Estimation of Reservoir Simulation Responses for Different Relative Permeability Curves using Machine Learning

Yong Do Kim

1 Introduction

In oil and gas reservoir simulators, using reasonable relative permeability curves is crucial to properly represent reservoir models and predict future reservoir performance. Initial relative permeability curves can be obtained from core flooding experiments or estimated using various correlations. However, applying the initial relative permeability often leads to significant errors between simulation results and actual production history data due to inherent uncertainties in the initial relative permeability.

In practice, the initial relative permeability curves are modified in history matching to decrease discrepancy between simulation results and actual production history. History matching process requires running a large number of simulations, of an order of 100-10,000. Performing such a large number of simulations is challenging because running each reservoir simulation is computationally expensive.

In this study, linear regression and neural network fittings are explored to estimate reservoir simulation responses (output) for different relative permeability curves (features). The trained model enables fast computation of approximate solutions to various relative permeability curves without running expensive reservoir simulations. Thus, it can be used to accelerate history matching process.

2 Related Work

Implementing machine learning algorithms is becoming popular in the oil and gas industry applications. Trehan [7] constructed error indicators from a small portion of reservoir realizations for upscaled models using high-dimensional regression techniques (random forest and LASSO). By using the error indicators from a few realizations, he corrected coarse model solutions from other remaining realizations. Tian and Horne [6] used kernel ridge regression to predict reservoir pressure at different well locations based on observed flow rate or to estimate missing flow rate history based on measured pressure data. Yeten et al. [8] and Guyagulera et al. [2] conducted well placement optimization using Artificial Neural Network (ANN) proxies. They demonstrated using ANN proxy reduced the number of reservoir simulation runs required. Silva et al. [4] conducted history match of horizontal permeability, porosity, and net to gross ration using proxies generated by ANN. The history matching result obtained with the proxy provided the similar result from the reservoir simulator.

3 Dataset and Features

In this study, water and oil relative permeability are calculated based on Corey-type relationship as follows:

$$k_{rw}(S_w) = krw^o(\frac{S_w - S_{wir}}{1 - S_{wir} - S_{or}})^{n_w}, \qquad k_{ro}(S_w) = kro^o(\frac{S_w - S_w - S_{or}}{1 - S_{wir} - S_{or}})^{n_o} \tag{1}$$

Here, $k_{rw}(S_w)$ and $k_{ro}(S_w)$ are relative permeability of the water and oil phase at specific water saturation, S_w, respectively. S_{wir} is the irreducible water saturation, S_{or} is the residual oil saturation, k^o_{rw} is maximum k_{rw} measured at $S_w = 1 - S_{wir} - S_{or}$, k^o_{ro} is maximum k_{ro} measured at at $S_w = 1 - S_{wir} - S_{or}$, and n_w and n_o are Corey water and oil exponents. In this study, n_w, n_o, S_{wir}, S_{or}, k^o_{ro}, and k^o_{rw} are uncertain relative permeability parameters. The goal of this study is to find an appropriate hypothesis using linear regression and neural network fitting to predict reservoir simulation responses for different relative permeability parameters. We defined feature $x^{(i)}$ as follows:

$$x^{(i)} = \begin{bmatrix} 1 \\ n_w \\ n_o \\ S_{wir} \\ S_{or} \\ k^o_{ro} \\ k^o_{rw} \end{bmatrix}, \quad x^{(i)} \in \mathbb{R}^7 \tag{2}$$

Target $y^{(i)}_f$ is a vector with N entries containing amount of cumulative fluid f (f = oil or water) produced at N different times ($y^{(i)}_f \in \mathbb{R}^N$).

Using Stanfords reservoir simulator (AD-GPRS) [5], 250 set of reservoir simulation outputs with different relative permeability parameters are obtained. The first 50 set is used for linear regression and neural network training. The remaining 200 set is used to test estimation accuracy.

4 Methods

4.1 Linear Regression

Linear regression attempts to find linear relationship between features and output by fitting a linear line based on observed data. The cost function of linear regression model is:

$$J(\theta) = \frac{1}{2}\sum_{i=1}^{m}(\theta^T x^{(i)} - y_f^{(i)})^2, \tag{3}$$

The value of θ^* that minimizes $J(\theta)$ derived from closed form solution is:

$$\theta^* = (X^T X)^{-1} X^T Y, \tag{4}$$

Using θ^*, where $\theta^* \in \mathbb{R}^{7\times N}$ in this problem, the trained liner regression model predicts reservoir simulation output for different relative permeability parameters.

4.2 Artificial Neural Network

Artificial Neural Network (ANN) is a machine learning algorithm that imitates how human brain process information. ANN consists of an input layer, hidden layers containing internal neurons, and an output layer. In ANN, multiple pieces of information (features) are transmitted to internal neurons. The neural network performs complex calculations and generates outputs. In this study, ANN fitting is used to estimate reservoir simulation output using MATLAB Neural Network Toolbox [3]. The neural network implemented in this study contains one sigmoid hidden layer. We connect the seven features introduced in Equation (2) to three neurons in the hidden layer, and train the neural network using Levenberg-Marquardrdt [1] backpropagation algorithm.

5 Result

The two-dimensional 45 $\times$ 45 geological model depicted in Figure 1 is used in this study. Four production wells (red circle) are located around each corner of the reservoir model. The four production bottomhole pressure are maintained as 190 bar. One injector (blue circle) is placed at the center of the model. Water is injected with constant bottomhole pressure 210 bar for 2500 days. All simulation parameters are assumed to be accurate except for relative permeability data. Trained linear regression and neural network models estimate cumulative oil or water recovery every 100 days from day 0 to day 2500 ($y_f^{(i)} \in \mathbb{R}^{26}$).

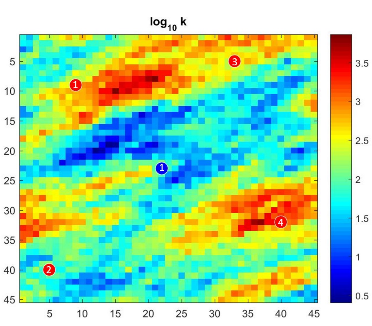
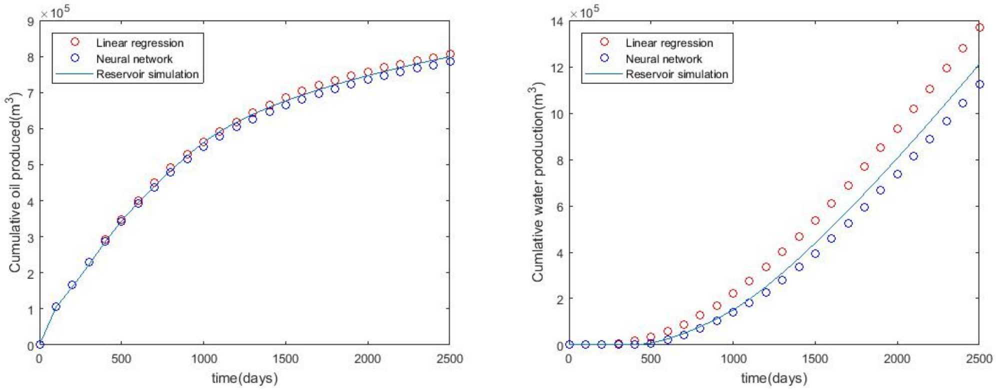


Figure 1: Permeability field ($\log_{10} k$, k in mD) and well configuration.

Figure 2 shows cumulative oil recovery and water recovery obtained from linear regression, neural network, and actual simulation output from one of 200 test set. In this case, the cumulative oil recovery and cumulative water recovery show great match between the estimation using the machine learning algorithms and simulation output.



(a) Cumulative oil recovery

(b) Cumulative water recovery

Figure 2: Estimation from linear regression and neural network and actual simulation output

We find the test set accuracy by comparing cumulative oil recovery solutions estimated using linear regression and neural network with the simulation output. In order to find error for this time varying parameter, we first Normalize the difference between machine learning based estimation, Q_{ML}, and simulation output, Q_{Sim} for different time steps using Equation (5).

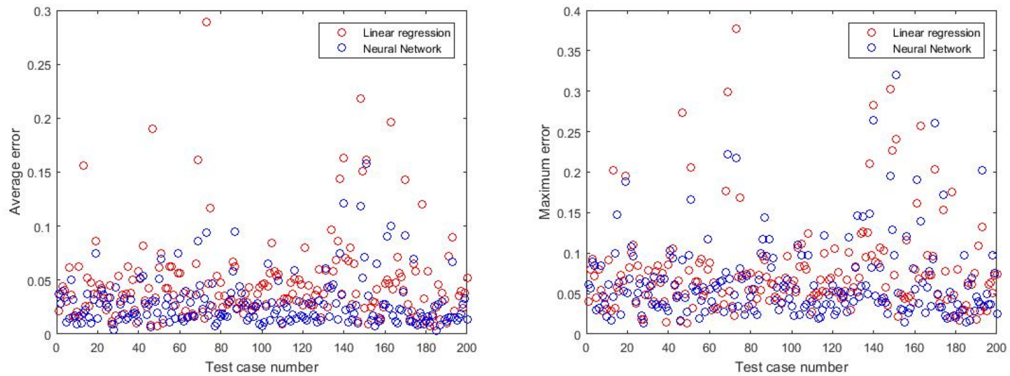
$$Error(t_n) = \frac{Q_{ML}(t_n) - Q_{Sim}(t_n)}{Q_{Sim}(t_n)} \tag{5}$$

Then, we report average $Error(t_n)$, Equation (6), and maximum $Error(t_n)$, Equation (7), for different test cases.

$$Ave.Error = \frac{\sum_{n=1}^{N}(Error(t_n))}{N} \tag{6}$$

$$Max.Error = \max_{n} Error(t_n) \tag{7}$$

Figure 3 shows the the average and the maximum $Error(t_n)$ from linear regression and neural network estimations models for each case. The error plots indicate that most of the test case shows acceptable error, less than 10 %. In addition, the estimation from neural network fitting is more reliable. 98% and 93.5% of test set from neural network and linear regression show the average error less than 10%, respectively. This might be because higher dimensional of the features are required to predict reservoir simulation output more precisely.



(a) Average error in cumulative oil recovery (b) Maximum error in cumulative oil recovery

Figure 3: Test set error

6 Conclusion

We apply linear regression and neural network fitting to estimate reservoir simulation response for different relative permeability parameters. For the example case considered here, both linear regression and neural network fitting show acceptable estimations from the most of test cases. The estimation from machine learning algorithms can be used in history matching process of relative permeability to provide significant computational savings.

References

[1] Howard B. Demuth, Mark H. Beale, Orlando De Jess, and Martin T. Hagan. *Neural Network Design*. Martin Hagan, USA, 2nd edition, 2014.

[2] B. Guyaguler, R.N. Horne, L. Rogers, and J.J. Rosenzweig. Optimization of well placement in a gulf of mexico waterflooding project. *SPE Reservoir Evaluation & Engineering*, 5(03):229–236, June 2002.

[3] Matlab Neural Network Toolbox, MATLAB R2017a.

[4] P. C. Silva, C. Maschio, and D. J. Schiozer. Application of neural network and global optimization in history matching. Petroleum Society of Canada. doi:10.2118/08-11-22-TN, November 2008.

[5] H. Tchelepi and K. Aziz. *AD-GPRS Manual*. Department of Energy Resources Engineering, Stanford University.

[6] C. Tian and R. N. Horne. Machine learning applied to multiwell test analysis and flow rate reconstruction. Paper SPE 175059 presented at the SPE Annual Technical Conference, Houston, Texas, 2015.

[7] S. Trehan. *Surrogate Modeling for Subsurface Flow: A New Reduced-order Model and Error Estimation Procedure*. PhD thesis, Department of Energy Resources Engineering, Stanford University, 2016.

[8] B. Yeten, L.J. Durlofsky, and K. Aziz. Optimization of nonconventional well type, location, and trajectory. *SPE Journal*, 8(13):200–210, 2003.