Evaluating Different Approaches to Permeability Prediction in a Carbonate Reservoir

Farhad Khoshbakht

Department of Reservoir Study and Field Development, Research Institute of Petroleum Industry (RIPI), Tehran, Iran

ABSTRACT

Permeability can be directly measured using cores taken from the reservoir in the laboratory. Due to high cost associated with coring, cores are available in a limited number of wells in a field. Many empirical models, statistical methods, and intelligent techniques were suggested to predict permeability in un-cored wells from easy-to-obtain and frequent data such as wireline logs. The main objective of this study is to assess different approaches to the prediction of the estimation of permeability in a heterogeneous carbonate reservoir, i.e. Fahliyan formation in the southwest of Iran. The considered methods may be categorized in four groups, namely a) empirical models (Timur and Dual-Water), b) regression analysis (simple and multiple), c) clustering methods like MRGC (multi-resolution graph-based clustering), SOM (self organizing map), DC (dynamic clustering) and AHC (ascending hierarchical clustering), and d) artificial intelligence techniques such as ANN (artificial neural network), fuzzy logic, and neuro-fuzzy.

This study shows that clustering techniques predict permeability in a heterogeneous carbonate better than other examined approaches. Among four assessed clustering methods, SOM performed better and correctly predicted local variations. Artificial intelligence techniques are average in modeling permeability. However, empirical equations and regression methods are not capable of predicting permeability in the studied reservoir. The constructed and validated SOM model with 6x9 clusters was selected to predict permeability in the blind test well of the studied field. In this well, the predicted permeability was in good agreement with MDT and core derived permeability.

Keywords: Permeability, Carbonate Reservoir, Clustering, Intelligent, Experimental Correlation

INTRODUCTION

In addition to being porous, a reservoir rock must be able to allow fluids to flow. The ability of rock to conduct fluids is defined as permeability. Core is the most reliable source of matrix permeability. Usually in a field, core is available in a limited number of wells because of costs and technical limitations. Permeability as one of the most important petrophysical properties of a reservoir is the main input to static and dynamic models. In order to construct a realistic model of a reservoir, accurate permeability data with an appropriate distribution ought to be available. 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 carbonate rocks are very heterogeneous and factors such as diagenesis (e.g., cementation), texture, grain size variation and fabric, cementation, and clay content control permeability [10]. Unlike core, wireline logs are available in almost all wells in a field. By establishing a robust model which relates permeability to wireline log responses permeability in un-cored wells can be predicted.

More than 40 percent of the world oil reservoirs are placed in carbonate rocks. Because of the economic importance of carbonates, their accurate characterization is vital for exploration and production. The ultimate goal of reservoir characterization is to improve production efficiency and oil recovery by understanding and modeling the reservoir. To correctly model the flow behavior in a carbonate reservoir, it is essential to understand the permeability profile [20].

Carbonate rocks are unstable and diagenetic processes chemically and physically alter a somehow homogeneous formation into a completely heterogeneous one. The permeability of a carbonate reservoir often is severely affected by tectonic and diagenetic phenomena like fracturing, dissolution, and leaching, which makes the prediction of permeability very difficult.

In the last decade, many researchers have tried to estimate permeability from wireline logs. Because of availability and frequency of wireline logs, they have become a popular source for estimating permeability. Archie (1942) introduced relationships which estimated permeability via core analysis data such as porosity and formation resistivity factor [3]. Porosity, water saturation, capillary pressure, formation resistivity factor, and NMR T1 & T2 parameters, derived from wireline logs, are used to estimate permeability. Leverett (1941), Tixier (1949), Wyllie and Rose (1950), Timur (1968), and Coates and Dumanoir (1974) developed correlations based on well log measurements to determine permeability [5, 21-23].

In recent years, artificial intelligence techniques have widely been used to estimate permeability from wireline logs. Balan et al. (1995), Mohaghegh et al. (1997), Lin et al. (1994), Zhang et al. (1996), and Huang et al. (1996) predicted permeability by means of neural networks. Cuddy (1998) and Finol et al. (2001) used the fuzzy approach to estimate permeability [4, 7-9, 13, 16, 25].

In this study, different approaches to permeability prediction are examined in order to establish a robust model to predict permeability in un-cored wells. The method with the best performance, i.e. SOM, was selected to predict permeability in un-cored wells. The predicted permeability in the blind test wells was compared with MDT-(Modular Formation Dynamics Tester) derived permeability and core permeability to ensure that the model worked properly.

MATERIALS AND METHODS

The studied field is located in Zagros fold-thrust belt, the southwest of Iran (Figure 1). Zagros is a part of Tethys Ocean and is one of the most important petroleum basins in the world [1]. This basin is located in the south west of Iran and north of Iraq. The geological history of this basin includes long time subsidence and deposition interrupted by short time uplift. Folding process of this basin occurred in Miocene and Pliocene and continued until now, which has formed long anticlines [17]. These anticlines constitute most of oil traps in this basin.

Fahliyan (the uppermost Jurassic-Cretaceous) is one of the major reservoir formations of Iranian oil fields. Fahliyan is a clean limestone comprised of massive oolitic or pelletal limestone [2]. In the studied field, Fahliyan formation is divided to eight reservoir zones.
In this study, the wireline and core data of four wells obtained from Fahliyan formation were used. The datasets of three wells (A, B, and C) were used for building models and the forth well (D) left out for the test. A suit of logs including neutron porosity (NPHI), sonic log (DT), resistivity log (LLD), photoelectric (PEF), and natural gamma ray (GR) was selected as the input to all the models (Figures 2 and 3). These logs record the properties of the reservoir, which control permeability such as shale content, effective porosity, water saturation, and lithology.

**Figure 2**: Cross plot of core permeability versus selected wireline logs (From left to right: NHPI, DT, LLD, PEF, and GR)

### Permeability Modeling

In order to predict permeability in un-cored wells, models should be constructed. The inputs to these models are well logs and the output is permeability. In wells with core data (A, B, and C), the models were constructed and tuned, and then they were used to predict permeability in un-cored wells. In the following, all the models are briefly described and the predicted permeability by means of each one is presented.

<table>
<thead>
<tr>
<th><strong>Figure 3</strong>: Statistics and histogram of permeability as the output of the models and selected wireline logs (NHPI, DT, LLD, PEF, and GR) as the inputs to the prediction models</th>
</tr>
</thead>
</table>

---

*Journal of Petroleum Science and Technology 2015, 5(1), 79-90*  
http://jpst.ripi.ir

© 2014 Research Institute of Petroleum Industry (RIPI)
Empirical Models

Empirical models are based on experiential observations and are introduced for particular reservoirs. For applying them to other reservoirs, with different properties, it is necessary to calibrate their parameters. Generally empirical equations relate permeability to porosity and water saturation. First, we calibrated the models by using the core data of train wells. In train wells, the inputs (porosity and irreducible water saturation) and the output of the model (permeability) were available and the only unknown of the equations (\(A\) and \(B\)) can be calculated. Then, in blind-test well the permeability was predicted using the calibrated models. In this study, we used two common empirical equations for predicting permeability: Timur and Dual-Water.

Timur (1968) found a relationship for estimating the permeability of sandstones from the in-situ measurements of porosity and residual saturation [21]. His model is applicable where the condition of residual saturation exists. The main source of uncertainty in Timur model is the high amount of error in determining the residual saturation. In our case, the residual water saturation of the studied reservoir was obtained from capillary pressure tests. The average value of residual saturation was 0.18. The Dual water model was developed by Coates and Dumanoir in 1974 [5]. They suggest this model for shaly formations. Total porosity, effective porosity, and irreducible water saturation are the inputs to this model. The equations and calibrated parameters are presented in Table 1.

Regression Analysis

Regression analysis is a statistical technique that identifies the relationship between two or more quantitative variables: a dependent variable whose value is to be predicted, and independent variable(s). Regression analysis is used to understand the statistical dependence of one variable on other variables.

We utilized simple and multiple regressions in order to estimate permeability from wireline logs. In the simple regression, the effective porosity (PHIE) was the input (independent variable). In the multiple regression, DT, LLD, NPHI, PEF, and GR were selected as the inputs. In the multiple regression, the data is constituted using the below equation:

\[
\log(perm) = a_0 + a_1\times(DT) + a_2\times(LLD) + a_3\times(NPHI) + a_4\times(PEF) + a_5\times(GR)
\]

The values between brackets are known and the coefficients \(a_0\), \(a_1\), \(a_2\), \(a_3\), \(a_4\), and \(a_5\) are unknown. A system of matrix \((Y=aX)\) representing the above linear equations is constructed and solved for unknown coefficients by minimizing the sum of the squares of the deviations of the data from the model (least-squares fit). The equation below represents the relationship between permeability and PHIE, which is obtained from the simple regression.

\[
Perm = 10^{(-1.748+0.00884115\times(DT)-0.132608\times\log(LLD)+1.64347\times(NPHI)+385849\times(PEF)+0.00575193\times(GR))}
\]

The relationship between permeability as the dependent variable and well logs as the independent variables is given by the below equation:

\[
Perm = 10^{(-1.748+0.00884115\times(DT)-0.132608\times\log(LLD)+1.64347\times(NPHI)+385849\times(PEF)+0.00575193\times(GR))}
\]

<table>
<thead>
<tr>
<th>Name</th>
<th>Equation</th>
<th>Constant ((A))</th>
<th>Exponent ((B))</th>
<th>(R^2) of prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timur</td>
<td>[ K = A \frac{\phi_e^{\frac{1}{B}}}{S_{wir}} ]</td>
<td>6.758595</td>
<td>1.230943</td>
<td>0.6028</td>
</tr>
<tr>
<td>Dual-Water</td>
<td>[ K = A \frac{\phi_e^{\frac{2}{B}} - \frac{V_{wir}}{V_{wir}}}{S_{wir}} ]</td>
<td>8.001080</td>
<td>1.176190</td>
<td>0.5787</td>
</tr>
</tbody>
</table>

Table 1: Used empirical equations and their parameters calibrated for the studied wells

© 2014 Research Institute of Petroleum Industry (RIPI)

http://jpst.ripi.ir
Clustering Methods

The aim of cluster analysis is to classify a dataset into groups that are internally homogeneous and externally isolated on the basis of a measure of similarity or dissimilarity among groups [12]. Clustering methods are widely used for electro-facies analysis and the prediction of petrophysical properties. By integrating clustering methods with intelligence techniques, some new methods such as SOM and MRGC were born [11,24]. In this study, we utilized two classic clustering methods, namely DC and AHC, and two new clustering methods, i.e. SOM and MRGC, to predict permeability from wireline logs.

Predicting permeability by means of clustering methods is carried out in three steps:
1) Partitioning log data of the train wells into somehow similar groups and determining the properties of each groups;
2) Identifying clusters of test data using cluster properties, determined in the previous step;
3) Predicting permeability in each cluster of the blind-test well by means of KNN method.

Each clustering method tries to partition data into stable and homogeneous groups on the basis of its particular algorithm. Selecting the optimal number of clusters is important, because variation in the number of clusters strongly affects the results. There are no universal criteria for choosing the optimal number of clusters in the examined methods. Therefore, we tried different values for the number of clusters and chose the optimal values via trial and error. NPHI, DT, LLD, PEF, and GR were the inputs to all the clustering methods.

MRGC is based on non-parametric K-nearest neighbor and graph data representation. MRGC automatically determines the optimal number of clusters. We assigned 4 and 35 respectively as the minimum and maximum of the desired clusters for MRGC. After executing MRGC, the log data of train wells were partitioned into six orders of clusters (5, 7, 10, 14, 16, and 18 clusters). The highest $R^2$ (0.5978) was obtained when MRGC model with 16 clusters was used for permeability prediction (Figure 4-a).

The SOM is a computational method for the visualization and analysis of high-dimensional data, especially experimentally acquired information [11]. We tested SOM networks with different dimensions to partition the datasets of train wells. Then, the permeability of the blind-test well was predicted using each model (Figure 4-b). As shown in Figure 4-b, the optimal dimension of SOM network is 6×9, which gives the highest $R^2$ (0.6936) (Table 2 and Figure 5-f).

Dynamic Clustering or $k$-means clustering is a non-hierarchical method to classify a dataset on the basis of a pre-defined number of clusters. The algorithm assigns each object into a cluster by iteration. It minimizes the sum of distance from each object to its cluster centroid, over all clusters. This algorithm moves objects between clusters until the sum cannot be decreased further [15]. This clustering method is suitable for datasets with a large amount of data. We examined different values for the number of clusters and predicted permeability in the blind-test well. Permeability prediction on the basis of 12 clusters (Figure 4-c) has the highest $R^2$ (0.6631) among DC models with different numbers of clusters (Figure 5-g).

Hierarchical clustering method partitions data over a variety of scales by creating a cluster tree or dendrogram [15]. Ascending hierarchical cluster (AHC) analysis is a statistical method for finding relatively homogeneous clusters of cases based on measured characteristics [14]. AHC produces different levels of clusters and user can choose
the most appropriate level for the dataset. Generally, clustering of a dataset can be executed by means of AHC in three steps: 1) measuring the distance between every pairs of objects in the dataset, 2) linking pairs of objects which are very close together, and 3) cutting the hierarchical tree into clusters [15]. Hierarchical trees were cut at different levels and the results showed that AHC model with 8 clusters has the highest $R^2$ (Figure 4-d) among levels of clusters. AHC models with less or more than 6 clusters decrease the $R^2$ value (Figure 5-h). Increasing the number of clusters does not make sharp changes in the results, because AHC combines relatively similar high order clusters and makes low order clusters.

Artificial Intelligence Techniques

ANN, fuzzy logic, and neuro-fuzzy are three artificial intelligence techniques which were used in this study to predict permeability from wireline logs.

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

Figure 4: Variation of $R^2$ by changing the number of clusters for a) MRGC, b) SOM, c) DC, and d) AHC
In order to predict permeability using ANN, a back propagation neural network with two hidden layers have been designed. A trained network was used to predict the permeability of the test well. The $R^2$ of the predicted permeability by means of neural network versus core permeability is 0.6279 (Table 2 and Figure 5-i).

In this paper, a fuzzy approach proposed by Cuddy (2000) was utilized for permeability prediction [7]. This approach works by assigning a probability to the quality of the prediction from each log, then combines the probabilities, and predicts the most likely permeability [7]. We examined different numbers of classes and predicted the permeability for each case. Fuzzy model with 10 classes predicted permeability with the highest $R^2$ (0.6275) among models with different numbers of classes (Figure 5-j).

We used a locally linear neuro-fuzzy (LLNF) model which utilized local linear model tree learning algorithm (LOLIMOT) to predict the permeability. The fundamental approach to a locally linear neuro-fuzzy model is dividing the input space into small linear subspaces with fuzzy validity functions. Any produced linear part with its validity function is described as a fuzzy neuron. Thus the total model is a neuro-fuzzy network with one hidden layer, and a linear neuron in the output layer, which simply calculates the weighted sum of the outputs of locally linear models [18]. By plotting the mean square error for the train and test data, we found that the optimal number of hidden layer neurons is 29, and constructed our neuro-fuzzy model with 29 neurons. The $R^2$ of permeability prediction by means of neuro-fuzzy model was 0.5900 (Figure 5-k).

RESULTS AND DISCUSSION

The prediction of permeability in un-cored wells is one of the difficult tasks in reservoir characterization, especially in heterogeneous and complex carbonates. Wireline logs are frequent and easy-to-obtain data, which are usually available in almost all the wells. We launched this study to examine four common approaches to predicting permeability, namely empirical methods, statistical analysis, clustering methods, and artificial intelligence techniques. The method with the best performance was selected to predict permeability in un-cored wells. The predicted permeability in one of the un-cored wells of the studied field was compared with MDT- (Modular Formation Dynamics Tester) derived permeability to ensure that the model worked properly.

The first examined approach was empirical equation: Timur and Dual-Water models. The $R^2$ of the measured permeability versus predicted permeability using Timur and Dual-Water models are 0.6028 and 0.5787 respectively (Table 2). Even though the predicted permeability by means of Timur model is closer to the measured permeability than the one measured by Dual-Water model, both models performed weakly (Figures 5-a and 5-b). In spite of our efforts to properly calibrate the parameters of these models, the permeability prediction was not satisfactory. Because of the heterogeneity of the reservoir, empirical equations failed to predict permeability properly. Usually the use of empirical equations is restricted to the region(s) where the models are developed for, and in different reservoirs, imperial models should be employed with caution.

Regression analysis is simpler and easier to practice than other methods. The $R^2$ values of permeability prediction by means of simple and multiple regressions are 0.4920 and 0.5279 respectively (Table 2). The distribution shape of the predicted data via simple and multiple regressions are similar and both overestimate low permeabilities and underestimate high permeabilities.
Table 2: The $R^2$ of the predicted versus measured permeability for all the used methods

<table>
<thead>
<tr>
<th>Type</th>
<th>Model</th>
<th>Empirical</th>
<th>Regression</th>
<th>Clustering</th>
<th>Artificial Intelligence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Timur</td>
<td>Dual-Water</td>
<td>Simple</td>
<td>MRGC 16</td>
<td>SOM 6×9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Multiple</td>
<td></td>
<td>DC 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AHC 6</td>
<td>ANN 0.6279</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fuzzy 10 0.6275</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Neuro-Fuzzy 0.5900</td>
</tr>
</tbody>
</table>

The cross plots of permeability versus logs (Figure 2) show that there is not a simple and linear relationship between them. Consequently, we cannot expect methods like empirical equations and regression analysis to explore complex relationships. Empirical methods simplify problems by reducing the number of variables, while different factors influence the permeability of reservoirs. In addition, regression methods tend to average data and ignore local variations, which are important in thin layered reservoirs. Figures (5-c) and (5-d) present the plots of the predicted permeability versus the measured permeability by regression analysis.

The aim of cluster analysis is to classify a dataset into groups that are internally homogeneous and externally isolated on the basis of a measure of similarity or dissimilarity between groups [12]. For proper permeability prediction in complex reservoirs, different criteria have been proposed for data classification such as reservoir layers, litho-facies, rock types, and hydraulic flow units. But in order to predict permeability from logs, methods should be employed which can handle the dimensionality and complexity of log data. Because of the variety of wireline logs and the heterogeneity of carbonates, it is usually difficult to easily recognize natural groups of data.

Generally, clustering methods predict permeability better than the other examined methods. Even though they are not perfect, their results are acceptable. Selecting the optimal number of clusters is one of the important tasks in using clustering techniques, because variation in the number of clusters strongly affects the result of permeability prediction. In addition, there are no general criteria for choosing the optimal number of clusters in the examined methods. We examined different values for the number of clusters and predicted permeability in each case. Each method has its own value for the optimal number of clusters, e.g. the optimal number of clusters for MRGC, SOM, DC, and AHC are 16, 54, 12, and 6 respectively. Predicting permeability on the basis of SOM has the highest $R^2$ among the examined clustering methods, even though SOM is not much more robust than the other clustering methods; but, the results of the prediction are slightly comparable. However, SOM predicts low and high permeabilities better than the other methods. The performance of ANN and fuzzy logic for predicting permeabilities between 1 to 10 mD is acceptable, but both methods fail to predict permeabilities higher than 10 mD correctly. Contrary to our expectation neuro-fuzzy performance was weaker than the performance of ANN and fuzzy logic. Neuro-fuzzy only predicts very high values (~100 mD) properly, but for the range 1 to 10 mD, at which most methods do well, neuro-fuzzy was disappointing.
Figure 5: Cross-plots of the measured permeability versus the predicted permeability by means of different methods, the methods and the $R^2$ of the prediction are presented in the up-left box of each cross plot.

The main preference of clustering methods is their ability to predict local variations. Most examined methods, especially regression and empirical methods, ignore local variations. These methods predict medium permeabilities well, but the prediction of extremes is associated with high errors. Statistically the number of samples in the extremes is low and most techniques cannot learn the pattern of these parts of data in training. Consequently, they tend to average the data and ignore local variations.

The best performed method (SOM with 6×9 clusters) was used to predict permeability in one of un-cored wells of the studied field. In this well, the permeability derived from MDT test was available. As it is clear in the right column of Figure 6, the predicted permeability matches well with MDT permeability. This interval comprises limestone with a limited amount of shale and its porosity is almost constant. Utilizing empirical methods which relate permeability to porosity leads to somehow constant predicted permeability, assuming that irreducible water saturation is also constant in this interval.
Figure 6: Wireline logs and predicted permeability in the blind test well; From left: Track 1) CGR and GR, Track 2) depth scale, Track 3) lithology, Track 4) resistivity logs, Track 5) NPHI, DT & RHOB, Track 6) predicted permeability by means of SOM 6x9 (solid line) and MDT permeability (dots), Track 7) predicted permeability by means of SOM 6x9 (solid line) and measured permeability (square).
CONCLUSIONS

Numerous methods have been proposed by researchers to predict permeability in un-cored wells. Common approaches to predicting permeability are empirical models, regression analysis, clustering methods, and artificial intelligence techniques. We examined these methods to predict permeability in a carbonate reservoir. The study showed that among the examined methods, regression analysis was the worst and clustering techniques were the best in predicting permeability. Accordingly, among the four assessed clustering methods, SOM performed better and could properly predict local variations. Since the clustering methods classify the datasets into homogeneous subclasses and then predict the permeability in each sub-class, it is expected that the clustering methods perform better compared with the other methods which treat the dataset as a whole; in fact, in each subclass the relationship between permeability and the log responses is simpler.

Artificial intelligence techniques were average in modeling permeability. Empirical equations and regression techniques were not capable of predicting permeability in the studied reservoir. The successful prediction of permeability in an un-cored well, confirmed by MDT, justified the capability of SOM as a clustering technique.

REFERENCES


