Machine Learning Aided Production Data Analysis using Representative Type Wells

A Thesis Proposal

Submitted to

The Faculty of the Department of Petroleum Engineering

Texas A&M University

In Partial Fulfillment

of the Requirement for the Degree

MASTER OF SCIENCE

By

Liuyi Jin

August 2017

**Contents**

[1 Abstract 1](#_Toc496121427)

[2 Problem and its status 1](#_Toc496121428)

[2.1 Type Wells 2](#_Toc496121429)

[2.2 Machine Learning Application in Petroleum Industry 3](#_Toc496121430)

[3 Objective 5](#_Toc496121431)

[3.1 Source of production data for type well construction 5](#_Toc496121432)

[3.3 Application of type wells with involvement of neural networks 6](#_Toc496121433)

[4 Questions to be addressed 6](#_Toc496121434)

[4.1 Well Selection 6](#_Toc496121435)

[4.2 Data Selection 7](#_Toc496121436)

[4.3 Type Well Construction 7](#_Toc496121437)

[4.4 Machine Learning 7](#_Toc496121438)

[4.5 Result Assessment 7](#_Toc496121439)

[References 8](#_Toc496121440)

# 1 Abstract

Type wells, which are typical well production profiles based on analysis of existing well histories, are growing in acceptance in the industry as a means of forecasting production in low-permeability reservoirs. Although the conceptual understanding is becoming clearer, there are still many challenges when we try, as a means of forecasting production in low-permeability reservoirs, to apply type wells in practice, chief among them being how to build type wells that accurately reflect original well production data. In this project, we will evaluate several available methods of building representative type wells. At the same time, we will adopt an advanced machine learning principle, neural networks, as a significant method to classify type wells. By using production data from the Barnett Shale, this project will be a trial to apply both type well construction strategies and machine learning principles, to traditional oil and gas production data analysis. We will further apply probabilistic theory to classify the undrilled wells.

# 2 Problem and its status

Unconventional resources have become of dominant in the oil and gas world in recent years. As a carry-over from the era of conventional resources, decline curve analysis (DCA), which is dominated by the Arps decline model (Arps 1945), is the main method used to predict production of these resources.

Unconventional resources, including shale gas, shale oil, and coalbed methane, are gaining increasing attention from many researchers in the petroleum industry. As defined by Holditch (Holditch 2003), an unconventional reservoir is one “that cannot be produced at economic flow rates or that does not produce economic volumes of oil and gas without assistance from mass stimulation treatments or special recovery processes and technologies.” This definition manifests the difficulties of unconventional resources extraction. Unlike extraction of conventional resources, the lengthy transient flow production period in unconventional resources limits the applicability of the Arps decline model, which was designed and validated in the conventional resources era.

Figure 1 below used viscosity and permeability to distinguish conventional and unconventional resources. Besides the characteristics of the unconventional resources themselves, unconventional reservoirs themselves are relatively more heterogeneous. There are more geologic uncertainties as well (e.g., different geologic layering of reservoirs).

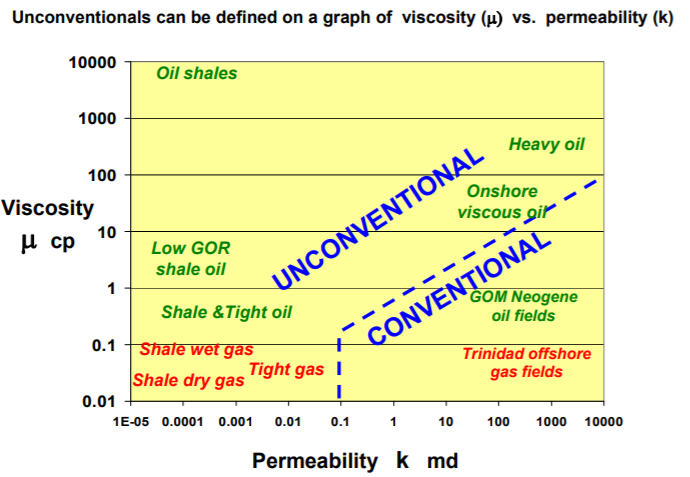


Figure 1 Conventional and Unconventional Resources (BP 2012)

These conditions have caused unexpected errors when applying conventional reservoir production forecasting methods to unconventional reservoirs. Based on Arps’ model (Arps 1945), some researchers have proposed alternative decline models to generate more accurate production data predictions. These modified models include Duong’s method, and stretched exponential and power law models, which can work well under certain constrained conditions for unconventional resources. In addition, the Arps model has been modified in an effort to provide more accurate forecasts. Unfortunately, all these approaches can result in ultimate error due to the many uncertainties in unconventional resources.

To tackle this issue, industry has begun to employ type wells, which should be distinguished from “type curves” in which a set of dimensionless flow rate versus dimensionless time curves are generated to predict future production.

## 2.1 Type Wells

The core idea of constructing type wells is to construct a well representing a set of wells that are being recognized as “analogous.” It aims to extract the inherent characteristics of multiple wells in a certain geologic area by producing one or a family of constructed wells which should be representative enough to represent all wells in an area of interest. The approach commonly employed in industry to construct type wells is to arithmetically average the production histories of a set of producing wells in the field. As pointed out by Freeborn (Freeborn et. al. 2012), there is a serious flaw in this approach. He maintains that “the type well calculation must include the SI well count as though those SI wells continue to produce at a rate of zero”. Figure 2 (Freeborn et. al. 2012), provides a graphical explanation of how to avoid this flaw.

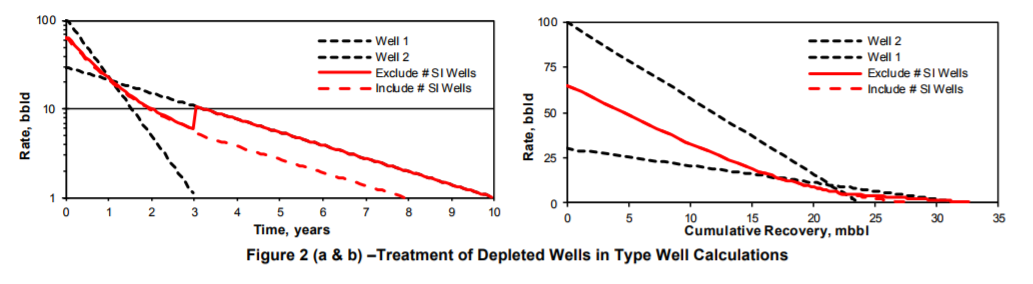


Figure Treatment of Depleted Wells in Type Well Construction (Freeborn et. al. 2012)

In his paper, Freeborn also addressed other issues that may be encountered when constructing type wells. Those issues include the date of first production in which he proposes to refer to the shifted data as having been normalized, multi well testing and type well using history data. To test the theory, Freeborn (Freeborn et. al. 2012) gave 4 different field examples. Rastogi (Rastogi and Lee 2015) extended the methodology of Freeborn in constructing type wells, and they also expressed their considerations on the method of normalizing production data and removing outlier data. Rastogi suggested applying a type well production profile generated in one area to another area in the same resource with different geologic characteristics, to determine the universality of type well method.

In terms of how to construct type well accurately and efficiently, Freeborn (Freeborn and Russell 2015) used aggregation methods, which he proposed as a way to construct type wells that “accurately simulate the average rate-time production profiles expected from a drilling program of specific size and certainty.” They also improved the normalization method by building several scaling rules that scale a rate-time profile. Apart from that, Chaudhary (Chaudhary and Lee 2017) chose to generate type wells by cluster-weighted modeling, in which one type well is constructed as one “global weighted probabilistic superposition of local models” where local models consist of several wells combined using simple cubic spline techniques. As researchers are increasingly focusing on type well methods to evaluate the future production of unconventional plays, engineers have reached a consensus that the current challenge of type well construction is concentrated in how to accurately create type wells that are representative of wells in a certain geologic area.

## 2.2 Machine Learning Application in Petroleum Industry

In March of 2016, AlphaGo (Wikipedia 2017) from DeepMind won the Go championship over Lee Sedol. From then on, artificial intelligence has made unprecedented progress. The AI industry, particularly the machine learning technique, has attracted more and more people to this area. Today, progress in the AI industry and machine learning is based mainly on deep learning, which is supported by the convolutional neural network work inspired by Fukushima’s neocognition proposed in 1980. Yann Lecun then proposed convolutional neural networks. Convolutional neural networks are now widely accepted and popularly employed in image processing. Figure 3 and figure 4 graphically illustrate neocognition and convolutional neural networks, respectively.

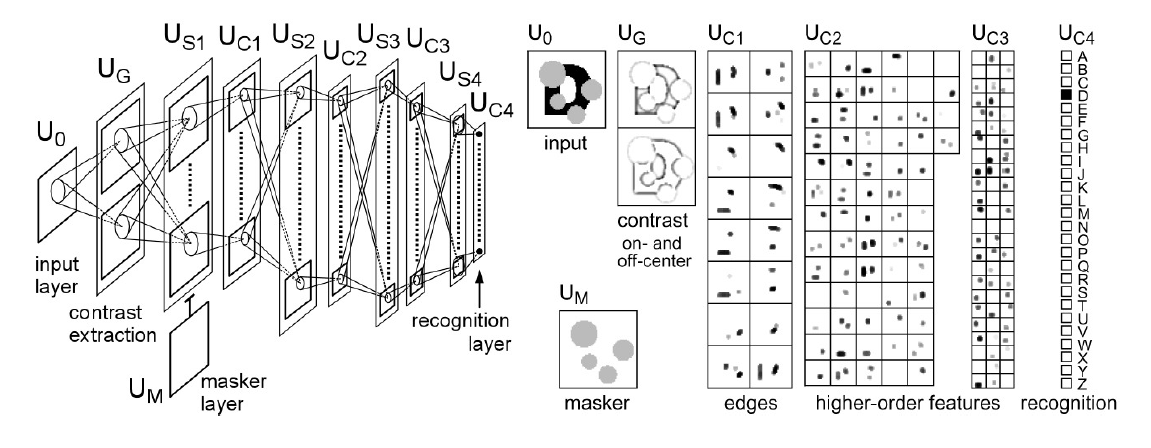


Figure Fukushima's Neocognitron

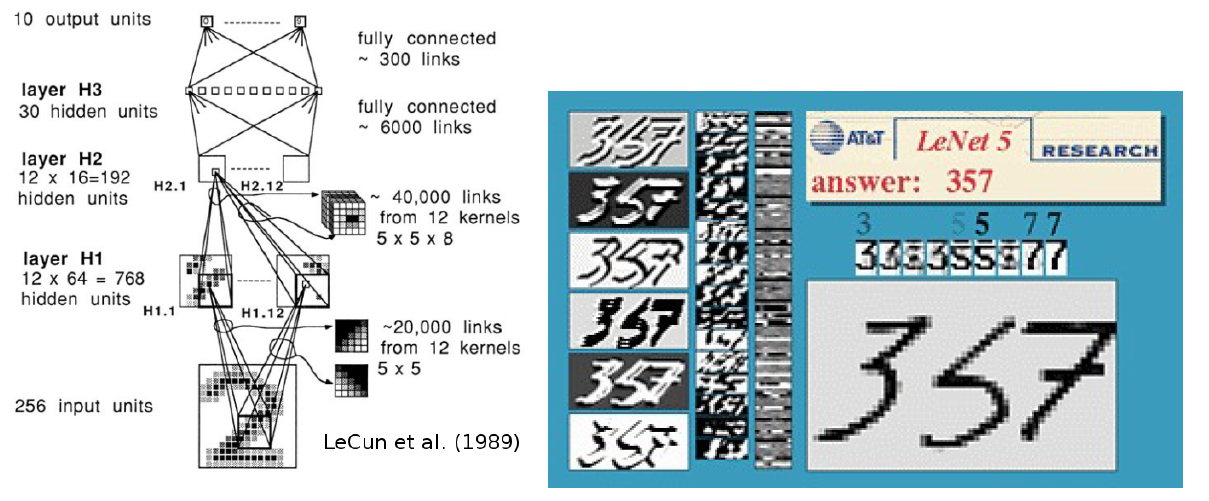


Figure LeCun's Convolutional Neural Nets

Recent researchers have applied AI, including machine learning techniques, to the oil and gas industry. Ali (Ali 1994) reviewed the application of neural networks, a major machine learning technique, in the petroleum industry at that time. He pointed out that the capabilities of neural networks include “pattern recognition, classification of noisy data, nonlinear feature detection, market forecasting and process modeling,” which makes them well suited to solve problems in petroleum industry. In particular, we know that neural networks are quite useful to deal with nonlinear relationships between various variables.

Mohaghegh (Mohaghegh 2000) reviewed general applications of virtual artificial intelligence and artificial neural networks in the oil and gas industry. Ramgulam (Ramgulam et al. 2007) developed a trained artificial neural network to specifically address the history matching problem. Liu (Liu and Horne 2011) adopted data-mining techniques to analyze permanent downhole gauges. Liu (Liu and Horne 2011) exploited the least-mean-squares method and used the stochastic gradient descent method to train the parameters in a given polynomial equation to evaluate the relationship between pressure and flow rate. In his papers (Liu and Horne 2013b), Liu continued to approach the problems in interpreting pressure and flow rate data from permanent downhole equipment using data mining and machine learning methods, respectively. Tian (Tian and Horne 2015) extended Liu’s work further. He applied the “kernel ridge regression based machine learning” to further interpret data from multi-well tests. In addition, he completed flow rate reconstruction.

The objectives of my research are to exploit the advantages of machine learning, and in particular, neural networks, to predict production rate profiles and use them in type well construction. Previous studies related to this area include the work of Subrahmanya (Subrahmanya et. al. 2014) who gave us a review of machine learning algorithms applied to production data pattern recognition. In this paper, he employed semi-supervised learning and active learning algorithm to analyze acquired data. More recently, Jia (Jia and Zhang 2016) used neural networks to forecast production in the Barnett shale, and he achieved much more accurate predictions than with conventional empirical models. Figure 5 illustrates the simple neural network structure that Jia employed (Jia and Zhang 2016).

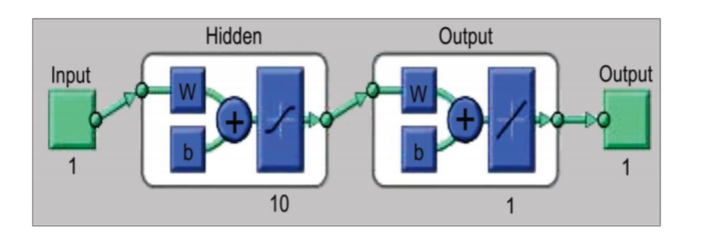


Figure Simple Neural Network Structure (Xinli Jia et. al. 2016)

# 3 Objective

The goal of this project is to build representative type well using data from the Permian Basin. By incorporating neural networks, our method to construct type wells will be innovative and potentially powerful. We will go through the following steps in this project.

## 3.1 Source of production data for type well construction

The Permian Basin is currently an active unconventional resource and much data available is recent. More specifically, data are available from 1958 to the present. The data used to construct type wells will be within this time range. To ensure generality, we will apply arithmetical averaging first to construct “plain vanilla” type wells with data specifically from Andrews county. I selected this county because its 10,000 wells outnumber those in any other county in the basin. In this process, we will avoid flaws Freeborn pointed out (Freeborn et. al. 2012).

3.2 Type well construction and evaluation

Currently, type wells are growing in importance, especially since 2008 when unconventional resources began to play an important role, but the method of constructing type wells has varied. We will consider several proposed methods to construct type wells. Each has its pros and cons, which we will identify. We will select the most suitable method for this project after sufficiently detailed study.

## 3.3 Application of type wells with involvement of neural networks

When the type wells have been constructed, we will regard them as standards. The neural network algorithm will then be applied to analyze production data from new wells. The advantage of neural networks over other machine learning algorithms is that neural networks can deal with any nonlinear relationships among various variables. In addition, the dependence between each pair of variables, the number of parameters and hyper-parameters do not generally need to be considered.

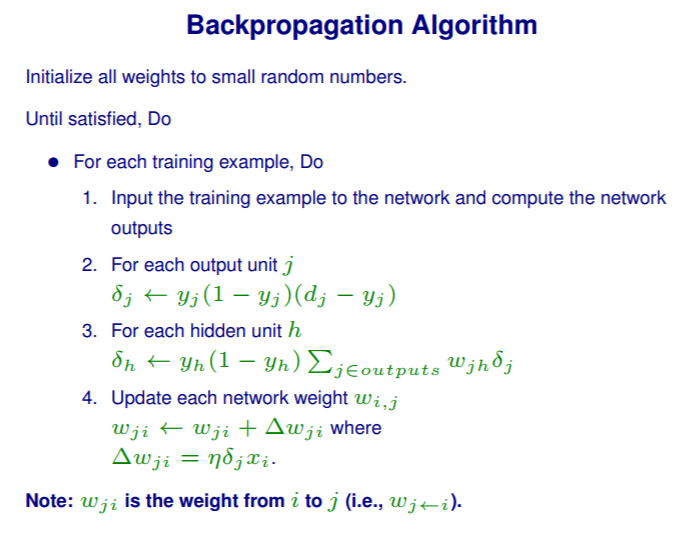


Figure Backpropagation Algorithm

The backpropagation algorithm is the core of the neural network algorithm. Figure 6 illustrates the back propagation algorithm schematically. It is based on a stochastic gradient descent algorithm. After wells have been classified through the training process, it will then be appropriate to classify our test wells. We will then examine wells with relatively short production histories, and identify the most analogous type well for these wells. When the type well has been identified, we can then easily predict the future production of a new well using the identified type well as a basis. In the training process, neural network performance can be improved by using various techniques such as cross validation.

# 4 Questions to be addressed

## 4.1 Well Selection

The wells within certain geographical areas will be selected to construct type wells. Which areas should we choose?

## 4.2 Data Selection

Which portion of data from wells selected to construct type wells should we use? All the data or only data within a specified time range? How should we bin wells? Can we scale for differing lateral lengths, vintage or other factors?

## 4.3 Type Well Construction

How should we use the data selected to build representative type wells? What method will we use? Why is that method superior to other methods?

In addition to the type well construction, how can we generate the complete probability distribution?

## 4.4 Machine Learning

After type well construction, what kind of machine learning techniques should we choose to use to identify the appropriate type well for a new well we are studying? Why is this method better than any alternative machine learning algorithm?

## 4.5 Result Assessment

How should we use the type well we generated to predict future production? How valid are the predictions? How can we determine whether the probabilistic distribution we generated is correct?

We will answer these questions in our work on the project.

# References

Ali, J.K. 1994. Neural Networks: A New Tool for the Petroleum Industry? . Presented at the European Petroleum Computer Conference, Aberdeen, United Kingdom, 15—17 March. SPE-27561-MS. <http://dx.doi.org/10.2118/27561-MS>.

Arps, J.J. 1945. Analysis of Decline Curves. SPE Journal 160 (1). SPE-45228. <http://dx.doi.org/10.2118/45228>.

BP 2012. AAPG , <http://www.searchanddiscovery.com/documents/2012/80217cander/ndx_cander>.

Chaudhary, N.L. and Lee, W.J. 2017. Constructing Enhanced Type Wells Using Cluster-Weighted Modelling. Presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Austin, Texas, 24—26 July. SPE-URTEC-2697554-MS. <http://dx.doi.org/10.2118/URTEC-2697554-MS>.

Freeborn, R. and Russell, B. 2015. Creating More Representative Type Wells. Presented at the SPE/CSUR Unconventional Resources Conference, Calgary, Canada, 20—22 October. SPE-175967-MS. http://dx.doi.org/10.2118/175967-MS.

Freeborn, R., Russell, B., and Keinick, W. 2012. Creating Analogs, Fact and Fiction. Presented at the SPE hydrocarbon, Economics and Evaluation Symposium, Calgary, Canada, 24—25 September. SPE-162630. <http://dx.doi.org/10.2118/162630>.

Holditch, S.A. 2003. The Increasing Role of Unconventional Reservoirs in the Future of the Oil and Gas Business. Journal of Petroleum Technology 55 (11): 34—79.

Jia, X. and Zhang, F. 2016. Applying Data-Driven Method to Production Decline Analysis and Forecasting. Presented at the SPE Annual Technical Conference and Exhibition, Dubai, United Arab Emirates, 26—28 September. SPE-181616-MS. http://dx.doi.org/10.2118/181616-MS.

Liu, Y. and Horne, R.N. 2011. Interpreting Pressure and Flow Rate Data from Permanent Downhole Gauges Using Data Mining Approaches. Presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, 30 October—2 November. SPE-147298-MS. <http://dx.doi.org/10.2118/147298-MS>.

Liu, Y. and Horne, R.N. 2013. Interpreting Pressure and Flow Rate Data from Permanent Downhole Gauges with Convolution-Kernel-Based Data Mining. Presented at the SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, 30 September—2 October. SPE-166440-MS. <http://dx.doi.org/10.2118/166440-MS>.

Mohaghegh, S. 2000. Virtual-Intelligence Applications in Petroleum Engineering: Part 1—Artificial Neural Networks. Journal of Petroleum Technology 52 (9). SPE-58046-JPT. <http://dx.doi.org/10.2118/58046-JPT>.

Ramgulam, A., Ertekin, T., and Flemings, P.B. 2007. An Artificial Neural Network Utility for the Optimization of History Matching Process. Presented at the Latin American&Caribbean Petroleum Engineering Conference, Buenos Aires, Argentina, 15—18 April. SPE-107468-MS. <http://dx.doi.org/10.2118/107468-MS>.

Rastogi, A. and Lee, W.J. 2015. Methodology for Construction of Type Wells for Production Forecasting in Unconventional Reservoirs. Presented at the Unconventional Resources Technology Conference, San Antonio, Texas, 20—22 July. SPE-URTEC-2152273-MS. http://dx.doi.org/10.2118/URTEC-2152273-MS.

Subrahmanya, N., Xu, P., El-Bakry, A. et al. 2014. Advanced Machine Learning Methods for Production Data Pattern Recognition. Presented at the SPE Intelligent Energy Conference&Exhibition, Utrecht, Netherlands, 1—3 April. SPE-167839-MS. <http://dx.doi.org/10.2118/167839-MS>.

Tian, C. and Horne, R.N. 2015. Machine Learning Applied to Multiwell Test Analysis and Flow Rate Reconstruction. Presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 28—30 September. SPE-175059-MS. <http://dx.doi.org/10.2118/175059-MS>.

Wikipedia 2017. <https://en.wikipedia.org/wiki/AlphaGo_versus_Lee_Sedol>.