Journal of Mining and Earth Sciences

Website: https://jmes.humg.edu.vn

Research on the application of artificial intelligence tools to predict the operating temperature and pressure of the pipeline transportation system at the Hai Thach - Moc Tinh field

Tuan Thanh Nguyen*, Thinh Van Nguyen, Truong Hung Trieu

Hanoi University of Mining and Geology, Hanoi, Vietnam

ARTICLE INFO

Article history:
Received 15th July 2023
Revised 14th Oct. 2023
Accepted 11th July 2024

Keywords:
Artificial intelligence (AI),
Hai Thach-Moc Tinh,
Machine learning,
Pipeline transportation system.

ABSTRACT

The observation of operating parameters for the single-phase or multi-phase flows of oil and gas pipelines always plays an important role and prerequisite in daily operation. The inlet and outlet parameters of the pipeline such as pressure and temperature will greatly affect the efficiency of the production and transportation process of fields. These parameters are very important. It must be required to observe carefully and strictly. In fact, pressure and temperature signal gauges will always be set up at the inlet and outlet of the pipeline for continuously variable transmission of the signal to the central control room, then the operator can observe regularly. Alarms will be triggered when the signal goes out of the setting operation area. However, the operational parameters transmitted from this gauge are only available after the actual fluid flow passes through the pipeline. This will limit when the operator needs to apply calculation methods to pre-predict the flow regime, and operating parameters of the fluid in the pipeline when fluid flows has not passed through the pipeline. From that approach, the paper presents the result of research on the application of machine learning algorithms (ML) to build models to predict operating conditions of three-phase flows in the pipeline at Hai Thach- Moc Tinh field based on input parameters such as the open of wellhead flow control valve of wells located in wellhead platform WHP-MT1, WHP-HT1, and the commercial gas volume. The results of the research show that the ML calculation method gives the result which is only +/-2 barg/2^{0}C difference compared to the field data obtained from pressure (HT1-PT-0911) and temperature (HT1-TI-0911) signals set up at the outlet of the transportation pipeline at the Hai Thach-Moc Tinh field.

*Corresponding author
E - mail: nguyenthanhtuan@humg.edu.vn
DOI: 10.46326/JMES.2024.65(4).04

1. Overview of the prediction of operating conditions (Pressure/Temperature) of the transportation pipeline of the Hai Thach-Moc Tinh field

The fluid transported in the pipeline has different flow structures, depending on factors such as fluid component, flow rate of transportation as well as liquid-gas ratio, diameter, and material of the pipeline, etc. In general, the flow regime can be divided into two main types based on the geometrical structure of the pipeline such as the flow regime in pipeline by horizontal and vertical directions. There are many names given to different flow regimes but in general, there are four regimes: slug flow, laminar flow, annular flow, and dispersed bubble flow. In general, it is difficult to determine exactly which flow regime is taking place in the pipe section Based on that it is possible to select an appropriate experimental coefficient for flow applying to the equation for calculating the pressure drop due to friction. Calculations are often repeated with different flow regimes. From that, it can be chosen the most suitable flow regime which has the smallest error. It takes time consumption, and the calculation error is quite high due to the flow regime chosen not being suitable. Therefore, calculation using simulation software is always the best choice. However, it is still difficult to determine the input parameters for the Hysys model (temperature, pressure, input flow component of the fluid) because of estimation by manual calculation from the characteristic curve and fluid component of each well. In some cases, especially when the liquid-gas ratio of the fluid flow increases, the calculation by Hysys software also gives an error "Not solved" because the thermodynamic model given by the user in Hysys does not calculate for this pipe section. Therefore, the Machine Learning (ML) algorithm has been studied based on historical data for the opening of choke size of wells, the gas flow rate output to get the pressure and temperature parameters at the 20th kilometer of the Hai Thach-Moc Tinh (HT-MT) pipeline and ignore the basic calculation steps above.

Today, there are many research works in the world applied to Machine Learning algorithms for pipelines and calculating pressure drop. The research of "Gas Gathering System Modeling the Pipeline Pressure Loss Match" carried out by authors R.G MCneil and D.R Lillico (McNeil and Lillico, 2005) is the most similar. In this research, the authors built up a mathematical model to calculate the pressure drop of the gas-gathering system. The authors created a gas-gathering system model that can calculate the backpressure with high accuracy and repeatability. It is able to predict future operating conditions when changing the operating information of wells or in case of adding new wells. The result of this research showed that with this mathematical model, it did not need to apply the complicated traditional pressure drop calculation methods. This model requires only the modeller to have a reasonable approach to research problems. The imperative approach starts with three rules: (1) divide the problem into manageable parts, (2) choose an appropriate pressure drop correlation and trust it without using any correction factor, and (3) always take a field trip to resolve the difference between the pipeline pressure drop by measuring and calculating.

Jane Ozi and Ayoade Kuye (2020) investigated pressure loss to predict the operational capability of pipeline systems, particularly for cross-border gas transportation pipelines. When pressure loss increases and reaches a point where the pressure drop in the pipeline exceeds the specified operating pressure limit or falls below the required minimum delivery pressure, it can impact the system's performance negatively.

2. Methods used to predict the operating conditions of transportation pipelines

2.1. Traditional calculation method

For single-phase flow in a pipeline, the pressure drop over a distance L is represented by the thermodynamic equilibrium equation:

$$\int VdP + \frac{g}{g_c}\Delta X + \frac{(\Delta v)^2}{2g_c} = -W_f - W \qquad (1)$$

Where V is the volume of fluid; P is the pressure of fluid; ΔX is the variation in height of fluid flow; Δv is the variation velocity of fluid flow; W_f is the energy loss due to friction; W is the energy of the system; g is the acceleration due to gravity (9.81 m/s^2); g_c is the mass/force conversion coefficient.

Formula (1), representing the friction force in a pipe length dL is proportional to the contact surface of the fluid, relative to the square of the velocity and the density of the fluid. When adding the friction coefficient f, the equation (1) is given as:

Frictional resistance = (f)(dL)(πd)($v^2/2g_c$)(ρ) (2)

Where L is the length of pipe (m); d is the inner diameter (m); v is the velocity of the fluid (m/s); ρ is the density of the fluid (kg/m^3); f is the size coefficient of pipe.

The mass of liquid in the pipe is the length dL multiplied by the cross-section of the pipe and the density of fluid. All work losses due to friction are calculated by the frictional resistance passing through the length dL. Combining the above concepts into equation (2) gives dWf.

$$dW_f = \frac{(f)(dL)(\pi d)\left(\frac{v^2}{2g_c}\right)(\rho)dL}{(\frac{\pi}{4})(d^2)(\rho)(dL)} \quad (3)$$

Simplify:

$$dW_f = \frac{2fv^2dL}{g_cd} \quad (4)$$

Integrating from 0 to W_f and 0 to L, we get:

$$W_f = \frac{2fLv^2}{g_cd} \quad (5)$$

Equation (5) is called the Fanning friction coefficient equation. Other forms of the equation have been published differently only in the value correction factor “f”.

Combining equations (1) and (5) gives:

$$\int VdP + \frac{g}{g_c}\Delta X + \frac{(\Delta v)^2}{2g_c} = -\frac{2fLv^2}{g_cd} - W \quad (6)$$

Equation (6) is also defined as the fundamental equation of flow or Bernoulli's equation.

With the pipe section without work used; W is removed from equations (6) to solve the problem between the pressure drop and flow rate, to get equation (7):

$$\int VdP + \frac{g}{g_c}\Delta X + \frac{(\Delta v)^2}{2g_c} = -\frac{2fLv^2}{g_cd} \quad (7)$$

The friction coefficient f depends on the roughness and diameter of the pipe, and the characteristics of fluid. Normally, the f-factor will be looked up according to the Moody's chart based on the Reynolds coefficient of the fluid (Campbell, 1992).

Normally for three-phase flows in pipeline transportation methods used to calculate the pressure drop without the help of computer tools will be quite complicated. In the Hai Thach - Moc Tinh field, this method is not usually used because of the large size of the pipeline (12 inches) and giving significant errors when applying Eaton's empirical equation for pipeline three-phase flow (because this method applies only to pipes having diameter less than 50 mm) or Beggs & Brill (<1.5 inches). The calculation results in a deficiency of liquid hold up compared to the reality obtained when applying Dukler's correction coefficient (Campbell, 1992). Therefore, this traditional method is not applied to calculate the pressure loss for the gathering pipeline at the Hai Thach - Moc Tinh field.

2.2. Calculation method by using simulation software

Currently, there are many simulation software applying these calculation methods and it allows the users to forecast future operating conditions such as HYSYS, OLGA, MSI,... In Vietnam, MSI software is available applied at the Nam Con Son pipeline (NCSP) taking the "what if" tool can predict the “settle out pressure”, the “landing pressure”. Based on the fact that operation team can adjust the inlet flow to gas owners. This software supports also NCSP to predict the "settle out pressure", and "landing pressure" when PVGas - the gas buyer announces an expected gas demand about a decrease or an increase in the next few hours/day. It helps the operation team to make requirements of appropriate input flow “production Guidance” for gas owners who are bringing gas into the NCSP pipeline system. The “what if” tool is also used in predicting the results (settle out pressure) of special operating situations such as: increasing or decreasing flow rate, pigging, and when stopping/restarting operation of offshore platforms. However, the licensing cost for this MSI software is expensive.

With the calculation of temperature and pressure drop using the Hysys model, in some cases of high Condensate-Gas Ratio (CGR), we will obtain results which do not suitable after many loops (Figure 1). On the other hand, compared with the database of the field, pressure drop calculation using the Hysys model gives also results of output pressure which is lower than reality itself (over the predicting pressure loss) because the thermodynamic model Peng-Robinson or SRK in Hysys uses the Beggs & Brill empirical equation to calculate the pressure loss. However, this equation is not suitable for pipes having a greater diameter (12 inches) of the Hai Thach-Moc Tinh pipeline. Specifically, with a gas flow rate of 4.0 million m^3/day from the wells at the WHP-MT platform, Hysys gives the results in pressure/temperature loss of 40barg/34^{0}C respectively. It is higher than 34barg/ 28^{0}C comparing the reality measured by pressure gauges HT1-PT0911 and temperature HT1-TI0911 located on the pipeline.

2.3. Research method using AI to predict the operating conditions of the pipeline

There are different types of machine learning (ML) algorithm models including linear and non-linear regression models, decision tree models, support vector machines, and artificial neural network. Some machine learning models are very flexible in architectural design, allowing unlimited model selection. Besides, each model has a large number of hyperparameters that can be optimized to maximize model performance. Therefore, model selection together with

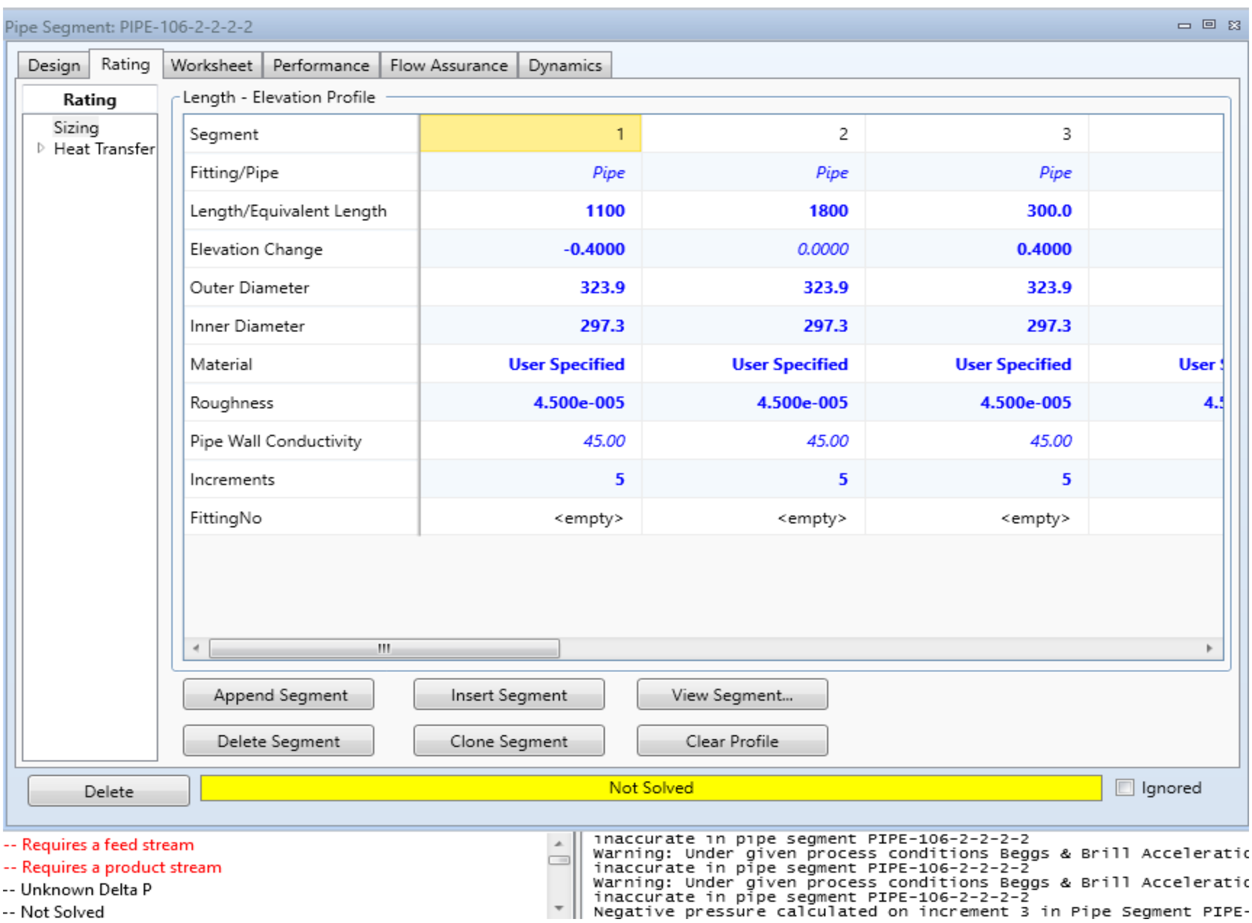

Figure 1. Hysys error message when pressure loss can not be calculated due to failure of solving thermodynamic equation with Begg and Brills calibration coefficient.

hyperparameter selection has become a difficult task in optimizing the use of machine learning models to solve certain problems. This section aims to estimate the basement theory of major machine learning models that have the potential to be applied to the current objective. In addition, the advantage and disadvantage of each model are also clarified, so that different models can be combined to build the most efficient model possible to solve the current problem. In particular, it needs a model covering the declining trend of the data (both gas condensate ratio and wellhead pressure) and predicting the outlet future dataset that is outside of the historical data zone. In addition, models that can perfectly adapt to small amounts of gas condensate ratio and small amounts of fluid components are required.

Linear regression may be the simplest machine learning model because it searches only to optimize the weights w_i concerning linearly related between the objective y of each sample and the corresponding characteristics of x_i mentioned by Montgomery et al. (2021):

$$y = w_0 + \sum_{k=1}^{N} w_i x_i \quad (8)$$

Where N is the number of features and w_0 is the intersection coefficient. By training the linear regression model, machine learning searches for the w_i component with i valeur varies from 0 to N to minimize the mean square error (applied all over samples of the dataset) between the target calculated by using the equation above and the measurable objective. A part of extending the linear regression model is the non-linear regression model, which also takes into account the nonlinear factors of the features such as square, cube, product between features, and so on. This corresponds to a linear regression model in which new features are generated from the basic properties through basic mathematical equations. In the case of small datasets, the parameter should be added to the linear regression model to avoid the problem of overfitting. To clarify this point, a pseudo example of a dataset having only two data points is considered. This dataset can be perfectly matched by a linear regression model going directly through these two points. However, this data-fitted model is not optimal for predicting a third data point that is not on the path through the original two points. Obviously, it is necessary to build a model that does not take into account for fitting the first two points perfectly but it can predict better than the third data point. Using the penalty parameter in the linear regression model to avoid the problem of overfitting is the main idea of the Ridge model (McDonald, 2009). In this model, we can set a hyperparameter (α parameter in the famous Scikit learning package) that controls the slope of the regression hyperplane, i.e. It controls the sensitivity of the target to properties variation. The higher the α parameter, the greater the sensitivity of the target varies to the characteristics and opposite. Similarly, Lasso or Elastic-Net models consider the regularization parameter to minimize the number of non-zero weights (Zhao and Yu, 2006). However, Lasso and Elastic-Net models which are suitable for reducing the number of features may not be a good choice for this problem having only three properties (wellhead pressure, wellhead temperature, and opening valve of production process) (Jia and Yu, 2010).

By training the artificial neural network, machine learning will find ways to optimize the weights and deviations to minimize the error between the predicted target and the measuring target. That error is called the loss function. For an actual regression problem, a typical loss function might be the mean squared error between the predicted gas condensate ratio and the actual measured gas condensate ratio. Mathematical algorithms for minimizing the loss function are called optimizers based on the random gradient reduction technique. These techniques aim to find out, in a multi-step process, the optimal direction in a multi-dimensional space in which the attenuation of the loss function is greatest. Then, the weights and deviation of the network are modified step by step to reduce the loss function in that optimal direction. The learning rate and momentum parameters are considered in such modification to stabilize the attenuation value of the loss function towards its overall minimum. Such modifications are made by starting from the last hidden layer back to the first hidden layer. This opposite change of weight and bias is called the backpropagation algorithm. By optimizing the artificial neural network model, the network architecture is searched for optimization (optimizing the number of hidden layers and the number of hidden nodes in each hidden layer), optimization of the choice of functions activation, loss function, and optimization associating with related learning ratio. Randomly disabling a certain number of hidden nodes is also a wide technique applied in a real condition to optimize the performance of an artificial neural network model. Such optimization is controlled by validating the trained model with the dataset that has not been used during the training process of the model. This cross-validation step, just like in any other machine learning model, is very important to avoid the problem of overfitting that causes the model to overfit the training dataset and it can work only on this dataset. A surface artificial neural network model having a small number of hidden nodes would be an appropriate choice for modelling the dataset of gas condensate ratios due to the small amount of data. However, it is not possible to apply the artificial neural network model to the fluid component dataset because the amount of data is very small while many features belong to this dataset.

The decision tree model is based on a fairly simple concept for data classification and regression. It uses conditions based on features to split the original dataset into categories. For a regression problem, the predicted value for each category is the mean of that category taken from the training dataset. For example, the condensate gas ratio dataset can be divided into two subsets: one of which having an opening valve is greater than 40% and the other having a discharge opening valve is less than 40%. Then each subset can be divided into two smaller subsets based on the wellhead temperature and the wellhead pressure respectively. A feature can be used multiple times to subdivide a data set. Finally, each subset corresponds to a leaf of a decision tree, where the regression value is the mean of the corresponding subset. By training the decision tree, machine learning will find a way to optimize the partition thresholds in order to minimize the error between the predicted target and the measured target. The structure of the decision tree must also be simplified to avoid the overfitting. This can be solved by considering regulatory parameters such as the minimum number of data points in each leaf of the maximum number of branches coming from the root to the furthest leaf (maximum depth of tree). A single decision tree may not work well on a complex dataset. However, a set of many decision trees can become a very powerful machine-learning model. The random forest model is an example of the decision tree model working very well on different types of tabular data. This model considers random starter datasets built from the original dataset and selects random features from each starter dataset to build and optimize multiple random decision trees. The final result of the model is the average value of the results obtained from the trees. Extra trees are another high-performance synthetic decision tree model based on a different concept from the random forest model. It considers the random split value at each decision node to build the tree. This technique helps to reduce the variance and increase the deviation of the model. This method works very quickly and it can run better than the random forest model or the others coming from well-known machine learning models in some applications. The motivation tree model can be considered as an extension of the random tree model. This modelling layer does not build trees at random, but each new tree is built by learning from the previous tree's errors. The Adaboost model could be considered the simplest motivation tree model considering the trees having only two leaves. The models having progressive amplitude consider more complex tree structures to the depth and have more leaves. The XGBoost model is a synthetic reinforcement tree model that had won in plural machine learning competitions concerning tabular dataset development. This model considers the specified parameters to remove and simplify the tree structure in order to avoid the problem of overfitting. It is affirmed better than other machine learning models such as support vector machines or artificial neural networks in many industrial applications.

The support vector machine model uses a hyperplane to divide the dataset into subcategories. It uses a core function to transform the original dataset into a space where the hyperplane exists. Comparing this model with the artificial neural network model and tree model for different tabular datasets can be found in the research of Nguyen-Sy et al. (2020) and Nguyen-Sy et al. (2021). In practical experience, artificial neural networks and high-level tree models often work better than support vector machine models.

3. Results and discussion

In this section, the research is carried out by optimizing and comparing machine learning models for hydrodynamic datasets of internal gathering pipeline three-phase flows at the Hai Thach-Moc Tinh field to solve the current problem of temperature and pressure prediction. To evaluate the model's performance based on predictive data, 20% of the total data points are kept to test the trained model. The rest of the data is used for training. To do this, the train_test_split function provided by the Scikit learning package was applied.

The root mean square error index (RMSE) and the mean absolute error (MAE) were calculated separately for the training and testing datasets. These two parameters will be used to compare machine learning models to evaluate predictive performance between models.

The linear regression model is the first model selected to test the prediction of temperature and pressure conditions of pipelines from operating parameters of production wells. The temperature and pressure results are evaluated on the training and testing dataset. It can be found that the linear regression model, although it is a simple model, works quite well. Regarding temperature parameters, the RMSE and MAE indexes are at 2.09 and 1.5^{0}C for the training set; 2.01 and 1.47^{0}C for the test set. Regarding pressure parameters, RMSE and MAE are at 0.75 and 0.59 barg for the training set; 0.73 and 0.56 barg for the test set (Figure 2).

The Support Vector Machine is the next artificial intelligence model to be tested. The hyperparameters of the model are optimized by sensitivity analysis. Regarding temperature parameters, the RMSE and MAE indexes are at 1.41 and 0.96^{0}C for the training set; 1.73 and 1.12^{0}C for the test set. Regarding pressure parameters, RMSE and MAE are at 0.6 and 0.51 barg for the training set; 0.69 and 0.55 barg for the test set (Figure 3).

Next, a more complex artificial intelligence model is tested. The authors use an artificial neural network model with two hidden layers of

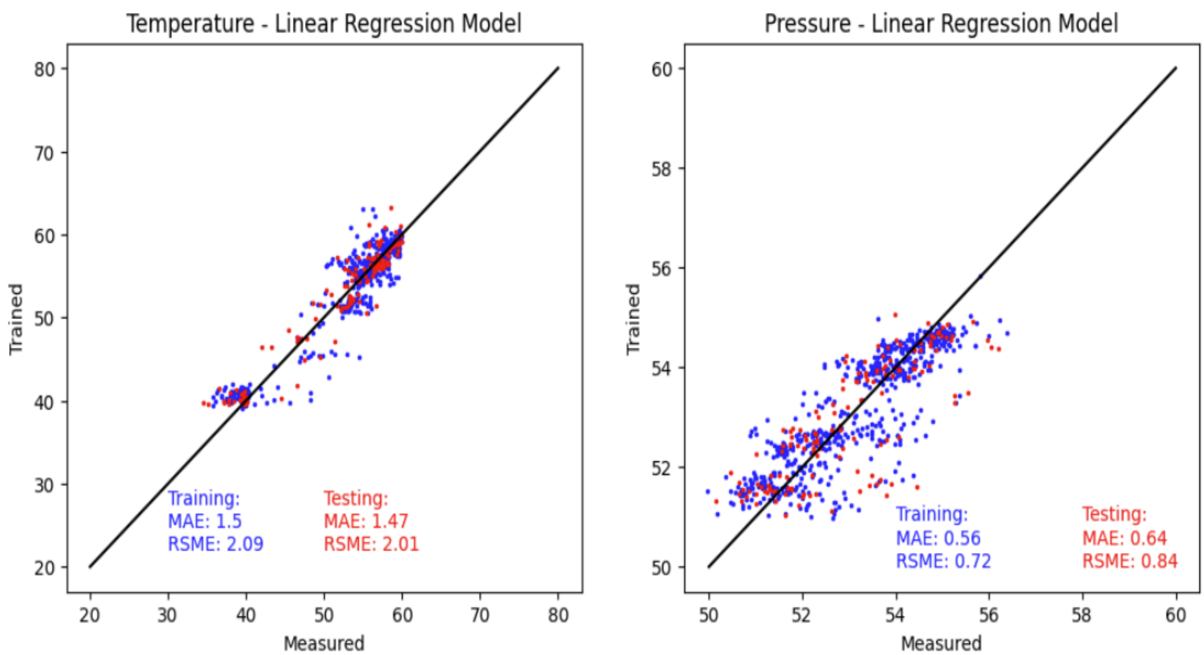

Figure 2. Comparison between predicted temperature and pressure with measured values for training and test subsets by linear regression model.

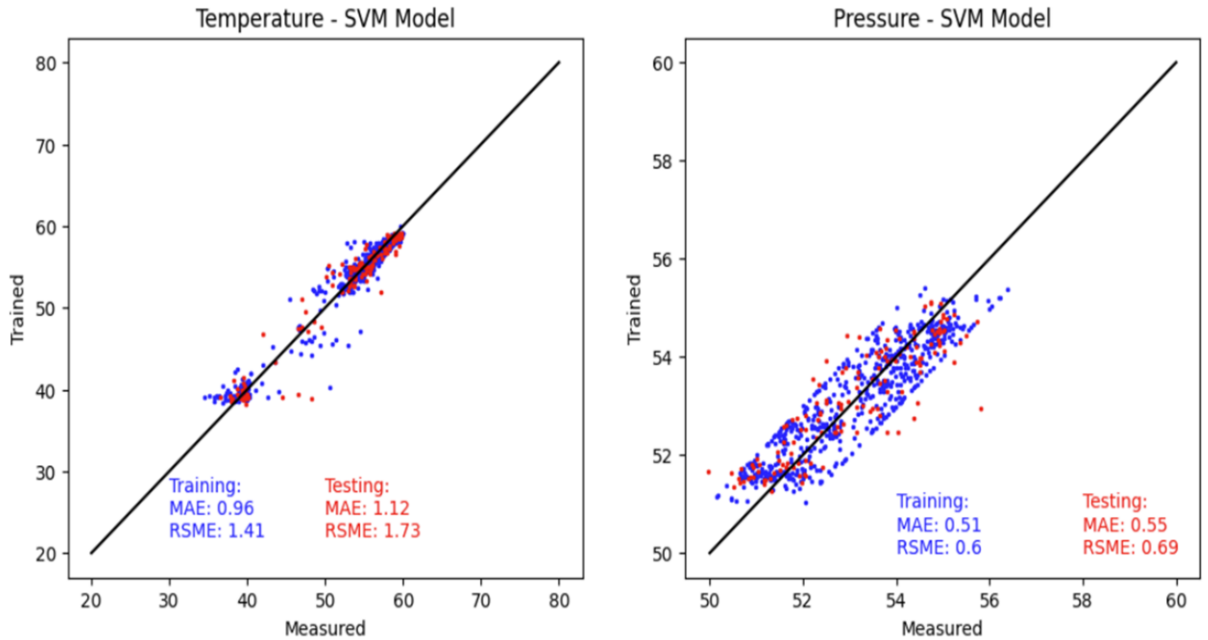

Figure 3. Comparison between predicted temperature and pressure with measured values for training and test subsets using the Support Vector Machine model.

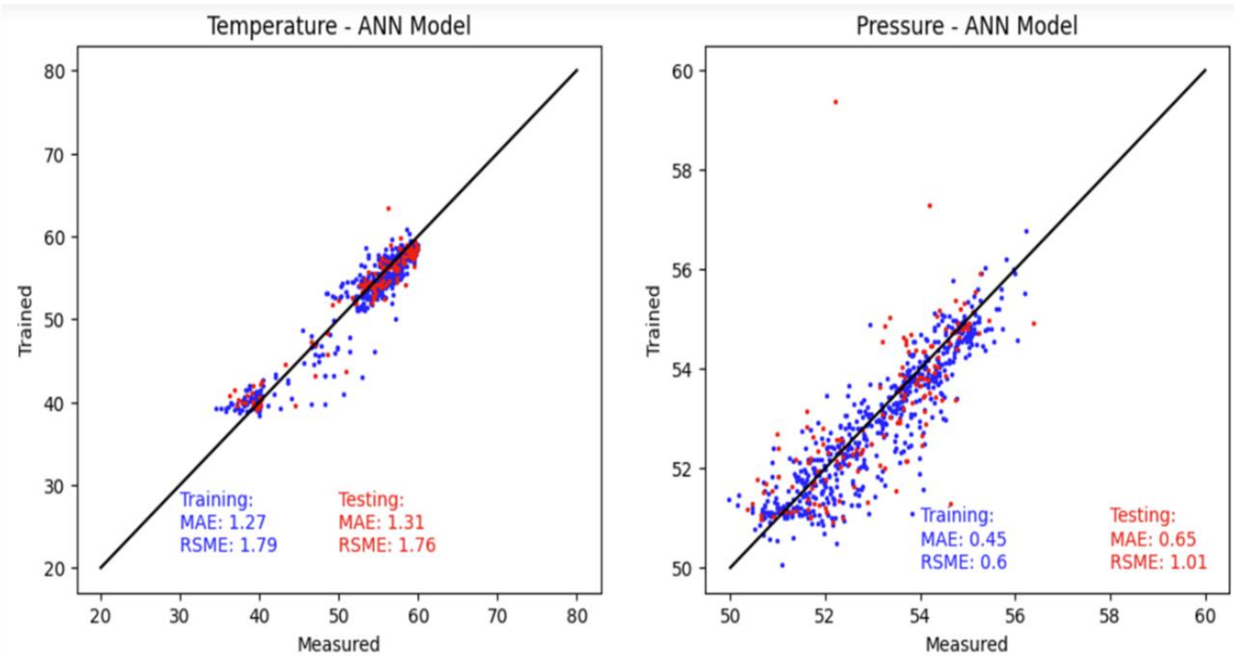

Figure 4. Comparison of temperature parameters and predicted pressure with measured values for training and test subsets using an artificial neural network model.

16 nodes. The hyperparameters of the model are optimized by sensitivity analysis. Regarding temperature parameters, the RMSE and MAE indexes are at 1.79 and 1.27^{0}C for the training set; and 1.76 and 1.31^{0}C for the test set. Regarding the pressure parameters, the RMSE and MAE indexes are at 0.6 and 0.45 barg for the training set; 1.01 and 0.65 barg for the test set (Figure 4).

The XGBoost model is also considered to solve the current problem because it has won consecutively in many machine learning competitions relating to tabular dataset development. The model's learning_rate superparameter is optimized by performing sensitivity analysis. Regarding temperature parameters, the RMSE and MAE indexes are at 0.3 and 0.21 ^{0}C for the training set; 1.2 and 0.73 ^{0}C for the test set. Regarding the pressure parameters, the RMSE and MAE indexes are at 0.3 and 0.22 barg for the training set; 0.57 and 0.41 barg for the test set (Figure 5).

The two metrics RMSE and MAE are compared to evaluate the predictive performance of machine learning models. Then different models will be tested based on aggregate results of these two metrics. The comparison results are shown in Table 1 and Figure 6. From this comparison, it can be found that the XGBoost model gives the best prediction score and is selected to apply to the production operation activities in the field.

4. Evaluate the prediction results of machine learning tool

In order to evaluate the accuracy of the machine learning tool in predicting the temperature and pressure of the internal gathering pipeline at the Hai Thach-Moc Tinh field during the operation production the authors observed and compared the predicted results obtained from the machine learning model to the traditional simulation results (the method combining the LINEST regression analysis tool from Microsoft Excel with the Hysys simulation software) and the actual results measured from the pipeline.

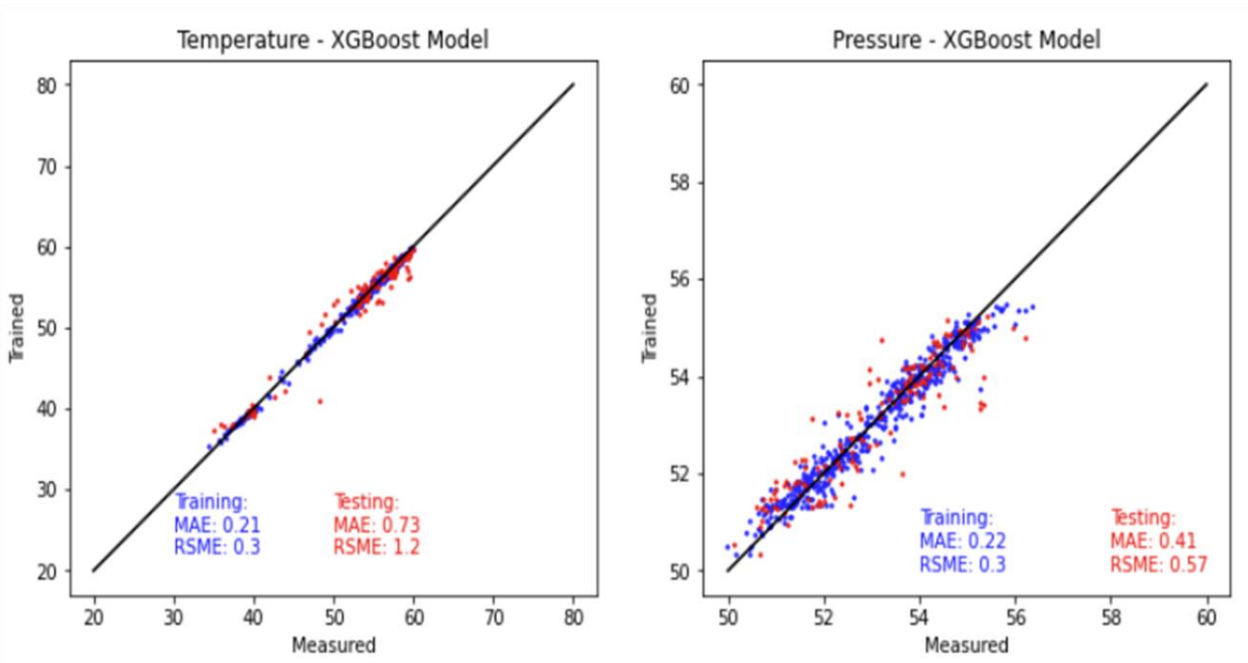

Figure 5. Comparison between predicting pressure and temperature with measured values for training and test subsets using XGBoost model.

Table 1. Comparison of predictive performance between models.

Models	Teperature		Pressure	
	RMSE (^{0}C)	MAE (^{0}C)	RMSE (barg)	MAE (barg)
Linear regression	2.01	1.47	0.84	0.64
Support Vector Machine	1.73	1.12	0.69	0.55
Artificial Neural Network	1.76	1.31	1.01	0.65
XGBoost	1.20	0.73	0.57	0.41

Based on this observation, the authors found that the machine learning model gives prediction results that are quite accurate with the actual parameters of temperature and pressure of the internal gathering pipeline at the Hai Thach-Moc Tinh field. At the same time, the machine learning model gives more accurate predictions than traditional simulation methods (Table 2).

With algorithms built from Machine Learning tools, production engineers or operators working at the central control room of the PQP-HT platform can predict the pressure at the 20th kilometer of the MT-HT pipeline easily. even if Hysys software is not available. The ML algorithm gives accurate results that help to save the working time for the production team during the production process.

For daily operations, the ML algorithm also supports well to operation teams because of the accurate and convenient results obtained from the model. In cases of the gas buyer requires a sudden flow rate reduction, the use of the ML algorithm will give quick results of pressure/temperature prediction of the Hai Thach-Moc Tinh pipeline. That helps the operator decide to check and select opening or stopping any production wells in order to keep the operating parameters of the Hai Thach-Moc Tinh pipeline within the allowable operating limits while reducing the total gas flow based on the request of the gas buyer. Especially, in the rainy season when the gas demand is low, the PQP_HT platform is often maintained at the lowest possible production limit or stopped operating activities to prevent the NCSP's pressure system from over 130 barg when the output gas consumption is reduced. In this case, the application of ML will help the operators to

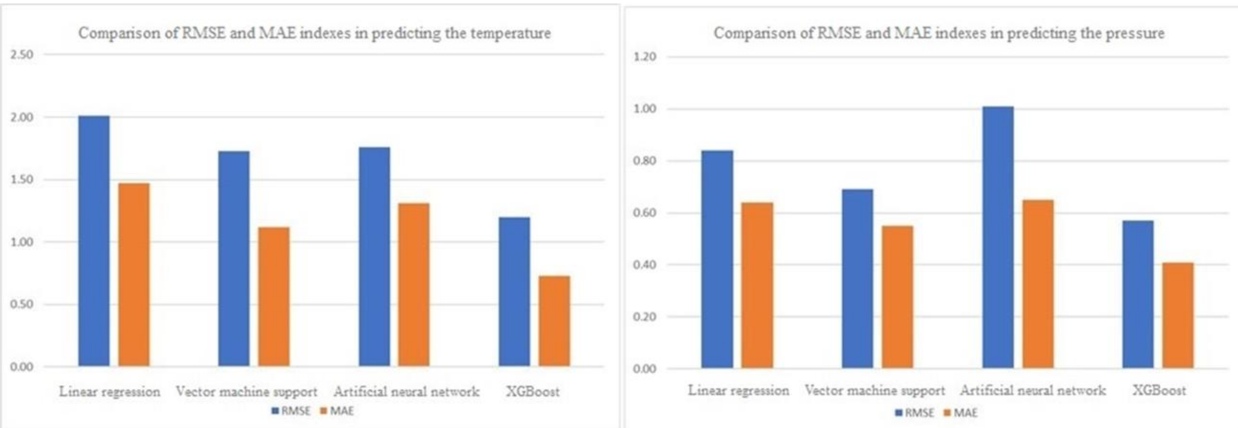

Figure 6. Comparison of predictive performance between models.

Table 2. Comparison of predictive performance between models.

N°	Temperature of the Hai Thach-Moc Tinh pipeline					Pressure of the Hai Thach-Moc Tinh pipeline				
	Field (⁰C)	Machine learning model		Traditional simulation methode		Field (barg)	Machine learning model		Traditional simulation methode	
		Predicting results (⁰C)	Errors	Predicting results (⁰C)	Errors		Predicting results (barg)	Errors	Predicting results (barg)	Errors
1	59.5	59.9	0.4	53.5	6	53.1	52.7	0.4	49.1	4
2	56.7	56.1	-0.6	51.3	5.4	52.2	52.55	-0.35	47.6	4.6
3	36.8	36.0	-0.8	40.1	-3.3	50.9	50.46	0.44	44.8	6.1
4	36.2	36.7	0.5	39.4	-3.2	50.7	50.31	0.39	44.2	6.5
5	50.5	49.9	-0.6	45.2	5.3	53.6	53.8	-0.2	56.6	-3.0

select promptly the appropriate production wells while still keeping the operating conditions of the Hai Thach-Moc Tinh pipeline within safe operating limits without shutdown of production operations.

5. Conclusions

The application of the ML algorithm for gathering pipeline three-phase flows at the Hai Thach-Moc Tinh field has solved the difficulties in time and errors of the current calculation method. The ML algorithm gave accurate prediction results quickly and conveniently. That helps the operator be active in selecting the suitable optimal well opening mode for each gas flow rate level based on the requirements of the buyers.

Predicting the operating conditions of the pipeline is important to the success of daily operations, especially the plan of production and preparation of spare chemicals. Today, there are many simulation softwares in the world to manage pipeline operations. However, the licensing cost of these softwares is expensive and it may not be available due to geopolitical reasons. In this case, ML algorithms can be applied that will help to save human costs (do the calculation based on the traditional assumption/loop method until the minimum errors are given) or the cost of licensing simulation software.

Contributions of authors

Tuan Thanh Nguyen - methodology, writing & editing, data analysis; Thinh Van Nguyen - writing, review & editing, supervision; Truong Hung Trieu - review & editing, supervision.

References

McNeil, R., G. and Lillico, D. R. (2005). Gas Gathering System Modelling the Pipeline Pressure Loss Match, Paper presented at the Canadian International Petroleum Conference, Calgary, Alberta. https://doi.org/10.2118/2005-230.

Jane, O., Tobinson., Briggs, A. K. (2020). Prediction of Pressure Drop in Subsea Pipeline Using Pipesim Software. Industrial Engineering Letters, vol 10, pp 62-81. doi: 10.7176/iel/10-2-06.

John, M., Campbell,. (1992). Gas conditioning and processing. Vol. 2 , Campbell Petroleum Series.

Montgomery, D. C., A. P. lizabeth, G. G. Vining., (2021). Introduction to Linear Regression Analysis, 6th Edition.

John Wiley & Sons.McDonald, G. C. (2009). Ridge regression. WIREs Computational Statistics, vol. 1, no. 1, pp. 93-100. doi: https://doi.org/10.1002/wics.14.

Zhao, P and Yu, B. (2006). On model selection consistency of LASSO. Journal of Machine Learning Research, vol. 7, pp. 2541-2563.

Jia, J and Yu, B. (2010). On Model Selection Consistency of the Elastic Net When p>>n. Statistica Sinica 20, pp. 1-19.

Nguyen-Sy, T., To, Q.d., Ngoc-Vu, M., Duong-Nguyen. T., and Trung- Nguyen, T. (2021). Predicting the electrical conductivity of brine-saturated rocks using machine learning methods. Journal of Applied Geophysics, vol. 184, p. 104238. https://doi.org/10.1016/j.jappgeo.2020.104238.T.

Nguyen-Sy, T., Wakim, J., Dong-To,Q., Ngoc- Vu, M., Duong-Nguyen,T., Trung-Nguyen, T. (2020). Predicting the compressive strength of concrete from its compositions and age using the extreme gradient boosting method. Construction and Building Materials, vol. 260, p119757. https://doi.org/10.1016/j.conbuildmat.2020.119757.