Prediction of Electricity Production in Micro-Hydro Power Plants (PLTMH) Using LSTM, and Linear Regression Models

Sandhi Ading Wasana, Rinaldi Munir.
Email: adingism@gmail.com, rinaldi-m@stei.itb.ac.id
School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia

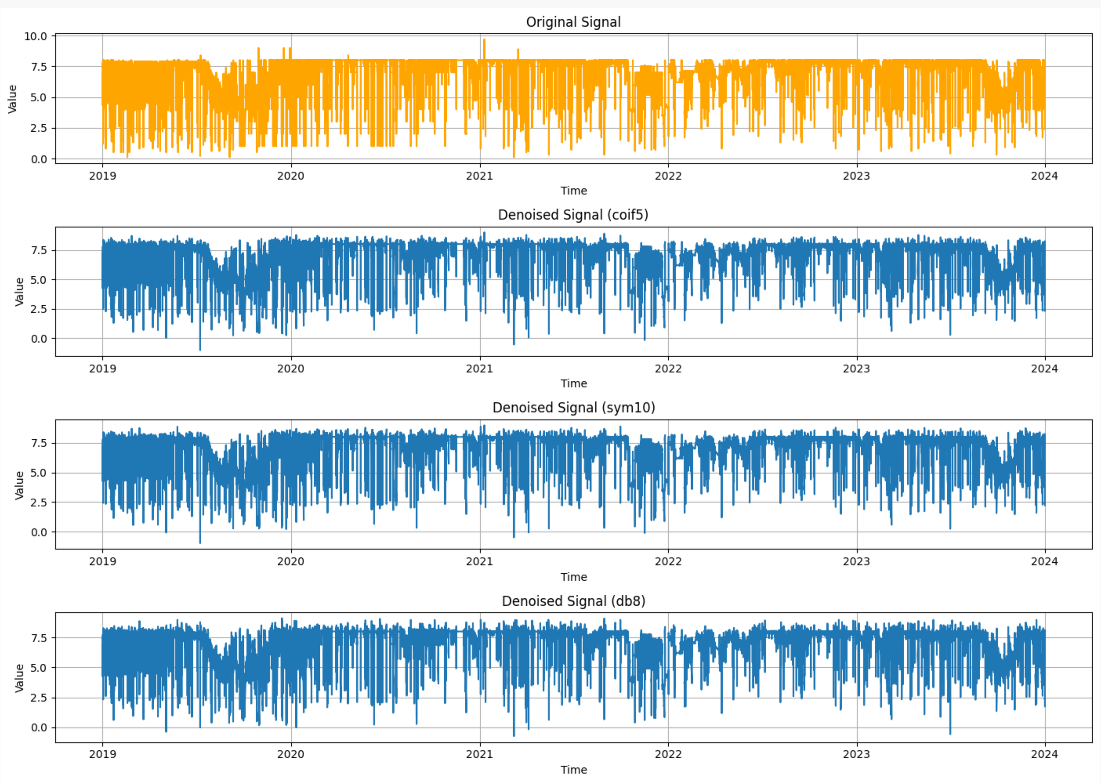
Abstract

The demand for clean energy from renewable sources has become a global challenge to combat climate change and global warming. Indonesia has committed to achieving 31% utilization of renewable energy by 2050. Among the potential renewable energy sources, Microhydro Power Plants (PLTMH) stand out due to their capability to harness Indonesia's abundant water resources, with an estimated potential of 75 GW, of which 26.3 GW is viable for development. However, PLTMH electricity production faces significant challenges due to its dependency on environmental factors, such as weather and water flow rates, leading to instability compared to fossil-based power generation. Accurate forecasting of electricity production is crucial for optimizing the integration of PLTMH into the grid and supporting sustainable energy adoption.
This research explores electricity production forecasting for PLTMH using machine learning methods, including LSTM, and Linear Regression. By combining historical production data with weather parameters, such as rainfall, this study aims to improve forecasting accuracy and provide valuable insights for the planning and management of renewable energy resources. The findings will aid PLN in maximizing the absorption of environmentally friendly electricity from PLTMH, contributing to Indonesia's renewable energy transition.

Dataset

Electricity Production Dataset
The dataset is recorded daily at 30-minute intervals. The unit of measurement used in the data is Megawatts (MW). Two micro-hydro power plants (PLTMH) are used as training data: Lubuk Gadang and Sangir Hulu. The data spans from January 1, 2019, to December 31, 2023.

Rainfall Data
Rainfall data is recorded every 10 minutes. The data is measured in millimeters (mm) and captures rainfall occurrences in the area surrounding the power plant.



Example of Data

Metodology

Data Collecting
Combining two datasets from different sources requires careful alignment and handling to ensure consistency and compatibility. Below is a step-by-step guide to combine your electricity production and rainfall data

Data Preparation
Perform smoothing using wavelets to reduce noise in the electricity production dataset.

Modeling
Modeling is conducted using a multilayer LSTM for the combined dataset of electricity production and weather data. In the linear regression model, only electricity production data is used.

Training
Training is performed with several hyperparameter tuning adjustments.

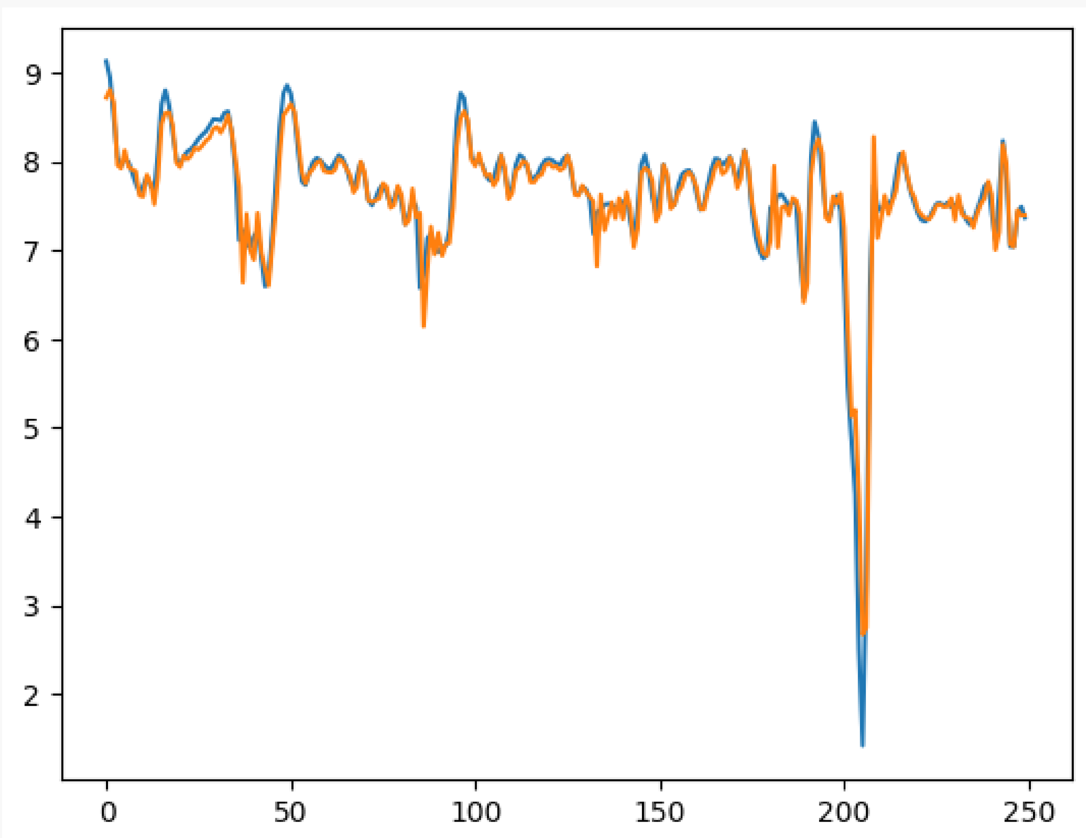
Evaluation
The evaluation of the training results is conducted using RMSE, R², and MAE metrics.

Implementation And Result

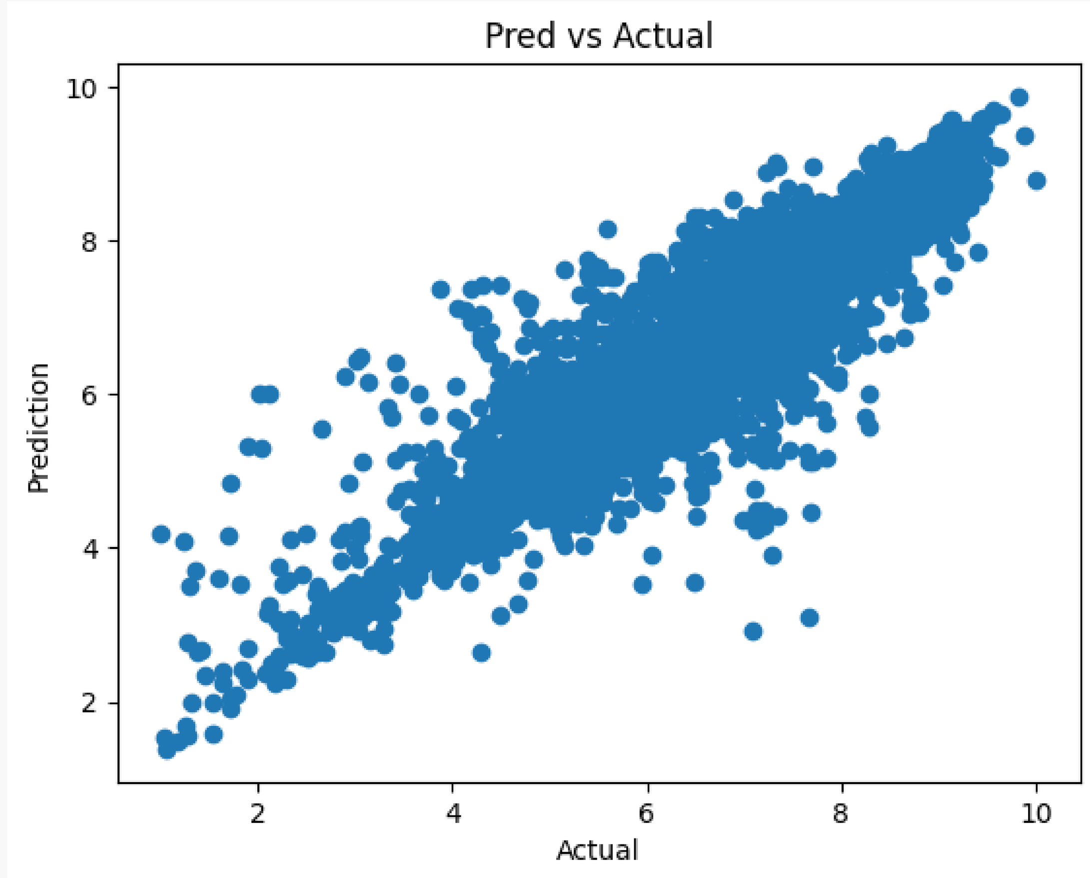
In this study, a total of 128 experiments were conducted. These experiments involved hyperparameter tuning. The dataset combination was performed using both electricity production data without weather data and a combination of both datasets. The following is an example of the models that were developed:

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_2 (LSTM)               (None, 12, 256)           265216

 dropout_3 (Dropout)         (None, 12, 256)           0

 lstm_3 (LSTM)               (None, 128)               197120

 dropout_4 (Dropout)         (None, 128)               0

 dense_2 (Dense)             (None, 32)                4128

 dropout_5 (Dropout)         (None, 32)                0

 dense_3 (Dense)             (None, 1)                 33

=================================================================
Total params: 466497 (1.78 MB)
Trainable params: 466497 (1.78 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

The following is a sample data test comparing actual data and prediction data. The graph shows that the prediction results closely align with the actual data.



The following is a plot graph comparing prediction data and actual data. The graph shows results concentrated along the x=y axis. However, there are some data points that are not accurately predicted, contributing to minor errors. This indicates room for further improvement in future research and development.



The following is a summary table of the experimental results. The table shows that incorporating weather data into the dataset can improve the accuracy of the model in predicting electricity production.

No	Pembangkit	RMSE		R2		MAE	
		a	b	a	b	a	b
1	Sangir Hulu	0.404	0.313	0.912	0.947	0.291	0.187
2	Lubuk Gadang	0.312	0.272	0.895	0.921	0.149	0.118

a: Data Production
b: Data Production + Rainfall

The results in the table above indicate that combining dataset smoothing to reduce noise and incorporating weather data can significantly improve the model's accuracy.

Conclusion

This study reinforces the importance of integrating external environmental factors, such as rainfall data, into predictive models for electricity production in micro-hydro power plants (PLTMH). The results clearly demonstrate that including rainfall data significantly enhances the model's performance across all evaluation metrics. Specifically, for Sangir Hulu, the RMSE decreased from 0.404 to 0.313, and the R² increased from 0.912 to 0.947, indicating a substantial improvement in prediction accuracy. Similarly, for Lubuk Gadang, incorporating rainfall data reduced the RMSE from 0.312 to 0.272 and improved the R² from 0.895 to 0.921.The inclusion of rainfall data not only reduced prediction errors but also resulted in lower MAE values, showcasing the robustness and reliability of the combined dataset approach. These findings emphasize that accurate predictive models, which incorporate relevant environmental variables, can significantly contribute to optimizing renewable energy production and improving grid reliability.This research highlights the potential of advanced machine learning techniques like LSTM when combined with essential environmental data, paving the way for the development of highly accurate predictive models that support the transition to a sustainable energy future.

References

Adiwarman, A. S., Pahlevi, R. R., and Nuha, H. H. (2023): Microhydro Power Plants with IoT-based Monitoring and Prediction using Linear Regression, 2023 International Conference on Artificial Intelligence, Blockchain, Cloud Computing, and Data Analytics, ICoABCD 2023, Institute of Electrical and Electronics Engineers Inc., 7–12.

Cao, R., Ye, H., Fang, Z., Chen, L., and Zhang, S. (2023): A Regional Monthly Electricity Consumption Forecast Method Based on HP Filter and Seq2Seq, 2023 IEEE 6th International Conference on Automation, Electronics and Electrical Engineering, AUTEEE 2023, Institute of Electrical and Electronics Engineers Inc., 515–521.

Chi, D. (2022): Research on electricity consumption forecasting model based on wavelet transform and multi-layer LSTM model, Energy Reports, 8, 220–228.

Gaboitaolelwe, J., Zungeru, A. M., Yahya, A., and Lebekwe, C. K. (2022): Electricity Load Prediction Using Machine Learning, 2022 International Conference on Smart Applications, Communications and Networking, SmartNets 2022, Institute of Electrical and Electronics Engineers Inc.

Marino, D. L., Amarasinghe, K., and Manic, M. (2016): Building Energy Load Forecasting using Deep Neural Networks. Prakash, S. A., Shah, D., Jayavel, K., and Mtonga, K. (2022): Hydropower Energy Generation Prediction Model: A Machine Learning Approch, 2022 International Conference on Computer Communication and Informatics, ICCCI 2022, Institute of Electrical and Electronics Engineers Inc.

Rahayu, L. N., and Windarta, J. (2022): Tinjauan Potensi dan Kebijakan Pengembangan PLTA dan PLTMH di Indonesia, Jurnal Energi Baru Dan Terbarukan, 3(2), 88–98.