PROCEEDINGS, 46th Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 15-17, 2021
SGP-TR-218

A New Machine Learning Algorithm for Production Well Analysis

Muchamad Harry, Jantiur Situmorang, and Welly Prabata,

AILIMA, CIBIS Business Park, TB Simatupang, Jakarta, Indonesia

harry@ailima.co.id, jantiur@ailima.co.id, welly@ailima.co.id

Keywords: well decline, machine learning, ANN, decision tree regressor, JIWA Flow.

ABSTRACT
Artificial intelligence or machine learning has been one of the buzzwords recently. The rapid development in the technology sector has greatly affected our lives. They are of great help to humankind. In terms of geothermal, production well is an asset that must be well managed. One of the challenges faced by reservoir engineers in managing the geothermal field is detecting problems in production well early. AILIMA, as a company, try to solve this problem by utilizing machine learning. Using WHP and production mass flowrate data with or without reservoir pressure data, some machine learning algorithms have been trained and tested to estimate the production decline rate. JIWA Flow wellbore simulator has been used to generate synthetic production history datasets for the machine learning test. Fifteen of existing machine learning algorithms tested, no one can provide a good result compared to a wellbore simulator. Therefore, a new algorithm, namely AILIMA-ONE is proposed. This paper describes the application of AILIMA-ONE to estimate production decline. The comparison against ANN (the most complex algorithm) and Decision Tree Regressor (the highest R^2 during training process) is also provided.

1. INTRODUCTION
One of the crucial things in managing geothermal field is monitoring production wells performance. To evaluate the performance of a production well, currently, the engineer would perform a complex wellbore simulation that may require few days or longer in the process. The complicated process in wellbore simulation is generally related to data preparation, model development, and manual iterative trial and error in the calibration process. In many cases, production issues are noticed when the well has dropped significantly.

On the other hand, in terms of technology, there have been many artificial intelligence and machine learning applications in helping human work. Machine learning can be designed to find out correlations between variables and can predict unseen things. It would be beneficial if machine learning could be implemented in the geothermal industry so that it could help to evaluate a well's performance more quickly.

There are so many machine learning algorithms that have been developed, such as linear regression, polynomial regression, decision tree regressor, SVR, naïve Bayes, k-NN, random forest, GBR, AdaBoost, VAR, GPR, stochastic gradient descent, XGBoost, feed forward neural networks, and RNN. All of those algorithms have been tested using synthetic production data generated from JIWA Flow wellbore simulator. Due to the unsatisfied result obtain from the algorithms above, a new algorithm, namely AILIMA-ONE, has been developed by modifying the existing algorithm and incorporating reservoir engineer fundamentals. This algorithm will be used in JIWA GENEWS, an intelligent machine that designed for geothermal power generation early warning system.

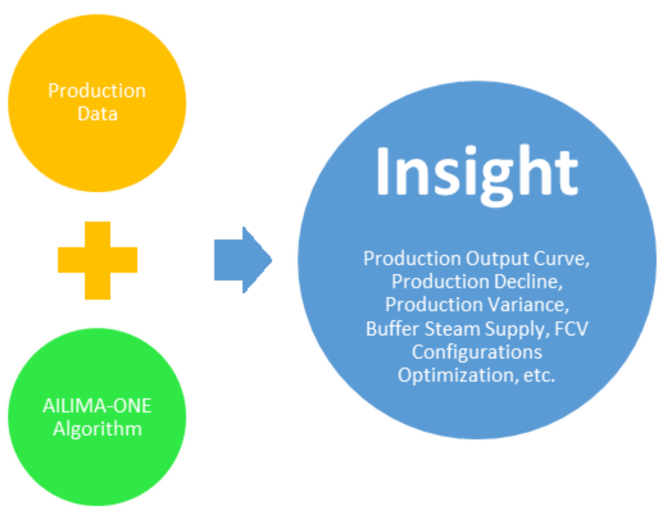

Figure 1: JIWA GENEWS Workflow

This study focuses on comparing AILIMA-ONE algorithm's performance to the other machine learning algorithms in terms of analyzing production anomaly and calculating production decline rate. This approach aims to assist reservoir engineers in managing their production wells and field production supply as well as a better understanding of the performance of the well more efficiently and effectively.

2. LITERATURE REVIEW

Machine learning has been applied in many aspects in the oil and gas or geothermal sector. In the oil and gas sector, Horne and Aljubran (2020) trained a neural network to predict multilateral well performance using ICV configurations and flowrate data. The data generated using homogenous and heterogenous numerical reservoir model allows for accurately predicting flow profiles at all possible ICV configuration. In exploration, Elmousalami and Elaskary (2020) developed a reliable classification machine learning model for drilling pipe stuck. Based on that research, the exploration team can predict the potential stuck pipe before the drilling started. So, they can prepare various kinds of mitigation to avoid the issue.

In line with machine learning development in oil and gas, research on machine learning implementation to solve many geothermal problems also increases. Haklidir F and Haklidir M (2019) use linear regression, linear SVM, and deep neural network model to predict geothermal reservoir temperature using geochemical data. Meanwhile, in the geothermal sector, Ariturk (2019) has implemented MLP to predict flowrate in production well. In our case, we use generated geothermal production data to predict decline using an algorithm that was developed by AILIMA.

3. METHODOLOGY

Production data used in this paper is generated using a wellbore simulator, namely JIWA Flow. The generated data, including time, WHP, mass flow rate, and with or without reservoir pressure, are used to train the AILIMA-ONE algorithm before it is used to predict production decline. For comparisons, the generated data are also used to train the other machine learning algorithms, including decision tree regressor and artificial neural network (ANN). In this study, we developed 3 cases of data. Case 1 is for dry steam well without scaling and cooling problems. Case 2 is for dry steam well with a scaling problem in the reservoir, and case 3 is for a two-phase well with a cooling problem.

4. DATA CONSTRUCTION

4.1 Case 1

Dry steam well is assumed to have casing configurations as provided in Table 1 below:

Table 1: Well design of case 1

MD	Type	Size
0-1200	Production Casing	13 3/8 inch
1200-1600	Perforated Liner	10 3/4 inch

A single feed zone is assumed in 1400 mMD, 38 bar reservoir pressure, and a *kh* of 6.35E-11 m^3. To generate 7000 daily production data, wellbore simulator is run 7000 times from 1/1/2000 using constant parameters as above except for the reservoir pressure which is assumed to decline 2%/year. The generated data from Case 1 is shown in **Figure 2**.

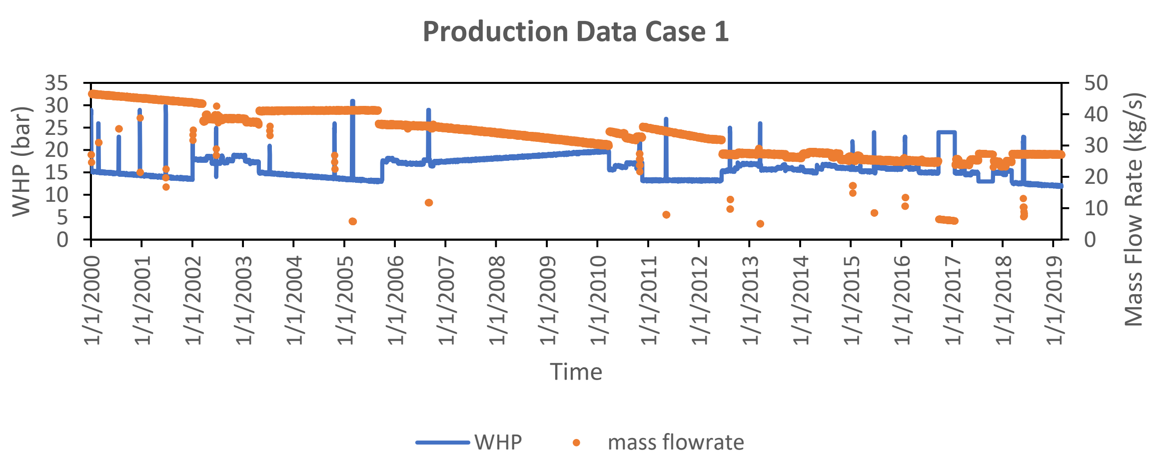

Figure 2: Generated production data for the case 1

4.2 Case 2 (Scaling Problem)

The well casing configuration and reservoir pressure in case 2 are assumed to be the same as case 1. The productivity index of case 2 is shown in **Figure 3**. The generated data from this case is provided in **Figure 4**.

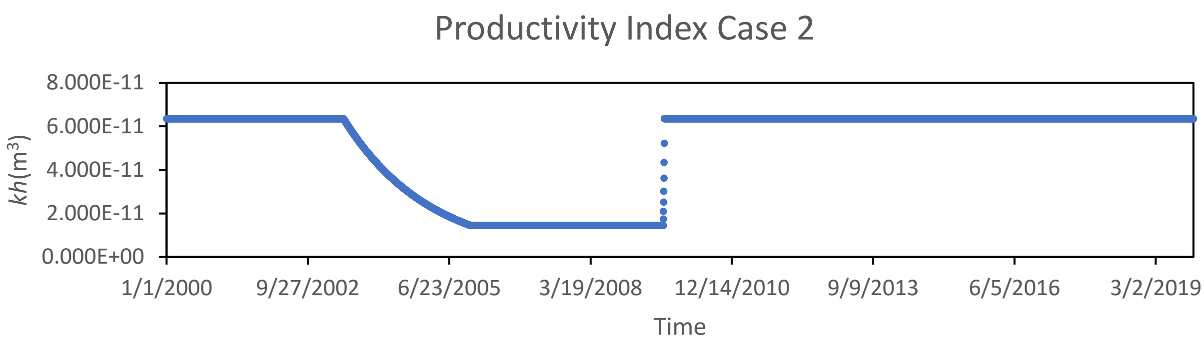

Figure 3: Productivity Index profile of case 2

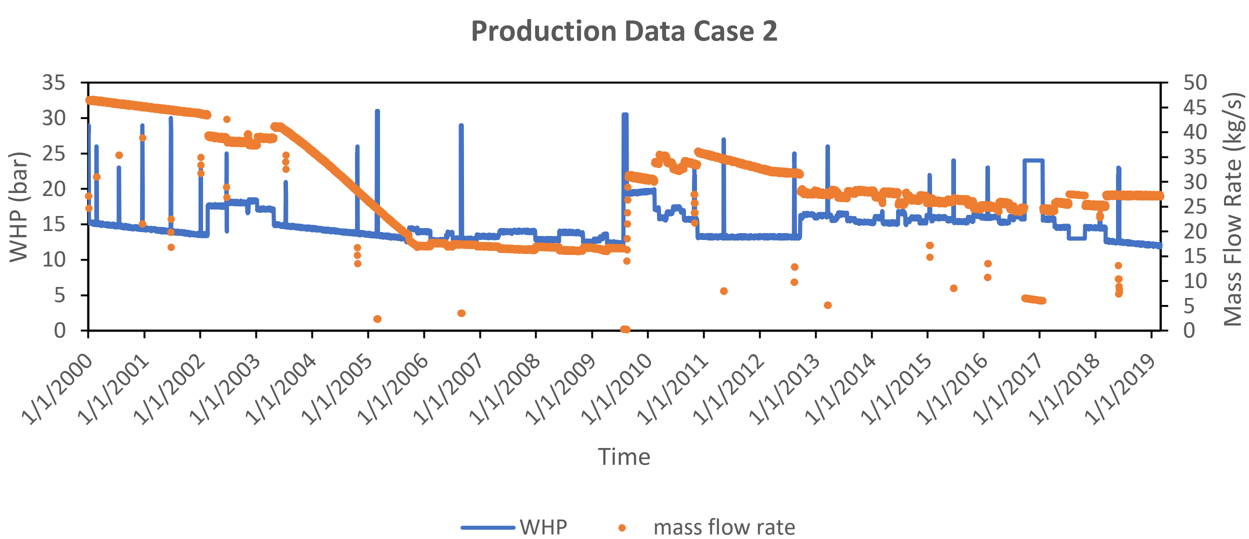

Figure 4: Production data case 2 (dry steam well with scaling problem) from wellbore simulator

4.3 Case 3 (Cooling Problem)

Case 3 is a two-phase well with casing configuration as follows:

Table 2: Well design of case 3

MD	Type	Size
0-1200	Production Casing	13 3/8 inch
1200-1600	Perforated Liner	10 ¾ inch

Two feed zones are assumed with the following parameters:

Table 3: Reservoir parameters of case 3

MD	Reservoir Pressure	Enthalpy	*kh*
1200	100 bar	1340 kJ/kg	0.5E-12 m^3
1500	126 bar	1370 kJ/kg	1.2E-12 m^3

Wellbore simulator has been run 7000 times under different specified WHP from 1/1/2000 using constant parameters as above except for enthalpy which, assumed to decline at 0.5%/year. The constructed data from this case as shown in **Figure 5.**

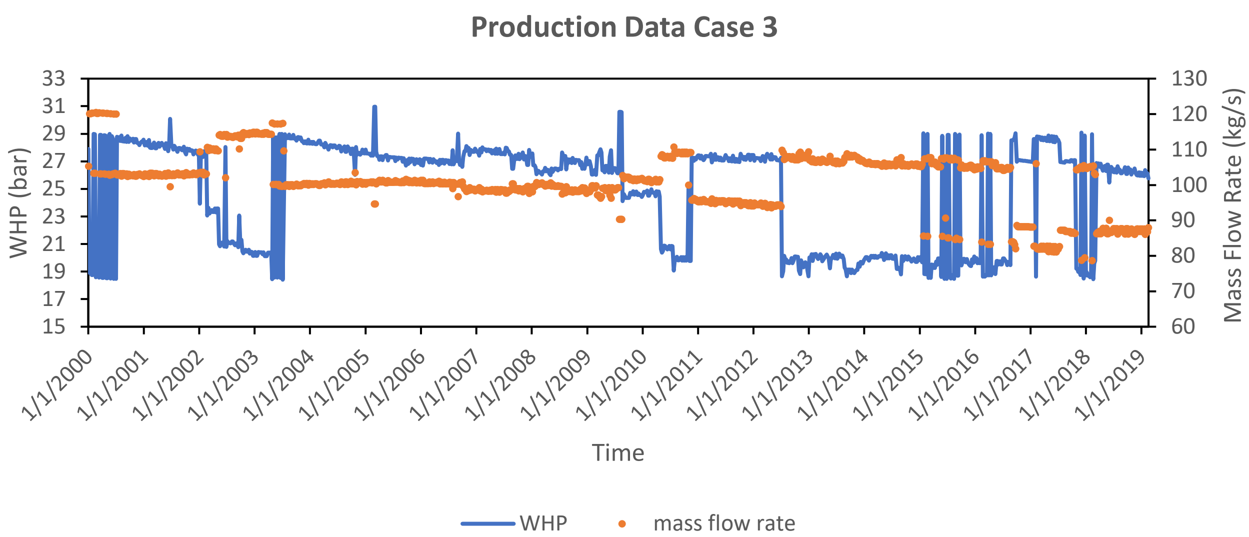

Figure 5: Production data case 3 (two-phase well with cooling problem) from wellbore simulator

5. MACHINE LEARNING

To calculate production decline, the production data are normalized at specified constant WHP. Before the algorithms are trained with generated data, the generated datasets are divided into training and testing datasets with a proportion of 80:20 randomly. Once the datasets are divided, they are used to train and test the algorithms. One thing that we must avoid is overfitting. Overfitting is a condition when the algorithm can predict accurately in the training dataset but cannot predict accurately in the testing dataset.

5.1 Artificial Neural Network (ANN)

ANN is used to train production data and predict mass flow rate at certain WHP. The ANN architecture applied in this test has 4 hidden layers with 32, 16, 8, 8 neurons. The results of the training process as follows:

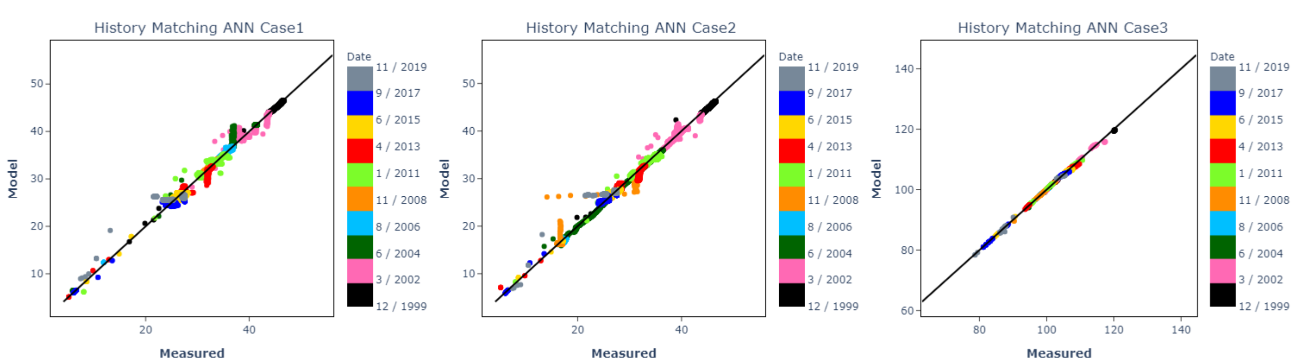

Figure 6: The results of training process using ANN in case 1, case 2, and case 3.

The R squared of the model for case 1, case 2, and case 3 are 0. 0.9904, 0.9949 and 0.9985, respectively.

5.2 Decision Tree Regressor

Decision tree regressor is one of the famous algorithms in machine learning. The results of the training process are as follows:

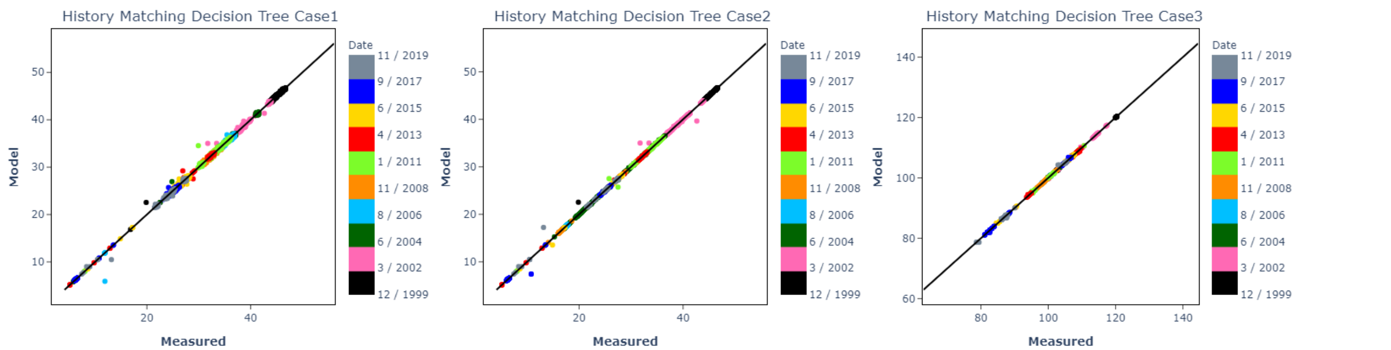

Figure 7: The results of training process using decision tree regressor in case 1, case 2, and case 3.

The R squared of the model of case 1, case 2, and case 3 are 0.9995, 0.9998 and 0.9998, respectively.

5.3 AILIMA-ONE Algorithm

The results of the training process using AILIMA-ONE algorithm are as follows:

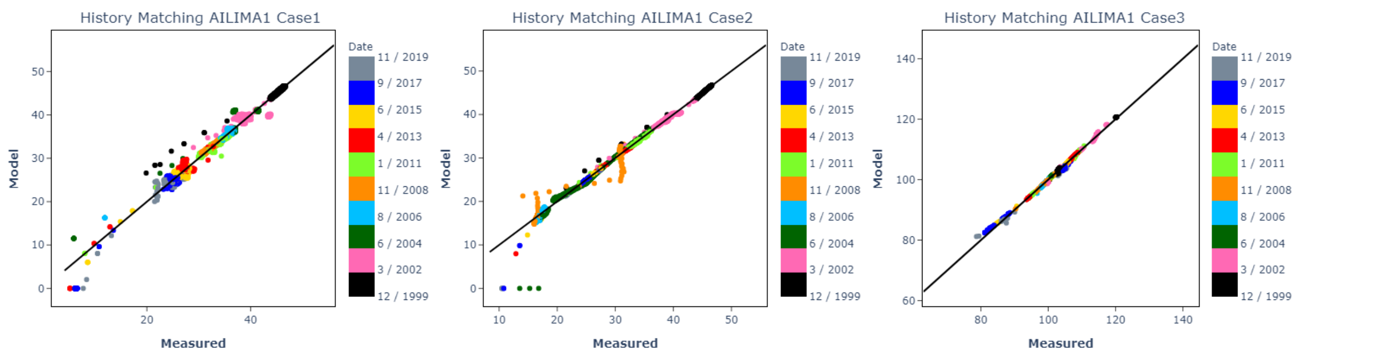

Figure 8: The results of training process using AILIMA-ONE algorithm in case 1, case 2, and case 3.

The R squared of the model for each case is 0.9839, 0.9884 and 0.9963, respectively.

6. DECLINE CALCULATION

After completing the training process, the machine learning models are used to normalize the production data at constant WHP. The normalized production flowrates are compared against the generated data from the wellbore simulator.

6.1 Case 1

Mass flow rate normalization of the AILIMA-ONE algorithm gives the best match with the generated data compared with the other algorithms as shown in **Figure 9**. Meanwhile, the prediction of the decision tree regressor, which has the greatest R squared, has significant errors. The results suggest that the decision tree regressor algorithm quite overfit as it cannot predict the unseen data accurately. The discontinuity of the decision tree regressor also becomes a concern here. In line with the decision tree regressor, the prediction from ANN is also not good enough.

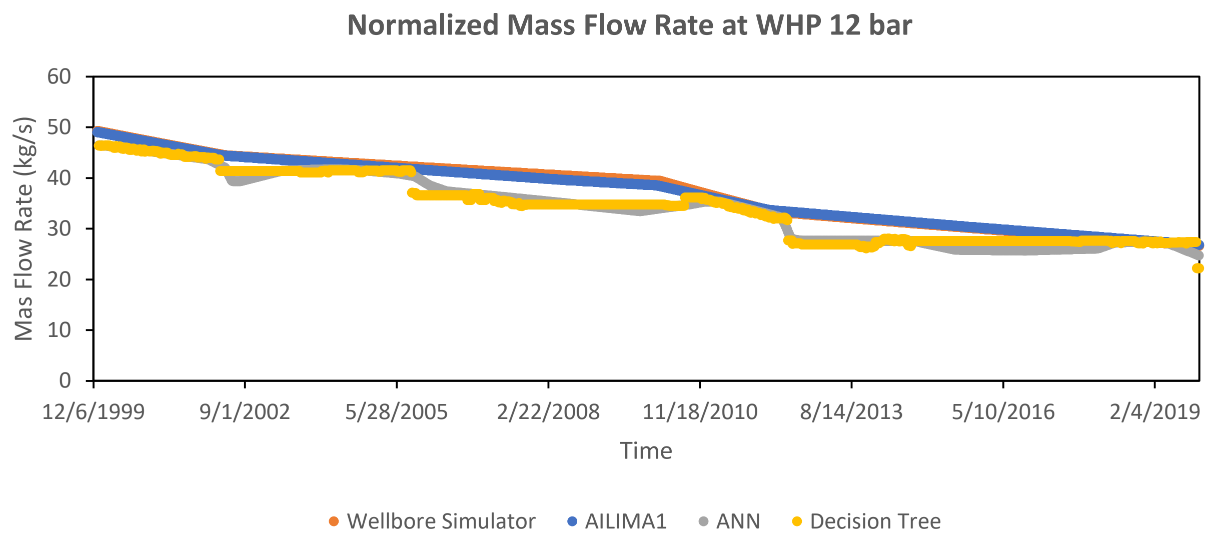

Figure 9: Normalized mass flow rate comparison between wellbore simulator, AILIMA-ONE algorithm, ANN, and decision tree regressor of case 1.

6.2 Case 2

In line with case 1, the AILIMA-ONE algorithm also gives the best result. The scaling process in the well can be detected by AILIMA-ONE algorithm more accurately as shown in **Figure 10**. ANN and decision tree regressor give unstable normalization results, such as in 2002 and 2010-2013.

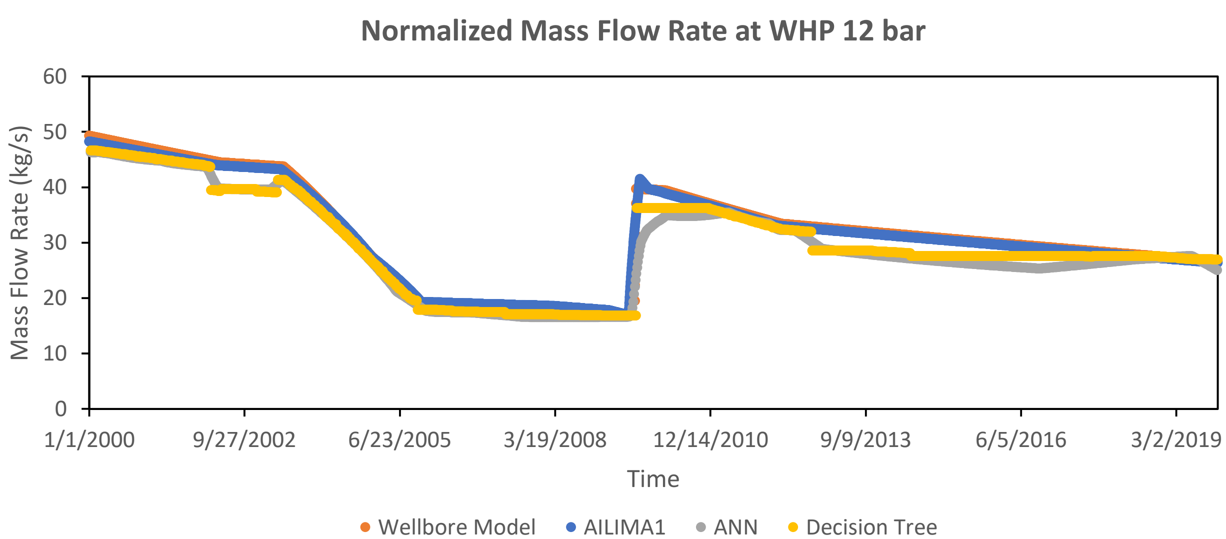

Figure 10: Normalized mass flow rate comparison between wellbore simulator, AILIMA-ONE algorithm, ANN, and decision tree regressor in case 2.

6.3 Case 3

Case 3 represents the two-phase well with a cooling problem. Theoretically, a two-phase well has more complexity than a steam well. But here, both AILIMA-ONE and ANN give excellent results as shown in **Figure 11**. It can match the wellbore simulator accurately. However, in the earlier time, ANN gives a little pessimistic prediction than the wellbore simulator. Meanwhile, the decision tree regressor algorithm still cannot predict normalized mass flow rate consistently even though it has the greatest R squared during the training process.

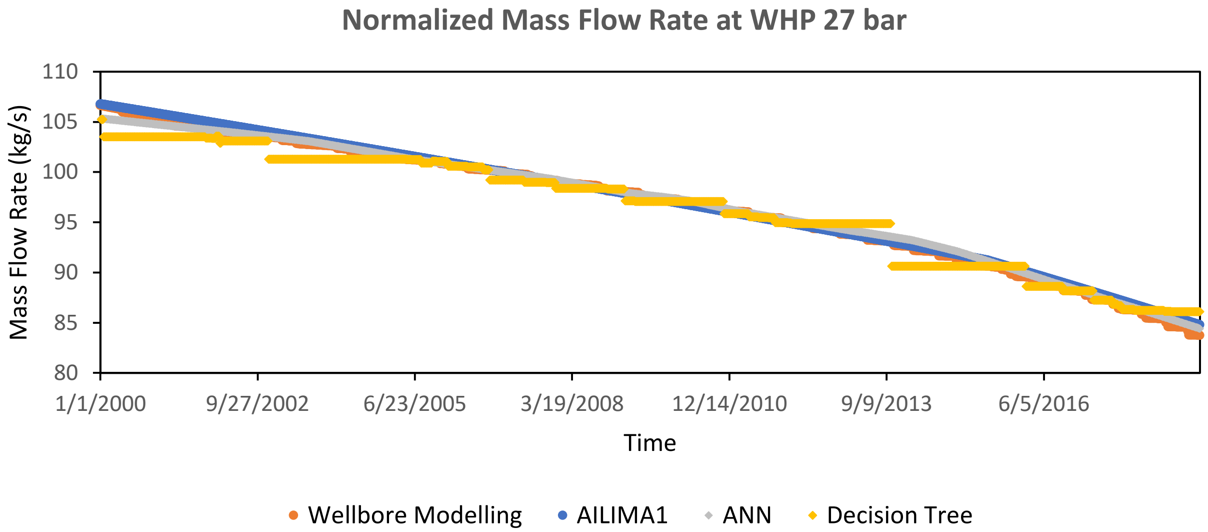

Figure 11: Normalized mass flow rate comparison between wellbore simulator, AILIMA-ONE algorithm, ANN, and decision tree regressor in case 3.

7. CONCLUSION

From the results above, we can see that machine learning can evaluate production performance in geothermal wells. Overall, ANN and decision tree regressor have an excellent result in the training process, but the prediction of normalized mass flow rate is quite far from the wellbore simulator. ANN and decision tree regressor give more pessimistic results than wellbore simulator. The other weakness of the decision tree regressor is, this algorithm provides discontinuity in the prediction. The other weakness of ANN is that it gives turbulent or wavy results (shown in **Figure 9** and **Figure 10**) compared to the wellbore simulator.

Meanwhile, even though its training R squared is lower than the others, AILIMA-ONE gives very great results in predicting the normalized mass flow rate. So, an excellent statistical value in the training process does not guarantee it will be consistent while working with unseen data. The results suggest that the machine equipped with reservoir engineering fundamentals is better than the machines with pure statistics/mathematics. Without incorporating reservoir engineering principles, the results from the machine can be misleading. It is also worth noting that the other benefit from the application of machine learning is, that the production decline analysis can be performed in a more significantly efficient manner than the wellbore simulator.

ACKNOWLEDGEMENT

This proceeding was made with the support and supervision of PT. Anugerah Indonesia Lima (AILIMA). The resource model was simulated using JIWA Flow by PT. Anugerah Indonesia Lima (AILIMA).

REFERENCES

Ariturk, Muhammet Salih.: Optimizing the Production and Injection Wells Flow Rates in Geothermal Field Using Artificial Intelligence, Graduate Theses, Dissertations, and Problem Reports, 3772 (2019).

Aljubran, M. and Horne, R.: Prediction of Multilateral Inflow Control Valve Flow Performance Using Machine Learning, SPE Annual Technical Conference and Exhibition, Calgary, Alberta, Canada (2020).

Elmousalami, H.H., Elaskary, M.: Drilling stuck pipe classification and mitigation in the Gulf of Suez oil fields using artificial intelligence. J Petrol Explor Prod Technol 10, 2055–2068 (2020).

Tut Haklidir, F.S., Haklidir, M.: Prediction of Reservoir Temperatures Using Hydrogeochemical Data, Western Anatolia Geothermal Systems (Turkey): A Machine Learning Approach. Nat Resour Res 29, 2333–2346 (2020). https://doi.org/10.1007/s11053-019-09596-0