Development of Water flood model for Oil production enhancement

Msc Student J.J Tsvaki
NNTC Dir PhD D.T Tailakov
Prof E.N Pavlovskiy

Faculty of Mathematics and Mechanics
Novosibirsk State University

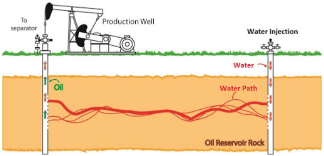
April 12, 2020

Outline

- Introduction
- Objectives
- Literature review
- Methodology
- Results
- Conclusions and recommendations
- References

Introduction

Waterflood Model



In oil production, waterflooding is by far the most widely used secondary recovery method in the oil industry.

Water flooding involves pumping water through an injection well into the reservoir. The water then forces itself through the pore spaces and sweeps the oil toward another set of wells known as producers.

As a result, there is an increment in the total oil production from the reservoir. Water injection increases oil production (known as the recovery factor) and maintains production rate of a reservoir over a longer period.

Research Problem (contd...)

Determining water injection rates which maximizes oil production

Setting proper water injection rates for the injection wells is a key factor to successfully operate an oil field under water flooding.

The success of such activity could

- (a) reduce water cycling at field, section and pattern levels;
- (b) improve water/oil ratio (WOR)
- (c) improve oil production and recovery by directing water injection to specific zones and areas; and
- (d) reduce OPEX by improving water utilization.

Objectives

- To provide a model for water flood field using artificial neural networks
- To provide model that explains daily water injection and daily oil production relationship.
- Providing daily water injections rates which maximizes daily oil production.

Literature Review

Waterflood Model

There are a number of artificial neural networks which have gained d popularity in time series modelling in recent studies such as [Sigurd Øyen. (2018), Pavel Temirchev. (2018), Chukwuka G. Monyei, Aderemi O. Adewumi, Michael O. Obolo, (2016), Elmabrouk E. Shirif R. Mayorga, (2014)].

They are as following:

- Multi-Layer Perceptron
- Convolution Neural Networks
- Long Short Term Memory
- Gated Recurrent Neural Networks

Model Improvement (contd...)

Model Improvements

- Data Augmentation
- Regularization

Data Augmentation (Mixup)
Mixup is a neural network training method that generates new samples by linear interpolation of multiple samples and their labels.

Mixup constructs virtual training examples

$$\hat{x} = \lambda x_i + (1 - \lambda) x_j \quad (1)$$

$$\hat{y} = \lambda y_i + (1 - \lambda) y_j \quad (2)$$

where x_i, x_j are raw input vectors and y_i, y_j are one-hot label encodings

(x_i, y_i) and (x_j, y_j) are two examples drawn at random from our training data, and $\lambda \in [0, 1]$

Water flood Optimization (cont...)

An optimization problem refers to maximizing or minimizing a real function by systematically selecting input values from a domain set and providing the value of the function.

A generalized model is as follows in mathematical terms finding n variables $(x_1, x_2, ..., x_n)$ which (maximizing or minimizing) the objective function $f(x)$,where

$$f(x) = f(x_1, x_2, ..x_n)$$

Algorithms for unconstrained nonlinear optimization are as follows

- Quasi Newton
- Nelder-Mead Method
- Trust Region Method

Methodology

The CRISP-DM "Cross Industry Standard Process for Data Mining" approach for data mining project management was used

Waterflood data was rearranged using sql operation from daily data into tables with respective rates by dates.

Data to be used in the study were from 21 December 2005 to 31 December 2016 where all the variables contained values.

Given that daily oil production is affected by 3 days' daily water injections prior, 1 lagged and 2 lagged water injection rates were selected for modeling

Methodology(cont...)

Daily water production and daily gas production data was also included to fully explain water flood activities. 36 events that took place during that time such as acid treatment, perforations, cement squeeze, leak repairs, etc. were also included in the model

Data Modeling

- MLP (3 fully connected,normal initializer and ReLu activation function,adam optimizer,mean absolute error- loss function,accuracy and mean squared error-metrics)
- CNN (2layers of 1 dimensional CNN,2 dropout,reLu,1 layer maxpooling, 2 dense layers,loss,metrics same as above)
- LSTM - (2 layers LSTM, 2 dropout,2 dense layers, RMSProp optimizer,loss,metrics same as above
- GRU RNN (2 layers GRU, 2 dropout,2 dense layers, RMSProp optimizer,loss,metrics same as above

Model Review:
To improve the performance of the models tabular data mixup was done in all models to increase datasets as neural network models work effectively and efficiently with huge data samples.

The interpolating factor is 0.6 increasing samples 5.6 times. Same model architectures were used to check improvements on model performances.

Water Injection Oil Production Model:
The complexity of nonlinear relationship between daily water injection and daily oil production is described by a two layered multilayered perceptron whose model is given by:

$$f(x) = G(b^{(2)} + W^{(2)}(s(b^{(1)}) + W^{(1)}x)))$$

where G and s are ReLu activation functions, $G = max(0, x)$, $s = max(0, x)$, parameters are $biases = b^{(1)}, b^{(2)}$ and $weights = W^{(1)}, W^{(2)}$

Methodology(Optimization)

Stage I: Develop the function

Water Injection Oil Production Model $f(x)$

$$f_M{}^{oil} \longrightarrow max$$

$$M - model(MLP)$$

$$f_M{}^{oil}(x_1, x_2, ..., x_n) = oil$$

Stage II: Maximize or minimize the function.
To maximize the function, the Nelder-Mead algorithm is used which is an optimization method unconstrained multidimensional variable without derivatives. It is initiated using scipy in python at 1 at all vertices. In order to maximize the function, the original function was negated and then minimized.

Results

A. Waterflood model

MLP	loss: 0.1425 - accuracy: 0.5879 - mean_squared_error: 0.0768
MLP Augmented	loss: 0.1257 - accuracy: 0.5480 - mean_squared_error: 0.0621
CNN	loss: 0.1249 - accuracy: 0.6480 - mean_squared_error: 0.0678
CNN Augmented	loss: 0.1194 - accuracy: 0.5543 - mean_squared_error: 0.0581
LSTM	loss: 0.1229 - accuracy: 0.6542 - mean_squared_error: 0.0660
LSTM Augmented	loss: 0.1212 - accuracy: 0.5565 - mean_squared_error: 0.0566
GRU RNN	loss: 0.1221 - accuracy: 0.6566 - mean_squared_error: 0.0665
GRU RNN Aug	loss: 0.1199 - accuracy: 0.5564 - mean_squared_error: 0.0620

GRU RNN out performed other models in accuracy and mostly importantly mean square error which is the proper metric for regression problems.

Data augmentation in all models helped in lowering the mean squared error and the loss function which are the main metrics of regression problems although it lowered the accuracy.

Results(cont...)

Water Injection Oil Production Model
The input features were daily water injection rates and output were the daily oil production rates. The MLP was 63.88 percent accurate and 13,5 mean absolute error (not normalized).

Optimization The objective of optimization task was to find daily water injections which maximizes daily oil production.

After the Nelder Mead optimization algorithm the the daily water injection rates for 577 wells which maximizes daily oil production for the 1377 oil production wells were recorded.

Results(cont...)

The prediction for the first entry was 1201.5874999999999 whereas it was actually 1296.776000 and using the proposed model

After calculating sum of daily oil production rates. The maximum oil produced per day was 27494.710000000003. Comparing with the estimated for maximum production. Our daily injection rates(with max of 160 vol per well) shows to be superior to 90.443 percent to the entire data set with regard to oil production

```
x_hat = x_hat.reshape(1,556)
```

```
y_hat = model.predict(x_hat)
```

```
y_hat =y_hat[np.nonzero(y_hat)]
```

```
y_hat.sum()
```

```
25124.812100000003
```

Conclusion and Recommendations

Combinations of different types of neural networks can produce better results as they noted in their report that complex methods led to greater accuracy.

It was also noted that most of the top performing models were combinations of statistical methods implying there is still potential to maximize artificial neural frameworks which is not utilized with how they have been able to solve complex and almost impossible tasks in the past in time series data.

More data and variables might be needed to fully explain waterflood fields as they have other factors that affect their operations such as rock types, porosity, etc. which might improve the quality of the model

References

Due to its length could not be presented on this slide.They will be provided in the paper