SPE-180984-MS

Data Driven Production Forecasting Using Machine Learning

Q. Cao, R. Banerjee, S. Gupta, J. Li, W. Zhou, and B. Jeyachandra, Schlumberger

This paper was prepared for presentation at the SPE Argentina Exploration and Production of Unconventional Resources Symposium held in Buenos Aires, Argentina, 1–3 June 2016.

Abstract

Forecasting of production in unconventional prospects has gained a lot of attention in the recent years. The key challenges in unconventional reservoirs have been the requirement to put online a) a large number of wells in a short period of time, b) well productivity significantly driven by completion characteristics and that c) the physics of fluid flow in these prospects still remain uncertain. In this paper, machine learning algorithms are used to forecast production for existing and new wells in unconventional assets using inputs like geological maps, production history, pressure data and operational constraints. One of the most popular Machine Learning methods – Artificial Neural Network (ANN) is employed for this purpose. ANN can learn from large volume of data points without assuming a predetermined model and can adapt to newer data as and when it becomes available. The workflow involves using these data sets to train and optimize the ANN model which, subsequently, is used to predict the well production performance of both existing wells using their own history and new wells by using the history of nearby wells which were drilled in analogous geological locations. The proposed technique requires users to do less data conditioning and model building and focus more on analyzing what-if scenarios and determining the well performance.

Prologue

Imagine a field with 100 wells. These wells have produced for five years. A few well logs and a static geological model for this field is also available. Now if someone were to ask you two simple questions – ‘Where should I drill the next well?’ And ‘how much will that new well produce?’ What will your answer be? Unfortunately the answer to these two seemingly simple questions is not so straightforward. The geological model and the well logs have to be tied to each other. The static model has to be upscaled. The numerical simulation case has to be built. History matching has to be performed at well and field level and then the new well is placed into the model to determine how much hydrocarbon it will produce. The well position has to be sensitized to determine the best location. This is a complex process and can take a substantial amount of time and effort which often results in bottleneck. Therefore, there is an interest for alternative techniques in the industry and we explore one such technique in this paper.

Introduction

The global energy demand is expected to increase by 1.4% annually with the majority of the growth coming from developing countries like China, India and Indonesia (Stark, et al., 2008). While some believe that production from conventional reservoirs will decrease annually at 4.5% (Stark, et al., 2008) and others project the production to increase due to new discoveries and improving technology (Richard, 2010), the gap between the increasing demand and decreasing supply of energy is still expected to grow with time. It is expected that there may be times where supply exceeds demand, however, these would be of relatively shorter duration and unconventional resouces will play a major role in the longer term in filling this gap.

Given the nature of these resources it will be very challenging, if not impossible, to characterize these unconventional resources using conventional methods. In an unconventional play, the industry currently lacks resources and information to understand how the reservoir works. There are a few key issues which engineers have to deal with:

- The unconventional formation is heterogeneous
- The shale is ultra tight
- The physics which determine production such as rock properties, fluid flow mechanism is uncertain

With the uncertainties of governing physics and massive amount of data from hundreds of wells, as discussed in the prologue, engineers need a quick solution to effectively predict the production.

Naturally it is considered that reservoir characterization of unconventional resources should be studied by using more complex mathematical models. However inaccuracies in complex models arise due to the uncertainties associated with additional information required by the model (Nikravesh, et al., 2001).

On the other hand artificial intelligence methodologies like fuzzy logic, neural network, reinforcement learning etc. started gaining popularity in oil and gas industry since mid ninties. Some of them have been widely tested to study and mimic experimental data in laboratories and some extensive work have been carried out in the field like well test analysis (Dakshindas, et al., 1999), predicting density logs using vertical sesmic data (Artun, et al., 2005) and enhanced oil recovery (Surguchev, 2000). A top-down modeling approach (Shahab D. Mohaghegh, et al., 2011) leveraging artificial intelligence and data mining techniques has been used to analyze production of shale formation while there are still major challenges remaining in the modeling and simulation because of the complexities of shale formation (Shahab D. Mohaghegh, et al., 2014). This type of solution is referred to as data driven because it is ″learned″ or driven directly from data without assuming a predetermined equation as a model.

Given the unknowns in unconventional reservoirs, instead of model driven solution, data driven solution becomes more feasible and drew our attention. Considering its appetite of data driven and non-dependency on the governing physics, machine learning becomes a natural choice to solve this problem.

Machine Learning

Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed. It has been widely employed in a range of computing tasks when it’s infeasible to define the rules by hand.

One of the most popular Machine Learning methods – Artificial Neural Network (ANN) is employed for this purpose. In computer science, artificial neural networks are forms of computer architecture, inspired by biological neural networks (the central nervous systems of animals, in particular the brain) and are used to estimate or approximate functions that can depend on a large number of inputs and are generally unknown. Artificial neural networks are generally presented as systems of interconnected

″neurons″ which can compute values from inputs, and are capable of machine learning or pattern recognition thanks to their adaptive nature. See Figure 1 and 2 for basic biological neuron structure and representation of artificial neuron.

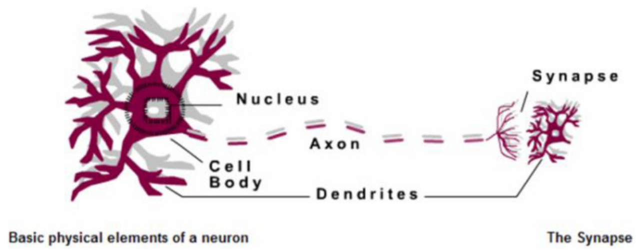

Figure 1—Basic physical elements of a biological neuron

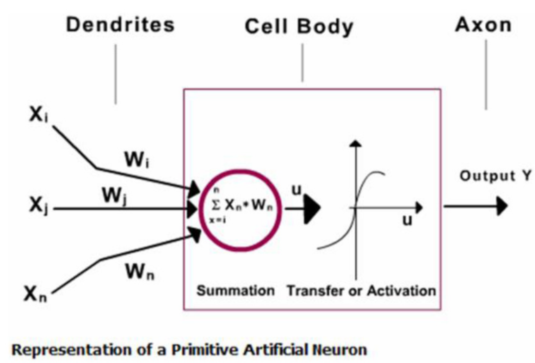

Figure 2—Representation of neuron in Artificial Neural Network

Data Sources/Inputs

The unconventional data source used in this paper is from Eagle Ford shale with 2-4 years of production history. The production data and the tubing head pressures are recorded daily. We treat the production history as a signal from the reservoir which carries information on how the well will respond to different operational constraint like tubing head pressure in the past. Understanding this behavior can help us predict its capability of producing with differing operational constraint in the near future.

The geological data will provide the information about the reservoir productivity from the well. Geology data from well logs plays a considerabily important role in this study. It includes thickness of the zones in the reservoir model, average porosity, average clay content and density.

Other well information, like the coordinates, can be used to evaluate the production interference effect of nearby wells. In this study, we use the normalized X and Y coordinates to the origin of target well location which will be discussed in forecasting a new well scenario.

Workflows

This methodology has been applied on two scenarios:

- **Production of an existing well**: Predict future production for a well based on its existing production history.
- **Production of a new well**: Predict production of a new well based on historical production from nearby wells in similar geological regimes.

The workflows for these two scenarios are as Figure 3 and 4, respectively.

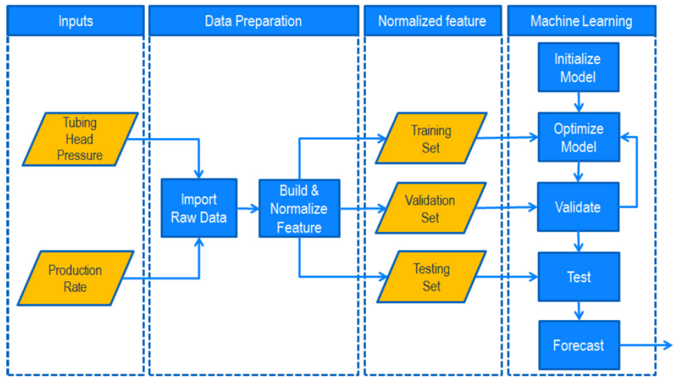

Figure 3—Workflow solution: forecasting for an existing well

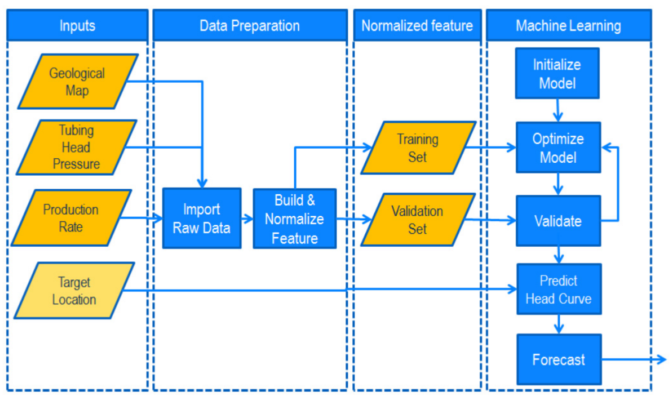

Figure 4—Workflow solution: forecasting for a new well

In the first scenario, since the geological data and coordinates are static in the scope of one single well. The NN model inputs are comparably simple. Data will be pre-processed and the normalized feature inputs will be divided into different groups for training, validation and testing. More details will be discussed in the NN model design session.

In the second scenario, the NN design is similar as previous one since it is using production history as inputs to do training and forecasting. However, the difference is this one will take various geological map data inputs and coordinates, representing different but nearby reservoir characteristics. The NN will learn how the normalized geological map data has impacted the production of nearby wells and apply that to the target location.

In this case the NN will need some history production of a well as a starting point, since the target well does not exist, it does not have any production history at any time. Therefore we need another NN model to firstly predict what the first phase of the production curve, here we call it head curve prediction. It will use the nearby well's geological data, tubing head pressure, history production and coordinates to predict the first 3 months production of the target well, as Figure 4. In parallel, the forecasting NN model has been trained using all the nearby wells production performance in the past. The head curve of target well will be passed as history production to get the whole forecasting

Neural network model design

Architecture In designing the architecture of an ANN model, the number of hidden layers and the number of neurons in each layer both play important rules regarding to the final performance. Unfortunately, there's no golden rule that can be applied to all the cases. Basically, as the network gets deeper and wider, the representation ability will increase accordingly. On the other hand, the training in complicated networks becomes more difficult than shallow and narrow networks, in which case the optimization is more likely to converge to some useless local optima. Ideally, we would like to design a model of reasonable complexity but powerful representation for the data we feed into it. Moreover, to avoid overfitting the model, the size of the training data has also to be considered in the designing. Therefore, taking all these concerns into account and after several trials on the validation dataset, we found one layer of hidden layer with the proper number of neurons fits our problem the best. An example of a one layer ANN architecture is as shown in Figure 5.

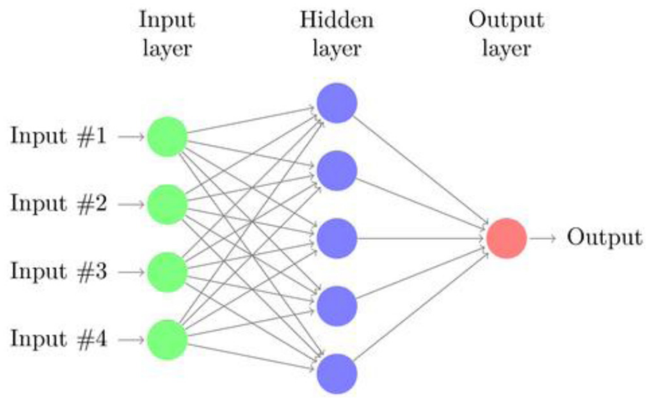

Figure 5—A simple Artifical Neural Network system structure

Inference

$$f(x, W, b) = W_{ho}\sigma(W_{ih}x + b_{ih}) + b_{ho} \tag{1}$$

where $\boldsymbol{x}$ is the input vector, $W = [W_{ih}, W_{ho}]$ and $b = [b_{ih}, b_{ho}]$ are respectively weight matrix and bias. σ is the activation function which is a sigmoid function, i.e. $\sigma(x) = 1/1+e^{-x}$. The subscript ″ih″ and ″ho″ represents inputs layer to hidden layer matrix and hidden layer to output layer matrix respectively.

Since the parameters of $\boldsymbol{W}$ and $\boldsymbol{b}$ are learned from the training, thus the inference is straightforward simply by feeding the pre-processed data into the neural network. Note that the data transformation is important in both inference and learning due to different scales of input features.

Learning In the training step, we will learn the ANN parameters $\boldsymbol{W}$ and $\boldsymbol{b}$ in Equation (1) from the training data. We use squared sum as the loss function and L2 norm as the regularization or so called weight decay. Thus, the objective loss function to be minimized in the training is as following:

$$L(W, b) = \sum_{i=1}^{N}(y_i - f(x_i, W, b))^2 + C||W||_2^2 \tag{2}$$

where $\boldsymbol{C}$ is the balancing coefficient between the loss term and the regularization term. This hyper-parameter is tuned on validation data. When $\boldsymbol{C}$ is small, the model tends to weight more on the correctness of prediction on training data while it might sacrifice its generalization on test data. $\boldsymbol{C}$ cannot be too large either, in which case the model could not accurately forecast the result of our interest.

To minimize Equation (2) with respect to $\boldsymbol{W}$ and $\boldsymbol{b}$, we use the efficient and effective method first order gradient descent. The gradient in the last layer is straightforward to compute while for those weights between input layer and hidden layer, the gradient is derived by utilizing the chain rule which is known as back-propagation.

To further increase the model generalization to test data, we deploy the early stop technique in the training. More precisely, before reaching the max predefined number of iterations, we keep checking the performance of the learned model on validation dataset in every $\boldsymbol{K}$ iterations. If the accuracy number is detected to drop in consective times on validation, then we stop the training even though the objective function of Equation (2) is still decreasing since the model is thought to be overfit the training data.

Hyperparameters In our experiments, we attempt different configurations of the hyper parameters of the model, trying to best fit the data in both scenarios. According to the evaluation on the validation dataset, we use 15 neurons in the hidden layer. The regularization coefficient $\boldsymbol{C}$ in Equation (2) is set 0.001. The initialization of the weights are randomly sampled from a uniform distribution $Unif(-1,1)$. The learning rate in the gradient descent is determined by line search strategy.

Validations

Case 1: Production of an existing well

Figure 6 is an existing well's actual production for 2 years. We take the first 80% of data (inputs and expected outputs pairs) to train the ANN model, get the optimized results for this same training inputs which is shown in orange in the Figure 7. As we can see in this figure, the optimized simulation result is very close to the actual data.

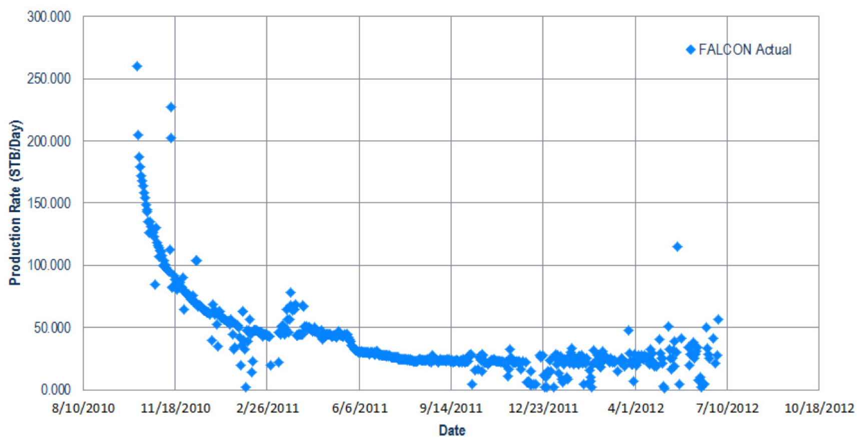

Figure 6—Actual production of an existing well

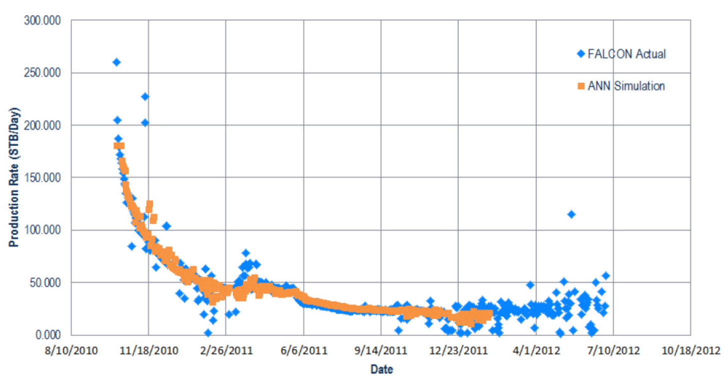

Figure 7—Compare actual production of an existing well and optimized simulation results

Now we have the ANN model. For validation purpose, if we take the remaining 20% data which is previously disregarded for the training process, pass the inputs to the ANN model and get our test data result in green in Figure 8. The overlay between the test result and actual result for these 20% samples proves how good the trained ANN model can perform.

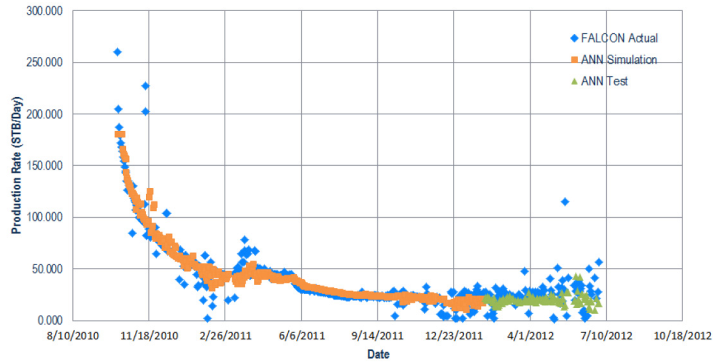

Figure 8—Compare actual production, ANN simulation and test results

More results for different wells are shown in Figure 9.

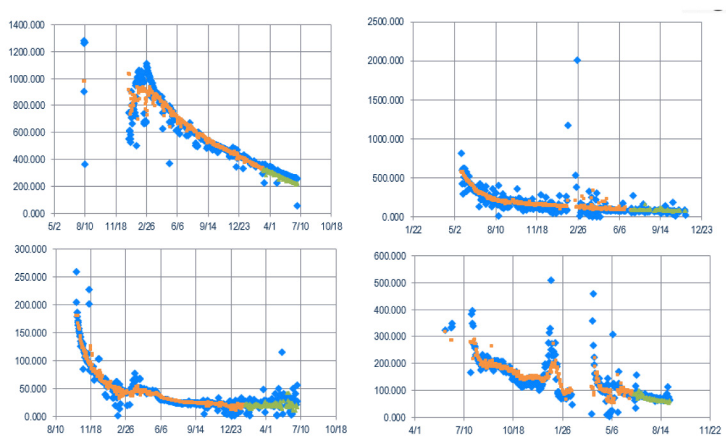

Figure 9—More results for production of an existing well

Case 2: Forecasting for a new well

Provided we already have a few wells, drilled and producing in the unconventional field, we can use their historical production data, to predict the production of a new well that is yet to be drilled.

After the training with existing wells data, in order to be able to validate our simulation result with actual data, we pick one existing well's location which is previously unknown to this ANN model; take its geological data, tubing head pressure and historical production rate as inputs to the test. Shown in Figure 10 is the test forecasting result we get for this new location. Comparison with the actual production history of this well is seen in Figure 11. As we can see the trend of testing results and actual production match closely. Figure 12 shows more results for different wells.

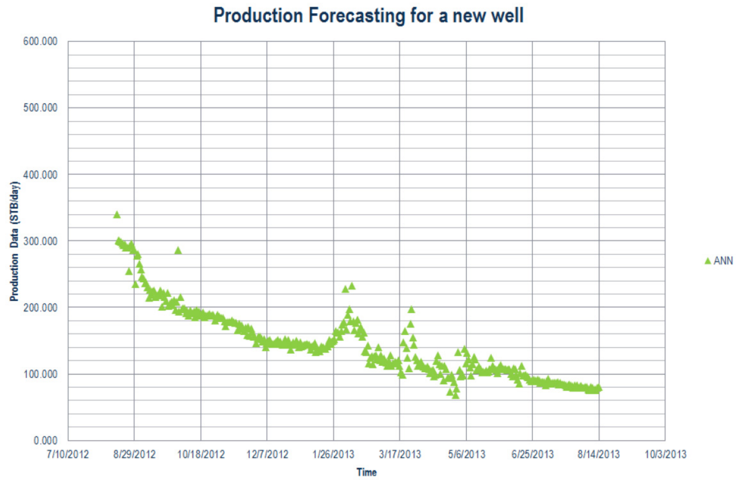

Figure 10—Test results for a new well using neighboring wells production

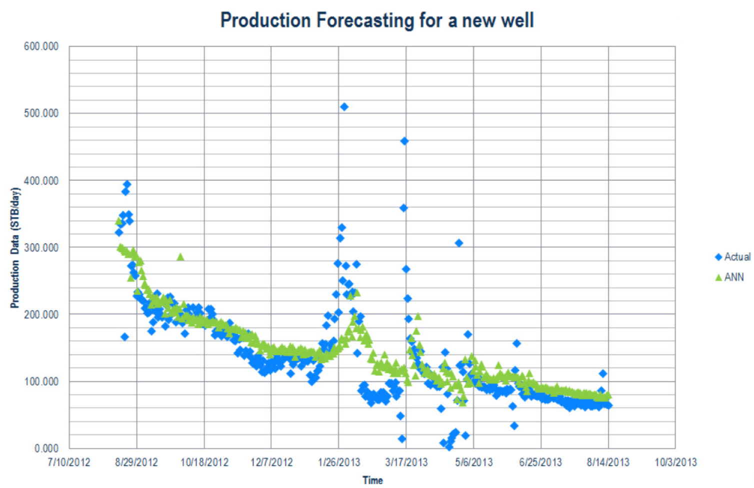

Figure 11—Compare actual production and test results for a new well

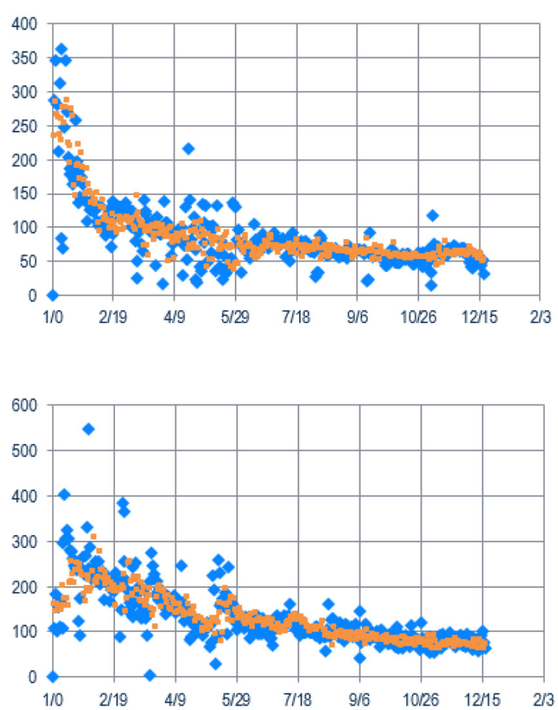

Figure 12—More results for production of a new well

Comparison of Forecasting

For the first case, in Figure 13, forecasting results from ANN has been compared with other decline curve techniques, namely Arps, Duong, SEPD and Power law methods. The last 20% of the data points were used for validation section in which the actual tubing head pressure is taken as inputs for the ANN model. The validation results shows that the ANN model can predict close results to actual data and reflect the variance of tubing head pressure change while decline curve results do not.

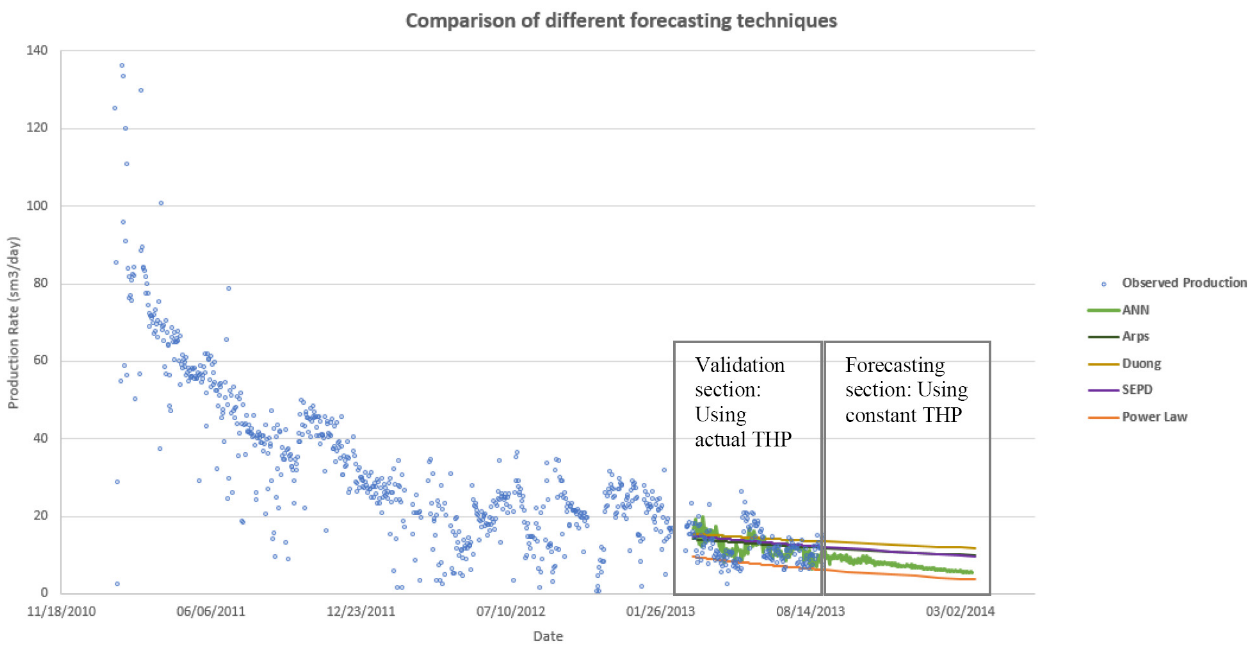

Figure 13—Compare NN model and DCA results for production of a new well

When it comes to the forecasting section, constant tubing head pressure calculated from the average of the last 7 days is used as input. It can be seen that ANN forecasting lies between the optimistic forecast from Duong and pessimistic forecast of Power law. It can also be seen that the trend matches the data better than any of the decline curves. Similar behavior was seen on many other wells that were analysed.

Conclusion

In this methodology, we are proposing a data driven approach to evaluate the well production in two scenarios, forecasting for existing well and forecasting for a new well yet to be drilled. This approach is more comprehensive compared to decline curve analysis in that it, includes more data inputs like geological map data, production constraint like tubing head pressure and locations that reflect complex reservoir characterization of unconventional wells without giving a predetermined model.

On the other hand, using the numerical simulation to do production forecasting using equally rigorous inputs can take a long time to process. With massive data, engineers can hit bottleneck in optimizing the models. The data driven approach provides a middle road to the engineers looking for a robust and fast technique to evaluate the wells performance in seconds.

ANN forecasting is by no means a replacement for production forecasting using empirical or numerical simulation. Rather ANN forecasting can be used to provide more confidence to these forecasting techniques by also tying in the data driven approach.

Reference

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

Arps, J.J.: 1956. Estimation of Primary Reserves. *Trans. AIME* **207**: 24–33.

Artun, E., Mohaghegh, S., Toro, J., Wilson, T., & Sanchez, A. (2005). Reservoir characterization using intelligent seismicinversion. Eastern Regional Meeting. Morgantown, W.V.: SPE. 126

Dakshindas, S. S., Ertekin, T., & Grader, A. S. (1999). Virtual well testing. Eastern Regional Conference and Exhibition. Charleston, West Virginia: SPE.

Duong, Anh N.: ″An Unconventional Rate Decline Approach for Tight and Fracture-Dominated Gas Wells,″ paper SPE 137748 presented at the Canadian Unconventional Resources & International Petroleum Conference held in Calgary, Alberta, Canada, 19–21 October 2010

Ilk, D., Rushing, J.A., Perego, A.D. et al. 2008. Exponential Vs. Hyperbolic Decline in Tight Gas Sands: Understanding the Origin and Implications for Reserve Estimates Using Arps’ Decline Curves. Presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA, 21–24 September. SPE-116731-MS. http://dx.doi.org/10.2118/116731-MS.

J. Nocedal, S. J. Wright, Numerical Optimization (2nd edition) p30–p37, 2006

L. Prechelt, Early Stopping — But When? Neural Networks: Tricks of the trade, pp 55–69, *Lecture Notes in Computer Science 1524,* Springer Verlag, Heidelberg, 1998

Nikravesh, M., & Aminzadeh, F. (2001). Past, present, and future intelligent reservoir characterization trends. *Journal of Petroleum Science and Engineering,* **31** (2-4), 67–79.

Richard, N. (2010). Annual Energy Outlook 2011Reference Case. Washington, DC. Rosenblatt, F. (1962). *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Michigan: Spartan Books. Russum, D. A. (2010).

Evaluating unconventional reservoirs: The questions we need to answer. Canadian Unconventional Resources & International Conference. Calgary, Alberta: CSUG/SPE.137

Shahab D. Mohaghegh SPE, Intelligent Solutions and West Virginia University, and Ognjen Grujic, SPE, Seed Zargari, SPE, and Masoud Kalantari, West Virginia University: Modeling, History Matching, Forecasting, and Analysis of Shale-Reservoir Performance Using Artificial Intelligence, SPE Digital Energy Conference and Exhibition, the Woodlands, Texas, 19-21 April. SPE-0911-0060-JPT.

Shahab D. Mohaghegh, SPE, Intelligent Solutions and West Virginia University: ″A Critical View of the Current State of Reservoir Modeling of Shale Assets″, 2013 SPE Eastern Regional Meeting, Pittsburgh, Pennsylvania, USA, 20-22 August. SPE-0714-0098-JPT

Stark, P., Chew, K., & Jackson, P. (December 2008). Importance of unconventional oil resources in shaping the Far East Energy future. International Petroleum Conference. Kuala Lumpur, Malaysia.

Surguchev, L. (2000). IOR evaluation and applicability screening using artificial neural networks. SPE/DOE Improved Oil Recovery Symposium. Tulsa, Oklahoma: SPE.139

Valko, P.P. and Lee, W.J. 2010. A Better Way to Forecast Production from Unconventional Gas Wells. Presented at the SPE Annual Technical Conference and Exhibition, Florence, Italy. 19–22 September. SPE-134231-MS. http://dx.doi.org/10.2118/134231-MS.