RISK REDUCTION WITH A FUZZY EXPERT EXPLORATION TOOL

Semi-Annual Report
March 15, 2001-March 15, 2002

By:
William W. Weiss

Date Published: June 2002

Work Performed Under Contract No. DE-AC26-99BC15218

New Mexico Institute of Mining and Technology
Socorro, New Mexico

**National Energy Technology Laboratory
National Petroleum Technology Office
U.S. DEPARTMENT OF ENERGY
Tulsa, Oklahoma**

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government.

This report has been reproduced directly from the best available copy.

Risk Reduction with a Fuzzy Expert Exploration Tool

By
William W. Weiss
Ron Broadhead

June 2002

Work Performed Under DE-AC26-99BC15218

Prepared for
U.S. Department of Energy
Assistant Secretary for Fossil Energy

Jim Barnes, Project Manager
National Petroleum Technology Office
P.O. Box 3628
Tulsa, OK 74101

Prepared by
New Mexico Petroleum Recovery Research Center
New Mexico Institute of Mining and Technology
801 Leroy Place, Kelly Bldg.
Socorro, NM 87801

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government.

This report has been reproduced directly from the best available copy.

Risk Reduction with a Fuzzy Expert Exploration Tool

By
William W. Weiss
Ron Broadhead

June 2002

Work Performed Under DE-AC26-99BC15218

Prepared for
U.S. Department of Energy
Assistant Secretary for Fossil Energy

Jim Barnes, Project Manager
National Petroleum Technology Office
P.O. Box 3628
Tulsa, OK 74101

Prepared by
New Mexico Petroleum Recovery Research Center
New Mexico Institute of Mining and Technology
801 Leroy Place, Kelly Bldg.
Socorro, NM 87801

Table of Contents

Table of Contents	iii
List of Tables	iv
List of Figures	v
Executive Summary and Objectives	1
Progress and Discussion of Results	3
Computational Intelligence	3
Web-Based Data Management System (WDMS).....	20
Engineering	22
Geology	26
Technology Transfer	33
Problems Encountered	33
Next Year's Tasks	34
References.....	36

List of Tables

Table 1. List of Variables Selected as Optimal.....	10
Table 2. Goodness of the Relationship between Commodity Future Price and Oil Price	32

List of Figures

Fig. 1. Histogram of average hydrocarbon equivalent. Value is average monthly production over the first year of production for Brushy Canyon wells.	37
Fig. 2. Log filtered version of histogram from Fig. 1. Notice that the data is much better behaved and follows a rough gaussian curve. It is easier to see outliers with the data in this format.	37
Fig. 3. Sample fuzzy membership functions drawn for two points of a 100-point random data distribution.	38
Fig. 4. Sample fuzzy membership functions drawn for two points of a 100-point random data distribution with a trend added.	38
Fig. 5. Fuzzy curves for the data sets in Figs. 3 and 4. Fuzzy curves are drawn by doing a weighted sum of all the fuzzy membership functions.	39
Fig. 6. Operation of second stage fuzzy ranking program FuzzyOnline. A) is a sample input file. In this case we wish to correlate to each of 36 data found in the database for 520 wells and the average monthly production we wish to correlate to. B) shows a progressive reduction in the number of possible inputs (initially 36) through several iterations. In each iteration the best fitting variable is identified, then the rest of the data is studied for uniqueness with respect to that best variable. In this case five iterations identified the four strongest and unique inputs. C) The generated output file shows the strength and ID number of the best inputs for regression analyses.	40
Fig. 7. Crossplot of neural network training using the inputs identified via FuzzyOnline.....	41
Fig. 8. Crossplot of testing data for the trained neural network of Fig. 7.....	41
Fig. 9. Map of predicted production potential based on the trained and tested neural network regression.	42
Fig. 10. This figure demonstrates the benefits of neural network regressions for complex problems where forecasts need to be made. Assuming only the five points enclosed by diamonds are known to start, it is quite easy to fit a polynomial that exactly fits the training data but poorly fits the true data distribution. Neural networks can bypass this problem if a sufficient sample exists for training by creating a generalized solution.	43
Fig. 11. The original schematic for the fuzzy expert system shell.	44
Fig. 12. More complicated system which breaks the analysis into several separate categories to simplify calculations and customization.....	44

Fig. 13. Initial rules for trap assessment (part I).	45
Fig. 14. Initial rules for trap assessment (Part II).	45
Fig. 15. Initial rules for trap assessment (Part III).	46
Fig. 16. Initial rules for source rock assessment (Part I).	46
Fig. 17. Initial rules for source rock assessment (Part II).	47
Fig. 18. Initial rules for source rock assessment (Part III).	47
Fig. 19. Initial rules for source rock assessment (Part IV).	48
Fig. 20. Initial rules for source rock assessment (Part V).	48
Fig. 21. Initial rules for regional assessment (Part I).	49
Fig. 22. Initial rules for regional assessment (Part II).	49
Fig. 23. Fuzzy membership functions for total organic carbon	50
Fig. 24. Fuzzy membership functions for porosity.	50
Fig. 25. Overtraining cartoon demonstrates that large errors are possible with 100% correlation coefficient.	51
Fig. 26. Small core dataset trains to almost 100% correlation coefficient.	51
Fig. 27. Training to almost 100% correlation coefficient produces an extremely poor blind test.	52
Fig. 28. Typical logs and core measurements through 200 ft of an algal mound reservoir.	52
Fig. 29. Shallow resistivity log fuzzy curve.	53
Fig. 30. Gamma ray log fuzzy curve.	53
Fig. 31. Trained neural network BVO; pseudo-log vs. core measurements.	54
Fig. 32. Trained neural network BVG; pseudo-log vs. core measurements.	54
Fig. 33. Trained neural network BVW; pseudo-log vs. core measurements.	55
Fig. 34. Trained neural network aspect ratio; pseudo-log vs. core measurements.	55

Fig. 35. Blind test of DD# 12 BVO neural network.	56
Fig. 36. Blind test of DD# 12 BVG neural network.	56
Fig. 37. Sum of pseudo-log BVO; log predictions vs. initial oil rate.	57
Fig. 38. Sum of pseudo-log BVG; log predictions vs. initial gas rate.	57
Fig. 39. Sum and average neural network predictions vs. actual initial oil rate.	58
Fig. 40. Sum and average neural network predictions vs. actual initial gas rate.	58
Fig. 41. Location of Brushy Canyon oil fields, fields with core utilized for petrographic and petrophysical study (hatch), and areas designated for mapping of clay content of lower Brushy Canyon sandstones. Lavender-colored well symbols indicate wells that unsuccessfully tested lower Brushy Canyon sandstones.	59
Fig. 42. Log profile through core of lower Brushy Canyon Formation, Nash Draw field.	59
Fig. 43. Plot of gamma ray intensity vs. petrographically determined clay content of lower Brushy Canyon sandstones.	60
Fig. 44. Area of lower Brushy Canyon sandstone in eastern project area (See. Fig. 41) that has at least 25 ft of sandstone with less than 10 percent clay.	60
Fig. 45. Initial oil production greater than 200 BOPD (shaded area) in eastern Brushy Canyon project area (See Fig. 41 for location of this map).	61
Fig. 46. Oil and gas fields productive from Siluro-Devonian carbonate reservoirs in southeast New Mexico.	61
Fig. 47. Stratigraphic column of Ordovician, Silurian and Devonian strata in southeastern New Mexico.	62
Fig. 48. Wells penetrating the Siluro-Devonian or pre-Devonian section in southeastern New Mexico.	63
Fig. 49. Geologic structure on Siluro-Devonian carbonates in southeastern New Mexico.	63
Fig. 50. Thickness of Siluro-Devonian strata in southeastern New Mexico.	64
Fig. 51. Actual WTI oil price vs.30-day future price. Time format.	64
Fig. 52. Actual WTI oil price vs. 30-day future price. Crossplot format.	65
Fig. 53. Actual WTI oil price vs. Euro/Yen 90-day future price.	65

Executive Summary and 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 have demonstrated beneficial results. A state-of-the-art exploration “expert” tool, relying on a computerized database and computer maps generated by neural networks, is being developed 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 FEE Tool can be beneficial in many regions of the U.S. by enabling risk reduction in oil and gas prospecting as well as decreased prospecting and development costs. In the 1998–1999 oil industry environment, many smaller exploration companies lacked 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 U.S. as reserves are depleted. The pool of experts is much reduced today. The FEE Tool will benefit a diverse group in the U.S., leading to a more efficient use of scarce funds and lower product prices for consumers.

This third of five annual reports contains a summary of progress to date, problems encountered, plans for the next year, and an assessment of the prospects for future progress. The emphasis during the April 2001 through March 2002 period was directed toward development of rules for the fuzzy system.

Introduction

In the first three years of the Fee Tool Project, an immense amount of data on the Delaware Basin has been accumulated. Data on geology, structure, production, regional information such as gravity as well as local data, such as well logs. This data, organized and cataloged into several online databases, is available for the Expert System and users as needed and as appropriate in analyzing production potential. A preliminary map of production potential for the basin has been generated and stands ready to be modified by rules defined both by human experts in exploring the Delaware basin, and by statistical rules defined by the database. We have generated a number of new and useful tools and technologies to support these efforts, including online useable interfaces for neural network analysis (PredictOnline), ranking of potential inputs using fuzzy logic (FuzzyOnline), a preliminary Expert System able to make rudimentary drilling decisions, and a web interface for accessing the databases and Expert System software.

In the next two years we will provide a finished Expert System that will run remotely from a browser on nearly any computer. The system will be able to aid in development and drilling decisions for both the Brushy Canyon and Devonian plays by providing readily accessible public information. An interactive and customizable questionnaire plus relevant analyses will produce an "Expert" opinion of a prospect in a short time to enhance the work of a human explorationist.

Progress and Discussion of Results

Computational Intelligence

Regional Data Analysis

A key component to this study is the analysis of the regional data, not only to provide baseline data to correlate with production potential, but also to provide a source of heuristic rules for the expert system. Four major categories of regional data (gravity, aeromagnetic, structure, and thickness) were selected, compiled and finalized during the previous year. Regional gravity surveys cover the entire area of the Delaware basin and have been compiled with an accuracy of a few milligals. The survey measurements are on the order of a few thousand feet apart, but sample point locations are highly variable as gravity is measured in easily benchmarked locations, such as along roadways. Gravity measures variations in density and tends to highlight large-scale regional structures in basement materials and if structure has an impact on maturation, migration or trapping of hydrocarbons in the basin useful information can be obtained.

Regional aeromagnetic data, primarily collected via over-flights, with one-mile spacing, re-gridded to 0.296 miles longitude and 0.346 miles latitude, also exists for the region. Aeromagnetic data highlights contrasts in the magnetic susceptibility between rocks and can help indicate basement blocks, large-scale faults, and possible large-scale alluvial deposits.

The structure of the lower Brushy Canyon was picked on 700 wells in the basin covering a geographically large area. Large-scale maps of structure covering the region were constructed with a kriging algorithm using this data. Structure can play more than one role in trapping and migration of hydrocarbons; two potentially helpful attributes

used in this study are structural highs and flexures that may be caused by subsurface faulting.

Finally, the wells used to compute structure were used to generate an isopach map for the Brushy Canyon in the region. Thickness may indicate areas of greater potential production and also can indicate pinch-outs and other nonstructural features that may form hydrocarbon migration pathways or traps.

A number of attributes were calculated from the four regional data types. These attributes are first and second derivatives along latitude and longitude; dip azimuth and magnitudes; and curvature azimuths and magnitudes. These values were computed to expose finer scale features in the basic data types that might be useful for correlating with a production indicator. A total of 36 maps were generated using the Zmap tool of Landmark Graphics Release 98 plus interpretation package.

Each of these maps was gridded at a scale of 1320-ft (quarter section) because that is the regulatory spacing for wells in the Brushy Canyon in New Mexico. The gridded data was exported and loaded into the project production database. Our current production database is a subset of the ONGARD database (courtesy of the SW PTTC) containing production information on all New Mexico wells. In this database we have also identified Brushy Canyon wells and using grid locations from the Zmap maps we were able to correlate producing wells with grid numbers. The gridding algorithm we selected is based on a 40 ac minimum well spacing and thus consists of points 1320 ft apart. This gridded data set essentially allows regressions to be formed using the production data as control points (training and testing) and the attribute data as variables.

Regressions formed in this manner can be used to predict production in all 60,478 40-ac bins in the basin.

There are two primary considerations in forming regressions. The first involves the quality of the data being predicted with the generated regression model. To address this, we focused initially on overall data quality and tried to filter the data to improve the likelihood of obtaining robust correlations. The second consideration deals with the choice of attributes or variables that will be used in forming the regression model. An optional consideration is the type of models used: linear models (least squares regression) or more complicated non-linear solutions such as polynomial regressions or neural networks.

For the filtering attempt an average of the first 12 producing months' Hydrocarbon Equivalent ($BO + MCF/6$) calculated at each well was chosen as the data to be modeled. Figure 1 shows a histogram of average hydrocarbon equivalent produced per month in barrels for the 2257 identified Brushy Canyon wells. The trend of the histogram is approximately an exponential decay function. A more ideal data distribution that simplifies modeling is data that follows a Gaussian distribution. The production data was conditioned with a log10 filter; Fig. 2 shows a histogram of the production indicator after log10 conversion. The bulk of the data now follows a roughly Gaussian distribution with some notable outliers on the low end. It may be desirable to remove outliers from the training data if those data are not significant to the solution. In this case, a cutoff of 50 barrels of oil per month was applied to remove the outliers. The data can then be unfiltered once the outliers are removed and used as normal in regression analysis as the filtered data is well conditioned for either linear or non-linear regression analyses.

However, the 2257-well data set was found to be unwieldy and time-consuming to run regression experiments with and instead we used a smaller subset of the wells that had been examined closely, but not filtered, for geologic data as a training data set.

For the second consideration, choice of attributes, there are a number of ways to determine which of a set of inputs (attributes) would best be used to form a regression for a particular output. Simply crossplotting each input against the output can give an indication of the quality of potential linear models, but more complicated schemes are required for nonlinear problems. The next subtopic deals with the software developed to address this concern.

Fuzzy Ranking Revisited

A fuzzy curve solution to the problem of identifying important neural network input variables from a large sparse database was discussed in the second annual report. This first attempt at fuzzy ranking required a certain expertise to use effectively; for this project we prefer to use as much automated software as possible, to avoid variances between interpretive styles. For completeness this "first stage" fuzzy ranking concept is reviewed. In a later section this initial fuzzy ranking concept is expanded and developed into useable, fully automated software.

First stage fuzzy ranking. There are a number of ways to select the best set of inputs to be used to form a regression for a particular output. Simply crossplotting each input against the output can give an indication of the quality of linear or multiple linear regression models that could be formed. Software was developed based on a single stage fuzzy-ranking algorithm to select inputs best suited for predicting the desired output. The

global algorithm statistically determines how well a particular input could resolve a particular output with respect to any number of other inputs using fuzzy curve analysis.

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

$$F_i(x) = \exp\left(-\left(\frac{x_i - x}{b}\right)^2\right) * y_i \dots\dots\dots \text{eq. 1}$$

Sample fuzzy membership functions are shown in Figs. 3 and 4. Here, $b=0.1$, since b is typically taken as about 10% of the length of the input interval of x_i . A fuzzy curve was constructed using a summation of all individual fuzzy membership functions in (x_i, y_i) , and this final curve can prioritize a set of inputs for linear or non-linear regressions. The fuzzy curve function $FC(x)$ is defined below:

$$FC(x) = \frac{\sum_{i=1}^N F_i(x) * y_i}{\sum_{i=1}^N F_i(x)} \dots\dots\dots \text{eq.2}$$

where N is the size of the data set or the total number of fuzzy membership functions. Figure 5 shows the curves for the data sets shown in Figs. 3 and 4. This simple example illustrates the ability of the fuzzy ranking approach to screen apparently random data for obscure trends such as the correlation between seismic attributes and reservoir properties.¹

Second stage fuzzy ranking. More information is needed to advance this analysis from the art of reading these fuzzy curves to a more robust and systematic elimination of

less useful inputs. Therefore, work began on a two-stage fuzzy ranking code.² The two-stage fuzzy ranking (TSFR) has two improvements: 1) Reduction of input variable space through random characterization and 2) Setting hard rules for selection of best-input variables. TSFR introduces second stage fuzzy curves, with first and second stage fuzzy surfaces that select the most important and independent input variables for modeling, while removing the input variables that show random characteristics.

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

$$P_c = \frac{P_{stage1}}{1 + P_{stage2}} \dots\dots\dots \text{eq. 3}$$

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

$$P_{c,N} = \frac{P_c}{P_{c,R}} \dots\dots\dots \text{eq. 4}$$

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

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

$$P_s = \frac{P_{stage1}}{1 + P_{stage2}} \dots\dots\dots \text{eq. 5}$$

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

$$P_{s,N} = \frac{P_s}{P_{s,R}} \dots\dots\dots \text{eq. 6}$$

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

Case study. For this study each of the 36 data and data attributes calculated and loaded into the database were analyzed using fuzzy ranking.³ It is both statistically dangerous and not computationally feasible to use all 36 attributes to form a regression relationship, so two-stage fuzzy ranking software was used to select attributes best suited for predicting production indicators. The algorithm statistically determines how well a particular input (regional data or data attribute), using fuzzy curve analysis, could resolve a particular output (production indicator) with respect to any number of other inputs. Using the newly refined fuzzy ranking tool each data attribute was ranked for its ability to predict the production potential at these well locations. Figure 6 shows the progressive elimination of non-unique and irrelevant data. The four best attributes selected were dip azimuth of gravity, second latitude derivative of thickness, longitude derivative of gravity, and longitude derivative of structure, as shown in Table I.

Table 1. Variables Selected as Optimal		
<i>RANKING</i>	<i>VARIABLE</i>	<i>PC/PCR</i>
1	6	0.8584718
2	32	0.8490006
3	2	0.8779765
4	20	0.88185640

PredictOnline Webware Revisited

For in-house neural network analysis, a web-based software, PredictOnline, was developed. 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. [*] 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.

Since PredictOnline is "web-based" it 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 our server machines. Thus, the users are not limited by the speed of their local computers, as neural network training can be computationally demanding.

Multivariate Regression

Using PredictOnline, a regression relationship was formed between these four inputs and the average first month's production for the first year at each of 520 carefully examined Delaware wells. When forming regressions, it is best to hold out a randomly selected sample of the data for testing. This data is used for testing the ability of the regression to accurately predict data not used in forming the regression. For this study a

520 well subset of the available 2434 wells in the basin was used to train the neural network. These 520 wells were selected because they were verified to have produced only from lower Brushy Canyon, and because they included dry holes in which a completion effort was made to generate production. In addition, they were distributed fairly evenly across the basin. Of these 520 wells, 466 were used to form the regression while 54 were held out for blind testing. A 4-10-10-10-1 neural network with 250 weights provided an excellent solution with a 2:1 ratio of data to weights, CC=0.90 for the training data, and CC=0.81 for the test data. For simplicity, BOPM (average barrels of oil per month expected in first year) at all 60478 40ac-bins in the basin were predicted using this model, including nearly 2000 other wells with Brushy Canyon production. Training and testing crossplots are shown in Figs. 7 and 8, respectively.

Results

The calculated BOPM for each 40-ac bin in the New Mexico portion of the Delaware basin, Brushy Canyon formation was used to generate a map (Fig. 6) to highlight potential areas of exploration. Neural network analysis necessarily results in a non-crisp solution; examination of the cross-plots in Figs. 7 and 8 demonstrate that there is some possible error in the maps, though in general the high cross-correlation means the overall fit is good. Therefore it would be inappropriate to expect that any given drilling locations would produce exactly as mapped, as there are simply too many variables and the algorithm is designed to form generalized solutions, as seen in Fig. 10. Our goal is to use this potential oil rate map as the basis for the expert system that will then quantify the risk associated with each prospect by answering questions often posed by human experts

exploring in the Brushy Canyon, as well as questions posed by statistical analyses of the data itself. The Map and associated database also provides a rich source of valuable heuristic rules for the expert system.

Expert System Rules

Basic design changes. The original design entailed the use of a single massive expert system to make decisions about a prospect's potential as a well site, Fig. 11. As we have investigated the process of designing and running expert systems, it has become apparent that a multi-tiered system, with components running in parallel, would be both more efficient and more versatile in actual usage. Figure 12 shows the current design structure for implementing and accessing the various expert systems needed to evaluate production potential. The new design is more efficient for several reasons. First, it will be faster to code the rules and the resulting code will run faster. Second, parallel expert systems will allow the user to consider only the data types they feel are most influential, and ease customization to their personal philosophies. Third, database entry, IO, from the system, will occur in numerous small packets instead of large chunks and extraneous data transfers will be reduced.

Implementation. Figure 12 shows the basic layout of the FEE Tool project. Tier 1 is a user interface that allows selection of an area or prospect of interest. Users can select the types of data they are interested in, and can review that data online with their browsers. Tier 2 in Fig. 12 represents the access of the user's browser with our online database. Advanced users can manipulate the transferred data for personal use. This data will reside on the user's computer and will not be generally available or affect the permanent database in any way. This allows the use of proprietary information with the

system. Once the data is accepted or modified, the next step is to run the appropriate expert systems using the available data to answer heuristic questions and accepting user input to answer other questions that “experts” tend to ask when evaluating Brushy Canyon prospects. In Tier 3, there are five expert systems that can be applied based on user wishes. These address *Regional Indications*, *Trap Assessment*, *Formation Assessment*, *Improved Recovery*, and *Oil Price*. Specifics and starting rules for these five systems are discussed below. Some users may elect to not factor in certain aspects, or to hard wire their own values for future oil price.

Types of rules. Two main types of rules are implemented. Heuristic rules are derived directly from our analysis of regional and local data. These rules are interpreted from the data using algorithms, such as distance relationships, and will be based on publicly available data. Heuristic rules include elements like proximity of mature source rocks, structural pinchouts, nearest producing well, and formation thickness. Expert rules will come from interviews with Delaware explorationists and will mimic questions they ask when evaluating prospects. Expert rules may include information about position on structure, porosity or permeability ranges, and production at analogous sites. In addition, heuristic rules can be replaced if the user has more detailed knowledge than is publicly available. Both types of rules may be fuzzy, as appropriate. Generally, for sites with less information heuristic rules will be more important and will provide a first estimate of production potential. For sites with sufficient specific or proprietary information Expert rules will be dominant.

Heuristic rules. One source of rules for the Fuzzy Expert Exploration Tool is statistical analyses of gridded data in our databases. Currently the regional database has

four basic data types for the Brushy Canyon: Gravity, Aeromagnetic, Structure, and Thickness. An additional eight attributes for each of those four basic types has been calculated: DX, DY, DX2, DY2, dip azimuth, dip magnitude, curvature azimuth, and curvature magnitude. Additional data include location information in latitude/longitude, oilfield X-Y coordinate systems, and a numeric grid number that also functions as a database key. Additionally, in grids that contain a Brushy Canyon well, there is relevant production information for oil, water, and gas. One factor that complicates working with the databases is the fact that the grid is not square: rather, it runs linearly from north to south increasing by integer amounts from the top of the study area to the bottom. The grid then steps over to the next “column.” Each gridpoint is separated by a physical distance of 1320 ft that corresponds to an area of 40 acres contained by four adjacent (squared) gridpoints. The gridding system looks something like this:

08	13	18
04	09	14 19 23
01	05	10 15 20 24 27
02	06	11 16 21 25 28
03	07	12 17 22 26 29

The primary use of the regional database thus far has been to organize the regional data, determine which bins contain production information, and then to calculate a “first guess” map of production potential using the data with the highest fuzzy rank to predict production. With this “first guess” map we enter a new phase, where we want to modify the probability of a successful well prediction using the “first guess” as a baseline. To do this we are using an expert system programmed in Jess and designed to

ask the same sort of questions a human explorationist would, when looking at a prospect. Is it high on structure? Where is it in the basin? Are there stratigraphic traps; is it close to other producing wells? A lot of these questions can be answered directly from the database itself, which speeds the process of examining a prospect. As a bonus, gridded heuristic answers (e.g. closeness to nearest producing well) also provide us with a way to quantify the value of these answers using fuzzy ranking.

Simple heuristic rules. The simplest questions to ask of the database are related to distances. How close is the nearest producing well? How close to structural highs or lows? Where on structure is the location? Questions about the thickness and existence of potential pay zones at the location of interest can also be answered in this manner. The following new data is being generated for the database:

- Distance to nearest oil show
- Distance to nearest production
- Distance to nearest production >500
- Distance to nearest production >1000
- Distance to nearest production >1500
- Distance to nearest production >2000 etc
- Nearness to structural high/low
- Where is the prospect on structure? Scale between local minima and maxima (0–100)
- Distance to nearest pinchout (structural trap) defined by sand thickness.
- Direction of nearest pinchout (updip downdip, across trend)
- Relative sandstone thickness (% distance between local maxima and minima)

- Sandstone thickness relative to nearest producing well (greater or lessser)
- What part of the basin does the prospect lie in? (defined as distance from key points?)
 - West edge
 - Center
 - North edge
 - East edge

Scoring of rules. Each of the subexpert systems will assign a numerical score based on the answers to individual questions. The score can come in several varieties: binary, or off/on flagging, assigned percentage values, or fuzzy-based distributions. Most rules in the subsystems will likely be assigned numerical values based on analysis of training data or fuzzy distributions based on data analysis. When combined to form the global relationship, fuzzy distributions or other functions will be applied.

Grading heuristic rules. One function of the new fuzzy ranking software is that it numerically grades, and sorts by uniqueness. We can now take suites of data, compare them to some desired quantity, and rapidly determine which are better (numerically ranked higher), and which add no new information (uniqueness test). Suites of heuristic rules can be compared to production and a numeric grade assigned based on the fuzzy rank and level of uniqueness.

Trap assessment rules. Initial trap assessment rules as programmed into the initial system are graphically illustrated in Figs. 13–15 in flowchart form.

Formation assessment rules. Initial source rock assessment rules as programmed into the initial system are graphically illustrated in Figs. 16–20 in flowchart form.

Regional indication rules. Initial regional assessment rules as programmed into the initial system are graphically illustrated in Figs. 21–22 in flowchart form.

Improved recovery rules. As the Brushy Canyon sands are a relatively new play, little information is available on improving existing production. Also, as the play is believed by some to be water-wet, waterfloods are high risk anywhere in the basin. Flooding with CO₂ is a possibility for advanced recovery techniques, but Brushy Canyon CO₂ field application data is scarce. Horizontal wells appear to be the most likely candidate technology for improved recovery.

Oil price rules. Oil price will be available in three main formats.

- *User entered:* the user enters a fixed price per barrel based on internal projections or company philosophy.
- *Predicted by neural network:* a neural network is being developed to project the price of oil using the futures markets. The status of the oil price predictive method is discussed later in the report.
- *Standard posted price for the day:* the calculation is made is the third option.

Numerical Fuzzy Rules

Introduction. Fuzzy set theory is a mathematical approach for working with imprecise data and measurements. In exploration, relevant data such as porosity is sometimes approximated or interpolated from data collected at nearby wells. This example shows how principles of fuzzy set theory are used along with expert opinions to

compute a value for a well's potential. The steps involved are: determining the input parameters and obtaining approximate numerical values, developing the linguistic values, fuzzifying the input parameters, firing the appropriate expert defined rules, and defuzzification of the output parameter. Each of these steps is discussed in detail below.

Input parameters. In this example, two variables will be used as input parameters. The variables, total organic carbon (T) and porosity (Φ) are variables for which it is sometimes difficult to get a precise value, and measurements may have to be used from nearby wells. For each of these variables, linguistic values will be defined based on the following criteria:

T=Total Organic Carbon

T: ZERO if $0 \leq T < 0.5$
T: LOW if $0.5 \leq T < 1.0$
T: MEDIUM if $1.0 \leq T < 1.5$
T: HIGH if $1.5 \leq T$

Φ =Porosity (percentage)

Φ : ZERO if $0 \leq P < 5$
 Φ : LOW if $5 \leq P < 10$
 Φ : MEDIUM if $10 \leq P < 15$
 Φ : HIGH if $15 \leq P$

For this example, 0.72 will be used as the best available value for TOC, and 13% will be used for the best available porosity. These two inputs will be used to develop a value for R, the prospect potential on a scale of 1 to 100.

Fuzzification of input parameters. The next step in the process is to “fuzzify” the input parameters. In order to do this, we will define fuzzy membership values for each of the sets; zero, low, medium and high, using a set diagram called a fuzzy membership curve that graphically defines each of the linguistic values. There are many curves that can be used in this process but the simplest is a trapezoidal graph, which we will use

here. The process is repeated for each of the input parameters. Figure 23 illustrates the process for the variable T. The value of 0.72 is plotted on the x-axis, corresponding to the following values of membership in each of the linguistic sets:

$$\begin{aligned}T(\text{Zero}) &= 0 \\T(\text{Low}) &= 56 \\T(\text{Medium}) &= 44 \\T(\text{High}) &= 0\end{aligned}$$

The process is repeated for the porosity (Figure 24), using the best value of 13%.

$$\begin{aligned}\Phi(\text{Zero}) &= 0 \\ \Phi(\text{Low}) &= 0 \\ \Phi(\text{Medium}) &= 40 \\ \Phi(\text{High}) &= 60\end{aligned}$$

Rules. Once the input parameters have been fuzzified, the linguistic sets with non-zero membership can be used to fire a set of rules determined by an expert. The rules for this example are

1. If T is zero then R is zero
2. If Φ is zero then R is zero
3. If T is low and Φ is low or medium, then R is low
4. If T is low and Φ is high then R is medium
5. If T is medium and Φ is low then R is low
6. If T is medium and Φ is medium or high, then R is medium
7. If T is high and Φ is low or medium then R is medium
8. If T is high and Φ is high then R is high

We use the non-zero memberships from the fuzzification process to determine that rules 3, 4 and 6 are applicable.

Defuzzification. The next step in the process is to determine the strength of each of the fired rules using the set theory operators min for “and” and max for “or”. Beginning with rule 3, we have T low with membership value of 56, Φ low with

membership value of 0 and Φ medium with membership value of 40. So, Φ is low *or* medium with a membership value of 40. Rule 3 is then “fired” with a strength of 40, using $\min(56,40)$ to arrive at this value.

Following this process for the two other rules, rule 4 and 6, we have rule 4 fired with a strength of 56 and rule 6 fired with a strength of 44. Rule 4 and 6, however, both result in R being medium, so we combine the two using the max operator. In the final results, R is medium with strength of 56 and low with strength of 40.

To obtain a numerical value for R, on a scale of 1 to 100, we consider the median values of 10 for low, 50 for medium and 90 for high. Then using the strengths computed above, we calculate R as follows:

$$R = 0.40*(10)+0.56*(50) = 32$$

This is a simple example of how the fuzzy set theory approach can be used to determine potential. In a more complex example, multiple input parameters may be used, and the curves used to determine the memberships may be more complex than the trapezoidal curves used here. The basic ideas are the same, however, and can be used to build the framework for computer codes that compute potential based on rules written by experts in the field.

WEB-Based Database Management System (WDMS)

A key component to the success of this project is the development of a dynamic web-accessible database for storing, managing, accessing, and analyzing data, including the development of heuristic fuzzy rules and operating the inference engine. As the data

files can be quite large, the system must be efficient, and usable by persons with varying degrees of computer literacy.

Several important advances in the web system have taken place in this semi-annual period. Key data definitions, data flows, data processing methods and user interfaces with WDMS have all progressed. A new version of PredictOnline has been coded, user management software has been developed and debugged, and a beta version of a two-stage fuzzy ranking code, FuzzyOnline, utilizing the more advanced software, has been developed.

Considerations of security for both users and potentially proprietary data, as well as for the integrity of the databases, led to the development of *Web-Based Account Management*. Users can now register, login, create and delete accounts, change user's privilege and see statistics on their personal usage of the system in a secure environment. In addition, system administrators can locally or remotely manage all accounts.

Improvements to PredictOnline include the removal of the java policy file. In the previous version, a java policy file was needed on the user's computer to allow PredictOnline to access data via applets. Upload and download functions without java policy files have been developed in JSP. Also, the neural network algorithm that we favor (conjugate gradient algorithm) was compared with a newer variant, which improved efficiency by nearly an order of magnitude for a test case, but did not blind-test very well; thus we are using our current time-tested algorithm in PredictOnline.

Java applets codes were designed to implement the user-side interface for a new FuzzyOnline software package. JSP codes were designed to implement the server-side

functions. JSP codes, called Fortran executable codes, have been successfully tested, though the overall program still requires development.

Engineering

Lower Brushy Canyon Clastics

Prior to the advent of a precise sidewall coring procedure, noncommercial Brushy Canyon completions were a problem. A new method based on artificial intelligence to predict bulk volume oil from conventional logs was developed

The methodology was developed with the Nash Draw #23 log-core information and was detailed in the Second Annual Report. In brief, it is known that operators rely on core analyses of porosity and oil saturation to make their Brushy Canyon completion decisions. Artificial intelligence tools, fuzzy logic and neural networks, were used to select the logs and to develop correlations with core measurements of the product of porosity and oil saturation (bulk volume oil), with the potential of reducing or eliminating the need for core information.

Fuzzy ranking was used to prioritize the logs on a global basis that best correlated with bulk volume oil, BVO, as measured in a 200 ft whole core from the Nash Draw Field. The density and neutron porosity logs, along with the deep and shallow resistivity logs, were then used as inputs to a neural network to develop a multivariate regression equation (trained neural network) that correlated the four logs with BVO. The Nash Draw 23 neural network was found to be robust by blind-testing it with logs and core data from 18 additional wells in the Nash Draw Field and 14 wells in the North Lee Field. The blind testing generally fit the BVO patterns evident in the limited core data.

The trained and tested Nash Draw 23 neural network was used to generate pseudo-BVO logs on 34 Lower Brushy Canyon wells in the NM Delaware basin. To reduce the risk of non-commercial completions, the 34 BVO pseudo-logs were compared to the average oil rate of each of the wells' first 12 producing months. A simple neural network architecture (3-4-1) was developed to correlate the statistical parameter's sum, average, and standard deviation of the pseudo-BVO log with the production indicator. Hence, using the sequential neural network approach, given the four logs through the Lower Brushy Canyon, a pseudo-BVO log can be generated and the statistical description of the log can be used to estimate the well's first year monthly production.

In the third year reporting period, the methodology was applied to a new Lower Brushy Canyon dataset (Poker Lake Field) and to a carbonate reservoir dataset. The Poker Lake Field dataset demonstrated the pitfalls that can occur with limited training data.

Two Poker Lake wells had core data for calculating BVO. Well 89 had 19 core measurements (samples) while well 80 had 79 samples. The well 89 dataset nicely illustrates the pitfall of insufficient data. An insufficient number of samples can result in "over-training." Over-training is illustrated by the cartoon in Fig. 25. The fictional 11-sample dataset illustrated in Fig. 25 was divided into a training set and a testing set. The polynomial best-fit line trains to a 100 R^2 using the five data points. The best-fit straight line trains to an 88 R^2 using same five data points. Using the trained polynomial curve (dashed line) to predict values could result in large errors, including negative values. The straight-line model predicts values with reasonable error throughout the limits of the dataset. A real example of this problem occurred with the Poker Lake 89 dataset. The

density and neutron porosity logs along with the deep and shallow resistivity logs were used as input to train a neural network with the limited (19 available) BVO measurements. The training was almost perfect, as shown in Fig. 26. The Poker Lake 89 neural network was used to predict the Nash Draw 23 BVO log shown in Fig. 27. Notice that the predicted high BVO values are in areas where the core values (truth) are low. Experience to date suggests that the ratio of samples to neural network weights (node tie lines) should exceed 2 to minimize the over-training problem. The more global correlations resulting from the Nash Draw 23 dataset adequately predicted the core BVO values of Poker Lake 89 even though the fields are about 10 miles apart.

Carbonate

Data is currently being assimilated on the Devonian carbonate that exists regionally throughout the Delaware basin. During the initial six months of data collection, no whole core information was uncovered. However, a consortium member was interested in applying the log analyses methodology to a carbonate system of algal mounds. The dataset serves as a test of the method in a carbonate reservoir.

Six wells with a full suite of logs (typical logs shown in Fig. 28) and core data throughout a dolomitic carbonate reservoir were made available to the project. The log suite consisted of caliper, gamma ray, photoelectric effect, laterolog deep resistivity, laterolog shallow resistivity, microspherically-focused resistivity, density porosity, and neutron porosity. Fuzzy curves were generated to rank the association between the log and the respective core-measured BVO. The caliper, shallow resistivity, and the density porosity were found to have a strong relationship with core-measured values of BVO. Examples of good and poor fuzzy ranking, defined as the maximum minus the minimum

values of the fuzzy curve plus the R^2 value of a best fit straight-line through the curve, are shown in Figs. 29 and 30.

The caliper, shallow resistivity, and the density porosity logs from the Dagger Draw 12 well were used to develop a 4-8-4-1 neural network that trained to a 90% correlation coefficient with 200 core-measured values of BVO. The training is illustrated in Fig. 31. The photo-electric factor, shallow resistivity, and the density porosity logs were used to model a BVG log defined by the core values of $1-S_w-S_o$. The goodness of the 4-6-4-1 neural network training (85% correlation coefficient) is seen in Fig. 32.

The correlations of caliper, shallow resistivity, and density porosity logs with bulk volume water (BVW) ranked highest and were used to develop a 4-6-4-1 neural network to produce a pseudo-BVW log. The goodness of the 85% correlation coefficient is seen in Fig. 33.

The same three logs were used to train a 4-6-6-1 neural network to produce an aspect ratio pseudo-log. The aspect ratio logs are defined by Φ/k as measured in the 200-core values. The goodness of the 80% correlation training is seen in Fig. 34.

There are an additional five wells with logs and cores available to train and test neural network architectures. The BVO neural network architecture and the best training correlation coefficient for each well is shown in Table 2 along with the number of training samples included in the each wells dataset. Notice that the Dagger Draw 12 well with the 4-8-4-1 neural network produced the best training correlation coefficient. The same is true for the BVG and BVW neural networks.

The Dagger Draw 12 BVO and BVG neural networks were used to generate pseudo-BVO and -BVG logs for the five wells with core information. As a blind test the

pseudo-logs were compared to the core measurements in Figs. 35 and 36. The statistical parameter's sum, average, and standard deviation for each pseudo-log were calculated and crossplotted with the average of the first producing year's monthly oil and gas production. The crossplots are shown in Figs. 37 and 38.

Lastly, two of the statistical parameters were used to generate correlations to predict oil, gas, and water production. The dataset is limited to the six wells previously discussed, so a simple neural network architecture (2-1-1) with Sum and Avg as the inputs was used to develop the production correlations shown in Figs. 39 and 40 used to predict oil and gas rates. Water production predictions were poor.

Geology

During the reporting period from October 2001 through April 2002, geologic data acquisition and analysis was completed on the Brushy Canyon Formation and work progressed on the Siluro-Devonian carbonates.

Lower Brushy Canyon

The following tasks were accomplished during the reporting period on the Brushy Canyon petroleum system:

1. *Logging and stratigraphic description of productive and non-productive cored sandstone intervals of the lower Brushy Canyon Formation.* Previous work on this project determined that some thick sections of Brushy Canyon sandstones in trap configuration are productive while others are non-productive despite having been extensively tested. One hypothesis for this phenomenon is that grain size and/or pore

size may be larger in the productive sandstones, thereby resulting in lower capillary entry pressures that permitted entry of migrating hydrocarbons while the finer sandstones, with higher capillary entry pressures effectively acted as seals for the hydrocarbons. Productive and nonproductive sandstones from core in the Nash Draw field shown in Fig. 41 were described macroscopically in order to provide baseline data on productive and nonproductive sandstones.

2. *Correlation of petrographic properties of sandstones with logs properties in the lower Brushy Canyon Formation.* Several petrographic properties related to capillary pressure properties and oil migration and emplacement were analyzed and related to log properties. The purpose of this work was to determine if factors affecting and/or limiting hydrocarbon migration/emplacement could be mapped in order to establish areal limits to potentially productive thick accumulations of sandstone. Such properties as percent macroporosity, percent microporosity, grain size, and clay content were quantitatively related to log properties (Figs. 42, 43). Similar work has proven successful in the Cherry Canyon Formation.⁴ We showed that the gamma-ray intensity of lower Brushy Canyon sandstones related linearly to clay content (Fig. 43). The difference in porosity given by the neutron logs and the density logs ($\phi_n - \phi_d$) also showed correlation to gamma ray intensity. Where the density logs read a higher porosity than the neutron log (gas effect = hydrocarbon-filled reservoir), the gamma-ray intensity is low, signifying low clay content. Where the neutron log gives a significantly higher porosity than the density log (no gas effect = water-filled reservoir), the sandstones have a high gamma-ray reading, and therefore high clay content. Petrographic examination of Brushy Canyon sandstones revealed that the

clays are mostly present in pores and will act to inhibit migration and emplacement of oil in the reservoir. During the next quarter we will verify these results on additional core and will map clay content via gamma ray logs in productive and nonproductive porous sandstones that are in trap configuration. The additional cores come from the Poker Lake field, from the southern part of the basin and from a nonproductive well adjacent to the Parallel field in the northern part of the basin (Fig. 41).

3. Detailed subsurface mapping of depositional elements of lower Brushy Canyon sandstones in areas proven to be productive and areas proven to be nonproductive.

These maps show sandstones in trap configuration and do not seem to be able to distinguish between thick productive sands in trap configuration and thick, non-productive sands in trap configuration. Therefore, steps 1 and 2 above were undertaken to acquire additional data that should be relevant to oil prediction.

Controls on Hydrocarbon Production

Additional tasks were undertaken in order to refine analysis of controls of hydrocarbon production in the lower Brushy Canyon. Work was concentrated on two areas (Fig. 41), an eastern area with abundant thick sandstones that are oil-productive and a western area with abundant, thick sandstones that have been extensively tested but are barren of oil.

Detailed core descriptions were completed and thin sections were taken of both productive and nonproductive sandstones. Core descriptions indicated that a majority of the lower Brushy Canyon Formation consists of finely interbedded siltstone and shale turbidites. Other lithologies include a fine siltstone that is gray in color and has high clay content. Tan siltstone also occurs, which has low clay content. The tan siltstones

correspond to the productive intervals within the wells (Fig. 42). Detailed petrographic descriptions confirm observations seen in the cores. Thin sections revealed that the nonproductive siltstones have a very high clay and calcite cement constituent. The high clay and calcite cement reduces the size and amount of porosity, therefore increasing pore entry pressures and reducing the ability of hydrocarbons to enter the pore space. Graphing point count data against well log values revealed a direct relationship between the gamma ray curve and the actual clay content within the siltstones (Fig. 43). As the clay content of the siltstones increased the gamma ray value increased. This relationship was then applied to digital well log curves and a value for average clay content was calculated. This enabled the mapping of average clay content throughout the study area (Fig. 44). These maps were then compared to production maps, isopach maps, isolith maps, and maps of average gamma ray value. All of these maps indicate a strong relationship with production (Fig. 45). It appears that the most productive regions correspond to regions with low clay content, low gamma ray and thick sands.

Siluro-Devonian Carbonates

Devonian and Siluro-Devonian carbonates produce from numerous oil and gas fields in southeastern New Mexico (Fig. 46). The 122 Siluro-Devonian fields in southeast New Mexico had produced a cumulative 443 MMBO by 1995⁵, 10 percent of the oil produced from southeast New Mexico. Production is from a number of zones within the Silurian and Devonian sections (Fig. 47). A variety of mechanisms form traps, most notably anticlines, faulted anticlines, and subunconformity pinchouts.⁶

Geologic data acquisition continued on the Siluro-Devonian carbonates during the reporting period. One thousand six hundred wells in southeast New Mexico that have

penetrated the Devonian (Fig. 48) had been identified during the previous reporting period and entered into a database; longitude and latitude have been calculated for these wells. During the present reporting period, a regional network of cross sections was constructed in order to establish correlation control and to provide quality assurance of data. Logs from a total of 220 wells were correlated throughout southeast New Mexico. Correlated wells were selected from all productive Siluro-Devonian fields and non-productive areas as well in order to provide a dataset that fully contrasts productive areas with non-productive areas. In each well, the top of the Siluro-Devonian carbonate section was correlated in addition to the top of the Fusselman Formation and the top of the Montoya Formation. Several prominent marker beds within the Siluro-Devonian section were correlated as well.

Similar to our earlier work on the Brushy Canyon Formation, we used our correlated data to produce geologic structure maps (Fig. 49), and isopach (thickness) maps (Fig. 50) of Siluro-Devonian carbonate strata. We will relate these to production/nonproduction in both visual and artificial intelligence settings. The goal is to use our artificial intelligence system to predict trap configurations in Devonian strata.

We will also construct regional maps of source rocks. The chief source rock unit is the Devonian Woodford Shale, which directly overlies the Devonian carbonates (Fig. 47). The Silurian Simpson shales are source rocks that underlie the Siluro-Devonian carbonate section. The Woodford is thought to be the chief source unit for lower Paleozoic reservoirs in the Permian Basin and the Simpson is a secondary but still important source unit.^{7,8} As with the Brushy Canyon, we will map regional distributions of source rock maturity and quality and relate these to oil and gas distribution. With the

depth of the Woodford varying from less than 7000 ft in the northern part of the basin in Chaves County to more than 15,000 ft in the southern part in Lea and Eddy Counties, we expect to encounter thermal maturity variations across the oil window/gas window boundary that will relate to the distribution of oil reservoirs and gas reservoirs. The map of Siluro-Devonian oil and gas fields (Fig. 46) indicates that most gas fields are located in the more deeply buried southern parts of the basin where thermal maturity of source rocks should be higher. These relationships will be quantified for use in the artificial intelligence system. Thermal maturity variations may also help in the prediction of gas-oil ratios and therefore relate to aspects of recovery efficiency.

Oil Price Forecast

Future oil price is the fundamental parameter in assessing the risk of any oil or gas production venture. One aggregate of speculation by experts on the purchasing and selling of crude oil is the Goldman Sachs future's price of oil. These experts try to reasonably predict the price of a barrel of oil a month in advance. Shown in Fig. 51 is the 30-day average of these experts' predictions preceding the strike date. The average is compared to the Navajo refinery posted price of West Texas Intermediate on the strike data. Notice that from June 1996 through October 1997 the experts did an excellent job. However, from November 1997 to March 1999 the error is approximately 33% or about \$5/bbl more than the posted price. Also from May 1999 to December 2000 the error is about 25% or \$7/bbl less than the posted price. The predictions presented in Fig. 51 are crossplotted in Fig. 52. Notice that R^2 is 0.72, yielding an 85% overall correlation coefficient. For the FeeTool to successfully predict economic success a better model of future oil price is needed.

The future price of commodities other than oil is also the subject of speculation by experts. The 90-day future price of the Euro vs. the Yen was examined using crossplots similar to the oil futures versus posted price, the hope being that Euro/Yen traders might indirectly have insight to the future price of oil. Figure 53 shows the crossplot of the WT Posted Price of oil vs. the 90-day average of the Euro/Yen. Based on the R^2 of 0.96 (98% correlation coefficient) there appears to be a strong relationship between the Euro/Yen future price and the posted price of WT intermediate, at least for this small dataset.

The relationships between the three-month future price of the Euro/Pound, 30-year Treasuries, 10-year Treasuries, and the Future Dow Jones with the posted price of WT intermediate oil were also examined in a similar manner. The results based on the goodness of the best-fit line are summarized in Table II.

Table 2. Goodness of the Relationship between Commodity Future Price and Oil Price	
Commodity	Correlation Coefficient, %
Goldman Sachs future price of oil	85
Euro/Yen	98
Euro/Pound	96
30-year Treasuries	94
10-year Treasuries	94
Future Dow Jones	93

The results of this preliminary analysis, albeit with a small dataset, suggests that the future price of some commodities other than oil will be more useful when predicting oil price. Assuming that the correlation coefficient describes experts' ability to forecast the future, it appears that the futures price of the Euro/Yen and Euro/Pound are closely tied to the price of oil. This is not as surprising as one might think, as both the Euro and Yen (for example) are currencies for nations and regions which on a whole are dependent on oil imports to fuel their industries, and hence their economies.

Technology Transfer

During this twelve-month period (April 2001–March. 2002) the following ten papers or presentations were made to disseminate the results of the project:

1. Weiss, W.: "Mining Regulatory Files with Artificial Intelligence to Predict Waterflood Recovery," presented to the City Different Petroleum Club, Santa Fe, NM, May 3, 2001.
2. Weiss, W. W., Stubbs, B.A. Balch, R.S.: "Estimating Bulk Volume Oil in Thin-Bedded Turbidites," SPE 70041 Paper presented at the SPE Permian Basin Oil & Gas Recovery Conference, Midland, Texas, 14-17, May 2001.
3. Weiss, W.: "Risk Reduction with a Fuzzy Expert Exploration Tool," SPE 70054, Poster presented at the SPE Permian Basin Oil & Gas Recovery Conference, Midland, Texas 14-17, May 2001.
4. Weiss, W. W., Wo, S., Weiss, J.: "Data Mining at a Regulatory Agency to Forecast Waterflood Recovery," SPE 71057 Presented at the 1999 SPE Rocky Mountain Technical Conference, Keystone CO, 21-23 May 2000.
5. Weiss, W.W.: "Neural Networks and Geostatistics Used to Characterize a Reservoir in a Coastal Dune Environment," Short Course Presented at the Annual Petroleum Engineering Summer School, Workshop No. 8 *Geomodeling In Exploration and Production of Oil and Gas*, Dubrovnik, Croatia, 4-8 June, 2001.
6. Additional presentations were made at PTTC seminars in Jackson MS, Evansville IN, Washington PA, and Tyler TX demonstrating the application of fuzzy ranking and neural network to solve oilfield problems.
7. Broadhead, R.F., and Justman, H.A., 2002, Regional controls on oil accumulations and regional sand trends, lower Brushy Canyon Formation, southeast New Mexico: presented to Roswell Geological Society, February 11, 2002.

Problems Encountered

The acquisition of regional seismic lines continues to be a problem due to the value of the data. Local datasets are available such as those from the DOE-funded Nash Draw project. The processed data from this 3D data set was used to develop new methods of interpreting the distribution of thickness, porosity, water saturation and depth throughout the survey area. The methodology can be applied throughout the Delaware Basin.

Coding of the required web interface algorithms has been an ongoing problem. New graduate students in place six months ago at the time of the semi-annual have

departed. Currently, REACT group professionals are educating themselves in JAVA and are continuing development in what should be a more sustainable fashion.

Next Year's Tasks

Continue Expert System Development

Currently the Expert System has been programmed to use some basic rules in several different categories. About 100 rules are in the current version and this preliminary system can be run on a laptop computer and will be of great assistance in giving human explorationists a feel for the system. Work is ongoing in numerically grading questions.

Log Analysis

As Devonian log information becomes available it will be correlated with core data to develop pseudo-logs. Production will then be correlated with the pseudo-log statistics.

Geology

September of 2001 marked the halfway point in the project schedule at which time the geologic focus of the project shifts to the Devonian carbonate. During the next 12 months, we expect to:

1. Finish correlating significant Siluro-Devonian marker beds in wells throughout the Permian Basin in southeastern New Mexico.
2. Produce appropriate computer-contoured structure and thickness maps of variables that should be relevant to the accumulation and entrapment of oil (for example,

structure on top of the Devonian carbonates, thickness of significant productive, porous units).

3. Identify which Siluro-Devonian strata produce from which fields.
4. Select appropriate wells for source rock analysis of the Woodford Shale and, to a lesser extent, the Simpson shales and perform source rock analyses with emphasis on thermal maturity analyses to differentiate gas vs. oil regimes.

Oil Price Forecast

The commodity variables having the strongest relationship with the future price of oil will be examined using the two-stage fuzzy ranking algorithm. A neural network architecture will be developed using the top ranked variables as inputs to correlate the future price of commodities with oil price.

References

1. Balch, R.S., Stubbs, B.S., Weiss, W.W., and Wo, S.: "Using Artificial Intelligence to Correlate Multiple Seismic Attributes to Reservoir Properties," paper SPE 56733 presented at the 1999 SPE Annual Technical Conference, Houston, Oct. 3–6.
2. Lin, Y., Cunningham, G.A., Coggeshall, S.V., and Jones, R.D.: "Nonlinear System Input Structure Identification: Two Stage Fuzzy Curves and Surfaces," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* (1998) **28**, No. 5, 678–684.
3. Lin, Y., and Cunningham, G.A.: "A New Approach to Fuzzy-Neural System Modeling," *IEEE Transactions on Fuzzy Systems* (1995) **3**, No 2, 190-198.
4. Spain, D.R.: "Petrophysical Evaluation of a Slope Fan/Basin-Floor Fan Complex: Cherry Canyon Formation, Ward County, Texas," *AAPG Bull.*, (1992) **76**, 805–827.
5. Broadhead, R.F., and Speer, S.W.: "Oil and Gas Fields in the New Mexico Part of the Permian Basin," in: *A Symposium of the Oil and Gas Fields of Southeastern New Mexico*, (1995 supplement) Roswell Geological Society, 32-49.
6. Speer, S.W.: PP-6; Siluro-Devonian, in: *Atlas of Major Rocky Mountain Gas Reservoirs*: New Mexico Bureau of Mines and Mineral Resources (1993) 163.
7. Dow, W.D., Talukdar, S.C., and Harmon, L.: "Exploration Applications of Geochemistry in the Midland Basin, Texas" (abstract), *AAPG Bull.*, (1990) **74**, 644-645.
8. Katz, B.J., Dawson, W.C., Robison, V.D., and Elrod, L.W.: "Petroleum System of the Central Basin Platform, West Texas, U.S.A.," in: Magoon, L.B., and Dow, W.G., eds., *The Petroleum System from Source to Trap*, American Association of Petroleum Geologists, Memoir 60 (1994) 453-461.
9. Read, A., Broadhead, R., Lopez, A., Fleming, E., and Watrous, J.: "New Mexico Oil and Gas Pool Maps," New Mexico Bureau of Mines and Mineral Resources (2000) Circular 209.
10. New Mexico Bureau of Mines and Mineral Resources: *Atlas of Major Rocky Mountain Gas Reservoirs*, Socorro, NM (1993)

Fig. 1. Histogram of average hydrocarbon equivalent. Value is average monthly production over the first year of production for Brushy Canyon wells.

Fig. 2. Log filtered version of histogram from Fig. 1. Notice that the data is much better behaved and follows a rough gaussian curve. It is easier to see outliers with the data in this format.

Fig. 3. Sample fuzzy membership functions drawn for two points of a 100-point random data distribution.

Fig. 4. Sample fuzzy membership functions drawn for two points of a 100-point random data distribution with a trend added.

Fig. 5. Fuzzy curves for the data sets in Figs. 3 and 4. Fuzzy curves are drawn by doing a weighted sum of all the fuzzy membership functions.

a) Input: gravity dx-grav dy-grav ...36 production

```
0.4862760 0.1512304 0.2664834 3.0988186
0.4851089 0.2415437 0.2647132 3.0725581
0.4843222 0.1666667 0.2183791 3.1368229
0.4832898 0.1777388 0.2751238 3.1901707
0.4798365 0.2047192 0.2633718 3.4063671
0.4799727 0.2045356 0.1696531 3.0689730
0.4832426 0.1794870 0.2642204 3.3659540
0.3837681 0.1541102 0.3653276 3.2306246
0.3824985 0.1532802 0.2636087 3.0127803
0.4995902 0.2215138 0.2701701 3.1323701
```

c) Results

```
Ranking= 1 variable= 6 Pc/Pc1=8.8564718
Ranking= 2 variable= 32 Pc/Pc1=8.8493086
Ranking= 3 variable= 2 Pc/Pc1=8.8773365
Ranking= 4 variable= 20 Pc/Pc1=8.8913564
```

b) Processing

```
1 0.1512304 0.2664834 3.0988186
2 0.2415437 0.2647132 3.0725581
3 0.1666667 0.2183791 3.1368229
4 0.1777388 0.2751238 3.1901707
5 0.2047192 0.2633718 3.4063671
6 0.2045356 0.1696531 3.0689730
7 0.1794870 0.2642204 3.3659540
8 0.1541102 0.3653276 3.2306246
9 0.1532802 0.2636087 3.0127803
10 0.2215138 0.2701701 3.1323701
11 0.1512304 0.2664834 3.0988186
12 0.2415437 0.2647132 3.0725581
13 0.1666667 0.2183791 3.1368229
14 0.1777388 0.2751238 3.1901707
15 0.2047192 0.2633718 3.4063671
16 0.2045356 0.1696531 3.0689730
17 0.1794870 0.2642204 3.3659540
18 0.1541102 0.3653276 3.2306246
19 0.1532802 0.2636087 3.0127803
20 0.2215138 0.2701701 3.1323701
21 0.1512304 0.2664834 3.0988186
22 0.2415437 0.2647132 3.0725581
23 0.1666667 0.2183791 3.1368229
24 0.1777388 0.2751238 3.1901707
25 0.2047192 0.2633718 3.4063671
26 0.2045356 0.1696531 3.0689730
27 0.1794870 0.2642204 3.3659540
28 0.1541102 0.3653276 3.2306246
29 0.1532802 0.2636087 3.0127803
30 0.2215138 0.2701701 3.1323701
31 0.1512304 0.2664834 3.0988186
32 0.2415437 0.2647132 3.0725581
33 0.1666667 0.2183791 3.1368229
34 0.1777388 0.2751238 3.1901707
35 0.2047192 0.2633718 3.4063671
36 0.2045356 0.1696531 3.0689730
```

It 1: 22 non-random inputs
(14 random inputs)

```
32 0.8488006
5 0.8787322
29 0.8850168
28 0.8851412
35 0.8851412
20 0.8922889
30 0.9035545
27 0.9038997
12 0.9021485
26 0.9053842
33 0.9061842
24 0.9061842
15 0.9351038
2 0.9412536
17 0.9544592
14 0.9720233
7 0.9788513
37 1.0000000
```

It 2: 17 non-random inputs
independent
of
#6

```
2 0.8779765
5 0.8846758
15 0.9038422
28 0.9118645
12 0.9138505
38 0.9246746
17 0.9278167
36 0.9339433
26 0.9339433
14 0.9526650
23 0.9566760
7 0.9612614
29 0.9667241
27 0.9724635
37 1.0000000
```

It 3: 14 non-random inputs
independent
of
#6 and #32

```
28 0.8818864
29 0.8906134
26 0.9072476
12 0.9172460
15 0.9188021
35 0.9210323
26 0.9210323
38 0.9266816
5 0.9465737
14 0.9572384
27 0.9587365
7 0.9736186
17 0.9829877
37 1.0000000
```

It 4: 13 non-random inputs
independent of
#6, #32, and #2

```
1 0.1512304 0.2664834 3.0988186
2 0.2415437 0.2647132 3.0725581
3 0.1666667 0.2183791 3.1368229
4 0.1777388 0.2751238 3.1901707
5 0.2047192 0.2633718 3.4063671
6 0.2045356 0.1696531 3.0689730
7 0.1794870 0.2642204 3.3659540
8 0.1541102 0.3653276 3.2306246
9 0.1532802 0.2636087 3.0127803
10 0.2215138 0.2701701 3.1323701
11 0.1512304 0.2664834 3.0988186
12 0.2415437 0.2647132 3.0725581
13 0.1666667 0.2183791 3.1368229
```

It 5: no non-random independent variables

Fig. 6. Operation of second stage fuzzy ranking program FuzzyOnline. A) is a sample input file. In this case we wish to correlate to each of 36 data found in the database for 520 wells and the average monthly production we wish to correlate to. B) shows a progressive reduction in the number of possible inputs (initially 36) through several iterations. In each iteration the best fitting variable is identified, then the rest of the data is studied for uniqueness with respect to that best variable. In this case five iterations identified the four strongest and unique inputs. C) The generated output file shows the strength and ID number of the best inputs for regression analyses

Fig. 7. Crossplot of neural network training using the inputs identified via FuzzyOnline.

Fig. 8. Crossplot of testing data for the trained neural network of Fig. 7.

Fig. 9. Map of predicted production potential based on the trained and tested neural network regression.

Fig. 10. This figure demonstrates the benefits of neural network regressions for complex problems where forecasts need to be made. Assuming only the five points enclosed by diamonds are known to start, it is quite easy to fit a polynomial that exactly fits the training data but poorly fits the true data distribution. Neural networks can bypass this problem if a sufficient sample exists for training by creating a generalized solution.

Fig. 11. The original schematic for the fuzzy expert system shell.

Fig. 12. More complicated system, which breaks the analysis into several separate categories to simplify calculations and customization.

Fig. 13. Initial rules for trap assessment (Part I).

Fig. 14. Initial rules for trap assessment (Part II).

Fig. 15. Initial rules for trap assessment (Part III).

Fig. 16. Initial rules for source rock assessment (Part I).

Fig. 17. Initial rules for source rock assessment (Part II).

Fig. 18. Initial rules for source rock assessment (Part III).

Fig. 19. Initial rules for source rock assessment (Part IV).

Fig. 20. Initial rules for source rock assessment (Part V).

Fig. 21. Initial rules for regional assessment (Part I).

Fig. 22. Initial rules for regional assessment (Part II).

Fig. 23. Fuzzy membership functions for total organic carbon.

Fig. 24. Fuzzy membership functions for porosity.

Fig. 25. Overtraining cartoon demonstrates that large errors are possible with 100% correlation coefficient.

Fig. 26. Small core dataset trains to almost 100% correlation coefficient.

Fig. 27. Training to almost 100% correlation coefficient produces an extremely poor blind test.

Fig. 28. Typical logs and core measurements through 200 ft of an algal mound reservoir.

Fig. 29. Shallow resistivity log fuzzy curve.

Fig. 30. Gamma ray log fuzzy curve.

Fig. 31. Trained neural network BVO; pseudo-log vs. core measurements.

Fig. 32. Trained neural network BVG; pseudo-log vs. core measurements.

Fig. 33. Trained neural network BVW; pseudo-log vs. core measurements.

Fig. 34. Trained neural network aspect ratio; pseudo-log vs. core measurements.

Fig. 35. Blind test of DD# 12 BVO neural network.

Fig. 36. Blind test of DD# 12 BVG neural network.

Fig. 37. Sum of pseudo-log BVO; log predictions vs. initial oil rate.

Fig. 38. Sum of pseudo-log BVG; log predictions vs. initial gas rate.

Fig. 39. Sum and average neural network predictions vs. actual initial oil rate.

Fig. 40. Sum and average neural network predictions vs. actual initial gas rate.

Fig. 43. Plot of gamma ray intensity vs. petrographically determined clay content of lower Brushy Canyon sandstones.

Fig. 44. Area of lower Brushy Canyon sandstone in eastern project area (See. Fig. 41) that has at least 25 ft of sandstone with less than 10 percent clay.

Fig. 45. Initial oil production greater than 200 BOPD (shaded area) in eastern Brushy Canyon project area (See Fig. 41 for location of this map).

Fig. 46. Oil and gas fields productive from Siluro-Devonian carbonate reservoirs in southeast New Mexico.

Fig. 47. Stratigraphic column of Ordovician, Silurian and Devonian strata in southeastern New Mexico.

Fig. 48. Wells penetrating the Siluro-Devonian or pre-Devonian section in southeastern New Mexico.

Fig. 49. Geologic structure on Siluro-Devonian carbonates in southeastern New Mexico.

Fig. 50. Thickness of Siluro-Devonian strata in southeastern New Mexico.

Fig. 51. Actual WTI oil price vs.30-day future price. Time format.

Fig. 52. Actual WTI oil price vs. 30-day future price. Crossplot format.

Fig. 53. Actual WTI oil price vs. Euro/Yen 90-day future price.