Society of Petroleum Engineers

SPE-190104-MS

Comparison of Decline Curve Analysis DCA with Recursive Neural Networks RNN for Production Forecast of Multiple Wells

J. Sun, X. Ma, and M. Kazi, CSE ICON

This paper was prepared for presentation at the SPE Western Regional Meeting held in Garden Grove, California, USA, 22-27 April 2018.

Abstract

Production forecast can significantly influence field development planning and economic evaluation. Traditional methods including numerical simulations and decline curve analysis models (DCA) requires extensive domain knowledge or lack of flexibility in modeling complex physics. However, data-driven techniques using recursive neural networks (RNN) have proven very efficient and accurate in time-series forecasting related applications. This study implemented and compared RNN with DCA in production forecast of single and multiple wells.

A typical RNN based long short-term memory (LSTM) models were first developed with various input and output sequences. Then, well-known DCA models such as Duong, Stretched Exponential Decline (SEPD), Power Law Exponential Decline (PLE) were implemented as reference solutions. Moreover, data cleaning process involves preparation of history production rates and well constraints for existing wells. For multiple wells, similar input parameters were aggregated together for adjacent wells before declining forecast using the former model. Finally, hold-out training and validation were performed, followed by comparison of model accuracy and efficiency.

Various LSTM based sequence-to-sequence models such as one-to-one, many-to-one, and many-to-many were successfully implemented for production forecast. Feature engineering was performed to generate additional features to facilitate training process. It was observed better agreement for the blind-forecasting validation dataset (i.e., last 20% of the given history) between LSTM model prediction and history production than DCA based models. LSTM models captured the overall trend whereas DCA only produced smooth curves. In addition, LSTM based models yielded good matches for all three-phase rates whereas DCA was usually limited to a certain phase. Moreover, for multiple wells, a group of neighboring wells with variable history lengths were used for training the model to forecast the production rates, where the modeling process is similar as character translation in natural language processing. Finally, it was demonstrated that the developed RNN based sequence-to-sequence models will be readily extended to model other time-series related problems such as condition-based maintenance and failure prediction.

This study proposed a novel approach to model time-series related problems (e.g., production forecast) using the RNN based sequence-to-sequence models. The developed data-driven approach makes the process of history matching and forecasting efficiency and accurate for assets with or without decent operation

history information. In addition, the algorithms and case studies herein were developed with open-source libraries, which could be readily incorporated into either in-house or commercial packages.

Background

Natural gas and oil contained in shale formations are an important part of a balanced and sustainable national energy strategy, especially after horizontal drilling and hydraulic fracturing technology facilitated economic production in recent years. The global demand is still growing and there is very clear that the unconventional resources is playing major role to supply this increasing demand (EIA,2017).

There always been a challenge to predict production and give an estimation for unconventional reservoirs. The complexity of geological and reservoir data, combined with dynamic operation management events, and rapid production decreasing rate because of low-permeability, these characteristics make it hard to use conventional methods to give production predictions. Decline Curve Analysis (DCA) is traditionally and recent popular method that used to provide deterministic estimates for future performance and remaining reserves for shale reservoirs (Aprs, 1944; Fetkovich, 1980; Duong, 2011). But unfortunately, this deterministic estimate contains significant uncertainty and far from the actual future production trend.

Even though there are lots of unknown and uncertainty to describe how the unconventional reservoir works, there are thousands of new wells drilling in several shale plays, such as Eagle Ford, Permian Basin. And these wells provide us massive amount of data, either statics or dynamics type, which is very helpful to build up reservoir models and history matching jobs as complex numerical simulations (Valkó, 2009).

However, build a reliable reservoir numerical model for a field with hundreds of wells based on geological and physics principle are time-consuming and cost-inhibitive. We are looking for an alternative technique which could give a quick solution to predict production more accuracy.

Thus, the implementation of data-driven models and workflows that feed into a data mining methodology is a feasible method for us to test, since there already have massive amount of static and dynamic data from shale fields (Mohaghegh, 2009, 2011). Machine learning models is crucial to determine a probabilistic production forecasts and estimates for well performance.

Machine Learning

Machine learning is still a new field of topic to oil and gas industry because of its data-driven base and non-dependency on governing physics. The machine learning model has an ability to learn and predict without explicit assigned algorithms, which is already widely employed in other industries. There are lots of machine learning model built for regression or classification problems which focuses on prediction-making using computers. In this paper, we will focus on a very popular method - neural networks for the production prediction model.

Neural Network

Neural network(NN), or artificial neural networks (ANN), are a family of deep learning models inspired by biological neural networks in machine learning and cognitive science and are used to estimate or approximate functions that can depend on a large number of inputs and are generally unknown (Schmidhuber, 2015). A typical neural network has artificial neurons called units arranged in a series of layers, each of which connects to the layers on either side (Figure 1). The input units contain various information from the outside world, and process to output units which shows how the information this network learned. There are one or more layers of hidden units between the input units and output units that form most of the artificial brain. The connections between one unit and another are represented by a number called a weight, which could be represent by a general formula:

$$a^l = \sigma(W^l a^{l-1} + b^l) \tag{1}$$

Where a^{l-1} and a^{l} are the neurons in layer l and l-1, W^{l} is weight matrix for each layer l and connect with all neurons in this layer, b^{l} is the bias vector and σ is the activation function define outputs on or off depending on inputs. This expression gives an intuitive way about how the activations between layers: we just apply the weight matrix to the neuron cells, then add the bias vector, and finally apply the activation function.

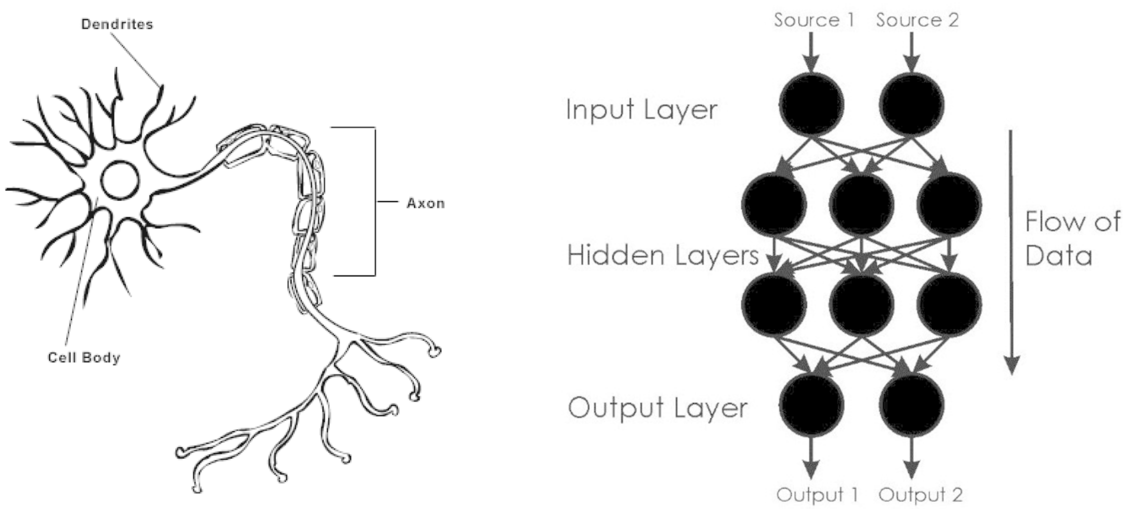

Figure 1—A typical neural network representation based on a biological neuro.

Neural network could solve regression and classification problems and has already shown great potential in terms of efficiency and accuracy based on special feature modeling. Here we use recurrent neural network (RNN) to build production prediction model, since RNN is very suitable to exhibit dynamic temporal behavior where connections between units form a directed cycle to transform time series information (Schmidhuber, 1993).

RNN-LSTM

Unlike regression predictive modeling, time series prediction problems are difficult with the complexity of a sequence dependence among the input variables. RNN is a powerful design to handle sequence dependence but has timeframe length limitation (Das, 1992) and the Long Short-Term Memory network (LSTM network) (Hochreiter, 1997) is a type of RNN used in deep learning which has very large architectures can be successfully trained.

It could fully utilize the capacity of the full history data due to the improved implementation of the "gates". There are three gates to control information process, forget gate, input gate and output gate. These gates can be turned on or off for adding result to the calculation of the current layer depends on output at this layer reaches the threshold or not. In addition, not as the traditional neural networks, where the inputs are mapped to outputs alone, the LSTM can learn a mapping function for the inputs over time to an output, i.e., a temporal dependence relationship.

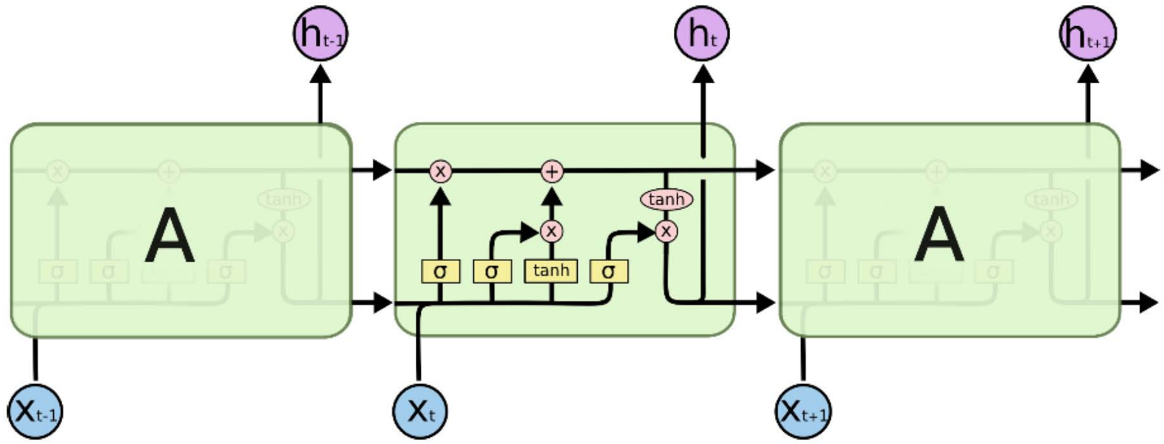

Figure 2—The repeating module in an LSTM contains four interacting layers (Olah, 2015).

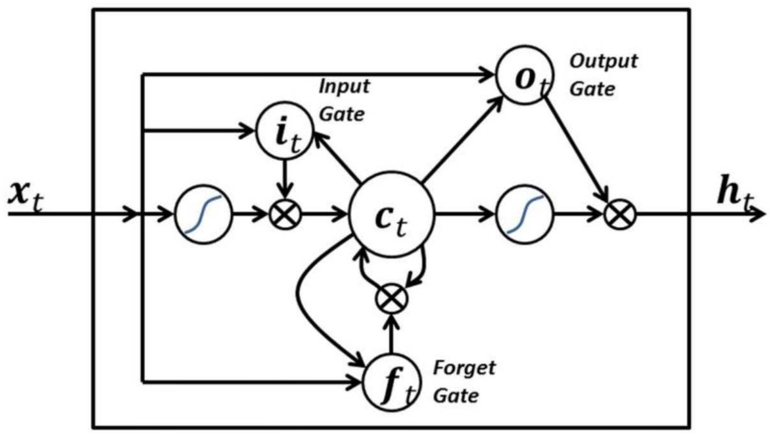

Figure 3—A peephole LSTM unit with input, output, and forget gates.

In this paper, we implement time series models for forecasting either short-term or long-term well production based on available LSTM algorithms in open-source libraries (Python & Keras Theano). Since production history are streams of data in a time series, LSTM is applied to build time series regression model to predict future production based on well performance history and other operational constraints.

Model Setup

Input and Output

The production dataset used in this paper is from Eagle Ford Shale play in west Texas with around 600-800 days history. Daily oil, gas and water production data are all recorded consist with the tubing head pressure as operation constraint. These production data are the main input variables for the neural network model, and it gives multi-phase production predictions in the near future as model output. Well head pressure is also considered as input variable since it contains information that what reservoir performance response to different operation constraints.

The base model is built for a single well prediction forecast. Moreover, an aggregate well scenario is also considered to improve the model with several nearby well production histories added as more input variables. Same outputs are selected to investigate whether the developed models show consistent results in terms of accuracy and efficiency or not.

Workflow

In this paper, there are two scenarios for neural network models: first, multi-phase production prediction for a well based on itself existing production history; second, production prediction for a well based on itself and nearby wells' production history and operation constraints. Here is the general workflow sketch for these scenarios:

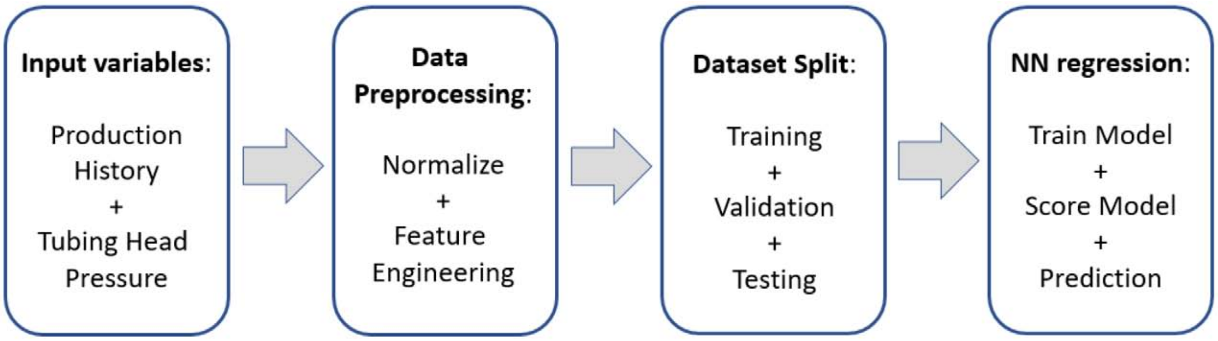

Figure 4—Workflow for production prediction implemented neural network model.

For each scenario, the input variables are the production history from a single well or multi-wells, and plus the tubing heard pressure as operation constraint. These data should be preprocessed and normalized before neural networks built. And the dataset will be split in to three groups, training, validation, and test for building ANN models. Here we choose 80% data as training, 10% to validate model and 10% for the blind test. A regression model with future multi-phase production rates along time series will be calculated from a good trained ANN model.

Build Model and Hyperparameters Setup

To train RNN-LSTM model, we need consider the information from time series and adapt it into the input variables. Here we set inputs to be three dimensions, which contains the number of training samples, the multiple lag number of timesteps, and the number of input features. In this model, we consider 3 days as previous time steps to reconstruct the inputs.

Return to Eq. 1, W and b are model weight and bias matrix which could be learning during the model training step. And to train good model, we choose the activation function "relu", which is f(x) = max(0,x). This function is used for the input layer and two hidden layers only except the output layer. It could speed up the training neural networks with simple gradient computation (either 0 or 1 depending on the sign of x).

The hyperparameters are tested several runs to fit the model with good simulation results. We use 30 and 18 neurons for the hidden layers, and dropout rate is 0.2 to avoid the overfitting. Also, we use mean squared error (MSE) as the loss function and L2 as the kernel regularization constraint.

$$\mathrm{MSE} = \frac{1}{n}\sum\left(y^{\mathrm{pred}} - y^{\mathrm{actual}}\right)^2 \tag{2}$$

Single Well Results and DCA Comparison

Figure 5 and 6 gives a comparison between actual production rate and simulated production rate by ANN model, which shows very close values and trend.

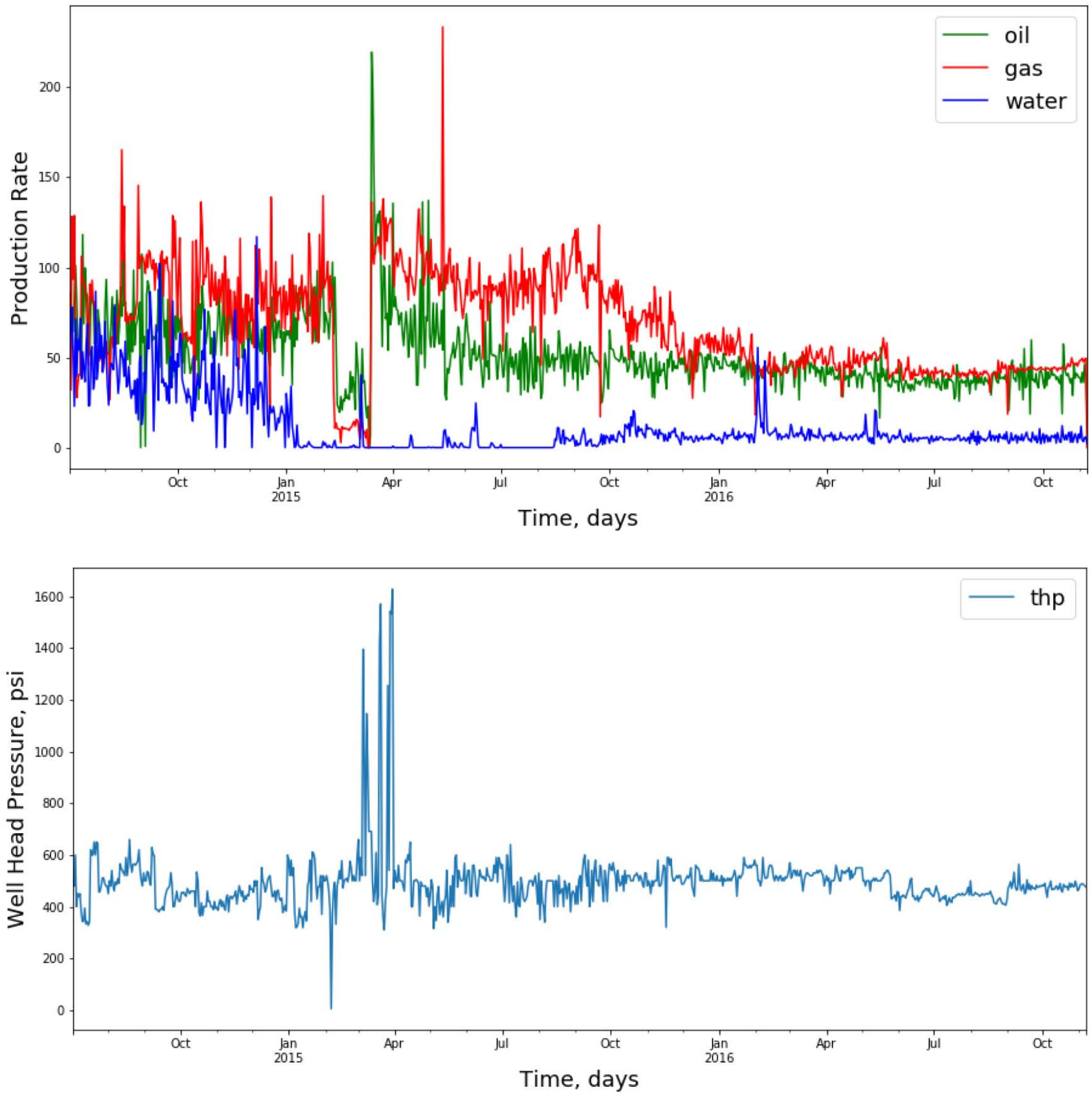

Figure 5—Actual production history and operation constraint for a single well.

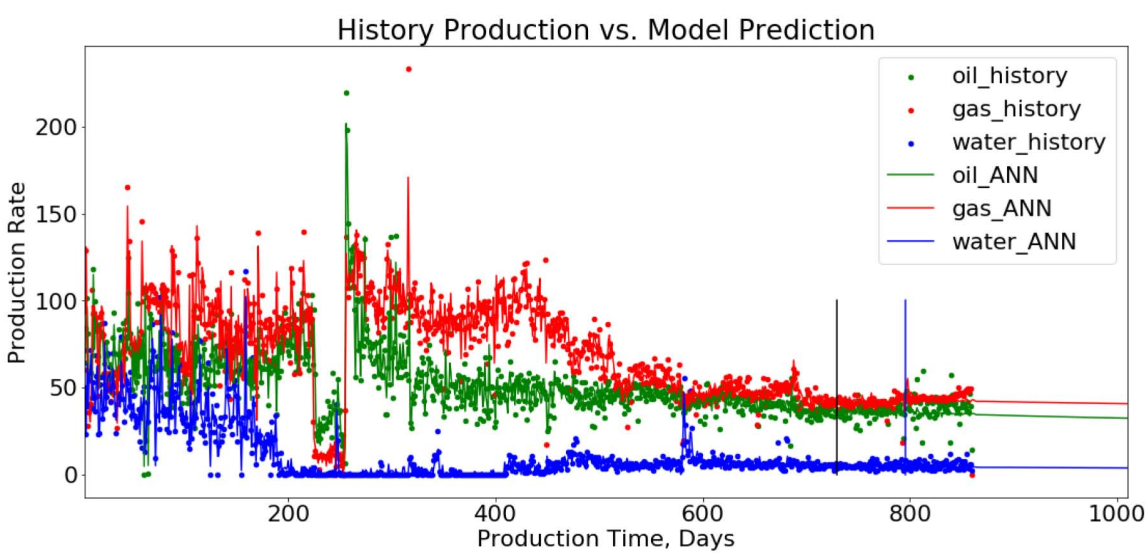

Figure 6—imulated production history from the ANN model for a single well.

Black line and blue line shown in Figure 6 are label lines, the data in the left side of this black line is training dataset and the data between these two lines is validation dataset. The data in the right side of the blue line is testing data, which gives forecast of multi-phase production rate. Here we assume a constant tubing head pressure which is the average pressure value in the last week as input for blind forecast section. And it could also give more than 5 months blind forecast and reasonable trend that is very useful for future production calculation.

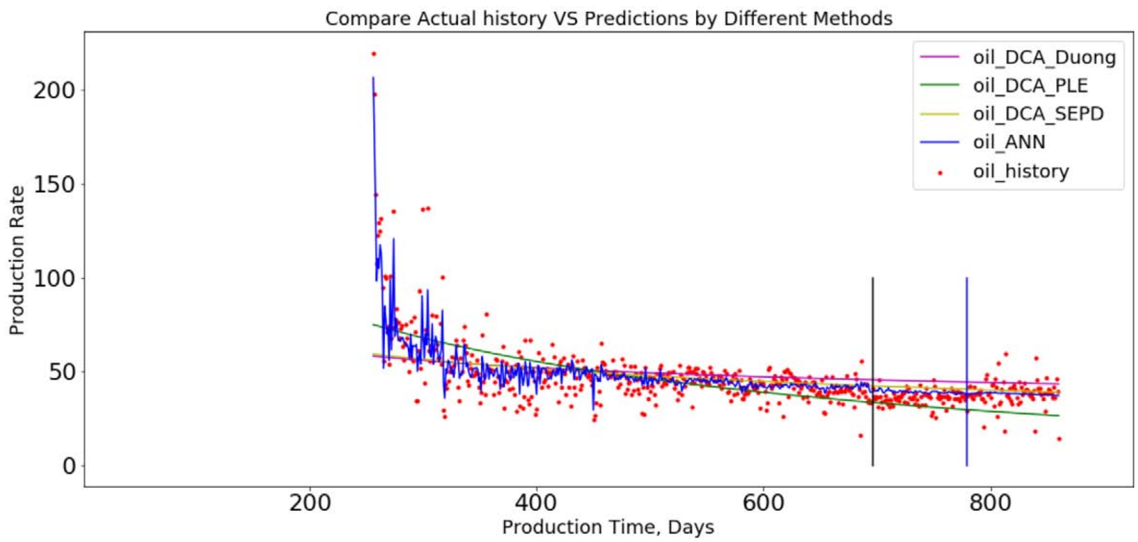

Figure 7—Compare ANN simulation results with traditional DCA methods.

We also compare the NN results with several traditional decline curve analysis (DCA) methods. ANN forecasting gives much better trend catching than DCA methods and much smaller errors compared with actual production data (Table 1).

Table 1—Compare forecast errors for certain blind prediction time

Blind Forecast Time	Actual Oil history	DCA_Duong Prediction /error%	DCA_PLE Prediction /error%	DCA_SEPD Prediction /error%	ANN Prediction /error%
at 10th day	38.98	44.51 / 14%	29.49 / 24%	40.60 / 4%	38.71 / 1%
at 20th day	38.36	44.38 / 16%	29.08 / 24%	40.40 / 5%	39.00 / 2%
at 50th day	38.20	44.03 / 15%	27.92 / 27%	39.84 / 4%	38.01 / 0.5%
Average of 30 days	38.57	44.45 / 15%	29.29 / 24%	40.50 / 5%	38.73 / 0.4%
Average of 60 days	38.39	44.07 / 15%	28.07 / 27%	39.9 / 4%	38.37 / 0.1%

Here is one more case tested for showing the capability of NN method (Figure 8), and we also compute and compare the mean square error (RMSE) for the blind forecast dataset which shown in Table 2. From these examples, it shows the better production prediction by NN method and more flexible capability for single well performance.

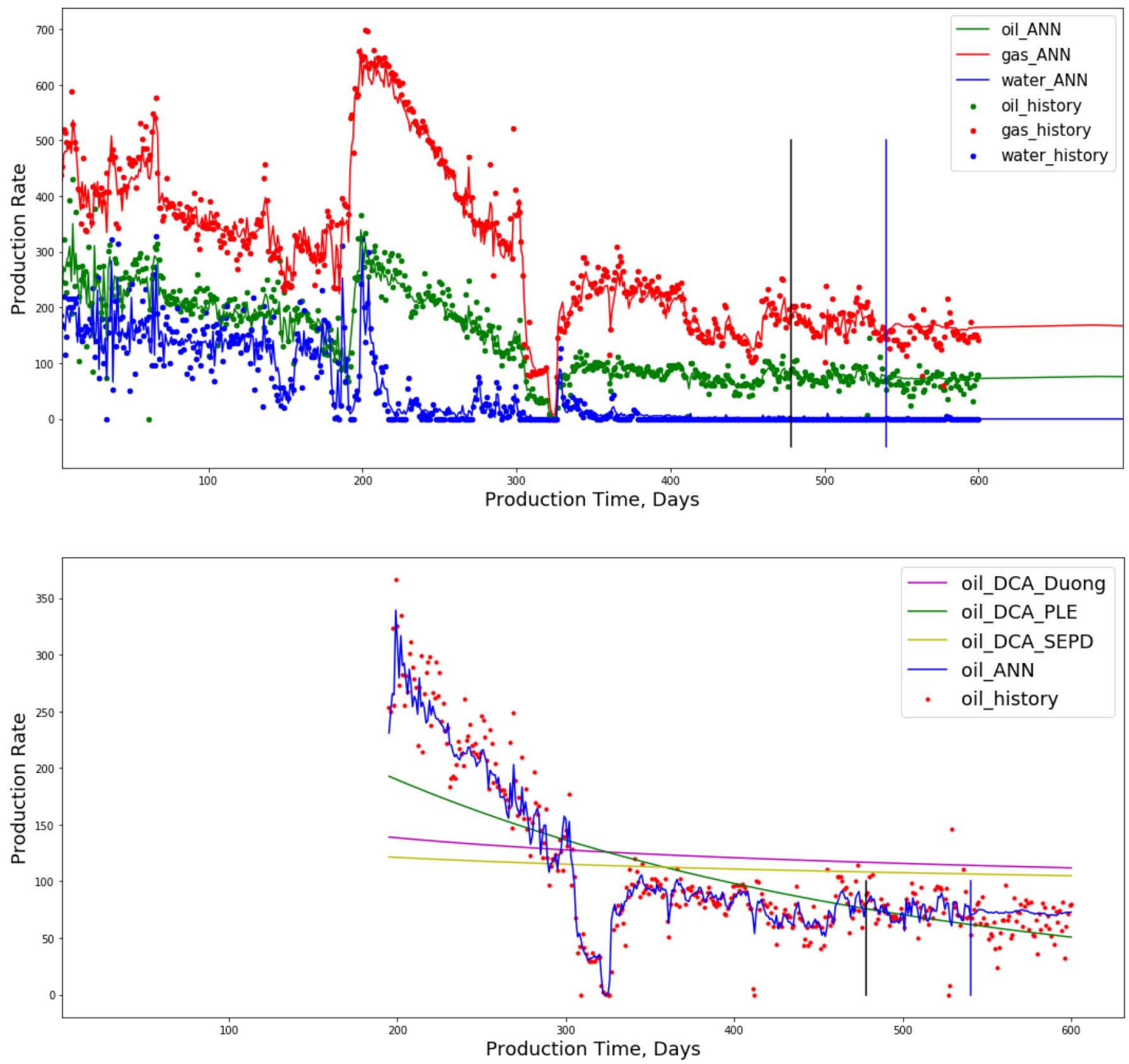

Figure 8—One more case to compare ANN simulation results with traditional DCA methods.

Table 2—Compare root mean square error (RMSE) for only the blind prediction data set

RMSE	DCA_Duong	DCA_PLE	DCA_SEPD	ANN
Case 1	124.23	128.14	91.41	81.65
Case 2	517.72	215.8	441.98	151.27

Aggregate Well Results

Assume we already have a few wells produced in the unconventional field for few years, we could use their production history to predict a new well which just start produced in a short time. Here we choose one existing well as a test target, and 4 nearby wells as reference wells. The production rates in the first one and half years and tubing head pressure are the inputs to train model, with the other half year production rates as validation part. NN model has one more layers compared with single well model to handle and transform more information. From Figure 10, it shows very close matching between the testing results and actual production rates of this target well. And it gives very better prediction than DCA curves.

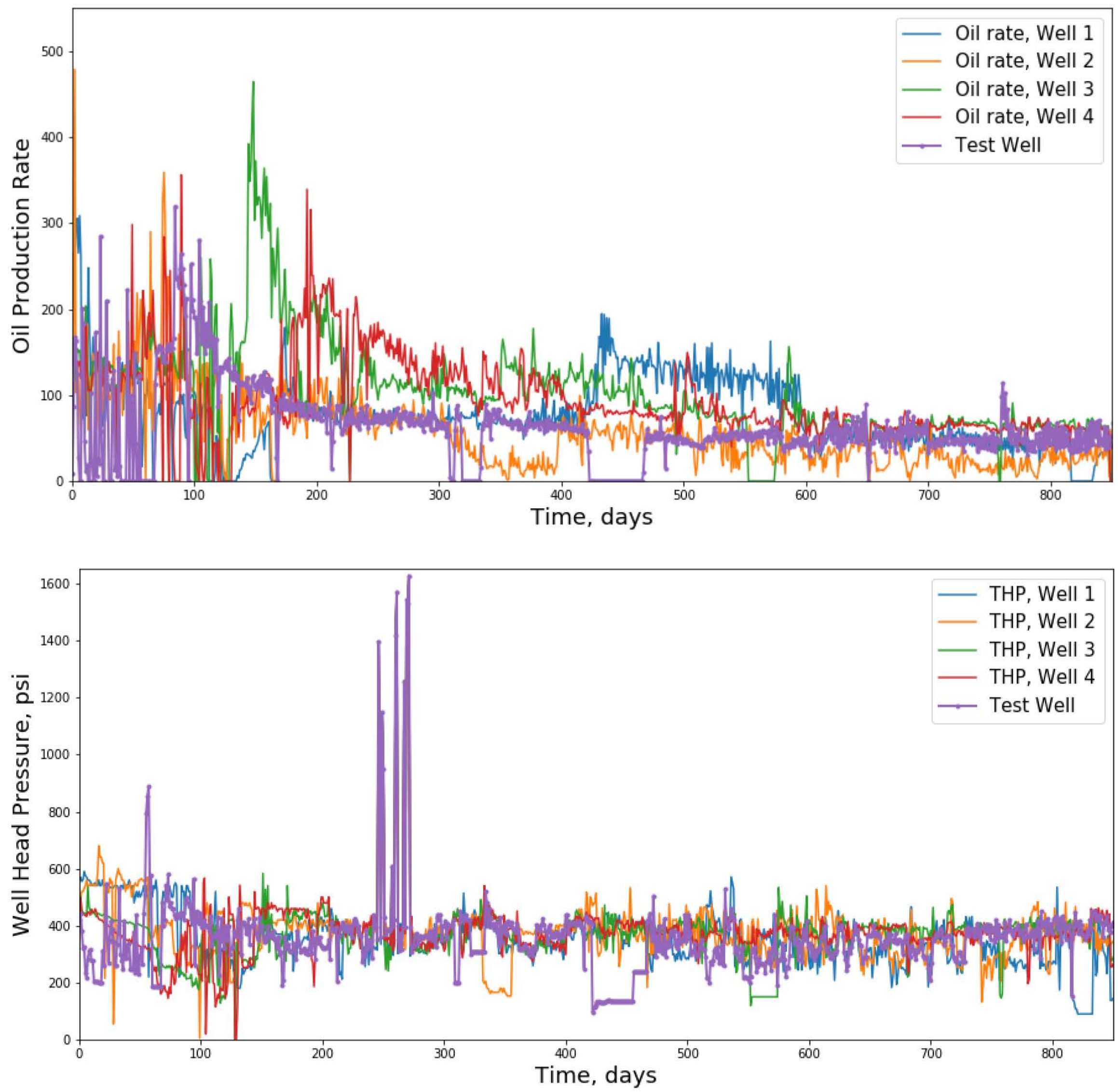

Figure 9—Actual production rate and operation constraints from the target well and other reference wells.

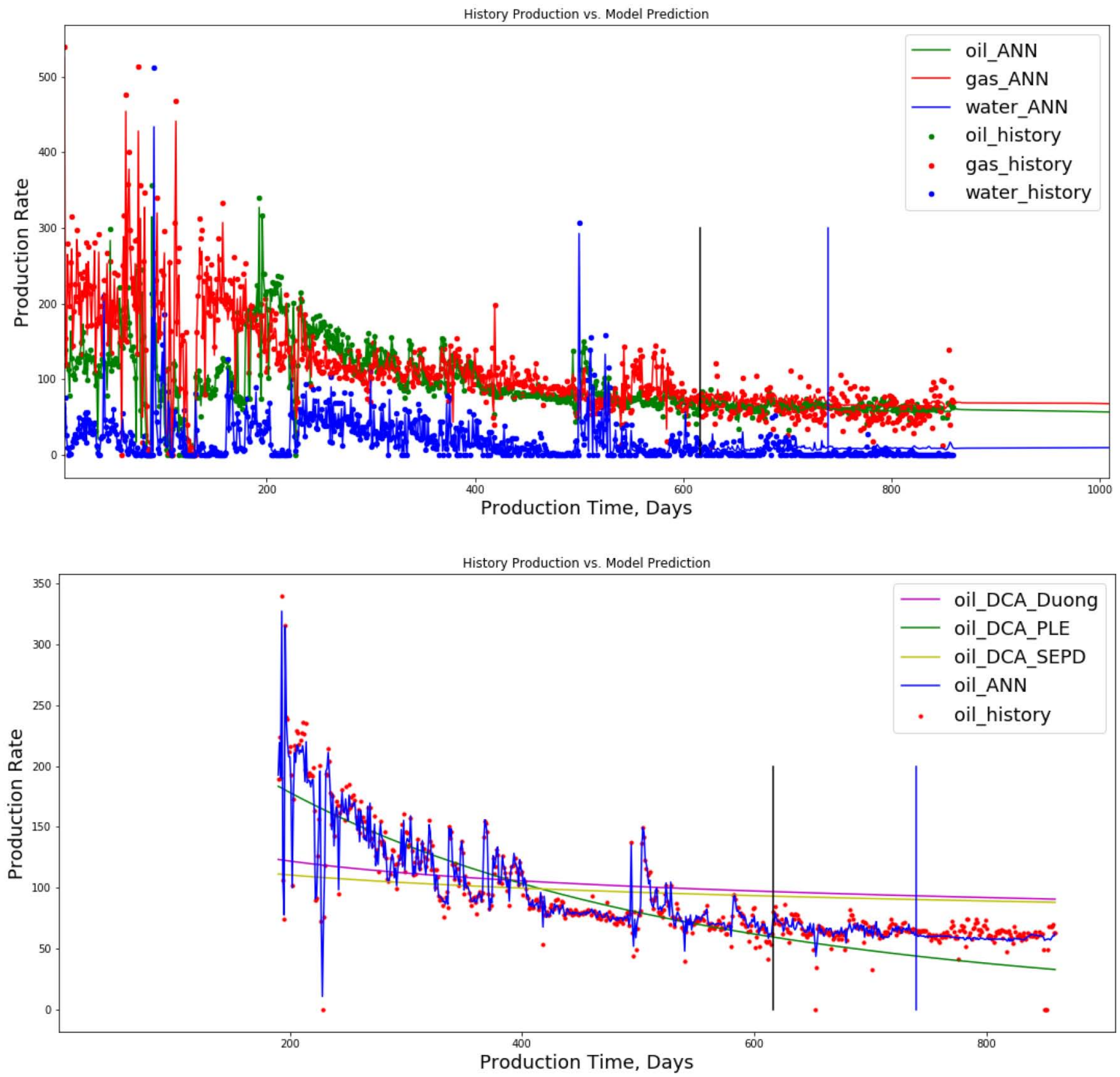

Figure 10—Production results from the ANN model for the target well by aggregate wells.

Conclusions

In this project, we implement a data driven approach to evaluate well production rates for single and aggregated well time series models based on available LSTM algorithms in open-source libraries. This approach is validated and compared with actual data history and traditional DCA methods, and shows much better trend forecasting and less error calculated. It also considers complex operation scenarios that could reflect more unconventional reservoir characterization which DCA methods cannot implement.

Compared with numerical simulation physics-based model, this NN approach doesn't heavily depend on engineering principle calculations, and gives solutions faster. Although this data-driven approach is physic-absence, it is a good alternative way to provide quick and robust results beside of empirical and numerical simulations.

References

EIA, 2017, Annual Energy Outlook 2017 with projections to 2050: https://www.eia.gov/outlooks/aeo/pdf/0383(2017).pdf

Arps, J.J., 1944. *Analysis of Decline Curves*. SPE-945228-G. *Trans. AIME* **160**: 228–247.

Das, S., Giles, C.L., Sun, G.Z., 1992, Learning Context Free Grammars: Limitations of a Recurrent Neural Network with an External Stack Memory, *Proc. 14th Annual Conf. of the Cog. Sci. Soc.*, p. 79.

Duong, A. N., 2011, *Rate-Decline Analysis for Fracture-Dominated Shale Reservoirs*. SPE 137748. doi:10.2118/137748-PA.

Fetkovich, M.J. 1980. Decline Curve Analysis Using Type Curves. *J Pet Technol* **32**(6): 1065–1077. SPE 4629-PA doi: 10.2118/4629-PA

Hochreiter, s., and Schmidhuber, J., 1997. Long short-term memory. *Neural computation*, **9**(8):1735–1780.

Holdawayk, K.R., Fabbi, L., Lozie, D., Unconventional Data-Driven Methodologies Forecast Performance In Unconventional Oil and Gas Reservoirs, *SAS Institute Inc (Paper 1910-2015)*. doi: 10.1.1.696.8225

Kanfar, M. S.A., 2013, *Comparison of Empirical Decline Curve Analysis for Shale Wells*, Master of Science Thesis, Texas A&M.

Mohaghegh, S.D., 2009, *Artificial Intelligence and Data Mining: Enabling Technology for Smart Fields*. SPE-0309-014-TWA. doi: 10.2118/0309-014-TWA

Mohaghegh, S.D., Grujic, O.S., Zargari, S., Dahaghi, A.K., 2011, *Modeling, History Matching, Forecasting and Analysis of Shale Reservoirs Performance Using Artificial Intelligence*, SPE 143875, doi: 10.2118/143875-MS.

Olah, C., 2015, *Understanding LSTM Networks*: http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Schmidhuber, J., 1993. *Habilitation thesis: System modeling and optimization*.

Schmidhuber, J., 2015. Deep Learning in Neural Networks: An Overview. *Neural Networks*. **61**: 85–117. doi: 10.1016/j.neunet.2014.09.003.

Valkó, P.P. 2009. *Assigning Value to Stimulation in the Barnett Shale: A Simultaneous Analysis of 7000 Plus Production Histories and Well Completion Records*. SPE 119639. doi: 10.2118/119639-MS.