R01

PRRC 99-14

Risk Reduction with a Fuzzy Expert Exploration Tool

(First Quarterly 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:	July 1, 1999
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 through June 30, 1999

Objectives

Incomplete or sparse information on types of data such as geologic or formation characteristics introduces a high level of risk for oil exploration and development projects. "Expert" systems developed and used in several disciplines and industries, including medical diagnostics, have demonstrated beneficial results. A state-of-the-art exploration "expert" tool, relying on a computerized data base and computer maps generated by neural networks, is proposed for development through the use of "fuzzy" logic, a relatively new mathematical treatment of imprecise or non-explicit parameters and values. Oil prospecting risk can be reduced with the use of a properly developed and validated "Fuzzy Expert Exploration (FEE) Tool."

This tool will be beneficial in many regions of the US, enabling risk reduction in oil and gas prospecting and decreased prospecting and development costs. In the 1998-1999 oil industry environment, many smaller exploration companies lack the resources of a pool of expert exploration personnel. Downsizing, low oil prices, and scarcity of exploration funds have also affected larger companies, and will, with time, affect the end users of oil industry products in the US as reserves are depleted. The proposed expert exploration tool will benefit a diverse group in the US, leading to a more efficient use of scarce funds and lower product prices for consumers.

This first of twenty quarterly reports contains an account of the progress, problems encountered, plans for the next quarter, and an assessment of the prospects for future progress.

Summary of Technical Progress

To ensure data-driven research, the Brushy Canyon interval of the Delaware sands (Fig. 1) in the Delaware Basin was selected as the focus of the expert exploration tool development.

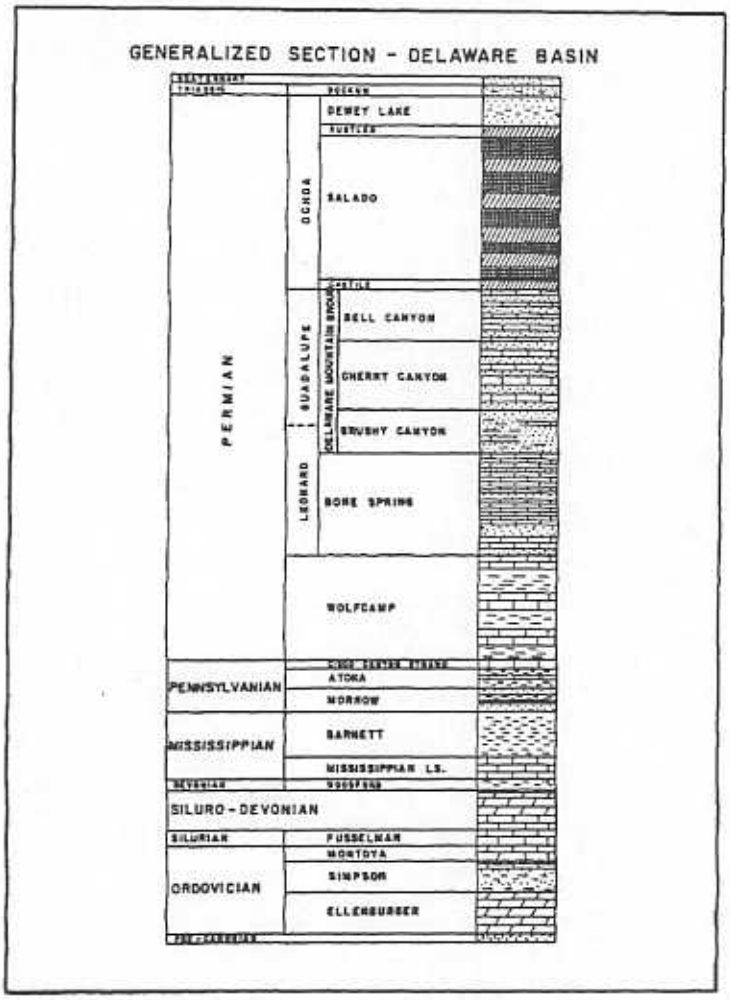

Fig. 1. Delaware Basin generalized stratigraphic column.

The Delaware Basin is a region where oil and gas exploration and development has continued through downturns in oil prices. Thus, the research will be driven by a problem with real field utility. Progress was made on the three major tasks required to meet the objective.

- Data gathering consisting of production history, wireline logs, gravity and aeromagnetic data, and seismic information was initiated and will continue throughout the project.
- Four expert system shells were found on the WWW that are candidates for developing the FEE Tool. Computational intelligence work commenced with the evaluation of a number of neural network architectures to evaluate Brushy Canyon logs. A first look at applying fuzzy logic to log evaluation was completed.
- Technology transfer consisted of introducing the project on the World Wide Web and preparation of three technical papers.

Data Assimilation

Production history, wireline log data, scout tickets, aeromagnetic, gravity, and seismic information are the initial components of the exploration database. The problem of compiling the data into a form suitable for searching by the expert system was addressed. Two database engines were considered. Oracle Corporation's *Oracle* is currently the most powerful of the commercial database engines that were reviewed. Microsoft offers two phases to developing a database engine. The first is the low-level product *Access,* followed by the more powerful *SQL Server.*

Two types of data are being compiled. Well data such as production history, oil shows, and interpreted log information are considered as single-point information. Spatially continuous information such as seismic, aeromagnetic, gravity, and regional structure maps is the second type of data, which is stored using a relational database.

Access will be the initial database engine. The software is inexpensive, relatively user-friendly, and is suitable for perhaps a 200-well database. Development of the prototype expert exploration tool will use *Access*, and will be accessible to consortium members via the Internet. As the project storage requirements grow, the *Access* tables will be transferred to *SQL Server,* recognizing that the entity relationships will have to be rewritten. *Oracle* does have a factor-of-ten speed advantage over *Access*, but the cost is also tenfold that of *Access*, and is not warranted at this stage of the research program. The PTTC Southwest Region is using a similar strategy to develop an internet-accessible production database for all wells in New Mexico. It is anticipated that the PTTC project will be completed in 2002, and the FEE Tool will link to the faster PTTC database engine for production records.

Wireline Log Data

Regional cross-sections at key places in the basin were used to select the initial 100 wells for inclusion in the database. The top and bottom of the Brushy Canyon interval in key wells were identified for the regional cross sections. The geologically significant markers are:

- Top of Brushy Canyon
- Top of Lower Brushy Canyon
- Base of Brushy Canyon

The markers have been used by others[1] and are relatively easy to recognize and correlate. The markers are the templates for other wells in the basin. These data will be

used in construction of structure maps, isopach maps, and are necessary for acquisition of porosity and other reservoir data as well as for source rock calibration/assessment.

Quality control of this critical part of the data acquisition process is assured by having all picks cross-checked by the co-PI who is a geologist. For consistency and quality control no more than two geologists will correlate the tops; additional interpreters could result in the confusion commonly encountered when reviewing scout ticket information.

The picks on four to six wells per township are being resolved and recorded. This well density is sufficient for the initial structure and isopach maps. Improved well density will result when the project goal of one well per section in the producing fields (about 500 wells) is reached. A photocopied record of the interval is being maintained for digitizing.

Production Data

The past seven years of production history from 2092 Delaware formation wells in 184 fields was ported to *Access* for querying from a 1.1 gigabit database with 330,000 entries. Cumulative production from the Delaware formation vs. time is shown in Fig. 2.

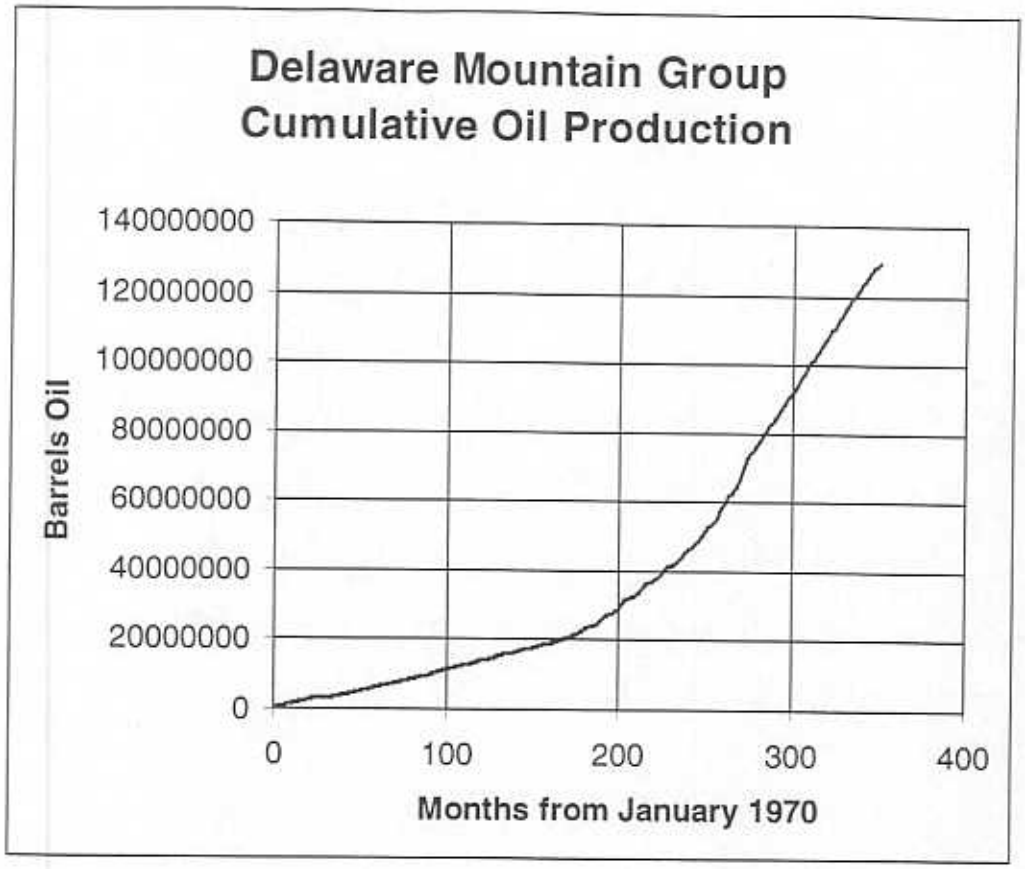

Fig. 2. Delaware zone production history.

Seismic Information

The commonly held idea that 3D seismic technology had displaced the need for 2D surveys was dispelled when attempts were made to induce owners of old 2D seismic lines to place their surveys in the public domain. The 2D lines shot by Western Geophysical and Permian Exploration Corporation are shown in Fig. 3, which includes Delaware zone producing fields.

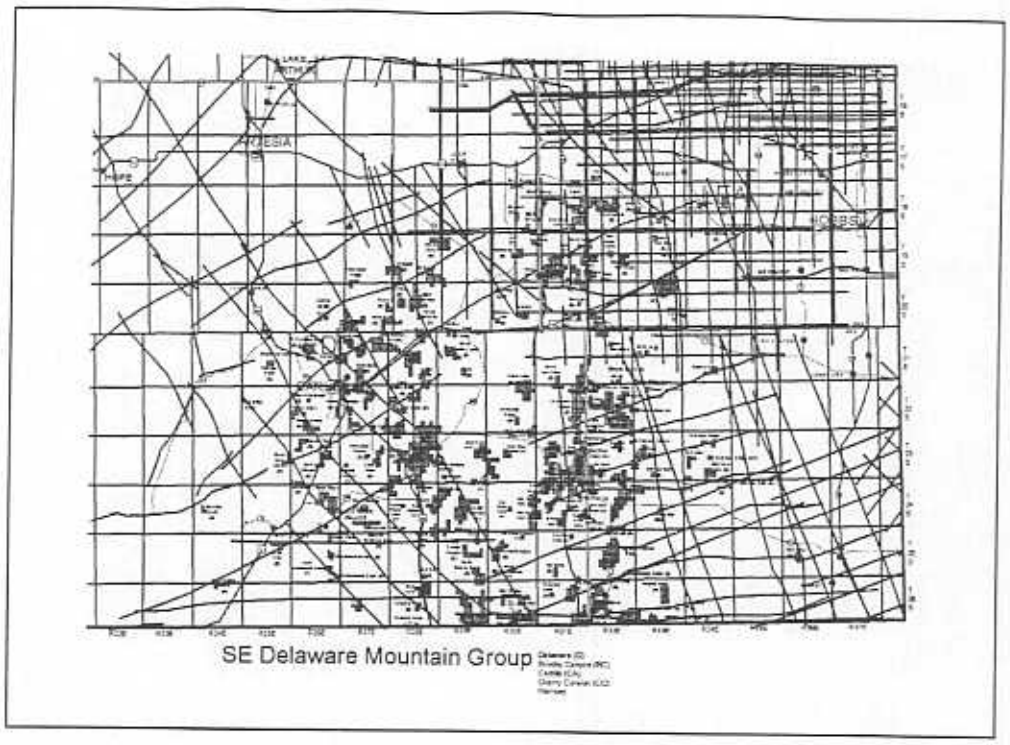

Fig. 3. 2D Seismic Lines collected by Western Geophysical and Permian Exploration with locations of Delaware zone producing fields.

An owner of a large number of 2D lines in the New Mexico Permian Basin stated that current 2D seismic license fees amount to $1,000,000 per year. Nevertheless, two operators have expressed a willingness to provide 2D information sufficient to establish regional lines through the New Mexico Delaware Basin. Additionally, Strata Production Company is providing the results of the Nash Draw Brushy Canyon interval 3D-seismic attribute analyses performed earlier by members of this research group.

Gravity Data

Gravity surveying measures the average density of the earth below each measurement. Gravity data was acquired from the National Geophysical Data Center in

Denver and mapped as shown in Fig. 4. Anomalies in the distribution of these points may correlate with Delaware fields.

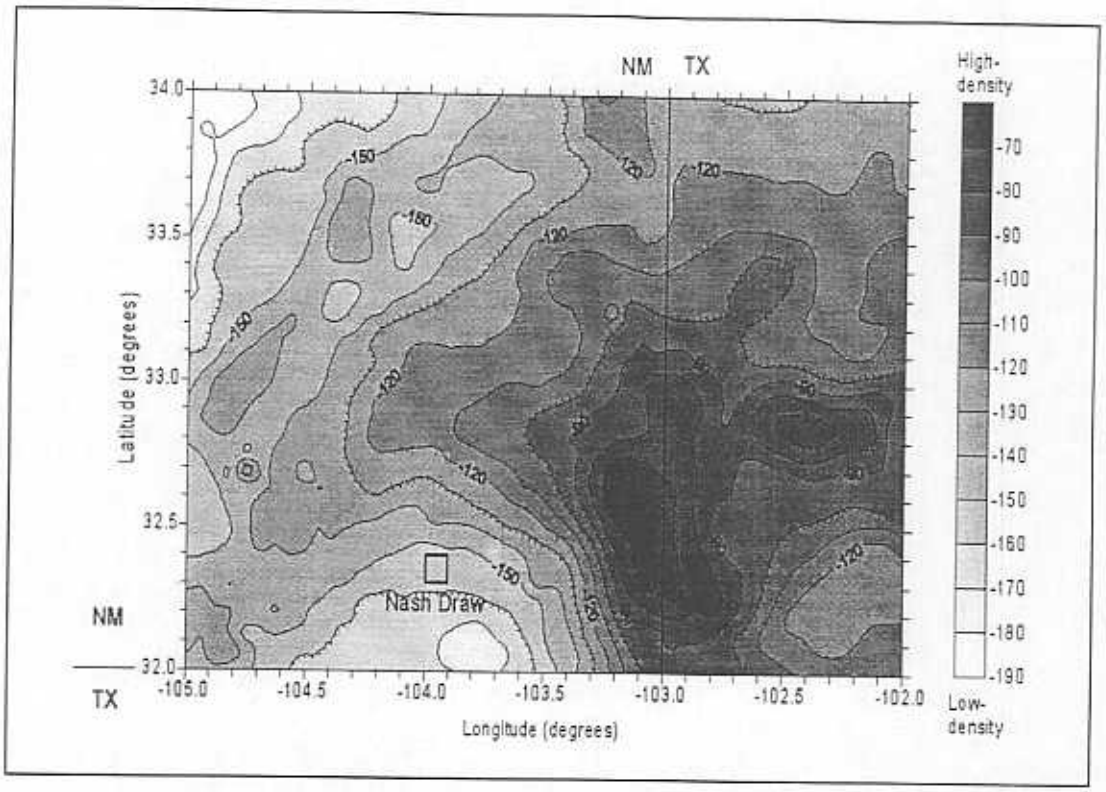

Fig. 4. Distribution of gravity anomalies across the Delaware Basin with respect to Nash Field.

The Bouguer gravity anomaly can be understood through a simple explanation on a global scale. The Bouguer anomalies around the world for sea level are approximately equal to zero. In regions of high elevations the Bouguer anomalies are negative, while for ocean regions the anomalies are positive. These large-scale effects are due to density variations in the crust, indicating higher density material beneath the oceans and lower density material in areas of elevated land. This effect is well demonstrated in the gravity map from southeastern New Mexico. The Central basin platform is an elevation high for

the region. It also has a less negative Bouguer anomaly for the region as well. This indicates an area of higher density material within an area of relatively lower density rock material. Therefore, the Bouguer anomalies tend to reflect variations within basement structures as well as imply more subtle trends within the overlying sedimentary structure.

Aeromagnetic Data

Researchers at the NASA-funded Pan American Center for Earth and Environmental Studies (University of Texas at El Paso), headed by Dr. Randy Keller, recently compiled and merged data sets from west Texas and southeastern New Mexico into an aeromagnetic database for the region (Fig. 5). The data was obtained from several aeromagnetic surveys flown over these areas during the 1950s and 60s.

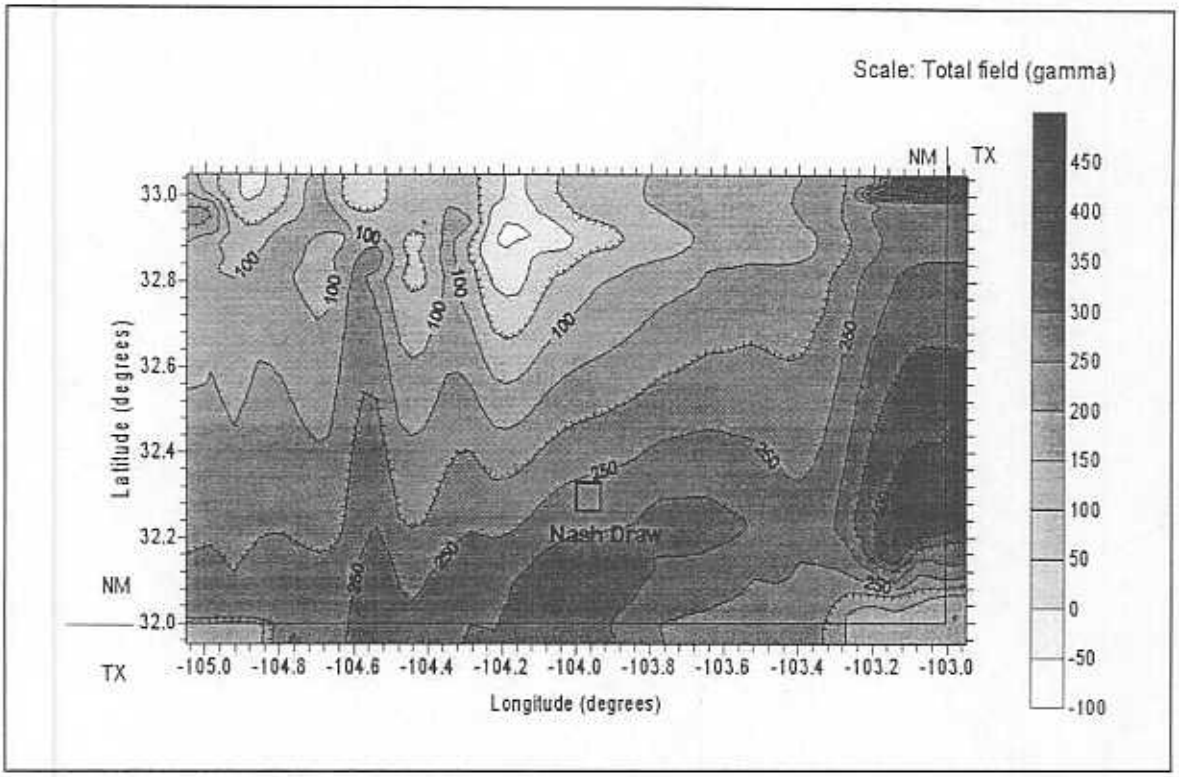

Fig. 5. Merged aeromagnetic surveys showing location of Nash Draw Unit.

The magnetic data included in the FEE Tool 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.

Computational Intelligence

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 I). The choices are to develop an in-house system based on CLIPS such as other vendors have created, or to begin with one of the other systems.

Table I
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

Log Interpretation

Several concepts based on forward or inverse 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[3] and resistivity modeling.[4]

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[3] 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[5] 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:

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

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 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 Permian 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 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 it and the output variable, as shown in Table 1.

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, Sw). 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 below in Fig. 6 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 II). Because the neural network can overtrain, it is trained to several different correlation coefficients to determine which produces the highest testing correlation coefficient.

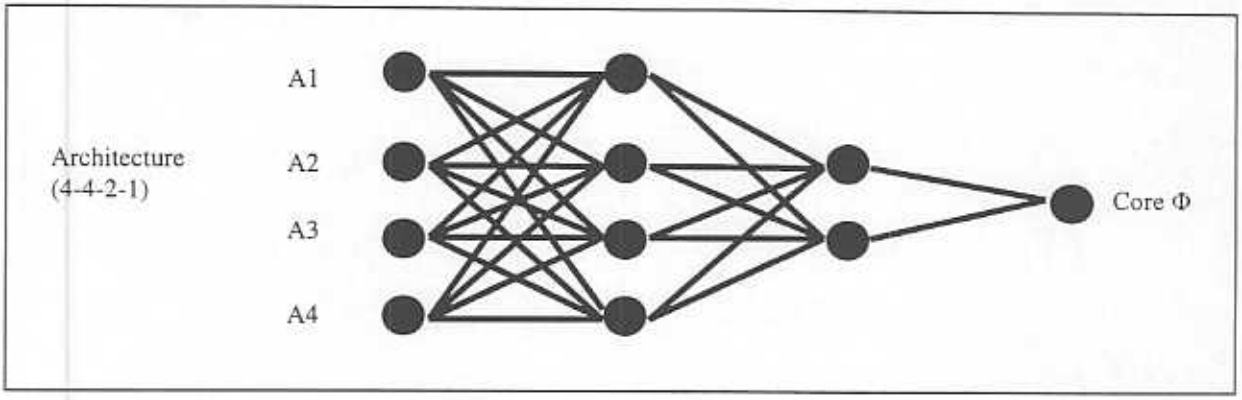

Fig. 6. A 4-4-2-1 neural network architecture. This architecture is four-layered with one input layer, two hidden layers and one output layer.

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.

Table I. Results of fuzzy ranking. Data set is 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

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 permeability (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 II. 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, Table II. Reasonably good correlations also exist between LLD and core permeability (0.51 cc) and DPHI (0.68 cc), Table II.

Table II. 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))

Results of the neural network training are compiled in Table III. 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 (left 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 (left 161 training points and 53 test points) instead of every fifth point (left 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 III.

Although the neural network provides reasonably good 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. 7.

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

Poor correlations were also obtained between log measurements and core So by itself (training to 0.56 and testing to 0.27 cc), Table III. 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 So (training to 0.62 and testing 0.52 cc), Table III.

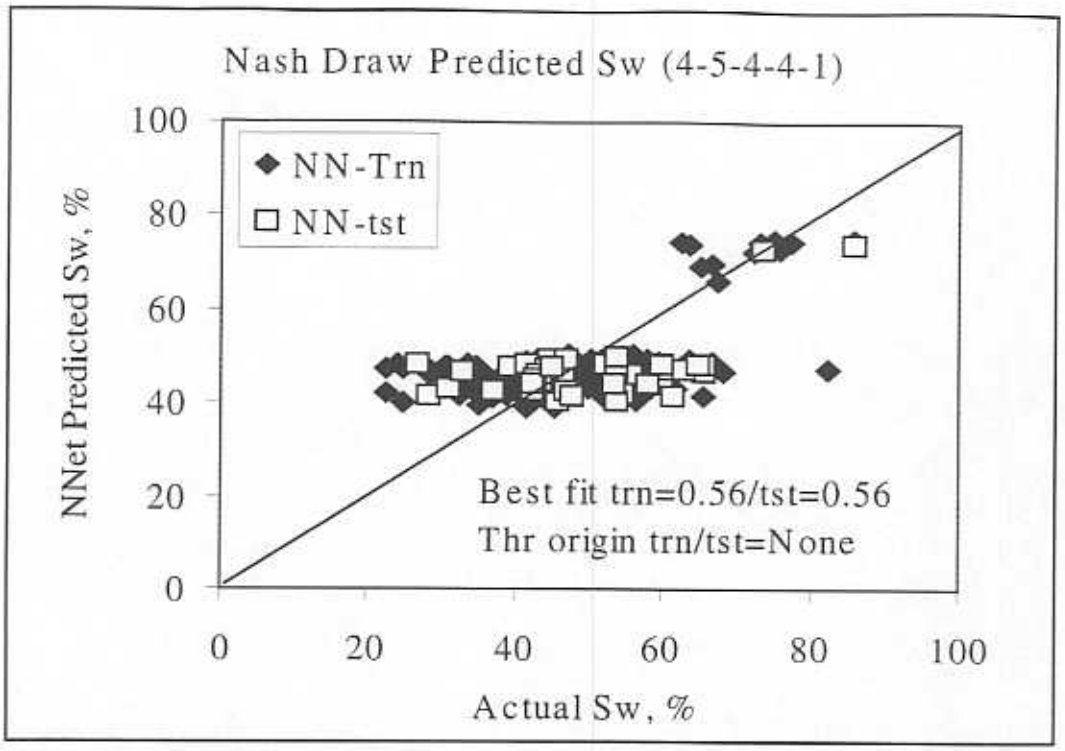

Fig. 7. No correlation with visual examination.

Table III. Nash Draw neural network correlations. The neural net program calculates cc 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)

Fuzzy Rule Log Interpretation

Wireline logs are widely used by subsurface geologists in exploration, and they provide a vital source of information in prospect risk assessment. Many different types of logs are in use today. Log curves may be interpreted directly by human experts, and software tools based on different decision algorithms have been developed to assist in the task. Log interpretation, however, remains a complicated and labor-intensive task. As a beginning part of our project to develop a system for risk assessment, we will investigate an advanced technique for log interpretation to fully exploit the available log data in the Delaware Basin.

In this study, we have developed a fuzzy logic-based algorithm to interpret log curves. Since there is almost always uncertainty, ambiguity, even inconsistency in the information inferred from log curves, fuzzy logic presents itself as a basis for developing better log interpretation methods. There are several reasons for this:

A. fuzzy systems have proved to be effective in many applications characterized by informational uncertainty[6,7,9];

B. neural network-based methods in log interpretation have achieved some degree of success, and equally powerful fuzzy systems can be built, in theory, without the need of network training[6,8,9];

C. a fuzzy logic-based interpretation system is much more adaptable and less costly than neural networks.

For ease of illustration of our initial research, we concentrate here on determining formation types from the interpretation of two logs, density and gamma ray log curves.

Log Curves

Logging provides an indirect, less expensive (than coring) way of gathering the subsurface geological information and log data that will be used extensively in our risk assessment procedure. The technique that we have developed, briefly described in the following section, will allow users to use all types of log data. As such, it will complement the method of importance ranking described earlier.

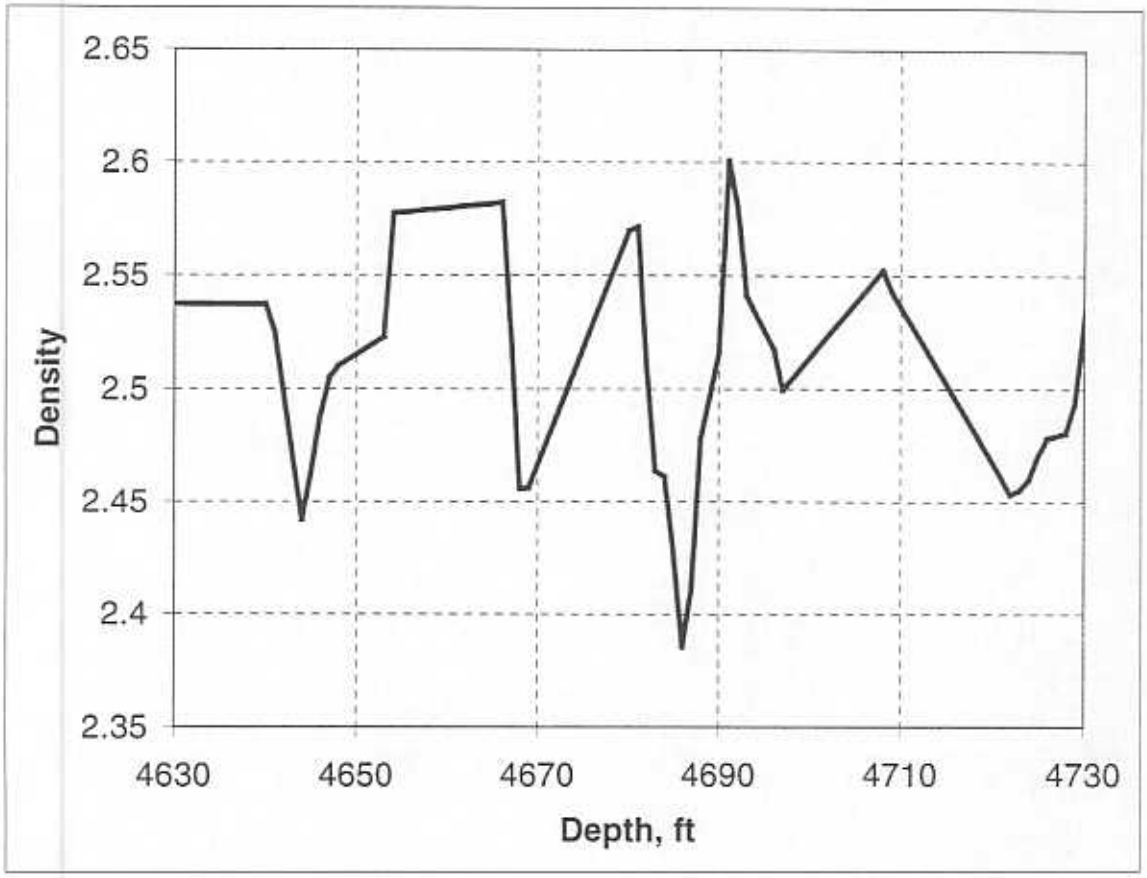

Fig. 8. A portion of a density curve.

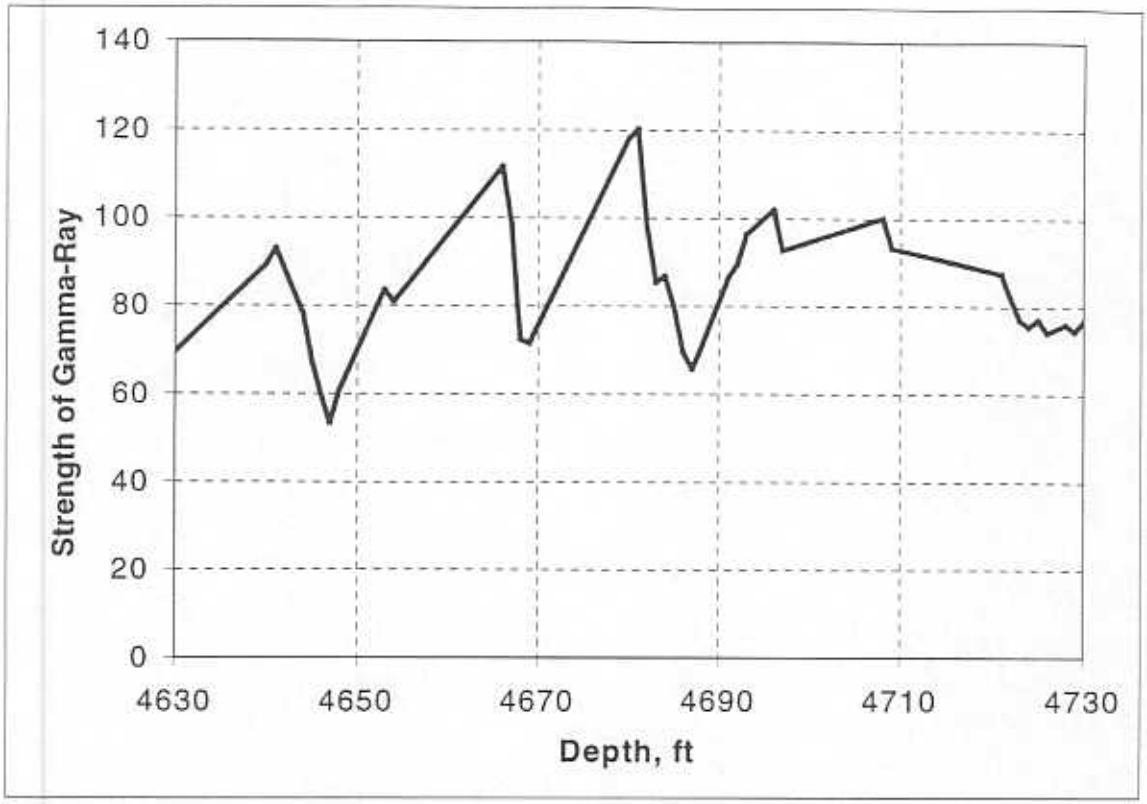

Fig. 9. A portion of a gamma ray curve.

Figure 8 is a portion of a density log curve and Fig. 9 is a portion of a gamma ray curve. Suppose we wish to interpret the curves to determine formation types. Typical densities in gram/cm^3 of common pure rocks are given in Table IV. This table will be used as the basis for interpreting the density log curve. Typical relative strengths of gamma ray for the rocks, normalized to the range of [0, 10], are also given in Table IV. This will be the basis to interpret the gamma ray log curve.

Table IV. Densities and gamma ray strengths of pure rocks		
Rock Type	Density (g/cm^3)	Gamma Ray (relative)
Dolomite	2.83 – 2.87	2 – 4
Limestone	2.71	2 – 4
Sandstone	2.65	4 – 7
Shale	2.3 – 2.7	> 7
Gypsum	2.33	2 – 3
Salt	2.08	1 – 3

An actual formation in a reservoir, however, may consist of one or more pure rocks; that is, the formation is a mixture of various rocks. Therefore, the density value at a certain depth on the density log curve may not necessarily be equal to any density values given in Table IV. A similar phenomenon occurs with the gamma ray log curve. This is one source of uncertainty in the interpretation of the log curves.

In addition to the uncertainty that resulted from the presence of different rocks in a formation, another uncertainty exists in the interpretation of the log curves. It can be seen from Table IV that density, even for some pure rocks such as dolomite and shale, is not a unique value; and that the density of gypsum falls right in the density range of shale. The situation is even worse for gamma ray values: for each rock listed in the table, its gamma ray value is not unique but falls in a range; and the ranges of different rocks overlap. How do we deal with this difficulty in log curve analysis?

Fuzzy Logic-Based Interpretation

Based on the knowledge of experts, we can formulate a "decision credibility matrix," containing pairs of weights (of credibility) that indicate the degree of plausibility of inferred formation types from analyzing the density log curve and the gamma ray log

curve. Table V, the decision credibility matrix for our example, is a sort of "rule base" that is obtained after an investigation of the rules used by a human expert. For simplicity, we will consider only the rocks listed in Table IV.

Table V. Rule base for interpretation of the log curves

Density Log / Gamma Ray Log	Dolomite	Limestone	Sandstone	Shale	Gypsum	Salt
Dolomite	X	DL(1.0)	DL(1.0)	GR(1.0)	DL(1.0)	DL(1.0)
Limestone	DL(1.0)	X	DL(1.0)	GR(1.0)	DL(1.0)	DL(1.0)
Sandstone	GR(0.3) DL(0.7)	GR(0.8) DL(0.2)	X	GR(0.9) DL(0.1)	GR(0.6) DL(0.4)	GR(0.6) DL(0.4)
Shale	GR(0.7) DL(0.3)	GR(1.0)	GR(0.9) DL(0.1)	X	GR(0.7) DL(0.3)	GR(0.7) DL(0.3)
Gypsum	DL(1.0)	DL(1.0)	GR(1.0)	GR(1.0)	X	DL(1.0)
Salt	DL(1.0)	GR(0.1) DL(0.9)	GR(1.0)	GR(0.8) DL(0.2)	DL(1.0)	X

In Table V, DL denotes the density log curve and GR denotes the gamma ray log curve. Figures within parentheses next to DL or GR in the table denote the degrees of credibility of the corresponding conclusion. For instance, DL(0.7) means that a weight of 0.7 is given to the conclusion arrived at from the density log curve; while GR(0.3) means that a weight of 0.3 is given to the conclusion reached from the gamma ray log curve. In other words, DL(0.7) means that the likelihood that "the actual formation type is identical to that interpreted from the density log curve" is 70%; similarly, GR(0.3) means that the likelihood "the formation type is identical to the result interpreted from the density log curve" is 30%. A zero weight for either DL or GR is omitted; and X's in the table mean that results inferred from the two log curves are identical, that is, both DL = 1.0 and GR =1.0. Therefore, the result inferred from either log curve is correct.

Two groups of fuzzy sets are defined for the density log curve and the gamma ray log curve, based on the data given in Table IV and the opinion of an expert. One group of fuzzy sets for each of the six rocks, according to each of the two log curves is shown in Figs. 10 and 11, respectively. In Fig. 10, the range of density values for each of the six variables is symmetrically extended 0.1 on both sides of the range. Thus, we have the fuzzy set below for salt, where d is a density value.

$$D_{\text{sal}}(d) = \begin{cases} 10(d - 1.98) & \text{if } d = [1.98, 2.08) \\ 10(2.18 - d) & \text{if } d = [2.08, 2.18] \\ 0 \text{ otherwise} \end{cases}$$

Similarly, we define density-based fuzzy sets for the other five variables: gypsum, shale, sandstone, limestone and dolomite.

Likewise, the range of values of relative gamma ray strength values for each of the six variables is symmetrically extended by 1.0 on each side of the range. Thus, we have the fuzzy set definition below for salt in terms of gamma ray strength, g. Similarly, we define fuzzy sets for the other five variables: gypsum, shale sandstone, limestone and dolomite.

$$G_{\text{sal}}(g) = \begin{cases} g & \text{if } g = [0,1) \\ 1 & \text{if } g = [1,3] \\ 4 - g & \text{if } g = (3,4] \\ 0 & \text{otherwise} \end{cases}$$

Similarly, gamma ray strength fuzzy sets for the other five variables are defined.

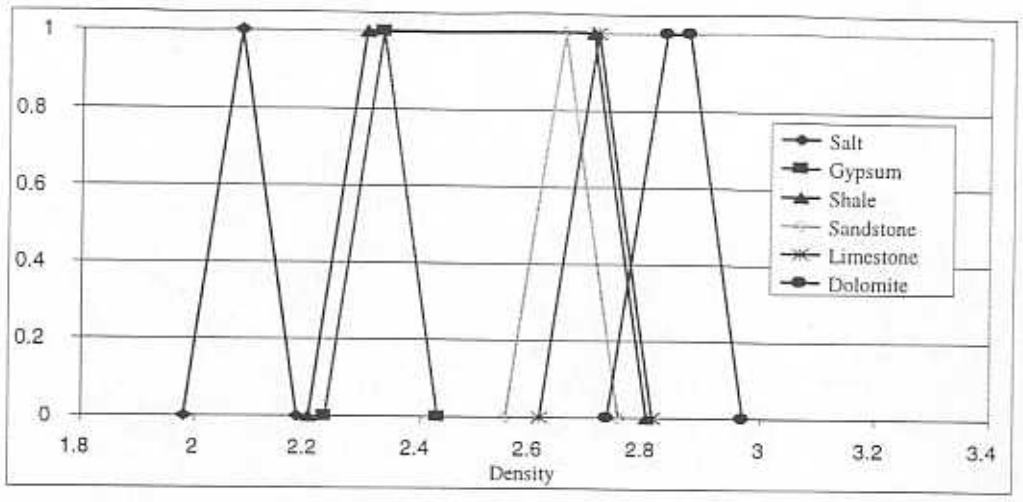

Fig. 10. Fuzzy sets for densities of rocks.

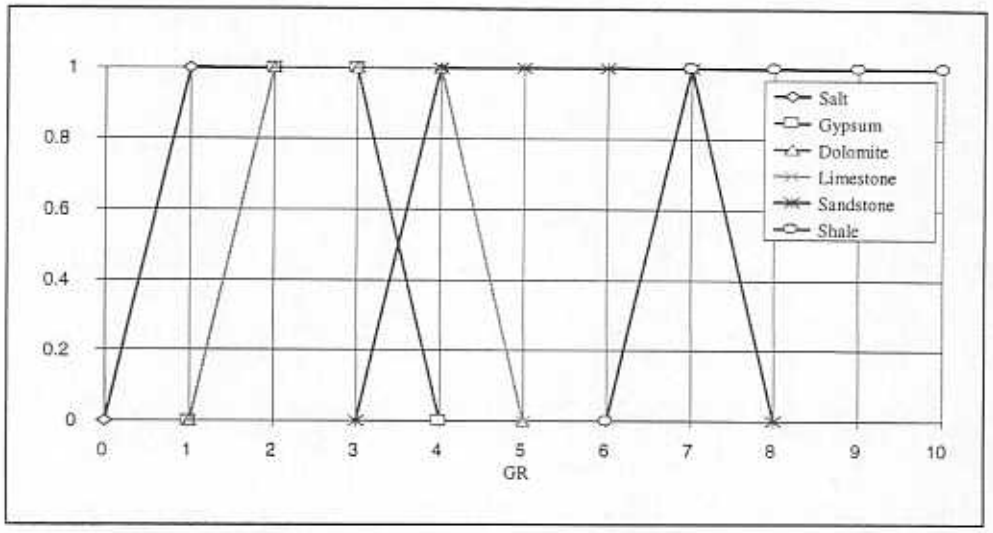

Fig. 11. Fuzzy sets for gamma ray strengths of rocks.

From the above two groups of fuzzy sets defining the densities and gamma ray strengths of the six rock types, and given a density d and a gamma ray strength g, we obtain two vectors giving the membership grades in the six rock types, according to, respectively, density and gamma ray strength:

$$\mathbf{D}(d) = <D_1(d), D_2(d), D_3(d), D_4(d), D_5(d), D_6(d)>,$$

$$\mathbf{G}(g) = <G_1(g), G_2(g), G_3(g), G_4(g), G_5(g), G_6(g)>$$

Equivalently, $\mathbf{D}(d)$ may be viewed as representing the following fuzzy set:

$$\mathbf{D}(d) = / \text{ dolomite} + D_2(d) / \text{ limestone} + D_3(d) / \text{ sandstone}$$

$$+ D_4(d) / \text{ shale} + D_5(d) / \text{ gypsum} + D_6(d) / \text{ salt}$$

Given a specific density reading d from the density log, the meaning of $\mathbf{D}(d)$ is that d points to dolomite with degree of $D_1(d)$, d points to limestone with degree of $D_2(d)$, etc.

Fuzzy Inference Method

Our objective is to apply a fuzzy logic-based reasoning method to determine formation types from the log curves. When human experts interpret log curves, they often face two problems:

A. How to reconcile the multiple conclusions that may be inferred from a single log curve,

B. How to determine the formation type if the results inferred from the density log curve and the gamma ray curve differ.

Here we propose an inference method for this problem based on the decision credibility matrix, as given in Table V. (It is noted that our Table V. is based on the knowledge of a single expert and needs to be modified to reflect the collective opinion of a group of experts for better interpretation results)

The item GR(0.8)/DL(0.2) at row Salt and column Shale in the table, for example, would be interpreted as

If	*density determines it is shale*	***and***	*gamma ray determines it is salt*
Then	*the degree of credibility of shale is 0.2*	***and***	*the degree of credibility of salt is 0.8.*

From Table V, we can extract two matrices, the density possibility matrix and the gamma ray possibility matrix, given below:

$$\begin{bmatrix} DL(1,1) & \cdots & DL(1,6) \\ \cdots & \cdots & \cdots \\ & DL(i,j) & \\ \cdots & \cdots & \cdots \\ DL(6,1) & \cdots & DL(6,6) \end{bmatrix} \qquad \begin{bmatrix} GR(1,1) & \cdots & GR(1,6) \\ \cdots & \cdots & \cdots \\ & GR(i,j) & \\ \cdots & \cdots & \cdots \\ GR(6,1) & \cdots & GR(6,6) \end{bmatrix}$$

The inference method we adopt consists of three steps. The first step is to calculate the two vectors $\mathbf{D}(d)$ and $\mathbf{G}(g)$ for density d and gamma ray g, respectively, from the two given log curves. These two vectors of membership grades in different rock types are evaluated from the corresponding group of fuzzy sets, as described earlier. In the second step, the rules in Table V are "fired", and we achieve two groups of fuzzy sets for rock type represented by the following two equations:

$$\mathbf{RD}(i) = \sum_{j=1}^{6} DjDL(i,j) \quad \text{for } i = 1,\cdots,6$$

$$\mathbf{RG}(j) = \sum_{i=1}^{6} GiGR(i,j) \quad \text{for } j = 1,\cdots,6$$

Where Dj is the jth component of $\mathbf{D}$, and Gi is the ith component of $\mathbf{G}$. $\mathbf{RD}(i)$ the ith (vector) fuzzy set for rock type derived from the evidence of density, and $\mathbf{RG}(j)$ is the jth (vector) fuzzy set for rock type derived from the evidence of gamma ray.

The final step gives the conclusion, **FT**, for rock type, which is calculated using the formula:

$$\mathbf{FT} = \sum_{i=1}^{6} \mathbf{RD}(i) + \sum_{j=1}^{6} \mathbf{RG}(j)$$

The **FT** is somewhat similar to the *scaled fuzzy set* in fuzzy logic controllers; it indicates the formation types in the reservoir, i.e., the different types of rocks that exist in a formation and the likelihood of each's presence. Thus, **FT** gives the interpretation we seek for the formation types.

Technology Transfer

During the first quarter the REACT group homepage was revamped to include the Fuzzy Expert Exploration Tool proposal. The WWW address for this proposal is http://baervan.nmt.edu/REACT/reacthomepage.htm, and it can also be accessed through the PRRC homepage at http://baervan.nmt.edu/.

A paper[10] was presented on the application of computational intelligence to log analysis at the SPE 1999 Rocky Mountain Regional Meeting in Gillette Wyoming.

A presentation[11] concerning the application of new seismic attribute analysis technology to a Brushy Canyon pool was made at the NPTO/FETC 1999 Oil & Gas Conference.

Future

Geology

Basically, the more wells that are included in the database, the better it will be. Sometime toward the end of the summer, correlated wells will be used to make additional cross-sections throughout the basin as an additional means of QA/QC.

Production

Digital data from 1970 to 1999 will be included in the *Access* database. The database will be queried to produce field production profiles of oil, water, and gas as well as their derivatives GOR and WOR.

Ultimate Delaware zone production will be estimated and a Delaware Hubbert curve will be developed.

Computational Intelligence

We will construct a shell for the Microsoft *Access* database in which all data will be entered for computer/digital analysis including neural network analysis. A decision on the expert system shell will be made by the fourth quarter.

Seismic

Current research using seismic attributes is exploring the possibility of generating depth maps using seismic attributes and computational intelligence. Most depth maps are made using wells as control points, and interpolating between wells and extrapolating beyond areas with well control. If a function to compute depth using seismic attributes at Nash Draw can be obtained, attributes along the major basin transecting 2D lines may also be calculated accurately. These "depth" lines will provide support for the geologic maps interpolation schemes that predict depth away from well bores.

Technology Transfer

An initial consortium meeting is tentatively scheduled for September. A paper[12] concerning the application of computational intelligence to Brushy Canyon zone 3D seismic attributes has been accepted and prepared for the SPE 1999 Annual Fall Meeting in Houston Texas.

Assessment of the prospects for future progress

Data assembly is progressing on schedule during the early part of the project; however, it appears that the number of public domain 2D lines will be restricted. By the end of the second quarter all the graduate students should be employed on the project and development of the fuzzy rules should commence. Fuzzy rule development will include interviews with explorationist who have discovered Brushy Canyon pools.

Successful development of computational intelligence tools for log interpretation will speed populating the database. In a similar manner, research into correlating seismic attributes to formation velocity variations, resulting in better depth maps, should speed integration of disparate seismic information by the fifth quarter.

An initial version of *Predict* should be available for consortium members during third quarter. Once the FEE Tool database is sufficient to justify organization of the Consortium, a conference will be held. The first meeting is tentatively scheduled for the project's forth quarter.

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. Hoylman, H. W. "Evaluation of Magnetics in the Delaware Basin," *World Oil,* **133** no.3, (1951) 91-98.

3. 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).

4. 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 & June 1999).

5. Gaymard, R. and Poupon, A.: "Response of Neutron and Formation Density Logs in Hydrocarbon Bearing Formations," *The Log Analyst* (Sept.-Oct. 1968).

6. Klir, G.J. and Yuan, B.: *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Upper Saddle River, NJ (1995).

7. Kosko, B. *Fuzzy Engineering*, Prentice Hall, Upper Saddle River, NJ (1997).

8. Hertz, J., Krogh, A., and Palmer, R.G.: *Introduction to the Theory of Neural Computation*, Addison-Wesley, Redwood City, CA (1991).

9. Jang, J-S. R., Sun, C-T., and Mizutani, E.: *Neural-Fuzzy and Soft Computing*, Prentice-Hall, Upper Saddle River, NJ (1997).

10. Weiss, W.W., Wo, S., and Balch, R.S.: "Integrating Core Porosity and S2 Measurements with Log Values," paper SPE 55642 presented at the 1999 SPE Rocky Mountain Regional Meeting, Gillette, May 15-18.

11. Weiss, W. W.: "Advanced Oil Recovery Technologies for Improved Recovery from Slope Basin Clastic Reservoirs, Nash Draw Brushy Canyon Pool, Eddy County, New Mexico," paper presented at the 1999 U.S. DOE Office of Fossil Energy, Federal Energy Technology Center and National Petroleum Technology Office Oil & Gas Conference, Dallas, June 28-30.

12. Balch, R.S., Stubbs, B.S. Weiss, W.W., Wo S.: "Using Artificial Intelligence to Correlate Multiple Seismic Attributes to Reservoir Properties," paper SPE 56733, to be presented at the 1999 SPE Annual Fall Meeting, Houston.