International Journal on Advanced Science Engineering Information Technology

Vol.13 (2023) No. 6
ISSN: 2088-5334

Application of Artificial Intelligence in Predicting Oil Production Based on Water Injection Rate

Diyah Rosiani [a,*], Muhamad Gibral Walay [a], Pradini Rahalintar [a], Arya Dwi Candra [a], Akhmad Sofyan [a,b], Yesaya Arison Haratua [c]

[a] *Department of Oil and Gas Production Engineering, Politeknik Energi dan Mineral Akamigas, Cepu 58315, Indonesia*
[b] *Department of Mineralogy, Geochemistry and Petrology, Faculty of Earth Science, University of Szeged, Szeged, 6722, Hungary*
[c] *PT Pertamina Hulu Rokan Zona 4, Jl. Jend. Sudirman No. 3, South Sumatra, 31122, Indonesia*
*Corresponding author: *diyahrosiani@gmail.com*

Abstract—**The utilization of artificial intelligence (AI) has become imperative across various domains, including the oil and gas industry, which covers several fields, including reservoirs, drilling, and production. In oil and gas production, conventional methods, such as reservoir simulation, are used to predict the oil production rate. This simulation requires comprehensive data, so each process step takes a long time and is expensive. AI is urgently needed and can be a solution in this case. This research aims to apply AI techniques to forecast oil production rates based on water injection rates from two injection wells. Three wells are connected with a direct line drive pattern. Three different AI methods were applied, including multiple linear polynomial regression (PR), multiple linear regression (MLR), and artificial neural networks (ANN) in constructing oil production rate prediction models. Actual field data of 1180 data are used, including water injection rate data from two injection wells and oil production history data from one production well. The dataset has been split randomly into 80% training and 20% allocated for testing subsets. The training data is used to build predictive models, while the testing data is used to validate model performance. Comparative analysis selects the model with the lowest root mean square error (RMSE) and the highest R^2 test value. Results demonstrate that the ANN model achieves the smallest Root Mean Square Error (RMSE) of 0.142 and the highest R^2 test value of 16.2%, outperforming the PR and MLR methods. The ANN prediction model provides a rapid and efficient approach to estimating oil production rates.**

Keywords—**Artificial intelligence; multiple linear regression; polynomial regression; artificial neural network; oil production.**

Manuscript received 15 Aug. 2023; revised 30 Oct. 2023; accepted 17 Nov. 2023. Date of publication 31 Dec. 2023.

I. INTRODUCTION

AI has become a hot issue in its application in the current era. Nearly every field adopts this method, making it crucial [1]. Accelerated processing time and precise predictions are the prominent outcomes popularized by the utilization of AI compared to conventional methods [2]-[5]. The oil and gas sector, being a capital-intensive sector with technological advancements and inherent risks, necessitates the implementation of AI in all its domains. Various studies have demonstrated the application of AI in the petroleum industry to expedite operational processes [6]-[8].

One such application of AI in the drilling field was conducted by Syah et al. [9]. Several AI methods utilized include PSO-ANFIS (particle swarm optimization-adaptive neuro-fuzzy inference system), ANFIS (adaptive neuro-fuzzy inference system), RBF (radial basis function) algorithm, and LSSVM-GA (least square support vector machine-genetics algorithm). These techniques are utilized for the purpose of forecasting drilling fluid density and choosing the most suitable model. The RBF model emerged as the most effective method in predicting the fluid, exhibiting validity and accuracy consistent with experimental data. Another researcher aimed to swiftly predict cutting settling velocity by developing a model using artificial neural networks (ANN). This model yielded accurate results, proving its practicality and potential implementation in the field [10].

In the reservoir domain, Artificial intelligence has been harnessed to attain remarkably precise forecasts of the compressibility factor (Z factor) in gas condensate reservoirs, specifically for two-phase conditions. [11]. The best AI model was developed using a feedforward neural network and Bayesian-Regularization algorithms. Rosiani et al. [12] employed principal component analysis and multivariate quality control to construct a screening model for CO_2-EOR, considering the interdependency of nine screening parameters. The resulting screening model expedited decision-making processes and was suitable for implementation in real-world field settings.

In the production domain, reservoir simulation is one of the current conventional methods used to predict oil production rates, which requires a large amount of data and significant costs. Additionally, it involves a lengthy execution time for running each of the simulation steps [13],[14]. AI implementation can help solve this problem.

Ghorbani et al. [15] have utilized AI for oil flow rate prediction by applying five machine learning methods: Radial Basis Function, Adaptive Neuro-Fuzzy Inference System, Multilayer Perceptron, Least Squares Support Vector Machine, and Gene Expression Programming. Multilayer Perceptron yielded a more accurate model for predicting oil flow rate based on 830 training data points. Another study used a high-level neural network (HONN) has been used in predicting cumulative oil production. This study results in good model forecasting with high accuracy [16]. ANN has proven useful in predicting hydrocarbon, liquid flow rates, and water from artificially lifted oil wells using electric submersible pumps (ESP) in Egypt. Four ANN models were developed using input parameters such as wellhead pressure, fluid properties, variable speed drive (VSD) sensor parameters, and downhole ESP sensor measurements. The ANN models demonstrated simplicity, efficiency, and cost-effectiveness [17].

In addition to ANN, regression methods are well-known and easy-to-apply prediction methods. Popular regression methods for multivariate data are MLR and PR [18]. Previous studies have used MLR as an effective prediction method for predicting the performance and quality of biodiesel [19]. MLR also performs well in predicting crude oil pricing variability [20]. Ajona et al. [21] used MLR and PR to develop a model for the sustainable biodegradation process of crude oil. MLR and PR are also used to model the cetane number of biodiesels [22].

This study aims to apply AI to predict the oil production rate in a single active well influenced by the water injection rates from two injection wells. Many previous studies have yet to perform this particular study. Actual field data, totaling 1180, are used to construct a predictive model with two input variables (water injection rate from well 1 and water injection rate from well 2) and one output variable (oil production rate). The prediction methods are models that can cover linear and nonlinear relationships between input and output variables. This study uses ANN as a prediction method for nonlinear relationships and is widely applied in the petroleum industry. The MLR and PR methods are also used as practical and fast statistical prediction methods. MLR can establish a linear relationship between response variables and predictors. Meanwhile, PR is employed as a prediction method for data with nonlinear relationships. Each of these three methods has its limitations. Therefore, they are applied, and the method that yields a predictive model more aligned with actual conditions is selected.

II. Materials and Method

The data utilized in this study is the actual field data of oil sites located in South Sumatra, Indonesia. The field is operated by PT Pertamina Hulu Rokan Zone 4. Data collected from PT Pertamina Hulu Rokan Zone 4 daily production report 2021. The average daily oil production from the field is 7799 barrels of oil per day (BOPD), with a total gas production of 136.9 million standard cubic feet per day (MMSCFD) and a total water production of 65807 barrels of water per day (BWPD). The water produced from the well is reinjected into the reservoir through injection wells. When the water is injected into the reservoir, it provides two advantages: pressure maintenance and sweep improvement. These events collectively enhance oil recovery. The production wells used in this study employ an artificial lift method known as a gas lift.

The flow diagram of this study can be seen in Figure 1. This study used actual data in the form of daily historical production data from a single production well and historical water injection rate data from two injection wells. These three wells are interconnected using a direct line drive pattern. A total of 1180 data points were gathered, spanning from February 2020 to May 2023. The independent variables consist of the rate of water injection from injection well 1 (x_1) and the water injection rate from injection well 2 (x_2). The dependent variable is the oil production rate from the single production well (y). An overview of the raw data is presented in Table 1.

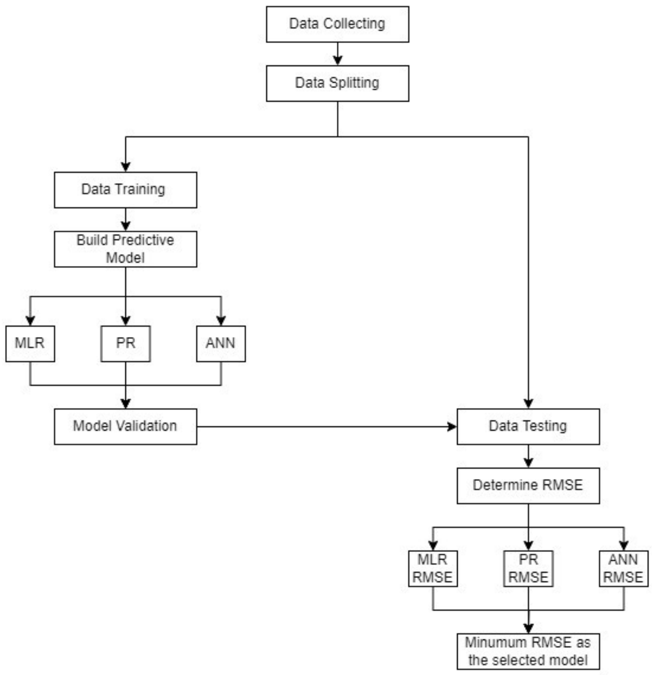


Fig. 1 Study Workflow

Three methods were employed to predict the rate of oil production in the production well. The MLR method was used to model linear trends, while the PR and ANN methods were suitable for constructing prediction models for nonlinear cases. The prediction models of the three methods were developed using the Python programming language. A dataset of 1180 observations was divided randomly into 80% allocated for training and 20% for testing purposes.

TABLE I

RAW DATA

Daily Data	Q_Oil, BOPD (y)	Q_Inj1, BWPD (x_1)	Q_Inj2, BWPD (x_2)
1	35	2015	2765
2	52	2043	2804
3	52	1984	2723
…	…	…	…
1180	35	1684	1286

The prediction models were constructed based on the training data. These three methods resulted in three distinct prediction models. Subsequently, the models were compared, and the method that produced a highly accurate prediction model was selected based on the testing data. The accuracy of the prediction was evaluated using RMSE in equation (1) [23] and the R^2 value in equation (2) [24]. A lower RMSE value indicates better performance of the model in prediction. A higher R^2 value approaching 1.0 indicates a better model fit.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2} \tag{1}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i-\hat{y})^2}{\frac{1}{N}\sum_{i=1}^{N}(y_i-\bar{y})^2} \tag{2}$$

A. Multiple Linear Regression

MLR is applied to establish a linear connection between multiple independent variables (x1, x2, ..., xn) and a dependent variable (y). The general form of MLR is given in equation (3) [25]-[28].

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \tag{3}$$

In this context, y represents the dependent variable, x_1, x_2, ..., x_n are the independent variables, β_0, β_1, … β_n are regression coefficients, and n represents the number of variables [29]. Least-squares methods estimated the value of the parameters.

B. Polynomial Regression

Polynomial regression is a specific instance of multiple regression where there is only a single independent variable, denoted as X. The model of polynomial regression with one independent variable can be mathematically represented by the following equation (4) [30].

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \cdots + \beta_k x_i^g + e_i \tag{4}$$

for i = 1, 2,…n, where g represents the polynomial degree. This corresponds to the order of the model. Effectively, it is equivalent to having multiple models with $X_1 = X$, $X_2 = X^2$, $X_3 = X^3$, and so on [31]. The coefficients of the polynomial regression, β, as shown in equation (9), are operated in matrix form as present in equations (5), (6), (7), and (8) [32].

$$\mathbf{X}(g) = \begin{bmatrix} n & \sum x_i & \ldots & \sum x_i^g \\ \sum x_i & \sum x_i^2 & \ldots & \sum x_i^{g+1} \\ \vdots & \vdots & & \vdots \\ \sum x_i^k & \sum x_i^{k+1} & \ldots & \sum x_i^{2g} \end{bmatrix} \tag{5}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \tag{6}$$

$$\mathbf{Y}(g) = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \vdots \\ \sum y_i x_i^g \end{bmatrix} \tag{7}$$

PR equation in matrix form can be written as:

$$\mathbf{X}(g)\boldsymbol{\beta} = \mathbf{Y}(g) \tag{8}$$

Therefore,

$$\boldsymbol{\beta} = \mathbf{X}(g)^{-1}\mathbf{Y}(g) \tag{9}$$

C. Artificial Neural Network

ANN is an information processing system that demonstrates performance characteristics similar to biological neural networks [33], [34]. ANN consists of three elements: network architecture, weighting factor, and activation function. Figure 2, as a single-layer ANN consists of one input layer and one output layer. The input layer receives signals and comprises several neurons connected by weights (w), summed and fed to a nonlinear activation function to the output layer in a forward flow.

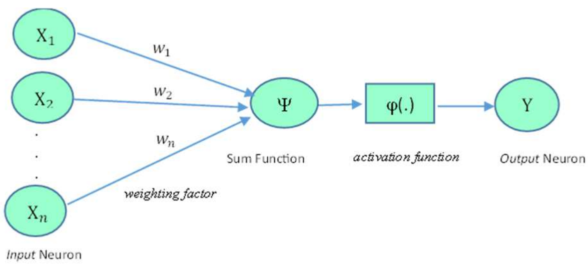


Fig. 2 Single-layer ANN

ANNs are predictive models that, learning from data, can identify complex nonlinear patterns without understanding how the variables are related. ANN can identify different patterns in nonlinear systems that cannot be resolved with general statistical methods. Different patterns adjusted parameters, namely synaptic weight and bias, during training [35], [36].

Typical neural networks are multi-layered systems with a sole input layer, one or more concealed layers, and a sole output layer. Most applications performed by ANN have been trained using supervised training techniques. During the training phase, the input layer receives input signals, which are processed through the hidden layers until they reach the output layer. This process is known as feedforward. The resulting output is compared to the desired output, allowing for the error calculation. The error information is then propagated backward (backpropagation), and the weights are adjusted to control the network. Adjust weight by minimizing the error function as a solution to learning problems. The training process is repeated until the desired level of accuracy is achieved. Subsequently, the network undergoes a testing phase to evaluate its designed capabilities. The trained network is tested using data independent of the training data [37]-[42].

In this study, the backpropagation method of ANN was applied. The ANN utilized was a multilayer network in Figure 3. The input layer consisted of two neurons (x_1 and x_2). Subsequently, the hidden layer, which could have more than one layer, was determined to achieve the optimal number of hidden neurons, thus ensuring optimal network performance. The layer output consisted of a single neuron (y). A nonlinear activation function, specifically the hyperbolic tangent (tanh) function in equation (10) [33], was utilized in the concealed layer, while a linear function was employed in the output layer.

$$f(\Psi) = \frac{e^{\Psi} - e^{-\Psi}}{e^{\Psi} + e^{-\Psi}} \tag{10}$$

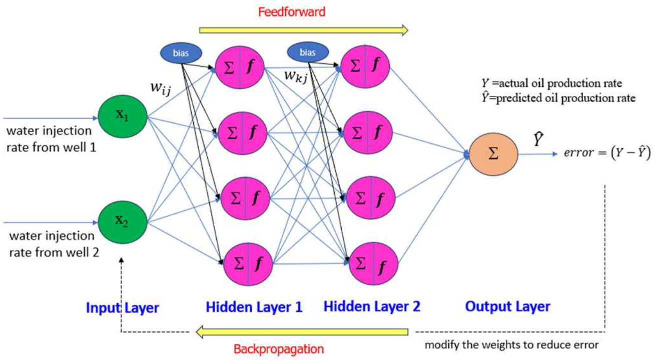


Fig. 3 Multilayer ANN structure

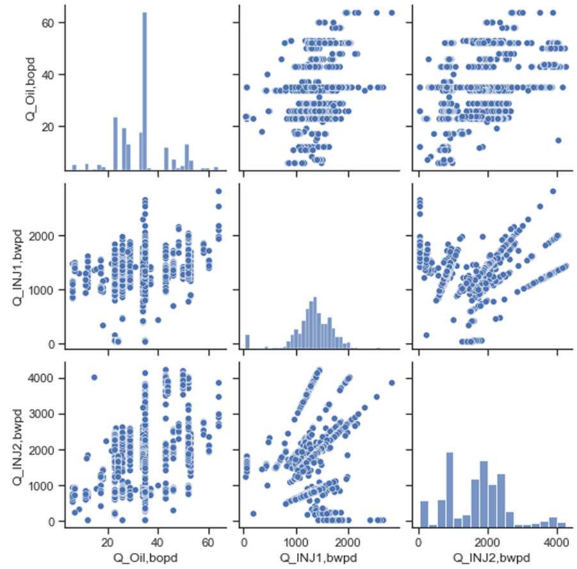
III. Result and Discussion

This research aims to create a predictive model for oil production rate using AI, with two independent variables, x_1 and x_2, and one dependent variable, y. The descriptions of the three variables are provided in Table 2. The daily oil production average rate from a production well is 33.90 bbl, and the water injection rate in well 1 is lower than in well 2, which is 1339.78 bwpd.

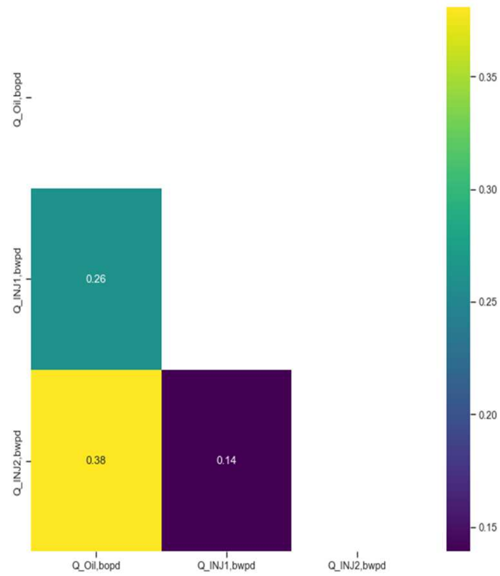
TABLE II
DATA DESCRIPTION

	Q_Oil, BOPD (y)	Q_Inj1, BWPD (x_1)	Q_Inj2, BWPD (x_2)
Count	1180	1180	1180
Mean	33.90	1339.78	1687.21
Minimum	6	46.43	23.76
Maximum	64	2826.35	4232.16

Based on the 1180 data points for each variable, predictive models were developed using MLR, PR, and ANN. A nonlinear trend can be observed upon observing the scatter plot in Figure 4a. Additionally, the low correlation results among the variables in Figure 4b further support the initial indication of non-linearity in the data. However, in this study, both linear (MLR) and nonlinear methods (PR and ANN) were employed to forecast the oil production rate, allowing for a comparison of the results.



(a)



(b)

Fig. 4 Trend of Data Description (a) Scatter Plot (b) Linear Correlation

A. *Model Prediction*

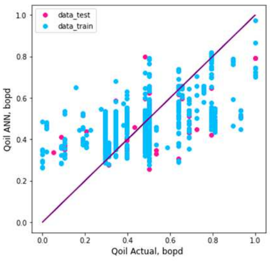
The study resulted in three different predictive models. The selected model is the one that best suits the actual conditions. The predictive model generated using the MLR method is as follows in equation (11).

$$y = 19.012 + (0.006 \times x_1) + (0.004 \times x_2) \tag{11}$$

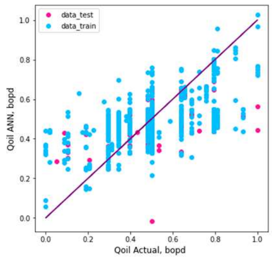
The predictive model using the PR method with a second order is as follows in equation (12).

$$\begin{aligned} y = 22.33 &+ (0.0012 \times x_1) + (0.0031 \times x_2) \\ &+ (0.000002 \times x_1^2) \\ &- (0.00000015 \times x_2^2) \\ &+ (0.0000003 \times x_1 \times x_2) \end{aligned} \tag{12}$$

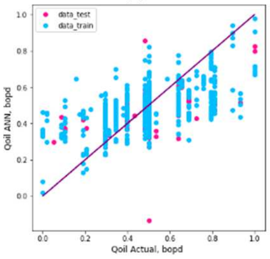
The predictive model using the ANN method in this study was built using Python. It underwent training through a total of 50,000 iterations using a learning rate of 0.001. The ANN's peak performance was reached by iteratively training the network with varying combinations of hidden layers, hidden neurons, and activation functions until the network weights reached the minimum acceptable error level. The ANN model produced several architecture models, as shown in Figures 5 and 6.



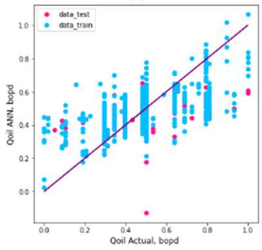
R^2 test= 0.153,
R^2 train= 0.312
(a)



R^2test= 0.116
R^2train= 0.372
(b)



R^2 test= 0.136
R^2 train= 0.374
(c)

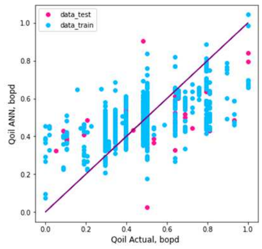


R^2 test= 0.092
R^2 train= 0.384
(d)

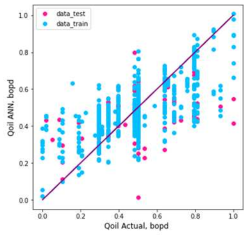
Fig. 5 ANN Model Architecture with a Single Hidden Layer. (a) 2-4-1 Model (b) 2-6-1 Model (c) 2-8-1 Model (d) 2-10-1 Model

The selected model was obtained by trying out multiple architecture models of the ANN. The chosen model had the smallest RMSE value and the highest R^2 value on the testing dataset. In Figures 5a, 5b, 5c, and 5d, there are four ANN architectures with two neurons in the input layer, one hidden layer with varying numbers of hidden neurons, and one neuron in the output layer. An example of 2-4-1 model has ANN architectures with two neurons in the input layer, four in the second hidden layer, and one in the output layer. In Figures 6a, 6b, 6c, and 6d, there are two neurons in the input layer, two hidden layers with different numbers of hidden neurons, and one neuron in the output layer. An example of 2-10-10-1 model has ANN architectures with two neurons in the input layer, ten neurons in the first hidden layer, ten neurons in the second hidden layer, and one neuron in the output layer.

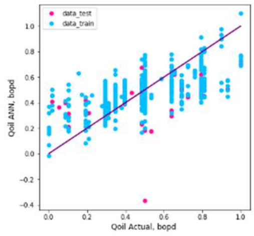
Eight architecture models of the ANN were compared RMSE (Figure 7) and R^2 test values (Figure 5 and Figure 6). The model with the lowest RMSE and highest R^2 test on the testing dataset was selected. The chosen model exhibited an RMSE of 0.142 and an R^2 test value of 0.162. This model, characterized by the architecture (2-4-4-1), features two neurons in the input layer, four neurons in the first hidden layer, four neurons in the second hidden layer, and one neuron in the output layer.



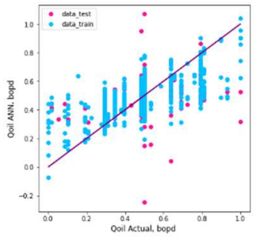
R^2 test= 0.162
R^2 train= 0.357
(a)



R^2 test= 0.076
R^2 train= 0.444
(b)



R^2 test= 0.094
R^2 train= 0.443
(c)



R^2 test= 0.098
R^2 train= 0.481
(d)

Fig. 6 ANN model architecture with two hidden layer. (a) 2-4-4-1 model (b) 2-6-6-1 model (c) 2-8-8-1 model (d) 2-10-10-1 Model.

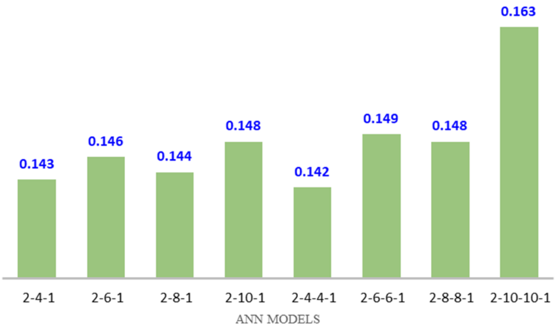


Fig. 7 RMSE for ANN Models

B. *Model Selection*

The predictive models generated using the ANN, MLR, and PR methods were compared. The RMSE and R^2 test values obtained from the three methods are presented in Figure 6. The model that will be chosen as the predictive model is the one with a small RMSE value and a large R^2 test value. From Figure 8, it can be observed that the RMSE value of the ANN model is the smallest, at 0.142, followed by the PR model at 8.407, and the MLR model has the largest RMSE value of 8.422. A smaller RMSE value indicates a lower error, making the ANN model the best among the three.

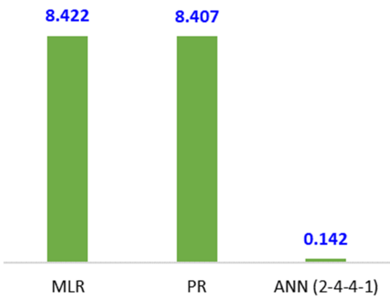
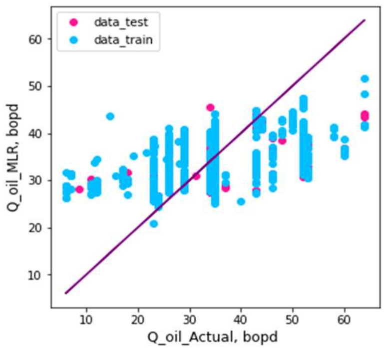
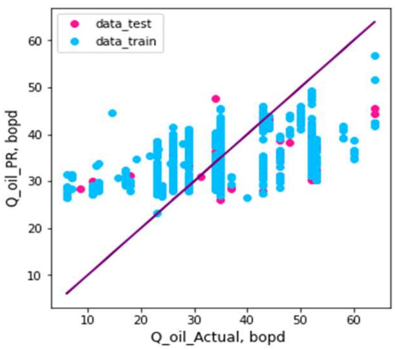


Fig. 8 RMSE of MLR, PR, and ANN (2-4-4-1) models

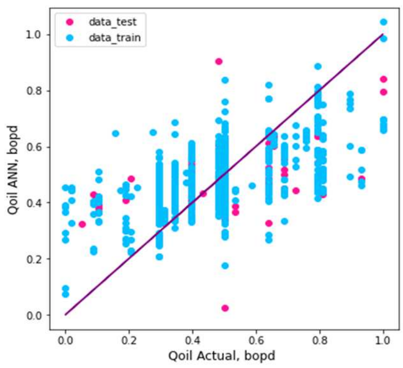
Additionally, in terms of R^2 test values, the ANN method achieved the highest score at 16.2%, surpassing the PR model with 12.8% and the MLR model with 12.5%. The highest R^2 value of the ANN method underscores its better-fitting capability. A visual examination of the scatter plot comparing the predicted oil production values to the actual data from the testing set reveals that the ANN model's predictions closely follow the purple line, signifying a closer match to the actual values.(Figure 9).



R^2 test= 0.125, R^2 train= 0.201
(a)



R^2 test= 0.128, R^2 train= 0.206
(b)



R^2 test= 0.162, R^2 train= 0.358
(c)

Fig. 9 Comparison of models (a) MLR (b) PR (c) ANN (2-4-4-1)

Based on these comparison results, the ANN method outperforms the MLR and PR methods in predicting oil production rates. The MLR and PR are general statistical methods with linear and nonlinear model limitations. On the contrary, ANN has shown its power in this study, which can build a nonlinear model with complex patterns and generate less error. In oil and gas production, ANN can predict the oil production rate in South Sumatra fields quickly and effectively compared to conventional methods, which require a lot of time and costs. The prediction model obtained is a fast method that can be implemented later in estimating oil production rates in fields in South Sumatra.

IV. Conclusion

One of the conventional methods of predicting oil production from a production well is reservoir simulation. The need for many data and running in conducting reservoir simulations is a problem because it requires time and cost. The implementation of AI in predicting oil production rates was carried out in this study using different methods, namely MLR, PR, and ANN. The ANN model yielded the smallest RMSE value and the largest R^2 test value, indicating that the ANN model is the most suitable predictive model. Future studies can explore other AI methods to improve predictive models further. Additionally, incorporating other influential variables besides the water injection rate could enhance the accuracy of oil production predictions.

Acknowledgment

The authors are genuinely grateful to the Indonesian Ministry of Energy and Mineral Resources, especially Politeknik Energi dan Mineral Akamigas, for supporting this study.

References

[1] K. Lichun, L. He, R. Yili, L. Kai, S. Mingyu, S. Jian, and L. Xin, "Application and development trend of artificial intelligence in petroleum exploration and development," *Petrol. Explor. Develop.*, Vol. 48, pp.1–14, February 2021, doi:10.1016/S1876-3804(21)60001-0.

[2] N. Salleh, S. S. Yuhaniz, and N. F. M. Azmi, "Modeling Orbital Propagation Using Regression Technique and Artificial Neural Network," *IJASEIT*, Vol. 12, pp. 1279-1286, 2022, doi:10.18517/ijaseit.12.3.15366.

[3] R. A. Azim, "A new correlation for calculating wellhead oil flow rate using artificial neural network," *Artificial Intelligence in Geosciences*, vol. 3, 2022, doi: 10.1016/j.aiig.2022.04.001.

[4] A. Sircar, K. Yadav, K. Rayavarapu, N. Bist, and H. Oza, "Application of machine learning and artificial intelligence in oil and gas industry," *Petroleum Research*, vol. 6, pp. 379-391, 2021, doi:10.1016/j.ptlrs.2021.05.009.

[5] H. H. Alkinani, A. T. T. Al-Hameedi, S. Dunn-Norman, and R. E. Flori, "Applications of Artificial Neural Networks in the Petroleum Industry: A Review," presented at the SPE Middle East Oil and Gas Show and Conference held in Manama, Bahrain, 2019, doi:10.2118/195072-MS.

[6] S. Bahaloo, M. Mehrizadeh, and A. Najafi-Marghmaleki, "Review of application of artificial intelligence techniques in petroleum operations," *Petroleum Research*, vol. 8, pp.176-182, 2023, doi:10.1016/j.ptlrs.2022.07.002.

[7] P. Solanki, D. Baldaniya, D. Jogani, B. Chaudhary, M. Shah, and A. Kshirsagar, "Artificial intelligence: New age of transformation in petroleum upstream," *Petroleum Research*, vol. 7, pp. 106-114, 2022, doi: 10.1016/j.ptlrs.2021.07.002.

[8] P. Tontiwachwuthikul, C. W. Chan, F. Zeng, Z. Liang, T. Sema, and C. Min, "Recent progress and new developments of applications of artificial intelligence (AI), knowledge-based systems (KBS), and Machine Learning (ML) in the petroleum industry," *Petroleum*, vol. 6, pp. 319-320, 2020, doi: 10.1016/j.petlm.2020.08.001.

[9] R. Syah, N. Ahmadian, M. Elveny, S.M. Alizadeh, M. Hosseini, A. Khan, "Implementation of artificial intelligence and support vector machine learning to estimate the drilling fluid density in high-pressure high-temperature wells," *Energy Reports*, vol 7, 2021, doi:10.1016/j.egyr.2021.06.092.

[10] O. E. Agwua, J. U. Akpabioa, and A. Dosunmub, "Artificial neural network model for predicting drill cuttings settling velocity," *Petroleum*, vol. 6, 2020, doi: 10.1016/j.petlm.2019.12.003.

[11] F. Faraji, J. O. Ugwu, and P. L. Chong, "Modeling two-phase Z factor of gas condensate reservoirs: Application of Artificial Intelligence (AI)," *Journal of Petroleum Science and Engineering*, vol. 208, 2022, doi: 10.1016/j.petrol.2021.109787.

[12] D. Rosiani, A. K. Permadi, H. P. S. Siregar, A. Y. Gunawan, and T. Ariadji, "A new CO_2-EOR methods screening model based on interdependency parameters," *Applied Sciences*, vol. 12, pp. 1-17, 2022, doi: 10.3390/app12083937.

[13] S. Bhattacharyya and A. Vyas, "A novel methodology for fast reservoir simulation of single-phase gas reservoirs using machine learning," Heliyon, vol. 8, 2022, doi: 10.1016/j.heliyon.2022.e12067.

[14] C. S. W. Ng, M. N. Amar, A. J. Ghahfarokhi, L. S. Imsland, "A Survey on the Application of Machine Learning and Metaheuristic Algorithms for Intelligent Proxy Modeling in Reservoir Simulation," *Computers and Chemical Engineering*, vol. 170, 2023, doi: 10.1016/j.compchemeng.2022.108107.

[15] H. Ghorbania, D. A. Woodb, A. Choubinehc, A. Tatard, P. G. Abarghoyie, M. Madanic, and N. Mohamadian, "Prediction of oil flow rate through an orifice flow meter: Artificial intelligence alternatives compared," *Petroleum*, vol 6, 2018, doi: 10.1016/j.petlm.2018.09.003.

[16] N. C. Chakra, K. Y. Song, M. M. Gupta, and D. N. Saraf, "An innovative neural forecast of cumulative oil production from a petroleum reservoir employing higher-order neural networks (HONNs)," *Journal of Petroleum Science and Engineering*, vol. 106, pp.18-33, Jun. 2013, doi: 10.1016/j.petrol.2013.03.004.

[17] A. Sabaa, M. A. El-Ela, A. H. El-Bandi, and M. H.M. Sayyouh, "Artificial Neural Network Model to Predict Production Rate of Electrical Submersible Pump Wells," *SPE Prod & Oper*, vol. 38, pp. 63-71, 2023, doi: 10.2118/212284-PA.

[18] P. Sagar, P. Gupta, and I. Kashyap, "A forecasting method with efficient selection of variables in multivariate data sets," *Int. j. inf. tecnol.*, vol. 13, Jun. 2021, pp. 1039-1046, doi: 10.1007/s41870-021-00619-9.

[19] G. A. Ngige, P. E. Ovuoraye, C. A. Igwegbec, E. Fetahi, J. A. Okeke, A. D. Yakubu, and P. C. Onyechi, "RSM optimization and yield prediction for biodiesel produced from alkali-catalytic transesterification of pawpaw seed extract: Thermodynamics, kinetics, and Multiple Linear Regression analysis," *Digital Chemical Engineering*, vol. 6, pp. 2772-5081, Nov. 2022, doi:10.1016/j.dche.2022.100066.

[20] N. A. Carlson, A. Singh, M. S. Talmadge, Y. Jiang, G. G. Zaimes, S. Li, T. R. Hawkins, L. Sittler, A. Brooker, D. J. Gaspar, R. L. McCormick, and M.M. Ramirez-Corredores, "Economic analysis of the benefits to petroleum refiners for low carbon boosted spark ignition biofuels," *Fuel*, vol. 334, Nov. 2022, doi: 10.1016/j.fuel.2022.126183.

[21] M. Ajona, P. Vasanthi, and DS Vijayan, "Application of multiple linear and polynomial regression in the sustainable biodegradation process of crude oil," *Sustainable Energy Technologies and Assessments*, vol. 54, Dec. 2022, doi: 10.1016/j.seta.2022.102797.

[22] Y. Kassem, H. Çamur, G. E. Duke, and A. H. Abdelnaby, "Artificial Neural Networks, Quadratic Regression, and Multiple Linear Regression in Modeling Cetane Number of Biodiesels*," in Proc. International Conference on Data Science and Application*, Feb. 2023, pp. 217-224, doi: 10.1007/978-981-19-6634-7_16.

[23] N. A. Hazrin, K. L. Chong, Y. F. Huang, A. N. Ahmed, J. L. Ng, C. H. Ko, K. W. Tan, M. Sherif, and A. El-shafie, "Predicting sea levels using ML algorithms in selected locations along coastal Malaysia," *Heliyon*, vol. 9, Aug. 2023, doi: 10.1016/j.heliyon.2023.e19426.

[24] A. S. Yusuff, N. B. Ishola, A. O. Gbadamosi, T. M.Azeez, and M. O. Onibonoj, "An artificial intelligence approach to model and optimize biodiesel production from used cooking oil using CaO incorporated zeolite catalyst," *Energy Conversion and Management: X*, vol. 20, Sep. 2023, doi: 10.1016/j.ecmx.2023.100452.

[25] L. A. E. Gomez, J. A. G. Gonzalez, and J. M. M. Navarro, "What is happening in shrinking medium-sized cities? A correlational analysis and a multiple linear regression model on the case of Spain," *Cities*, vol. 134, Jan. 2023, doi: 10.1016/j.cities.2023.104205.

[26] A. Yalçınkaya, I. G. Balay, Birdal, and S. Enoğlu, "A new approach using the genetic algorithm for parameter estimation in multiple linear regression with long-tailed symmetric distributed error terms: An application to the Covid-19 data," *Chemometr Intell Lab*., vol. 216, June 2021, doi: 10.1016/j.chemolab.2021.104372.

[27] E. Aissaoui, Y. E. A. E. Madani, L. Oughdir, A. Dakkak, and Y. E. Allioui, "A Multiple Linear Regression-Based Approach to Predict Student Performance Ouafae," *AISC*, vol. 1102, pp. 9-23, Jan. 2020, doi: 10.1007/978-3-030-36653-7_2.

[28] S. Rath, A. Tripathy, and A. R. Tripathy, "Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, pp. 1467-1474, July 2020, doi:10.1016/j.dsx.2020.07.045.

[29] M. R. Elkharbotly, M. Seddik, and A. Khalifa, "Toward sustainable water: prediction of non-revenue water via artificial neural network and multiple linear regression modelling approach in Egypt," *Ain Shams Engineering Journal*, vol. 13, September 2022, doi:10.1016/j.asej.2021.101673.

[30] H. Alahmer, A. Alahmer, R. Alkhazaleh, and M. I. Al-Amayreh, "Modeling, polynomial regression, and artificial bee colony optimization of SI engine performance improvement powered by acetone–gasoline fuel blends," Energy Report, vol. 9, pp. 55-64, Jan. 2023, doi: 10.1016/j.egyr.2022.12.102.

[31] S. Sarwar, H. Mokhlis, M. Othman, H. Shareef, L. Wang, N. N. Mansor, A. S. M. Khairuddin, and H. Mohamad, "Application of polynomial regression and MILP for under-frequency load shedding scheme in islanded distribution system," Alexandria Engineering Journal, vol. 61, Jun 2021, doi: 10.1016/j.aej.2021.06.028.

[32] R. J. Freund, W. J. Wilson, and P. Sa, *Regression Analysis Statistical Modeling of a Response Variable*, London, UK: Academic Press, 2006.

[33] L. Fausett, *Fundamentals of neural networks architectures, algorithms, and applications*, Prentice Hall, Inc., London, pp. 320-323, 1994.

[34] A. Ghaemi, M. K. Dehnavi, and Z. Khoshraftar, "Exploring artificial neural network approach and RSM modeling in the prediction of CO_2 capture using carbon molecular sieves," *Case Studies in Chemical and Environmental Engineering*, vol. 7, 2023, doi:10.1016/j.cscee.2023.100310.

[35] P. Zafari and A. Ghaemi, "Modeling and optimization of CO2 capture into mixed MEA-PZ amine solutions using machine learning based on ANN and RSM models," *Results in Engineering*, vol. 19, Jun. 2023, doi: 10.1016/j.rineng.2023.101279.

[36] J. V. Tu, "Advantages and Disadvantages of Using Artificial Neural Networks versus Logistic Regression for Predicting Medical Outcomes," *Journal of Clinical Epidemiology*, vol. 49, March 1996, doi: 10.1016/S0895-4356(96)00002-9.

[37] E. M. Shokir, M. H. Sayyouh, and K. A. Fattah, "Selection and evaluation EOR method using artificial intelligence," In Proc. Annual International Conference and Exhibition, August 2002, doi:10.2118/79163-MS.

[38] S. A. Irfan, F. A. Abdulkareem, A. Radman, G. Faugere, and E. Padmanabhan, "Artificial neural network (ANN) modeling for CO2 adsorption on Marcellus Shale," *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 1003, 2022, doi: 10.1088/1755-1315/1003/1/012029.

[39] P. Singh, P. Bera, S. Bhardwaj, and S. K. Singh, "Predictive models for bond strength of reinforced concrete with the application of ANN," *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 889, 2021, doi:10.1088/1755-1315/889/1/012080.

[40] P. Wellyantama and S. Soekirno, "Temperature, pressure, relative humidity and rainfall sensors early error detection system for automatic weather station (AWS) with artificial neural network (ANN) backpropagation," *J. Phys.: Conf. Ser.*, vol. 1816, 2021, doi:10.1088/1742-6596/1816/1/012056.

[41] K. Khan, M. A. M. Johari, M. N. Amin, M. I. Khan, and M. Iqbal, "Optimization of colloidal nano-silica based cementitious mortar composites using RSM and ANN approaches," *Results in Engineering*, vol. 20, 2023, doi: 10.1016/j.rineng.2023.101390.

[42] H. Cao, B. Cao, C. Xia, Z. Meng, H. Bian, and M. Ding, "Prediction of heat transfer coefficients for steam condensation in the presence of air based on ANN method," *International Journal of Advanced Nuclear Reactor Design and Technology*, vol. 5, 2023, doi:10.1016/j.jandt.2023.07.001.