Proceedings World Geothermal Congress 2026, Calgary, Canada, June 8—11, 2026

Implementation of Machine Learning to Predict Cooling Tower Capability at Lahendong Geothermal Power Plant

Fikrul Jihad, Agung Dwi Susanto, Budi Wigiarto
PT Pertamina Geothermal Energy Tbk. Area Lahendong

INTRODUCTION

The cooling tower is one of the many critical equipment at Lahendong Geothermal Power Plant (GPP) Unit 5&6. Its function is to cool the hot water coming out of the condenser. To ensure optimal cooling tower capability, Pertamina Geothermal Energy (PGE) regularly conducts performance test based on CTI ATC 105 done by certified third-party company.

There are several independent variables to calculate cooling tower capability such as Wet Bulb Temperature (WBT in °C), Dry Bulb Temperature (DBT in °C), Hot Water Temperature (HWT in °C), Cold Water Temperature (CWT in °C), and Water Flow Rate (Q in l/s). As the result of the calculation and also acts as the dependent variable, there is cooling tower capability in percentage (%) unit.

In model development phase, we use data obtained from the performance test conducted in March 2025 to achieve a model with a very low Root Mean Square Error (RMSE). The model was built using MATLAB, which provides a Regression Learner feature to create models and select the best one with the lowest RMSE.

The model development used 840 independent data points and 168 dependent data points that were obtained from measurements during the performance test. We also used the developed model to calculate the cooling tower capability by using the data from Distributed Control System (DCS). Finally, we compare both results and see if the DCS data can be used to predict cooling tower capability with the developed model.

METHODS

Machine learning is a rapidly developing field of study. One of the most commonly used types of machine learning is linear regression. Linear regression is a method used to create a model that can predict an output value using a new set of data.

Linear regression consists of two types of data: a dependent variable and an independent variable. The dependent variable is the output (result), which strongly depends on the value of the independent variable (input). The dependent variable is the value to be predicted using the model.

The Shapley value is a concept introduced by Lloyd Shapley in 1953. This concept is used as a method in cooperative game theory to fairly distribute rewards among players based on their respective contributions. In our application, the Shapley value is used to identify which variables play the most important role in the model we have developed.

Mean Absolute Percentage Error (MAPE) is an evaluation metric used to calculate the average percentage difference between predicted and actual values. In other words, MAPE measures the average prediction error as a percentage of the actual value. In this case, MAPE is used to validate whether the model has a large or small prediction error.

RESULTS

Figure 1 shows the best models we obtained, where the lowest RMSE value of 0.162 was achieved using the Medium Neural Network method.

The cold water temperature plays an important role in determining the cooling tower capability value. In addition, we also performed statistical analysis using the Pearson correlation, which showed that the relationship between cold water temperature and cooling tower capability is a strong negative correlation, with a Pearson correlation coefficient of –0.99332.

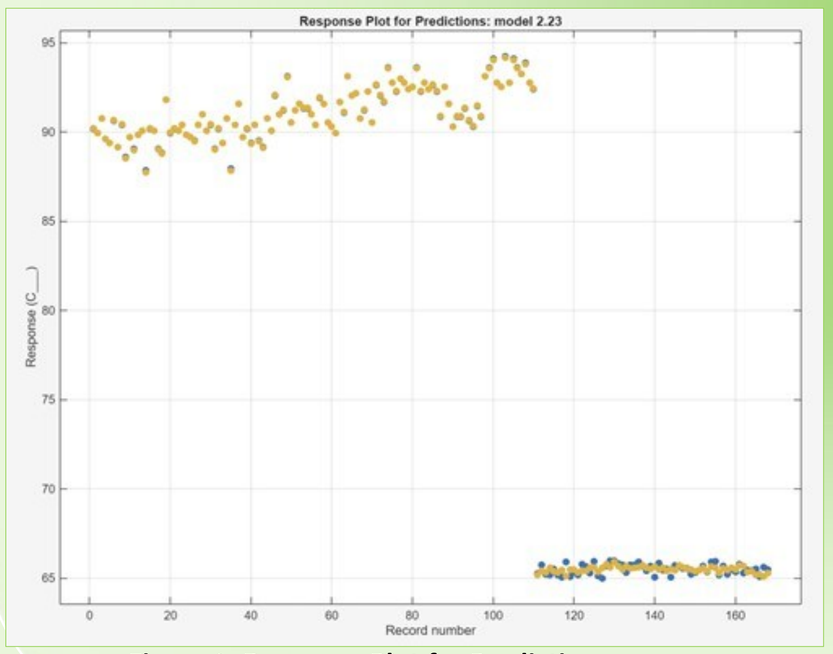


Figure 1. Response Plot for Predictions.

In Figure 2, we compared the generated result from performance test data with the DCS data to determine whether the DCS data can be used to estimate the cooling tower capability.

The average MAPE value of the DCS data compared to the calculation data was 6.57%, with a maximum MAPE of 9.4% and a minimum of 4.04%.

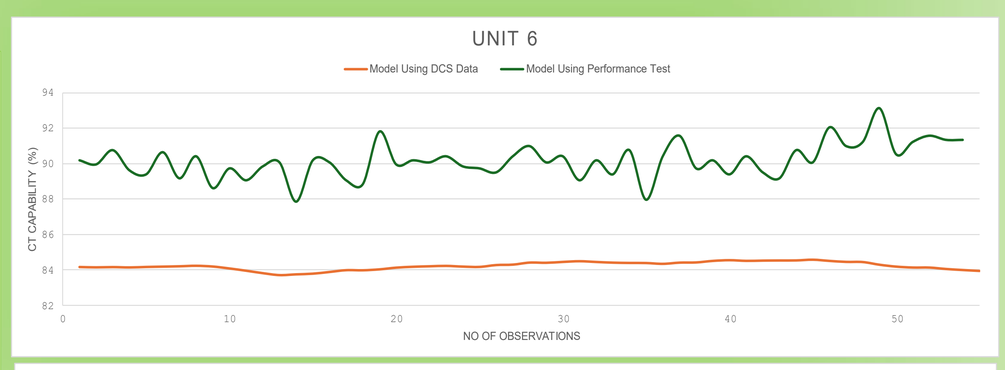


Figure 2. Comparation between model using performance

test data and DCS data from GPP Unit 6

As comparison, Figure 3 shows the generated result for Unit 5 by using DCS data and the developed model, where the average MAPE obtained was 11.97%, with a maximum value of 14.77% and a minimum value of 8.84%.

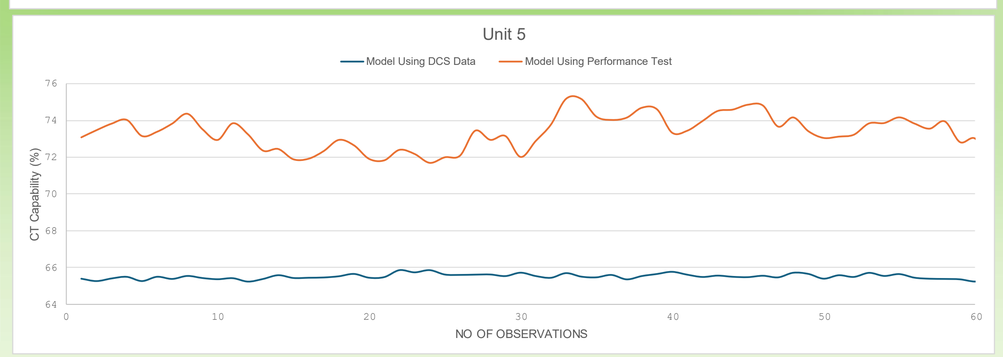


Figure 3. Comparation between model using performance

test data and DCS data from GPP Unit 5

DISCUSSION

The experimental study was conducted in two phases. The first phase is to make model by using performance test data taken during the period of March 2025 at Cooling Tower Unit 5&6. We use those data and regression learner from MATLAB to make a model.

The second phase is we used the developed model to predict dependent variable which is the cooling tower capability by using the sample data from DCS that collected at the same period as the performance test data.

The result shows the Medium Neural Network has the best RMSE around 0.162 so we use this model to predict the cooling tower capability. Furthermore, the results generated from both phases and data show the same pattern even though the DCS data is taken from different sampling points compared to performance test data.

CONCLUSION

Prediction of cooling tower capability by using calculated data and predicted data from DCS for Unit 6 has a better result, with a MAPE of 6.57%, indicating a very good model. On the other hand, Unit 5 has a different result, with a MAPE of only 11.97%, indicating a moderate model.

Several factors influenced these results:

1. Different sampling point or source data between performance test data and DCS data.
2. Cold water temperature (CWT) had poor data during testing. The data changes due to changes in environmental temperature, not due to operational processes, so to get a more accurate result, we have to repair the temperature transmitter. It's important to remember that CWT plays a critical role in determining cooling tower capability.
3. We can also improve the result if we have dedicated flow meter to measure the water flow coming to the cooling tower.

Acknowledgements

We would like to acknowledge the support from our management at PT Pertamina Geothermal Energy Tbk. Area Lahendong, our families, and friends.

We also appreciate the assistance of SPIG and PT Babcock & Wilcox Asia during performance test at Lahendong Cooling Tower Unit 5&6.

REFERENCES

1. Cooling Technology Institute, Acceptance Test Code for Cooling Towers, CTI ATC-105 (2022).
2. SPIG, Performance Test Report Cooling Tower Unit 5&6 (2025).
3. Application of Artificial Neural Network to Exergy Performance Analysis of Geothermal Power Plant (2020)
4. The Implementation of multiple Linear Regression for Swimming Pool Facilities: Case Study at Joa, Norway
5. Linear Regression With MATLAB: A Tutorial Introduction to the Mathematics of regression Analysis