Screening reservoir candidates for enhanced oil recovery (EOR) in Angolan offshore projects.

RAMOS, G.A.R., ELIAS, B. and YATES, K.

2020
Screening Reservoir Candidates for Enhanced Oil Recovery (EOR) in Angolan Offshore Projects

Geraldo Ramos, Bruno Elias and Kyari Yates

1. Introduction

The success of an EOR implementation requires effective planning on the selection of the appropriate technique for the field or reservoir under investigation. This process involves integration of a set of parameters governing technical and economic performance of a reservoir but not limited to the environmental, commercial, political and governmental factors.

Around the world, researchers, operators and service companies have conducted and published several studies on conventional, advanced and geological EOR screening of reservoir candidates for EOR projects. These include: conventional screening methods or data analysis by using tables and graphs; laboratory work and advanced methods or artificial intelligence.

Taber performed the first study on conventional screening criteria for EOR selection whereas, the first study on advanced methods (AI) was published by Guerillot. The study performed by Taber was updated by Goodlett et al. and later improved by Adasani and Balsamic. Dickson et al., , and . Subsequent to the work published by Guerillot, several works have been published to improve the quality and accuracy of the models. These models are based on fuzzy-logic (FL) and expert system approach, artificial neural network (ANN), least square support vector machine (LSSVM), and very recently the combination of both fuzzy-logic (FL) and neuro-fuzzy (NF).

In this work, an AI based on neuro-fuzzy algorithm approach was employed in the technical selection of a suitable EOR technique for Block B in offshore Angola. The model is based on the five-layered feedforward-backpropagation technique and combines both searching potential of fuzzy-logic (FL) and the learning capability of neural network (NN) to make a prior decision. Three hundred and sixty-five (365) data set from multiple successful thermal, miscible gas, chemical and biological EOR projects worldwide were used in the model. A total of a hundred and twenty-one (121) field data set are mined from Angola Block B; which consists of one field, five reservoirs and seven wells.

2. Neuro-fuzzy data handling Procedures

Neuro-fuzzy (NF) data handling is crucial for accuracy and efficiency of the model. The data is distinguished in input and output data in which the NF is highly dependent. Good training data plays a key role on model optimization which are related to the number and type of the input parameters. To ensure the quality and accuracy of the model, comprehensive data analysis was performed prior to development of a NF model. Most of the data handling system employed in most of the AI models include data acquisition and pre-processing.

A. Data acquisition

Data used in the model was obtained from the successful EOR projects and investigated data from Block B of Angolan oilfield. The input data and predicted output which are the data for successful EOR field projects, was collected from the EOR field data published biannually by Oil and Gas Journal. This includes the reservoir rock-fluid properties in which six (6) parameters: reservoir depth, permeability, porosity, oil viscosity, oil saturation and oil API gravity were considered as input data of the model.

The testing data set (actual output) is collected from Angolan oilfield, Block B as highlighted in Fig.1. The success of the model depends on the quality and quantity of the data available for training, validation and testing process defined mainly by the input and output data set. Fewer data set, sometimes may lead to false results such as over-fitting of the model solution that may not reflect the expected outcome of the model. It should be noted that the training process is crucial for model optimization.

Hence, the data set used for the training process (Table 1) might fill all rows of the data set in the model to avoid errors resulting from missing values in the simulation.

A. Data pre-processing

Data pre-processing is most of the times used in AI models to transform data into a format that fits the purpose of the model. This includes data cleaning, data normalization (organizes data for more efficient access), data sampling (selects a representative subset from a large sample) and data denoising (removes noise from data). The cleaning and sampling are the two-pre-processing steps used in our investigation. The data used in the model was not normalized by the fact that the model uses raw data. To avoid discrepancy in computation, the output value of the objective
function is divided by the standard deviation.

Table 1: Worldwide successful EOR data

<table>
<thead>
<tr>
<th></th>
<th>API</th>
<th>D, m</th>
<th>K, md</th>
<th>D, %</th>
<th>So, %</th>
<th>μ, cP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steam</td>
<td>145</td>
<td>145</td>
<td>134</td>
<td>145</td>
<td>138</td>
<td>141</td>
</tr>
<tr>
<td>CO₂</td>
<td>131</td>
<td>130</td>
<td>129</td>
<td>130</td>
<td>107</td>
<td>128</td>
</tr>
<tr>
<td>Miscibl Gas</td>
<td>37</td>
<td>37</td>
<td>36</td>
<td>37</td>
<td>33</td>
<td>36</td>
</tr>
<tr>
<td>Polymer</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>Combustion.</td>
<td>16</td>
<td>16</td>
<td>14</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Surfactants</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Nitrates</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Microbial</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Hot Water</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Acid Gas</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>364</td>
<td>363</td>
<td>348</td>
<td>362</td>
<td>321</td>
<td>352</td>
</tr>
</tbody>
</table>

Where D= diameter, K=permeability, Φ=porosity, So=oil saturation and μ=viscosity

3. Neuro-Fuzzy Model Development

The NF model adopted in this work is a five (5) layered perceptron feedforward-backpropagation neural networks (Fig. 2). The first and last layers are input and output layers, the intermediate layer, is called the hidden layer and their neurons[33]. The number of neurons for the input and output layers are dependent on the type of problem and the number of input and output variables[33]. The number of neurons and hidden layers are based on the accuracy of the model[33].

The input layer represents the input variables, whilst the output layer (defuzzification) represents the output decision signals. For the defuzzification, center of gravity (COG) and minimum of maximum (MOM) were employed. In the hidden layers, layer two (2) nodes are functioning as input and output membership functions, and layers three (3) and four (4) nodes act as fuzzy logic rules AND, OR respectively[30, 31, 25].

The operation is performed in many simple individual processors called neurons. On each layer, each neuron is connected to the neurons in the preceding layer by direct links, which have their own special weight[34, 28]. Each neuron applies an activation function to its net input to produce its output after receiving signal from the preceding neurons, whilst x represents the input signal to a node; f is an activation function[34, 28].

The operation is performed in many simple individual processors called neurons. On each layer, each neuron is connected to the neurons in the preceding layer by direct links, which have their own special weight[34, 28]. Each neuron applies an activation function to its net input to produce its output after receiving signal from the preceding neurons, whilst x represents the input signal to a node; f is an activation function[34, 28].

Where mᵢ and σᵢ are, respectively, the center (or mean) and the width (or variance) of the jth term of the ith input linguistic variable x. In all three neurons, the three membership functions are implemented which corresponds to the one generating the least error is considered. All the inputs of the layer 2, corresponds to three different grades, namely low (L), medium (M) and High (H) using triangular, trapezoidal, and Gaussian membership functions[26, 27].

![Fig. 2: A typical neuro-fuzzy framework developed in this work representing a 5-layer feed-forward neural network. Adapted from Akanji and Sandrea[30].](image)

The activation functions for each layer are described from Eq. (1) to (5) in which the number of each equation represents the corresponding layer in Fig. 2. The link weight in layers 1, 3 and 4 is unity (1) whereas layers 2 and 5 are mᵢ and mᵢ,αᵢ, respectively.

\[ f = \begin{cases} x, & \text{if } x \leq \mu_i \\ e^{-\frac{x-\mu_i}{\sigma_i}}, & \text{otherwise} \end{cases} \]  
\[ f = e^{\frac{x^2-\mu_i^2}{2\sigma_i^2}} \]  
\[ f = \min(x_1, x_2, ..., x_p) \]  
\[ f = \min(1, \sum_{i=1}^{n} x_i) \]  
\[ f = \frac{\sum_{i=1}^{n} (m_i \alpha_i) x_i}{\sum_{i=1}^{n} x_i} \]

During the learning process, the knowledge extracted from the NF system can be expressed in the form of fuzzy rules by computing weights, number of rules and fuzzy set parameters. These parameters are computed by machine learning process from the EOR data with the input fuzzy sets determined by the fuzzy clustering algorithm. The aforementioned parameters can also be determined by engineers and experts in the field.

The backpropagation algorithm developed by Nauck and Kruse[29] is used to tune all parameters where the error is propagated from the output towards the input units. The mean square error is expressed by Eq. 6[29, 24]

\[ E(\theta, d) = \frac{1}{2} \| \hat{y}(\theta) - d \|^2 \]  
\[ W(t+1) = W(t) - \alpha \frac{\partial E(\theta, d)}{\partial W(t)} \]

Where \( \hat{y}(\theta) \) is the desired output and d is the current actual output. \( \alpha \) represents a learning rate coefficient, set in simulations to 0.01 after error validation sensitivity.

The \( \frac{\partial E(\theta, d)}{\partial \theta} \) for the input and output of the Layer five (5) and two (2) can be determined as described in the work published by Akanji and Sandrea[30]. Hence, the updated value of w can be determined, and the root mean square error (RMSE) and non-dimensional error index (NDEI) is used to evaluate the predicted error defined as:

\[ RMSE = \frac{1}{N} \sum_{i=1}^{N} | \hat{y}(\theta) - d |^2 \]  
\[ NDEI = \frac{RMSE}{\bar{y}(\theta)} \]
This model has some advantages compared to other models due to the fact it uses raw without normalization and does not need to make assumption but matches the pattern from the reservoir field under investigation to the least error data from the successful EOR projects\[31]. It provides the degree of suitability of a typical EOR project obtained from the model prior to full field implementation as well as permits to segregate more oil properties and reservoir characteristics that could impact on EOR projects\[31]. To reduce the cost of function that may lead to prediction that are less robust by using raw data\[32, 33], NDEI was used as decision making in testing process. The model has good performance when run with enough training, validation and testing data sets. Inadequate data may result in over-fitting leading to unexpected results.

4. modelling process

A NF model was designed and developed in C, C++ object-oriented programming platform with the ANFIS\[34]. The mode used a feedforward neural network structure implementation of TSK (Takagi-Sugeno-Kanga) and Mamdani approach\[24, 25] for simulating and technically evaluating the EOR potential of candidate field. The Mamdani approach has the advantage of linguistic interpretability while TSK is computationally efficient (accuracy). Three types of Membership functions were used in this study: Gaussian, triangle and trapezoidal where the leftmost and rightmost values were shouldered. The input and output of the system are represented by X and Y variables where the input vector w(t) measured at time, t, comprises of N components, w(t), i=1, ..., N in which each crisp input variable corresponds to a linguistic variable x_i and partitioned into several overlapping regions labeled with linguistic values\[26]. Each of the variable function is approximated based on prior knowledge using a system of fuzzy rules allowing appropriate initialization with the remaining rules determined by learning from worldwide successful EOR.

The modelling process consists of training, validation and testing stages. The input variables for the NF model consist of training functions where the hidden layer nodes are varied to obtain the lowest root mean square error (RMSE) for training and validation process and non-dimensional error index (NDEI) for testing process. During the training process, sensitivity analysis was performed by employing the developed model based on TSK approach against the Mamdani approach incorporated in the model\[24, 25]. During the training and validation stage, the weights are estimated to minimize the deviations between the actual and predicted outputs, whilst the testing data are used for checking the performance of the model\[26]. The object function of the model is the root mean square error (RMSE) with the threshold designed to be 0.01 and the number of epochs per each training case is set to a maximum number of 2,000. The accuracy of the model was examined by the least RMSE which also leads to a least NDEI\[26].

A. Training and validation process

The available data set from successful EOR projects was randomly divided into two sub-data sets of training (80%) and validation or prediction (20%). This data was used to construct and optimize the model parameters by using RMSE\[26]. The training sub-data sets are the input while the validation sub-data sets are the output.

Two approaches employed in the model were used to perform the sensitivity analysis of the model for the six parameters investigated (depth, porosity, permeability, viscosity, saturation and oil gravity) and five EOR techniques (steam, misc. gas, CO2, combustion and polymer). Besides the six variables and five EOR techniques, three membership functions (triangular, trapezoidal and Gaussian) were employed in training and validation process to determine the optimal model for testing purposes.

The method is based on the TSK and Mamdani approach, generating 1350 runs in which 15 runs for each variable, 90 runs for each technique as described in Ramos & Akani\[29, 30] and Ramos\[31]. This process was performed for other variables and techniques. Subsequently, the best validated data set (with the least RMSE) from the model is then used as predicted output in the testing process. Figs. 3, 4, 5, 6 and 7 illustrate five (5) out of fifteen (15) runs for plots of API for weighted training data, prediction or validation data and associated error for steam using TSK approach.

![Fig. 3: Plots of API for the weighted training data, prediction data and associated error versus number of patterns for steam-option 1](Image)

![Fig. 4: Plots of API for the weighted training data, prediction data and associated error versus number of patterns for steam-option 2](Image)

![Fig. 5: Plots of API for the weighted training data, prediction data and associated error versus number of patterns for steam-option 3](Image)
The error computation is critical to ensure that the NF technique is suitable for the EOR process or technique under investigation. The developed model performed satisfactorily when run with enough training, verification and testing data sets. The degree of suitability of a typical EOR project obtained from the model prior to full field implementation as well as permits to segregate more oil properties and reservoir characteristics that could impact on EOR projects. The formation type is not included in the model, however, this can be determined by screening criteria from the successful EOR worldwide field data set published in the literature.

B. Testing process

The testing process was performed by using the available data set from Angolan oilfield (Blocks B), the actual output, and the predicted output (the available data set of the parameters from the successful EOR field projects), represented by the least RMSE from the training and validation process while the decision making was based on NDEI. In this stage, the investigation of the model performance and accuracy is employed. During the testing process, the best validation data set generated during the training and validation process is used as validation data set or predicted output in the testing process.

Three scenarios were considered and investigated in this project: (1) NDEIs ≤ 10%; (2) 10 < NDEIs ≤ 20%; and (3) 20 < NDEIs ≤ 30%. This procedure was performed for the six variables investigated (API, depth, porosity, saturation, permeability and viscosity) for five EOR techniques (miscible gas, steam, CO₂, polymer and combustion). The defined scenarios cannot be considered as threshold as this is not a binary decision operation, and engineering knowledge of the process is required in decision-making.

As an example, variables like viscosity and depth for thermal processes (steam and hot water), pressure for gas and steam injection, temperature for chemical and hot water are very sensitive and critical[31]. Permeability is not a critical variable for gas injection[11].

The NF results from the simulation for Block B illustrated in Table 2 shows that the steam and CO₂ fail for the investigated EOR techniques for all scenarios. In contrast of miscible hydrocarbon gas, polymer and combustion that are found to be suitable for all scenarios. However, laboratory test, simulation and pilot test are required to confirm the suitability of the results obtained from the NF screening model.

<table>
<thead>
<tr>
<th>Description</th>
<th>Steam</th>
<th>Gas</th>
<th>CO₂</th>
<th>Polymer</th>
<th>Combustion</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDEIs ≤ 10%</td>
<td>X</td>
<td>V</td>
<td>X</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>10 &lt; NDEIs ≤ 20%</td>
<td>X</td>
<td>V</td>
<td>X</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>20 &lt; NDEIs ≤ 30%</td>
<td>X</td>
<td>V</td>
<td>X</td>
<td>V</td>
<td>V</td>
</tr>
</tbody>
</table>

5. Conclusion

A five layer feedforward-backpropagation model based on TSK and Mamdani has been trained and validated to obtain the optimal model for testing oilfield data from Block B in offshore Angola. The sensitivity analysis of the two approaches was employed using the successful EOR data and the least RMSE from the trained and validated model for each parameter and NDEI was used in the testing process with TSK approach being more accurate than Mamdani approach.

The results obtained from the NF model show that this model presents strengths that can be considered as robustness required for screening reservoir candidates. Several advantages compared to other models have been identified due to the fact of the model uses raw data and no assumptions needed but matches the pattern from the data under investigation.

The result of the screening process within the five EOR techniques presented miscible gas, polymer and combustion as the most suitable techniques whereas CO₂ and steam injection were not suitable within the investigated range. Therefore, additional investigation such as laboratory tests, simulation and pilot tests are recommended to confirm the results obtained from the model.

Acknowledgment

This study is sponsored by the Angolan National Oil Company (Sonangol EP) and the authors are grateful for their support and the permission to use data and publish this manuscript.

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

References


