Using Machine Learning Techniques for Enhancing Production Forecast in North Malay Basin

Truc (Trevor) Doan, Bear and Brook Consulting, Brisbane, Australia; and **Minh Van Vo***, Chevron Unocal East China Sea Ltd., Chengdu, China

Abstract

The geological environment in the Gulf of Thailand (GOT) is very complicated, with thousands of small discontinuous reservoirs, in part due to a high density of faults. With this geological characteristic, hundreds of required slim-hole wells will be planned and drilled annually to ensure enough capacity to meet gas demand in gas sale agreement. Hence, production forecast plays key role to deal with drilling schedule and operations planning and installing surface facilities. An accuracy degree of production forecast is required highly based on production data from existing produced wells and future produced wells.

In recent years, machine learning techniques have been widely applied to the oil and gas industry, in this case, the Gulf of Thailand. Machine learning methods of Support Vector Regression (SVR) and K-Means Clustering have been applied effectively for removing outliers or noisy data. By using a huge dataset of production data from thousands of wells in this area, a solution can be rapidly made to automatically eliminate unreliable data in given historical data from existing produced wells.

In brief, this study provides an automated approach to apply machine learning algorithms to assist technical teams in improving the quality of data in production data analysis, with the aim of enhancing reliable production forecast, optimizing drilling schedule and saving operating costs.

Introduction

A several thousands of wells have been drilled with slim-hole well technologies and the commingle production strategy has been used in North Malay Basin since the early 1980s. For a typical well in this area, with an average of 10 pay sands, the reservoir properties of each individual sand tend to vary by depth and by location. **Figure 1** is used as an excellent illustration for a drilling program with a complex subsurface picture, which includes many small and discontinuous gas reservoirs in the high density of faults (Pinto et al. 2004).

Production forecast plays a critical role to provide production throughput information that can be used for facilities capacity design, drilling sequence/schedule, and then economic evaluation during development of a Field Development Plan FDP. Therefore, a reliable degree of production forecast is highly required, which is based on historical production data from both the existing producing wells and future producing wells. **Figure 2** explains how importance of the production forecast in scheduling drilling projects (Doan and Vo 2019a).

DOI:10.14800/IOGR.1190
Received April 15, 2021; revised July 12, 2021; accepted August 12, 2021.
*Corresponding author: minh.vo@chevron.com

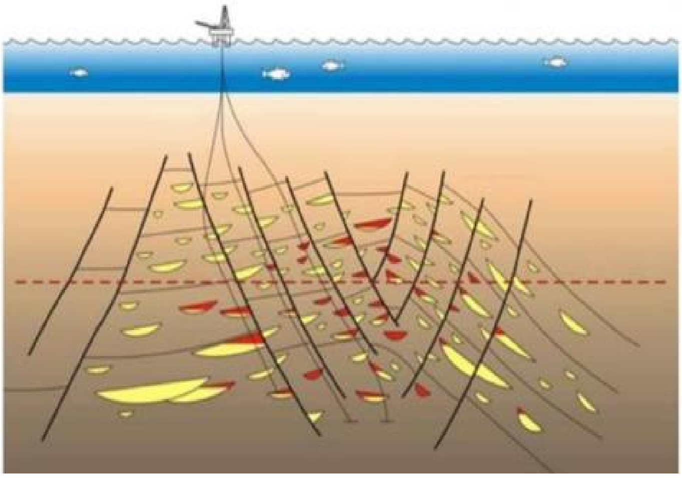

Figure 1—Reservoir structure and complexity in the North Malay Basin.

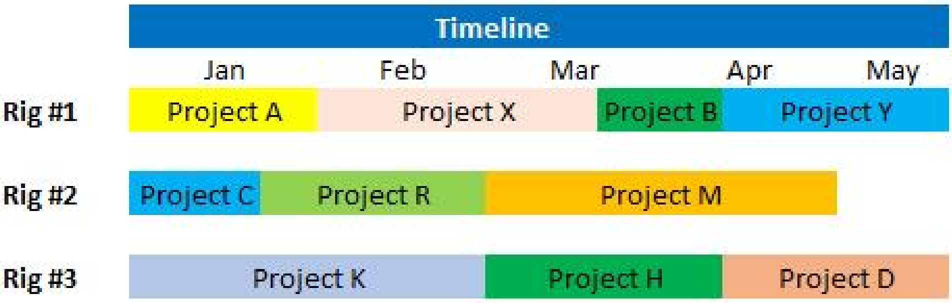

Figure 2—An illustration of drilling project timeline and number of rig allocation schedule.

In recent years, machine learning techniques have been widely applied in the oil and gas industry, and in this case, the North Malay Basin in the Gulf of Thailand GOT. Machine learning methods of Support Vector Regression (SVR) and K-Means Clustering have effectively been applied for removing outliers or noisy data. By applying the machine learning techniques, the production forecast can be enhanced with the high quality of forecasting in short term and long term.

With geological characteristics in North Malay Basin, hundreds of wells are annually required to drill to ensure enough capability to meet the market gas demand (**Figure 3**). Time was known to be the top priority because any delay or failure to deliver gas to gas buyers would make a financial loss to the company and its partners. Hence, a robust solution to improve production forecast is very necessary to develop oil and gas in this geological environment.

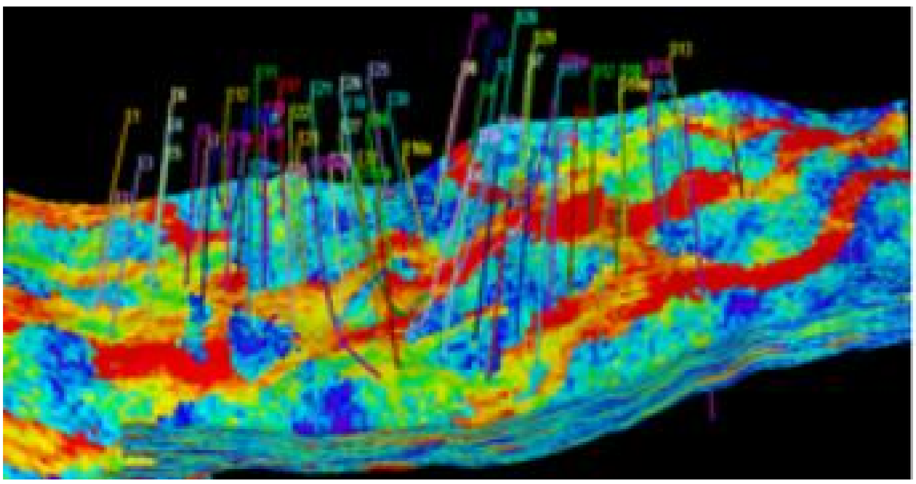

Figure 3—Example with multiple wells drilled in the North Malay Basin.

Methodology

The methodology is developed which mainly integrates the machine learning-based techniques with the approach of decline curve analysis approach from a huge dataset of the existing produced wells. In production operation, measured data are frequently contaminated with anomalies (noise or outliers) that can be generated internally from the measurement device or come in from external sources. In particular, the solution will be built to automatically eliminate unreliable data points in given historical production data by using machine learning methods and then to estimate remaining hydrocarbon reserves and predict production performance by using the decline curve analysis.

Decline Curve Analysis

The approach of Decline Curve Analysis (DCA) is used to generate production profiles and to allocate gas rate from every single well/platform to meet the gas sales contract requirements. This approach is very powerful in estimating ultimate gas recoveries and in predicting field development performance from the analysis of long-term gas production data, either from individual wells or from entire fields. It is proper when large uncertainty limits the data to justify a complex reservoir simulation. **Figure 4** below describes the typical production profile for a well, when the initial production is controlled at the maximum plateau rate for a period and then the production will naturally decline.

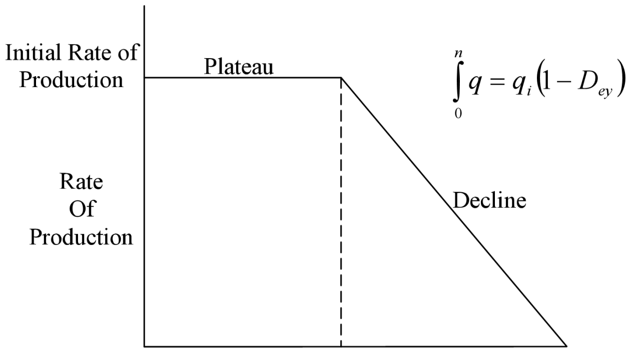

Figure 4—A rate model by using decline curve analysis.

Machine Learning Methods

K-Means Clustering is an unsupervised machine learning very popular which can be applied to detect any anomalies in dataset. It is a centroid-based algorithm, or a distance-based algorithm, where the distances is calculated to assign a point to a cluster (Doan and Vo 2020). In this study, it is used as a first step of data preprocessing to remove any outliers in a given dataset before analysis, as shown in **Figure 5**.

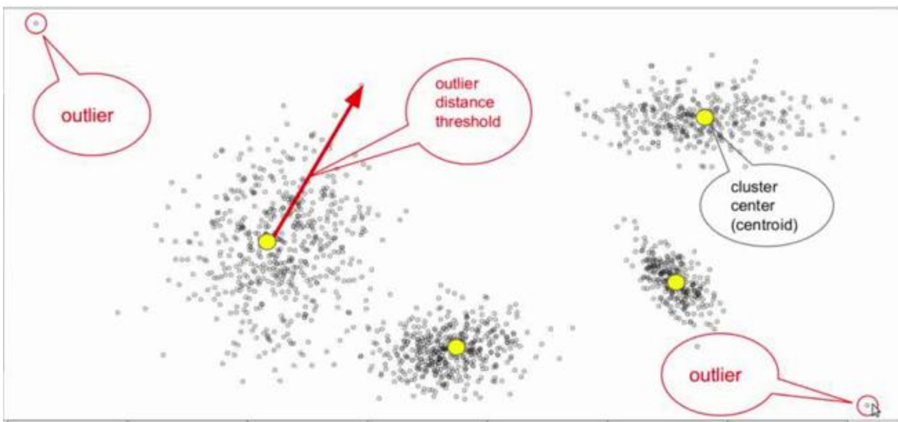

Figure 5—A K-means clustering method to detecte anomalies or outliers.

Support Vector Regression (SVR) uses the same principle as Support Vector Machine. However, SVR allows to be changed easily in defining how much error is good enough in selecting our model and will find an appropriate line (hyperplane in higher dimension) to fit the data, as shown in **Figure 6**. It has been applied effectively to remove noise data in the analysis. In this study, the method is used to make a more reliable rate prediction for the existing produced wells.

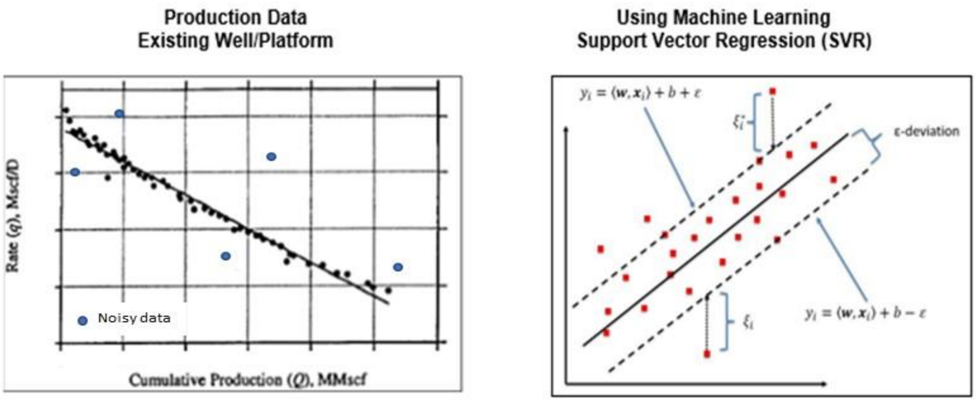

Figure 6—An example of SVR method for removing noisy data (Doan and Vo 2019b)

Based on the methodology, a workflow for this approach is divided into five steps as follows:

Step 1: Data Pre-processing

This is the first step to identify and reduce any outliers of historical data, including reservoir pressure, water cut, wellhead pressure, choke size and rate, which are collected from a huge number of producing wells in GOT by using the K-means clustering method before moving to Step 2.

Step 2: Dataset

This step is to categorize and arrange the data in Step 1 into a standard dataset before proceeding in Step 3.

Step 3: Prediction Model

This step is to apply the SVR method to remove any noise data before forecasting production rate at the well level at acceptable quality using decline curve analysis.

Step 4: Production Forecast

The main task of this is to integrate production profile of the existing producing wells into the production forecast model (Doan and Vo 2019a) for generating many different development scenarios.

Step 5: Results

At the final step, many alternatives of perforated sands sequencing in the new wells are automatically made for economic evaluation.

The diagram in **Figure 7** below illustrates the workflow on how to improve production forecast with machine learning methods.

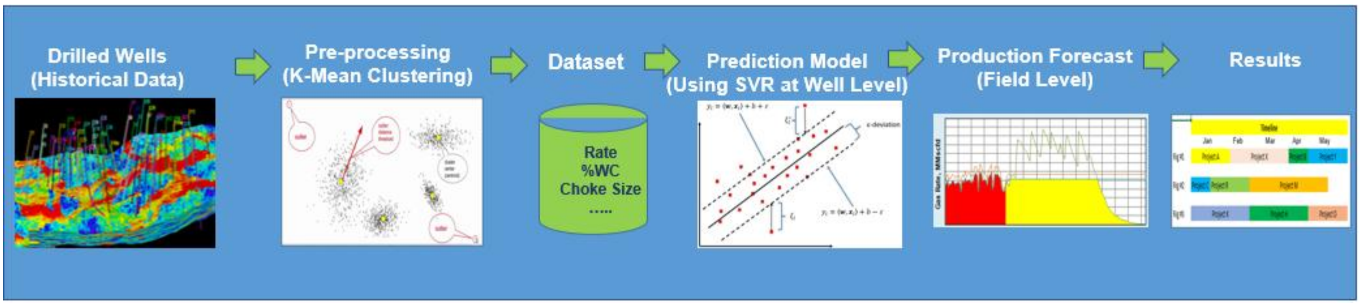

Figure 7—Workflow in the Application of Machine Learning to Improve Production Forecast.

Application and Results

Based on the methodology and workflow, Doan and Vo (2019a) have developed an application to improve production forecast by integrating Python scripts (using for machine learning techniques) into Excel VBA scripts. **Figure 8** describes a user-friendly interface of the application.

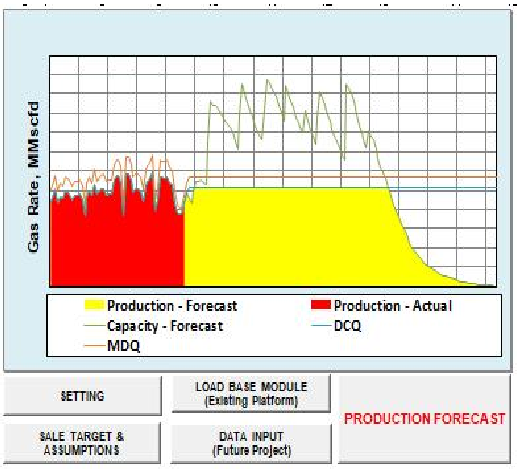

Figure 8—An interface of production forecast modelling.

Key features of the model are:

- User-friendly interface;
- Structure of a simple and small model file;
- Production forecast support for both modes (manual and automatic); and
- Rapid execution to give the forecasting results and economic evaluations.

Production forecast model will automatically allocate gas flow rate of each platform which contributes to meet the market demand. It also provides the installation time and the number of platforms to be installed or wells to be drilled, which could be necessary to satisfy the demand as shown in **Figure 9**.

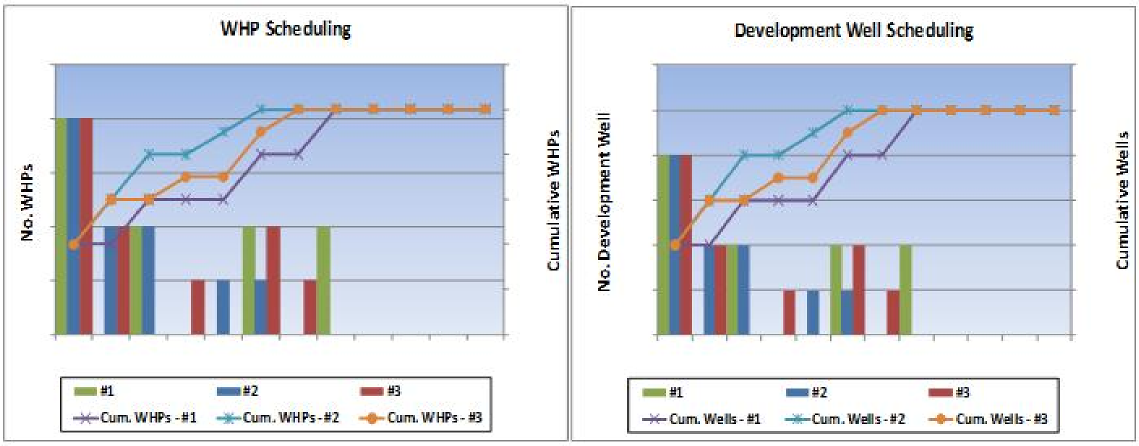

Figure 9—Drilling project sequence and number of required WHP/development wells.

By applying machine learning approaches, the results have provided an effective sequence and drilling schedule for future drilling projects and aligned with the current production capacity of existing producing wells. The drilling project timeline and number of rigs allocation schedule have refined to be more realistic in compared with the previous scenarios (**Figure 10**).

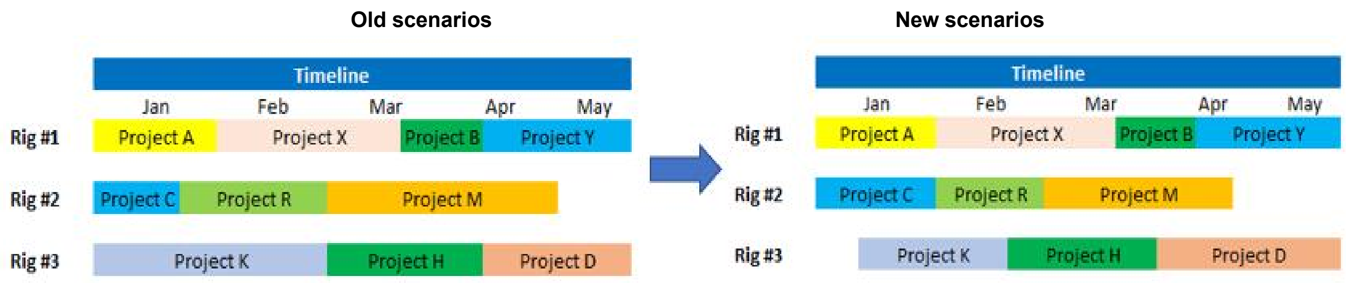

Figure 10—Drilling project timeline and number of rig allocation schedule.

Furthermore, the outputs provided not only the drilling schedule for future drilling projects, but also risk mitigation in project planning and management to optimize the logistic work (mobilize/demobilize, long lead items, and labor costs).

Conclusions

This study provided an integrated solution using machine learning methods to improve the quality of production forecast and drilling schedule. The results have effectively supported engineers to make decisions for both short-term and long-term asset development and to make significant savings in operating costs. It is strongly believed that the model can be applied to solve many similar engineering challenges in the oil and gas field, especially for oil and gas fields that have many wells and a complicated dataset in the North Malay Basin.

Nomenclature

DCA = Decline Curve Analysis
SVR = Support Vector Regression
GOT = Gulf of Thailand
WHP = Well Head Platform
ML = Machine Learning
VBA = Visual Basic Application

Conflicts of Interest

The author(s) declare that they have no conflicting interests.

References

Pinto, C.J., Pendleton, L.E., Dick, J.L., et al. 2004. Ultrafast Drilling in the Gulf of Thailand: Putting Science into the Design Process. Paper presented at the IADC/SPE Drilling Conference, Dallas, Texas, 2-4 March. SPE-87173-MS.

Doan, T.T. and Vo, M.V. 2019a. A Rapid Modelling Approach to Optimize Drilling and Production for A Complex Field Development in North Malay Basin. Paper presented at International Petroleum and Petrochemical Technology Conference, Xi'an, China, 3-5 July.

Doan, T.T. and Vo, M.V. 2019b. Using Machine Learning Techniques to Evaluate Performance for Existing Waterflood Projects in the Gulf of Thailand. Paper presented at SPE Workshop: Water Injection Excellence, Kuala Lumpur Malaysia, 3-5 March.

Doan, T.T. and Vo, M.V. 2020. Application of Machine Learning for Initial Completion Plan of Multiple Zone Gas Wells in North Malay Basin. Proceedings of the International Petroleum and Petrochemical Technology Conference & Exhibition, Beijing, China, 11-13 September.

Truc (Trevor) Doan, SPE, is a senior Petroleum Engineer at Bear and Brook Consulting. Before that, he had worked as Petroleum Engineering lecturer for 2+ years at Ho Chi Minh City University of Technology HCMUT and as Petroleum Engineer for Chevron for 10+ years at various locations. He holds bachelor's in petroleum engineering from HCMUT, master’s degree in petroleum engineering from AIT, and currently, has been studying for PhD Degree at School of Engineering, Edith Cowan University in Australia.

Minh Vo, SPE, is currently subsurface manager in Unocal East China Sea Ltd., where he has worked for the last 6+ years. He has had 25+ years of experience in the oil and gas industry with multiple global locations. His research interests are in reservoir engineering, production optimization, and systems engineering. He holds several master’s degrees from UNSW in petroleum engineering, from RMIT in systems engineering, and MBA from NYU.