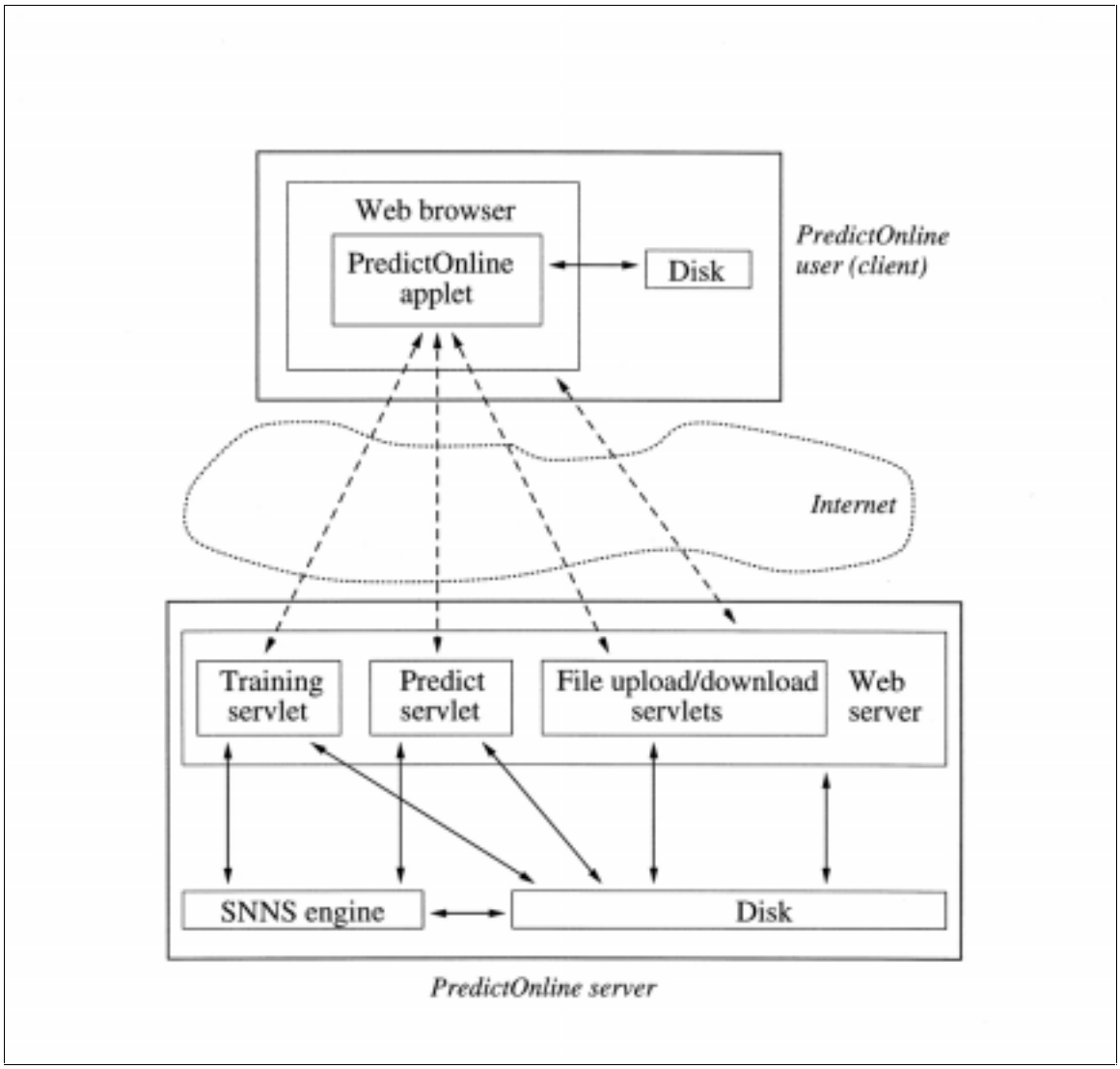

Figure 44. Detailed architecture of *PredictOnline*.

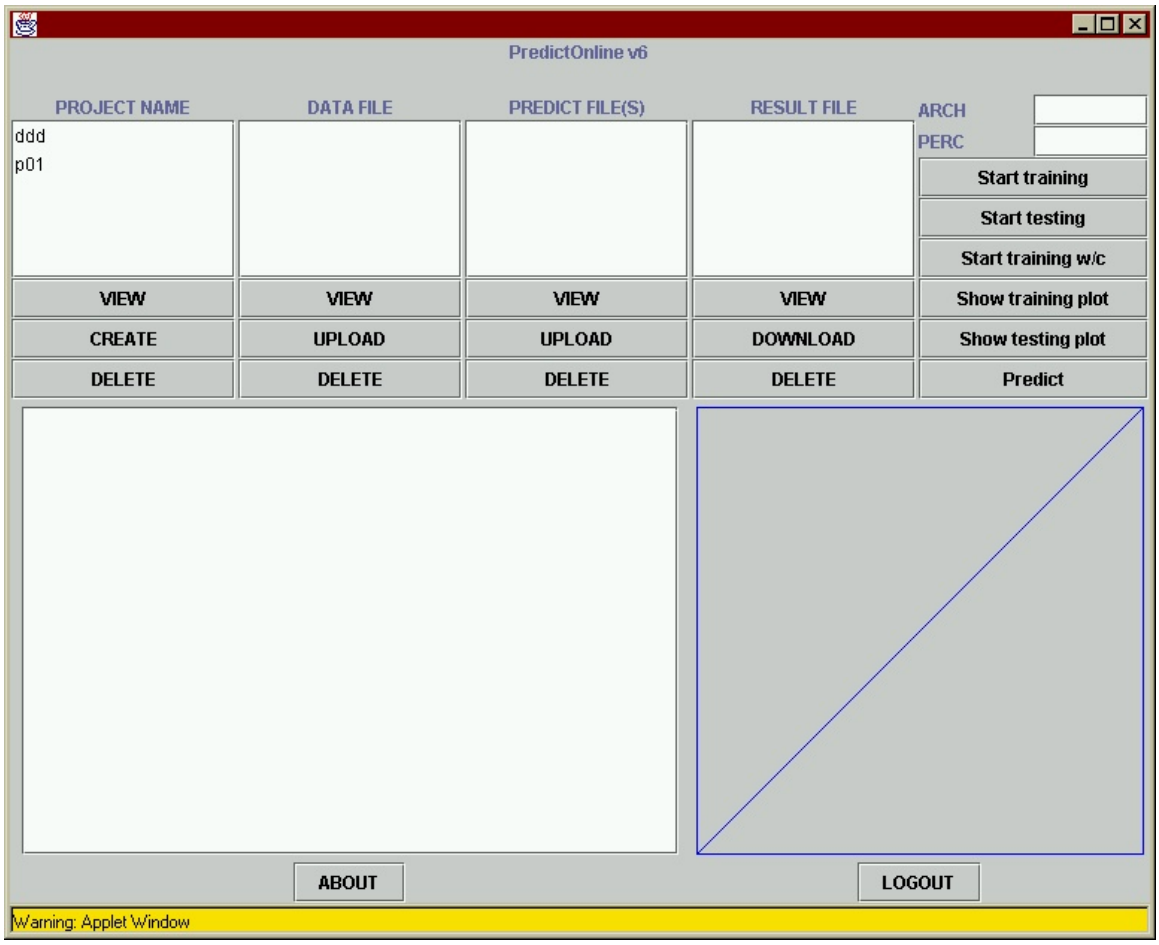

Figure 45. *PredictOnline* main window.

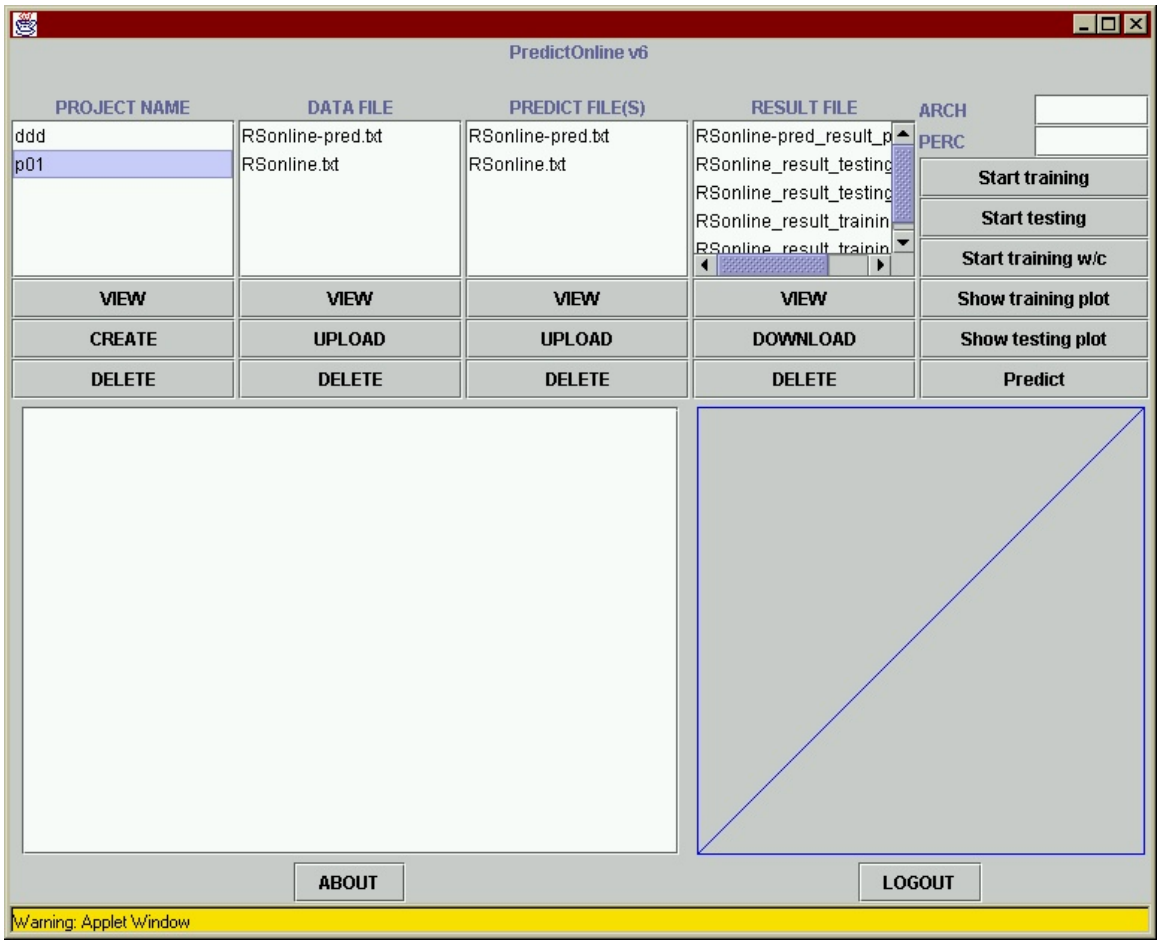

Figure 46. *PredictOnline* main window—project selected.

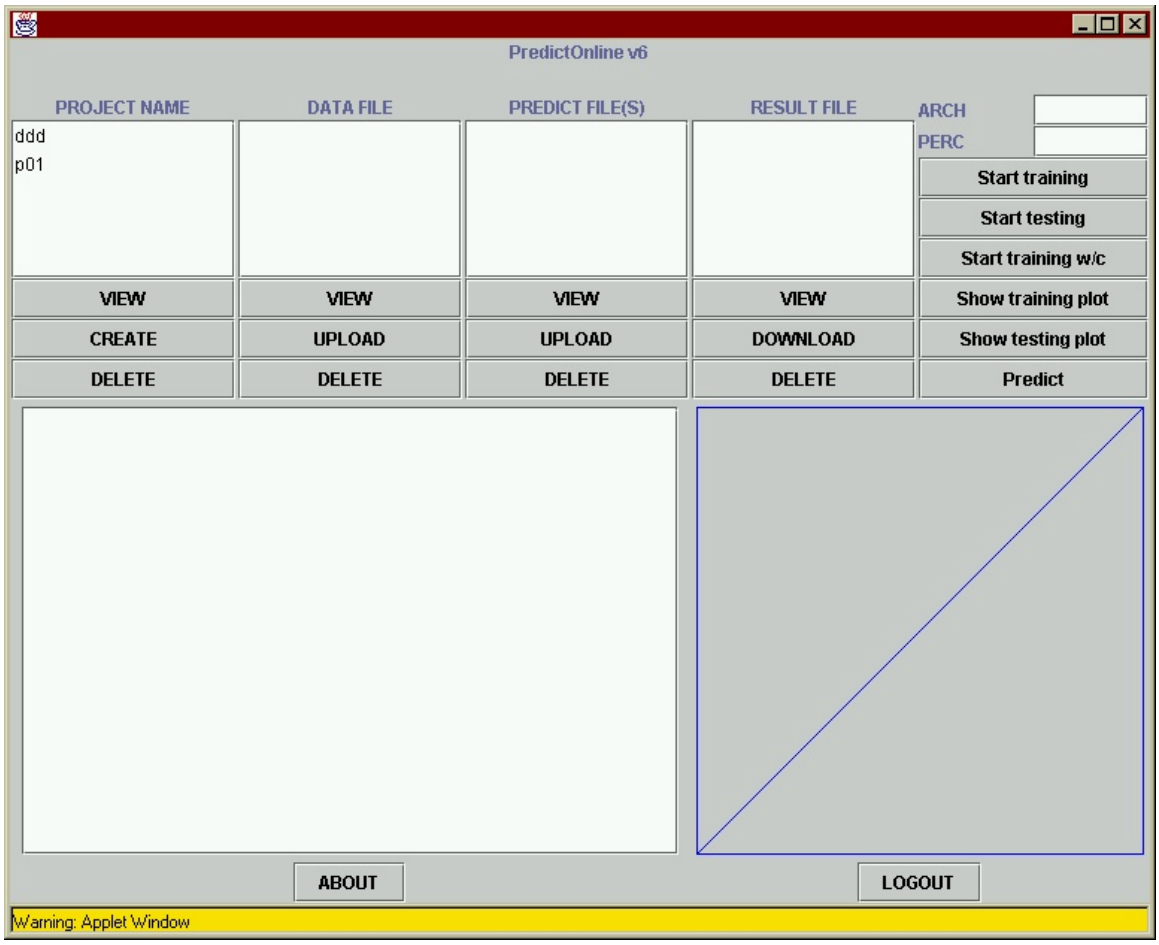

Figure 47. *PredictOnline* main window—project selected.

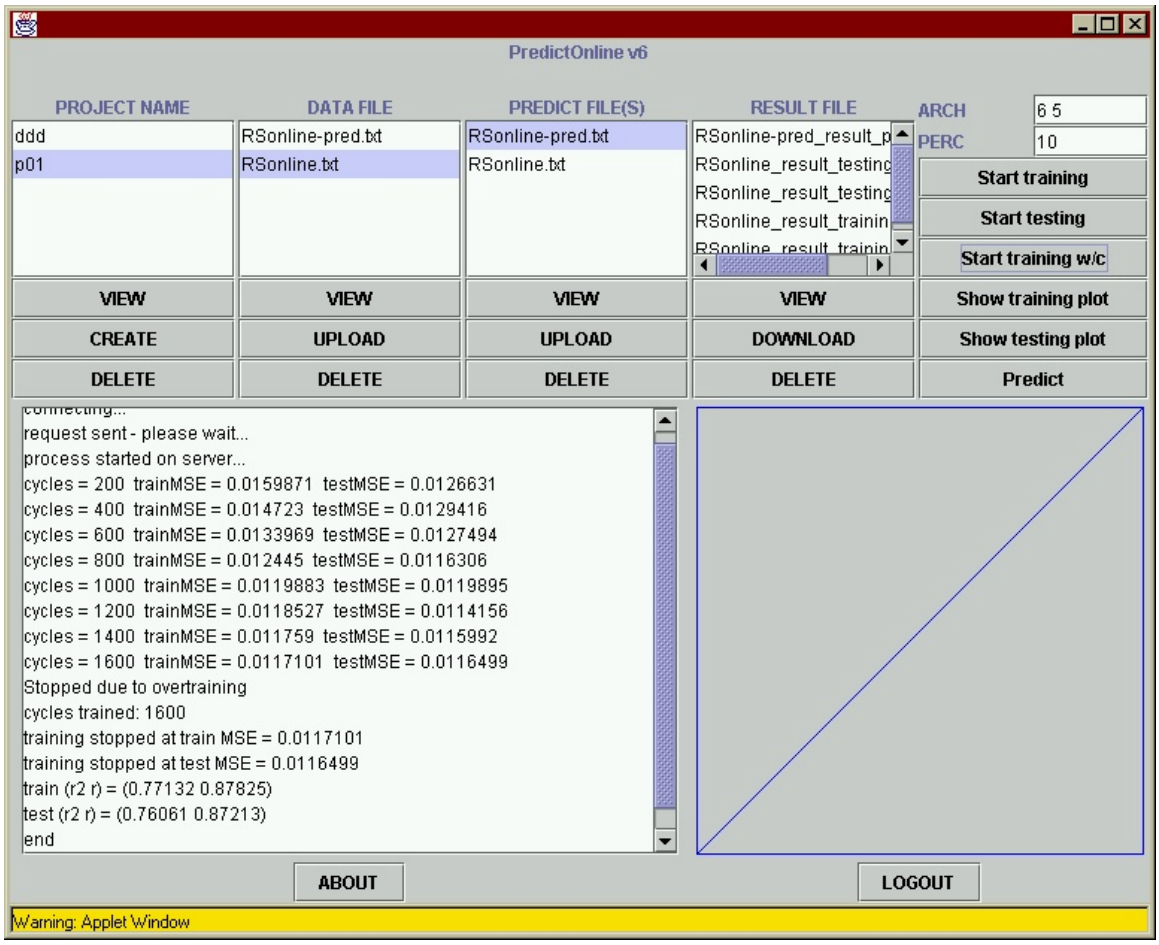

Figure 48. *PredictOnline* main window—training.

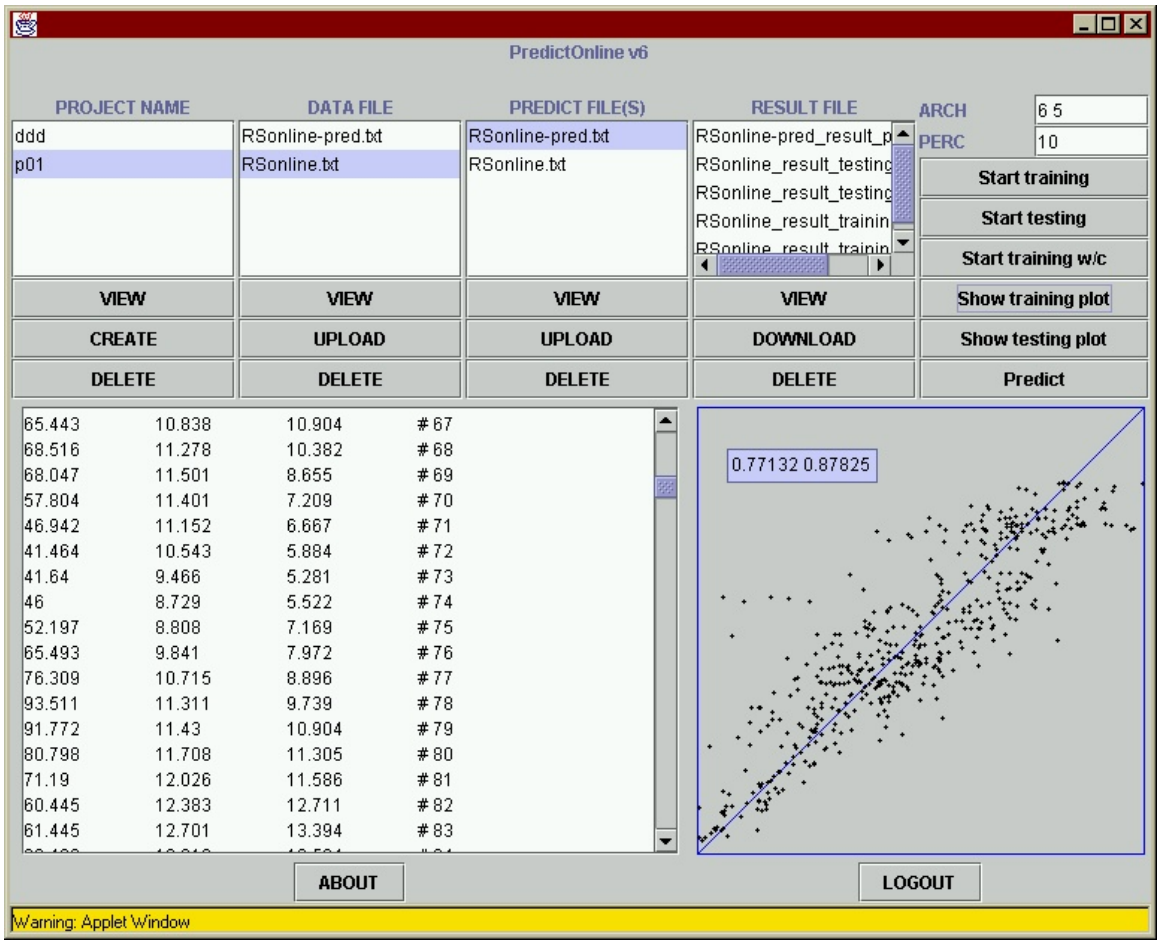

Figure 49. *PredictOnline* main window—training plot.

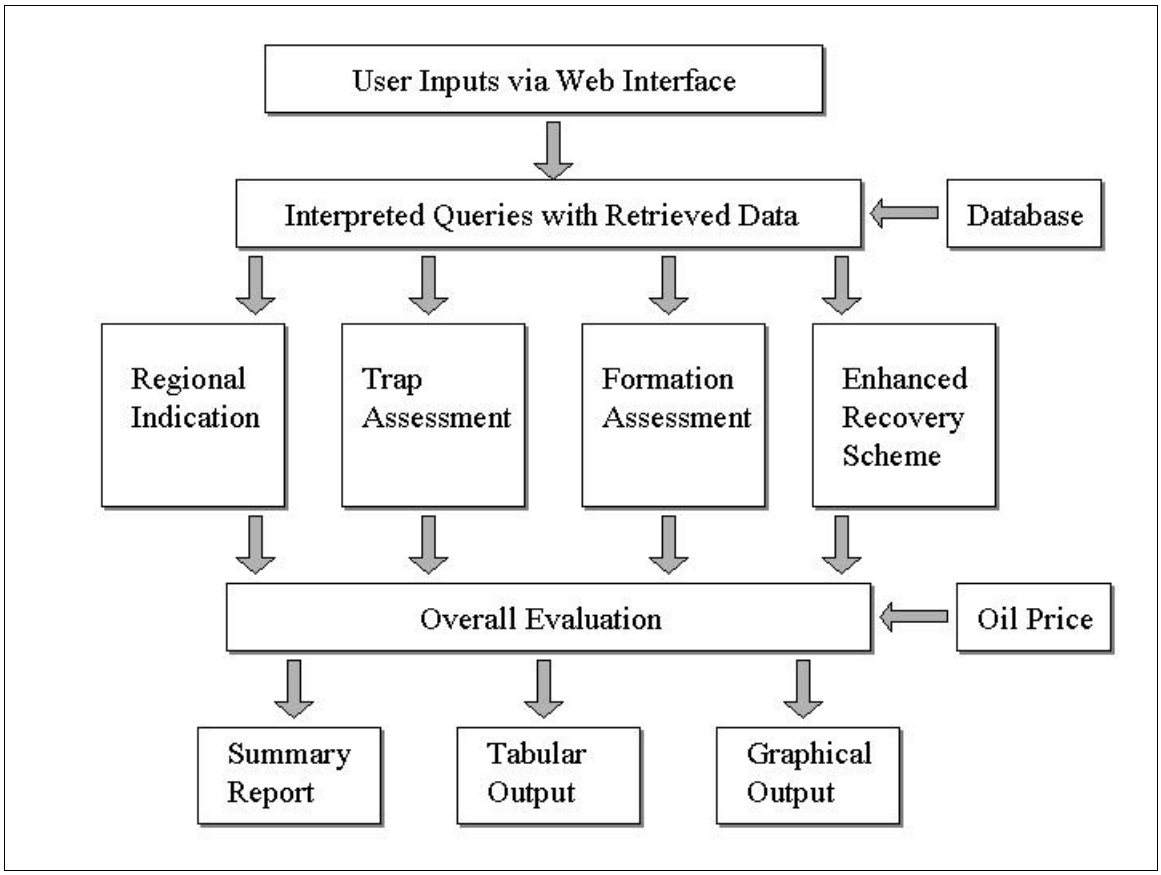

Figure 50. Inference workflow diagram.