

Hybrid method for porosity classification in carbonate formations

Ildar Batyrshin^a, Leonid Sheremetov^{a,*}, Mikhail Markov^a, Alexandra Panova^b

^a*Mexican Petroleum Institute, Eje Central Lazaro Cardenas, 152, Col. San Bartolo Atepehuacan 07730, Mexico, D.F., Mexico*

^b*Kazan State Power Engineering University, Kazan, Russia*

Received 14 September 2004; accepted 29 November 2004

Abstract

Recently the methods of intelligent computing (IC) have been applied for the interpretation of well log data. This is due to the necessity to process well logs in the situations when complete information about them cannot be obtained. In this case, hybrid methods based on statistical analysis, fuzzy logic and evolutionary algorithms could be very useful. This paper presents such hybrid analysis of logging data of the wells from the Cantarell Oil Complex, in the Zonda of Campeche, Mexico. Different methods, such as principal component analysis, factor analysis, fuzzy classification and evolutionary optimization are used for analysis of the structure of porosity space given by primary, cavernous and micro-fractures porosity classes. Comparison and analysis of the obtained results show that IC methods can substantially compensate for the absence of exact information without losing the precision of data analysis and at the same time decrease the costs of well logging.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Petrophysics; Intelligent computing; Intelligent well log analysis; Fuzzy clustering; Secondary porosity

1. Introduction

The log interpretation process implies the solution of geophysical problems where all the available data and information are combined to in the attempt to reliably relate log measurement to the sought petrophysical parameters by using appropriate mathematical models. Several log interpretation techniques have been used to detect hydrocarbon-bearing zones

and to estimate their properties (Bassiouni, 1994; Huenges et al., 1997; Doveton, 2000; Verga and Viberti, 2002).

Many wells are logged during or after the drilling process. Logging data constitute thus an invaluable information source for predicting rock parameters in oil and gas reservoirs. The goal of this paper is to apply different intelligent computing (IC) methods for analyzing pore microstructure for carbonate formations based on minimal complex of logging measurements. This goal is motivated by the following reasons.

1. Realization of the total complex of logging measurements is very expensive. For cost reduc-

* Corresponding author.

E-mail addresses: batyr@imp.mx (I. Batyrshin), sher@imp.mx (L. Sheremetov), mmarkov@imp.mx (M. Markov), panova1@front.ru (A. Panova).

tion it is expedient to optimize the measurement complex and to exclude the measurements which are not indispensable at the specific conditions.

2. There exists a large data base of log measurements carried out in exploration holes many years ago which used a minor complex of methods than it is accepted now. Often a necessity of reprocessing of existing information about these exploration holes for reconsideration of previously made decisions arises.

Formation evaluation is meant to help providing the information to elaborate the most suitable technical and economical strategies for reservoir exploration. This paper presents an analysis of the logging data of the wells from the Cantarell Oil Complex, in the Zonda de Campeche, Mexico. The dispersion of the data can be taken into consideration using empirical relationships between pairs of parameters. Such covariance of parameters was calculated by the principal component and factor analyses (PCA and FA). Another purpose of this paper is to evaluate different IC technologies applied to the problem of secondary porosity classification.

Registration data from several wells were used in this study for porosity classification. Using the statistical and IC methods the classification of the rock according to the type of the pore space was realized. First, the method of description of porosity classes by fuzzy classes composed of several fuzzy granules is applied. Such granulation of classes gives a possibility to describe intersected porosity classes in the space of considered attributes. The second method applied for porosity classification uses separation of classes by planes in three-dimensional spaces of attributes. The evolutionary algorithm developed for construction of optimal planes is discussed.

The methodology of data mining consisting in several stages is described in Section 2. First, we apply traditional statistical methods such as principal component and factor analysis to determine the main components in the logging data set and importance of each factor (Section 3). Further, the main attributes are used for description of a porosity structure of log data in terms of fuzzy classes (Section 4). Finally, the description of porosity classes as separated areas in

attribute space is obtained by means of evolutionary algorithms aimed on obtaining the global optimum of optimization problem (Section 5). Results interpretation and future work are summarized in the concluding section.

2. Methodology of analysis

The quantitative estimation of the secondary porosity in carbonate formations using the integration of acoustic and electrical well log data was proposed by [Brie et al. \(1985\)](#). In this method the double porosity formation is presented as a homogeneous host rock (with primary porosity) that contains spherical secondary pores. The assumption of a low pore concentration was applied. As characteristics of the host rock Wyllie's relation for acoustic properties and Archie's law for electrical resistivity were accepted. The method was applied for the determination of oomoldic and vugular porosity in carbonate formations ([Brie et al., 1985](#); [Dutta et al., 1991](#)). Unfortunately, the assumption of the secondary-pore's form as spheres does not allow applying this method for fractured formations.

Recently, in the work by [Kazatchenko et al. \(2003\)](#), a technique of the secondary porosity estimation by acoustic core and well log data was presented. The elastic moduli of the double porosity medium are considered as functions of both primary and secondary porosities. The process of the secondary porosity estimation includes the determination of the matrix parameters (velocity–primary porosity relationship, grain and pore aspect ratios) and the inversion of the measured P-, S-wave velocities to find secondary pore aspect ratio and value. The inversion of the measured elastic velocities is performed for two variables: the concentration and aspect ratio of secondary pores. The aspect ratio of secondary inclusions indicates type of secondary-pore system. The cavernous are characterized by the aspect ratios 0.4–1 and the aspect ratios in the range 0.001–0.05 correspond to the micro-fractures. Classification results obtained with this method are used in this paper for data analysis (Section 3.1). Unfortunately, the measurements of S-wave velocity are not always available that

is why in this paper, another classification methods are developed.

In this paper, a standardized log data set was used for porosity classification. The following parameters: Delta-T Comp-P and S (DT4P), Bulk Density (RHOB), Thermal Neutron Porosity (Ratio Method) (NPHI), Gamma Ray (GR), MSFL Resistivity (MSFL) were considered. Other parameters such as Delta-T Shear (DTSM), Delta-T Shear-P and S (DT4S), for monopole and dipole acoustic sources, Bulk Density Correction (DRHO), Enhanced Vertical Resolution Density (NRHO), Density Porosity (DPHI), Corrected Near Thermal Counting Rate (CNTC), Thermal Neutron Porosity (TNPH), Enhanced Thermal Neutron Porosity (NPOR), Caliper (CALI), Corrected Far Thermal Counting Rate (CFTC), Sonic Porosity (SPHI), etc., were ignored because of their high correlation with the former ones or their low significance. As an example, the correlation between bulk density parameter and density porosity can be clearly seen from Fig. 1.

The methodology developed in this paper first uses traditional PCA in order to identify the most important

parameters. PCA and FA described in details in the following section, were applied to data using Statistical Toolbox from the MATLAB software (MATLAB, 2000). Later on, fuzzy classification and evolutionary optimization were used for porosity classification in parameter spaces. These methods of fuzzy classification and evolutionary optimization in analysis of log data porosity classes are developed based on fuzzy modelling, evolutionary computations and more traditional methods of cluster analysis and pattern classification (Anderberg, 1973; Batyrshin et al., 2002; Duda et al., 2000; Goldberg, 1989; Jang et al., 1997; Kosko, 1997; Klir and Yuan, 1995).

First, for porosity classification, the number of parameters was decreased to the following 5 parameters: p_1 —DT4P, p_2 —RHOB, p_3 —NPHI, p_4 —GR, p_5 —MSFL. Furthermore, we restricted the data set only to 50 m basis to apply the proposed methodology. The total number of 326 points of well log data was used containing respectively 59, 128 and 139 points from classes C_1 , C_2 and C_3 , respectively. These methods were studied on well logs from the same carbonate formation. The classification of porosity

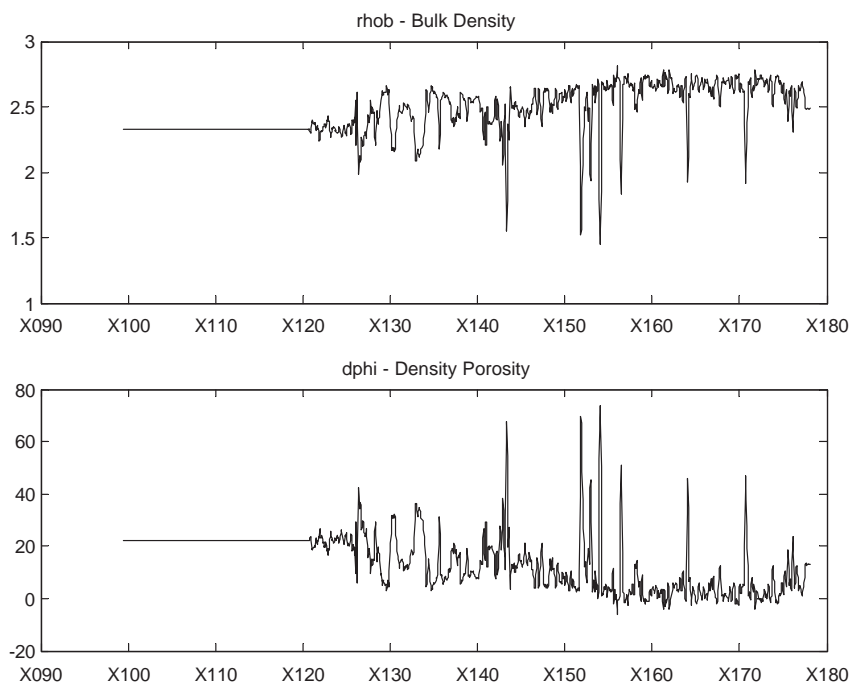


Fig. 1. Relationship between density porosity and bulk density.

classes was based on known classification of Cantarel XX74 log data on three porosity classes C_1 (cavernous porosity), C_2 (primary), C_3 (micro-fractures) shown in Fig. 2. This well log data were used as a training data while the log data from well XX95 with the known classification on porosity classes C_1 , C_2 and C_3 were used as testing data.

Fuzzy rules obtained by fuzzy granulation of porosity classes at the previous step, form the basis for the further classifications of the pore space for well logs. At the next stage, both based on the PCA/FA and fuzzy classification, only DT4P and GR are used for porosity classification as the most important parameters. Additionally, NPHI is added as a value of total porosity. Based on the analysis of the mutual location of points from porosity classes C_1 , C_2 and C_3 in well XX74, a hypothesis that classes C_1 and C_3 may be separated from other classes by planes in the 3D space was formulated and tested. A genetic algorithm is developed to generate the parameters of these planes used for final porosity classification.

In the following sections we describe each step of this methodology in more details.

3. Reduction of the parameter space

3.1. Principal component analysis

The PCA technique is used to establish the amount of original information contained in the log data. Principal component analysis examines the total variability represented in the data and describes this variability in terms of a set of factors. Each factor will account for a proportion of the original variability and will not be correlated with other factors to be found. Principal components are orthogonal. Generally, the bulk of the variability is described by fewer principal factors than there are variables. The fundamental theorem of factor analysis is given by the Eq. (1):

$$Z = A * P \quad (1)$$

with $Z=(i*j)$ matrix of standardized values:

$$z_{ij} = \frac{x_{ij} * x_{mj}}{\sigma_j} \quad (2)$$

of i cases (depth points) and j variables, $A=(j*f)$ factor loading matrix of j variables and f factors ($f \leq j$) and $P=(f*i)$ factor matrix of f (orthogonal) factor

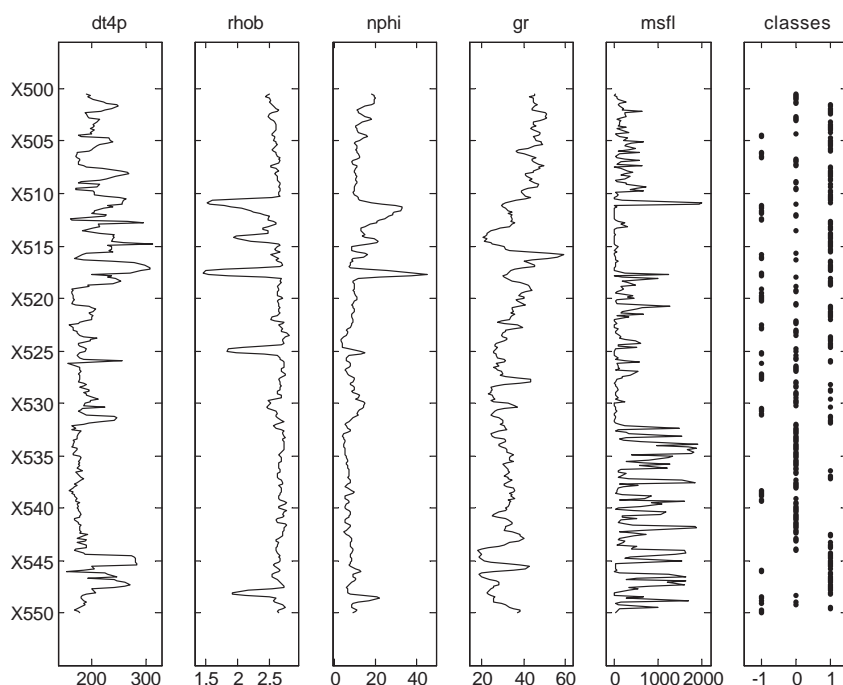


Fig. 2. Parameter fluctuations and corresponding porosity classification (–1: cavernous, 0: primary, 1: micro-fractures classes).

scores at i depth points. In Eq. (2) x_{ij} denote measured values of variable j at depth i , σ_j is a standard deviation of variable j and x_{mj} is a mean value of variable j of the complete data set. The mean of the standardized value is zero and the standard deviation equals to one.

To achieve a sufficient common data base for principal component analysis, the data were normalized. This is appropriate because the measurements are made on different scales. From this data set, three independent initial factors were extracted which explain about 80% of the variability of the given data set.

Table 1 shows the percent of variability of the given data set explained by factors. As it can be seen, 45% of the variability is explained by factor 1, 20% by factor 2, and 14% by factor 3. After the factors have been gained they are usually transformed into ‘simple structure’ to render easier interpretation of their significance. Each principal component is a weighted sum of all the variables.

Table 2 shows the factor loading. The communality, the variance of single variables, which are represented by the 3 factors, varied from 49% (Bulk Density) to 75% (Delta-T Comp-P and S).

The factor loadings quantify the relationship between parameters and the factors (see Eq. (1)). They are either positively or negatively correlated. High values of factor loadings indicate also highly correlated loading variables. The first order background variable factor 1 (explaining more than 45% of the variability) is loaded by Bulk Density, which coincide with high values of Thermal Neutron Porosity (Ratio Method) (NPHI) and Corrected Near Thermal Counting Rate (CNTC). CNTC is positively correlated with the Bulk Density. Factor 2 is loaded

Table 2

The first three principal component vectors

Parameter	1	2	3
dt4p	0.3051	0.1823	0.7565
Rhob	−0.4886	−0.2630	0.0873
CNTC	−0.5662	0.0224	0.1713
Nphi	0.5514	−0.0182	−0.0721
Gr	0.0547	−0.7296	0.4852
Msfl	−0.2013	0.6037	0.3875

positively with MSFL Resistivity and negatively with Gamma Ray. Factor 3 is loaded by Delta-T Comp-P and S (DT4P).

Rock classification according to the structure of porosity space: primary, cavernous porosity, micro-fractures was used to analyze the obtained results. As mentioned above, these classes were defined using classical full-wave form acoustic data processing (Kazatchenko et al., 2003). As shown in Fig. 3, this classification can be clearly seen in three-dimensional space of the principal components.

Visual analysis of Fig. 3 shows the possibility to separate three classes in the space of three principal components by three-dimensional planes. This idea will be used later on in Section 5 for porosity classification. On the other hand, from Table 2 it can be seen that these components are defined by DT4P, GR and NPHI/CNTC, respectively.

3.2. Factor analysis

The same data set was also processed with FA. As a matter of fact, both PCA and FA are dimension-reduction techniques, in the sense that they can be used to replace a large set of observed variables with a smaller set of new variables. In this paper, we use PCA to summarize and approximate log data using fewer dimensions (or to visualize it, as shown in Section 3.1), while in the explanatory model of correlation among log data, FA should be used.

Since the estimated loadings from an unrotated FA fit can have a complicated structure, we shall analyze the factors after rotation. The goal of factor rotation is to find a parameterization in which each variable has only a small number of large loadings, i.e. is affected by a small number of factors, preferably only one. Factor loadings after rotation are shown in Table 3.

Table 1

The percent of the total variability explained by each principal component

Factors	Variances	Percent explained	% explained by first K factors
1	2.7310	45.5160	45,5160
2	1.2038	20.0629	65,5789
3	0.8388	13.9799	79,5588
4	0.7623	12.7043	92,2631
5	0.3726	6.2108	98,4739
6	0.0916	1.5260	99,9999

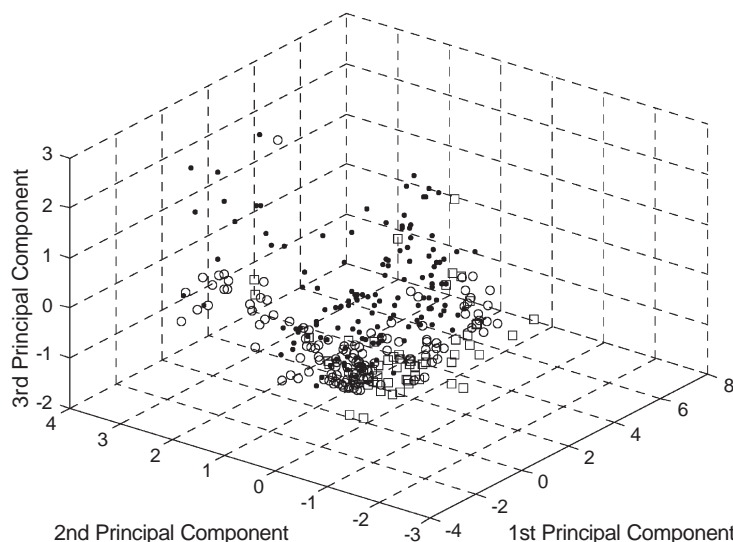


Fig. 3. Log data projected onto the first three principal components. Porosity classes: \square —cavernous, \circ —primary, \bullet —micro-fractures.

The rows of the loadings matrix can be considered as coordinates in M-dimensional space, and then each factor corresponds to a coordinate axis. Factor rotation is equivalent to rotating of these axes and to computing new loadings in the rotated coordinate system. For this example, we rotate the estimated loadings by using the promax criterion, a common oblique method.

A specific variance (Table 4) of 1 would indicate that there is no common factor component in that variable, while a specific variance of 0 would indicate that the variable is entirely determined by common factors. That means that Delta-T Comp-P and S (DT4P), and Gamma Ray (GR) are the most important parameters for the future analysis (Section 5). This result is completely justified by the fact that sonic travel time is the most sensible to the porous space structure.

4. Fuzzy granules for porosity classification

The visualization of data obtained in the space of the first three principal components (Fig. 3) shows that the three porosity classes are partially mixed and partially separated and may be considered as mutually intersected fuzzy classes. We look for the description of these porosity classes C_1 , C_2 and C_3 by fuzzy classes G_1 , G_2 , G_3 as follows. Each fuzzy class G_i was considered as a union of three fuzzy granules: $G_i = G_{i1} \cup G_{i2} \cup G_{i3}$. The fuzzy granules G_{ij} were obtained as a result of intersection of correspondent membership functions MF_{ijk} defined on the domains of parameters p_k . We used generalized bell membership functions (Jang, 1997) $MF(x) = 1/(1 + ((x-c)/a)^{2b})$ determined by three parameters (a, b, c) . The total number of MF parameters for all fuzzy sets was equal to $3 \times 3 \times 5 \times 3 = 135$.

Table 3
The first three factor vectors after rotation

Parameter	1	2	3
dt4p	0.2872	−0.0224	−0.1596
rhob	−0.6346	0.0088	0.2516
CNTC	−0.9588	0.1028	−0.1394
nphi	1.1743	0.1088	0.4204
gr	0.2159	−0.0823	0.4836
Msfl	0.0189	0.9927	−0.0781

Table 4
Specific variances

Parameter	Specific variances
dt4p	0.8422
Rhob	0.3692
Cntc	0.1344
Nphi	0.0050
Gr	0.8040
Msfl	0.0050

The criterion of good classification was defined as follows. For each point x which belongs to class C_i in training data set the separability value was calculated as follows: $Q(x) = MF_i(x) - \max(MF_k(x), MF_j(x))$, where $MF_i(x)$, $MF_k(x)$ and $MF_j(x)$ are membership values of object x in fuzzy classes G_i , G_k , G_j ($k, j \neq i$). The higher is the value $Q(x)$, the better is the classification of the object x . For each class C_i the average value of separability Q_i of its points was calculated. General criterion of classification was equal to $Q_s = Q_1 + Q_2 + Q_3$. The maximization of this criterion was used to find optimal parameters of membership values. Optimization algorithm of constraint nonlinear minimization from MATLAB Optimization Toolbox was applied for classification of given 326 points of well log data. Really, because the Matlab optimization procedures search the minimum of the function, we used for optimization $F = -Q_s$ criterion. For training data set, 90% of points from each class were randomly selected. Other points were used in testing data set. Finally, fuzzy rules of the following type were generated based on the obtained granular classification of data:

If (p_1 is MF_{i11} and ... p_5 is MF_{i15})
 or (p_1 is MF_{i21} and ... p_5 is MF_{i25})
 or (p_1 is MF_{i31} and ... p_5 is MF_{i35})
 then data point is C_i (3)

where p_1 —DT4P, p_2 —RHOB, p_3 —NPHI, p_4 —GR, p_5 —MSFL.

The true classification on training data set was equal to 74%, 65% and 88% for classes C_1 , C_2 and C_3 , respectively. This classification on testing data set was equal to 33%, 62% and 86%, respectively. The classification of porosity classes C_2 (primary) and C_3 (micro-fractures) may be considered as sufficiently good taking into account the high mixture of classes in some regions of the parameters space. Unfortunately, an extension of the obtained results on another well was problematic. For this reason, other models based on fuzzy granulation of classes were studied.

First, we changed the criterion of classification. Instead of the average value of separability Q_i of the points for each class C_i , simply the sum of these values was calculated. For fuzzy classes obtained by new optimization procedure the probabilities of true classification of points by fuzzy classes G_1 , G_2 , G_3

were equal to 0.79, 0.80 and 0.75, respectively. These probabilities mean that, i.e. the point randomly selected from the class G_1 belongs to the class C_1 with the probability 0.79 and so on, because 79% of points from G_1 belong to the class C_1 , etc. The obtained fuzzy classes G_1 , G_2 , G_3 cover, respectively, 71%, 64% and 93% of points from the classes C_1 , C_2 and C_3 . For the constructed fuzzy classes the following probabilities of true classification were obtained on testing data from the well XX95: 0.47, 0.55 and 0.73 for porosity classes C_1 , C_2 and C_3 , respectively.

The analysis of obtained granules composing the fuzzy classes G_1 , G_2 , G_3 shows that really only two granules may construct each of these classes because one granule in each fuzzy class does not contain points or contain only few points. So, in the next model we looked for more simple representation of porosity classes by two granules $G_i = G_{i1} \cup G_{i2}$. As initial fuzzy granules for each fuzzy class in new optimization procedure, we used the best two granules obtained by anterior optimization procedure for three granules model. These initial fuzzy granules were used in considered above optimization problem with optimization criterion based on the sum of separability values Q_i . The new optimal fuzzy classes G_1 , G_2 , G_3 each composed of two fuzzy granules were obtained. These classes cover, respectively, 75%, 65% and 90% of points from the classes C_1 , C_2 and C_3 with the probabilities of correct classification 0.73, 0.76, and 0.80, respectively.

The obtained results are comparable with the results of fuzzy classification based on the use of three granules in each fuzzy class but the description of the fuzzy classes now is more simple because each of them are based only on two granules. The membership functions obtained by optimization procedure for each fuzzy granule are shown in Figs. 4–6. The testing of the obtained classes on the log data from well XX95 gives the following probabilities of true classification: 0.55, 0.52 and 0.74 for the classes C_1 , C_2 and C_3 , respectively. These testing results are better than testing results obtained for fuzzy classification based on three granules. We should note that in both cases we have only some locally optimal solutions of optimization problem.

The transfer of fuzzy granules constructed on training data of XX74 well and testing log data of

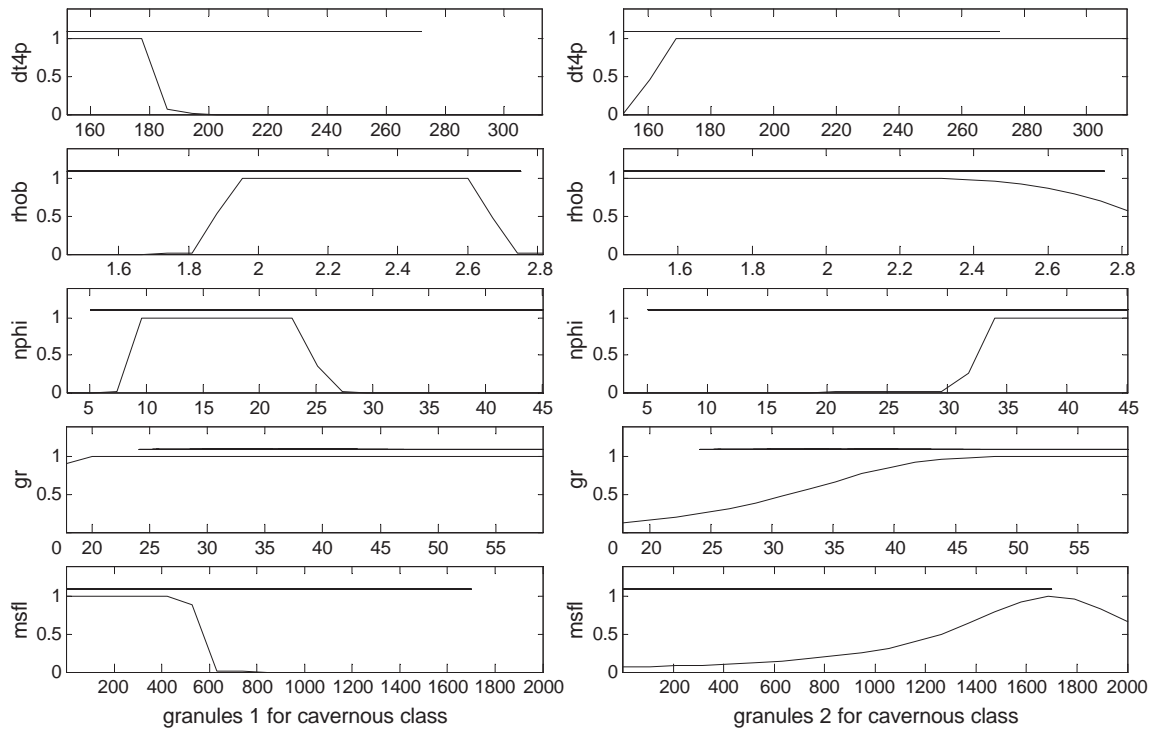


Fig. 4. Membership functions defined on the domains of parameters p_1 – p_5 for the granules G_{11} (left) and G_{12} (right) of the cavernous porosity class G_1 . The lines over membership functions show the domains of parameter values.

XX95 well shows that these log data have different domains for some parameters. Due to such difference some fuzzy granules obtained on the border of the domain of the training data and which can be retranslated as “greater than N_1 ” after transfer on the more extended domain of XX95 log data are presented as fuzzy intervals with interpretation “greater than N_1 but less than N_2 ”. Fig. 7(a,b) shows such situation taking place for fuzzy set defined on the domain of GR parameter.

On the domain of training data this fuzzy set may be roughly translated as “approximately greater than 35” but the same fuzzy set on the domain of testing data can be translated as “approximately between 35 and 75”. It was supposed that a correction of fuzzy sets obtained on training data which is preserving the same linguistic interpretation on the domain of testing data will increase the classification capability of fuzzy solution. For this reason such fuzzy granules obtained after optimization procedures on training data were artificially extended in the direction of the borders of the domain of XX95 log data to give the similar

interpretation “greater than N_1 ” on the domain of XX95 log data. Corrected in such a way, the membership function for GR parameter is shown in Fig. 7(c).

After testing of the obtained corrected fuzzy classification on the well XX95 log data, the following probabilities of true classification were obtained: 0.60, 0.52 and 0.75 for the classes C_1 , C_2 and C_3 , respectively. These testing results are slightly better than results obtained for fuzzy classification with non-corrected fuzzy granules.

The final description of fuzzy classes is given by the rules of the following type:

$$\begin{aligned} &\text{If } (p_1 \text{ is } MF_{i11} \text{ and } \dots p_5 \text{ is } MF_{i15}) \\ &\text{or } (p_1 \text{ is } MF_{i21} \text{ and } \dots p_5 \text{ is } MF_{i25}) \\ &\text{then data point is } C_i. \end{aligned} \quad (4)$$

Fuzzy granules can describe not just crisply separated classes existing in many practical situations. The results obtained by fuzzy granulation of porosity classes describe a fuzzy nature of these porosity

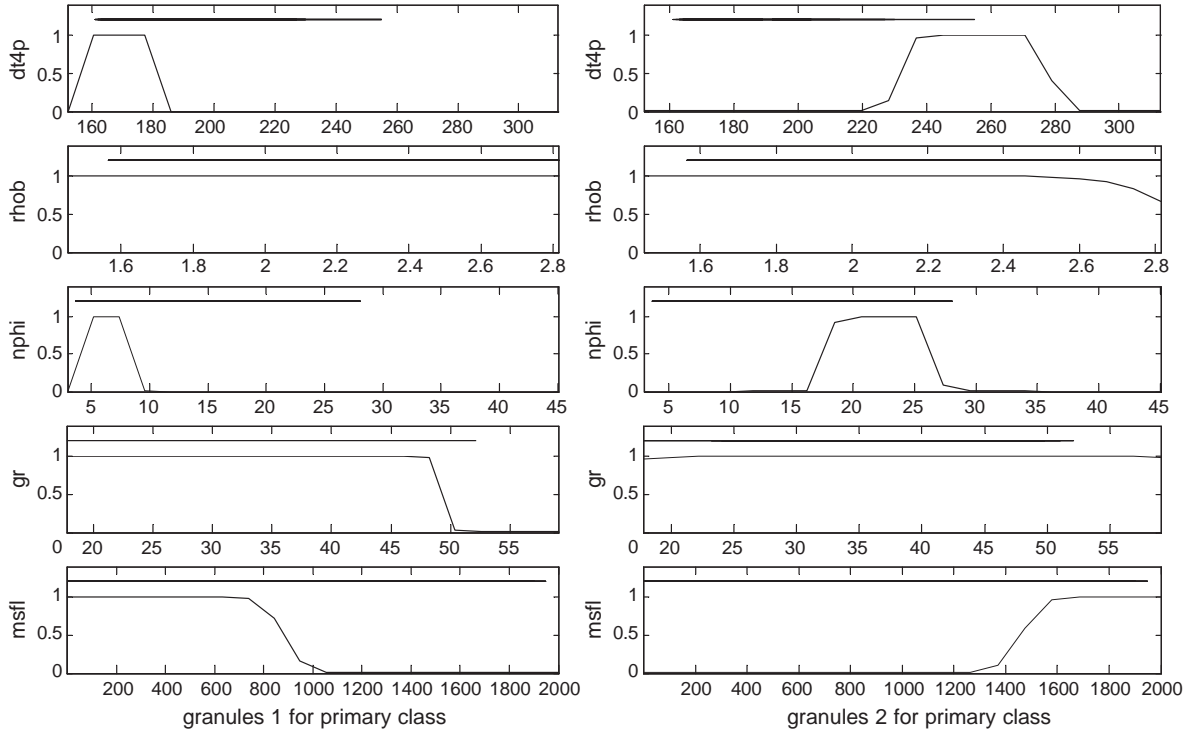


Fig. 5. Membership functions defined on the domains of parameters p_1 – p_5 for the granules G_{21} (left) and G_{22} (right) of the primary porosity class G_2 .

classes in considered set of parameters. The obtained fuzzy classification gives sufficiently good classification of micro-fractures porosity class in carbonate formations, which is the most important porosity class for oil reservoir exploration. These rules form the basis for the further classifications of the pore space based on well logs.

5. Evolutionary algorithm for porosity classification

Based on the PCA and FA results (Table 2 of Section 3.1 and Table 4 of Section 3.2), DT4P and GR are used for further porosity classification as the most important parameters. Additionally, NPHI is added as a value of total porosity. We looked for 2 planes separating the points of the porosity classes from the points of other classes. As a criterion of good separability for each plane the maximization of the difference between the numbers of correctly and incorrectly classified points is considered. The cases

when the points of corresponding class are located above and below the plane were considered. The following equations of planes were used: $z_3 = a_3x_1 + b_3x_2 + c_3$ for class C_3 and $z_1 = a_1x_1 + b_1x_2 + c_1$ for class C_1 where x_1 and x_2 are the values of the parameters X_1 and X_2 . For example, for the class C_3 located below the plane z_3 the maximized criteria had the form

$$F_3 = N_3 - (N_1 + N_2),$$

where N_1 , N_2 and N_3 are the numbers of points from classes C_1 , C_2 and C_3 , respectively, located below the plane. The similar criterion was considered for the second plane:

$$F_1 = N_1 - (N_2 + N_3),$$

where N_1 , N_2 and N_3 are the numbers of points from classes C_1 , C_2 and C_3 , respectively, located above the plane. The point with coordinates (x_1, x_2, x_3) is located below the plane z_3 if $x_3 < a_3x_1 + b_3x_2 + c_3$ and this point is located above the plane z_1 if $x_3 > a_1x_1 + b_1x_2 + c_1$. An evolutionary algorithm was developed to find

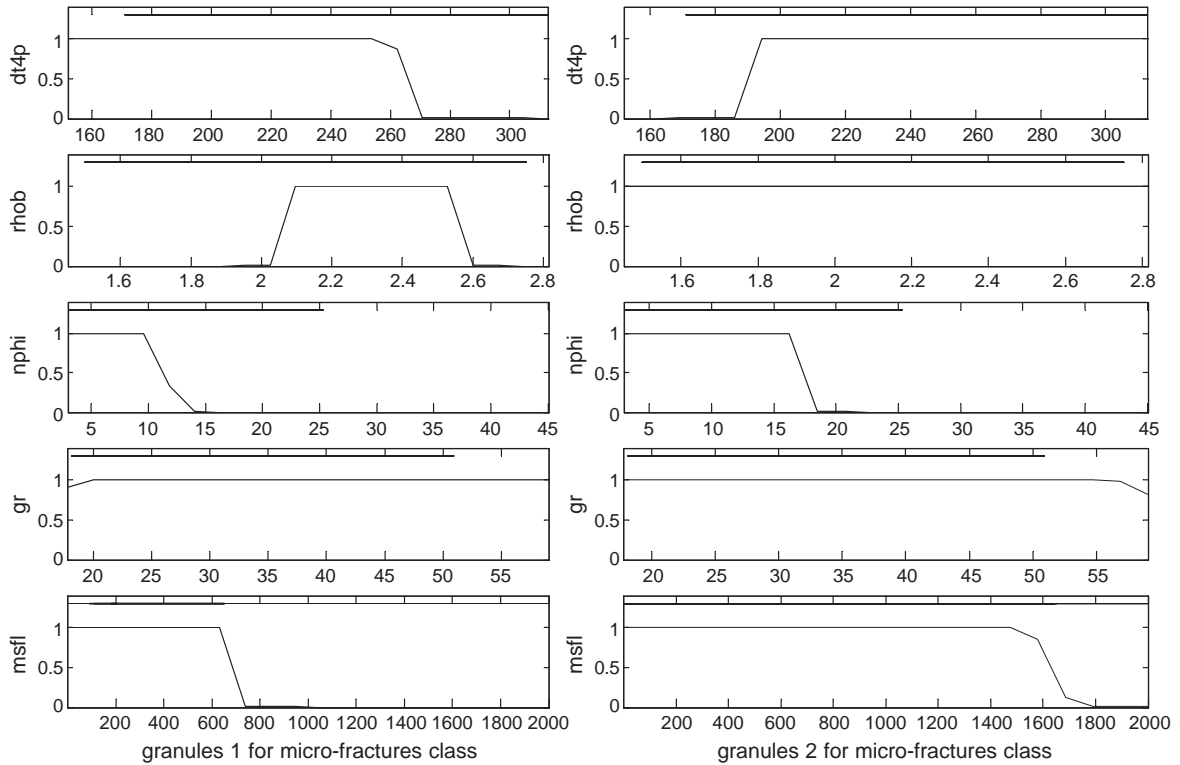


Fig. 6. Membership functions defined on the domains of parameters p_1 – p_5 for the granules G_{31} (left) and G_{32} (right) of the micro-fractures porosity class G_3 .

optimal parameters of separating planes. The optimal parameters (a_1, b_1, c_1) and (a_3, b_3, c_3) of planes z_1 and z_3 were found independently and in a similar way.

Denote $[x1min, x1max]$, $[x2min, x2max]$, $[x3min, x3max]$ the domains of parameters x_1 , x_2 and x_3 , respectively. These three intervals define the cube of

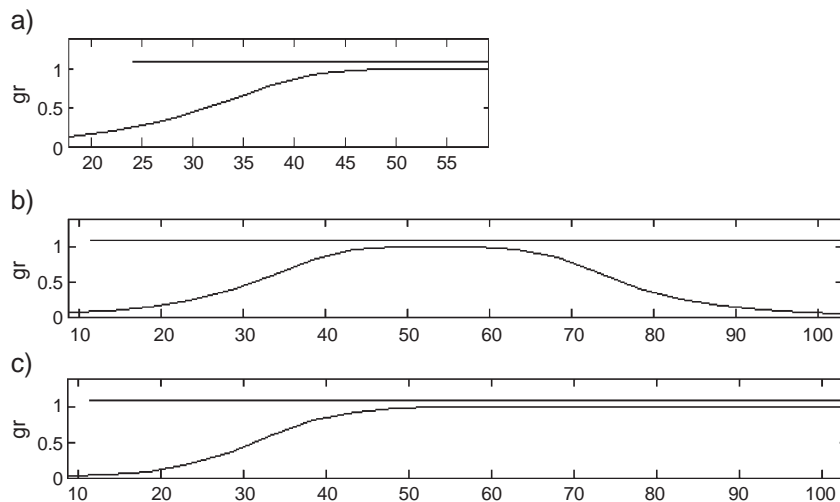


Fig. 7. Example of correction of membership functions for granule G_{12} transferred from the domain of training log data a) to the domain of testing log data b) and corrected on this domain c) to preserve the same linguistic interpretation “greater than approximately 35” as in a).

possible values (x_1, x_2, x_3) . For generating initial population for genetic algorithm, the parameters (a, b, c) of 12 different planes crossing three vertexes of this cube (except the planes orthogonal to the plane $X_1 \times X_2$) were calculated. For example, the parameters (a, b, c) of the plane passing through the vertexes $v_1=(x1min, x2min, x3min)$, $v_2=(x1max, x2max, x3min)$, $v_3=(x1min, x2max, x3max)$ were obtained as a solution of the system of three equations:

$$x3min = a*x1min + b*x2min + c,$$

$$x3min = a*x1max + b*x2max + c,$$

$$x3max = a*x1min + b*x2max + c.$$

These planes are used for generation the initial population of planes because with high probability some of them are crossing the set of given points and as result, they can give more or less good approximation of the optimal solution. The parameters (a, b, c) of these 12 planes were used for generation of new triplets of parameters in initial population of genetic algorithm as follows. From each triplet (a, b, c) , $K_1=100$ triplets of parameters were generated by multiplying their values by the values M_1*r , where $M_1=500$ and r chosen from uniform distribution on $[0,1]$. For each triplet (a, b, c) from initial population, the value of criteria F_3 (or F_1) called a fitness function was calculated and $K_2=20$ triplets with the maximal value of this criterion were selected in elite. From this elite, a new population was generated, and K_2 best triplets from this population were selected in a new elite, etc. $K_3=20$ generations of such new populations were used.

From each elite a new population was generated by crossover and mutation operations as follows. For two “parent” triplets (a_1, b_1, c_1) and (a_2, b_2, c_2) selected from elite a convex combination of parameter values $a=a_1*r+a_2*(r-1)$ is applied, where r is chosen from uniform distribution on $[0,1]$. Similarly the parameters b and c are generated. The parameters were used for generation of new triplets (a, b_1, c_1) , (a_1, b, c_1) , (a_1, b_1, c) , (a, b_2, c_2) , (a_2, b, c_2) , (a_2, b_2, c) , (a, b, c) . These “children” triplets are generated for all combinations of two “parents” selected from elite. In more traditional approaches to evolutionary and genetic algorithms, these parents are selected with some probability; moreover the elements of populations

are presented by binary code (Goldberg, 1989). These “children” and elite triplets constitute a first part of a new population. The second and the third parts of new population are obtained as a result of mutation of the first part, more exactly, to the values of the triplets (a, b, c) from this part mutation values M_2*r and M_3*p are added respectively, where $M_2=100$, $M_3=1$, and r, p are chosen from a normal distribution with zero mean, variance and standard deviation equal to one. The large mutation value is used for the extension of the search area and the small mutations are used for fine tuning of the good solutions. Each new population contains $\{K_2+K_2*(K_2-1)*0.5*7\}*3=K_2*(10.5*K_2-7.5)$ triplets containing previous K_2 elite triplets and $K_2*(10.5*K_2-8.5)$ new triplets. The total number of generated triplets equals to $12*K_1+K_3*K_2*(10.5*K_2-8.5)$ triplets. For selected values of parameters it equals to 81,800 generations of new triplets.

The values of parameters of the plane z_3 separating micro-fractures class C_3 obtained by evolutionary algorithm are as follows: $a=5.7246$, $b=-18.0$, $c=-895.0$. The value of criterion F_3 equals to 115. The points (x_1, x_2, x_3) located below this plane, i.e. satisfying inequality

$$x_3 < 5.7246*x_1 - 18*x_2 - 895$$

are classified in C_3 class. This rule covers 88.5% of points from XX74 well log data for class C_3 and probability of true classification with this rule equals to 0.94, i.e. randomly selected object of XX74 well log data lying below this plane with the probability 0.94 will belong to the class C_3 . Fig. 8 shows this plane in three-dimensional space of parameters DT4P, NPHI and GR. Fig. 9 shows the curve of fitness function values obtained during the search of plane separating micro-fractures porosity class from other porosity classes. The optimal triplet was found after 7 generations of evolutionary algorithm.

The values of parameters of the plane z_3 separating cavernous class C_1 obtained by evolutionary algorithm are: $a=9.8701$, $b=-25.4205$, $c=-1388.3$ with the value of fitness function $F_1=32$. The points (x_1, x_2, x_3) located above this plane such that

$$x_3 > 9.8701*x_1 - 25.4205*x_2 - 1388.3$$

are classified in C_1 class. This rule covers 63% of cavernous porosity class given in training log data with probability $p=0.88$ of true classification. Fig. 10

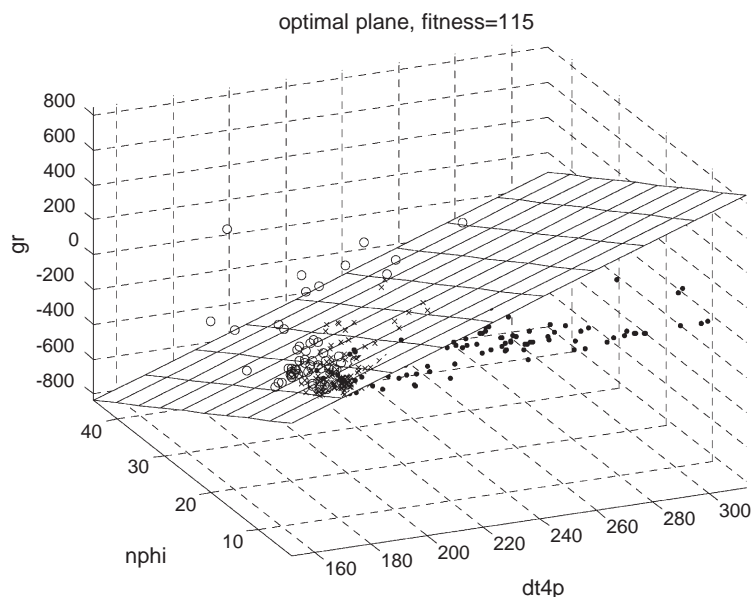


Fig. 8. The optimal plane separating the micro-fractures porosity class (points ● located under the plane) from cavernous (○) and primary (×) porosity classes in training log data.

shows this plane in three-dimensional space of parameters DT4P, NPHI and GR.

The obtained classification rules were checked on well XX95 log data. The first separation plane covers 87% of micro-fractures porosity class with the probability 0.83 of true classification. The results are

presented in Fig. 11. The second separation plane covers 42% of cavernous porosity class with the probability 0.83 of true classification. The results are presented in Fig. 12.

6. Discussion and conclusions

Carbonate formations are characterized by a complex pore microstructure that includes a primary porosity (matrix pores) and a secondary porosity (cavernous and micro-fractures) (Archie, 1952; Lucia, 1999). In these formations the determination of the type and value of secondary porosity has a significant influence on the correct permeability prediction and evaluation of hydrocarbon reserves. This paper describes a methodology based on the use of hybrid methods, such as principal component and factor analysis, fuzzy classification and evolutionary optimizations for analysis of well logs and for qualitative pore structure classification in carbonate formations. Aminzadeh et al. (2000) use a similar methodology of principal component analysis before a neural network transformation for reservoir parameter estimation. Nikraves and Aminzadeh (2001a,b) and Tamhane

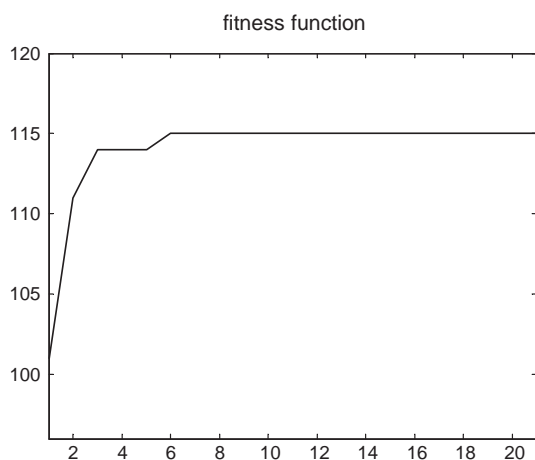


Fig. 9. The curve of the fitness function values obtained during the search of the plane separating micro-fractures porosity class from other porosity classes.

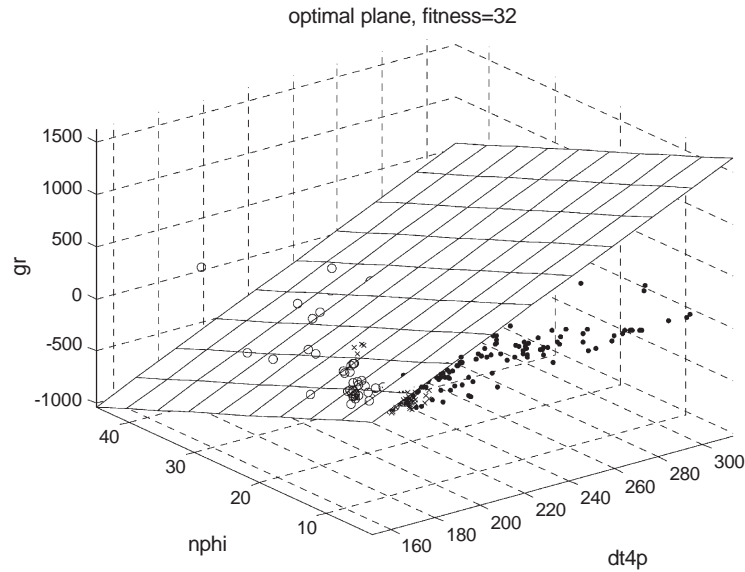


Fig. 10. The optimal plane separating the cavernous porosity class (points (O) located above the plane) in training log data.

et al. (2002) also use similar approaches for similar problems.

Although classification is a traditional approach in geophysics, many ambiguities exist in finding exact classifications. In a mathematical or statistical environment a value may or may not belong to one class.

In porosity classification problem there are usually imprecise conditions, some measurements are missing and therefore fuzzy methods seem to be more suitable than crisp ones. Fuzzy techniques (Jang et al., 1997; Batyrshin et al., 2002; Nikraves et al., 2003) may facilitate standardization of classification routines and

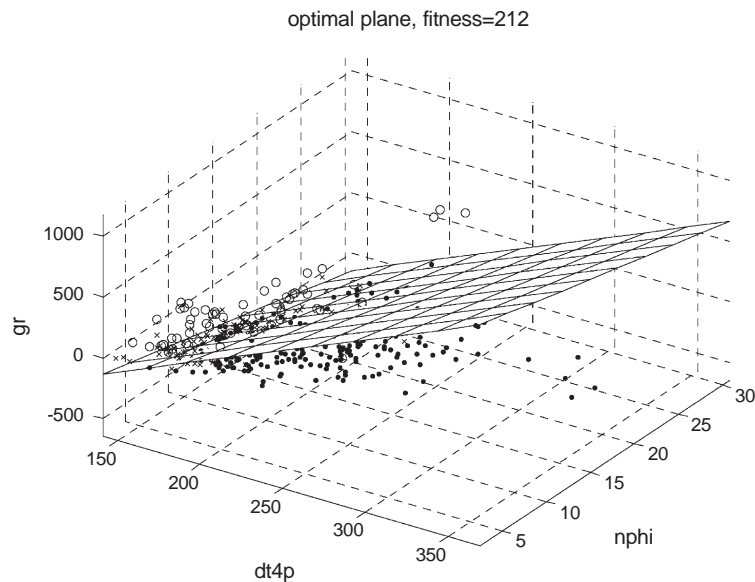


Fig. 11. The optimal plane separating micro-fractures porosity class (points (●) below the plane) in the testing log data.

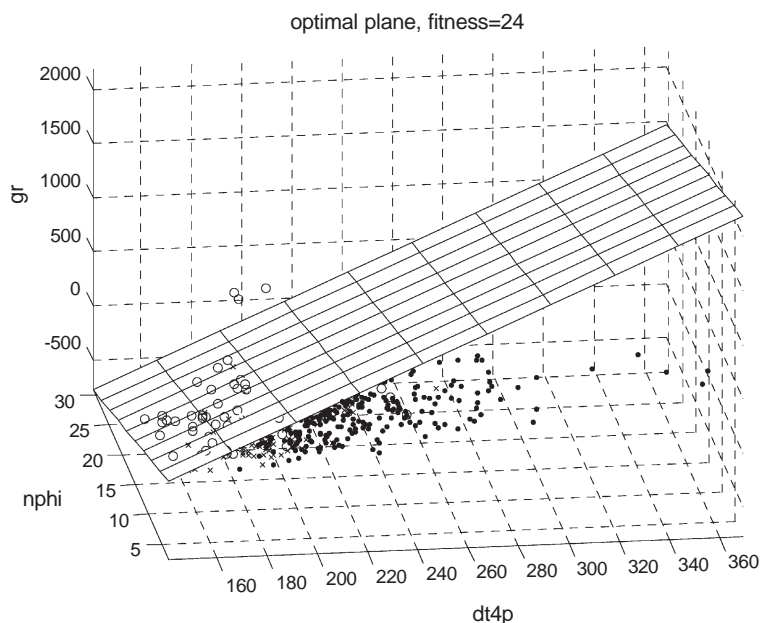


Fig. 12. The optimal plane separating cavernous porosity class (points (O) above the plane) in the testing log data.

hence can be important supportive tools for the geophysical engineer in practice as well as valuable tools in quality control and training. This classification is useful under the conditions of a deficiency of the information, for example during data processing about the old boreholes. For such boreholes there is no available data for many contemporary methods of logging. Therefore it is important to estimate the petrophysical properties of rocks under the conditions of incomplete information.

The absence of measurements or absence of reliable data often leads to inadequate understanding of reservoir behaviour and to poor performance prediction. Risk and uncertainty assessment using intelligent computing techniques have become the major issues in reservoir characterization and improved oil recovery (Wong and Nikraves, 2001; Aminzadeh and Wilkinson, 2004; Wong et al., 2002). Different types of artificial neural networks (ANN) have been used for reservoir characterization and also for well log interpretation (Pezeshk et al., 1996). Recently fuzzy methods also found wide application in petroleum engineering to handle the uncertainty, imprecision and errors in facies classification (Wong et al., 1997; Finul et al., 2001a), biostratigraphical data analysis (Wakefield

et al., 2001), etc. Fuzzy rules extraction for permeability prediction is described by Nikraves and Aminzadeh (2001a,b) and for facies classification by Finul et al. (2001a).

Among the advantages of the use of fuzzy rules for porosity classification is that they permit to complement for the future analysis the rule set constructed in this paper with the lithological descriptions of core samples usually available as qualitative or linguistic statements. The attempts to use these descriptions of core samples for porosity classification have been reported by Gedeon et al. (2001) where expert systems, supervised clustering (using k-prototype algorithm) and ANN were used in order to describe 4 classes of porosity (very poor, poor, fair and good). For all these three methods, the precision of classification (% of correct classifications) varies from 35% to 90.9% with the overall values from 60% to 68%. So, it looks viable to conduct the experiments with a mixed fuzzy rule base while the core information is available.

The developed approach does not require any previous assumptions to construct a model from the measured data set. Obviously, as a pattern recognition technique, it needs a good data set, valid and representative of different features present in the reservoir under study.

The constructed porosity classification models can be applied to make predictions for the wells from the oil reservoirs pertaining to micro-fractured carbonate formations. This approach can be extended to the rock type characterization in the absence of adequate geological information. As shown by Pawar et al. (2001) and Molina and Belanche (2005), permeability as another important reservoir property is strongly influenced by porosity and even in the absence of facies data, porosity data can be used to construct permeability distributions. In Finul et al. (2001b), it is shown that fuzzy models like Takegi–Sugeno–Kang fuzzy rules (Jang et al., 1997) can be used to predict porosity and permeability with a high degree of accuracy. Even with one input variable (porosity obtained from core density porosity was used to predict permeability) good predictions were obtained. These properties can be successfully used for the prediction of the productive volumes surrounding different wells and thus about reservoir storage capacity and fluid flow capability.

Another important application of the developed methods is the reduction of costs via the selection of the rational combination of the set of geophysical methods designed to solve the problems of the production process under particular geological and technical conditions.

Acknowledgments

The authors would like to thank Mexican Petroleum Institute (IMP) for support of this work within the project D.00006 (Distributed Intelligent Computing) and within the Naturally Fractured Reservoirs (YNF) Research Program.

References

- Aminzadeh, F., Wilkinson, D., 2004. Soft computing for qualitative and quantitative seismic object and reservoir property prediction: Part 2. Fuzzy logic applications. *First Break* 22, 69–78.
- Aminzadeh, F., Barhen, J., Glover, C.W., Toomanian, N.B., 2000. Reservoir parameter estimation using a hybrid neural network. *Computers and Geosciences* 26, 869–875.
- Anderberg, M., 1973. *Cluster Analysis for Applications*. Academic press, New York.
- Archie, G.E., 1952. Classification of carbonate reservoir rocks and petrophysical considerations. *AAPG Bulletin* 36 (2), 278–298.
- Bassiouni, Z., 1994. Theory, measurement, and interpretation of well logs, vol. 4. SPE Textbook Series, Richardson, TX, 372 pp.
- Batyrshin, I., Kaynak, O., Rudas, I., 2002. Fuzzy modeling based on generalized conjunction operations. *IEEE Transactions on Fuzzy Systems* 10 (5), 678–683.
- Brie, A., Johnson, D.L., Nurmi, R.D., 1985. Effect of spherical pores on sonic and resistivity measurements”, paper W presented at the SPWLA 26th Annual Logging Symposium.
- Doveton, J.H., 2000. Geologic log analysis using computer methods. *AAPG Computer Applications in Geology*, vol. 2. The American Association of Petroleum Geologist, Tulsa, OK.
- Duda, R., Hart, P., Stork, D., 2000. *Pattern Classification*. Wiley.
- Dutta, D.J., Madhavan, S., Sundaram, K.M., 1991. Characterization of dual porosity system in carbonates using sonic and resistivity measurements, paper C presented at the SPWLA 40th Annual Logging Symposium.
- Finul, J.J., Guo, Y.K., Jing, D., 2001a. Fuzzy partitioning systems for electrofacies classification: a case study from the Maracaibo basin. *Journal of Petroleum Geology* 24 (4), 441–458.
- Finul, J.J., Guo, Y.K., Jing, D., 2001b. A rule based fuzzy model for the prediction of petrophysical rock parameters. *Journal of Petroleum Science Engineering* 29, 97–113.
- Gedeon, T.D., Tamhane, D., Lin, T., Wong, P., 2001. Use of linguistic petrographical descriptions to characterise core porosity: contrasting approaches. *Journal of Petroleum Science Engineering* 31, 193–199.
- Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Ontario.
- Jang, J.-S.R., Sun, C.-T., Mizutani, E., 1997. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Prentice Hall.
- Huenges, E., Bucker, C., Lippmann, E., Lauterjung, J., Kern, H., 1997. Seismic velocity, density, thermal conductivity and heat production of cores from the KTB pilot hole. *Geophysical Research Letters* 24 (3), 345–348.
- Kazatchenko, E., Markov, M., Mousatov, A., 2003. Determination of Primary and Secondary Porosity in Carbonate Formations Using Acoustic Data”, paper SPE 84209 presented at the SPE Annual Technical Conference and Exhibition.
- Kosko, B., 1997. *Fuzzy Engineering*. Prentice-Hall, Upper Saddle River, NJ.
- Klir, G.J., Yuan, B., 1995. *Fuzzy Sets and Fuzzy Logic Theory and Applications*. Prentice Hall.
- Lucia, F.J., 1999. *Carbonate Reservoir Characterization*. Springer-Verlag, Berlin. 226 pp.
- MATLAB v. 5.2., 2000. The MathWorks.
- Molina, L.C., Belanche, L., 2005. Representing a relation between porosity and permeability based on inductive rules. *Journal of Petroleum Science and Engineering* 47, 23–34.
- Nikraves, M., Aminzadeh, F., 2001a. Mining and fusion of petroleum data with fuzzy logic and neural network agents. *Journal of Petroleum Science and Engineering* 29, 221–238.
- Nikraves, M., Aminzadeh, F., 2001b. Past, present and future intelligent reservoir characterization trends. *Journal of Petroleum Science Engineering* 31, 67–79.

- Nikravesh, M., Aminzadeh, F., Zadeh, L.A., 2003. *Soft Computing and Intelligent Data Analysis*. Elsevier.
- Pawar, R.J., Edwards, E.B., Whitney, E.M., 2001. Geostatistical characterization of the Carpinteria Field, California. *Journal of Petroleum Science Engineering* 31, 175–192.
- Pezeshk, S., Camp, C.V., Karprapu, S., 1996. Geophysical log interpretation using neural network ASCE. *Journal of Computing in Civil Engineering* 10 (2), 136–142.
- Tamhane, D., Wong, P.M., Aminzadeh, F., 2002. Integrating linguistic descriptions and digital signals in petroleum reservoirs. *International Journal of Fuzzy Systems* 4 (1), 586–591.
- Verga, F., Viberti, D., 2002. Uncertainty evaluation in well logging: analytical or numerical approach. SPWLA 43rd Annual Logging Symposium, Oiso, Japan.
- Wakefield, M.I., Cook, R.J., Jackson, H., Thompson, P., 2001. Interpreting biostratigraphical data using fuzzy logic: the identification of regional mudstones within the Fleming field, UK North Sea. *Journal of Petroleum Geology* 24 (4), 417–440.
- Wong, P.M., Nikravesh, M., 2001. Field applications of intelligent computing techniques. *Journal of Petroleum Geology* 24 (4), 381–387.
- Wong, P.M., Gedeon, T.D., Taggart, I.J., 1997. Fuzzy ARTMAP: a new tool for lithofacies recognition. *AI Applications* 10 (3), 29–39.
- Wong, P.M., Aminzadeh, F., Nikravesh, M., 2002. *Soft Computing for Reservoir Characterisation and Modeling*, *Studies in Fuzziness and Soft Computing*, ed. Physica-Verlag, Springer-Verlag.