SID

ابزارهای
پژوهش

سرویس ترجمه
تخصصی

کارگاه‌های
آموزشی

بلاگ
مرکز اطلاعات علمی

سامانه ویراستاری
STES

فیلم‌های
آموزشی

سامانه ویراستاری (ویرایش متون فارسی، انگلیسی، عربی)

کارگاه‌ها و فیلم‌های آموزشی مرکز اطلاعات علمی

روش تحقیق کمی

روش تحقیق کمی

آموزش مهارت‌های کاربردی در تدوین و چاپ مقالات ISI

آموزش مهارت‌های کاربردی
در تدوین و چاپ مقالات ISI

آموزش نرم افزار Word برای پژوهشگران

آموزش نرم‌افزار Word
برای پژوهشگران

Assessment of Clustering Methods for Predicting Permeability in a Heterogeneous Carbonate Reservoir

F. Khoshbakht* and M. Mohammadnia

Department of Petrophysics, Research Institute of Petroleum Industry (IRPI)
khoshbakhtf@ripi.ir

Abstract

Permeability, the ability of rocks to flow hydrocarbons, is directly determined from core. Due to high cost associated with coring, many techniques have been suggested to predict permeability from the easy-to-obtain and frequent properties of reservoirs such as log derived porosity. This study was carried out to put clustering methods (dynamic clustering (DC), ascending hierarchical clustering (AHC) self organizing map (SOM) and multi-resolution graph-based clustering (MRGC)) into practice in order to predict the permeability of a heterogeneous carbonate reservoir in southwest of Iran. In addition, the results are compared with three conventional approaches, empirical models, regression analysis, and ANN. The performance of all the examined methods was compared in order to choose the best approach for predicting permeability in un-cored wells of the studied field. For all clustering methods, selecting the optimal number of clusters is the most important task. The optimal values for the number of clusters are selected by iteration. The optimal number of clusters for MRGC, SOM, DC, and AHC are 7, 9, 9, and 8 respectively. Empirical equations and regression analysis weakly predict permeability and the value of R^2 parameters of both approaches are around 0.6. Generally the performance of clustering techniques is acceptable in Fahliyan formation. These techniques predict permeability between 1 and 1000 mD very well and just overestimate permeability values less than 1 mD. SOM performed the best among all examined techniques ($R^2=0.7911$). The constructed and validated SOM model with 9 clusters was selected to predict permeability in one of un-cored wells of the studied field. In this well, the predicted permeability was in good agreement with MDT derived permeability.

Keywords: Permeability, Prediction, Clustering Methods, Carbonate Reservoir

Introduction

In addition to being porous, a reservoir rock must have the ability to allow fluids to flow through it. The ability of rock to conduct fluids is defined as permeability. One of the most reliable sources for determining the matrix permeability of a reservoir is core. Usually core is available in a limited number of wells in a field because core is very expensive. Thus, it is essential to predict permeability from easy-to-obtain and frequent data such as logs. The prediction of permeability is not straightforward. The difficulty of predicting permeability is due to the fact that carbonates are very heterogeneous and petrophysical variations rooted in diagenesis, texture, grain size variation, cementation, clay content and etc. [1]. Unlike core, wireline logs are available in almost all wells in a field. By establishing a robust model which relates permeability to porosity, permeability in un-cored wells can be predicted. Log derived porosity shows good correlation with core porosity as reference porosity (Figure 1) and in un-cored wells, log derived porosity can be

introduced to models to predict permeability where no core is available.

Figure 1: Core porosity in respect to log porosity

Since the production of oil from the first oil well, many researchers have tried to estimate the permeability of reservoirs by means of different methods. Core analysis data and wireline logs are two main sources of data for predicting permeability. Archie introduced relationships to estimate permeability from core analysis data such as porosity and formation resistivity factor [2]. Because of the availability and frequency of wireline logs, logs are becoming more popular source for estimating permeability. Porosity, water saturation, capillary pressure, formation resistivity factor, and NMR's T1 and T2, derived from wireline log, can be used to estimate permeability. Leverett, Tixier, Timur, and Coates and Dumanoir developed correlations based on well log measurements to determine permeability [3-6].

In recent years, artificial intelligence techniques have been widely used to estimate permeability from wireline logs. Balan et al., Lin et al., and Zhang et al. predicted permeability by means of neural network [7-9]. Cuddy used fuzzy approach to estimate permeability [10].

The studied formation (Fahliyan) is one of the carbonate reservoirs of Iranian oil fields in the southwest of Iran. Fahliyan comprised of laminar to massive oolitic limestone with shale interbeds. This formation is a heterogeneous carbonate with complicated por-perm relation (Figure 2-a). In order to establish robust relationship between porosity and permeability to predict permeability, an efficient clustering technique is needed to classify heterogeneous por-perm data into homogeneous clusters. Although litho-types are geological groups with similar sedimentological properties, their petrophysical properties are different. In por-perm cross plot, most litho-facies overlap each other (Figure 2-b). Hence litho-facies is not a suitable criterion to classify por-perm data into homogeneous clusters and presumably predicting

permeability based on litho-facies fails to obtain satisfactory results.

In this study, we examined different techniques of data clustering in order to improve the quality of permeability prediction. MRGC, SOM, DC, and AHC are clustering methods which are utilized to predict permeability. The results are compared with other approaches of permeability prediction, empirical equations, regression analysis, and ANN. The method with best performance will be selected for predicting permeability in un-cored wells. The predicted permeability in un-cored wells was compared with permeability obtained by MDT (modular formation dynamics tester) to insure that the models worked properly.

Materials and Methods

Complexity of carbonate makes it difficult to treat all the data obtained from reservoirs as a unit. Usually partitioning the data into homogeneous groups leads to better results with less error. Clustering methods are widely used for the electro-facies analysis and petrophysical property prediction. In this study, we examined MRGC, SOM, DC, and AHC as the main clustering techniques to predict permeability.

In order to predict permeability in un-cored wells of the studied field, first, datasets are divided into two groups; the core data of three wells are selected for generating models (training) and the core data of the forth well was left out for the blind test. Then, based on the abovementioned clustering techniques, optimum models were constructed and por-perm data of core were classified into homogeneous clusters. After that, for validating the constructed models, these models were blind-tested in the forth wells to investigate how much the predicted permeability fitted with measured permeability.

Figure 2: The por-perm data distribution for the different litho-types, a) the measured data and b) approximate location of the por-perm data in each litho-type

The input of these models is core porosity and the output is permeability. The predicted permeability in the blind test step was compared with other common methods of permeability prediction (empirical models, regression analysis, and ANN). Finally, in un-cored wells, by introducing log porosity as input to models, permeability was predicted. The predicted permeability in un-cored wells was compared with the permeability obtained by MDT to insure that the models worked properly.

The models were established and validated using porosity and permeability of core. The model with best performance in predicting permeability was selected to predict permeability in un-cored well using log derived porosity. The used clustering techniques classify porosity and permeability of core in two-dimensional space into stable clusters. MRGC, SOM, and DC are iterative optimization methods. These algorithms first randomly initialize the kernel (center of each cluster) and then calculate the distance between each data point and the kernels. Each data point is associated to its nearest kernel. After processing all data points, each kernel is thus surrounded by a cloud of points. Each kernel is then moved to the center of its cloud. This process is iterative until the kernels are stabilized, i.e. the coordinates of kernel change less than a prescribed amount on the next iteration [11]. These methods are different in optimization algorithm; for example, SOM uses neural network to optimize, but DC uses K-nearest neighbor algorithm. The characteristics of each method are shortly mentioned in the next section. The details of these algorithms can be found in the references.

In the generated models, the permeability of each kernel is calculated by distance-weighted averaging of the permeability values in the cluster. The permeability of un-cored but logged zones is assigned by distance-weighted averaging of the average permeability of "K" nearby kernels.

The next step is selecting the optimal number of clusters. Variation on the number of clusters strongly affects the result. There are no universal criteria for choosing the optimal number of clusters in examined methods; therefore, we tried different values for the number of clusters and chose the optimal values which produced the best results.

Multi-resolution Graph-based Clustering (MRGC)

Ye and Rabiller proposed multi-dimensional dot-pattern recognition (MRGC) as a new clustering method for electro-facies analyses [12]. MRGC is based on non-parametric K-nearest neighbor and graph data representation. In this novel method, the disadvantages of other clustering methods such as prior-knowledge about the number of clusters, initial parameters, and the reliability of results were eliminated. MRGC is a tool which analyzes the structure of the complex data and partition natural data groups into different shapes, sizes, and densities [12].

MRGC automatically determines the optimal number of clusters.

Clusters produced by MRGC are organized in a hierarchical manner so that the high order clusters are the sub-cluster of low order clusters. Consequently, depending on the needed resolution, we can select an appropriate number of clusters.

We assigned 4 and 35 as the minimum and maximum of the desired clusters for MRGC. After executing MRGC, six models with different numbers of clusters (7, 10, 13, 16, 19, and 26 clusters) were generated. In the lowest order, 7 clusters and in the highest order, 26 clusters were partitioned. Then, each model was used to predict permeability.

The predicted permeability on the basis of each model was compared with core permeability. The model with the maximum value of R^2 was selected as the optimal number of clusters for MRGC model. Figure (3-a) presents the cross-plot of the value of R^2 of the predictions versus the number of clusters used in MRGC model. It is clear that by increasing the number of clusters, the error of the prediction increases, and the largest values of R^2 were obtained when MRGC model with 7 clusters was used for the permeability prediction.

Distance between the natural clusters of por-perm data is very short and most clusters overlap. By increasing the number of clusters, illusive clusters are produced and the prediction takes place in an unreal condition. So, the minimum number of clusters (7 clusters) is optimal for using in the prediction (Figure 4-a). MRGC model with 7 clusters was used to predict permeability in blind test well. The permeability predicted by MRGC was compared with core permeability in blind test well (Figure 5-e).

Self-Organizing Map (SOM)

The self-organizing map (SOM), commonly also known as Kohonen network, is a computational method for the visualization and analysis of high-dimensional data, particularly experimentally acquired information [13]. In this method, the number of clusters depends on the dimension of initial network; for instance, a 3×3 initial network of SOM produces 9 clusters. We examined 45 networks with different dimensions which made 4 to 100 clusters.

All the generated SOM models were used to predict permeability. The values of R^2 of prediction are plotted versus the different dimensions of SOM models in Figure (3-b). As expressed in the Figure (3-b), by increasing the values of SOM network dimensions from 2×2 to 3×3 , the value of R^2 of the prediction increases and the maximum value of R^2 is obtained when SOM model with 3×3 neurons is used. The values of R^2 decrease as the dimensions of SOM networks increase further, even though the R^2 curve shows local increments. The optimal dimension of SOM model (3×3) makes 9 clusters (Figure 4-b).

Figure 3: Relationship between the number of clusters and error in predicting permeability by a) MRGC, b) SOM, c) DC, and d) AHC

Figure 4: Clusters of train data, partitioned by different methods; grey dots are centers of clusters (kernel).

Dynamic Clustering (DC)

Dynamic clustering is a non-hierarchical method to partition data based on a pre-defined number of clusters. Algorithm places initial centers of clusters randomly and computes the distance between samples and centers; then, the samples are assigned to the nearest center. By iteration, the location of centers changes until all the points are assigned to the nearest center and all the centers are stabilized. We examined different numbers of clusters and generated models with different clusters. Permeability prediction on the basis of 9 clusters (Figure 4) has a higher value of R^2 ($R^2=0.7572$) (Figure 3-c). The value of R^2 increases by increasing the number of clusters until reaching up to 9 clusters, and then it decreases by further increasing the number of clusters.

Ascendant Hierarchical (AHC)

Hierarchical cluster analysis is a statistical method for finding relatively homogeneous clusters of cases based on measured characteristics [14]. It starts with each case in a separate cluster and then combines the clusters sequentially, reducing the number of clusters at each step until only one cluster is left. For choosing the optimal number of clusters, different values for the numbers of clusters were tested. After partitioning data into different orders of clusters, on the basis of each order, permeability was predicted. AHC model with 8 clusters produced the highest value of R^2 (Figure 3-d and Figure 4-d). The value of R^2 increases by increasing the number of clusters, goes through a maximum, and then decreases. The general trend of R^2 curve of AHC is similar to DC.

Regression Analysis

Linear regression estimates the coefficients of the linear equation, involving one independent variable that best predict the value of the dependent variable [15]. In this study, the dependent variable is the logarithm of permeability and the core porosity is an independent variable. The obtained equation for training data is as follows:

$$\text{Perm}=10^{(-0.096189+5.6647 \times \text{Por})}$$

The obtained equation is used to estimate the permeability of blind test well. Correlation coefficient between the measured and predicted permeability via regression analysis is 0.7443 (Figure 5-a). Although, this method is simple and easy to use, its result is comparable with clustering methods.

ANN

Neural network is a modeling technique, which models systems in a brain-like way [16]. The main feature of neural network is that it can learn the internal characteristics of a system by analyzing datasets.

In order to predict permeability, a back propagation neural network with two hidden layer designed. Porosity of three train wells was taken as input and permeability as output for training designed network. Trained network was used to predict permeability of test well on the basis of porosity. Correlation coefficient of the predicted permeability by means of neural network versus core permeability is 0.7581 (Figure 5-d). The predicted permeability by means of ANN is well matched for permeability values higher than 1 mD, but for cases with values less than 1 mD, ANN overestimates permeability. Possibly samples with very low values of permeability are erroneous and their permeability is not determined correctly because of the resolution limitation of laboratory instruments.

Results and Discussion

The prediction of permeability in un-cored wells is one of the difficult tasks of reservoir characterization, especially in heterogeneous and complex carbonates. We performed this study to examine the ability of different clustering methods in predicting permeability. We compared the results of these methods with other approaches of permeability prediction such as empirical equations, regression analysis, and ANN.

Partitioning por-perm data into stable and homogeneous groups was our first milestone. The optimal value for the number of clusters was selected by iteration. For each clustering technique, models with different numbers of clusters were generated. There were no general criteria for selecting the optimal number of clusters. The optimal number of clusters for MRGC, SOM, DC, and AHC was obtained to be 7, 9, 9, and 8 respectively (Figure 2). The trend of R^2 curve versus the number of clusters for MRGC and SOM is similar and the shape of R^2 curve of AHC is similar to DC. The size and shape of clusters which AHC and DC determined were comparable and it seemed that the performance of these two methods were approximately similar regarding to our data. The size of the clusters of AHC and DC were uniform but SOM and MRGC produced clusters with different sizes. Clusters of SOM were completely different from the other methods. Because of the nature of this method, increasing the number of clusters did not cause sharp changes in the results. AHC combined relatively similar high order clusters and made low order clusters that showed approximately same behavior. Generally clusters were very close together and they were difficult to be determined visually (Figure 4).

Table 1: Comparison of the results of different performed methods; Number of clusters and R^2 value of the prediction

Type	Empirical		Regression	ANN	Clustering			
Model	Timur	Dual Water	Simple	-	MRGC 7	SOM 3×3	DC 9	AHC 8
R^2	0.6821	0.7189	0.7443	0.7581	0.7580	0.7911	0.7572	0.7515

Figure 5: Core permeability versus the predicted permeability by means of: a) Timur correlation, b) Dual water correlation, c) regression, d) ANN, e) MRGC, f) SOM, g) DC, and h) AHC; the inset at the top left is the correlation coefficient of the prediction

Empirical equations and regression analysis could not predict permeability well. In spite of our effort to modify the parameters of empirical equations, these methods overestimate the permeability values. It seems they are not suitable for using in heterogeneous carbonates because they can not consider the complexity of this formation. Generally, the performance of clustering techniques was acceptable regarding to Fahliyan formation. These techniques predict permeability values between 1 and 1000 mD very well and just overestimate permeability values less than 1 mD. SOM is the best among the examined techniques ($R^2=0.7911$) (Table 1). Mismatch between the core and predicted permeability in low permeability intervals might occur because of the initial errors associated with measuring permeability in lab, because the measured permeability for low permeable cores are not reliable. Another possible explanation for this phenomenon might be the weakness of the used

clustering methods in extrapolating the extremes. On the whole, clustering techniques did not perform perfectly but in comparison with other used method they worked better. Clustering methods especially SOM and MRGC are multi-dimensional techniques and their performance boosts when they are used for partitioning multi-dimensional data. In this study, the input data was 2 dimensional, and hence SOM and MRGC could not show their capability compared with the other simple methods like experimental correlations or single regression.

The constructed and validated SOM with 9 clusters was selected to predict permeability in one of un-cored wells of the studied field. In this well, log and MDT are available. MDT is a kind of formation tester which measures mobility (k/μ) of formation in-situ. The results of predicting permeability in this well and its comparison with the permeability obtained from MDT is presented in the Figure 6.

Figure 6: Comparison of the predicted permeability versus MDT permeability in track 2 from right

Conclusion

For all clustering methods, selecting the optimal number of clusters is the most important task. In this study, the optimal values for the number of clusters were selected by iteration. The optimal number of clusters for MRGC, SOM, DC, and AHC is 7, 9, 9, and 8 respectively.

The empirical equations and regression analysis weakly predict permeability, and R^2 values of both methods are around 0.6. Generally, the performance of clustering techniques is acceptable in Fahliyan formation. These techniques predict permeability values between 1 and 1000 mD very well and just overestimate permeability values less than 1 mD. SOM performs best among all the examined techniques ($R^2=0.7911$). It is also worth mentioning that the quality of predicting permeability by means of ANN is higher than empirical equations and regression, but it is lower than the examined clustering methods. The successful prediction of permeability in an un-cored well of the studied field, confirmed by MDT, confirms the capability of SOM as a clustering technique. We recommend that these clustering methods should be utilized in the fields with more core data because the low number of core samples is one of the reasons for the moderate performance of clustering methods. In addition, the prediction of permeability using multidimensional data likes wireline logs can boost the performance of clustering methods. Clustering methods can be used for classifying complicated features of reservoirs such as fractures and reservoir rock types.

References

- [1]. Jennings J. W., and Lucia F. J., "Predicting Permeability From Well Logs in Carbonates With a Link to Geology for Interwell Permeability Mapping", SPE reservoir Evaluation and Engineering, 2003.
- [2]. Archie G. E., "The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics", Petroleum Transactions, Journal of Petroleum Technology, 1942.
- [3]. Leverett M. C., *Capillary Behavior in Porous Solids*, Trans, AIME, 1941.
- [4]. Tixier M. P., *Evaluation of Permeability From Electric-Log Resistivity gradients*, Oil and Gas J., 1949.
- [5]. Timur A., "An Investigation of Permeability", Porosity, and Residual Water Saturation Relationships for Sandstone Reservoirs, The Log Analysts, vol. 9, no. 4, 1968.
- [6]. Coates G. R., and Dumanoir J.L. *A New Approach to Improved Log Derived Permeability*, The Log Analyst, 1974.
- [7]. Balan B., Mohagheg, S., and Ameri, S. "State of the Art in Permeability Determination from Well Log Data: A Comparative Study, Model Development", paper SPE 30978 presented at SPE Eastern Regional Conference & Exhibition held in Morgantown, West Virginia, U.S.A., 1995.
- [8]. Lin J. L., and Salisch H. A., "Determination from Well Logs of Porosity and Permeability in a Heterogeneous Formation", SPE 28792, presented at the SPE Asia Pacific Oil & Gas conference, Melbourne, Australia, 1994.
- [9]. Zhang Y. J., and Lollback P. A., "A Methodology for Estimating Permeability From Well Logs in a Formation of Complex Lithology", SPE 37025, presented at the SPE Asia Pacific Oil & Gas Conference, Adelaide, Australia, Oct. 28-31, 1996.
- [10]. Cuddy S. J., *The Application of the Mathematics of Fuzzy Logic to the Geosciences*, SPE Paper 49470, 1998.
- [11]. Rabiller P., and Ye S. J., "The iterative use of clustering and modeling to improve permeability prediction", SPWLA 42th Annual Logging Symposium, 2001.
- [12]. Ye S. J., and Rabiller P., "A new tool for electro-facies analysis: multi-resolution graph-based clustering", SPWLA 41th Annual Logging Symposium, 2000.
- [13]. Kohonen T., *Self-organizing maps*, Third, extended edition, Springer, 501 p, 2001.
- [14]. Lukasova A., "Hierarchical agglomerative clustering procedure", Pattern Recognition, vol. 1, pp. 365-381, 1979.
- [15]. Lee S. H., and Datta-Gupta A., "Electro-facies Characterization and Permeability Predictions in Carbonate Reservoirs: Role of Multivariate Analysis and Nonparametric Regression", SPE No. 56658, 1997.
- [16]. Mohaghegh S., Balan B., and Ameri S., *Permeability Determination from Well Log Data*, SPE formation evaluation, 1997.

SID

ابزارهای
پژوهش

سرویس ترجمه
تخصصی

کارگاه‌های
آموزشی

بلاگ
مرکز اطلاعات علمی

سامانه ویراستاری
STES

فیلم‌های
آموزشی

سامانه ویراستاری (ویرایش متون فارسی، انگلیسی، عربی)

کارگاه‌ها و فیلم‌های آموزشی مرکز اطلاعات علمی

روش تحقیق کمی

آموزش مهارت‌های کاربردی
در تدوین و چاپ مقالات ISI

آموزش نرم افزار Word
برای پژوهشگران