The Application of Fuzzy Logic and Genetic Algorithms to Reservoir Characterization and Modeling

S.J. Cuddy[1,2] and P.W.J. Glover[2]

[1] Petro-Innovations, An Caisteal, 378 North Deside Road, Cults, Aberdeen, UK.

[2] Department of Geology and Petroleum Geology, University of Aberdeen, UK.

Abstract. A 3D model of oil and gas fields is important for reserves estimation, for cost effective well placing and for input into reservoir simulators. Reservoir characterization of permeability, litho-facies and other properties of the rocks is essential. A good model depends on calibration at the well locations, with cored wells providing the best data. A subset of wells may contain specialized information such as shear velocity data, whereas other wells may contain only basic logs. We have developed techniques able to populate the entire field database with a complete set of log and core data using fuzzy Logic, genetic algorithms and hybrid models. Once the gaps in the well database have been filled, well logs can be imported to a 3D modeling software package, blocked and upscaled to match the geocellular model cell size.

Litho-facies typing and permeability are important for understanding sedimentological controls on reservoir quality distribution as well an input to 3D reservoir models. Litho-facies and permeability prediction have presented a challenge due to the lack of borehole tools that measure them directly. We demonstrate, using several field examples, how these new predictive methods can be applied in a variety of ways to enhance the understanding of rock physical properties. Examples include prediction of litho-facies, permeability and shear sonic logs. The new techniques give better predictions compared to conventional methods such as multiple linear regression and cluster analysis.

1. Introduction

In this paper we describe two soft computing techniques, fuzzy logic and genetic algorithms, for making predictions from electrical logs. These results are used to improve reservoir characterization and modeling.

The philosophy of the fuzzy logic technique discussed in Section 2. Section 3 explores these concepts further. Two important inputs for 3D modeling are litho-facies and rock permeability. Section 4 describes the mathematical implementation of fuzzy logic for the purpose of litho-facies prediction, and Section 5 applies fuzzy logic to litho-facies prediction in the North Sea. Sections 6 and 7 discuss the requirements for, and implementation of fuzzy logic for the purpose of permeability prediction, and Section 8 applies fuzzy logic to permeability prediction in the Ula field.

The philosophy of the genetic algorithm approach is outlined in Section 9. Section 10 applies the genetic algorithm technique to the prediction of shear velocity in the North Sea. Finally, Section 11 compares briefly the use of the fuzzy logic and genetic algorithm techniques with other more conventional methods used in the geosciences.

The soft computing concepts of fuzzy logic and genetic algorithms have been around since the 1960's, but have only recently been applied to reservoir characterization and modeling. This is mainly due the dramatic improvement in the speed of computers. The computer programs described in this paper take only a couple of minutes to run on a 400 MHz computer. A number of oil and service companies have confidential fuzzy logic and genetic algorithms software. It is hoped that this paper will introduce these topics to the public domain.

2 Reservoir Characterization using Fuzzy Logic

Fuzzy logic is an extension of conventional Boolean logic (zeros and ones) developed to handle the concept of "partial truth" – truth values between "completely true" and "completely false". Dr. Lotfi Zadeh of UC/Berkeley introduced it in the 1960's as a means to model uncertainty [1].

Science is heavily influenced by Aristotle's laws of logic initiated by the ancient Greeks and developed by many scientists and philosophers since [2]. Aristotle's laws are based on "X or not-X"; a thing either is, or is not. This has been used as a basis for almost everything that we do. We use it when we classify things and when we judge things. Managers want to know whether something is this or that, and even movies have clear goodies and baddies. Conventional logic is an extension of our subjective desire to categorize things. Life is simplified if we think in terms of black and white. This way of looking at things as true or false was reinforced with the introduction of computers that only use the bits 1 or 0. When the early computers arrived with their machine driven binary system, Boolean logic was adopted as the natural reasoning mechanism for them.

Conventional logic forces the continuous world to be described with a coarse approximation; and in so doing, much of the fine detail is lost. We miss a lot in the simplification. By only accepting the two possibilities, the infinite number of possibilities in between them is lost. Reality does not work in black and white, but in shades of gray. Not only does truth exist fundamentally on a sliding scale, it is also perceived to vary gradually by uncertainties in measurements and interpretations. Hence, a gray scale can be a more useful explanation than two end points. For instance, we can look at a map of the Earth and see mountains and valleys, but it is difficult to define where mountains start and the valleys end.

This is the mathematics of fuzzy logic. Once the reality of the gray scale has been accepted, a system is required to cope with the multitude of possibilities. Probability theory helps quantify the grayness or fuzziness. It may not be possible to understand the reason behind random events, but fuzzy logic can help bring meaning to the bigger picture. Take, for instance, a piece of reservoir rock. Aeolian rock generally has good porosity and fluvial rock poorer porosity. If we find a piece of rock with a porosity of 2 porosity units (pu), is it aeolian or fluvial? We could say it is definitely fluvial and get on with more important matters. But let's say it is probably fluvial but there is a slim probability that it could be aeolian. Aeolian rocks are generally clean (i.e., contains little or no clay minerals) and fluvial rocks shalier (i.e., contain clay minerals). The same piece of rock contains 30% clay minerals. Is it aeolian or fluvial? We could say it is approximately equally likely to be aeolian or fluvial based on this measurement.

This is how fuzzy logic works. It does not accept something is either this or that. Rather, it assigns a grayness, or probability, to the quality of the prediction on each parameter of the rock, whether it is porosity, shaliness or colour. There is also the possibility that there is a measurement error and the porosity is 20 pu not 2 pu. Fuzzy logic combines these probabilities and predicts that, based on porosity, shaliness and other characteristics, the rock is most likely to be aeolian and provides a probability for this scenario. However, fuzzy logic says that there is also the possibility it could be fluvial, and provides a probability for this to be the case too. In essence, fuzzy logic maintains that any interpretation is possible but some are more probable than others. One advantage of fuzzy logic is that we never need to make a concrete decision. In addition, fuzzy logic can be described by established statistical algorithms, and computers, which themselves work in ones and zeros, can do this effortlessly for us.

3 Why Fuzzy Logic can help the Geosciences

Geoscientists live with error, uncertainty and fragile correlations between data sets. These conditions are inherent to the geosciences, because of the challenge of designing and building sensors to measure complex formations in hostile environments. Even in the laboratory it is difficult to relate a log response to a physical parameter. Several perturbing effects such as mineralogy, fluids and drilling fluid invasion can influence a simple measurement, say porosity. Conventional techniques try to minimize or ignore the error. Fuzzy logic asserts that there is useful information in this error. The error information can be used to provide a powerful predictive tool for the geoscientist to complement conventional techniques. Fuzzy logic is now used routinely in formation evaluation [3][4]

Early investigators of natural science noticed that many seemingly random events fell into a pattern. These eighteenth century scientists found an astonishing degree of regularity in the variation of an observation about its mean or average value. These patterns or distributions were closely approximated by continuous curves referred to "normal curves of errors" and attributed to the laws of chance. Abraham De Moivre (1667 to 1745), Pierre Laplace (1749 to 1827), and Karl Gauss (1777 to 1855) first studied the mathematical properties of these normal curves. These curves are now called normal or Gaussian curves, and have a characteristic bell-shape. This distribution is the cornerstone of modern statistical theory [5].

The normal distribution is more than an accident of nature. It is a fundamental result of applied probability known as the Central Limit Theorem. This remarkable theorem states that a distribution that is the result of a number of underlying, relatively independent, variables will be normal in shape irrespective of the distribution shapes of the component variables. For instance if we take the porosity of a core-plug, each plug consists of numerous pores, each of which contribute to the pore volume. Many factors control an individual pore volume including grain shape, mineralisation and pore fluids. In addition, when we measure porosity the resulting errors are the combined effect of a large number of independent sources of error. The resulting porosity distribution will be normal as a direct result of the Central Limit Theorem, and this is confirmed by the empirical analysis of core-plugs.

Fuzzy logic does not require a normal distribution to work as any type of distribution that can be described mathematically can be used. Because of the prevalence of the normal distribution, supported by the Central Limit Theorem and observation, it is the best distribution to use in most cases. The normal distribution is completely described by two parameters, its mean and variance. As a consequence, core-plugs from a particular litho-facies may have dozens of underlying variables controlling their porosities but their porosity distribution will tend to be normal in shape and defined by two parameters - their average value (arithmetic mean) and their variance, which is a measure of the width of the distribution. This variance (the standard deviation squared) depends on the hidden underlying parameters and measurement error. This variance, or fuzziness, about the average value, is key to the method and the reason why it is called fuzzy logic.

To clarify the importance of the fuzzy term, take an example of two litho-types. Aeolian facies may have an average porosity of 20 pu and a variance, or fuzziness, of ± 2 pu. Fluvial facies may have an average porosity of 10 pu with a variance of ± 4 pu. If we measure the porosity of an unknown facies as 15 pu, it could belong to either litho-facies. However, it is less likely to be aeolian because the aeolian distribution is much tighter, even though its porosity is equally distant from the "most likely" or average porosity expected for each litho-type. Litho-facies prediction using fuzzy logic is based on the assertion that a particular litho-facies type can give any log reading although some readings are more likely than others.

4 The Fuzzy Mathematics of Litho-Facies Prediction

The normal distribution is given by:

$$P(x) = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}} \tag{1}$$

P(x) is the probability density that an observation x is measured in the data-set described by the arithmetic mean μ and the standard deviation σ.

In conventional statistics the area under the curve described by the normal distribution represents the probability of a variable x falling into a range, say between x_1 and x_2. The curve itself represents the relative probability of variable x occurring in the distribution. That is to say, the mean value is more likely to occur than values 1 or 2 standard deviations from it. This curve is used to estimate the relative probability, or fuzzy possibility, that a data value belongs to a particular data set. If a litho-facies type has a porosity distribution with a mean μ and standard deviation σ, the fuzzy possibility that a well log porosity value x is measured in this litho-facies type can be estimated using Equation (1). The mean and standard deviation are simply derived from the calibrating or conditioning data set; usually core data.

Where there are several litho-facies types in a well, the porosity value x may belong to any of these litho-facies, but some are more likely than others. Each of these litho-facies types has its own mean and standard deviation, such that for f litho-facies types there are f pairs of μ and σ. If the porosity measurement is assumed to belong to litho-facies f, the fuzzy possibility that porosity x is measured (logged) can be calculated using Equation (1) by substituting μ_f and σ_f. Similarly, the fuzzy possibilities can be computed for all f litho-facies. These fuzzy possibilities refer only to particular litho-facies and cannot be compared directly as they are not additive and do not sum to unity. It is necessary, therefore, to devise a means of comparing these possibilities.

We would like to know the ratio of the fuzzy possibility for each litho-facies to the fuzzy possibility of the mean or most likely observation. This is achieved by de-normalizing Equation (1).

The fuzzy possibility of the mean observation μ being measured is:

$$P(\mu) = \frac{e^{-(\mu-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}} = \frac{1}{\sigma\sqrt{2\pi}} \tag{2}$$

The relative fuzzy possibility $R(x_f)$ of a porosity x belonging to litho-facies type f compared to the fuzzy possibility of measuring the mean value μ_f is Equation (1) divided by Equation (2):

$$R(x_f) = e^{-(x-\mu_f)^2/2\sigma_f^2} \tag{3}$$

Each fuzzy possibility is now self-referenced to all possible litho-facies types. To compare these fuzzy possibilities between litho-facies, the relative occurrence of each litho-facies type in the well must be taken into account. This is achieved by multiplying Equation (3) by the square root of the expected occurrence of litho-facies f. If this is denoted by n_f, the fuzzy possibility of measured porosity x belonging to litho-facies type f is:

$$F(x_f) = \sqrt{n_f}\, e^{-(x-\mu_f)^2/2\sigma_f^2} \tag{4}$$

The fuzzy possibility $F(x_f)$ is based on the porosity measurement (log), x, alone. This process is repeated for a second log type such as the volume of shale, y. This will give $F(y_f)$, the fuzzy possibility of the measured volume of shale y belonging to litho-facies type f. This process can be repeated for another log type, say z, to give $F(z_f)$. At this point we have several fuzzy possibilities ($F(x_f)$, $F(y_f)$, $F(z_f)$ ….) based on the fuzzy possibilities from different measurements (x, y, z ….) predicting that litho-facies type f is most probable. These fuzzy possibilities are combined harmonically to give a combined fuzzy possibility:

$$\frac{1}{C_f} = \frac{1}{F(x_f)} + \frac{1}{F(y_f)} + \frac{1}{F(z_f)} + \ldots \tag{5}$$

This process is repeated for each of the f litho-facies types. The litho-facies that is associated with the highest combined fuzzy possibility is taken as the most likely litho-facies for that set of logs. The associated fuzzy possibility $C_f(max)$ provides the confidence factor for the litho-facies prediction. There are statistical techniques for combining probabilities based on Bayes Theorem. The fuzzy logic technique described in this paper has been developed by analysis of large data sets from many oil fields, and differs from Bayes theorem in two respects. The fuzzy possibilities in fuzzy logic are combined harmonically, whereas the Bayes approach combines probabilities geometrically. When comparing lithologies that are equally likely, with similar probabilities, the harmonic combination emphasizes any indicator, which suggests the lithology selection is unlikely.

Secondly, fuzzy logic weights the possibilities by the square root of the proportion in the calibrating data set whereas the Bayes approach uses the direct proportion.

Litho-facies prediction using fuzzy logic is based on the assertion that a particular litho-facies type can give any log reading although some readings are more likely than others. For instance, clean aeolian sand is most likely to have a high porosity, although there is a finite probability that the logging tool could measure a low porosity. It is important to have a consistent set of logs between wells, although accuracy is not essential. In practice the best curves to use are the porosity log (in pu units), as this can be calibrated to core, and the normalized gamma ray (in API units). The gamma ray can be normalized by creating a frequency distribution of the gamma ray readings within the reservoir formation. The five-percentile point is determined for each well, and this point is regarded as the clean point. This clean point plus a fixed number of API units (say 100 API) determine the shale point. The gamma-ray log can then be re-scaled between 0 and 100%.

Any number of curves can be used by the technique. However, the additions of further curves may not necessarily improve the prediction as the porosity and shaliness response to the litho-facies type generally controls other log responses. The photoelectric, nuclear magnetic resonance and resistivity log curves are possible exceptions to this rule.

5 The Application of Fuzzy Logic to Litho-Facies Prediction in the North Sea

Litho-facies typing is useful in well correlation, and is important for building a 3D model of the field by geostatistical or stochastic techniques. These models can be used for assessing oil volumes in the reservoir, well placing and reservoir engineering. Using fuzzy logic for litho-facies prediction makes no assumptions and retains the possibility that a particular facies type can give any log reading, although some are more likely than others. This error or fuzziness has been measured and used to improve the facies prediction in several North Sea fields.

The Viking area is located on the northern flank of the Permian Rotliegendes Sandstone in the Southern North Sea. The Viking field was developed in 1972 and to date has produced 2.8 Tcf of gas. Consideration has recently been given to tying back several smaller satellite pools. As part of the feasibility study, 13 exploration and production wells, drilled between 1969 and 1994, have been re-evaluated using fuzzy logic.

The reservoir was deposited in a desert by aeolian, fluvial, and lacustrine processes. Three major litho-facies associations have been recognized from core studies:

- Aeolian Dune. Aeolian sandstones have the best permeabilities by virtue of their better sorting and lack of detrital clays. Clean aeolian dune sandstones give the highest porosities in the reservoir, with an average around 16 pu. Dune base sandstones (wind ripple) give a lower average porosity of 12-14 pu, as they are less well sorted.

- Sabkha. Sandy sabkha has good porosity but the presence of detrital clay enhances compaction effects and thus reduces primary porosity. Muddy sabkha porosities and permeabilities are very low with no reservoir potential.

- Fluvial. The fluvial sandstones often have poorer permeabilities (<0.3 mD) and porosities (<10 pu) than the sandy sabkha sandstones. Their porosity is dependent on the detrital clay content and pore filling cements.

In addition, in all litho-facies, diagenetic overprint of pervasive fibrous illite clays severely reduces permeabilities. Only in the well-sorted grain-flow litho-facies that has a macro-porous network are moderate permeabilities retained. The object of applying fuzzy logic to this field was to differentiate litho-facies types in uncored wells and to help with building the reservoir model of the field and with future well placing.

One recent well with substantial core coverage was used to calibrate the litho-facies and permeability predictor for the older wells. The left track of Figure 1 shows the core-described facies from this well. There are several litho-facies described; aeolian, fluvial and sabkha. The aeolian facies is sub-divided into grainflow, wind-ripple and sand sheet sub-facies; the sabkha facies into sandy, mixed and muddy sub-facies; and the fluvial facies into cross-bedded and structureless sub-facies. The result of the fuzzy predicted litho-facies is shown in the second track. There is near perfect differentiation between aeolian, fluvial and sabkha rock types. In addition, the technique goes some way towards differentiating between sandy, mixed and muddy sabkhas. The right track shows the comparison of core derived and fuzzy predicted permeabilities. It must be remembered that the core descriptions themselves are from observations and can contain errors due to the subjective nature of the measurement. Consequently, sedimentologists can use predicted litho-facies as an aid to refining core interpretations. This example of a self-calibrated well has helped the sub-surface team develop the Viking satellite reservoirs pools. “Blind-testing” between wells can test the predictive ability of the technique in the same field. This was conducted on data from the South Ravenspurn field.

The South Ravenspurn gas field is located in the southern North Sea, 40 miles off the English coast. Gas reserves are around 1 Tcf, and current production is 200 mmscf/d. The field is developed by some 40 wells, in shallow water no more than 50 meters deep. Descriptions from 10 cored wells were used to derive facies in 30 uncored wells. The left well shown in Figure 2 shows the described and predicted facies types for one cored well in the field. The prediction success rate is over

86% compared to a random prediction rate of 13%. The prediction success rate is calculated as the number of correct predictions divided by the total number of possible predictions. When we are attempting to predict *X* facies types, say 10, a random prediction success rate would be around *1/X* or 10%. Any prediction method is expected to produce successful predictions greater than this threshold.

Using the fuzzy relationships between the described litho-facies and electrical logs, litho-facies were predicted in a second well shown on the right of Figure 2. The prediction success in this second well between the predicted facies and "hidden" but known and core-described facies is 73%, with the majority of the "failed" predictions falling into the next closest litho-facies type rather than one with completely different reservoir characteristics.

6 The Application of Fuzzy Logic to Permeability Prediction

Knowledge of permeability, the ability of rocks to flow hydrocarbons, is important for understanding oil and gas reservoirs. Permeability is best measured in the laboratory on cored rock taken from the reservoir. However coring is expensive and time-consuming in comparison to the electronic survey techniques most commonly used to gain information about permeability. In a typical oil or gas field all boreholes are "logged" using electrical tools to measure geophysical parameters such as porosity and density. Samples of these are cored, with the cored material used to measure permeability directly. The challenge is to predict permeability in all boreholes by calibration with the more limited core information.

In principle, determining permeability from electrical measurements is a matter of solving equations in rock physics. In practice, there are numerous complicating factors that make a direct functional relationship difficult or impossible to determine. One problem is that permeability is related to the aperture of pore throats between rock grains, which logging tools find difficult to measure. Several perturbing effects such as mineralogy, reservoir fluids and drilling fluid invasion can influence the permeability measurement. Litho-facies determination is a clear application of fuzzy logic as the litho-facies types are described in clear "bin" types such as aeolian or fluvial. These predicted litho-facies, in wells without core, have several uses from inter-well correlation to geostatistical modeling. One of the main drivers behind litho-typing is to predict permeability as the different litho-facies exhibit different permeabilities. It was soon realized that fuzzy logic could be used to predict permeability directly, by-passing the litho-facies step.

Permeability is a very difficult rock parameter to measure directly from electrical logs because it is related more to the aperture of pore throats rather than pore size. There is a weak correlation between the two that explains the spread of points on cross-plots of core porosity and permeability. Determining permeability from logs is further complicated by the problem of scale; many well logs have a vertical resolution of typically 2 feet compared to the 2 inches of core plugs. In addition to these issues, there are measurement errors on both the logs and core. When you add these problems together it is surprising that predictions can be made at all. The mathematics of fuzzy logic provides a way of not only dealing with the errors, but also using them to improve the prediction.

7 The Fuzzy Mathematics of Permeability Prediction

Fuzzy logic is used for litho-facies prediction by assigning a data bin to each litho-type. The challenge for litho-typing is how to combine the fuzzy possibilities between the litho-types as the litho-facies are not equally frequent in the cored section of the well. Predicting permeability using fuzzy logic, avoids this problem by ensuring, at the outset, that the bins are of equal size. First the core permeability values are scanned by the program and divided into ten (or more) equal bin sizes on a logarithmic scale. That is to say that the bin boundaries are determined so that the number of core permeabilities in Bin 1 represents the tenth percentile boundary of the permeability data. Bin 2 represents the twentieth percentile boundary and so on. In this example there are ten divisions in the data but there is no reason why there could not be twenty or more. Each one of these bins is then compared to the electrical logs. The log data associated with levels in the well corresponding to Bin 1 (very low permeability) are analyzed and their mean and standard deviation calculated. In this way, not only is the average or most probable log value associated with Bin 1 calculated, but also some idea of the uncertainty in the measurement is obtained. Again porosity and volume of shale are the best and first logs to try. Fuzzy logic asserts that a particular log porosity value can be associated with any permeability, but some are more likely than others.

This logic is clarified using Figure 3. For simplicity it shows only 5 bins that represent each of the 5 familiar decades for logarithmic permeability. The diagram shows only 2 axes (porosity and volume of shale) whereas the technique can use an unlimited number of bins in n-dimensional space. The mean value of porosity and volume of shale for each permeability bin is represented by the point at the center of each cross. For instance, for core permeability greater than 100 mD the average porosity and volume of shale are 26 pu and 12%, respectively. The vertical and horizontal lines through each point represent the error bar or standard deviation (fuzziness) of data in that bin. The error bars are different for each bin. The resulting permeability line, through the points, is field specific and is “S” shaped and shown without error. A real cross plot of log data would show considerable scatter about this curve. A single curve predictor would predict different permeabilities depending whether porosity or volume of shale was taken as the predictor. Take a log depth that has a porosity of 23 pu and volume of shale of 30% as shown on Figure 3. A porosity only predictor would estimate a permeability of 10-100 mD by extrapolating the point vertically. The volume of shale only predictor would give a permeability of 0.1-1 mD by extrapolating the point horizontally.

In contrast, fuzzy logic can deal with “shades of gray”. The point at 23 pu and 30% volume of shale would be compared to all permeability bins. Knowing the mean and standard deviation of each bin, the fuzzy possibility that the point lies in that bin can be calculated using Equation (3). It is not necessary to normalize the distributions because the permeability bins are of equal size. This is done separately for porosity and the volume of shale. Their fuzzy possibilities are combined to predict the permeability for that log depth with its associated fuzzy possibility or “grayness”. Figure 4 shows typical results of this analysis where each of the ten permeability bins has an associated fuzzy possibility. The highest fuzzy possibility is taken as the most probable permeability for that combination of log measurements. A predicted permeability is calculated as the weighted mean of the two most probable bins.

The program uses any number of permeability bins with any number of input curves. The distribution of bin boundaries depends on the range of expected permeabilities, as described above. The number of bins depends on the number of core permeabilities available for calibration, the statistical sample size. A reasonable sample size is around 30. Consequently the number of bins is determined so that that there are at least 30 sample points per bin. For a well with 300 core permeabilities it would be appropriate to use 10 permeability bins. The permeability prediction has also been attempted using genetic algorithms [6]. Vertical permeability can be predicted simultaneously by simply comparing the core vertical permeabilities with the logs in a similar manner.

8 The Application of Fuzzy Logic to Permeability Prediction in the Ula Field

The Ula field is 130 miles to the southwest of Norway and was discovered in 1976. The recoverable reserves of Ula are 435 million barrels of oil, 167 billion cubic feet of gas and 42.8 million barrels of NGL. The reservoir is late Jurassic sandstone at a depth of 3320 mtvdss. It has porosities of around 20 pu with average permeabilities of 300 mD. Fuzzy logic was recently used to update the reservoir model in order to unlock the potential of an upper unit using new drilling techniques. This interval contains potentially 50% of the remaining reserves and was initially ignored because of poor rock characteristics.

The right hand track of Figure 5 shows the comparison between core-derived and fuzzy-predicted permeabilities in one of the cored Ula wells. “Blind-testing” between wells was used to test the predictive ability of the technique. To test the fuzzy prediction, the technique was calibrated in a cored well and “blind-tested” in another well to see how well it fitted the actual core permeabilities. Figure 6 shows the second well where permeabilities were predicted using the calibration from the first well. The comparison between the predicted and cored derived permeabilities is good compared to the natural spread in permeability data.

9 Reservoir Characterization using Genetic Algorithms

Genetic algorithms (GAs) are models of computer learning, which derive their behavior from an analogy of the processes of evolution in nature. The individual organisms in this analogy are possible solutions to some given well-defined problem in reservoir characterization. The analogy is implemented by the creation within a computer of a population of individuals represented by GA-chromosomes that are analogous to the DNA chromosomes. These GA-chromosomes take the form of mathematical equations relating the solution to a set of input data. The individuals in the population then go through a process of evolution. Mutation, achieved through random number generation, can play an important part in the process. After a number of generations, the computer uses a fitness function to select individuals probabilistically to undergo genetic operations analogous to sexual reproduction. The fitness function assesses how close the individual comes to solving the problem. Genetic algorithms use stochastic processes, and as they are not random searches for a solution to a problem, they perform better than classical optimization routines. As in nature poorly performing individual die or their species become extinct, the computer discard poor solutions. The computer then iterates using the new population, with one iteration being one generation.

10 Shear Velocity Prediction using Genetic Algorithms

The measurement of shear velocities is important for understanding reservoir rock properties. Shear sonic data (*Dts*) is required for rock strength analysis to determine fracture propagation and formation breakdown characteristics, and for improved porosity prediction as *Dts* is largely unaffected by fluid type. Shear sonic data are also becoming important for enhanced seismic interpretation. Because the value of shear velocity data is only now being realized, and because such data is expensive to acquire, there is limited amount of information available in the North Sea. Genetic algorithms have been used to determine the shear velocities in oil wells based on calibrations elsewhere in the oil field. Not only have genetic algorithms determined the constant parameters of these calibrations, but GAs have also evolved the calibration equations themselves.

In a recent study, calibration data from 4 wells with shear velocity data were used to populate all the wells in a large field. This gave the oil company a cost effective method of building a 3D reservoir model that enabled improved location of oil wells. *Dts* can be acquired by dipole logging tools. If *Dts* data have not been acquired by logging, it can be estimated from other curve responses using genetic algorithms.

Shear velocities are related to porosity ϕ, formation resistivity Rt, and the volume of shale Vsh. Porosity is the measure of pore space in the rock matrix that is filled with reservoir fluids such as oil, gas and water. Formation resistivity is the inverse of the conductivity of the fluid-saturated rock. The volume of shale, in this context, is a normalized measure of the radioactivity of the rock matrix by measuring the formation gamma-ray background. Porosity ϕ, Rt and Vsh are measured by borehole electrical logs.

Our objective is to construct empirically a function $f(\phi,Rt,Vsh)$ which predicts shear velocities at each depth, i given ϕ, Rt, Vsh and at each depth. We are therefore searching for an appropriate function of the form:

$$Dts = f(\phi,Rt,Vsh) = [a\ \phi^{\,b}] \bullet_1 [c\ Rt^{\,d}] \bullet_2 [e\ Vsh^{\,g}] \bullet_3 [h] \tag{6}$$

where $\bullet_1$, $\bullet_2$, $\bullet_3$ etc. represent the algebraic operators addition and multiplication, a, c, e, and h are unknown constants, and b, d, and g are unknown constant exponents.

The next step is to provide a method for determining how good a given $f(\phi,Rt,Vsh)$ is as a predictor of Dts. The approach we adopt is to sum absolute errors in prediction over all depth levels for a given borehole. We seek a function of the form Equation (6), which minimizes this sum. A more standard way to do this might be to use least squares rather than absolute values of residuals. The reason for the approach that we take is that the borehole data is noisy and includes many “outliers”. These can only be removed by extensive manual editing of the data sets and rechecking of measurements. By using the absolute value of residuals, one diminishes the effect of noise and outliers and produces more appropriate predictor functions. Mathematically, the problem can be stated as:

$$\underset{f}{Minimise}: \quad {}_{i}\left|Vs_i - f(\phi_i, Rt_i, Vsh_i)\right| \tag{7}$$

The genetic algorithms were constructed as follows. An initial population of individuals is picked randomly in the solution space. Each individual has randomly chosen constants a,b,c,d,e,g,h and operators $\bullet_1$, $\bullet_2$, $\bullet_3$. The fitness criterion of each of these individuals is determined by Equation (7). The best existing algorithm for minimizing Equation (7) starts with a randomly generated f and uses local search by mutating the coefficients one at a time or flipping the operator between an addition and a multiplication. The mutation range is initially set very high in order that the individuals search all of the solution space. After a number of generations a pool of individuals is selected, by linear ranking, for mutating and cloning. Mating is achieved by coefficient merging. Some of the best individuals are cloned to add more individuals, where solutions are most promising. After a number of generations the mathematical operators are fixed and the percentage change in mutated coefficients is gradually reduced. The algorithm stops when the percentage improvement in evaluation reaches a predefined lower limit or a maximum number of iterations has been reached.

Each chromosome is a vector of length 10. Three alleles are binary integer values that represent the mathematical operators $\bullet_1$, $\bullet_2$, $\bullet_3$. The rest of the alleles are floating point values that represent the coefficients *a,b,c,d,e,g,h*. The initial population is generated by creating chromosomes with a random binary numbers for $\bullet_1$, $\bullet_2$, $\bullet_3$ and random floating point numbers for the coefficients *a,b,c,d,e,g,h*. If the allele represents the operator $\bullet$, its value is binary and it will be switched. If the allele represents one of the real variables, it will be modified by multiplication by a value randomly picked from the range 0.8 to 1.2. This range decreases in value as the number of generations increases. This provides a method that allows the search to become more local towards the end of the algorithm as better solutions emerge.

The prediction of shear velocities by GAFL (Genetic Algorithms and Fuzzy Logic) is show in Figures 7, 8 and 9. Track 3 of Figure 7 shows a comparison of the log-derived (measured) shear velocities with those predicted using fuzzy logic. Track 4 of Figure 7 shows a comparison of the log-derived shear velocities with those predicted using genetic algorithms. The comparisons are extremely good. Surprisingly, for the thin-bedded intervals, where there is a mismatch between the measured and predicted velocities, it is thought that the measured log values are incorrect. This is because the vertical resolution of the shear velocity measurement tool prohibits a correct measurement in thin beds. However, the fuzzy logic and genetic algorithm techniques have access to all measurements, including ones with better vertical resolution than the shear velocity tool. As a result, the measurements with good vertical resolution pick the facies type before the techniques predict the appropriate shear velocity, and the predicted shear velocity in the thin beds is likely to be a more appropriate value than actually measured by the shear velocity tool.

Figures 8 and 9 show the recorded data cross-plotted against the predictions for the fuzzy logic and genetic algorithm techniques. Data from 4 wells in the field are shown. The key point to notice is the good predictions by both techniques at the extremes of the scale. This is where methods such as linear regression fail. The extremes of the scale are often the most important in reservoir characterization. The “stratification” of the data in Figure 8 is a result of fuzzy logic being a binning technique. This effect is not noticed when the data is displayed as a continuous curve as in Figure 7. The fuzzy logic technique has the advantage of not requiring all the input data to make a prediction. Often curves are missing from some wells due to borehole problems. Genetic algorithms derive an equation which calculates a continuous curve. It therefore requires all the input curves to make a prediction, but does not show any stratification. In this example fuzzy logic and genetic algorithms compute shear velocities by two independent methods and therefore provide confidence in the predictions.

11 Comparison of Fuzzy Logic and Genetic Algorithms with Other Methods

Genetic algorithms and fuzzy logic (GAFL) are two ways, out of many, of making predictions from logs. Standard statistical techniques such as least squares regressions are essential tools of the geoscientist but are poor at predicting extremes, whereas fuzzy logic seeks these out. However, least squares regression has the ability to extrapolate and predict values outside the range of the conditioning data set whereas fuzzy techniques are confined to look only within the calibrating data set.

Neural networks are a promising technique, but require the correct amount of conditioning. In addition, neural networks are very hard to "figure out" and are therefore often regarded as "black boxes". By contrast, GAFL results are completely open and easy to understand, and relate to the problem at hand. Although interpreting fuzzy results is simple they often describe complex non-linear systems that would defy conventional logic. Cluster analysis works well but can have difficulty in dealing with data equidistant from cluster centers, and requires extensive user interaction *via* cross plots. Artificial intelligence and expert systems have clear decision logic and generally ignore or minimize the error in the data. These other methods have their place and are valuable to the geosciences. There is no reason why they should not incorporate elements of GAFL or complement the GAFL results.

In common with other techniques, GAFL can easily incorporate an unlimited number of input logs. It is equally fair to say that where dozens of curves are available, analysis shows that a couple of carefully picked curves contain most of the information controlling the correlations. GAFL requires little user intervention, as there are no cross-plots to make, or parameters to set. This is a useful feature for the busy geoscientist, as the technique can be applied to fields containing hundreds of wells in a matter of minutes.

12 Conclusions

Fuzzy logic and genetic algorithms have found several applications in reservoir characterization and modeling including the prediction of litho-facies, permeability and shear velocities. These are simple tools for confirming known correlations, or as powerful predictors in uncored wells. Litho-facies typing is used for well correlation and as input for building a 3D model of the field. Permeability prediction is useful to complement current technology and to gain insight into older wells without core or extensive logging programmes. The measurement of shear velocities is important for understanding reservoir rock properties.

The methods described here use basic log data sets such as porosity and density, which are cheap and easy to obtain, rather than depending on new and expensive logging technology. Over recent years, oil exploration has suffered due to erratic and often low oil prices. Oil producing countries are now struggling to meet demand, and there is an urgent need to find new reservoirs and make efficient use of existing resources. Fuzzy logic and genetic algorithms make an important contribution to this endeavor.

13 Nomenclature

x	= log variable
μ	= arithmetic mean
σ	= standard deviation
n_f	= expected occurrence of x in litho-facies f
μ_f	= arithmetic mean value of x in litho-facies f
σ_f	= standard deviation of x in litho-facies f
$P(x)$	= fuzzy possibility density of an observation x
$R(x_f)$	= relative fuzzy possibility of x
$F(x_f)$	= fuzzy possibility of x belonging to litho-facies f
C_f	= combined fuzzy possibility
Dts	= shear velocity
ϕ	= porosity
Rt	= resistivity
Vsh	= volume of shale
a,b,c,d,e,g,h	= numerical coefficients to be determined
$\bullet_1, \bullet_2, \bullet_3$	= operators representing either addition or multiplication

14 References

1. Zadeh, Lotfi., *Fuzzy Sets*, Information and Control 8:338-353. 1965.
2. Kosko, B., *Fuzzy Thinking - The New Science of Fuzzy Logic*, Hyperion, New York, NY. 1993.
3. Norland Ulf., "Formalizing Geological Knowledge - With an Example of Modeling Stratigraphy using Fuzzy logic". *Journal of Sedimentary Research* Vol. 66, No. 4, July 1996.
4. Cuddy, S.J., "The Application of Fuzzy Logic to Petrophysics", The Thirty-Eighth Annual Logging Symposium of the Society of Professional Well Log Analysts, June 1997.
5. Freund J.E,. Walpole R.E., *Mathematical Statistics*. Prentice - Hall International. p206. 1980.
6. Brown, D.F, Cuddy S.J., Garmendia-Doval A.B. and McCall J.A.W., "The Prediction of Permeability in Oil-Bearing Strata using Genetic Algorithms" Third IASTED International Conference Artificial Intelligence and Soft Computing, July 2000.

Fig. 1. Permeability and facies prediction.

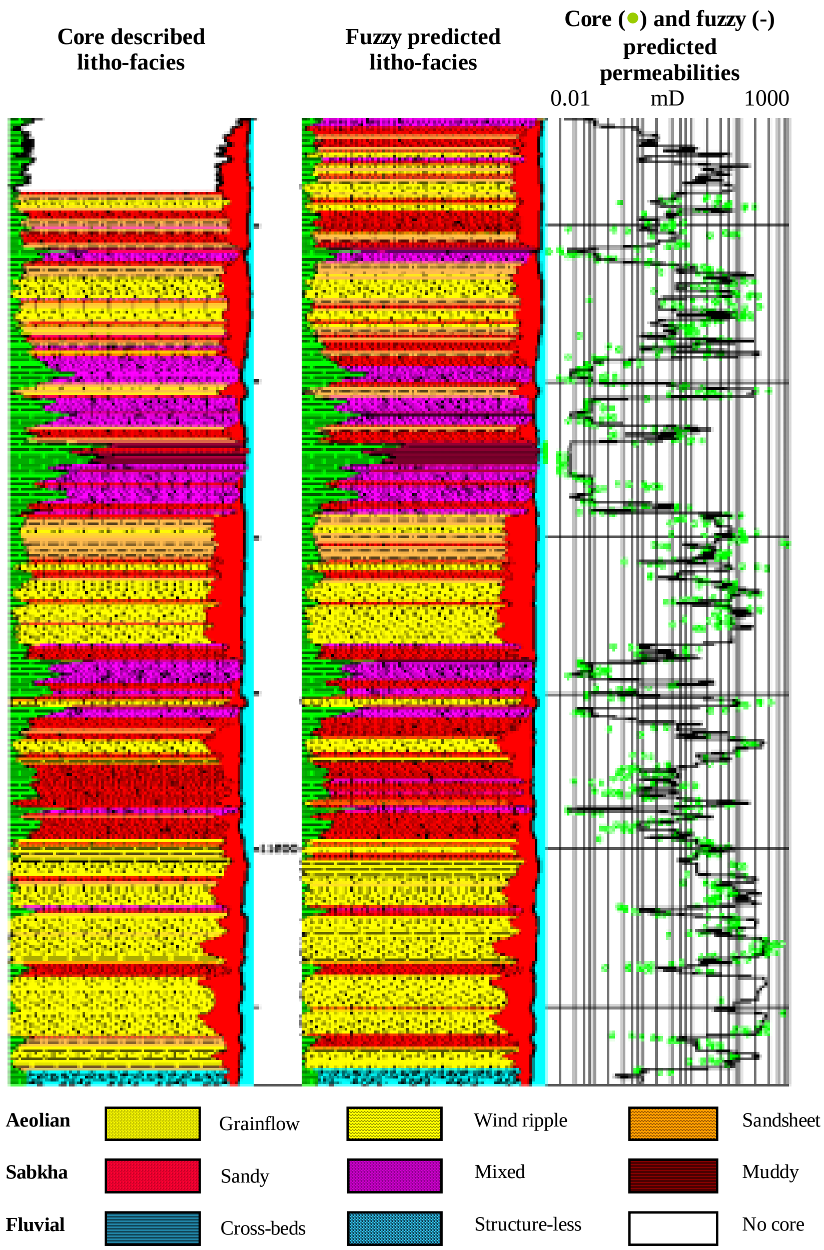


Fig. 2. Blind-testing predictions in the South Ravenspurn field.

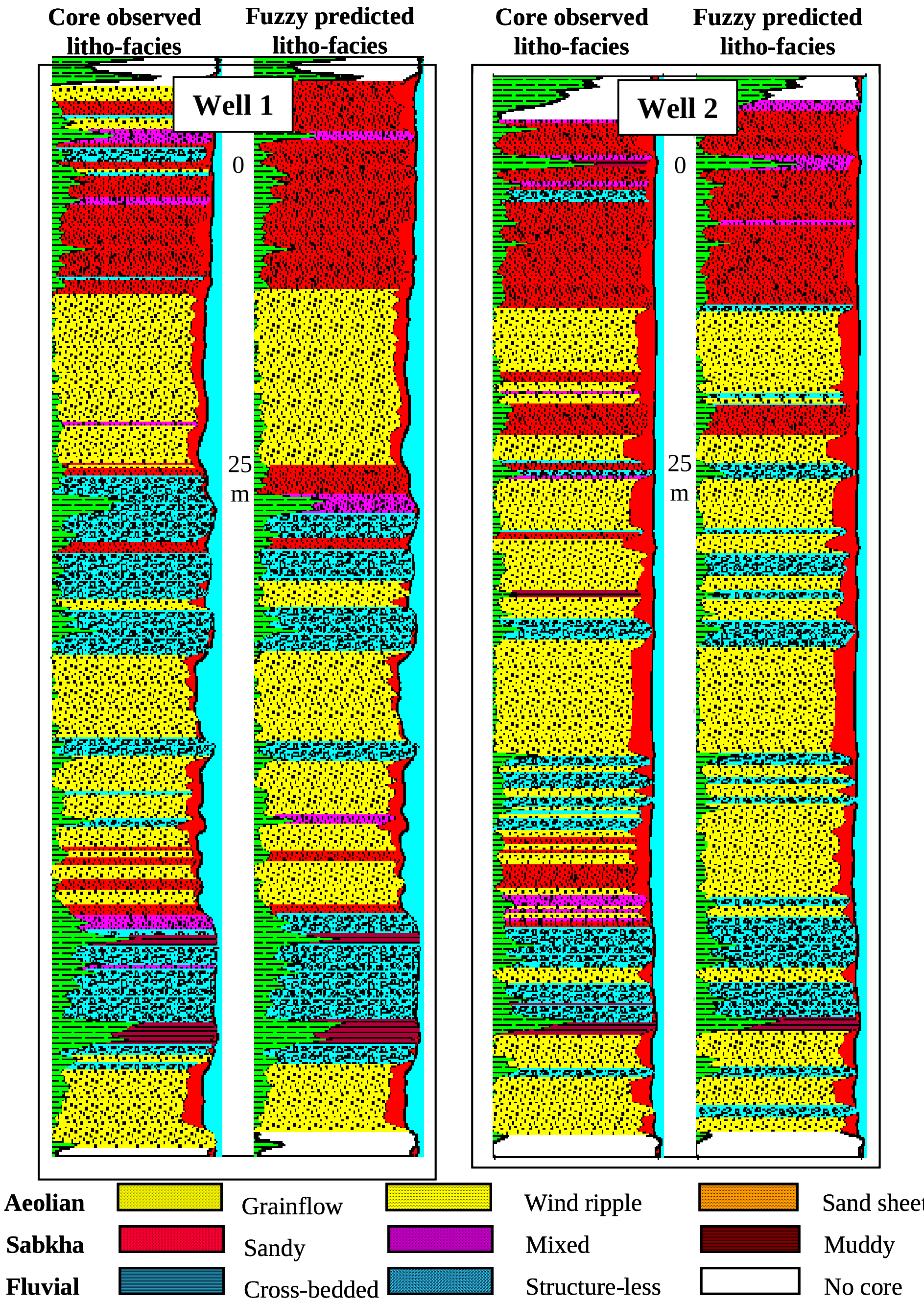


Fig. 3. Permeability bin determination.

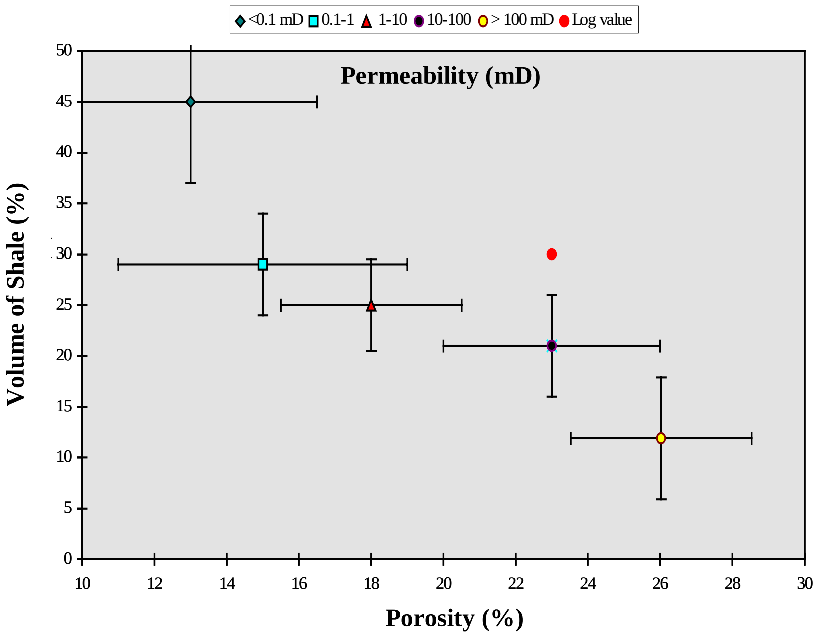
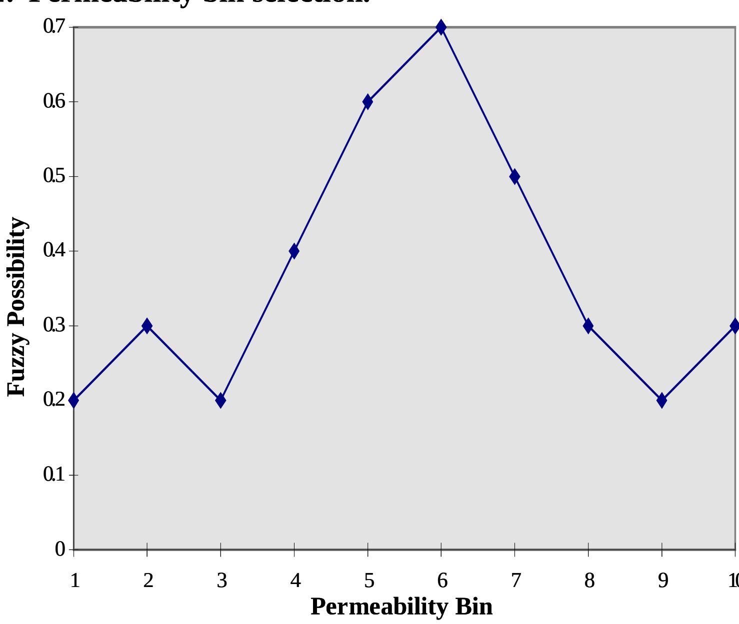


Fig. 4. Permeability bin selection.



Bin No.	mD
1	0.5
2	9.6
3	87
4	154
5	204
6	277
7	347
8	420
9	510
10	1320

Fig. 5. Permeability prediction in the Ula field.

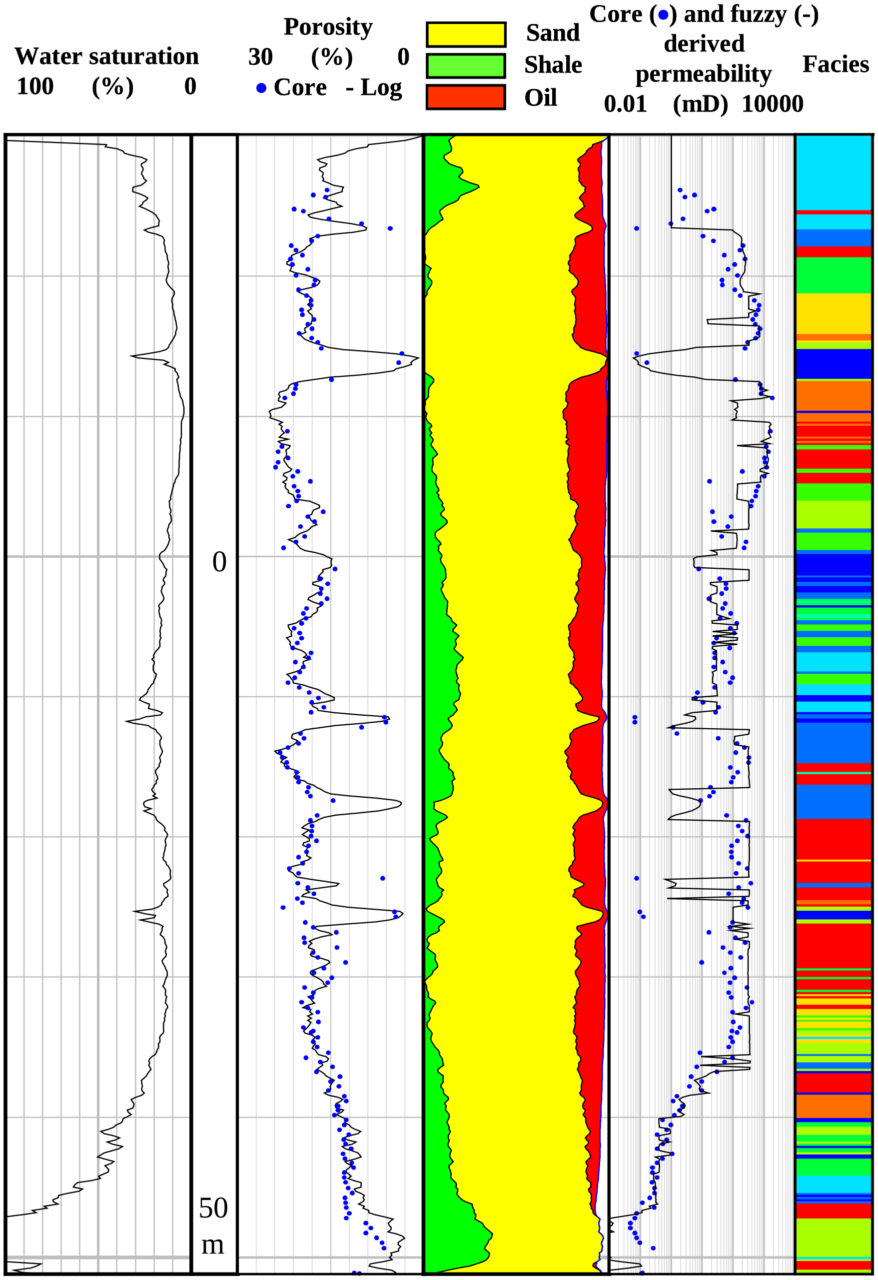


Fig. 6. Blind testing permeability prediction.

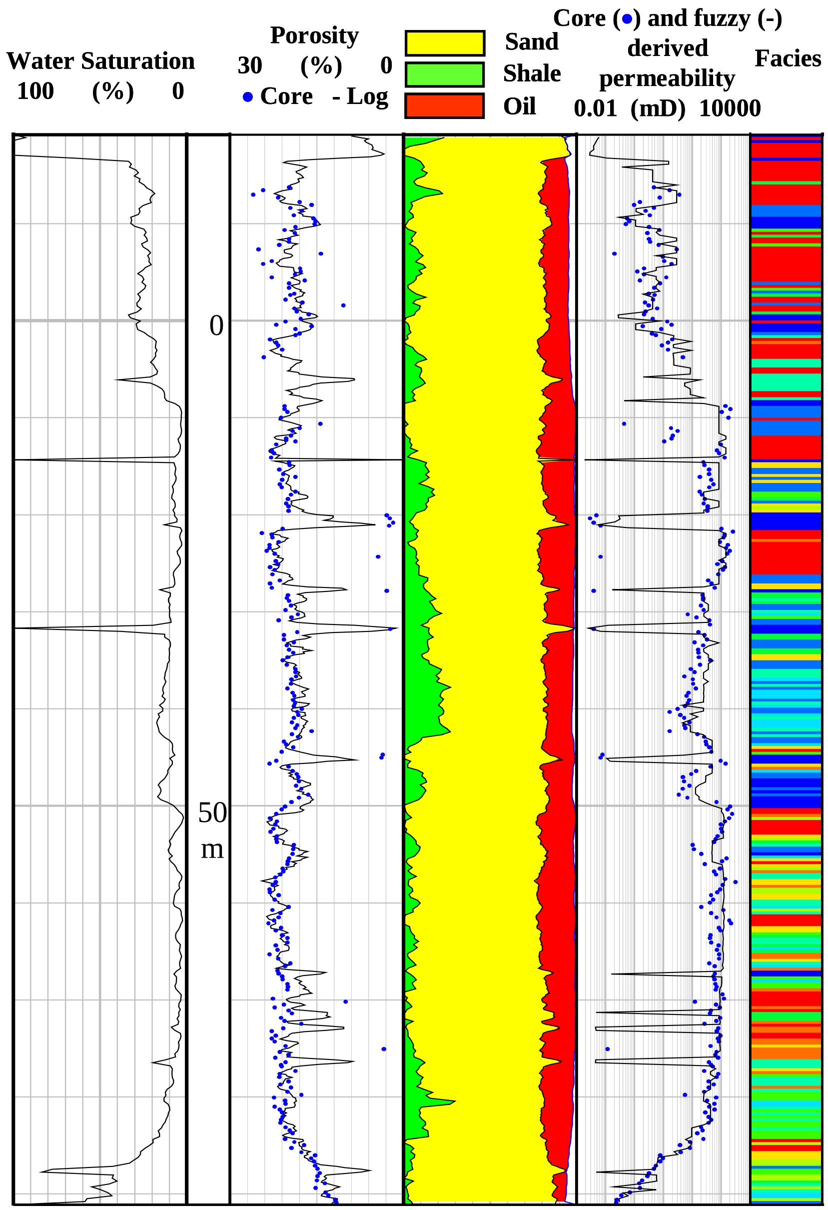


Fig. 7. The prediction of shear velocities.

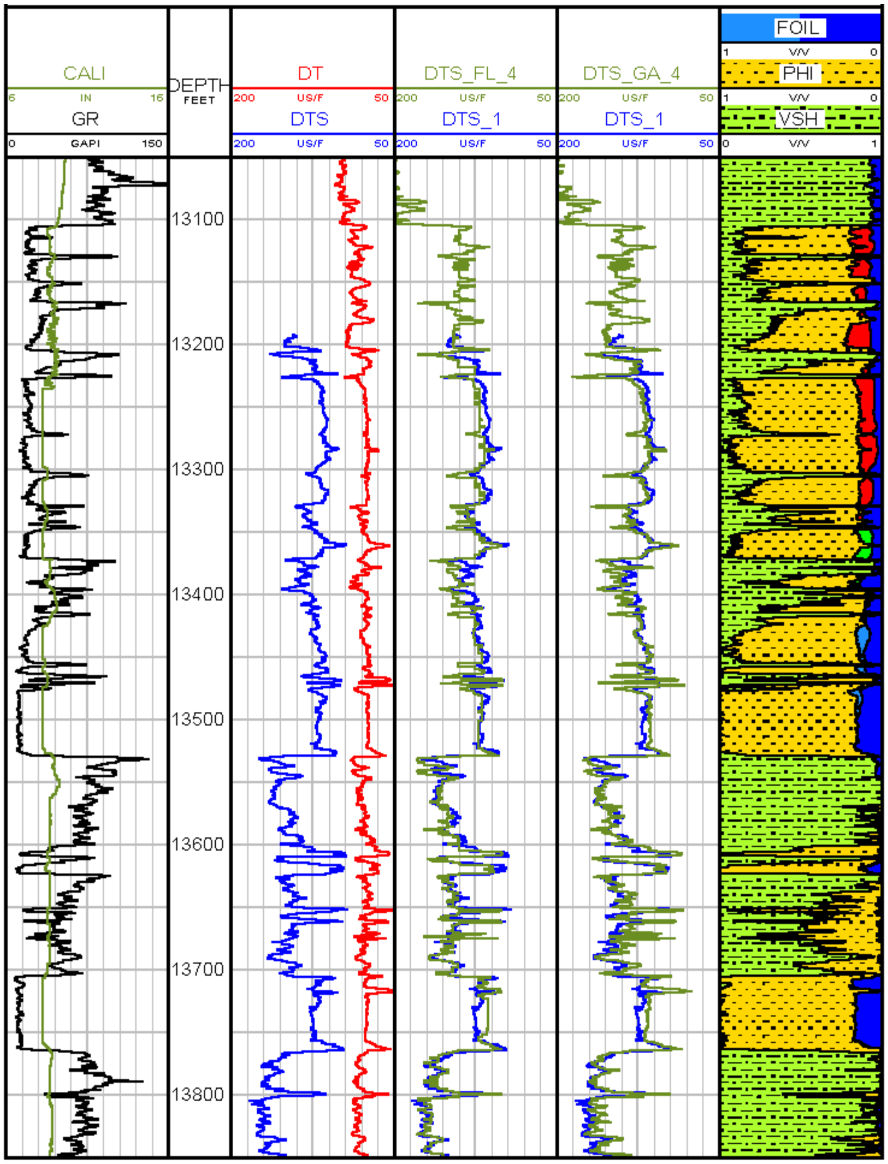


Fig. 8. Shear prediction using fuzzy logic.

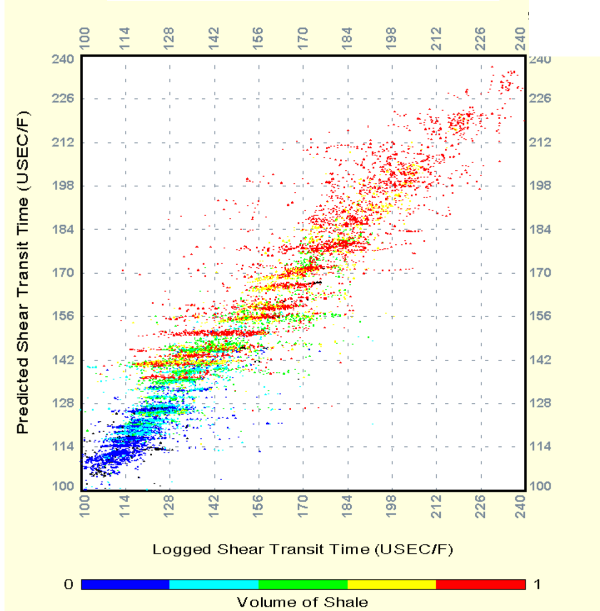
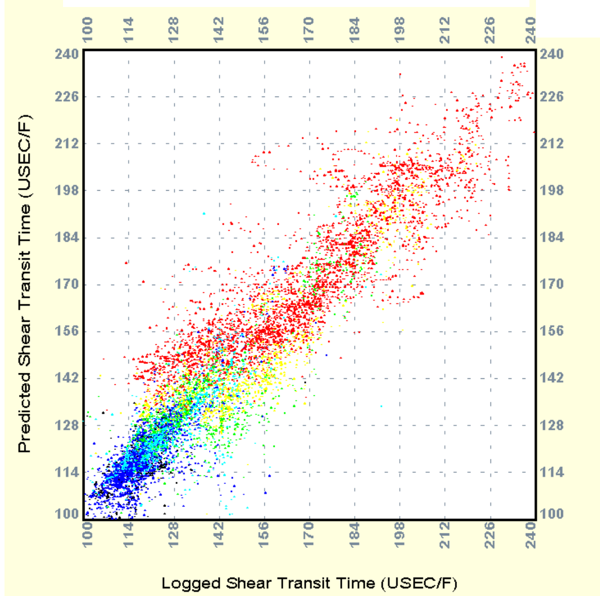


Fig. 9. Shear prediction using genetic algorithms.



Note. Transit time = 1/velocity