Evaluating Machine Learning Techniques for Carbonate Formation Permeability Prediction Using Well Log Data

Usama Alameedy¹*, Ahmed T. Almomen¹ and Najah Abd²
¹ Department of Petroleum Engineering, College of Engineering, University of Baghdad, Baghdad, Iraq
² Geology Department, College of Science, University of Baghdad, Baghdad, Iraq
* Correspondence: usama.sahib@coeng.uobaghdad.edu.iq

Abstract
Machine learning has a significant advantage for many difficulties in the oil and gas industry, especially when it comes to resolving complex challenges in reservoir characterization. Permeability is one of the most difficult petrophysical parameters to predict using conventional logging techniques. Clarifications of the work flow methodology are presented alongside comprehensive models in this study. The purpose of this study is to provide a more robust technique for predicting permeability; previous studies on the Bazirgan field have attempted to do so, but their estimates have been vague, and the methods they give are obsolete and do not make any concessions to the real or rigid in order to solve the permeability computation. To verify the reliability of training data for zone-by-zone modeling, we split the scenario into two scenarios and applied them to seven wells' worth of data. Moreover, all wellbore intervals were processed, for instance, all five units of Mishrif formation. According to the findings, the more information we have, the more accurate our forecasting model becomes. Multi-resolution graph-based clustering has demonstrated its forecasting stability in two instances by comparing it to the other five machine learning models.

Keywords: Machine learning; Well log; Permeability; Bazirgan; Carbonate formation

1. Introduction
Enhancing hydrocarbon production necessitates precise knowledge of reservoir features such as permeability. For its impact on flow unit identification, reservoir characterization, and perforation design, permeability is often regarded as the most important metric in reservoir management (Al-Yaseri et al., 2013). The competence to accurately anticipate a reservoir’s permeability values would greatly improve decision-making across a wide range of operational domains. This includes field and production planning, a reserve estimate, and problem solving during production (Talebkeikhah et al., 2021).

Borehole surveys never offer a comprehensive record of the permeability of the whole reservoir region. Permeability is governed by the interplay of geologic variables like deposition and diagenesis, which are not directly addressed by a single record (Rabiller et al., 2001). Impact of wireline-induced permeability Logs are often suggested over core for permeability profile evaluation because to their inexpensive cost and high success rate. Deterministic permeability estimates from logs and, more recently, descriptive or statistical techniques employing facies analysis are examples of the methodologies used (Alameedy, 2003). Since the facies profile is inferred from core data and a sedimentary model is presented before any facies isopach maps can be drawn, this technique places a
premium on a high degree of knowledge. Qualified permeability profiles could then be used to populate the reservoir models’ cells using geostatistical methods (Rabiller et al., 2001; Alameedy et al., 2022)

Sedimentary processes governed the original reservoir interior geometry, and these petrophysical qualities, such as capillary pressure, permeability, and porosity, may be described by the term "facies." The complex recognition of sedimentary facies is necessary for the establishment of a 3D geological model (Ye et al., 2000)

According to the description provided by Serra and Abbott (1982), an effective method of using a facies analysis is to establish a classification model that divides the log data into sets of log answers that describe a sediment and enable the sediment to be identified from others. Later on, in the survey process, the model is extended to include data from all of the wells in the same field or basin. Appropriate techniques must be employed to evaluate both the representativeness of the cores used to calibrate the model and the model's transferability to a different well (Thevoux-Chabuel et al., 1997).

The purpose of this scientific research is to make an accurate prediction of the permeability values of carbonate reservoirs in three units (Mb11, Mb12, and MB21) of the Mishrif formation in the Bazirgan oil field. Moreover, the goal is to demonstrate that the machine learning approach offers insights for issues related to the quality of the log data collection and the disparity in investigation scale between logs and core samples (plugs). Furthermore, different advanced statistical methods for upscaling high-resolution core measurements to extrapolate rock sample data to find extremely heterogeneous formations and accurately predict permeability in these zones, such as MRGC electrofacies clustering collected from multiple logs.

2. Geological, Geographical Descriptions, and Dataset of the Study Area

Near Iraq–Iran border, some 40 kilometers north-east of Amara, lies the Bazirgan oil field. Fig. 1 shows that the Bazirgan oilfield is consisting of an asymmetrical anticline fold that runs from NW to SE and has two domes, one to the north and one to the south, that are separated by a saddle (CNOOC, 2012).

Fig. 1. The Mishrif formation top structural contour map at the Bazirgan oil field (CNOOC, 2012)
The main oil reservoir in the Bazirgan oilfield is the Mishrif Formation, which is mostly consisted of limestone and dolomite with shale in between, especially at the top (Mohammed et al., 2022). The Mishrif Formation is divided into eight stratigraphic units (MA, MB11, MB12, MB21, MB22, MB23, MC1, MC2). These units are in turn separated by eight compacted zones. In contrast to the top part of the Mishrif Formation, the middle and lower regions, which correspond to the MB21 and MC2 units, have an excellent porosity range of 14% to 21%. The most recent studies confirm that these components are in fact, the Mishrif reservoir's primary oil-producing MB21, MC1, and MC2 units (Alhusseini and Hamd-Allah, 2023).

Specifically, for the five units, this research uses information acquired from seven oil wells, including conventional well logs and core measurements (porosity and permeability), in the Mishrif reservoir (MB11, MB12, MB21, MC, and MC1). Full-set log curves include gamma ray (GR), neutron porosity (NPHI), acoustic time (DT), and formation density (RHOB). After correcting for environmental borehole effects, the resistivity records (SN, ILD, MLL, MSFL, LLD, and LLS) are converted to true resistivity (Rt) and flushed resistivity (Rxo), which are then used as input parameters for the training data set with other logs. Table 1 provides a summary of these variables in the form of statistics. The permeability has been modeled using a large number of samples from a dataset of 1466 data points using five distinct computer models (intelligence approaches).

Table 1. Statistical analyses of the input and output variables

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable</th>
<th>Unit</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GR</td>
<td>API</td>
<td>1.14</td>
<td>66.39</td>
<td>19.23</td>
</tr>
<tr>
<td>2</td>
<td>RHOB</td>
<td>gm/cc</td>
<td>2.08</td>
<td>2.78</td>
<td>2.54</td>
</tr>
<tr>
<td>3</td>
<td>DT</td>
<td>US/F</td>
<td>40.01</td>
<td>88.58</td>
<td>60.8</td>
</tr>
<tr>
<td>4</td>
<td>NPHI</td>
<td>Fraction</td>
<td>0.0</td>
<td>0.37</td>
<td>0.097</td>
</tr>
<tr>
<td>5</td>
<td>Rt</td>
<td>OHMM</td>
<td>0.23</td>
<td>2212.5</td>
<td>56.79</td>
</tr>
<tr>
<td>6</td>
<td>Rxo</td>
<td>OHMM</td>
<td>0.22</td>
<td>630</td>
<td>22.37</td>
</tr>
<tr>
<td>7</td>
<td>Permeability (k)</td>
<td>Md</td>
<td>0.01</td>
<td>2017</td>
<td>35.33</td>
</tr>
</tbody>
</table>

3. Machine learning (ML)

A reservoir's formation, fluid content, and the best method for extracting that content are all aspects of the petroleum industry that may be better understood and explored with the use of geoscientific methods. Fluid storage capacity, petrophysical characteristics, and the reservoir's capability to transport these fluids to the wellhead are all factors that must be analyzed. From exploration and drilling through reservoir management and production to eventual field abandonment, reservoir characterization gives a comprehensive image of the formation necessary for efficient oil recovery planning. The oil and gas industry is increasingly relying on machine learning (ML) to forecast future production, formation assessment, drilling, stimulation, and reservoir characterization activities. Rocks' permeability, porosity, and other petrological parameters may be described using ML (Talebkeikhah et al., 2021; Tian et al., 2021; Al-Ateya et al., 2022; Alameedy et al., 2022).

Talebkeikhah et al. (2021) uses machine learning techniques such as artificial neural networks to estimate permeability values from a complete set of well logging data and core measurements by clustering the dataset using SVR, RBF, and MLP modules. In order to enhance permeability estimates, (Mohebian et al., 2021) developed a hybrid algorithm that combines MRGC with rock typing techniques. To figure out how permeable a rock is, the results from the three rock-typing methods are added together and then insert into Multi-Resolution Graph-based Clustering (MRGC).
3.1. The Function of Electrofacies and Clustering

Electrofacies are a group of log responses that characterize a sediment and enable for differentiation between different sediments and rock types (Davis, 2018). Facies, rock types, or clusters that are created or differentiated based on variations in log responses are collectively referred to as "Electrofacies."

Clustering is a method used in Machine Learning to organize large amounts of data into more manageable chunks. Using a clustering technique, we can divide a collection of data points into distinct categories. In principle, data points belonging to the same group should have comparable attributes and/or features, whereas data points belonging to other groups should display strikingly distinct characteristics. Unsupervised learning approaches, such as clustering, are widely utilized in the realm of statistical data analysis.

The criteria associated with each clustering method may or may not be suitable for the actual structure of the data. For many uses, the resultant divisions are meaningless unless the results have been verified.

The following is a list of the unsupervised clustering methods that were used in this study:
- Dynamic Clustering (DYN) (Mourot, 1993).
- Ascendant Hierarchical Clustering (AHC) (Guha et al., 1998)
- Self-Organizing Map (SOM) (Kohonen, 1990).
- Multi-Resolution Graph-based Clustering (MRGC) (Ye et al., 2000).

Class knowledge is not required to do an unsupervised classification, allowing you to discover hidden patterns and characteristics in your data. Typically, the researcher does not know the number of clusters in advance, which is why cluster approaches are needed. Electrofacies and petrophysics predictions may be made with the MRGC approach. Establishing the expected number of clusters is essential before applying SOM, DYN, or AHC to the data set. References to the Alameedy et al. (2022) research may be found for anyone interested in learning more about the theory and specifics of these models.

4. Processing Using ML and ANN Methodologies

Using Electrofacies and log predict (Fig. 2) demonstrates the whole modeling process's workflow. The question is "which wells should be considered in order to establish the reference wells?" for the interval to be processed must be asked and answered before any further interpretation can take place in electrofacies workflow diagram and log-predict modeling technique. Different log combinations may reveal unique aspects of the rock, such as its sedimentation, diagenesis, or pore network activity; thus, it's important to look at each set of logs separately to determine its true worth.

Reference with the most comprehensive log suite, the capacity to cross-validate logs and cores, and a wealth of geological information (in terms of expected facies and fluid zones), wells are seen as the gold standard. Separate models are developed for each epoch of interest in the stratigraphic record or fluid zone. The value of every possible combination of logs is not the same. Sedimentation, diagenesis, and the behavior of porous networks are just a few examples of how different log combinations may shed light on these and other aspects of the rock. As long as the core data used to create the model for prediction is representative, there is no reason why predicted core measurements cannot be used as input logs. During the training stage, you will establish a protocol for recording and describing core samples, as well as measuring their dimensions. The next stage is to evaluate the model's capability of decomposing log responses into meaningful electrofacies.

At any point in the modeling process, you may swap between models based on cluster similarity and neural network analysis of the training data. There will need to be several iterations of this calibration process until a good fit is reached.
The data is clustered using either a hierarchical or non-hierarchical clustering technique, both of which are used in this paper as part of the ANN strategy for combining objects with similarity in terms of examined properties into homogenous groups. Hierarchical approaches, which are represented by the clustering techniques, enable the user to choose the optimal number of clusters while still revealing the underlying structure of the data set. Another type of clustering is non-hierarchical clustering, in which the number of groups is set ahead of time and things are put into groups based on how similar they are (Puskarczyk, 2019). One of the first things done in a hierarchical analysis is to generate a distance matrix using either the Euclidean or Manhattan metrics. The first cluster is formed by looking for the items that are most comparable to each other (the closest ones). Distances between objects and the newly formed cluster are recalculated when the distance matrix is shrunk. Iterations of this procedure are performed until all items can be placed into a single cluster (Huang et al., 1996). A tree diagram is used to depict the effects of grouping visually (Fig. 3). The k-means method starts by assuming that there are a fixed number of clusters, and then it chooses cluster seeds at random.

The clusters' gravity means are determined, and then the objects’ distances to the cluster centroids are recalculated so that they may be moved to the nearest cluster. Afterwards, we move the items to the nearest cluster. This procedure is continued until all items have been placed in clusters with the shortest distance between their respective centers of gravity.
In practice, the k-mean method is used to divide large sets of data into groups of similar subsets. This is done by using a hierarchical method to figure out the number of clusters and the starting center of each cluster.

5. Model Verification

The last step in the process of creating a working ML model is to get it validated. This is done by feeding the model with data that wasn't utilized in the training process to see whether it can generalize or if it merely learned the characteristics that were present in the training data. The model is considered "real-world ready" if it performs well on the validation set. There are many error measurements that may be examined to evaluate the model's generalization capacity.

5.1. T-Tests and Minimum Error Analyses

Both provide test statistics, one of which is the t-value, which may be analyzed statistically. It calls for a test of the null hypothesis that both test samples have identical means. When the results of a t-test or regression analysis show that the means are different, the researchers reject the null hypothesis in favor of the alternative hypothesis. The test findings indicate how similar the two samples are and the significance of the discrepancy between the expected and measured mean permeability, using the following formula:

\[
 t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}
\]  

Where, \(x\) is the average, \(\mu_0\) the mean of the whole samples, \(s\) is the standard deviation, and \(n\) is the total data size.

A description of the relationship between the calculated and observed variables is required for data prediction. The expected log value is determined by adding up the individual members' log values and weighting them according to their membership status. Equations 2 and 3 are commonly utilized across many fields; mean absolute error (MAE) and the root mean squared error (RMSE) are used to evaluate the precision of forecasts (Karunasingha, 2022). The best model is picked based on its ability to predict values with the least amount of error.

\[
 RMSE = \sqrt{\frac{\sum (y_i - y_p)^2}{n}}
\]

\[
 MAE = \frac{\sum |y_i - y_p|}{n}
\]

\(y_i\) is real for measured sample, and \(y_p\) is calculated (permeability)

6. Results and Discussion

While many empirical relationships for estimating permeability have been provided thus far, the results of these relationships are often unsatisfactory in different contexts because of the interplay of multiple factors that affect permeability and the absence of certain variables from empirical relations. Because various reservoirs and formations have distinct lithologies and fluids, applying the same relationships to them may not provide the same results. While carbonate rocks make up the vast majority of Iraq's reserves, we still know very little about their petrophysical properties.

For the training data for the offset wells, the models for permeability prediction of the Mishrif formation would be created in two stages: first, the model would be targeted zone by zone (e.g., for each
unit of Mishrif separately, MB11, MB12, MB21, and MC), and then the second stage would be performed on all the units simultaneously. Then, a comparison should be conducted to determine whether the scenario is more accurate in predicting permeability by propagating the model to the well without incorporating reference wells beforehand. Error metrics that may be investigated to evaluate the model's generalization capability could be evaluated further.

6.1. Case 1

The statistics for the units MB11 and MB12 training data are contained in three wells: Number of samples is 341. The four wells will include model logs and the source data needed by Geolog 19 Software for MB11 and MB12 to construct the models. Fig. 4’s crossplots and histograms indicate the connection between the modeled logs (NPHI, RHOB, GR, Rt, Rxo, and permeability).

![Fig. 4. Well logging and permeability crossplots and histogram for units Mb11 and MB12](image)

The crossplots show how two logs relate to each other and how their relationship may indicate certain rock properties, such as lithology and porosity. The histograms provide a frequency distribution of the values for each log, giving an idea of the range and distribution of values in the reservoir. Crossplots and histograms can help find patterns and outliers in the data. The crossplots show the...
correlation between two log measurements, with each data point representing a single depth point in the well. For example, the crossplot of NPHI vs. RHOB indicates a clear positive correlation, with higher values of NPHI tending to correspond with higher values of RHOB. This relationship can be used to help estimate porosity and density in the formation. This gives a better idea of the properties of the reservoir and any problems that might come up during production. For example, if the crossplot of porosity and permeability shows a tight cluster of points, it may show a high-permeability zone that could be targeted for production. On the other hand, if the resistivity (Rt) histogram shows a bimodal distribution, it may mean that different types of fluid are in the reservoir. The histograms show the frequency distribution of a particular variable, which provides information on the range and distribution of values for that variable. For example, the histogram of permeability in Fig. 5 shows a skewed distribution with a long tail towards higher values, indicating that the permeability values are concentrated towards the lower end but have some high values as well.

This automatically generates the cluster model facies based on the clustering of the learning data values based on their facies attributes. Facies comprises one or more groups (Fig. 6) that are optimally formed based on our information. Facings have a designated name, shade, and fill pattern. When a cluster model is constructed, the software effectively creates faces for each grouping. Our research of the four MRGC-suggested regions, which contained 15, 18, 21, and 23 facies, led us to conclude that the optimal cluster consists of 15 facies based on the higher reliable probability and weighted logs. The larger the cluster size, the more precise the predictions; however, running the models is time-consuming and requires a high-speed processor, so it is advised that as little data as possible be employed to save
processing times and memory. Comparing the 15 and 23 cluster facies, the 15 cluster introduced the most reasonable, precise, and near-projected permeability.

Fig. 6. permeability predicted vs core by models (a) AHC, (b) ANN, (c) SOM, (d) MRGC, and (e) DYN for units Mb11 and MB12.
After the optimal clustering and facies processing model has been trained and selected, the data is propagated to the offset well to forecast permeability, keeping in mind the same interval and unit as MB11 and MB12. Table 2 displays the results of three different validation models as measured by the MAE, RMSE, and T test. The MRGC model outperforms the others with a stronger correlation coefficient ($R^2 = 0.94$) and a lower root-mean-squared error (RMSE) of 41.8 (meaning the error between predicted and observed values is as little as possible). Cross-plots of estimated permeability against core permeability by models (a) AHC, (b) ANN, (c) SOM, (d) MRGC, and (e) DYN for units MB11 and MB12 are depicted in Figs. 6 and 7.

### Table 2. Summary of the model’s validation.

<table>
<thead>
<tr>
<th>Metric\Approach</th>
<th>AHC</th>
<th>DYN</th>
<th>MRGC</th>
<th>SOM</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>98.9</td>
<td>108.1</td>
<td>41.8</td>
<td>144.1</td>
<td>162.3</td>
</tr>
<tr>
<td>MAE</td>
<td>26.9</td>
<td>29.9</td>
<td>12.6</td>
<td>38.6</td>
<td>37.8</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.67</td>
<td>0.57</td>
<td>0.94</td>
<td>0.23</td>
<td>0.62</td>
</tr>
<tr>
<td>T</td>
<td>21.39</td>
<td>21</td>
<td>22.7</td>
<td>22</td>
<td>0.43</td>
</tr>
</tbody>
</table>

*Fig. 7. Permeability profiles predicted vs core by models AHC, ANN, SOM, MRGC, and DYN for offset well.*

#### 6.2. Case 2

In this scenario, the training data set consisted of seven wells that spanned all four formation units (MB11, MB12, MB2, and MC). There were 1466 points in the input frame, and all of the logs were complete with the exception of the Rxo, which was lacking values in intervals for some wells. Estimated permeability versus core by models AHC, ANN, SOM, MRGC, and DYN for offset well are present in the Fig. 8, the results show that MRGC tends to be a high correlation between the predicted and observed permeability throughout the higher value of $R^2 (0.94)$. Full-set logs with estimated permeability versus core using the MRGC model for all offset well units are shown in Fig. 9. The majority of core points have fallen on the projected permeability line, showing an adequate level of forecasting accuracy.
7. Conclusions

Estimates of permeability are frequently poor in many circumstances due to the interaction of multiple factors that affect it. Due to the fact that not all reservoirs and formations have the same lithologies and fluids, utilizing the same relationships may not yield the anticipated results.
Creation of a two-stage model for estimating the permeability of the Mishrif formation using training data. While the first case only employed three wells (number of samples: 341) as opposed to the second case’s extensive sample count (number of samples: 1466), it performed better at predicting permeability.

When a cluster model is constructed, the software effectively creates facies for each group. This automatically generates the cluster model based on the clustering of the learning data values. The larger the cluster size, the more precise the predictions can be obtained. By suggesting regions with 15, 18, 21, and 23 cluster facies, the MRGC model outperformed the other models in both cases.
References


