Chapter 17

# A HYBRID GP-FUZZY APPROACH FOR RESEVOIR CHARACTERIZATION

*With a Gentle Introduction to Oil Exploration and Production*

Tina Yu, Dave Wilkinson and Deyi Xie
*ChevronTexaco Information Technology Company & ChevronTexaco Exploration and Production Technology Company*
*6001 Bollinger Canyon Road, San Ramon CA 94583 U.S.A.*

**Abstract**: A hybrid GP-fuzzy approach to model reservoir permeability is presented. This approach uses a two-step divide-and-conquer process for modelling. First, GP is applied to construct classifiers that identify permeability ranges. Within each range, ANFIS is employed to build a Takagi-Sugeno-Kang fuzzy inference system that gives permeability estimation. We applied this method to five well log data sets. The results show that this hybrid system gives more accurate permeability estimation than other previous works.

**Key words**: Oil Exploration and Production, Reservoir Characterization, Soft Computing, Permeability Estimation, Fuzzy Logic, Genetic Programming, Fuzzy Modelling.

## 1. Introduction

Where has all the oil gone? The image of oil shooting from the ground in some of the old western films gives us an illusion that if you drill it, it will come. However, this is quite far from reality. Not only is the oil reservoir difficult to find, sometimes the discovered reserve can “disappear” to the neighbouring properties. Oil hunting is therefore an adventure that continues to attract many people from different disciplines.

Oil exploration and production is a business of oil hunting. Currently, less than one third of proven reserve is recovered by oil and gas companies. How to reduce the amount of oil left behind (bypassed oil) is of great importance to the world's energy supply. One key step to increasing oil recovery is using data integration (such as seismic and well logs) intelligently to provide accurate reservoir characterization and modelling. With such information, better reservoir management decisions can be made to optimize oil production.

We first give a gentle introduction to the process of oil exploration and production (Section 2). The data integration step for reservoir characterization is then explained (Section 3). After that, we present a hybrid GP-fuzzy approach that uses well log and core data to model reservoir permeability (Section 4 & 5). The experimental results are then presented and analyzed (Section 6 & 7). Finally, we conclude the paper and give avenues of future work (Section 8).

## 2. Oil Exploration and Production

Oil exploration and production is the process of discovering and developing new petroleum or natural gas reservoirs. Typically, the process starts with a thorough analysis of the geology in a potential region, particularly the probability of finding hydrocarbons (oil or gas), and the economic factors such as risk and investment needed. If it is decided to explore for resources in the region, a seismic exploration survey is conducted. In a survey, a controlled sound source is used to set off sound waves. These waves penetrate the earth, propagate, reflect, refract and then reach back to the surface of the earth where they are recorded by geophones or hydrophones. Figure 1 gives the schematic of the seismic method. Typically, the sound wave is generated by a vibrator truck. A computer database system is connected to the geophones lying on the ground to record the data.

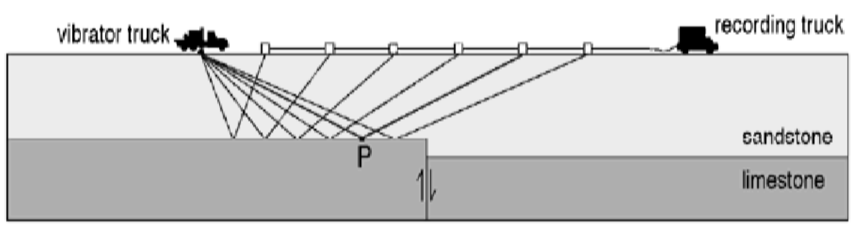


*Figure 1.* Schematic of the seismic method.

Seismic data require extensive processing to obtain an image of the subsurface, which is then carefully examined for potential accumulation points of hydrocarbons. Frequently, bright amplitudes or structural high curve shapes are indications of hydrocarbon accumulation points. For example, there are two bright spots in Figure 2 (around 10,000 feet underground). One is indicated with pressure anomaly and the other is next to where the sonic log is.

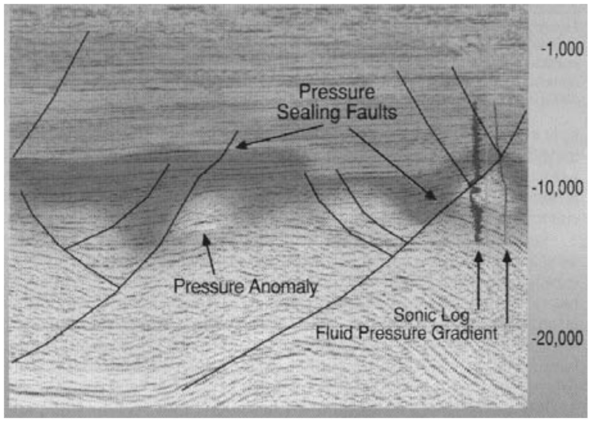


*Figure 2.* A seismic image with two bright spots.

Seismic bright spots are not sufficient evidence of the existence of hydrocarbons. To prove the presence of hydrocarbons, a well needs to be drilled. Current statistics shows that only one in three wells drilled encounters hydrocarbons; dry wells are not uncommon experiences in the oil exploration process.

Besides the presence of hydrocarbons, there are other criteria that are essential for a profitable oil production:

- First, the reservoir needs to be porous enough to contain large amounts of fluid.
- Second, these pores need to be mostly filled with hydrocarbons (instead of other substances, such as water).
- Lastly, these pores need to be connected with each other so that hydrocarbons can be extracted easily.

These three criteria are known as *porosity*, *hydrocarbon saturation*, and *permeability*. If these properties are uniform throughout the reservoir, measurements made in one exploratory well would be sufficient to determine the economic feasibility of the entire reservoir. Unfortunately, most reservoirs are heterogeneous. As a result, before deciding if an oil field is worthwhile, we need to estimate these properties for the whole area. The process of determining these spatially varying geological properties is known as reservoir characterization.

## 3. Reservoir Characterization

Reservoir characterization is a critical step in reservoir development and future production management. Knowing the details of a reservoir allows the simulation of different scenarios. The problem, however, is to define an accurate and suitable reservoir model including small-scale heterogeneity. Currently, the most abundant data about the reservoir, which is the seismic data, do not have enough resolution. The typical resolution of seismic data is on the order of 100 feet or more, which does not give enough detail of reservoir properties.

In contrast, well log data, which are collected by inserting sensing devices into an exploratory well, give an excellent description of the well at scales ranging from centimetres to hundreds of meters. However, due to its high cost, only a few well locations have log data. This scarce information is usually not sufficient to build a reservoir model that includes the small scale variations. In between wells, geological information needs to be estimated (see Figure 3).

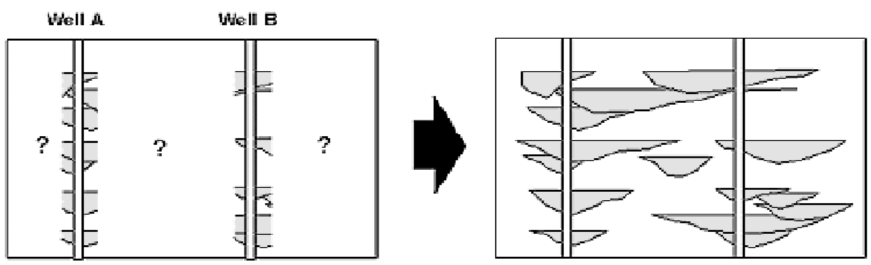


*Figure 3:* Geological information between wells needs to be filled to build a reservoir model.

A viable approach to construct a reservoir model is using both seismic and well log data. This approach first builds a framework using seismic scale and well log scale data, then fills in small details using statistical or soft computing methods. These methods are generally stochastic in nature, hence they generate many different realizations (reservoir models). This trait is actually an advantage, since the uncertainties of the different scenarios can be studied by generating a large number of equally possible reservoir models. Consequently, this approach of applying stochastic methods on seismic and well log data is widely used in reservoir characterization (Huang, Wong and Gedeon, 2000; Xie, 2001; Wong, Aminzadeh and Nikravesh, 2002; Nikravesh, Aminzadeh, and Zadeh, 2003). Currently, we are working on a research project that applies different stochastic methods to integrate seismic and well log data for reservoir characterization. This paper presents the part of the project that applies soft computing on well log data to model one of the reservoir properties: permeability.

### 3.1 Permeability Modelling

Permeability is a measure of fluid conductance in porous media. It is the most difficult reservoir property to estimate but of great importance to reservoir management decisions, such as drilling location. Two types of data were provided for this work: well log data and core permeability data. Well log data contain geological information, such as porosity and density. However, current logging technology is still unable to measure the permeability of a well. In order to obtain permeability information, rock samples (called cores) need to be removed from the exploratory wells and examined with various laboratory experiments. The permeability of the core is then calculated using Darcy's law (Scheidegger, 1974). The experimental method is described in (Yu and Lee, 2002).

It is commonly observed that permeability is related to other geological properties, although a theoretical proven mathematical equation to describe such a relationship does not exist. The objective of this work is to build a model that approximates such a relationship. This model can then be used to estimate reservoir permeability where well log data are available but core permeability data are not. In the future, it will also be used to estimate reservoir permeability where only seismic data are available. In this case, required well log properties are estimated from seismic data.

We carry out this modelling task using Genetic Programming (GP) (Koza, 1992) and Adaptive-Network-based Fuzzy Inference Systems (ANFIS) (Jang, 1993; Jang and Sun, 1997). We first describe the overall system (model) structure in Section 4. The modelling process is explained in Section 5.

## 4. The Hybrid GP-Fuzzy System

In general, developing a generic system that gives good permeability estimation for all types of reservoir is a difficult task. This is because permeability is highly dependant upon the rock formation (lithology) of a well. Moreover, different permeability ranges demonstrate different geological characteristics. We therefore adopt a divide-and-conquer approach to build the system:

– The first layer of the system identifies lithology group.
– The second layer of the system predicts permeability range.
– The last layer of the system estimates the permeability value.

Figure 4 shows the high-level structure of the 3-layer permeability estimation system.

The lithology analyzer identifies which of the five possible groups (sand, shaly sand, sandy shale, shale and high impedance sand) the rock belongs to. Within each group, a classifier predicts the range of permeability (high, medium or low). Depending on the range, a different fuzzy system is used to estimate the permeability value. In summary, given well log data inputs, the system analyzes the lithology, classifies the permeability range and then estimates the permeability value.

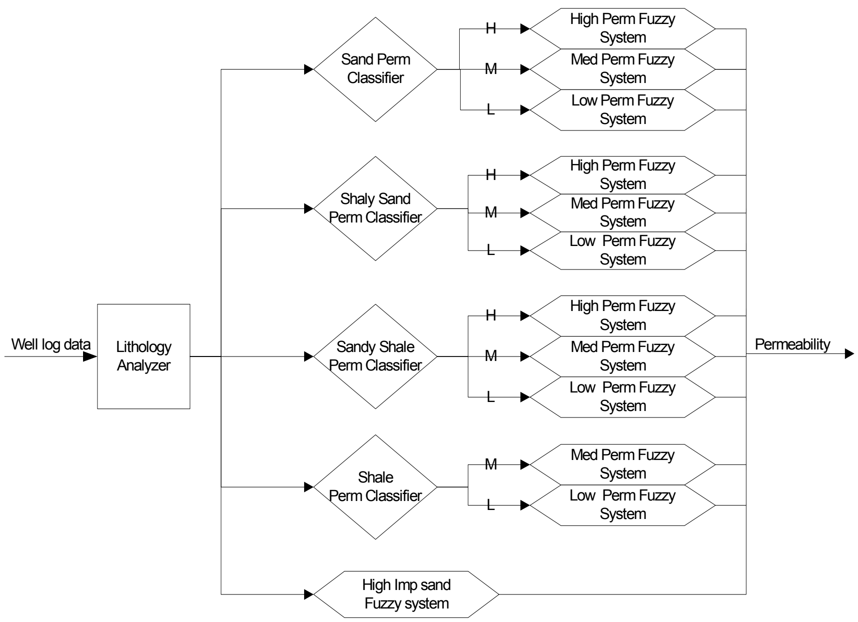


*Figure 4:* The hybrid system high-level structure.

## 5. Modelling Process

We use four types of well log data to build the system. They are *porosity*, *density*, *velocity* and *gamma ray*. We do not use log data from some of the contemporary logging technologies, such as Nuclear Magnetic Resonance (NMR), because they are not always available across all well locations. In order to maximize future usage of the system, only common types of well data are used. Another advantage of this approach is that the established methodology can be applied to build models for other oil fields.

## 5.1 Data sets

We have log data from 5 wells; all of them have matching core permeability data. Table 1 gives the number of data samples in each lithology group and permeability range. These groups and ranges are established based on geological knowledge and the analysis of the data sets. The details will be explained in Section 5.2 and 5.3. The total number of data samples is 827.

*Table1.*The number of data in each lithology groups and permeability ranges.

| | Sand | Shaly Sand | Sandy Shale | Shale | High Impedance Sand |
|---|---|---|---|---|---|
| High Perm | 148 | 81 | 83 | - | 43 |
| Medium Perm | 29 | 23 | 76 | 20 | - |
| Low Perm | 11 | 27 | 139 | 147 | - |
| Total | 188 | 131 | 298 | 167 | 43 |

## 5.2 Lithology Analyzer

One way to define lithology group is based on the percentage of shale in the rock. This information, called V-shale, can be computed using well log *gamma ray* data. Upon plotting V-shale vs. permeability (Figure 5), it is apparent that these two variables have an almost logarithmically linear relationship: the higher the V-shale value, the lower the permeability. The critical decision that needs to be made is how to partition these data into different lithology groups so that each group has similar geological properties in order to make it easier to model the data. An analysis of the s-curve suggests four lithology groups: sand, shaly sand, sandy shale and shale. The V-shale cut points for the four groups are 0.15, 0.4 and 0.75.

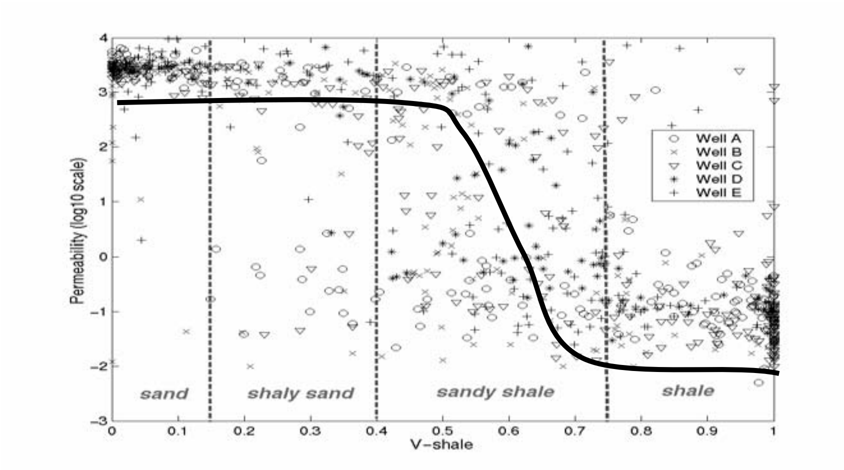


*Figure 5:* V-shale vs. permeability (log10 scale).

Among them, sandy shale data do not show any clear relationship pattern. The permeability in sandy shale is therefore the most difficult to estimate. It will be shown in Section 6 that this is indeed a challenge to our hybrid system.

High impedance sand, which has a *velocity* greater than 13,000, normally gives high permeability, regardless of the shale content. It is therefore made into a lithology group by itself. Moreover, it requires no classifier as it only has one permeability range: high (see Figure 4).

## 5.3 GP Classifiers

In each lithology group, we used GP symbolic regression to train the classifiers that predict permeability ranges (high, medium or low). The cut points of high, medium and low permeability are different in different lithology groups. For example, in sand, permeability greater than 2.5 is high and less than 1.0 is low. Anything in between is medium. Table 2 gives the permeability cut points for the four lithology groups. Note that shale has only one cut point since it has only two ranges: medium and low. Also, the values given are in log10 scale. This is a common practice in handling geological data.

*Table2.* Cut points of high, medium and low permeability in 4 lithology groups.

| Lithology Group | High vs. Med | Med vs. Low |
|---|---|---|
| Sand | 3.0 | 2.0 |
| Shaly Sand | 2.5 | 1.0 |
| Sandy Shale | 2.5 | 1.0 |

| Lithology Group | High vs. Med | Med vs. Low |
|---|---|---|
| Shale | | 1.0 |

The GP software we use, *Discipulus*(RML Technologies, 1998-2003), is only able to train binary classifiers. For lithology groups that have three possible classes, two classifiers are required. Figure 6 shows the operational flow using two classifiers to identify three permeability ranges.

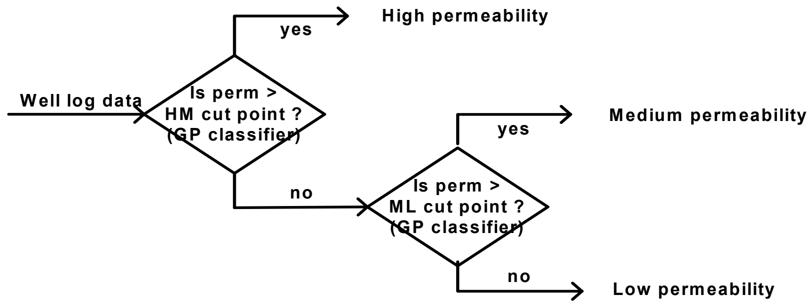


*Figure 6:* Using two classifiers to identify three permeability ranges.

Table 3 lists some of the GP parameters used to train the classifiers. The fitness is based on hit rate: the percentage of the training data that are correctly classified. During tournament selection, four candidates are randomly selected to compete for two winners. If two candidates are "tied" in their hit rates, the mean squared error measurement is used to select the winners. The "tied threshold" for mean squared error measurement is 0.01% in this work. If two classifiers are tied in both their hit rates and mean squared error measurements, one of them is selected as the winner randomly.

*Table 3*. Some of the GP parameters to evolve the classifiers.

| | |
|---|---|
| Objective | Generate a classifier that distinguishes high vs. medium or medium vs. low permeability. |
| Functions | addition; subtraction; multiplication; division; abs; data transformation |
| Terminals | Porosity ($\phi$), density (v), velocity ($1/\rho$) and constants. |
| Fitness | Hit rate then squared error fitness |
| Hit rate | The percentage of the training data that are correctly classified. |
| Selection | Tournament (4 candidates/2 winners) |

There is another set of GP parameters that are not fixed but selected by the software for each run. These parameters include *population size*, *maximum program size*, *crossover* and *mutation rates*. In the first run, one set of values for these parameters were specified. When the run does not produce improved solution for a number of generations, this run is terminated and a new run starts with a new set of parameter values. The system records the best solution found throughout the runs. The best classifier found in the runs is selected as the final winner. In this work, each classifier is set to continue for 100 to 120 runs depending on the solution quality, i.e. more runs are made if the solution found so far is not satisfactory.

The software also assembles the best classifiers into teams. A team may consist of any odd number of classifiers. The whole team decision is the majority vote of the team members. Since the number of classifiers in a team is odd, there is always a winner (i.e. no tie). The following is an example of a team:

$f_1(\phi, \nu, \rho) = \ldots \rho * \nu - \phi^2 / \rho \ldots$
$f_2(\phi, \nu, \rho) = \ldots \rho^{\nu} - \phi^2 / \rho + \nu + 1.2 \ldots$
$f_3(\phi, \nu, \rho) = \ldots \nu * \nu - \phi^2 / \rho - (\phi + 2.3) \ldots$
*cut point* = 2.5;
*Majority ((*$f_1$ *>cut point), (*$f_2$ *> cut point), (*$f_3$ *> cut point))*

## 5.4 ANFIS Fuzzy Inference System

ANFIS is a fuzzy modelling tool that generates Takagi-Sugeno-Kang (TSK) type fuzzy inference systems (Sugeno, 1985; Sugeno and Kang, 1986; Sugeno and Tanaka, 1986) based on the given input and output data. A TSK fuzzy system has the following structure:

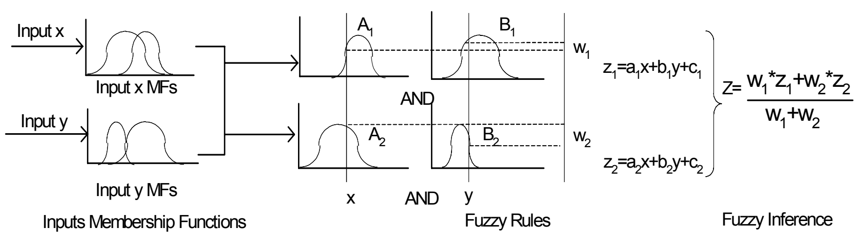


*Figure 7:* TSK fuzzy inference system structure.

The first component is a set of input membership functions (MFs). Membership functions map crisp inputs to linguistic values or labels. A membership function can have any shape, such as triangular, Gaussian and trapezoidal, as long as it varies between 0 and 1. In Figure 7, each curve defines a membership function. The transformation of a crisp input into degree (between 0 and 1) of match with one or more linguistic values is called "fuzzification".

Fuzzy rules are conditional statements in if-then format. The if-part consists of linguistic values and fuzzy operators (AND, OR, NOT). The then-part is a first order linear equation. The following is a TSK fuzzy rule:

*If porosity is high and density is low, permeability = -42572*porosity-53/velocity-115807*density + 260911.*

Fuzzy inference is a method to interpret the input values and, based on the fuzzy rules, assigns output values. In a TSK system, the output value is calculated based on the firing strength $w_i$ of each rule. The formula is given in Figure 7.

TSK is an abstract system model and can be implemented in many different ways. One implementation is ANFIS, which use a feed-forward network architecture to implement the TSK system. Figure 8 is the equivalent ANFIS representation of the TSK system in Figure 7. Layer 1 corresponds to the input membership functions while layer 2 corresponds to the fuzzy rules. Layer 3, 4 and 5 correspond to the calculation of the final output z.

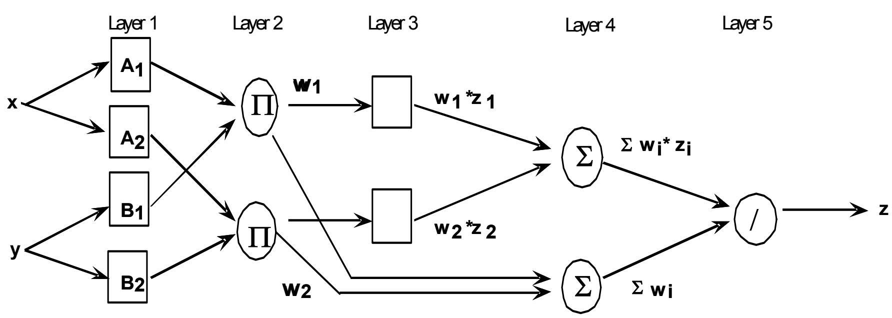


*Figure 8:* The ANFIS system architecture.

In the software we used, *matlab* (The MathWorks, Inc., 1994-2003), the construction of ANFIS is automatic through a two-step learning process:

- The network structure is built by a subtractive clustering algorithm (Chiu, 1994).
- The input membership function parameters (layer 1) are tuned by a back propagation algorithm while the output parameters (layer 3) are tuned by a least squares estimation.

The first learning step determines the number of fuzzy rules (layer 2). The clustering algorithm first partitions the data into groups and then generates a minimum number of rules to distinguish the fuzzy qualities associated with each of the groups. The shapes of the input membership functions (layer 1) are not learned but specified by users. Depending on the shapes of the input membership function selected, different parameters (e.g.

mean and standard deviation for Gaussian) are initialized. Meanwhile, the output parameters (e.g. the coefficients in the linear equations) are initialized.

The second learning step adjusts input and output parameters to minimize the error. More specifically, in the forward pass, training inputs go forward till layer 4 and the output parameters are identified by the least squared estimate. In the backward pass, the error rates propagate back and the input membership function parameters are updated by gradient decent.

The first learning step is a one-pass process. In contrast, the second learning step is an iterative one. Users can specify the “error tolerance” as the stopping criterion. The learning process stops after the learning error has reached the specified error tolerance.

## 6. Results

The accuracy of the GP classifiers is given in Table 4. As expected, sand and shale have simple, uniform property and are easy to model. In contrast, sandy shale has the most complex geological formation (discussed in Section 5.2). As a result, the classifiers have the lowest accuracy rate.

*Table4.* Classification accuracy of the GP classifiers.

| | Sand | Shaly Sand | Sandy Shale | Shale |
|---|---|---|---|---|
| High Permeability | 93.24% | 93.83% | 67.47% | - |
| Med Permeability | 82.76% | 78.95% | 78.95% | 85% |
| Low Permeability | 100% | 88.89% | 80.58% | 96.64% |
| Total | 92.02 | 90.08% | 76.51% | 97.01% |

The TSK fuzzy systems give permeability estimations that are well correlated with the core permeability for all lithology groups. The coefficient of variation values ($R^2$) are between 0.9 and 0.95.

We combined the GP classifiers with the TSK systems then ran the overall system on the data set. Figure 9 shows the results. Among the 5 lithology groups, sandy shale data have the worst match between estimated and core permeability. This is due to the misclassification of permeability range by the GP classifiers (Table 4). Once the permeability range is misclassified, it is very hard for the TSK system to give correct permeability estimation. This indicates that the accuracy of the classifiers is of vital importance to the overall system performance. We are currently investigating different techniques to improve these classifiers including different classifier ensemble methods and Receiver Operational Characteristics analysis for better cut point selection.

Figure 10 gives the permeability estimations of the current system. It is clear that the hybrid GP-fuzzy system outperforms the current system on all 5 wells' data. We also made detailed performance comparisons with previous systems using data from only one well (well C). Figure 11 shows that the hybrid system (Perm-FIS) gives permeability estimations that match the core permeability (Perm-Core) better than the two previous works (Perm-Log is the current system and Perm-Vsh is the system used when only shale data is available). For example, the Core-Perm ("+" in Figure 10) between depth 6360 and 6340 feet are with high values (> $10^3$). The value decreases between 6340 and 6335 feet and raises again at 6330 feet. The hybrid system gives permeability estimations that match this trend of high-low-high pattern very well. In contrast, Perm-Vsh gives permeability estimations around $10^{2.3}$ across this depth range, matching badly to the Core-Perm trend. Another previous work, Perm-Log, gives a better match than Perm-Vsh does. However, its estimations match poorly between the depth of 6340 and 6325 feet where Core-Perm drops. We are very encouraged by the results and convinced that this hybrid modelling approach is well suited for reservoir characterization.

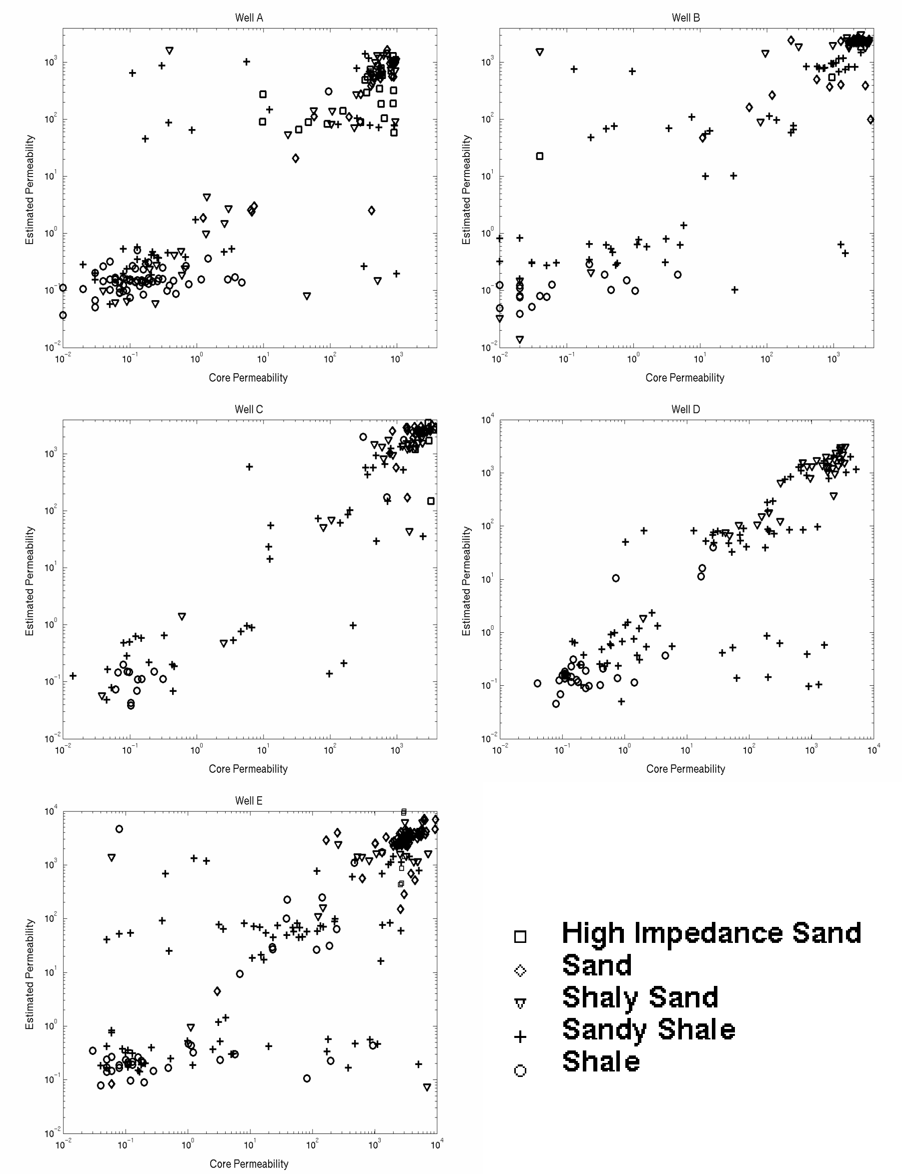


*Figure 9:* The overall hybrid system performance (estimations vs. core permeability).

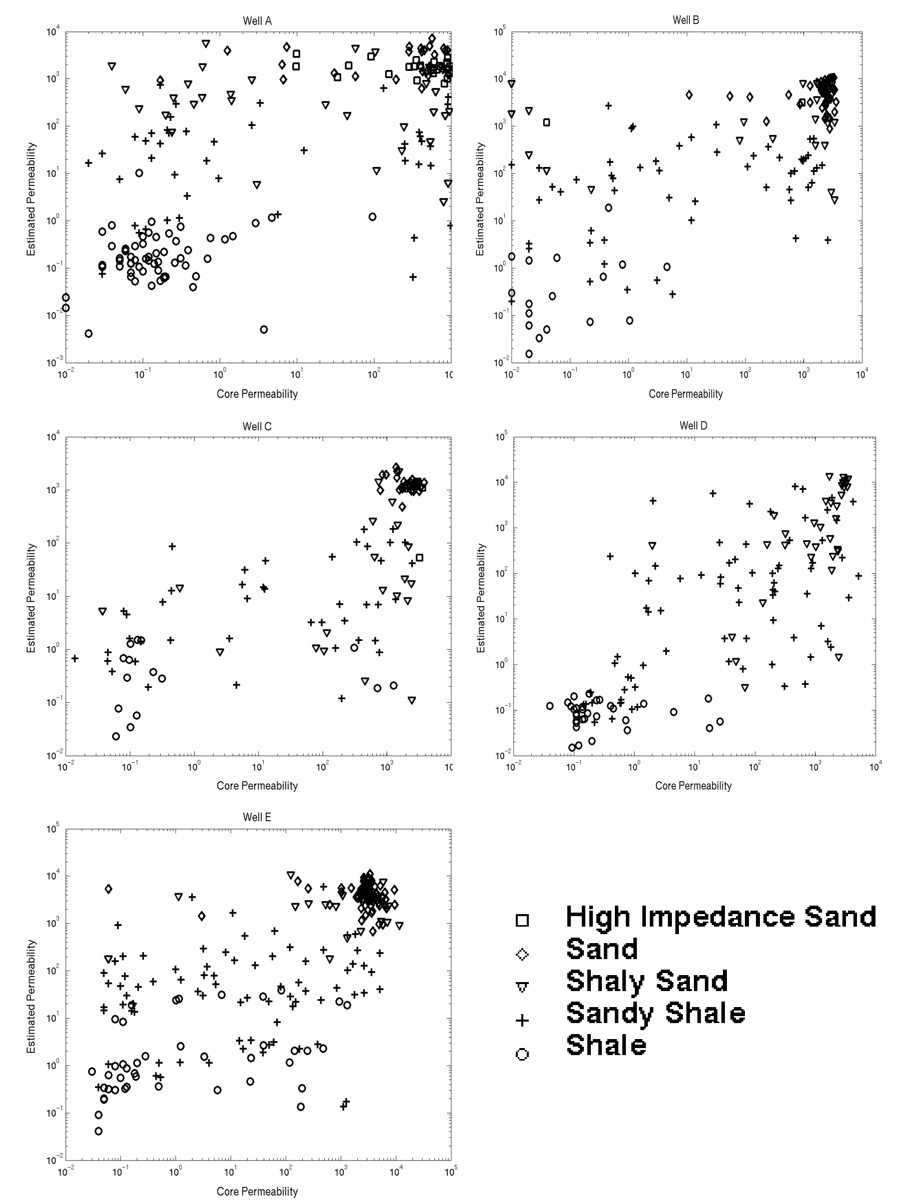


*Figure 10:* The current system performance (estimations vs. core permeability).

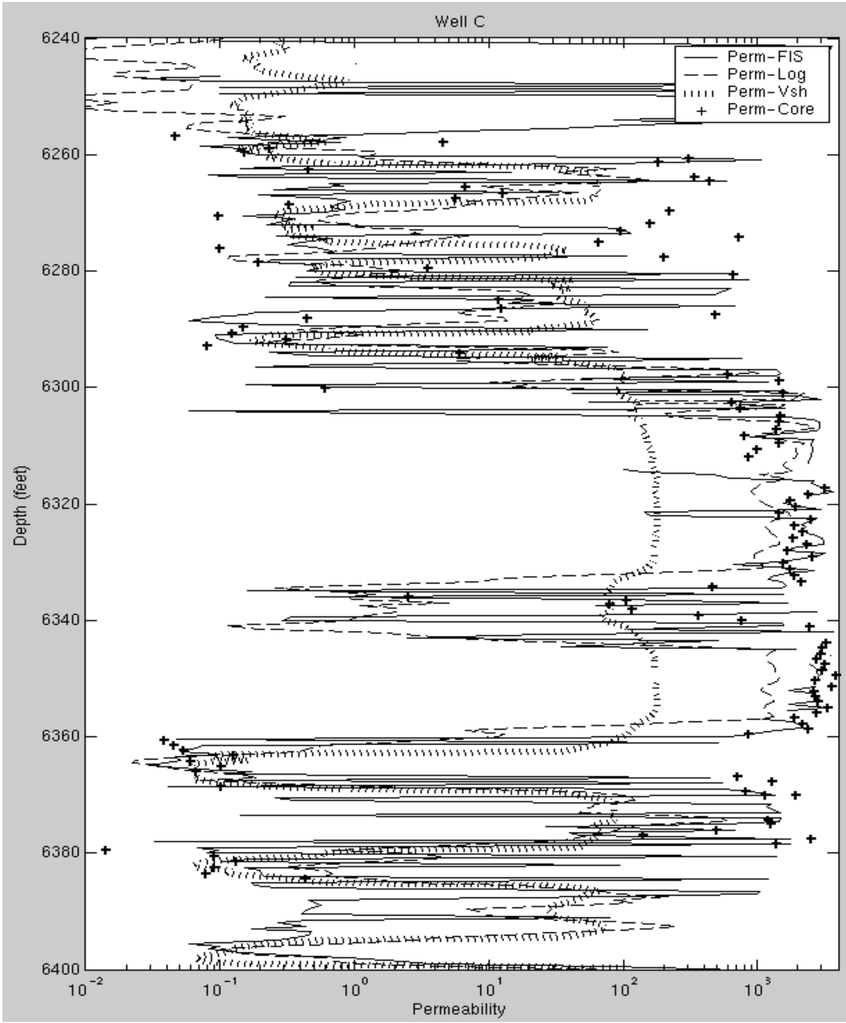


*Figure 11:* Performance comparison of the hybrid system with two previous systems.

## 7. Discussions

Real world applications are results driven. Meanwhile, there is time pressure to deliver. Since the classification accuracy rates were mostly satisfactory, we made little effort in GP theoretical analysis. The only exception is the sandy shale classifiers because they have much lower accuracy than other lithology groups and the misclassification impairs the overall hybrid system performance.

One investigation we made was on the sensitivity of the mean squared error (MSE) tied threshold. Recall that the fitness measurement is based on hit rate rather than MSE (see Section 5.3). We have tested two MSE thresholds (1% and 0.1%) using cut point 1.0 to evolve the classifier that separates low from high/medium permeability. Table 5 and 6 give the results. Note that sandy shale classifiers have a different structure from that in Figure 6. It first separates low from high/medium, and then distinguishes high from medium. In this way, both classifiers have a balanced set of positive and negative samples for training (see Table 1).

*Table 5.* Confusion matrix for sandy shale classifier using MSE tied threshold 1%.

| | Estimated Low | Estimated High/Medium | Total |
|---|---|---|---|
| Core Low | 126 (90.65%) | 13 (9.35%) | 139 |
| Core High/Medium | 36 (22.64%) | 123 (77.36%) | 159 |
| Total | 162 | 136 | 298 (83.56%) |

*Table 6.* Confusion matrix for sandy shale classifier using MSE tied threshold 0.1%.

| | Estimated Low | Estimated High/Medium | Total |
|---|---|---|---|
| Core Low | 116 (83.45%) | 23 (16.55%) | 139 |
| Core High/Medium | 21 (13.21%) | 138 (86.79%) | 159 |
| Total | 137 | 161 | 298 (85.23%) |

With MSE tied threshold 1%, the evolved classifier identifies more data as low permeability (162 vs. 137) while the classifier built using threshold 0.1% identifies more data as high/medium permeability (161 vs. 136). This indicates that many evolved classifiers have the same hit rate and using the second fitness measurement (MSE) for selection can impact the quality of the classifiers. In this case, a smaller MSE threshold trains a better classifier.

What is a good sensitivity measure for an effective classifier? Generally, we prefer robust solutions, hence would prefer a medium range tie threshold. However, in this case, the accuracy of this classifier is very important to the overall system performance because it is the first classifier applied to the data. Once a data point is misclassified, the second classifier can not correct it and the ANFIS fuzzy system would not be able to give the correct permeability estimation. We therefore select the classifier that was trained using the smaller MSE threshold. However, we decide not to train new classifiers using MSE thresholds smaller than 0.1% in order to preserve the robustness of the classifiers.

Evaluation of alternative technologies is also very important in this project. We had tested various modelling technologies on sandy shale data.

These include Decision Tree (C4.5) and Neural Networks. The results show that GP classifiers have the best accuracy rates.

Would theoretical work impact GP applications and vice versa? My view is yes. Most GP applications are multi-disciplinary projects, which provide the opportunity to investigate "what GP is doing" from the prospect of application domain. For example, while working on an engine oil blending project (Yu and Rutherford, 2001), a sub-equation with value between -1 and 1 was evolved during GP symbolic regression. The final viscosity increase of the engine oil is based on this value and adjusted by various additives used. With a thorough analysis, this might provide a different way to understand the GP process.

Theoretical works also help practitioners explaining GP to people who are not familiar with the technology. In the past few years, I have been working with researchers from many different fields. It is helpful to be able to discuss GP search process with them using fundamental theorems.

## 8. Conclusions

Optimization of oil production is becoming more important as the global energy supply is facing various challenges. We have presented a hybrid GP-fuzzy system for reservoir characterization and demonstrated its effectiveness on modelling reservoir permeability. The divide-and-conquer approach is very effective in dealing with the geological heterogeneity in a typical reservoir. This is shown in the exceptional performance of the hybrid systems: once the lithology group and permeability range are identified, estimation of permeability becomes more robust.

We are continuing this research in improving the classifiers, particular that of sandy shale. Meanwhile, we are testing the system on other well log data and investigating issues for future system deployment.

## Acknowledgements

I would like to thank Bill Langdon, Mark Kotanchek, Conor Ryan and Terry Soule for their comments and suggestions. Thanks to Rick Riolo for encouraging me to write up the discussions section. We also wish to acknowledge ChevronTexaco Overseas Petroleum, for providing data and granting permission for these results to be published.

## References


Chiu, S. (1994). A Cluster Estimation Method with Extension to Fuzzy Model Identification. In *Proceedings of Third IEEE International Conference on Fuzzy Systems*, pp. 1240-1244. IEEE Press.

Huang, Y., Wong, P. M., and Gedeon, T. D. (2000). "Permeability Prediction in Petroleum Reservoir using a Hybrid System." In *Soft Computing in Industrial Applications*, Y. Suzuki, S.J. Ovaska, T. Furuhashi, R. Roy and Y. Dote (eds.), Springer-Verlag, London, 437-446.

Jang, J-S R. (1993). ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on Systems, Man and Cybernetics 23( 3)*: 665-685.

Jang, J-S. R. and Sun, C-T. (1997). *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence.* Prentice Hall.

Koza, J. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, MA: MIT Press.

The MathWorks, Inc. MATLAB. Natick, MA, 1994-2003.

Nikravesh, M., Aminzadeh, F. and Zadeh, L. A. (2003). *Intelligent Data Analysis for Oil Exploration, Developments in Petroleum Science*, Elsevier, in press.

RML Technologies, Inc. Discipulus 3.0, Littleton, CO, 1998-2003.

Scheidegger, A. E. (1974). *The Physics of Flow through Porous Media*. University of Toronto Press.

Sugeno, M. (1985). *Industrial Applications of Fuzzy Control*, Elsevier Science Pub. Co.

Sugeno, M. and Kang, G. T. (1986). Fuzzy Modeling and Control of Multilayer Incinerator. *Fuzzy Sets and Systems 18:* 329-346

Sugeno, M and Tanaka, K. (1991). Successive Identification of a Fuzzy Model and Its Applications to Prediction of Complex Systems. *Fuzzy Sets and Systems 42:* 315-334

Wong, P. M., Aminzadeh, F. and Nikravesh, M. (2002). *Soft Computing for Reservoir Characterization and Modeling.* Series Studies in Fuzziness and Soft Computing, Vol. 80, Physical Verlag, Springer.

Xie, D. (2001) Thin-bed Reservoir Characterization using Integrated 3-D Seismic and Well Log Data: a Case Study of the Central Boonsville Field, Fort Worth Basin, North-Central Texas, PhD Dissertation, Michigan Technological University.

Yu, T. and Lee, S. (2002). Evolving Cellular Automata to Model Fluid Flow in Porous Media. In *Proceedings of the 2002 NASA/DoD Conference on Evolvable Hardware*, pp. 210-217. IEEE Computer Society.

Yu, T. and Rutherford, J. (2001). Modelling Sparse Engine Test Data Using Genetic Programming. In *Proceedings of Fourth Workshop on Mining Scientific Datasets,* pp. 42-48. Chandrika Kamath.