Prediction of Porosity and Water Saturation Using Neural Networks in Shaly Sand Reservoirs, Western Deseret, Egypt

Hamada, G.\textsuperscript{1}, Nyein, C.\textsuperscript{2} and Elsakka, A.\textsuperscript{3}

\textsuperscript{1} College of Engineering, The American University of Kurdistan, Duhok, Kurdistan, Iraq
\textsuperscript{2} Petroleum Engineering Department, Faculty of Geosciences and Petroleum Engineering, University Teknologi Petronas, Malaysia
\textsuperscript{3} Petroleum Engineering Department, Faculty of Geosciences and Petroleum Engineering, University Teknologi Petronas, Malaysia
*Corresponding Author ghareb.hamada@auk.edu.krd

Abstract

Petrophysical properties evaluation of shaly sandstone reservoirs is a challenging task in comparison to clean sand reservoirs. Logging derived porosity in shaly sands requires shale correction and Archie’s formula cannot be used in shaly sands for the determination of water saturation, therefore many water saturation models were proposed to get accurate water saturation of shaly sand reservoirs. In this paper, three water saturation models were used; two empirical models (Simandoux and total shale) and one theoretical model (effective medium model). Shale corrected density log was used in all models. The use of computer-generated algorithm, fuzzy log neural network is of increasing interest in the petroleum industry. This paper presents artificial neural network (ANN) as an effective tool for determining porosity and water saturation in shaly sand reservoir using well logging data. ANN technique utilizes the prevailing unknown nonlinear relationship in data between input logging data and output petrophysical parameters. Results of this work showed that ANN can be supplement or replacement of the existing conventional techniques to determine porosity and water saturation using empirical or theoretical water saturation models. Two neural networks were presented to determine porosity and water saturation using GR, resistivity and density logging data and adapted cut off for porosity and water saturation. Water saturation and porosity were determined using conventional techniques and neural network approach for two wells in a shaly sand reservoir. Neural network approach was trained for porosity and water saturation using the available well logging data. The predicted porosity and water saturation values have shown good matching with the core data in the two wells in comparison to the porosity and water saturation derived from the conventional techniques. This work showed that developed neural network (ANN) could provide an accurate porosity and water saturation in shaly sands reservoirs, it is subject to volume of available well logging data.

Introduction

Shaly sand layer is a sandstone with a considerable amount of clay minerals (shale) distributed in different forms depending geological depositional conditions. The term “shale” is defined and used differently in various disciplines in the petroleum industry. From a geologist’s perspective, shale is defined as a clastic sedimentary rock with a fine grain size below 0.031 mm and having > 35% clay minerals and silt grains; From a reservoir engineer’s point of view, shale is simply classified as low permeability or impermeable formation. The distribution of clay minerals in sandstone rock can be in three forms: dispersed, laminated, or structural with each form having different direct effects on porosity and indirect effects on water saturation. The relationship between different petrophysical properties and fluid saturation is well-known for clean sand reservoirs, developed by Archie 1942 [1]. The existence of shale in reservoir rock is, however, an extremely disturbing factor because it (a)
complicates the determination of oil-in-place, (b) considerably reduces the permeability of the reservoir rock for oil production, and (c) significantly affects the reservoir characterization of shaly sand producing formations. [1,2,3,4,5]

Shale volume or clay content has a significant impact on porosity, permeability and water saturation. In general, shale is indicated by high gamma-ray responses and low resistivity. However, as it was mentioned above, interpretation and estimation of the main petrophysical parameters from log responses with shale consideration is a complex matter. Over the decades, many corrections, correlations, and even new models have been proposed and published for the estimation of petrophysical parameters in shaly sand formations. Nonetheless, there is still, a certain extent of error in manual empirical calculations due to its non-linear nature and assumptions made for each model to account for uncertainties. [5,6,7]

Water saturation is usually derived from resistivity log, however, most of the water saturation models evolved from the very first and basic Archie’s equation with considerations of conductivity contributed by clay or shale presence. The core assumption of Archie’s equation is that the formation is considered as clean sands with no shale or clay presence. This implies that this equation assumes the matrix has zero contribution of conductivity to the formation resistivity. [1,2,5,8,9]

In shaly sand formations, the clay bound water has to be taken into account Peeters [10]. In general, the formation water is saline and have charged ions such as Cl- (anions) and Na+ (cations). The dipole nature of water molecules creates an exchange of cations and anions, conducting a current through the pore water. Thus, the presence of the bound water results in an extra conductivity which distorts the principles of standard log-derived calculations of water saturation. Alruwaili & Alwaheed [11] reported that shale type distribution must also be considered. Many empirical models have been developed for calculation of water saturation in shaly sands and they are dependent on input variables, and type of clay minerals and their distributions. Table 1 summarizes the most commonly used water saturation models that were developed for shaly sand formations. [5,12,13,14,15]

Table 1 Shaly Sand Application for the calculation of water saturation [15]
Analyzing shaly sand formations using water saturation shaly sand models is challenging for log analysts because there is no recommendations or clear limitations for using a certain model than others. There is also no complete sensitivity analysis to test the different water saturation shaly sand models against any errors in the following: porosity, volume of shale, true formation resistivity, shale resistivity, mud filtrate resistivity, water resistivity, spontaneous potential, temperature, and Archie's constants (a, m, n) values. In this regard, neural network could be taken as the rightful tool or model for accurate determination of water saturation and effective porosity of shaly sand layers. [3,13]

Neural network is a computational model of a largely distributed processor with parallel networks which uses input processing units called neurons to predict single output values for non-linear problems. It is a machine-representative of a human brain that can attain and store knowledge and learn through a training process to solve for problems with inputs that are unseen before [16,17]. As a human brain is made of nuclei, dendrites, and axons to convey signals, the imitated neuronal model is also made of 3 main components; (I) a set of synapses or connecting links, each characterized by its "weights" (II) An adder for combining or generalizing the input signals and (III) An activating function to limit the boundaries of an output signal.[18,19,20]

Multilayer Perceptron (MLP) networks can provide a non-linear relationship between inputs and outputs. Therefore, the MLP networks are the most suitable for function approximation. The additional advantage is the ability to learn and generalize with its built-in capability to adapt the synaptic weights to decrease the error. Moreover, this network structure shows great robustness and error tolerance due to its built-in redundancy; even if there are a few faults in its hidden components, the network's overall performance will not be affected, unlike Multiple Linear Regression (MLR) which also predicts relationship among variables, MLP does not tend to reinforce the algorithm for predicted values to lie close to the mean values and hence, it maintains the variable and non-linear nature of the data. [21, 22]

For the above-mentioned advantages, MLP architecture is the best candidate for estimating petrophysical parameters since these parameters cannot be simply derived directly from well log and core data due to their non-linear nature in complex reservoirs. The known petrophysical models and formula are tuned for such reservoirs with many assumptions and the results can mostly be distorted with errors, which, in extreme cases, are negligible. [16, 23, 24, 25, 26]

One of the challenges faced in training MLP is the over-fitting of data where the network memorizes the training data. The performance of the network is reduced when exposed to unseen testing data, to avoid this, data must be allocated into three separate groups for training, validation and testing. Moreover, regression line can be plotted to see the generalization performance of the network. Figure 1 illustrates the artificial neuron and Figure 2 illustrates the perception of the hidden layers used in the design of ANN algorithm. [21,27,28,29,30,31].

Figure 1 Model of Artificial Neuron,[28]
Neural Network Applications in Shaly Sand Evaluation, Elshafei & Hamada [32] achieved a successful result in its utilization of neural networks for predicting porosity, water saturation and hydrocarbon potential index of a shaly sand formation. Jozanikohan et al [33] also conducted a study on the application of MLP network for clay estimation in shaly sandstone reservoir, Shurijeh gas field in Northeastern Iran. The results showed a satisfactory success in estimating clay content using MLP network, and backpropagation algorithm with values close to laboratory data, the network is then validated with unseen core data. Moreover, this study as well conquered the conventional petrophysical and multiple linear regression models. Different network architectures require appropriate learning algorithms where the network structure involves determining: the number of hidden layers and the number of neurons in the layers, type of activation function in hidden and output layers, and selecting the optimization learning algorithm. As there are no clear limitations for using a certain model of the shaly sand models, the objective of this study is to help the log analyst to choose the suitable model for any formation under study based on sensitivity analysis taken place. Accurate prediction of water saturation from ANN depends on the accuracy of the training patterns, which are from the computer processed interpretation (CPI) logs, and the accuracy of the individual log measurements. The idea of using neural networks for fluid saturation is thus not to eliminate the careful petrophysical evaluation behind the CPI log, but to transfer into the neural network for future application the effort and expertise already imbedded in the petrophysical database. Comparison of Sw values of the neural network with those of CPI logs, in wells that are unknown to the network, indicates a standard error of less than 0.03. For porosity prediction we have made a study initially with a single neural network and then by the CM approach. [16, 17, 34,35,36]

This paper focuses on the application of artificial neural network (ANN) as an effective tool for determining porosity and water saturation in shaly sand reservoir using well logging data. ANN technique utilizes the prevailing unknown nonlinear relationship in data between input logging data and output petrophysical parameters. Results of this work showed that ANN can be supplement or replacement of the existing conventional techniques to determine porosity and water saturation using empirical or theoretical water saturation models. Two neural networks were presented to determine porosity and water saturation using GR, resistivity and density logging data and adapted cut off for porosity and water saturation. Water saturation and porosity were determined using conventional techniques and neural network approach for studied two wells in a shaly sand reservoir.

**Used Data and Methodology**

A set of logging data of two wells (Well A and Well B) of Upper Cretaceous shaly sand formation in the Western desert, Egypt was available. This includes five types of well log data: gamma-ray, laterolog deep, density, neutron and photoelectric factor. Core data was available.
for the section 8284 ft to 8373 ft with 90 data points in Well A and 8194 ft to 8378 ft with 179 data points in Well B. Using this data, evaluation of petrophysical properties, mainly porosity and water saturation were done by two approaches; conventional approach and neural network approach. This work starts by determining water saturation and porosity using well logging records of GR, prosody logs and resistivity logging records using conventional methods and in the second stage ANN will be tested and validated to predict porosity and water saturation of shale sand layers.

**Evaluation of Petrophysical Properties of Shaly Sandstone Reservoir (Conventional Approach)**

In this section, using conventional well logging records of porosity logs and resistivity logs, petrophysical properties mainly porosity and water saturation will be determined using conventional methods. Porosity and water saturation for both Well A and Well B were calculated using conventional empirical formula. The results were plotted against depth in comparison with the measured water saturation and porosity from core samples analysis and shown in the figures in following sections.

(a) **Porosity**

Log-derived total porosity was determined using RHOB and NPHI logs as input to the empirical formula as below.

\[ \varnothing_D = \frac{\rho_{ma} - \rho_b}{\rho_{ma} - \rho_f} \]  
(1)

where, \( \varnothing_D \) = porosity density, \( \rho_{ma} \) = matrix density (assumed 2.67 g/cc), \( \rho_b \) = bulk density, \( \rho_f \) = fluid density (assumed 1 g/cc).

\[ \varnothing_t = \frac{\varnothing_D + \varnothing_n}{2} \]  
(2)

where, \( \varnothing_t \) = total porosity, \( \varnothing_D \) = density porosity, \( \varnothing_n \) = neutron porosity (NPHI).

To calculate effective porosity, shale volume must be known. Shale volume was calculated using GR log and correlation [3 and 4], where:

\[ IGR = \frac{GR_{Log\ signal} - GR_{Clean\ rock}}{GR_{Shale} - GR_{Clean\ rock}} \]  
(3)

\[ V_{sh} = 0.33(2^{(2IGR)} - 1) \]  
(4)

Where, IGR = gamma ray index, \( GR_{Log\ signal} \) = gamma ray reading of zone of interest (API), \( GR_{Clean\ rock} \) = gamma ray reading of clean bed (API), \( GR_{Shale} \) = gamma ray reading of shale bed (API), \( V_{sh} \) = shale volume.

\[ \varnothing_{eff} = \varnothing_t (1 - V_{sh}) \]  
(5)

where, \( \varnothing_{eff} \) = effective porosity, \( \varnothing_t \) = total porosity, \( V_{sh} \) = shale volume.

Porosity was calculated for both Well A and Well B by using the empirical formula depicted in equations number (1) to (5). The results were plotted against depth in comparison with the core porosity. In Figure 3, from the plot of calculated porosity versus core porosity of Well A, it can be observed that the conventional approach gave an acceptable accuracy for sections with insignificant shale content. However, for sections with high shale content or very shaly sections, which are from 8310 ft to 8370 ft, the accuracy was very poor when validated with core data. Similarly, in Figure 4, the plot of calculated porosity versus core porosity of Well B, it showed that shaly sections which are from 8240 to 8260 ft and from 8280 to 8300 ft yielded very poor accuracy when validated with core data. This is believed to be due to the distortion in log responses by the presence of shale, which led to errors in linear relationship assumed in log-derived calculations using linear empirical equations.
(b) Water Saturation

Water saturation was determined by the conventional approach using three different models. The calculation of water saturation of shaly sand formations is considered a relatively large area of study with more than 30 commonly used models to be covered. Therefore, three suitable models were chosen to be covered according to the type of shaly sand of the field where the data was taken from, which is laminated shale.

**Simandoux Equation** This model was developed by Simandoux [37] to calculate water saturation of shaly sands based on laboratory experiments on physical reservoir models of artificial sands and clays. This equation works regardless of the type of shale distribution (Simandoux 1963).

\[
S_w = \frac{a R_w}{2 b m} \left[ \left( \phi_{e f f}^m \frac{V_{sh}}{R_t} + \frac{V_{sh}}{R_{sh}} \right) - \frac{V_{sh}}{R_{sh}} \right] 
\]

where, \( S_w \) = water saturation, \( R_t \) = true/deep resistivity (obtain from resistivity log), \( R_w \) = water resistivity, \( m \) = cementation exponent (assumed 2), \( n \) = saturation exponent (assumed 2), \( a \) = tortuosity factor (assumed 1), \( \phi_{e f f} \) = effective porosity, \( R_{sh} \) = shale resistivity, \( V_{sh} \) = shale volume.

**Total Shale Model** This model was developed by Schlumberger [4] based on the previous study by Simandoux and field experience in Niger Delta. This model also can be applied to any type of shale formation.
distribution. However, it does not consider the effect of cementation factor unlike Simandoux, and relatively can reduce the accuracy.

\[
\frac{1}{R_I} = \frac{\phi^2 S_w^2}{(1-V_{sh})^2 R_w} \frac{V_{sh} S_w}{R_{sh}} \tag{7}
\]

where, \( S_w \) = water saturation, \( R_I \) = true/deep resistivity (obtain from resistivity log), \( R_w \) = water resistivity, \( \phi \) = porosity, \( R_{sh} \) = shale resistivity, \( V_{sh} \) = shale volume.

**Effective-Medium Model**  Water saturation equation medium effective model was derived from the Hanai-Bruggeman (HB) equation. This is a theoretical model for water saturation calculation from resistivity and porosity measurements and can be used for laminated and dispersed shale distributions, the assumption in this model is that matrix and hydrocarbon can be treated together as developed by Berg [38]

\[
S_w^n = \left( \frac{R_{sh}}{R_T} \right) \left( \frac{R_T - R_d}{R_w - R_d} \right) \tag{8}
\]

where, \( R_d = \frac{R_s (1-\phi) S_w^n}{(1-\phi)} \tag{9} \)

\[
R_T = \frac{(1-\phi)}{(1-\phi_s)(1-V_{sh})} \frac{(1-\phi)V_{sh}}{R_{sh}} \tag{10}
\]

\[
m = m_{sw} \cdot (1-V_{sh}) + m_{sh} V_{sh} \tag{11}
\]

\( S_w \) = whole-rock saturation, \( R_I \) = partially saturated whole-rock resistivity, \( R_d \) = dispersed phase resistivity (the combination of matrix and hydrocarbons). Assumed values as, \( m_{sw} = 1.8, m_{sh} = 2.7, R_{sh} = 10 \text{Ωm}, R_{sw} = 100 \text{Ωm}. \)

Water saturation was calculated for both Well A and Well B by three shaly sand saturation models Simandoux, Total Shale and Effective-medium as depicted in equations (6), (7) and (8) respectively. In Simandoux equation, effective porosity was used as an input parameter. Thus, shale volume was determined first using equations (3) and (4). With calculated shale volume, effective porosity was determined using equation (5). The results were plotted against depth in comparison with the core data. **Figure 5** shows that Effective-medium model gave the best match with core data for Well A. However, it was difficult to select the best match for Well B since all three models produced similar trends with none of them being a perfect match with core data as seen in **Figure 6**. On the overall, the effective-medium model was considered as best choice according to the plots and logs. Thus, this model’s results will be used for comparison with neural network approach.

**Figure 5** Water Saturation calculated by Different Models, Well A
Determination of Petrophysical Properties of Shaly Sandstone Reservoir (Neural Network Approach)

This section presents application of ANN approach to estimate target petrophysical parameters; porosity and water saturation in shaly sands. Porosity and water saturation for both Well A and Well B were predicted by the neural network approach. The results were plotted against depth in comparison with the core data and conventional method. The two-layered feedforward networks were developed to predict porosity and water saturation, where for porosity, 25 neurons were used in the hidden layer and for water saturation, 30 neurons were used in the hidden layer. Both networks used five parameters as inputs which were well log data: gamma-ray, laterolog deep, density, neutron and photoelectric factor. (Abdideh 2012) reported that in the first network for porosity prediction, the target output used for training was the core porosity and similarly, in the network for water saturation prediction, the target output used for training was the core water saturation. The available data was allocated to three sets which were training, validation and testing: 70%, 15% and 15% respectively. The training algorithm used in this project was Levenberg-Marquardt as this algorithm is mathematically easy to understand and training is time-efficient, especially in data fitting and function approximation problems.

(a) Porosity

Porosity prediction for Well A yielded a mean squared error of 0.002209 and Well B yielded a mean squared error of 0.002771. As observed in Figure 7, the predicted porosity of Well A gave a good match with core data in shaly sections compared to the conventional approach. Similarly, in Figure 8, the predicted water saturation of Well B gave a good match with core data in shaly sections compared to the conventional approach.

(b) Water Saturation

Water saturation prediction for Well A gave mean square error of 0.02141 and Well B yielded a mean squared error of 0.02921. As observed in Figure 9, the predicted water saturation of Well A gave a good match with core data in shaly sections compared to the conventional approach. Similarly, in Figure 10 the predicted water saturation of Well B gave a good match with core data in shaly sections compared to the conventional approach.

Discussion of Results

This part focuses on comparison between water saturation and porosity derived from the conventional and the neural network approaches. For both porosity and water saturation, neural network approach achieved better accuracy, even in sections with high shale content.
(a) Porosity

- The results of porosity determined by both neural network approach and the conventional approach were plotted together in comparison with the core data. In Figure 7, a significant difference of the two approaches in terms of accuracy compared to core data can be observed in very shaly sections which are from 8280 ft to 8340 ft of Well A. Similarly, as seen in Figure 8, shaly sections of Well B which are from 8295 ft to 8335 ft and from 8360 ft to 8373 ft show that neural network approach is significantly more effective. Table 2 shows average porosity from conventional, ANN and core data for wells A & B over the same shaly sand section.

![Figure 7 Porosity from Neural Network and Conventional Approaches, Well A](image1)

![Figure 8 Porosity from Neural Network and Conventional Approaches, Well B](image2)
Table 2 Average porosity comparison from conventional, ANN for well A, B

<table>
<thead>
<tr>
<th>Well</th>
<th>Core (Average)</th>
<th>NN prediction</th>
<th>Calculated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well A</td>
<td>0.133</td>
<td>0.125</td>
<td>0.155</td>
</tr>
<tr>
<td>Well B</td>
<td>0.12</td>
<td>0.10</td>
<td>0.16</td>
</tr>
</tbody>
</table>

(b) Water Saturation

The results of water saturation determined by both neural network approach and conventional approach were plotted together in comparison with the core data. In Figure 9, a significant difference of the two approaches in terms of accuracy compared to core data can be observed in very shaly section ranging from 8284ft to 8340 ft of Well A. Figure 10 shows that neural network approach is significantly more effective in Well B for almost the whole depth section. Moreover, Table 3 shows average water saturation from ANN, three shaly models for well A, B and core data for the same shaly sand section for Wells A and B.

![Figure 9 Water Saturation from Neural Network and Conventional Approaches, Well A](image1)

![Figure 10 Water Saturation from Neural Network and Conventional Approaches, Well B](image2)
### Table 3
Average water saturation comparison conventional and ANN for well A, B

<table>
<thead>
<tr>
<th>Well</th>
<th>Core (Average)</th>
<th>NN prediction</th>
<th>MEF</th>
<th>Simandoux</th>
<th>Archie</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.34</td>
<td>0.32</td>
<td>0.39</td>
<td>0.45</td>
<td>0.58</td>
</tr>
<tr>
<td>B</td>
<td>0.52</td>
<td>0.51</td>
<td>0.55</td>
<td>0.57</td>
<td>0.65</td>
</tr>
</tbody>
</table>

### Conclusion
- Density porosity was determined using conventional methods and water saturation was determined using three water saturation models for shaly sands reservoir. Results have referred to core data values in studied wells in case study of shaly sandstone filed.
- Two neural networks were successfully developed, trained, validated and tested. The first network predicted porosity and second one predicted water saturation. Generated results have demonstrated excellent match with core values and it is found that Neural network with Levenberg-Marquardt Back-propagation algorithm is most reliable to predict porosity and water saturation in shaly sand layers.
- Data Scattering between ANN approach values and the core data has been observed and to minimize such scattering, it is recommended to have much data points in testing level to produce the most suitable ANN algorithm ending with good estimation of the targeted petrophysical parameters.

### Funding sources
This research received no external funding.

### Conflicts of interest
There are no conflicts to declare.

### References


