SPE 30202

Predicting Production Using a Neural Network (Artificial Intelligence Beats Human Intelligence)

Robert J. Boomer, Texaco Exploration and Production Inc.

SPE Member

This paper was prepared for presentation at the Petroleum Computer Conference held in Houston, TX, U.S.A., 11-14 June 1995.

Abstract

Petroleum professionals, that is engineers and geoscientist, routinely make recommendations to drill wells. This process requires the generation of a production profile, oil producing rate versus time, which is a key component in the final economic decision of whether or not to drill the well. The success of the drilling program, and possibly the professional's career, rests, to a high degree, upon the accuracy of this predicted production profile. The more accurate this prediction the more certain the economic value of the recommended drilling location. This paper looks at this predictive process as used in the Vacuum Field of New Mexico. It explores the predictive methods used in the past and reports on their accuracy. A new method, using artificial intelligence, is presented which improves upon the accuracy of this process.

Introduction

Acting upon a recommendation to drill a well, which usually involves hundreds of thousands if not millions of dollars, is a major decision. The final decision of whether or not to drill a well is based, in part, upon the economic evaluation of the anticipated performance of the well. A key component of this economic analysis is the predicted production profile, oil producing rate versus time. This forecasting of the production performance of a new well is one of the most important functions a petroleum professional performs. Both underestimating, predicting a producing rate that is lower than actual, and overestimating, predicting an oil producing rate that is greater than actual, the producing performance of the well is common. Both have a major impact upon the actual economic value achieved by drilling, or not drilling, the recommended well. At the extremes, either overestimating or underestimating oil or gas producing rates results in a potentially disastrous economic decision. Overestimating the production profile will result in the drilling of a well that is not economical while underestimating the production performance causes a well to not be drilled which would have been economical. Usually the result of an inaccurate production forecast is the drilling of a well whose economic performance does not yield the optimum economics for the available drilling investment.

References at the end of the paper.

But this is all part of the risk and uncertainty we face in the oil industry. The industry drilled wells in the past and will continue to drill wells in the future. We accept this uncertainty. However, it raises the question of just how accurate have the forecasts from the past been and can this accuracy be improved. For with improved forecasts come greater economic certainty and greater profitability.

Infill Development Drilling Study Area

A category of significant interest in the Permian Basin of West Texas and Southeast New Mexico is infill development drilling. Infill development drilling is defined as drilling within an existing secondary recovery project to reduce well spacing and consequently the waterflood injection pattern.

San Andres reservoirs in the Permian Basin have produced over 7.7 billion barrels of oil, the majority from secondary recovery projects. This is 46 percent of the total production from the entire basin. This large cumulative combined with low recovery, 30 percent of the oil-in-place, makes this reservoir an excellent candidate for improved development.[1] Since infill development drilling in existing waterflood projects provides many of the drilling prospects for the Permian Basin, it was selected as an area to study the overall accuracy of production performance predictions. Specifically, the Vacuum Field of New Mexico is a good area to study since it is a typical San Andres field and the individual production performance of every well in the field is readily available.

The Vacuum Field, see Figure 1, is located in the southeast corner of New Mexico in Lea County at a distance of approximately 120 miles northwest of the city of Midland, Texas.

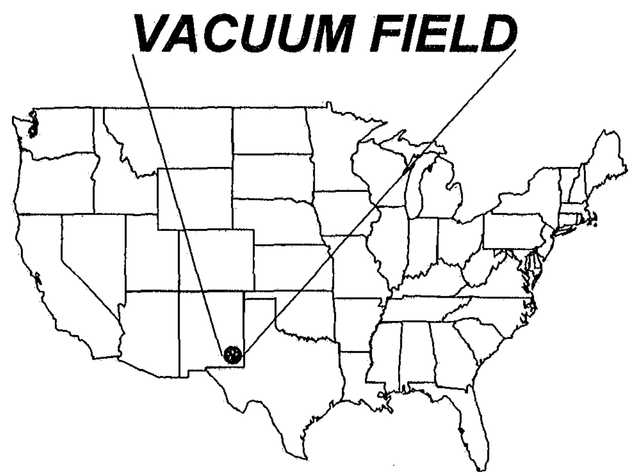

Figure 1 Field Location

The field lies at the north end of the Delaware Basin and Central Basin Platform along a productive east-west-trending shelf area of the Permian Basin. Production is primarily from the Permian Guadalupian age San Andres formation at a depth of 4400 feet. The San Andres is composed of cyclical evaporites and carbonates recording the many transgressings and regressings of sea level occurring around 260 million years ago in a climate very similar to the present day Persian Gulf.[2]

Nine secondary recovery projects constitute the majority of the acreage, wells, and production from the Vacuum Field. Figure 2 shows the outlines of these secondary recovery projects. These projects cover approximately 32 square miles. The field was discovered in 1929 but field development did not start until a pipeline to the field was laid in 1937.

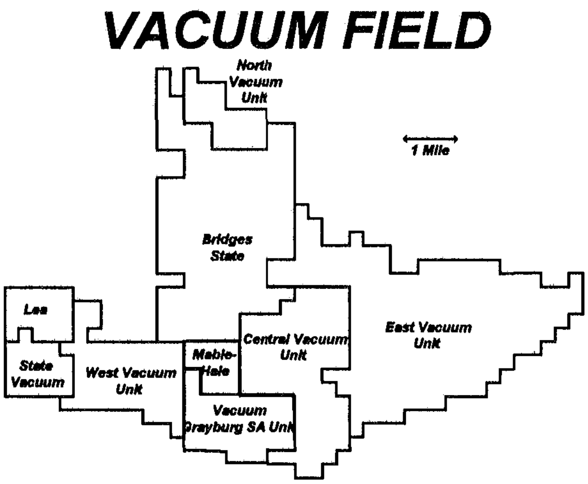

Figure 2 Secondary Recovery Projects

Figure 3 is a plot of monthly oil, water, water injection and well count from 1970 to present. Also shown is oil cut (% Oil), oil volume as a percentage of total produced fluids (oil plus water), and the injection-withdrawal ratio (I/W), water injection rate divided by the total fluid volume. The plot shows that over 300 infill development wells have been drilled in the field since 1970. Field development covers the start of water injection in the 1970's and the down spacing of injection patterns from 80-acre 5-spots patterns to 40-acre 5-spots and to even some 20-acre 5-spots.

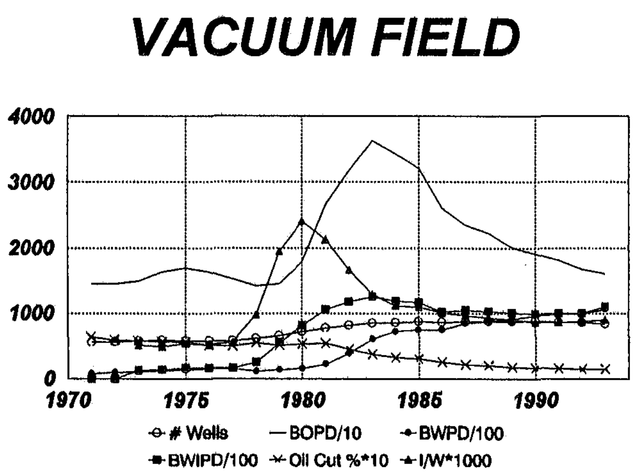

Figure 3 Field Performance

Past Methods of Making Predictions

The Professionals have at their disposal a myriad of tools to predict the production profile of an infill development well. Either a deterministic and/or a statistical method is typically used.

The typical deterministic method for predicting a production profile is a straightforward process. The Professional reviews any geological maps, cross-sections, and existing well logs of the prospect area to determine overall reservoir quality. Next, the Professional reviews the offset well's current and past production performance. Past production performance includes not just the rates and volumes of oil production but also the rates and volumes of water production and water injection. They may also review the production profiles of wells from past drilling programs. It is characteristic of infill drilling for the operator to drill wells in programs with each program including from 2 to 20 wells. In the end the Professionals use this knowledge to generate a production profile for a specific drilling location. Sometimes a numerical reservoir simulator is used to predict a production profile.

Statistical methods consider the actual variability of past drilling programs to generate a single production profile for all proposed drilling locations, knowing that averaging the production performance from every individual well in the drilling program will come close to this statistically generated profile. The actual production performance from each well in the most recent drilling program is first normalized. That is the oil producing rates are normalized with respect to initial completion date not calendar date, production for month number one is the production for the first month after initial completion and not for the month of January. Next these normalized rates are averaged on this monthly basis for every well in that particular drilling program. This resulting production profile is used for the next drilling program. Alternatively, instead of the most recent drilling program the most recent program of "like" wells is used. "Like" in the fashion of 40-acre wells versus 20-acre wells or geologically similar locations, top of structure versus down dip.

All of these approaches are successful. Successful in that we have used them for decades to recommend wells to be drilled and then to have them drilled. Still, the actual production profile does not match the predicted production profile. Many factors that are not associated with the drilling program itself will influence the long-term forecast, such as changing injection rates or altering injection pattern. If we constrain the time frame over which we compare the actual versus predicted production profile, we might get better accuracy. Monthly production for the first year is a good time frame to use since the profitability of a new well is heavily weighted by the first year's production profile. Accurate prediction of the rate of production for just the first year would improve the certainty of achieving the predicted economical performance of any drilling program. Just how accurate have these approaches been at forecasting oil production over the first twelve months of a well's life?

Prediction Accuracy

A review of the first year's production predictions versus the well's actual performance is enlightening. In the Vacuum Field of New Mexico over 250 producing wells have been drilled since 1970. The Professionals' forecasts were found by searching the well files and correspondence files where Texaco is the operator or where Texaco has a non-operating working interest for the original Authority for Expenditure (AFE) to drill the well. The AFE typically includes the predicted production profile for the proposed location. A sampling of 25 of the 250 wells was used to determine the overall accuracy of the predictions. The production profiles were not biased by a small group of Professionals making forecasts nor were they biased by any particular year's drilling results. Error calculations were based upon the difference between actual and predicted cumulative production, expressed as barrels of oil per day (BOPD), at the one-, three-, six-, and twelve-month time intervals.

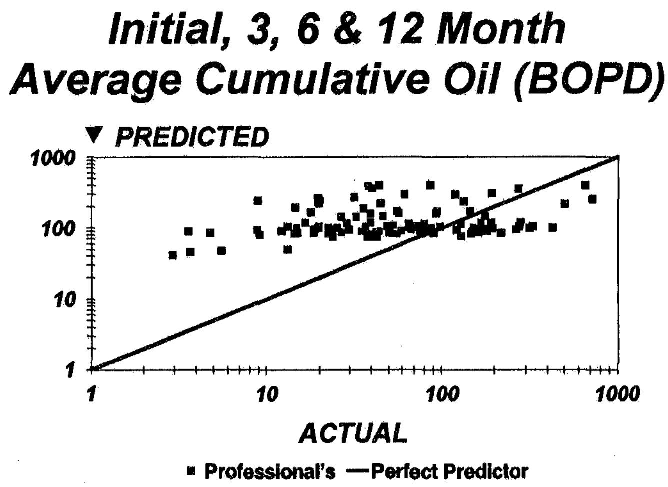

Figure 4 Professional's Performance

Figure 4 shows how accurate the forecasts are by plotting the actual performance versus the predicted performance. This type of plot is called a Confusion Matrix.[3] Included on this figure is a straight line labeled Perfect Predictor. If you had a perfect predictor, then all the points would fall on this line where the actual and predicted values are equal. These Professional predictions have a mean error of 39 BOPD and a mean absolute error of 99 BOPD. The mean absolute error is an excellent measure of the accuracy of the Professional's forecast. For any well their forecast will average within 99 BOPD of the actual rate.

Further statistical analysis reveals that the Professionals predict within ±20BOPD just 8 percent of the time and within ±10BOPD only 5 percent of the time. The Professionals overestimated production 75 percent of the time. It is obvious that the Professionals are not perfect predictors.

One explanation for this poor showing is because of the number of variables considered in making the predictions, variables such as oil rate, cumulative oil, injection rate, and cumulative injection. The Professionals limit the number of offset wells and the type of data considered from each well in their prediction process. It is very difficult for humans to consider more than a handful of variables at any one time. If a methodology could consider more offset wells and parameters, such as water production, injection withdrawal ratio, etc., then more accurate predictions could be made. What is needed is a tool that provides the professional the means to consider a larger number of variables. Neural Networks are ideally suited to this type of problem.

Neural Network Primer

A Neural Network, as defined by Dr. Robert Hecht-Nielsen (inventor of one of the first commercial neurocomputers), is "a computing system made up of a number of simple, highly interconnected processing elements, which process information by its dynamic state response to external inputs."[4] This definition reflects the basic structure of a Neural Network as that of a computer representation of a biological neuron (nerve cell) that is interconnected with other neurons (a brain).[5] Mathematically, Neural Networks can be viewed as a multi-variable non-linear regression. Neural Networks are not programmed to find a solution, instead they learn by example. Where most computer programs have the principles of chemistry and physics coded into the program itself as equations, Neural Networks do not. Neural Networks do not have any *a priori* knowledge coded into the program at all. Conventional programs apply the coded scientific principles in the program to the problem's inputs to calculate an answer. A Neural Network learns an empirical relationship between the inputs and the desired outputs. The network is presented with different examples of inputs and outputs and relationships are learned after reviewing the examples over and over again, as many as ten million times.[6] It is capable of seeing what Professionals would see if they had the time and inclination to review the inputs and outputs a million times. Professionals who have reviewed inputs and outputs over and over again and have learned a relationship have what is called years of experience. They are considered experts.

A Neural Network initially assumes a random relationship between all the inputs and the desired outputs. By comparing its first attempt at an answer to the desired output, it self-modifies this initial random relationship into a relationship that best fits the outputs. Figure 5 is a graphical representation of a typical Neural Network.

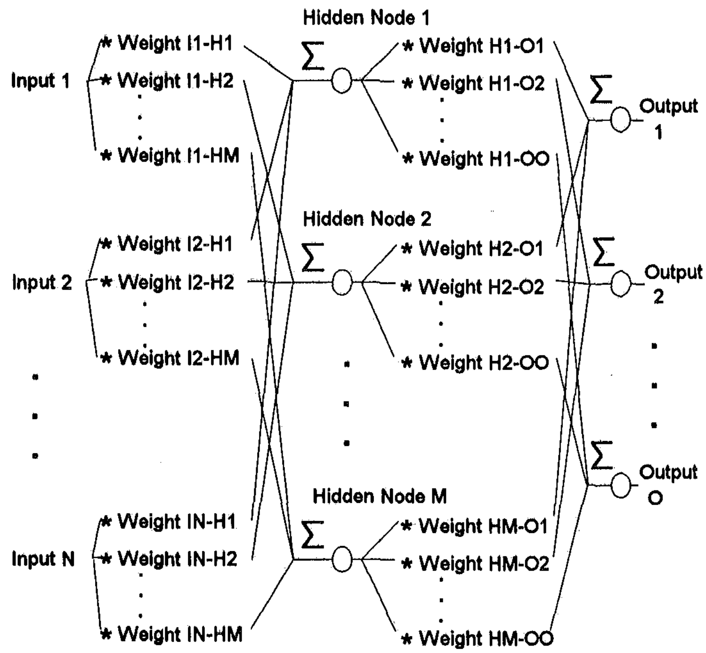

Figure 5 - Typical Neural Network Architecture

The network is made up of a group of interconnected nodes. Biologically, the nodes represent neurons and the interconnections represent the synapse, resulting in a simple brain. The largest network, biggest brain, the author has built has 158 nodes and a total of 5400 interconnections. In comparison, the human brain has over 100 billion neurons and some of these neurons are connected to 10,000 other neurons.

Every input and output is represented by an individual node. The nodes connecting the inputs with the outputs are called hidden nodes. They are called hidden nodes because the end user of the network does not know they exist. Every input node is connected to every hidden node and every hidden node is then connected to every output node. One name for this type of network is a "feed forward" network. It is called feed forward because the inputs are presented to the network on one side and they flow through the network to the outputs on the other side.

Every connection in a network is associated with a weight factor. These weight factors are independent of each other. The mathematical workings of a network are simple. First the possible range in values for every input and output is determined. Then all inputs and outputs are scaled for their range with 0 for the smallest possible value to 1 for the maximum possible value. At every node, any input to that node is multiplied by its weight factor and then summed with the product of the other inputs and their individual weight factors (see Figure 5). This resulting sum is processed by a transform function within the node.

Two of the more commonly used transform functions are the hyperbolic tangent and logistic. They are "squashing functions;" that is, they pass along a signal with a value between -1 to 1 and 0 to 1, respectively. This transform function provides non-linearity to the network and constrains the node's signal within a fixed range. Regardless of how "excited" an individual node gets, it has a fixed maximum signal that it passes on through the network, just like their biological neuron counterparts. Figure 6 shows a single node and its use of the logistic function.

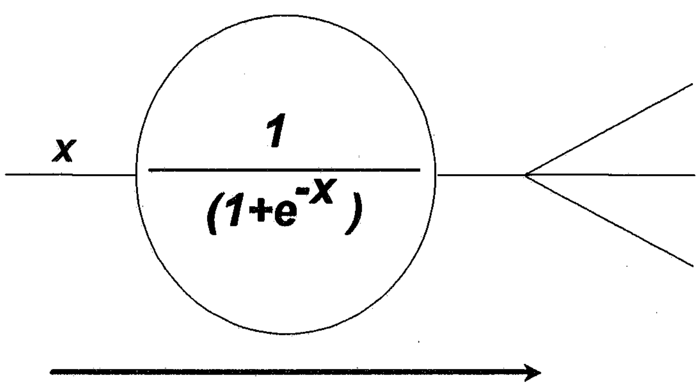

Figure 6 Node's Transform Function

The choice of which transform function and the number of hidden nodes is left to the Network designer, usually several combinations are tried and the one yielding the smallest error is finally chosen. The signals from the hidden nodes proceed in a similar fashion to the output nodes where the final signal is scaled-up to yield the network's answer. The first pass of data through the network produces an answer that is usually way off the mark since all the connection weights are initially random numbers and have not yet learned anything. The networks answer is compared to the correct answer and the network makes self-adjustments to all the connection weights to minimize this error.

This self-adjustment is the reason that another name for this type of network is "back propagation" network. The error is back propagated through the network by adjusting the connection weights in a proportional fashion so the error is minimized. This whole process, feed forward and back propagate, is repeated over and over again for every example the network is supposed to learn. Learning is finished when the average error for all the examples is minimized. This iterative learning may take as many as several million passes of each example through the network before an acceptable average error is reached.

Neural Networks are so good at learning that they can actually memorize all of the training examples. This usually produces a network that does not "understand," it will be poor at applying its newfound knowledge. It is just like the student who had a copy of the answer sheet and just memorized the sequence of answers. He may get a good grade on the test but he does poorly when having to apply this knowledge.

To keep the network honest requires an additional data set of examples for testing the network. Every so often the resulting best network to that point in training time is used on these examples to see how accurate the network is. As long as the network keeps improving on this test set, training will continue. When the network's error starts increasing on this test set, then it is futile to continue training as the network is just memorizing the answers.

The true test of the validity of a network is done by giving the network a "final exam." A third set of examples is used for this testing. The network error on this data set is the true measure of how good the Neural Network performs. The selection of data for the Learning, Testing, and Verifying data sets is not a matter to be taken lightly. To have a Neural Network that you have confidence in requires that the data in each of these three data sets be representative of the problem you are trying to solve. In terms of Artificial Intelligence, the data must cover the Domain of the problem.

It is not unusual to try various combinations of transform functions with a varying number of hidden nodes and selected inputs before deciding upon the final network architecture. This entire process—feed forward, back propagate, stop and test, and then continue training—is easily accomplished with commercially available software.

Production Prediction Neural Network

The first step in building any Neural Network is to develop a data model that captures the data that a professional would use to solve the problem. What inputs would an expert use to generate an output? The review of the past methods of predicting production is the starting point. Data used most often is either offset production data or reservoir characterization data. Since production data is readily available for all wells in the Vacuum Field and reservoir characterization data is not readily available for all the wells in the field, only production data is selected as inputs. The specific production data gathered for every offset well is shown in Table 1.

Current Rate Basis	Cumulative Basis
Oil	Oil
Water	Water
Percent Oil	Percent Oil
Injection	Injection
Inj. - Withdrawal Ratio	Inj. - Withdrawal Ratio

Table 1 Collected Production Data

The professional typically looks only at the nearest offset wells to a prospective drilling location. This usually is the nearest four to eight wells. Since a Neural Network can easily handle more wells than this, a Data Mask is utilized to capture data for wells that are not direct offset wells. The Data Mask is shown in Figure 7. The Data Mask is a 5x5 grid covering 1,000 acres. Each square grid (1,320 feet per side) equals 40 acres. This Data Mask is placed over the drilling location's 40-acre tract and the data listed in Table 1 is collected for every well within the boundaries of the Data Mask. Since the number of inputs in a Neural Network is a fixed number, the number of data values collected by the Data Mask have to be constant for every drilling location. This is impossible as the number of wells within this Data Mask varied between less than 10 to greater than 50. The solution is to divide the Data Mask into three concentric "rings." The "rings," labeled X, 10, and 9, form a bull's-eye target with the drilling location located at the center in the X-Ring. Data collected within each ring is then averaged so the data collected in Table 1 is used with the Neural Network on a per-ring basis instead of a per-well basis. The point in time that this current rate and cumulative data represent is the calendar month preceding the month that oil production for the newly drilled well is first reported to the state of New Mexico, i.e., initial completion.

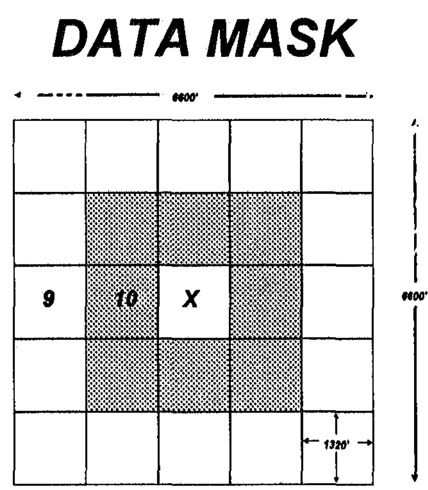

Figure 7 Data Mask

In summary, the data input items, data an expert might use to make a production forecast, for the Neural Network are the current rates and cumulatives for all the wells within a ¾-mile radius of any proposed drilling location expressed as an average for three distinct locations from the drilling location as of the month before the drilling well is initially completed. The outputs for the Neural Network are the one-, three-, six-, and twelve-month cumulative production volumes for the proposed drilling location. All this well data was collected for each of the 250 producing wells completed in the Vacuum Field since 1970. To gather the data for these 250 wells required that data for 7883 offset wells be captured. A schematic of the network is shown as Figure 8.

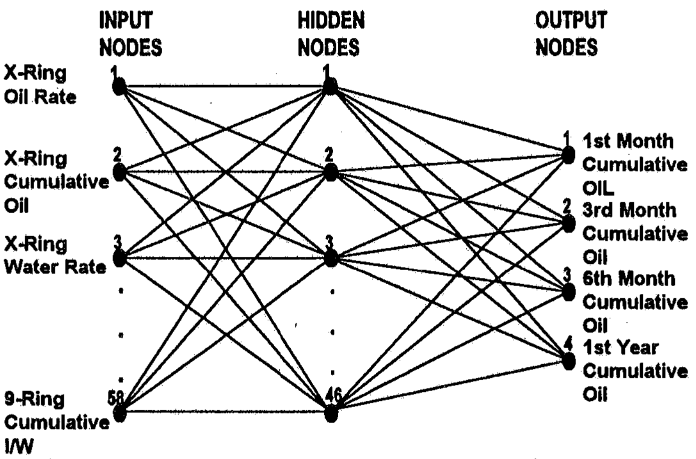

Figure 8 Neural Network

This network learned the relationship between all these variables for about 164 randomly selected wells. An additional 74 randomly selected wells were used to test the network to prevent memorization. The remaining 12 wells were used to verify the accuracy of the Neural Network. Over 30 different network configurations were tested before this network was selected as the best. Training time took approximately four hours on a 33mhz 486 PC.

Neural Network Prediction Accuracy

Figure 9 shows the same Confusion Matrix seen earlier but with the Neural Network's predictions. This is the verify data set, data the network had never seen while training. It shows that the Neural Network slightly underestimates production 38 percent of the time. It has a mean error of -19 BOPD and a mean absolute error of 27 BOPD. This network is over 3.5 times (a mean absolute error of 27 BOPD for the Neural Network versus and a mean absolute error of 99 BOPD for the Professionals) more accurate than the Professionals.

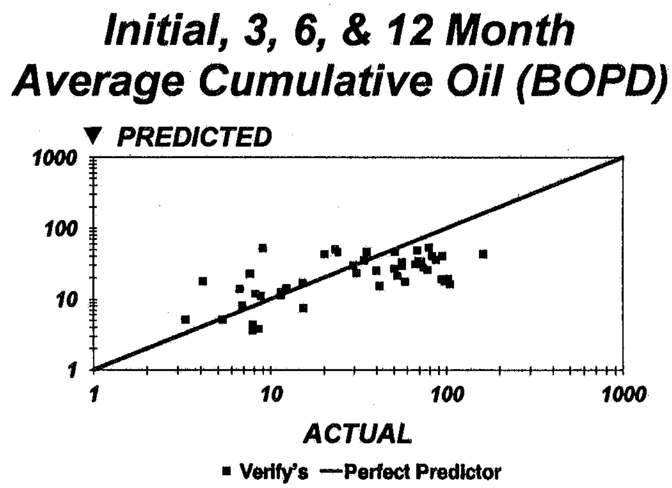

Figure 9 Neural Network's Performance

Conclusions

Statistical analyses of the networks' results are shown in Table 2 along with the analysis of the Professionals' results. As shown in Table 2, the Neural Network is within ±20BOPD 51 percent of the time and within ±10BOPD 39 percent of the time. In a direct comparison, the Neural Network out performed the Professionals 93 percent of the time.

	Accuracy (% of the Time)	
	PROFESSIONAL	NEURAL NETWORK
±20 BOPD	8	51
±10 BOPD	5	39
Optimum	7	93

Table 2 Performance Comparison

While this is quite an improvement, it still is not a perfect predictor. If the Neural Network was presented with additional data that correlates with production, then it would perform even better. An obvious source of this relevant data is geological/geophysical information. Just as the Professionals used this information in making their predictions, a Neural Network would also gain knowledge if it had geological/geophysical data as an input. Figure 10 shows the more familiar rate versus time production plot for an average well. It also shows the predicted profile for both the Professionals and the Neural Network.

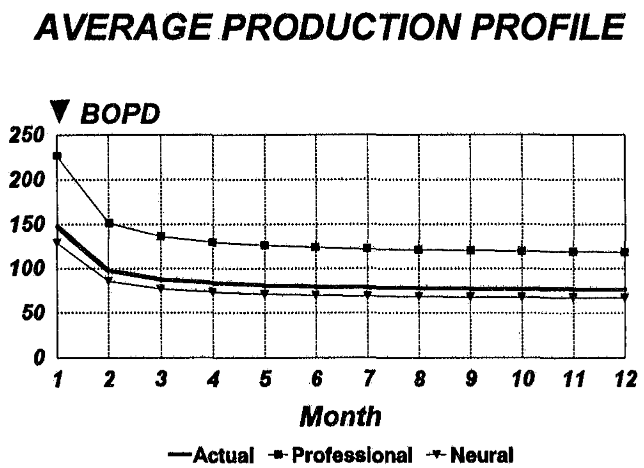

Figure 10 Typical Performance

Another advantage, besides accuracy, the Neural Network has over the Professionals is that it can quickly screen vast areas for those locations with the highest production potential, generating production profiles for all possible drilling locations. Figure 11 shows the results of such a screening that covers several select properties, about eight square miles, within the Vacuum Field. It would take a Professional weeks or even months to generate these production profiles while it takes only a couple of hours to build the data set and execute the Neural Network.

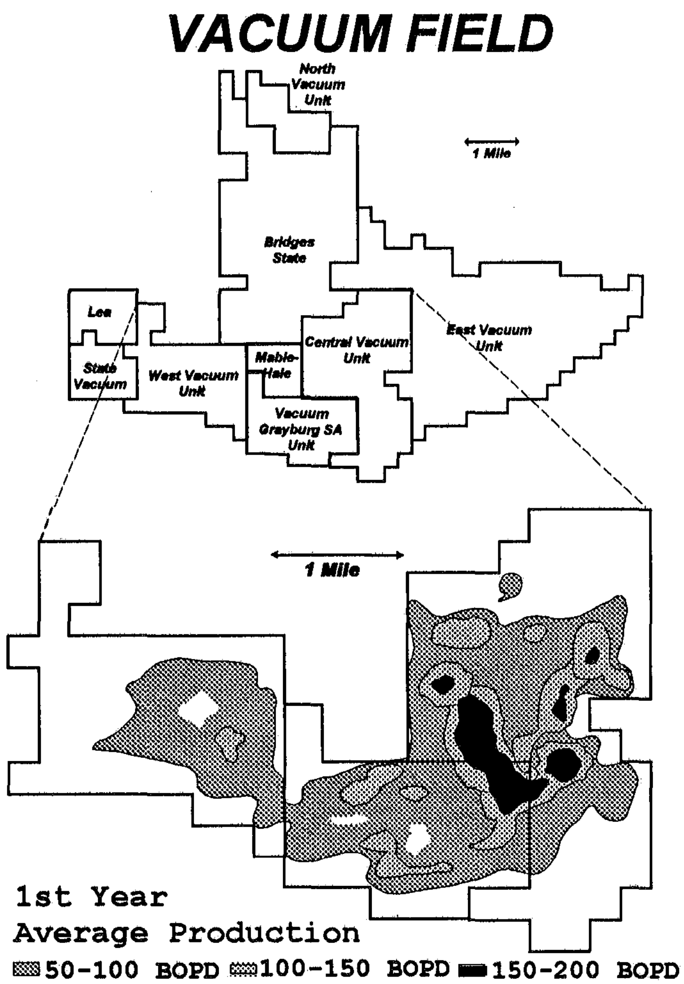

Figure 11 Neural Network Screening

Summary

A Neural Network that predicts production is a powerful tool that can reduce uncertainty in our predictions.

Acknowledgments

The author thanks Texaco Exploration and Production Inc. for permission to publish this paper and to my colleagues for their support during its preparation and review.

References

[1] *Geologic and Engineering Approaches in Evaluation of San Andres/Grayburg Hydrocarbon Reservoirs-Permian Basin*, D. G. Bebout and P. M. Harris (ed.), Bureau of Economic Geology, Austin, Texas (1990) iii

[2] Purves, William J.: "Reservoir Description of the Mobil Oil Bridges State Leases (Upper San Andres Reservoir), Vacuum Field, Lea County, New Mexico," *Geologic and Engineering Approaches in Evaluation of San Andres/Grayburg Hydrocarbon Reservoirs-Permian Basin*, D. G. Bebout and P. M. Harris (ed.), Bureau of Economic Geology, Austin, Texas (1990) 87-112

[3] Klimasauskas, Casimir C.: "Applying Neural Networks: Part III,: *PC AI* May/June 1991 **5 3**, 20-24
[4] Caudill, Maureen: "Neural Networks Primer Part I," *AI Expert* December 1987 **2 12**, 46-52

[5] Klimasauskas, Casimir C.: "Applying Neural Networks: Part I," *PC AI* January/February 1991 **5 1**, 30-33

[6] Wasserman, Phillip D.:*Neural Computing Theory and Practice*, Van Nostrand Reinhold, New York (1989) 1-26