Application of Artificial Intelligence in Forecasting Geothermal Production

Galih Kusumo Wardoyo\(^1\), Heru Berian Pratama\(^1\), Sutopo\(^1\), Ali Ashat\(^1\) and Yodha Yudhistira\(^2\)

\(^1\)Geothermal Engineering Master’s Program, Faculty of Mining and Petroleum Engineering, Institut Teknologi Bandung, Jl. Ganesha 10, Bandung 40132, Indonesia
\(^2\)PT. PLN (Persero), Jl. Trunojoyo Blok M-I No.135, Jakarta Indonesia

Email: galihkusumo99@gmail.com
https://orcid.org/0000-0002-8263-5960

Abstract. A reservoir is the main asset of a geothermal business. Decreasing reservoir performance affects the sustainability of production in the future. In planning future production strategies, forecasting production projections are used with the assumption of particular parameter values. A reservoir is a porous media with a heterogeneous nature with a high degree of uncertainty, so using a specific production forecasting method’s assumption parameter becomes inaccurate. This study aims to develop alternative methods for optimizing and estimating geothermal production by eliminating assumptions. Artificial intelligence (AI) is an alternative method that can be used to predict reservoir productions that has properties with a high degree of uncertainty and can be used to optimize production. The AI model was created using measurement data of production parameters at the steam dominated geothermal field in Patuha, Indonesia. AI models in estimating geothermal reservoir production are more accurate in calculating production projections because they use real field data and eliminate assumptions.

1. Introduction
Geothermal reservoir performance is essential in geothermal exploitation, so monitoring must be carried out correctly. A good reservoir monitoring model can be used to predict reservoir production in the future. The model uses certain assumptions in calculating the non-linear conditions and uncertainties of the geothermal reservoir properties. The reservoir is a porous media that is heterogeneous with a high degree of uncertainty, so using the assumption parameters in the production estimation method makes the production estimate less accurate. Artificial intelligence (AI) is an alternative method that can be used to predict reservoir performance that has properties with a high degree of uncertainty.

Artificial intelligence is a technique that adopts the intelligence system of humans to solve problems. One of these methodologies is an artificial neural network. An artificial neural network is a computational machine learning method inspired by biological neuron systems that could process information after weight adjustments are made through training mechanisms with specific algorithms and data. Artificial neural networks have been applied to petroleum engineering. Noshi [1] uses machine learning to predict oil and gas reservoir production. Input parameters used are production time, valve opening, oil volume, gas volume, water volume, and wellhead pressure. Olivares [2] developed a monitoring production method using artificial intelligence on APLT assets. The data used as input is
data from operating wells such as wellhead pressure, flowline pressure, temperature, choke size, bottom-hole pressure, and temperature. The data is used for the training process on artificial intelligence.

In the geothermal application, Serpen [3] applied an artificial neural network model for a Na/K geothermometer. A training process conducted the multilayer feed-forward neural network with Na and K values as input parameters, and the geothermometer temperature was used as output. The artificial neural network model was successfully applied to predict the reservoir temperature. Priyangga and Ruliandi [4] applied artificial neural networks in the geothermal power plant operating system in the Kamojang geothermal field to predict power plant performance. Ariturk [5] optimizes the production flow rate and injection wells in the geothermal field using artificial intelligence. Studies on the development of geothermal reservoir performance analysis using artificial intelligence to estimate geothermal production in Indonesia up to now are still rarely conducted. The development of artificial intelligence science opens opportunities to be applied to the geothermal reservoir performance analysis.

The artificial neural network model was created using measurement data of production parameters at the steam dominated geothermal field in Patuha, Indonesia. Production data is analyzed first to determine which data can be used as input and output on artificial neural networks. The analyzed data are used as input and output data on the artificial neural network architecture that has been compiled. The training process is carried out with several scenarios to determine the extent to which artificial neural networks can understand the characteristics of geothermal wells. The measure of artificial neural networks’ success in understanding geothermal wells’ characteristics is the error value’s magnitude (training error and verification error). The target error in this study is <0.01.

2. Field Description

The Patuha geothermal field is situated in the southern part of Bandung, West Java Province. Surface manifestations found in the Patuha geothermal field are fumaroles, hot springs, mud pools, and warm springs [6]. Based on the previous study [7, 8, 9, 10, 11], Patuha geothermal field is a vapor-dominated reservoir with a temperature of about 215 - 230°C.

The total geothermal potential in the Patuha prospect area (proven) is estimated at 190 MW [12]. PT Geodipa Energi manages the Patuha geothermal field. PT Geodipa is a State-Owned Company with the proportion of its shares is 93.3% owned by the Ministry of Finance, the Republic of Indonesia, and 6.7% is owned by PT PLN (Persero). In 2014, PT Geodipa Energi completed the construction of Patuha Geothermal Power Plant Unit 1 with a net capacity of 60 MW. PT Geodipa is currently planning the development of Patuha geothermal power plant Unit 2 and 3 with 55 MW each [13].

In general, Patuha geothermal is a steam dominated field with steam fractions above 98% except for the PPL-02 production well. PPL-02 has a lower dryness than other wells, presumably because of a mechanical problem, namely leakage of cement and casing, which causes cold fluid to enter the well. A summary of production well outputs is shown in Table 1.

| No | Well  | Steam output [ton/hour] | T Reservoir [°C] | Brine Output [ton/hour] | Dryness [%] |
|----|-------|-------------------------|-----------------|------------------------|-------------|
| 1  | PPL-01| 71                      | 216             | Small                  | 99.1        |
| 2  | PPL-02| 32-34                   | 238             | 14-16 (172°C)          | 65-70       |
| 3  | PPL-2A| 45.8                    | 232             | Small                  | 99.12       |
| 4  | PPL-03| 55.9                    | 232             | Small                  | 99.2        |
| 5  | PPL-3A| 15.3                    | 229             | Small                  | 97.3        |
| 6  | PPL-3B| 79.9                    | 226             | Small                  | 99.7        |
| 7  | PPL-04| 5.32                    | 241             | Small                  | 97.0        |
| 8  | PPL-05| 20.9                    | 224             | Small                  | 99.8        |
| 9  | PPL-06| 7.1                     | 221             | Small                  | 99.13       |
| 10 | PPL-07| 28.1                    | 227             | Small                  | 98.8        |

TOTAL 361.32  15
Steam produced at Patuha geothermal field is then distributed to a steam header and then to a 60 MW power plant. Production from the wells is used optimally, including production at PPL-02 well, which produces some brine that requires steam purification through the flashing process. The steam is then used to drive a steam turbine coupled with an electric generator. The electricity output from the electricity generator is transmitted through the 150-kV transmission line owned by PT PLN (Persero) [13].

The sale of electricity from PT Geodipa Energi to PT PLN (Persero) is stated in the Power Purchase Agreement (PPA) between PT Geodipa Energi and PT PLN (Persero) for 30 years. Good reservoir management is needed to maintain the sustainability of production during the contract period. Production planning, production optimization, and good injection strategy are necessary to maintain geothermal field production sustainability.

3. Method
The research methodology is briefly illustrated in (Figure 1). The methodology consists of data preparation, a built artificial intelligence model, and model verification.

**Figure 1. Methodology.**
3.1. Data Preparation
1. The data used in this study are production data, i.e., wellhead pressure, wellhead temperature, reservoir pressure, wellhead valve opening, mass flow rate. The data were acquired based on operator records from January 2016 to January 2018. The data were normalized to value 0 to 1 using the maximum and minimum value of data to synchronize with the transfer function value of artificial intelligence.
2. After the data acquisition is made, the correlation analysis is carried out to determine which data are related to each other and their correlation. This process was done to determine artificial intelligence input and output data.
3. Validation of input and output data is done by Lag Plot evaluation. The evaluation is carried out to determine the continuity of the existing data.
4. Input data are production data obtained from the field. These data are then processed and determined as input and output parameters on artificial intelligence. Input and output data used to train the artificial intelligence is divided into four, i.e., 50%, 60%, 80%, and 100% of total field measurement data. The remaining data for each scheme is then used for verification, as described in (Figure 2).

![Figure 2](image-url). Scenarios of data usage for training and verification.

3.2. Built Artificial Intelligence Model
The artificial intelligence model uses artificial neural network feed-forward backpropagation methods with Python as a programming language, Numpy as an interface for matrix operations, Matplotlib as an interface for graphical operations Neurolab for artificial intelligence interpretation. This model includes determining how many input and output neurons, how many layers, the amount of initial weight, determining the type of activation function, and the artificial neural network training algorithm. The configuration of artificial intelligence is shown in Table 2.

| Parameter                    | Remark                                           |
|------------------------------|--------------------------------------------------|
| Artificial Intelligence Method| Feed-forward neural network                      |
| Network configuration        | 4 (four) neurons in the input layer, 8 (eight) neurons in the hidden layer, 1 (one) neuron in the output layer |
| Activation function          | Tan-sigmoid                                      |
| Initial weight factor        | Random number                                    |
| Maximum epoch                | 2000                                             |
| Sum square error (target)    | 0.001                                            |
| Training algorithm           | Newton conjugate gradient method                 |
3.3. Model Verification

The model verification process is carried out by providing input to the model created to produce the model’s output. The validation process uses 20%, 40%, 50%, and 100% of the field measurements’ total data. The output from the model is then compared with measurement data obtained in the field. Comparisons are made by calculating the model data’s deviation with field measurement data that calculate using mean square error (MSE).

4. Result

This study was conducted at the PPL-01 Geodipa geothermal field well. The data obtained is based on records carried out by Patuha geothermal field operators. The primary data used in this study are:

- a. Steam rate/M (ton/hr)
- b. Wellhead pressure/WHP (bar)
- c. Wellhead Temperature/WHT (°C)
- d. Reservoir Pressure/PR (bar)
- e. Wellhead valve position/%CV (%)

The data obtained is then plotted on the graph shown in Figure 3.

![Cross plot diagram for measured data of PPL-01 well.](image)

Figure 3: Cross plot diagram for measured data of PPL-01 well.

Some data in the cross-plot diagram show a zero value because the operators might not record it or an error in the measurement tool. Correlation analysis using the Pearson Correlation method is shown in Table 3a. Based on the cross-plot diagram (Figure 3) and correlation analysis Table 3a show that the steam flow rate (M) has a high correlation with % CV and reservoir pressure. Correlation analysis is developed by deleting data with zero values shown in the circle dash marker on the cross-plot diagram. The correlation analysis results using the Pearson correlation method show that the mass flow rate (M) has a high correlation to %CV, wellhead temperature, and reservoir pressure.
Table 3. Pearson correlation matrix (a) before cleansing data (left) and (b) after cleansing data (right).

|       | %CV   | WHP   | WHT   | PR    | M    | %CV   | WHP   | WHT   | PR    | M    |
|-------|-------|-------|-------|-------|------|-------|-------|-------|-------|------|
| %CV   | 1     | -0.61 | -0.46 | 0.52  | 0.56 | 1     | -0.69 | 0.2   | 0.89  | 0.89 |
| WHP   | -0.61 | 1     | 0.72  | -0.81 | -0.91| WHP   | 1     | 0.18  | -0.67 | -0.82|
| WHT   | -0.46 | 0.72  | 1     | -0.43 | -0.54| WHT   | 0.2   | 0.18  | 1     | 0.17 |
| PR    | 0.52  | -0.81 | -0.43 | 1     | 0.98 | PR    | 0.89  | -0.67 | 0.28  | 1    |
| M     | 0.56  | -0.91 | -0.54 | 0.98  | 1    | M     | 0.89  | -0.82 | 0.17  | 0.97 |

Lag plots can be used to check for data without a pattern or randomness, serial correlation, and the possibility of analysis in time series. The lag plot of steam production data is generated to confirm the data continuity. Figure 4 shows data continuity in a linear trend. The linearity of data is suitable to apply in this study.

![Lag Plot](image)

Figure 4. The lag plot of mass flow rate data on the PPL-01 production well.

Based on the data analysis that has been performed, the selected data can be used as an input and output for artificial intelligence (Table 1). The data were normalized to value 0 to 1 using the maximum and minimum value of each data to synchronize with the transfer function value of artificial intelligence. In this study, we use tan-sigmoid transfer functions with a value between -1 and 1. The selection of normalized data range value and transfer function can affect artificial intelligence’s speed in the learning process.

Table 4. Input and output parameters.

| Input                        | Output          |
|------------------------------|-----------------|
| Wellhead pressure/WHP (bar)  |                 |
| Wellhead Temperature/WHT (°C)| Steam rate      |
| Reservoir Pressure/PR (bar)  |                 |
| Wellhead valve position/%CV (%)|                |

Artificial intelligence is set up in 4 scenarios to determine how artificial intelligence can understand PPL-01 production wells' characteristics. The difference between the 4 (four) scenarios is the percentage of data used for the artificial intelligence training process Table 5. This scenario includes 50%, 60%,
80%, and 100% of the total data used for the artificial intelligence training process, while the rest of the training process data is used to verify.

Table 5. Scenarios for each study.

| Parameter          | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 |
|--------------------|------------|------------|------------|------------|
| Training Data      | 50% Data   | 60%        | 80%        | 100%       |
| Verification Data  | 50% Data   | 40%        | 20%        | 100%       |
| Input Parameter    | Wellhead Pressure, Wellhead Temperature, Reservoir Pressure, %CV | Flow Rate |
| Output Parameter   |            |            |            |            |
| Network            | 4 (Input), 8 (Hidden Layer), 1 (Output) |            |
| Activation Function| Tan-Sigmoid |            |            |
| Initial weight factor| Random number |            |
| Epoch              | 2000       |            |            |
| Sum Square Goal    | 0.001      |            |            |
| Training Algorithm | Newton-Conjugate gradient method |            |

After preparing data and configuring artificial intelligence through programming in Python, the training process is carried out with the planned scenarios. The training process parameters’ constraints are the number of epochs and sum square goals. In this study, the epoch number is 2000, and the sum square goal is 0.001. The training process will run according to the specified epoch and sum square goals. If it has reached the specified sum square goal, then the training process will be terminated, and if it has reached the specified epoch while the sum square error has not reached the specified target, the training process will be stopped. Determination of the number of epochs is based on trials wherein the trial process can be known if at a specific value the value of the sum square error is constant to the change in the epoch, then it is assumed that there has been no change in the value of the sum square error so that the change in the epoch does not affect the change in the sum square error significantly. The value of the sum square error in each scenario is shown in (Figure 5).
Figure 5 shows the training error progress for each scenario (Scenario 1, 2, 3, and 4).

Figure 5 (a) and (d) shows that the error value has a constant value at the epoch close to 250. Figure 5 (c) and (d) shows that the error value has a constant value when the epoch is < 250. It happens because the training data used in Scenario 1 Figure 5 (a) has a narrower range of data compared to Scenarios 2, 3, and 4 Figure 5 (b), (c), and (d). The difference will be more clearly seen in the plot of the measurement value chart and the artificial neural network model’s simulation results shown in (Figure 6).

One of the challenges in this study is that the initial weight factor value on the artificial neural network is generated automatically by Python. The initial weight factor value is a random number and affects the speed of artificial neural network training. Figure 10 shows that artificial neural networks can understand the pattern of steam mass flow rate data from PPL-01 well in general. It is expressed in the comparison between the actual value (measurement) with the simulation value.
Figure 6. Training and simulation result.

Based on training and simulation results, 50% of the training process data shows unfavorable results than other scenarios. The considerable mean square error value also indicates this at the verification stage. The more the percentage of data used in the training process, the less the mean square error value during the verification process. The mean square error values for each scenario are shown in Table 6.
Table 6. The error of training and simulation result.

| Parameter                        | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 |
|----------------------------------|------------|------------|------------|------------|
| Training error (sum square error)| 0.0011     | 0.0044     | 0.0076     | 0.0038     |
| Verification error (mean square error) | 0.0459     | 0.0042     | 0.0014     | 0.000001   |

5. Conclusion

Studies in this paper show that artificial neural networks succeed in understanding PPL-01 geothermal well steam production characteristics. The proposed method using wellhead valve position, wellhead pressure, and temperature data was successfully applied in the artificial neural network training and verification process with an error value below 0.01 as the target of this study. The best result in the training process, regarding the training error, is Scenario 1 with configuration 50% data as training data and 50% data as verification data. However, Scenario 1 shows a considerable error value compared to other scenarios because the distribution of input and output data in Scenario 1 is relatively narrower than other scenarios (Figure 9). Based on the studies, the minimum proper data for artificial neural networks is 60% of the training process data and 40% of the verification process data. Alternatively, on the other hand, the artificial neural network can adequately predict 40% of total data in a given period.

References

[1] Noshi C I, Eissa M R, Abdalla R M and Schubert J J 2019 An Intelligent Data-Driven Approach for Production Prediction, OTC-29243-MS
[2] Olivares G, Escalona C and Gimenez E 2012 Production Monitoring Using Artificial Intelligence, APLT Asset SPE 149594, 27–29
[3] Serpen G., Palabiyik Y. and Serpen U 2009 An Artificial Neural Network Model for Na/K Geothermometer Proc., 34th Workshop on Geothermal Reservoir Engineering, (Stanford: Stanford University)
[4] Priyangga H Y and Ruliandi D 2018 Application of Pattern Recognition and Classification Using Artificial Neural Network in Geothermal Operation. Proc., 43rd Workshop on Geothermal Reservoir Engineering (Stanford: Stanford University)
[5] Ariturk M S 2019 Optimizing the Production and Injection Wells Flow Rates in Geothermal Field Using Artificial Intelligence, West Virginia University
[6] Elfina 2017 Updated Conceptual Model Of The Patuha Geothermal Field, Indonesia United Nations University Geothermal Training Programme Reykjavik
[7] Ashat A and Pratama H B 2018 Application of experimental design in geothermal resources assessment of Ciwidey-Patuha, West Java, Indonesia IOP Conf., Series: Earth and Environmental Science 103
[8] Ashat A, Pratama H B and Itoi R 2019 Comparison of resource assessment methods with numerical reservoir model between heat stored and experimental design: Case study Ciwidey-Patuha geothermal field IOP Conf., Series: Earth and Environmental Science 254
[9] Ashat A, Pratama H B and Itoi R 2019 Updating conceptual model of Ciwidey-Patuha geothermal using dynamic numerical model IOP Conf., Series: Earth and Environmental Science vol 254
[10] Ashat A, Ridwan R H, Judawisastra L H, Situmorang J, Elfajrie I, Atmaja R W, Iskandar C and Ibrahim R F I 2019 Conceptual Model and Numerical Simulation Update of Patuha Geothermal Field, West Java, Indonesia Conceptual Model and Numerical Simulation Update of Patuha Proc., 41st New Zealand Geothermal Workshop Auckland, New Zealand
[11] Pratama H B, Widiatmo J S and Ashat A 2020 Numerical Investigative Modeling of Changes Within the Patuha Geothermal Reservoir and Its Production Sustainability Under Two Different Conversion Technologies Natural Resources Research
[12] EBTKE 2017 Potensi Panas Bumi Indonesia Jilid I Jakarta, Indonesia
[13] Geodipa 2019 Paparan PT Geodipa Energi, Mengenal Lapangan Panas Bumi Patuha (unpublished)