PETROVIETNAM SCIENCE, TECHNOLOGY & INNOVATION PETROVIETNAM

APPLICATION OF MACHINE LEARNING TO PREDICT THE TIME EVOLUTION OF CONDENSATE TO GAS RATIO FOR PLANNING AND MANAGEMENT OF GAS-CONDENSATE FIELDS

Ngo Huu Hai[1], Trinh Xuan Vinh[1], Nguyen Ngoc Tan[1], Hoang Ky Son[1], Ngo Tuan Anh[1]
Tran Ngoc Trung[1], Tran Vu Tung[1], Nguyen Sy Tuan[2]
[1]Bien Dong Petroleum Operating Company (Bien Dong POC)
[2]Institute for Computational Science and Artificial Intelligence, Van Lang University
Email: sonhk@biendongpoc.vn
https://doi.org/10.47800/PVSI.2023.02-02

Summary

One of the most important parameters for the evaluation, forecast, and management of gas-condensate fields is the evolution of the condensate to gas ratio (CGR) over time. This parameter tends to decrease as reservoir pressure declines. In the conventional approach, gas and condensate samples are collected at the beginning of production and periodically later to conduct laboratory experiments on composition, CGR, and fluid properties. However, sample collection, transportation, and analysis require a lot of time and effort and could be very expensive. Likewise, dynamic models are also frequently used to predict CGR over time. However, these models could include many uncertainties due to ambiguous input data, including reservoir structures, fluid phase interaction, and reservoir property distribution. Therefore, the application of machine learning to predict the time evolution of CGR in this research could be a new and effective approach to supplement conventional methods.

Key words: Machine learning, condensate to gas ratio, production forecast, Hai Thach field.

1. Introduction

Predicting condensate to gas ratio (CGR) is an important task in gas-condensate field management. Whitson et al. showed that the main difference in managing dry gas reservoirs against gas-condensate reservoirs is the requirement to forecast CGR so that condensate production can be reliably estimated [1]. Therefore, this parameter is extremely important, and the prediction on its evolution over time must be reliable.

The conventional method to predict CGR during pressure depletion is based on PVT (pressure - volume - temperature) model. To generate this model, representative samples of reservoir fluid must be collected, and PVT experiments must be conducted. This process is both costly and time consuming. After the main parameters of the fluid are determined, the equation of state (EOS) is then generated to describe the characteristics of the reservoir fluids. Without the representative samples and the main parameters acquired by experiments, the EOS cannot be derived. Similarly, if fluid properties from different reservoirs are used as analogue, the EOS may not reflect the actual characteristics of the reservoir fluids.

Date of receipt: 21/7/2023. Date of review and editing: 02/8 - 24/11/2023.
Date of approval: 7/12/2023.

Another conventional method for prediction is dynamic simulation. However, for most of Hai Thach wells, CGR prediction using this method is difficult although history matching of wellhead pressure is acceptable as shown in Figures 1 and 2.

History matching of CGR is challenging in Hai Thach field due to many reasons. First, there are 9 separate reservoirs in Hai Thach field but only one representative sample could be collected. Therefore, there is insufficient data to generate PVT models for all reservoirs. For reservoirs without representative samples, it is assumed that they might have similar properties to the sampled reservoir; thus, PVT models for these reservoirs have many uncertainties. Furthermore, most Hai Thach wells are produced from commingled reservoirs with contribution among them varying over time. Therefore, history matching of CGR using dynamic simulation is

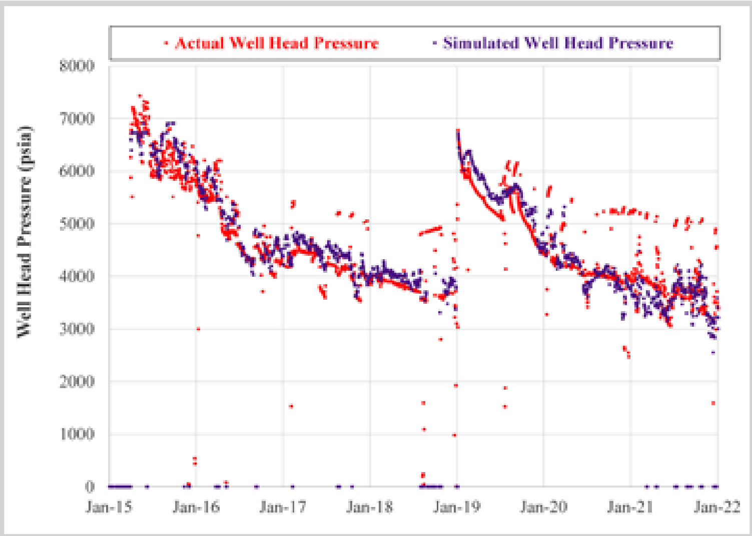

***Figure 1.** History matching results of well head pressure using dynamic simulation are reasonable for HT-Y.*

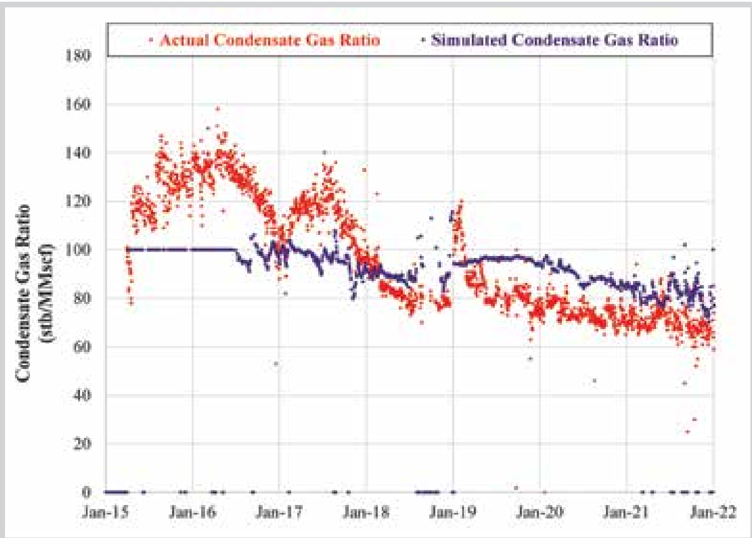

***Figure 2.** History matching results of CGR using dynamic simulation for HT-Y still have some difficulties.*

complicated and there is certain deviation from actual values, causing great difficulty in forecast. Regarding short-term forecast, this deviation has impact on condensate lifting because condensate production forecast would be too low or too high compared to reality, leading to high risks of shortage or tank top issues. With respect to long-term production forecasts, deviation in forecasting CGR results in inaccuracy of condensate production rate and cumulative condensate production and affecting the economic evaluation of the whole project. As a result, the necessity to develop an auxiliary method to precisely forecast CGR is very urgent.

In comparison with complex EOS modeling or dynamic simulation, machine learning can perform forecast with fewer input parameters. Therefore, it has been widely used to solve many forecasting problems. The application of this method in predicting the fluid properties of oil and gas reservoir has also been studied by a number of research groups, for example applying machine learning algorithms to predict the dew point pressure of gas condensate reservoirs [2 - 5], estimate CGR [6, 7], and fluid composition [8]. However, these studies have not concentrated on predicting the change of CGR over time as reservoir pressure declines during production.

For production wells, it can be observed that CGR is highly dependent on wellhead pressure, wellhead temperature, as well as choke size. Additionally, add-on perforation also showed great impact on CGR. Since the values of wellhead pressure, wellhead temperature, and choke size are monitored and recorded regularly, establishing the relationship between these parameters with those more difficult and expensive to obtain such as CGR will bring lots of practical values. Consequently, the application of machine learning to forecast CGR can be an alternative to overcome the difficulties of traditional methods.

However, machine learning also has some difficulties in forecasting over time, especially relatively long-term forecast, as reported in the following studies. Lee et al. constructed a long short - term memories (LSTM) network trained on the data of 300 wells to predict the production of 15 wells with good results but they are short term forecasts of just one month [9]. In another study, Zhan et al. used data from more than 300 unconventional oil wells to build two LSTM models, one to forecast the decline in production rate and the other to predict cumulative production [10]. For each well, the production history of the first 3 months was used to train the model and the remaining 21 months was used for testing. To overcome the problem of error accumulation in time series prediction and the challenge of capturing the steep production decline at the beginning, in addition to tubing pressure and oil rate, 12 wells most similar to the well being analyzed were selected from the database and translated into additional machine learning features. However, the results from the LSTM models still showed over-prediction for the production rate and under-prediction for the cumulative production. It can be observed that many machine learning models have certain

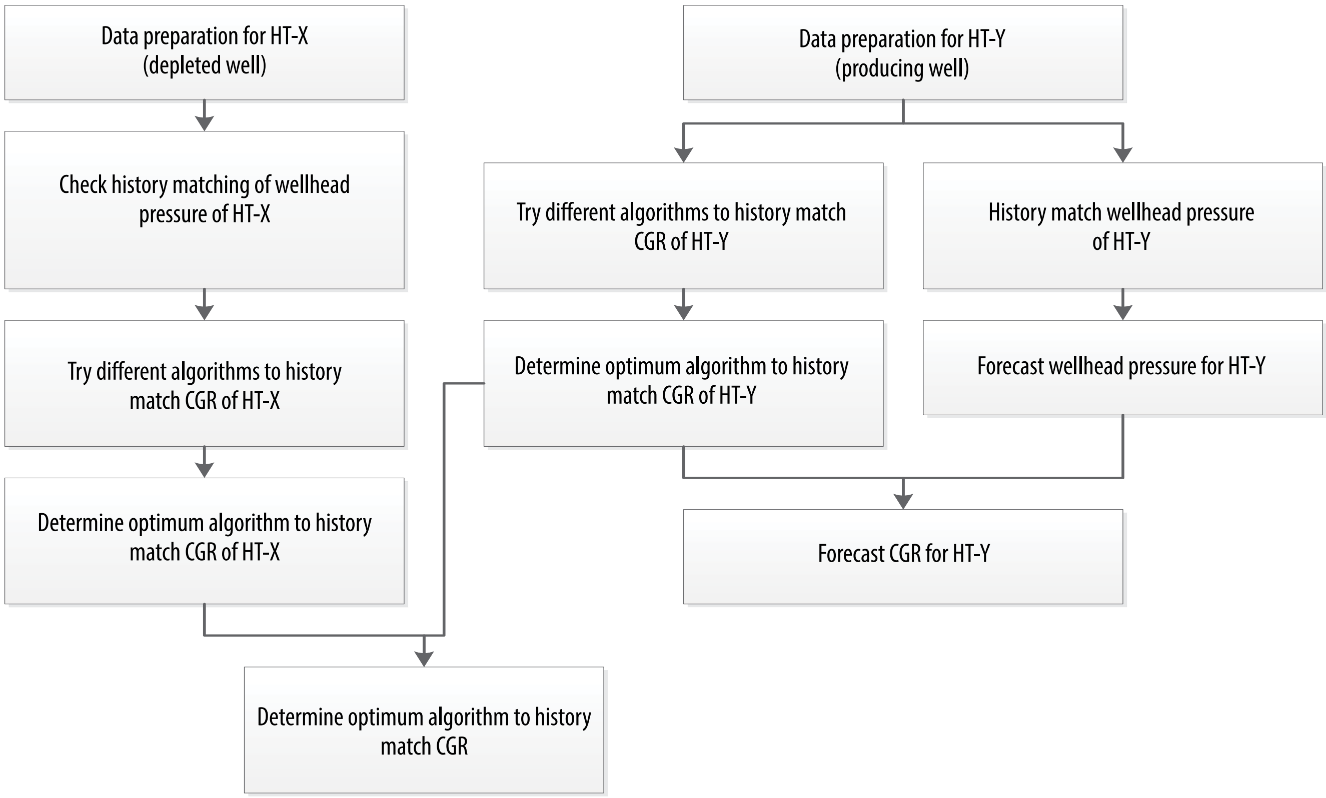

Figure 3. *Flow diagram of the study.*

Table 1. *HT-X WHP prediction using machine learning with different split ratios of training and testing dataset*

Split ratio of training and testing dataset	Mean_leaf	Training error (%)	Testing error (%)	Prediction results
50/50	20	-1.5 to 1.5	-15 to 0	Not very good
60/40	100	-15 to 15	-7.5 to 10	Relatively reasonable
70/30	50	-3 to 3	-10 to 7.5	Good forecast results
80/20	100	-8 to 8	-2 to 10	Good forecast results

difficulties in long-term forecasting. Furthermore, picking up abnormal changes in the production history is also a great challenge for wells with add-on perforation. Due to the above challenges, piecewise regression combined with linear regression and XGBoost is used to solve this forecasting problem.

2. Methodology

CGR is a parameter that depends on reservoir pressure, and therefore on wellhead pressure. Thus, this study was divided into two steps. The first step was to forecast the decline of wellhead pressure during production period. The next step was to forecast CGR according to the decline of wellhead pressure.

In the Hai Thach field, the HT-X well started production in 2015 with good deliverability. After 5 years of production, HT-X was depleted with wellhead pressure decreasing to process pressure and CGR decreased from initial value of 100 stb/MMscf to only 10 stb/MMscf. Since the historical data of wellhead pressure and CGR of HT-X was complete, this well is used to develop the machine learning model. The wellhead pressure data set consists of 1566 data points from daily production history of HT-X including uptime, choke size, and gas production rate. The CGR data set consists of 52 data points from flow tests including uptime, choke size, and wellhead pressure. Since the production history dataset used for forecasting CGR of HT-X is relatively small, the mean_leaf parameter will have a big impact on forecast results. The combination of piecewise regression and linear regression or piecewise regression and XGBoost will be used for the prediction of CGR of HT-X with different split ratios of training and testing datasets to find the optimal algorithm.

The same process is then applied to the prediction of well head pressure and CGR of HT-Y which is the main target of this study. HT-Y also started production in 2015 but this well has better performance and pressure was not declining as fast as HT-X. After 7 years of production, HT-Y is still the strongest gas producer in Hai Thach field. The forecast of CGR over time will help better manage HT-Y production. The wellhead pressure data set of HT-Y consisted of 1658 data points from daily production history including uptime, choke size, and gas production rate. The CGR data set consists of 132 historical data points of flow tests including uptime, choke size, and wellhead pressure. A big difference between HT-Y and HT-X is that HT-Y had add-on perforation that significantly changes the historical trend and that event would be used to check the capability of the algorithms.

The flow diagram of the study is shown in Figure 3.

3. Study result

The dataset used for forecasting wellhead pressure of HT-X is splitted into training and testing set with different ratios. The mean_leaf parameter is optimized based on the highest score of correlation factor by comparing forecast results and actual data on training dataset. The calculation results are shown in the following figures.

The combination of piecewise regression and linear regression is applied to history match wellhead pressure data of HT-X with different split ratios of training and testing dataset as shown in Table 1, and representative results shown in Figure 4.

The testing results show that history matching and

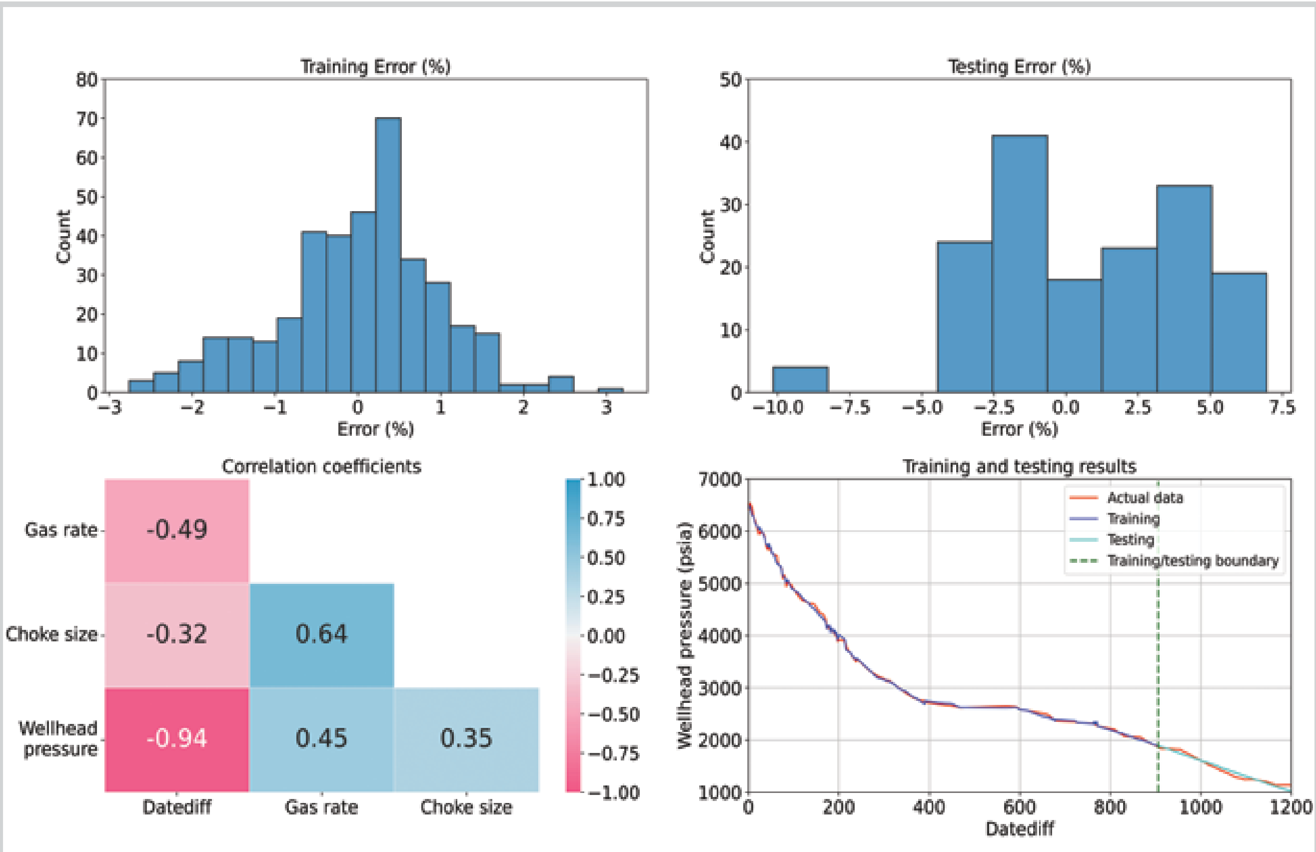

Figure 4. *HT-X WHP prediction using machine learning with split ratio of training and testing dataset of 70/30.*

Table 2. *HT-Y WHP prediction using machine learning with different split ratios of training and testing dataset*

Split ratio of training and testing dataset	Mean_leaf	Training error (%)	Testing error (%)	Prediction result
50/50	83	-5 to 5	-8 to 3	Reasonable
60/40	199	-8 to 8	-8 to 4	Reasonable
70/30	61	-8 to 8	-12 to 2	Reasonable
80/20	184	-8 to 8	-3 to 6	Reasonable

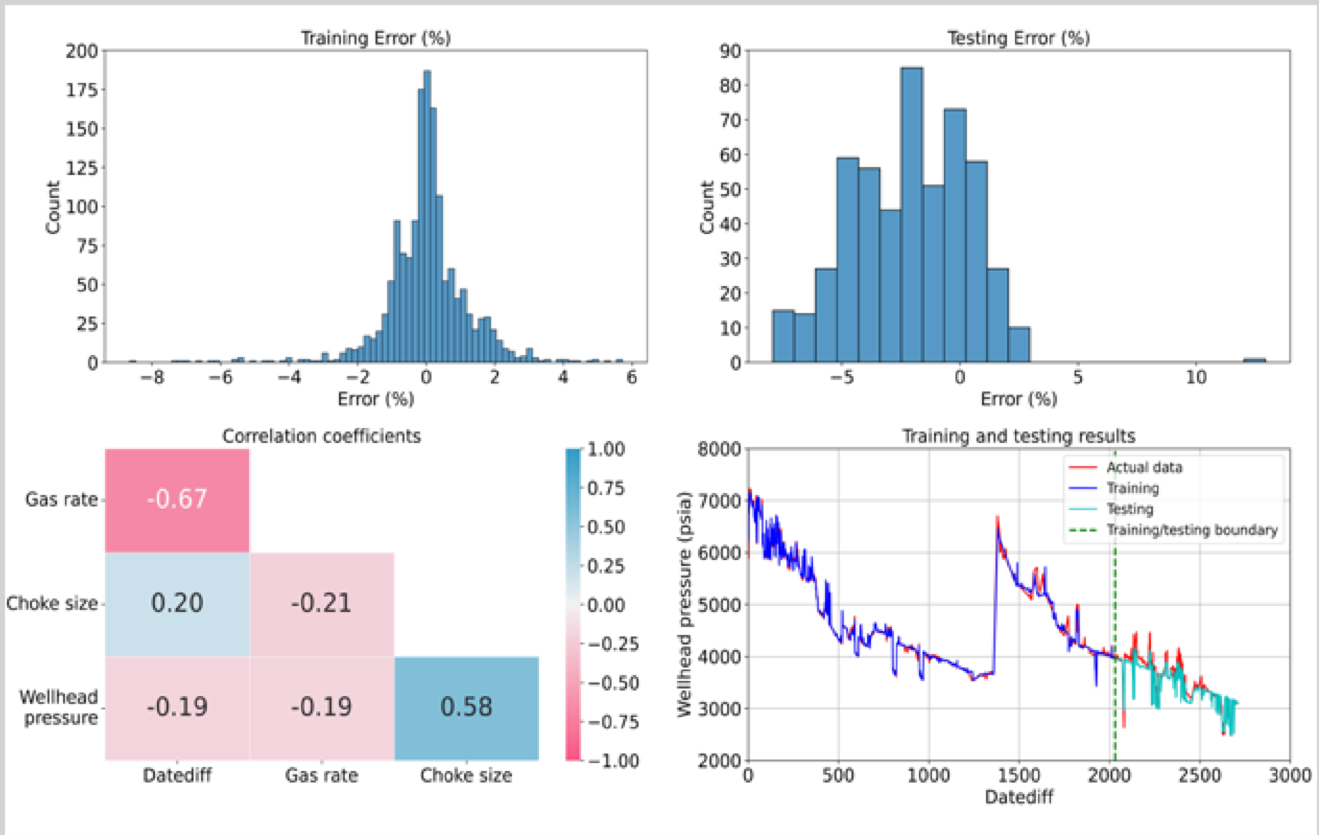

***Figure 5.** HT-Y WHP prediction using machine learning with split ratio of training and testing dataset of 50/50.*

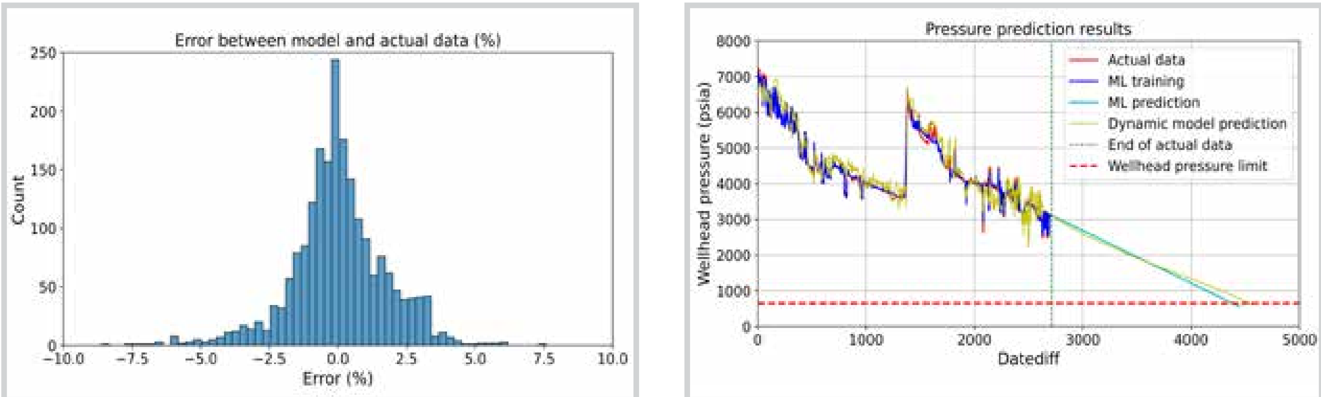

***Figure 6.** HT-Y WHP prediction using machine learning in comparison with dynamic simulation results.*

***Table 3.** HT-X CGR prediction using different split ratios of training and testing dataset*

Split ratio of training and testing dataset	Algorithm	Mean_leaf	Training error (%)	Testing error (%)	Prediction results
70/30	Piecewise regression combined with linear regression	6	-10 to 25	0 to 70	Reasonable
80/20		8	-15 to 25	-30 to 40	Reasonable
70/30	Piecewise regression combined with XGBoost	6	-10 to 25	0 to 70	Reasonable
80/20		8	-15 to 25	-40 to 40	Reasonable

forecast for wellhead pressure decline is feasible for HT-X when the ratio of historical data/forecast data is at least 60/40. The same process is then repeated for history matching and forecast of wellhead pressure of HT-Y. Since HT-Y wellhead pressure changes significantly after add-on perforation, the same split ratios of 50/50, 60/40, 70/30, and 80/20 are still applied, but only to the data after add-on perforation. The prediction results with different split ratios are summarized in Table 2 and representative results shown in Figure 5.

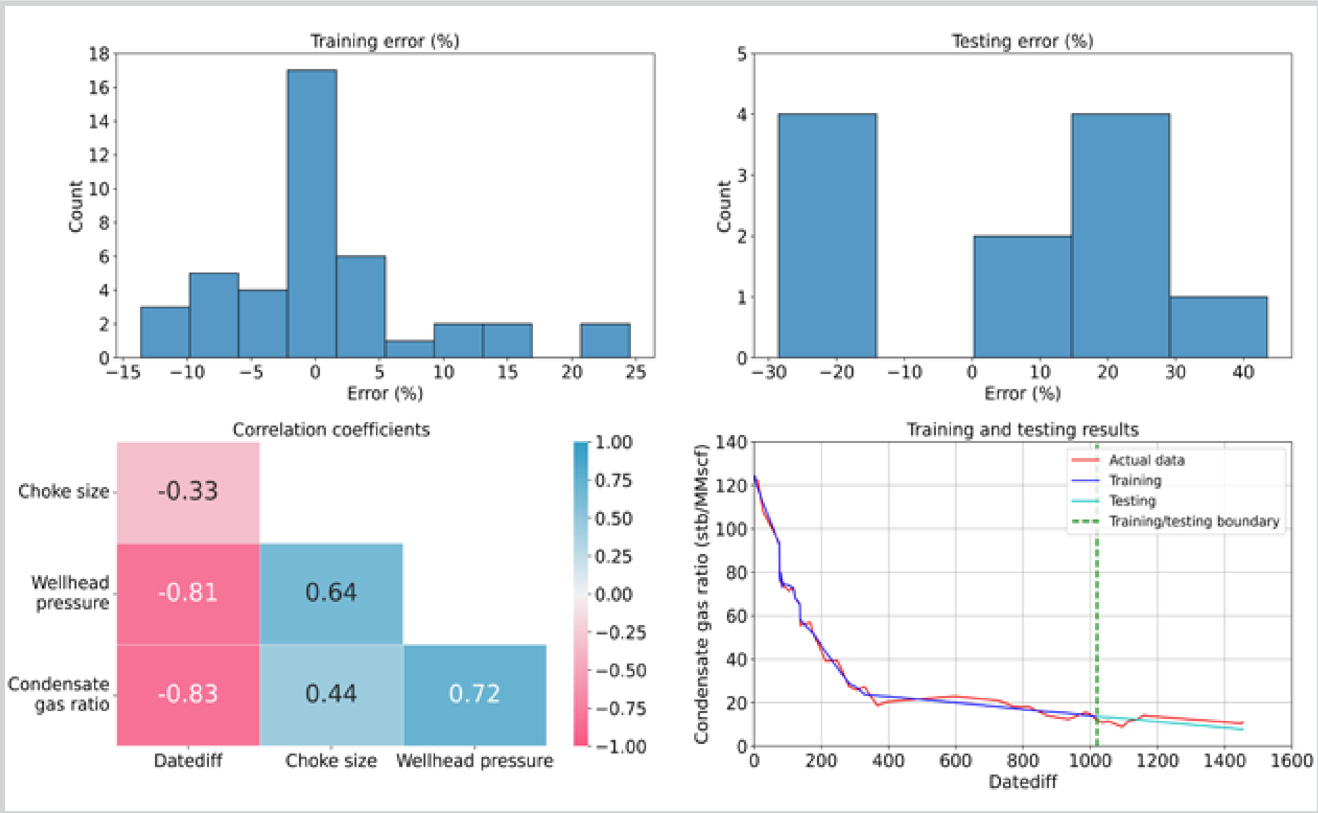

Figure 7. *HT-X CGR prediction using piecewise regression and linear regression with split ratio of training and testing dataset of 80/20.*

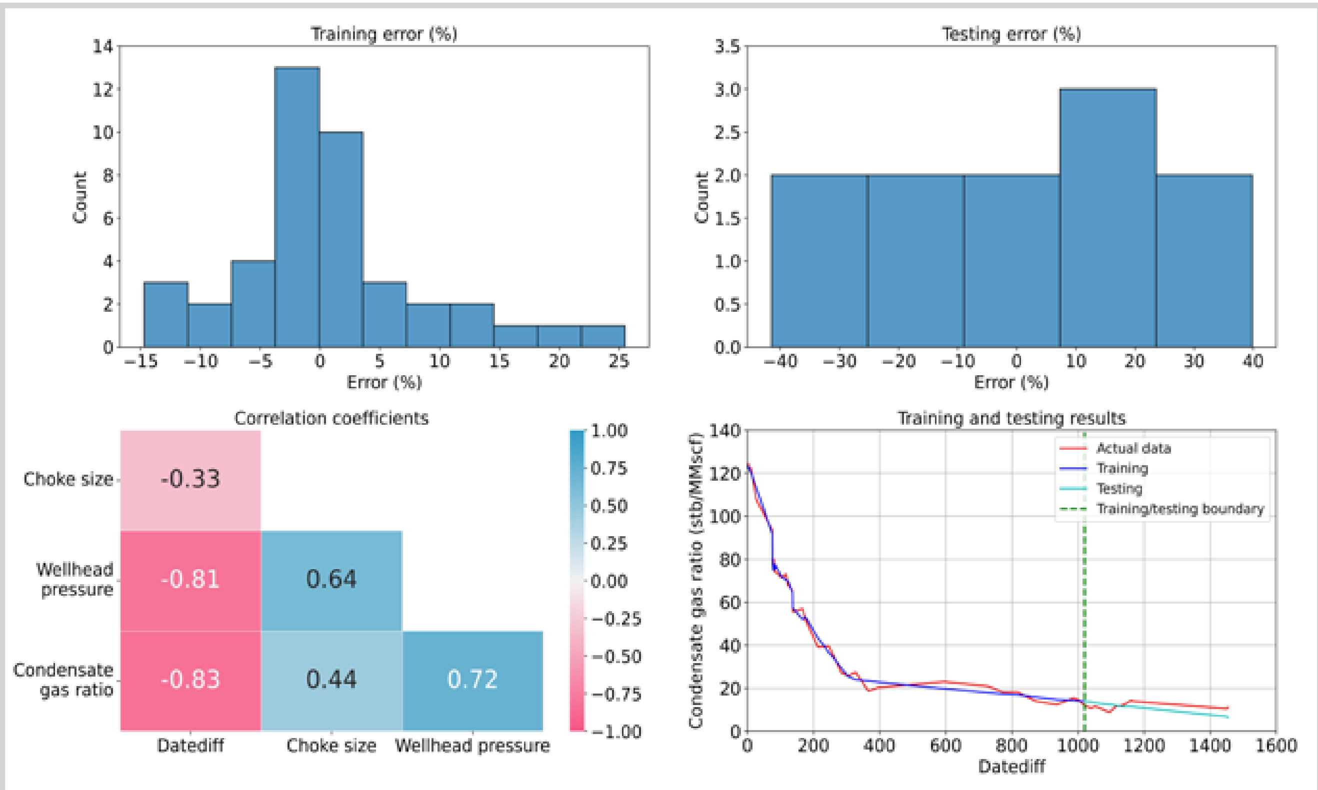

Figure 8. *HT-X CGR prediction using piecewise regression and XGBoost with split ratio of training and testing dataset of 80/20.*

The forecast results of wellhead pressure for HT-Y in the future by machine learning are compared to the results forecasted from dynamic simulation in Figure 6. The comparison of the results of the two forecasting methods shows that the use of machine learning for predicting HT-Y wellhead pressure is reasonable and slightly different compared to dynamic simulation results.

After the wellhead pressure is predicted with good accuracy, CGR is then derived by machine learning. Since the CGR dataset is relatively small compared to the wellhead pressure dataset, only 70/30 and 80/20 split ratios of training and testing datasets are applied. The forecast results for HT-X are summarized in Table 3 with representative results shown in Figure 7 and Figure 8.

For the CGR prediction of HT-X, piecewise regression combined with linear regression or XGBoost provide similar forecast results. Therefore, CGR of HT-Y was also predicted by both algorithms with split ratios of 70/30 and 80/20 as shown in Table 4.

Piecewise regression combined with XGBoost has better prediction results than piecewise regression combined with linear regression at 70/30 split ratio. Forecast results are similar for both algorithms at 80/20 split ratio, as shown in Figures 9 and 10.

In summary, piecewise regression combined with XGBoost has better and more stable forecast results, so this algorithm is used to predict future CGR of HT-Y and the results from machine learning are compared to the forecast results by dynamic simulation in Figure 11. The comparison of two methods shows that machine learning prediction of CGR for HT-Y is reasonable and can be used as a supplement of dynamic simulation forecast for production management.

Table 4. *HT-Y CGR prediction using different split ratios of training and testing dataset*

Split ratio of training and testing dataset	Algorithm	Mean_leaf	Training error (%)	Testing error (%)	Prediction results
70/30	Piecewise regression combined with linear regression	20	-15 to 10	-15 to 30	Overestimated
80/20		30	-15 to 10	-15 to 25	Reasonable
70/30	Piecewise regression combined with XGBoost	20	-15 to 10	-20 to 15	Reasonable
80/20		30	-15 to 10	-20 to 20	Reasonable

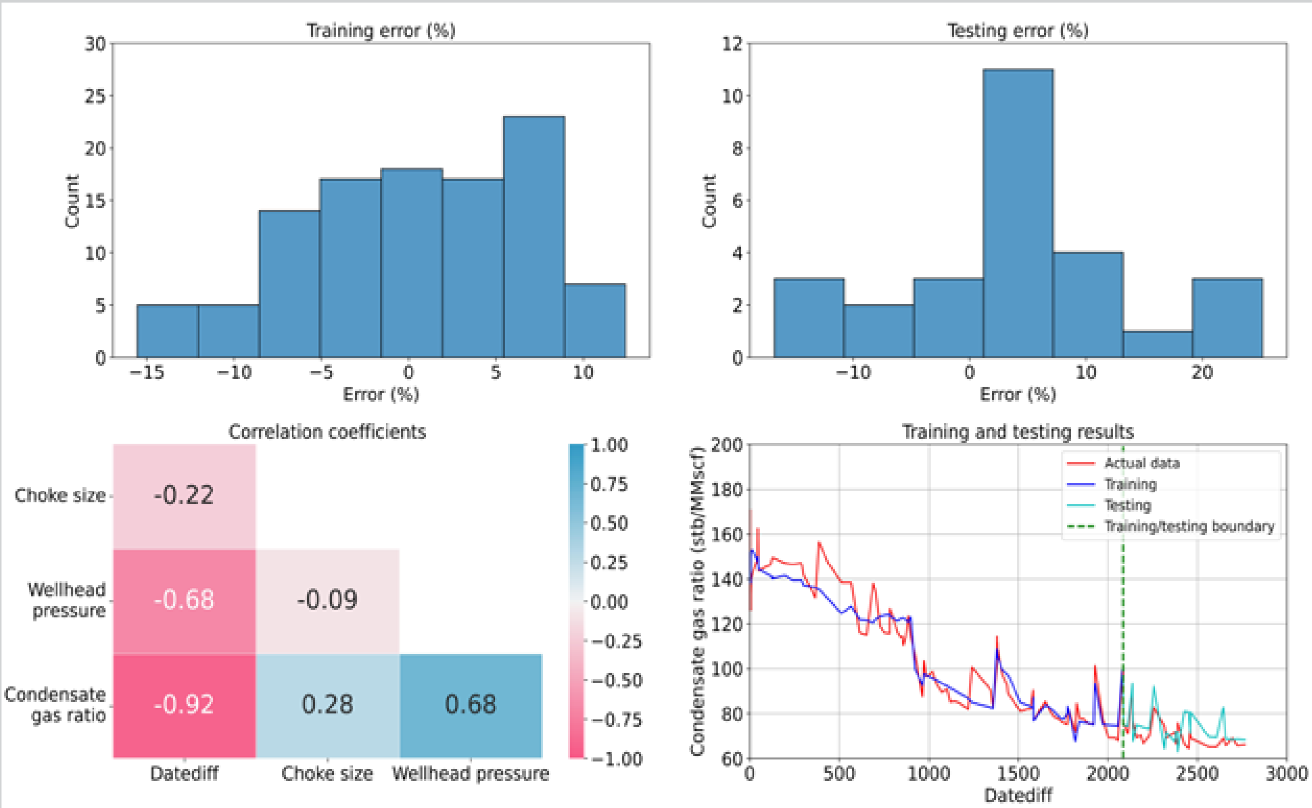

Figure 9. *HT-Y CGR prediction using piecewise regression and linear regression with split ratio of training and testing dataset of 80/20.*

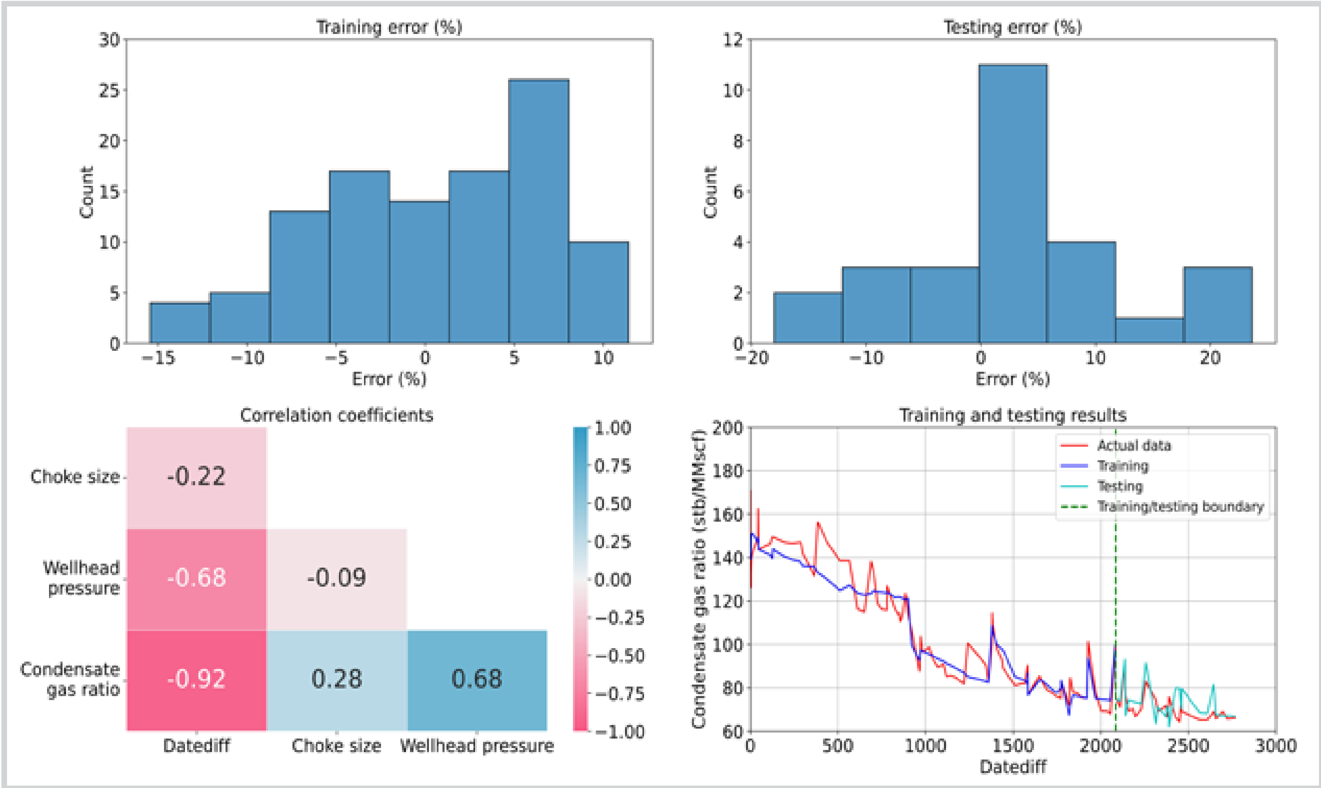

Figure 10. *HT-Y CGR prediction using piecewise regression and XGBoost with split ratio of training and testing dataset of 80/20.*

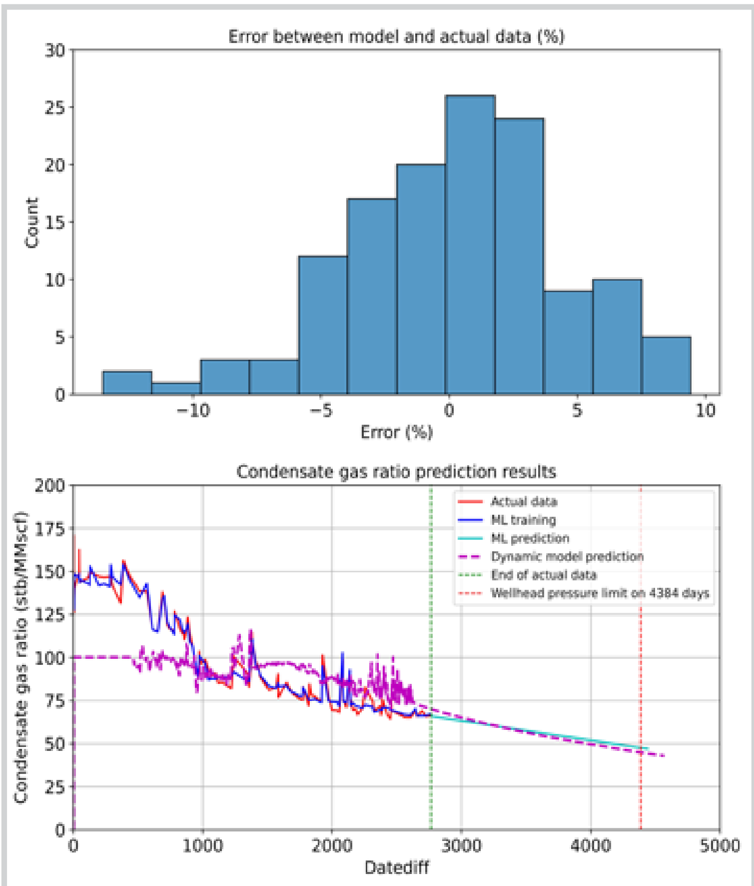

Figure 11. *HT-Y CGR prediction using piecewise regression combined with XGBoost in comparison with dynamic simulation results.*

4. Conclusions

The main conclusions of the study can be summarized as follows:

- Machine learning is applied successfully to predict the changes overtime of CGR which is one of the most important parameters for gas-condensate reservoirs but very challenging to forecast using traditional methods;

- About machine learning algorithm, piecewise regression combined with XGBoost provides reasonable and reliable forecast results for CGR;

- The successful application of machine learning in forecasting CGR during production provides significant support to the prediction of condensate production, thereby helping to better optimize production management of gas condensate fields.

Reference

[1] Curtis H. Whitson, Øivind Fevang, and Tao Yang, “Gas condensate PVT - What’s really important and why?”, *IBC Conference “Optimisation of Gas Condensate Fields”, London, 28 - 29 January 1999*.

[2] Meisam Karbalaee Akbari, Farhang Jalali Farahani, and Yaser Abdy, "Dewpoint pressure estimation of gas condensate reservoirs, using artificial neural network (ANN)", *SPE Europec/EAGE Annual Conference and Exhibition*, *London, United Kingdom, 11 - 14 June 2007*. DOI: 10.2118/107032-MS.

[3] Seyed Mohammad Javad Majidi, Amin Shokrollahi, Milad Arabloo, Ramin Mahdikhani-Soleymanloo, and Mohsen Masihi, "Evolving an accurate model based on machine learning approach for prediction of dew-point pressure in gas condensate reservoirs", *Chemical Engineering Research and Design*, Volume 92, Issue 5, pp. 891 - 902, 2014. DOI: 10.1016/j.cherd.2013.08.014.

[4] Zhi Zhong, Siyan Liu, Mohammad Kazemi, and Timothy R. Carr, "Dew point pressure prediction based on mixed-kernels-function support vector machine in gas-condensate reservoir", *Fuel*, Volume 232, pp. 600 - 609, 2018. DOI: 10.1016/j.fuel.2018.05.168.

[5] Princewill Ikpeka, Johnson Ugwu, Paul Russell, and Gobind Pillai, "Performance evaluation of machine learning algorithms in predicting dew point pressure of gas condensate reservoirs", *SN Applied Sciences*, Volume 2, 2020. DOI: 10.1007/s42452-020-03811-x.

[6] Sohrab Zendehboudi, Mohammad Ali Ahmadi, Lesley James, and Ioannis Chatzis, "Prediction of condensate-to-gas ratio for retrograde gas condensate reservoirs using artificial neural network with particle swarm optimization", *Energy & Fuels*, Volume 26, Issue 6, pp. 3432 - 3447, 2012. DOI: 10.1021/ef300443j.

[7] Mohammad Ali Ahmadi, Mohammad Ebadi, Payam Soleimani Marghmaleki, and Mohammad Mahboubi Fouladi, "Evolving predictive model to determine condensate-to-gas ratio in retrograded condensate gas reservoirs", *Fuel*, Volume 124, pp. 241 - 257, 2014. DOI: 10.1016/j.fuel.2014.01.073.

[8] Hana AlMatouq, Mohammed Alabbad, and Fatai Anifowose, "An artificial intelligence approach to predict molar compositions of reservoir fluid components", *SPE Gas & Oil Technology Showcase and Conference, Dubai, UAE, 21 - 23 October 2019*. DOI: 10.2118/198555-MS.

[9] Kyungbook Lee, Jungtek Lim, Daeung Yoon, and Hyungsik Jung, "Prediction of shale gas production at Duvernay formation using deep-learning algorithm", *SPE Journal*, Volume 24, Issue 6, pp. 2423 - 2437, 2019. DOI: 10.2118/195698-PA.

[10] Cheng Zhan, Sathish Sankaran, Vincent LeMoine, Jeremy Graybill, and Didi-Ooi Sher Mey, "Application of machine learning for production forecasting for unconventional resources", *Unconventional Resources Technology Conference, Denver, Colorado, USA, 22 - 24 July 2019*. DOI: 10.15530/urtec-2019-47.