Nafta-Gaz 2024, no. 7, pp. 408–417, DOI: 10.18668/NG.2024.07.02

The application of back propagation neural network for predicting production flow rate of oil wells in Hai Su Trang field, Vietnam

Zastosowanie sieci neuronowych z propagacją wsteczną do prognozowania wskaźnika wydobycia ropy w odwiertach naftowych na złożu Hai Su Trang w Wietnamie

Hong Duong Vu, Tien Hung Nguyen, Minh Hoa Nguyen

Hanoi University of Mining and Geology

ABSTRACT: Oil production flow rate prediction is a critical aspect of oil and gas exploitation operations. Currently, flow rate forecasting is often estimated using theoretical or empirical models. Theoretical models tend to provide predictions with a wide range of errors and require extensive input data. On the other hand, empirical models have limitations due to restricted data. The objective of this article is to develop a correlation for highly accurate forecasting of the production flow rate. In order to achieve this goal, this study applies an artificial neural network (ANN) for flow rate prediction. The backpropagation algorithm and the tansig function are selected in this study as a learning algorithm to forecast flow rate. The study considered 262 datasets collected from six wells in the Hai Su Trang field, Cuu Long basin used in the ANN model, with 70% for training, 15% for testing, and the remaining 15% for validation. This article evaluates the ability of ANN model to predict flow rate with different numbers of neuron. The predicted results obtained from the ANN model with eight neurons and backpropagation algorithm achieved high predictability when compared to empirical methods and multivariate regression model, with a strong correlation coefficient of 0.97 and a low *RMSE* of 32.54 bbl/d. Therefore, the developed ANN models have been shown to be an effective tool in production flow rate forecasting in oilfields.

Key words: artificial neural network, backpropagation algorithm, flow rate prediction, multivariate regression method, gas-lift.

STRESZCZENIE: Prognozowanie wskaźnika wydobycia ropy naftowej jest krytycznym aspektem eksploatacji złóż ropy naftowej i gazu ziemnego. Obecnie prognozowanie wydajności przypływu jest często szacowane przy użyciu modeli teoretycznych lub empirycznych. Modele teoretyczne zazwyczaj generują prognozy z wieloma błędami i wymagają obszernych danych wejściowych. Z drugiej strony, modele empiryczne wykazują ograniczenia ze względu na ograniczoną ilość danych. Celem tego artykułu jest opracowanie korelacji dla bardzo dokładnego prognozowania wskaźnika wydobycia ropy. W tym celu w badaniu zastosowano sztuczną sieć neuronową (ANN). Algorytm wstecznej propagacji i funkcja tansig zostały wybrane w tym badaniu jako algorytm uczenia się do prognozowania wskaźnika wydobycia ropy. W badaniu uwzględniono 262 zestawy danych zebranych z sześciu odwiertów eksploatacyjnych na złożu Hai Su Trang w basenie Cuu Long, które wykorzystano w modelu ANN, z czego 70% do szkolenia, 15% do testowania, a pozostałe 15% do walidacji. W niniejszym artykule oceniono zdolność modelu ANN do przewidywania wskaźnika wydobycia ropy przy różnej liczbie neuronów w warstwie. Przewidywane wyniki uzyskane z modelu ANN z ośmioma neuronami w warstwie i algorytmem wstecznej propagacji wykazały dużą przewidywalność w porównaniu z metodami empirycznymi i wielowymiarowym modelem regresji, z silnym współczynnikiem korelacji wynoszącym 0,97 i niskim *RMSE* wynoszącym 32,54 bbl/d. Dlatego też opracowane modele ANN okazały się skutecznym narzędziem w prognozowaniu wskaźnika wydobycia ropy naftowej ze złóż.

Słowa kluczowe: sztuczna sieć neuronowa, algorytm wstecznej propagacji, przewidywanie wskaźnika wydobycia ropy, metoda regresji wielozmiennej, gazodźwig.

Acronym's explanation

Q – Production flow rate [STB/day]
D_{64} – Choke size [inch]
P_{wh} – Wellhead pressure [kPag]
GLR – Gas/liquid ratio [SCF/STB]
$BS\&W$ – Basic sediment and water [%]
Q_{glift} – Injected gas-lift rate [MMSCFd]
P_{glift} – Injected gas-lift pressure [kPag]
ANN – Artificial Neural Network
R^2 – Correlation coefficient
$RMSE$ – Root Mean Square Error

Corresponding author: M.H. Nguyen, e-mail: *nguyenminhhoa@humg.edu.vn*

Article contributed to the Editor: 30.10.2023. Approved for publication: 18.06.2024.

Introduction

Estimating oil production flow rates is crucial for effective oil and gas operations. It enables planning for drilling, repairs, and interventions required to guarantee and maintain production, as well as to prompt monitoring of the well status. The inability to predict the rate of oil production can lead to challenges in estimating the lifespan and profitability of a hydrocarbon producing facility. Due to numerous production characteristics and field conditions, including wellhead pressure, choke size, gas/oil ratio, water cut, gas injection rate, and gas injection pressure, developing a flow forecasting model for well exploitation is a difficult and demanding task. Various theoretical and practical strategies have been devised to address this complexity.

The first theoretical investigation of multiphase flow through chokes was provided by Tangren et al. in 1949. Their method was only effective when the liquid was in the continuous phase. In accordance with Tangren's methodology, Gilbert (1954) created an empirical correlation using production well-test data and analyzed 268 data sets from Ten Section Kern County Oil fields of California for various choke sizes, ranging from 6/18 to 64/64 inches, to forecast production rates under critical flow conditions.

The equation is given by:

$$Q = \frac{P_{wh} \cdot S^b}{aR^c} \tag{1}$$

where:

Q – critical-flow liquid rate (Stock tank barrel per day),
P_{wh} –wellhead pressure [psia],
S – choke size (1/64 inch),
R –gas/liquid ratio (Standard cubic feet/Stock tank barrel),
a, b and c –empirical constants.

Several studies developed similar relations with different empirical constants for different fields (Baxendell, 1958; Ros, 1960; Achong, 1961). These relations are summarized in Table 1.

Al-Attar and Abdul-Majeed (1988) collected data from over 150 wells from East Baghdad oil field (Iraq). This dataset includes parameters such as gas/liquid ratio, wellhead pressure, choke size, production rate, and API oil gravity. The researchers conducted a sensitivity study to determine the best correlation for estimating production rates. With an average inaccuracy of 6.19%, their results showed that Gilbert's correlation provided a relatively accurate prediction of wellhead rates.

Al-Attar (2008) used 97 datasets from three wells of gas-condensate reservoir in the Middle East with different choke sizes to develop an algorithm for estimating choke performance under subcritical conditions.

Table 1. Summary of empirical constants for different correlations
Tabela 1. Zestawienie stałych empirycznych dla różnych korelacji

Correlation	Empirical constant		
	a	***b***	***c***
Gilbert	0.10000	1.89	0.546
Baxendell	0.10460	1.93	0.546
Ros	0.57400	2.00	0.500
Achong	0.26178	1.88	0.650

Osman and Doka (1990) proposed a correlation for calculating flow rates through chokes using the least square approach for gas condensate reservoirs located in the Middle East.

Beiranvand et al. (2012) developed a new formula for predicting liquid flow rates, incorporating a parameter not included in the Gilbert's correlation: free water, sediment, and emulsion.

$$Q = \frac{P_{wh} \cdot S^b \cdot \left(1 - \frac{BS\,\&\,W}{100}\right)^d}{aR^c} \tag{2}$$

where:

$BS\&W$ – basic sediment and water [%],
a, b, c and d – coefficients calculated based on sufficient data for a specific reservoir with $a = 0.0382$, $b = 2.151$, $c = 0.5154$, and $d = 0.5297$.

Espinoza (2015) developed a modified empirical correlation to estimate and anticipate liquid rates in oilfields with constant water-cut and naturally flowing wells. This method, which builds on modified correlations by Gilbert and Ros, relies on choke size, upstream wellhead pressure, and oil-gas ratio. Additionally, a new empirical coefficient was added to align historical production rate data from the studied field in the Caspian Sea. This coefficient must be recalculated with each new test.

Ghorbani et al. (2018) proposed an equation, which proved to be more successful than other models by other authors. Using 182 datasets from the Reshadat oil field on Lavan Island, Ghorbani et al. (2018) developed and proposed new coefficients, contrasting Beiranvand's method using coefficients a, b, c, and d.

The aforementioned empirical models are limited by the restricted data used in the studies and often require multiple parameter settings when flow conditional change. These models lack accuracy when applied to other fields and are not commonly used.

In order to address the flaws and limitations of both theoretical and empirical correlation methodologies, several researchers have recently employed artificial neural network (ANN) to forecast oil and gas production rates. Table 2 presents previous studies on using ANN to forecast liquid flow rates.

Table 2. Several machine learning applications for forecasting oil flow rate

Tabela 2. Przykłady zastosowań uczenia maszynowego do prognozowania wskaźnika wydobycia ropy naftowej

Authors	Machine learning method	R^2	RMSE/MSE/AAPE
Hasanvand and Berneti (2015)	ANN	0.96	1254
Gorjaei et al. (2015)	Least squares support vector machine-fuzzy logic	0.976	0.8
Al Ajmi et al. (2015)	Fuzzy logic	0.94	1392
Choubineh et al. (2017)	ANN	0.947	1227
Ghorbani et al. (2018)	Genetic algorithm and Excel's solver optimizer	0.997	303.1–562.52
Khan et al. (2018)	Support vector machine (SVM) and ANN	0.96-0.99	2.5618–3.7496
Barjouei et al. (2021)	Deep learning,	0.9969	196
Ibrahim et al. (2021)	Random forests and SVM	0.94–0.98	1.3–1.8
Azim (2022)	ANN	0.96	0.02
Kaleem et al. (2023)	Extra trees, Random Forest, Gradient Boosting, Decision trees…	>0.97	>60.873

In order to estimate surface oil rates, Liu et al. (2011) presented a neural network model based on the backpropagation technique. The neural network comprises ten neurons and one hidden layer. The input neurons include well numbers, well coordinates, cumulative production, the derivative of cumulative production, shut-in time, average distance to surrounding wells, average cumulative production of surrounding wells, and cumulative production days. The output layer has one neuron representing the cumulative production at time $t + 1$. The results demonstrate that the model can forecast the flow rate for a brief period.

Alakeely and Horne (2021) proposed a methodical analysis comparing the use of deep learning algorithms and Gilbert correlation for liquid flow prediction problems. Their research introduced a novel approach and illustrated its applicability to actual field data for predicting well liquid and multiphase constrained production flow rates using wellhead surface measurements.

These studies highlight the superiority of ANN in predicting production flow rates globally. ANN models provide highly accurate results in forecasting production flow rates.

In this study, the authors propose applying an ANN with a backpropagation algorithm to improve the prediction of production flow rates of gas-lift oil wells at Hai Su Trang Field, Cuu Long basin (Vietnam). The forecast results will be compared with experimental equations of other published authors to evaluate the superiority of the ANN model.

Field description

The Hai Su Trang Field is situated in the south-central region of block 15-2/01, nestled within the oil-rich Cuu Long basin, offshore Vietnam. Located approximately 120 km east of Vung Tau, the field features numerous stacked oil reservoirs in the Lower Miocene and the upper segment of the Upper Oligocene.

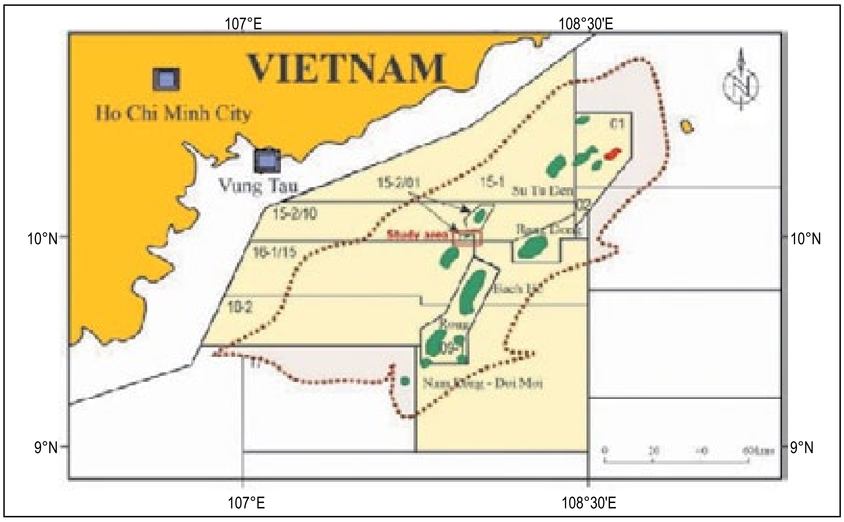

Figure 1. Location of the study area (red rectangle)

Rysunek 1. Lokalizacja obszaru badań (czerwony prostokąt)

The prevalent lithostratigraphic framework primarily comprises a sandstone clastic system, shaped by a fluvio-deltaic complex channel system and lacustrine deposits. The Miocene oil reservoirs exhibit commendable quality, with effective porosity ranging between 14–24% and permeability spanning 10–1000 millidarcies [mD]. Testing outcomes from the HST1 well have verified the Miocene reservoir's exceptional quality. Hai Su Trang field discovered oil in September 2006 from well HST-1X and began operations in mid-2013, with production rates from the first three wells (HST 1P, 2P and 3P) reaching 13,000 barrels per day. After more than 10 years of exploitation, the current production rate is approximately 2,960 barrels per day [bbl/d], and this number is rapidly decreasing. Production wells at Hai Su Trang Field are nearing the final stage with an average water cut of up to 86%. Therefore, it is necessary to assess and forecasts of oil flow rates in the near future to create effective exploitation plans as well as propose solutions to enhance oil recovery and restore production in this field.

Materials and Methodology

The Back Propagation Neural Network (BPNN) is a layered, forward-propagating artificial neural network that uses the backpropagation method for learning. Recognized as a prevalent network structure, the BPNN was chosen in this research to predict production flow rates of gas-lift oil wells at Hai Su Trang Field in the Cuu Long basin (Vietnam). The model was further refined using Genetic Algorithms (GA) to optimize the network setup. The construction process of the model encompasses three primary stages: data gathering, network model development and refinement, and performance evaluation of the refined networks. The entire BPNN framework was developed and fine-tuned using the network tools available in MATLAB software version 2021.

In this study, 262 datasets were collected from six wells in the Hai Su Trang Field. The available parameters included production flow rate (Q), choke size (D_{64}), wellhead pressure (P_{wh}), gas liquid ratio (GLR), basic sediment and water ($BS\&W$), injected gas-lift rate (Q_{glift}), injected gas-lift pressure (P_{glift}), as detailed in Table 3.

The dataset, spanning from 2019 to 2020, was divided into three parts:

- 70% of the data was allocated for training the model;
- 15% was used for testing the model;
- 15% was reserved for validation purposes.

When constructing a model, it is essential to purify the data sample to accurately reflect the issue at hand. Occasionally, a dataset may include extreme values that deviate significantly from the anticipated range and differ from the rest of the data.

Table 3. Data of 6 study wells

Tabela 3. Dane uzyskane z sześciu odwiertów eksploatacyjnych

Parameters		First dataset (2019–2020)	Second dataset (2021)
Number of samples		234	28
Production flow rate [STB/day]	minimum value	92.95	195.04
	maximum value	1059.73	897.35
	mean value	547.02	534.77
	standard deviation	276.60	242.00
Choke size [inch]	minimum value	103.15	136.80
	maximum value	140.52	161.49
	mean value	101.76	151.75
	standard deviation	16.30	10.05
Wellhead pressure [kPag]	minimum value	2381.47	2571.38
	maximum value	3531.01	3144.18
	mean value	2960.38	2903.65
	standard deviation	159.35	152.03
Gas liquid ratio [SCF/STB]	minimum value	143.63	398.27
	maximum value	890.05	890.65
	mean value	524.94	599.91
	standard deviation	141.27	157.19
Basic sediment and water [%]	minimum value	76.68	79.99
	maximum value	98.23	97.74
	mean value	87.58	91.11
	standard deviation	6.48	5.07

cont. Table 3/cd. Tabela 3

Parameters		First dataset (2019–2020)	Second dataset (2021)
Number of samples		234	28
Injected gas-lift rate [MMSCFd]	minimum value	1.48	1.90
	maximum value	4.11	4.27
	mean value	2.86	3.26
	standard deviation	0.54	1.01
Injected gas-lift pressure [kPag]	minimum value	9092.64	9488.96
	maximum value	12594.80	11752.09
	mean value	10984.33	10728.26
	standard deviation	620.76	603.32

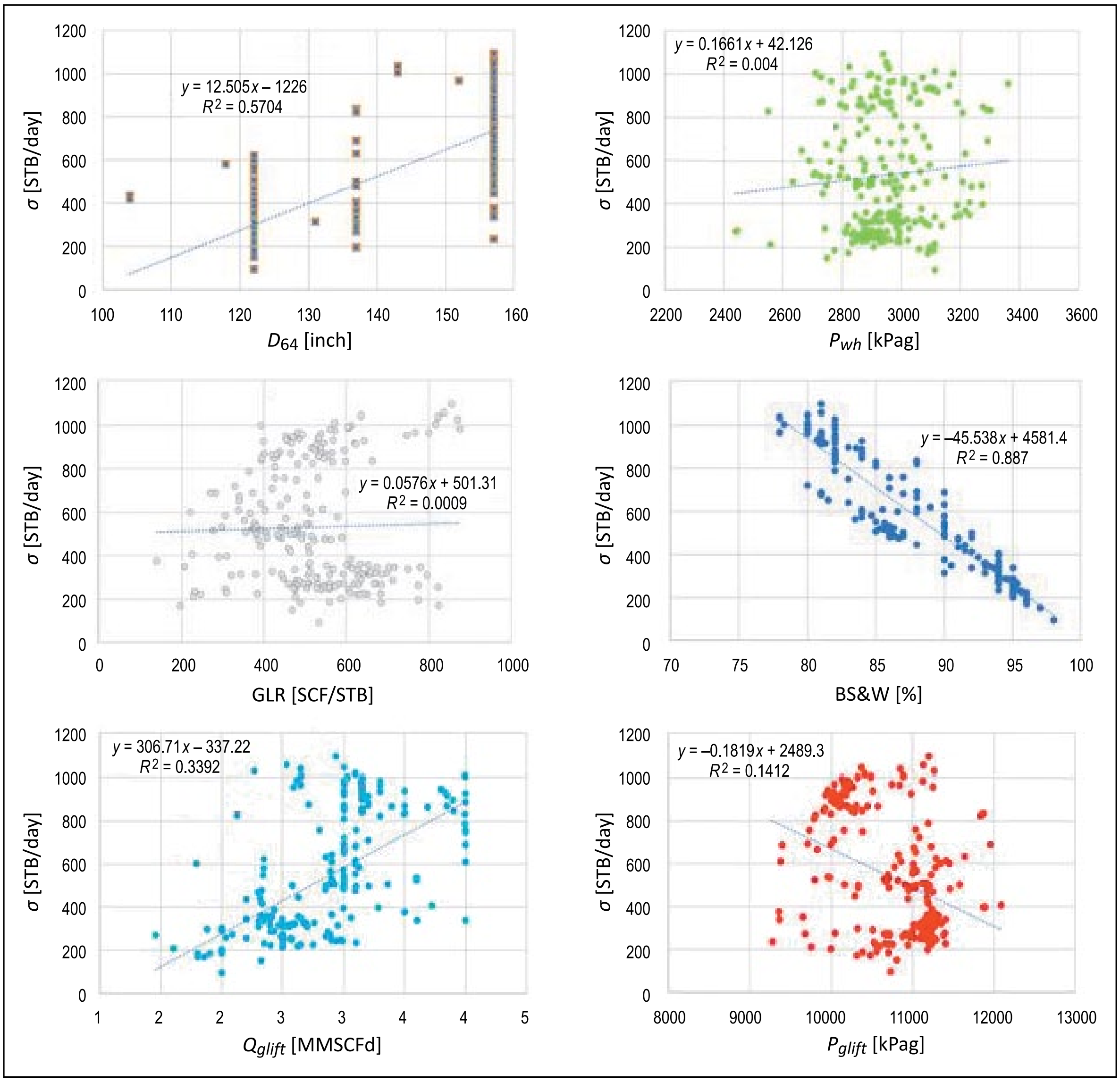

Figure 2. Correlation coefficient R^2 between production parameters and production flow rate
Rysunek 2. Współczynnik korelacji R^2 pomiędzy parametrami wydobycia a wskaźnikiem wydobycia ropy

Such values are known as outliers. Often, the performance of machine learning models and the overall model proficiency can be enhanced by recognizing and potentially eliminating these outliers. An outlier is a data point that stands out markedly from other observations. It could arise from measurement variability or potentially signal an error in the metering equipment. Identifying outliers is critical as they can negatively affect the Artificial Neural Network (ANN) model's efficacy, leading to overfitting or inadequate generalization. To purify the data for training and to confirm the reliability of the ANN model's outcomes, pinpointing outliers is imperative. In this research, the z-score method is employed to detect outliers. The z-score is calculated using the following equation:

$$z = (X - \mu)/\sigma \tag{3}$$

where:
X – individual data point,
μ – average value derived from the participants' results for the analyte,
σ – standard deviation of the dataset.

As per Tripathy et al. (2013), the z-score method is applied to identify outliers. A data point is generally classified as an outlier if its z-score exceeds +3 or falls below -3. Fortunately, in the dataset used in this study, after analysis, only 2 outliers appeared in the wellhead pressure parameter data. The other data points all met the requirements to be used as a training data set for the ANN model.

The precision and computational efficiency of an ANN model's forecasts hinge on the choice of input variables. Theoretical research by the aforementioned scholars suggests that the initial six input factors (referenced in Table 3) influence the production flow rate output. To corroborate this, literature reviews typically lean on the correlation coefficient R^2 to gauge the effect of these inputs on the output (illustrated in Figure 2). This process facilitates the precise determination of inputs for the ANN model's training.

The observed correlation coefficients between the production parameters and the production flow rate are all below 0.6. This implies that for crafting an ANN model with high accuracy in predicting production flow rates, a broad array of input parameters covering diverse aspects is essential. Consequently, all production parameters should be considered as inputs, deemed to contribute uniformly to the predictive model.

ANN model development for forecasting oil production flow rates

Prior to utilizing the specified parameters, it is essential to normalize them within a range from 0 to 1 using the normalization formula:

$$X_{Nor} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{4}$$

where:
X – observed value present in the dataset,
X_{Nor} – normalized data,
$X_{\min}$ – the minimum value in the dataset,
$X_{\max}$ – the maximum value in the dataset.

An ANN is a computational construct that mimics the signal transmission of biological neurons. It is composed of numerous neural units linked together to process data. A standard ANN typically includes three layers: Input, Hidden, and Output:

- Input layer: This is where the information enters the ANN. The input node categorizes and analyzes the data before forwarding it to the subsequent layer;
- Hidden layer: Data moves from the input layer to the hidden layer(s). An ANN may contain one or several hidden layers, each processing the data received from the preceding layer before sending it on;
- Output layer: This layer delivers the final processed data results from the ANN. It can consist of one or multiple nodes.

For predicting the production flow rate in gas-lift wells, this study's ANN model employs the backpropagation algorithm (Marfo and Kporxah, 2020). The production parameters are inputs for network training, with the production flow rate as the output. The neural network operates through two phases: forward and backward propagation. The forward phase transmits signals through neurons to compute output targets, while the backward phase generates an error vector between the actual and target values. This error is used to adjust the network's weighted connections. The process continues until the error reaches a set minimum threshold or a certain number of cycles is completed. Consequently, the neural network incrementally adjusts its output to more closely match the intended target output. This process of convergence is a fundamental aspect of the network's learning and adaptation capabilities.

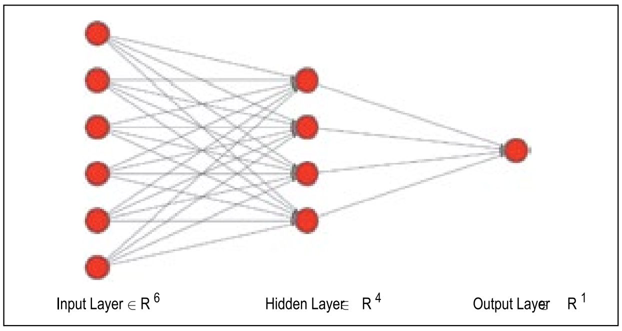

Figure 3. Structure of ANN
Rysunek 3. Struktura ANN

The quantity of neurons within the hidden layer is a crucial factor influencing both the precision and computation duration for ANN model predictions. Selecting the appropriate number of neurons is vital for generating precise forecasts that align with expected outcomes. To avert the risk of overfitting, which can occur with too many neurons, careful consideration is required. The findings from various models with different neuron counts in the hidden layer are presented in Table 4. An analysis of

Table 4. Summary of the results of R^2 and RMSE from different ANN models

Tabela 4. Zestawienie wyników R^2 i RMSE z różnych modeli ANN

No.	R^2			RMSE		
	training	validation	testing	training	validation	testing
4	0.95	0.95	0.92	52.64	67.41	62.32
5	0.94	0.96	0.95	48.16	58.07	64.46
6	0.97	0.93	0.94	43.37	53.38	58.52
7	0.95	0.94	0.95	32.86	45.89	51.95
8	0.97	0.96	0.96	32.54	43.56	45.65
9	0.95	0.96	0.95	32.97	46.46	46.08
10	0.97	0.96	0.93	31.46	45.97	46.96

the correlation coefficient (R^2) and Root Mean Square Error (RMSE) from Table 4 and Figure 4 indicates that increasing the neuron count in the hidden layer from 4 to 7 enhances the model's accuracy. However, further increasing the count from 7 to 10 does not yield significant improvements and may even lead to reduced accuracy, as seen in the 10-neuron model. Consequently, the researchers recommend using 8 neurons in the hidden layer for the ANN that predicts production flow rates in gas-lift wells at the Hai Su Trang field, as this number simplifies the model while preserving high forecasting accuracy.

The correlation coefficients, which compare the predicted production flow rates from the multivariate regression method and the ANN model against the actual production flow rate values, are depicted in Figure 5. Figure 5 illustrates that the ANN model's predictions are more accurate than those of the multivariate regression model (correlation coefficient R^2 of the ANN model is 0.9646, while the multivariate regression model has an R^2 of approximately 0.79).

While the ANN model yields highly accurate results with the 2019–2020 dataset, the authors opted to further validate

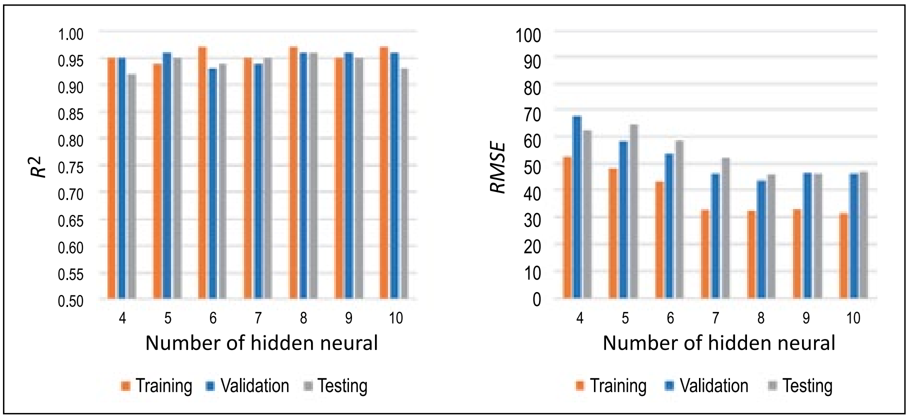

Figure 4. The results of R^2 and RMSE from different ANN models

Rysunek 4. Wyniki R^2 i RMSE z różnych modeli ANN

Results and discussions

To evaluate the effectiveness of the developed ANN models, the researchers conducted a comparison of the models' predictive accuracy against a conventional approach, namely the multivariate regression method, using the same dataset. This traditional method is articulated through the following equation 5:

$$Q_l = x_1 D_{64} + x_2 P_{wh} + x_3 GLR + x_4 BS\&W + x_5 Q_{Glift} + x_6 P_{Annulus} + y \tag{5}$$

where: x_1, x_2, x_3, x_4, x_5, x_6 and y – empirical parameters (Table 5).

Table 5. Coefficients of Equation 5

Tabela 5. Współczynniki równania 5

Parameters	Coefficients
Intercept (y)	1475.1878
x_1	4.2219
x_2	0.1416
x_3	0.3083
x_4	–25.5412
x_5	67.6298
x_6	–0.0027

its efficacy and superiority for future predictions or different wells. To this end, they applied the ANN model to a second dataset from 2021, comprising 28 data points, as shown in Figures 6 and 7.

Despite the fact that the production histories of wells HST 5 and HST 6 in 2021 were not part of the training dataset, the history matching between the ANN model's predictions and the actual data shows a strong correlation.

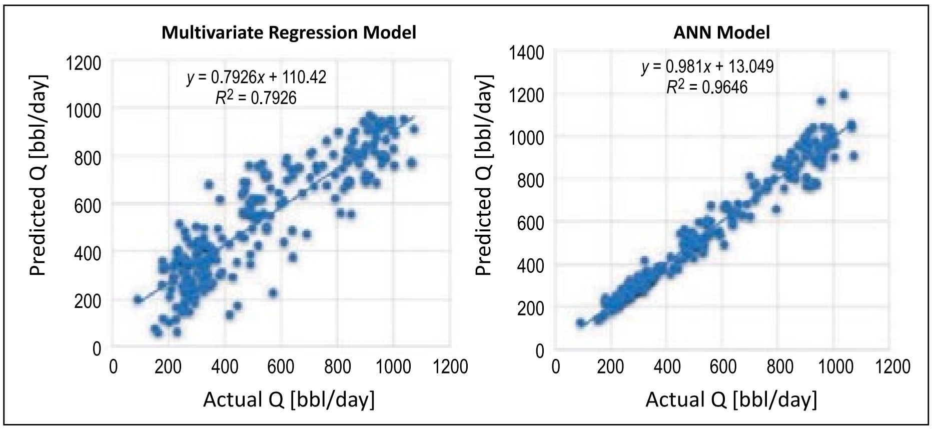

Figure 5. Cross-plot of predicted production flow rate and actual value

Rysunek 5. Wykres krzyżowy przewidywanego wskaźnika wydobycia i wartości rzeczywistej

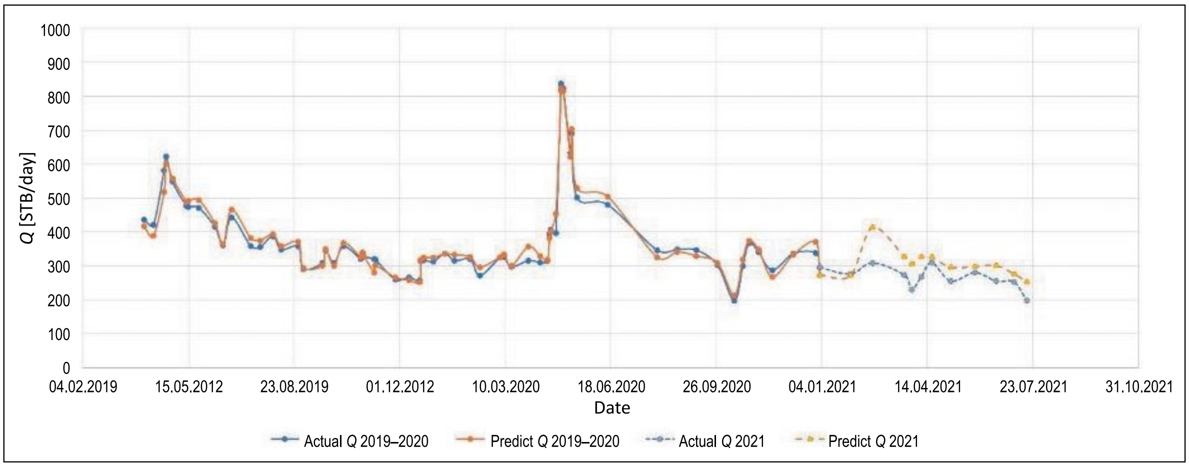

Figure 6. Comparison of actual production flow rate and predicted value from ANN model for HST5 well

Rysunek 6. Porównanie rzeczywistego wskaźnika wydobycia i wartości przewidywanej z modelu ANN dla odwiertu HST5

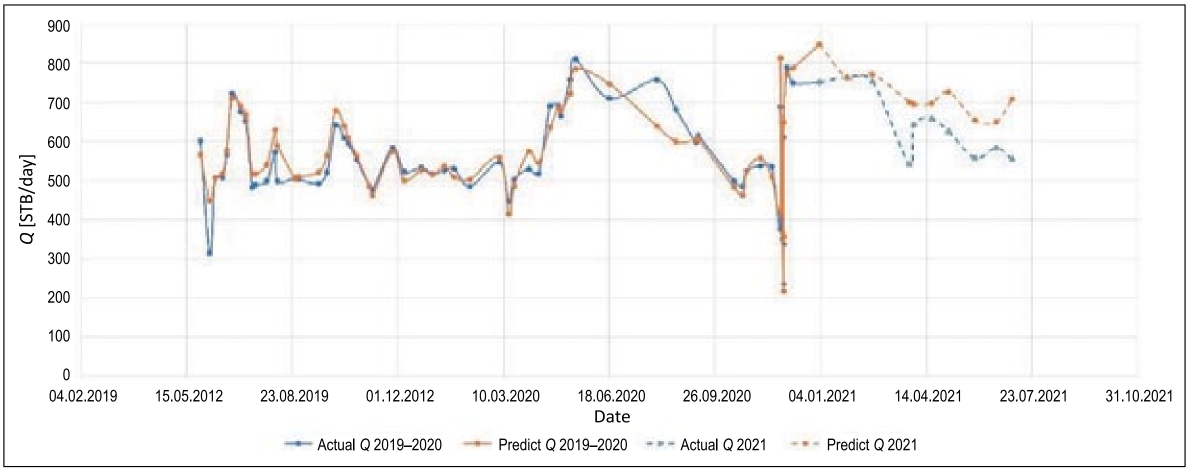

Figure 7. Comparison of actual production flow rate and predicted value from ANN model for HST6 well

Rysunek 7. Porównanie rzeczywistego wskaźnika wydobycia i wartości przewidywanej z modelu ANN dla odwiertu HST6

The predicted future production flow rates also seem to be in close agreement with the real data, underscoring the ANN model's potential as a robust tool for predicting oil production flow rates beyond the initial training set.

Conclusions

This study explored the use of an ANN method to forecast the oil production flow rate of wells in Hai Su Trang Field, Vietnam. The key takeaways are:

1. The constructed ANN model, which utilizes a backpropagation algorithm and includes 8 neurons in the hidden layer, delivers highly accurate production flow rate predictions that closely mirror actual data (with an R^2 of 97% and a low RMSE of 32.54 bbl/d).
2. The ANN model has shown promise as an effective instrument for forecasting production flow rates in oilfields. It not only delivers precise predictions for the oil production flow rate of each well but also captures the dynamic trends of the production flow rate over time. This confirms the ANN model's accurate depiction of the correlation between production variables and production flow rate, making it a valuable asset for predicting production flow rates in oilfields.
3. Predicting the oil production flow rate at the HST field is essential for monitoring well conditions and formulating prompt intervention strategies to sustain and secure production levels. Additionally, the ANN model aids in identifying suitable production parameters to regulate flow, preserve output, and enhance oil recovery processes.
4. For further refinement of the ANN model's accuracy, it is necessary to incorporate additional datasets from prior years and to update the model with new data for ongoing training.

Acknowledgements

This research is supported by Hanoi University of Mining and Geology under grant number T24-11.

Literature

Achong I.B., 1961. Revised bean performance formula for lake Maracaibo wells. *Shell Internal Report*.

Al Ajmi M.D., Alarifi S.A., Mahsoon A.H., 2015. In improving multiphase choke performance prediction and well production test validation using artificial intelligence: a new milestone. *SPE Digital Energy Conference and Exhibition*. DOI: 10.2118/173394-MS.

Alakeely A.A., Horne R.N., 2021. Application of deep learning methods to estimate multiphase flow rate in producing wells using surface measurements. *Journal of Petroleum Science and Engineering*, 205: 108936. DOI: 10.1016/j.petrol.2021.108936.

Al-Attar H.H., 2008. Performance of wellhead chokes during subcritical flow of gas condensates. *Journal of Petroleum Science and Engineering*, 60(3–4): 205–212. DOI: 10.1016/j.petrol.2007.08.001

Al-Attar H.H., Abdul-Majeed G.H., 1988. Revised bean performance equation for East Baghdad oil wells. *SPE Production &. Engineering*, 3(01): 127–131. DOI: 10.2118/13742-PA.

Azim R.A., 2022. A new correlation for calculating wellhead oil flow rate using artificial neural network. *Artificial Intelligence in Geosciences*, 3(3). DOI: 10.1016/j.aiig.2022.04.001.

Barjouei H.S., Ghorbani H., Mohamadian N., Wood D.A., Davoodi S., Moghadasi J., Saberi H., 2021. Prediction performance advantages of deep machine learning algorithms for two-phase flow rates through wellhead chokes. *Journal of Petroleum Exploration and Production Technology*, 11: 1233–1261. DOI: 10.1007/s13202-021-01087-4.

Baxendell P.B., 1958. Producing wells on casing flow-an analysis of flowing pressure gradients. *Petroleum Transactions*, 213(01): 202–206. DOI: 10.2118/983-G.

Beiranvand M.S., Mohammadmoradi P., Aminshahidy B., Fazebdolabadi, B., Aghahoseini S., 2012. New multiphase choke correlations for a high flow rate Iranian oil field. *Mechanical Sciences*, 3(1): 43–47. DOI: 10.5194/ms-3-43-2012.

Choubineh A., Ghorbani H., Wood D.A., Moosavi S.R., Khalafi E., Sadatshojaei E., 2017. Improved predictions of wellhead choke liquid critical-flow rates: Modelling based on hybrid neural network training learning-based optimization. *Fuel*, 207: 547–560. DOI: 10.1016/j.fuel.2017.06.131.

Espinoza R., 2015. In digital oil field powered with new empirical equations for oil rate prediction. *SPE Middle East Intelligent Oil and Gas Conference and Exhibition*. DOI: 10.2118/176750-MS.

Ghorbani H., Wood D.A., Moghadasi J., Choubineh A., Abdizadeh P., Mohamadian N., 2018. Predicting liquid flow-rate performance through wellhead chokes with genetic and solver optimizers: an oil field case study. *Journal of Petroleum Exploration and Production Technology*, 9(3): 1–19. DOI:10.1007/s13202-018-0532-6.

Gilbert W.E., 1954. Flowing and gas-lift well performance. *API Drilling and Production Practice*, 13: 126–157.

Gorjaei R.G., Songolzadeh R., Torkaman M., Safari M., Zargar G., 2015. A novel PSO-LSSVM model for predicting liquid rate of two-phase flow through wellhead chokes. *Journal of Natural Gas Science and Engineering*, 24: 228–237. DOI: 10.1016/j.jngse.2015.03.013.

Hasanvand H., Berneti S.M., 2015. Predicting oil flow rate due to multiphase flow meter by using an Artifical neural network. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, 37(8): 840–845. DOI: 10.1080/15567036.2011.590865.

Ibrahim A.F., Al-Dhaif R, Elkatatny S., Al Shehri D., 2021. Applications of artificial intelligence to predict oil rate for high gas–oil ratio and water-cut wells. *ACS Omega*, 6(30): 19484–19493. DOI: 10.1021/acsomega.1c01676.

Kaleem W., Tewari S., Fogat M., Martyushev D.A., 2023. A hybrid machine learning approach-based study of production forecasting and factor fluencing the multiphase flow through surfaces chokes. *Petroleum*. DOI: 10.1016/j.petlm.2023.06.001.

Liu Y.H., Zhang Z.B., Chen D.Z., Tang Z.P., Hu S.Q., 2011. Analysis of liquid-carrying capacity in high GLR gas wells with downhole choke. *Xinjiang petroleum geology*, 32(5): p. 495.

Khan M.R., Tariq Z., Abdulraheem A., 2018. In utilizing state of the art computational intelligence to estimate oil flow rate in artificial lift wells. *SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition*: SPE-192321-MS. DOI: 10.2118/192321-MS.

Marfo S.A., Kporxah C., 2020. Predicting oil production flow rate using artificial neural network and decline curve analytical methods. *Proceedings of 6th UMaT Biennial International Mining and Mineral Conference, Tarkwa, Ghana*, 43–50.

Osman M.E., Dokla M.E., 1990. Gas condensate flow through chokes. *European Petroleum Conference, The Hague, Netherlands*, SPE-20988-MS.

Ros N.C.J., 1960. An analysis of critical simultaneous gas/liquid flow through a restriction and its application to flow metering. *Applied Scientific Research*, 9: 374–389. DOI: 10.1007/BF00382215

Tangren R.F., Dodge C.H., Seifert H.S., 1949. Compressibility effects in two-phase flow. *Journal of Applied Physics*, 20(7): 637–645. DOI: 10.1063/1.1698449.

Tripathy S.S., Saxena R.K., Gupta P.K., 2013. Comparison of statistical methods for outlier detection in proficiency testing data on analysis of lead in aqueous solution. *American Journal of Theoretical and Applied Statistics*, vol. 2(6): 233–242. DOI: 10.11648/j.ajtas.20130206.21.

Hong Duong VU, M.Sc.
Lecturer at the Department of Geophysics
Hanoi University of Mining and Geology
18 Vien Street, Duc Thang Ward, Bac Tu Liem District, Hanoi, Vietnam
E-mail: *vuhongduong@humg.edu.vn*

Tien Hung NGUYEN, Ph.D. Candidate
Lecturer at the Department of Drilling and Production
Hanoi University of Mining and Geology
18 Vien Street, Duc Thang Ward, Bac Tu Liem District, Hanoi, Vietnam
E-mail: *khoankhaithac@humg.edu.vn*

Minh Hoa NGUYEN, Ph.D. Candidate
Lecturer at the Department of Petroleum Geology
Hanoi University of Mining and Geology
18 Vien Street, Duc Thang Ward, Bac Tu Liem District, Hanoi, Vietnam
E-mail: *nguyenminhhoa@humg.edu.vn*