Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.2, April 2025, pp. 333~340
e-ISSN: 2548-6861

Long Short-Term Memory as a Rainfall Forecasting Model for Bogor City in 2025-2026

Nur Anggraini Fadhilah [1]*, Muhammad Abshor Dzulhiij Rizki[2]*, Muhammad Ryan Azahran[3*], Siti Arbaynah[4*], Rakesha Putra Antique Yusuf[5*], Yenni Angraini[6*], Muhammad Rizky Nurhambali[7*]
* Statistika dan Sains Data, IPB University
873anggraini@apps.ipb.ac.id [1], ipbabshor@apps.ipb.ac.id [2], starbaynah@apps.ipb.ac.id [4], y_angraini@apps.ipb.ac.id [6], rizkynurhambali@apps.ipb.ac.id [7]

Article Info

Article history:

Received 2025-01-09
Revised 2025-03-10
Accepted 2025-03-13

Keyword:

Cross-validation,
Forecasting,
LSTM,
Rainfall.

ABSTRACT

Indonesia is a country with a tropical climate that has unique and changing weather patterns. Accurate rainfall prediction can help local governments, farmers, and the broader community plan activities that depend on rainfall patterns. This research aims to develop a rainfall prediction model for Bogor City using past rainfall data in Bogor City, which is known as an area with high rainfall levels and dynamic rainfall patterns. The analysis utilizes rainfall data recorded by the JAXA satellite from January 1, 2014, to December 31, 2024. The prediction method implemented in this research is the long short-term memory (LSTM). The LSTM modelling process evaluates various models by comparing RMSE, MAE, and correlation values through expanding window cross-validation, selecting the model with the lowest average RMSE and MAE with the highest correlation as the optimal choice. The best-performing model was achieved with 25 epochs and a batch size of 1, resulting in an average RMSE of 56.3340, MAE of 35.5223, and correlation of 0.3209. This best-performing model is then employed to predict rainfall for the next two years. The results show significant daily variations in the predicted rainfall but can capture existing seasonal patterns.

I. Introduction

Indonesia is a country with a tropical climate that has unique and changing weather patterns, as well as two main seasons, namely the dry season and the rainy season, which alternate every six months. However, global climate change over the past few years has disrupted the seasonal cycle, making rainfall patterns increasingly difficult to predict [1]. To support water resource management and natural disaster mitigation, such as floods and droughts, the development of rainfall prediction models is very important. Accurate rainfall prediction can help local governments, farmers, and the wider community in planning activities that depend on rainfall patterns.

Weather predictions, including rainfall, can be made using a time series method that utilizes historical data to project future values through specific mathematical models [2]. ARIMA models typically excel at capturing and predicting geographic seasonal trends in stable historical datasets. However, they often struggle with forecasting unforeseen changes. To address these limitations, deep learning techniques can be employed. Deep learning uses end-to-end concepts [3], where learning layers in deep learning make accuracy and performance better than other algorithms [4]. One of the deep learning methods commonly used for forecasting time series data is long short-term memory (LSTM). LSTM has hyperparameters that influence the reliability and performance of the model [5]. LSTM-based models have gained significant attention due to their ability to model nonlinear and temporal dependencies effectively. For example, [6] integrated LSTM with the WRF-Hydro model to improve streamflow predictions. Their research highlighted that LSTM significantly outperformed traditional statistical approaches, particularly in handling the nonlinearities and spatial-temporal complexities inherent in hydrological data. This underscores the potential of LSTM for rainfall prediction tasks, where capturing intricate temporal dynamics is crucial for accurate forecasting and decision-making.

Such findings reinforce the relevance of adopting LSTM for rainfall accumulation prediction in regions like Bogor City, characterized by dynamic and unpredictable rainfall patterns. The unique advantages of LSTM, such as its internal memory and ability to mitigate vanishing gradient problems, make it an ideal method to address the challenges posed by Bogor's high rainfall variability. This study builds upon prior advancements by applying LSTM to predict rainfall accumulation, with the goal of supporting flood mitigation efforts and optimizing water resource management in the area.

Long-short-term memory (LSTM) is one of the modifications of the Recurrent Neural Network (RNN) that was first introduced by Hochreiter and Schmidhuber in 1997 [7], which has the ability to capture long-term patterns in time-series data so that it is suitable for data with strong temporal dependencies, such as rainfall [8]. LSTMs are specifically designed to better model temporal sequences and long-term dependencies than conventional RNNs [9]. In its application, LSTM uses a repetitive backward learning method, which requires a long processing time [7]. However, with various gates in its architecture, LSTM is able to learn data over a very long period of time. The advantage of memory cells in LSTM is not possessed by the classical statistical prediction method. In addition, LSTMs have configurable hyperparameters to improve the reliability and performance of the model [5]. LSTMs have also proven effective for forecasting seasonal data that have diverse patterns. The unique internal memory in this type of RNN allows LSTMs to overcome the problem of vanishing gradients [10]. To optimize the performance of the LSTM model, the Adam optimizer is employed. Adam has been widely recognized for its computational efficiency and adaptability in managing sparse gradients and noisy data, which are typical characteristics of rainfall datasets [11]. By combining the benefits of AdaGrad and RMSProp, Adam offers faster convergence and improved stability, making it a reliable choice for deep learning tasks such as rainfall prediction.

This research aims to develop an LSTM-based rainfall prediction model for Bogor City. Bogor City is known as an area with high rainfall levels and dynamic rainfall patterns, so it is important to predict rainfall in this area as part of flood risk mitigation efforts and water resource planning.

II. Research Method

The research begins with data collection and continues to the data preprocessing stage. The creation of models from LSTM is adjusted according to its hyperparameters. Time series k-fold cross validation is also carried out for parameter optimization. The overall model will be evaluated according to this performance. All stages of this research were carried out using Phython 3 programming language through the Google Collab software application. The entire research flow can be seen in Figure 1.

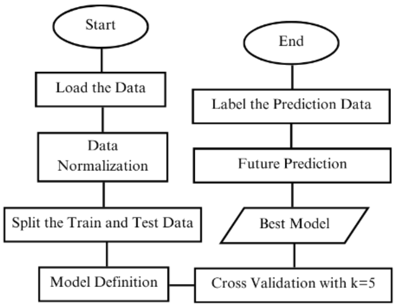

Figure 1. Research Flowchart

A. Data

The data used in this study is daily rainfall accumulation data for the Bogor area. The data is sourced from the Japan Aerospace Exploration Agency: Global Satellite Mapping of Precipitation (JAXA GSMaP) through the website https://sharaku.eorc.jaxa.jp/GSMaP/. The selection of data sources is based on [12], where GSMaP has been applied to many studies and has results similar to the observations of the Meteorology, Climatology, and Geophysics Agency (BMKG), which appear near real-time and have spatial accuracy by the GSMaP satellite making GSMaP increasingly used for research. Hourly data on the site from January 1, 2014, to December 31, 2024, was accumulated into daily data so that 4018 daily rainfall accumulation data were obtained.

B. Data Normalization

According to [13], data normalization is one of the processes carried out in the pre-processing phase. In normalization, the values are rescaled so that processing can be made easier. One of the normalization methods on the data is the MinMax Normalization technique by changing the actual value to an interval value between 0 and 1. Data normalization helps improve model performance by adjusting the weights in the model [14]. The equation for calculating MinMax Normalization, where x is a specific value to be normalized, x_{min} is the minimum value of an attribute, and x_{max} is the minimum value of an attribute.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Description:
x = a specific value to be normalized
x' = normalized result value
x_{min} = the minimum value of an attribute
x_{max} = the minimum value of an attribute.

C. Data Split

Data is divided into two, namely training data and testing data before processing further data. The data shift as an example in the 1st fold of the data train is rainfall data from day 1 to day 672 and continued with test data from day 673 to day 1341. In the 2nd fold, the data train starts from the 1st day to the 1341th day and continues with the test data from the 1342th to the 2010th day, which means that the train data and test data on the 1st fold become the data train on the 2nd fold, this is called overlapping data.

TABLE I
SPLITTING DATA

Fold	Data train	Data test
1	01/01/2014 - 03/11/2015	04/11/2015 - 02/09/2017
2	01/01/2014 - 02/09/2017	03/09/2017 - 03/07/2019
3	01/01/2014 - 03/07/2019	04/07/2019 - 02/05/2021
4	01/01/2014 - 02/05/2021	02/05/2021 - 02/03/2023
5	01/01/2014 - 02/03/2023	03/03/2023 - 31/12/2024

D. Model Definition

Construct four hyperparameter model scenarios with varying epoch values, while the batch size used remains constant at a value of 1 for all scenarios. Thus, the focus of the comparison lies on the variation in the number of epochs (25, 50, 75, and 100). The entire model scenario uses adaptive moment estimation (Adam) as an optimizer, allowing for more consistent control on aspects other than epoch. This scenario is prepared with reference to references from [15] for models with epochs 25, 50, and 100, and from [16] models with epochs 75. Each multiple of 25 epoch values is defined as Model 1 for epoch 25, Model 2 for epoch 50, Model 3 for epoch 75, and Model 4 for epoch 100. Other hyperparameters in this study are made default with the modules used.

E. k-Fold Cross Validation

After the LSTM models were trained to study rainfall data patterns, the models were stored and tested using test data through an expanding window cross-validation process. The selection of expanding window cross-validation is inseparable from rainfall that follows the season, so it has a seasonal pattern, as in [17]. In addition, according to [18], the expanding window method can capture long-term volatility patterns better than rolling window cross-validation. The time series k-fold cross-validation using expanding window with k=5 is implemented through the time step of 60. First, the dataset is divided into 5 overlapping folds like in Table 1. For each fold, the model uses the first $k-1$ folds as training data and the k-th fold as testing data. All models are tested on each fold created during this process. The performance of the models is then evaluated using the root mean squared error (RMSE), mean absolute error (MAE), and correlation (r) between actual test data and prediction test data to measure their accuracy. This process ensures a comprehensive assessment of model performance across different data splits.

F. Best Model

The performance of all models in each fold is evaluated after the epoch process by calculating the average RMSE, MAE, and correlation value across all fold of time series 5-fold expanding window cross-validation. The model with the best hyperparameters is determined as the one with the smallest average RMSE and MAE value with highest correlation [19]. The best model is the most consistent among the three evaluation metrics.

G. Future Prediction

After identifying the best model, forecasting is carried out. However, before proceeding, the data must be re-normalized to revert it to its actual values. This is necessary because the forecasting results are still in a normalized range of 0 to 1. An inverse transformation is applied using the formula:

$$x = \left(x_{normalized} * (max - \min)\right) + min \qquad (2)$$

The rainfall forecast is then performed for the next six months using the best model derived from the analysis.

H. Label the Prediction Data

Once the forecasted data is obtained, labels are assigned based on rainfall intensity categories defined by Indonesian Meteorological, Climatological, and Geophysical Agency (BMKG): 0 mm (cloudy or clear), 0.5–20 mm (light rain), 20–50 mm (moderate rain), 50–150 mm (heavy rain), and >150 mm (extreme rain). This labeling provides meaningful interpretation of the forecasted rainfall data.

III. RESULT AND DISCUSSION

A. Data Exploration

Data exploration is used to determine the pattern and distribution of data to be used in this research. Exploration is an important stage to think about what analysis should be done. Based on Figure 1, the distribution of rainfall data does not have a specific pattern, that is, it does not show a consistent pattern or does not show the presence of seasonal components. The volume of rainfall increases in certain periods, most likely during the rainy season, with significant peaks at the beginning of the year, such as in January 2023 and 2024.

Reporting from [20] **CNN** Indonesia on October 10, 2023, stated that the BMKG predicts the beginning of the rainy season in Jakarta's buffer cities such as Bogor, Depok, Tangerang, and Bekasi starting in October 2023 and reaching its peak until January 2024. In the plot there is a space that is empty due to a zero rainfall volume between the end of July and the beginning of October in 2023, as reported by [21] **Pakuanraya**, the Head of the BMKG Dramaga Bogor Climatology Station, Rahmat Prasetya explained that the peak of the dry season in 2023 occurred from August to October 2023 due to strong easterly wind conditions and reported by [22] **Kompas**, BMKG stated that the temperature or temperature anomaly in the Pacific Ocean at that time was increasing and had reached 0.8, if the temperature had touched 1, this condition could be said to be a Moderate El Nino and Indian Ocean Dipole (IOD) towards a positive phase that occurred from July to October 2023. [23] in her research stated that the ENSO (El Niño–Southern Oscillation) and IOD (Indian Ocean Dipole) phenomena have a significant impact on seasonal rainfall variations in Indonesia. Historical analysis indicates a reduction in rainfall ranging from 20% to 50% of the normal value in areas sensitive to ENSO, particularly during the months of July to October. This highlights the critical influence of these climatic phenomena on Indonesia's rainfall patterns, with implications for water resources, agriculture, and disaster management in the affected regions.

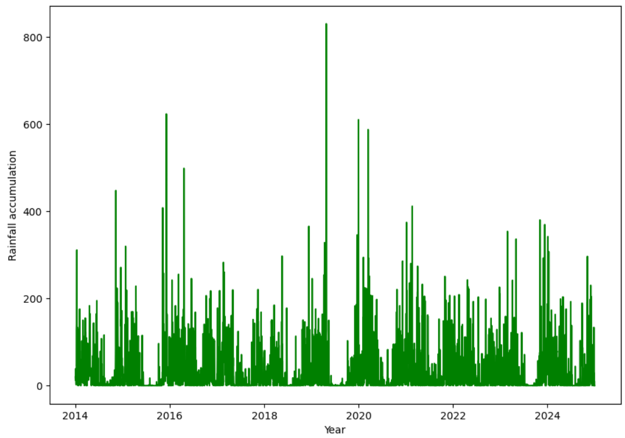

Figure 2. Bogor Daily Rainfall 2014 –2024 Graph

The characteristics in the data can be seen with their descriptive statistics. Descriptive statistics include mean, median and standard deviations. Table 2 shows that the average of rainfall accumulation in the past ten years is 32.8764 mm. In addition, the standard deviation in this data is 56.8390, this is because the value of rainfall data has varying values. This standard deviation can be used as a benchmark for evaluating the accuracy of the model. A good model is one that has a smaller RMSE value than the standard deviation value [24].

TABLE II
DESCRIPTIVE STATISTICS

Variable	Rainfall
Minimum	0.0000
Maximum	829.7251
Mean	32.8764
Median	7.1981
Standard Deviation	56.8390

B. Data Normalization

Normalization is needed to accelerate the process of building a machine learning model, normalization can optimize the LSTM model to be built [25]. Data normalization uses the min-max normalization technique which converts the data into a range of 0 to 1. The data is converted into a relative proportion of the entire data value range, which is calculated using Equation 1. The following in Table 3 is the normalized value of rainfall.

TABLE III
DATA NORMALIZATION

Date	Rainfall	Normalization Value
01/01/2014	19.5360	0.0235
02/01/2014	12.5916	0.0152
03/01/2014	12.3000	0.0148
:	:	:
29/12/2024	10.2297	0.0123
30/12/2024	2.2810	0.0027
31/12/2024	0.0000	0.0000

C. Time Series k-Fold Cross Validation

The *k*-fold cross validation process is used to optimize parameters in each model by using the value *k*=5. The division of each fold refers to the whole data, this is done so that each fold is representative. The model was built using an Adam optimizer with the same batch size of 1 but different epochs. Epoch consists of 25, 50, 75, 100 which aims to get the best-performing model with maximum accuracy value. The entire data is divided into Figure 2, Figure 3, Figure 4, Figure 5, and Figure 6. The model fold is a time series data shift working well as shown in Table 1.

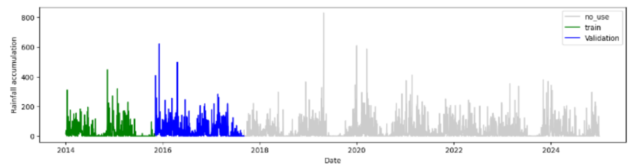

Figure 3. Fold 1

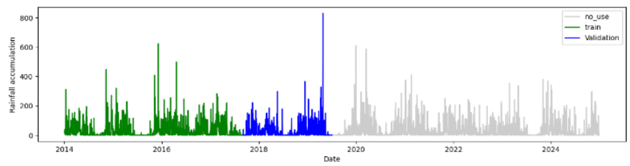

Figure 4. Fold 2

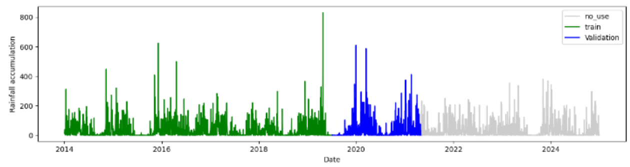

Figure 5. Fold 3

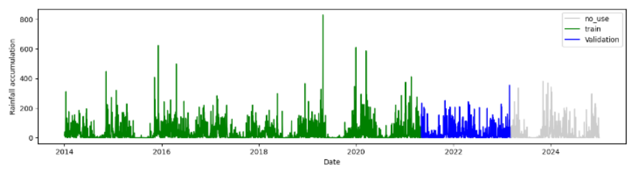

Figure 6. Fold 4

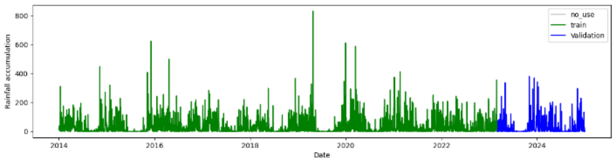

Figure 7. Fold 5

D. Data Denormalization and Model Evaluation of Cross Validation Results

Return data to the original data by using an inverse transformation on each previously normalized data. By denormalizing the data back to its original value. The denormalization process is carried out before the cross-validation process and before the model evaluation. The use of time steps 60 is one of the good influencing factors on each model. The time step in the LSTM model makes it possible to take into account information in the overall actual data when conducting the forecasting process [26].

Figure 8 and Figure 9 show the cross-validation results of each model that has been tested. It shows that the number of epochs affected the model accuracy. This is in line with research [27], where the addition of epochs does not improve the model's accuracy. The best model with the best average value is Model 1. Model 1 using the Adam optimizer scenario, epoch 25, and batch size 1 had the smallest RMSE, the smallest MAE, and the highest correlation values of 56.3340, 35.5223, and 0.3209. An LSTM model with an epoch value of 25 can train the entire data to obtain more detailed and complex patterns. This model can adjust weights and biases better so that the model can be forecasted with the expectation of more accuracy and high credibility.

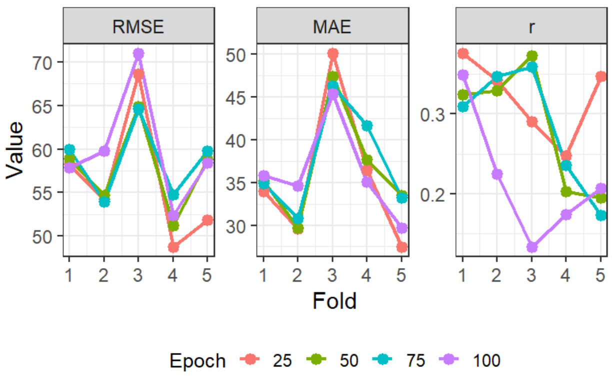

Figure 8. Cross-validation Results of Each Fold

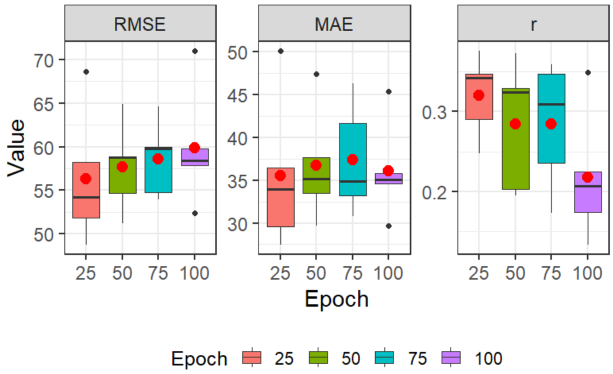

Figure 9. Distribution Cross-validation Results of Each Epoch

The prediction graph using cross-validation Model 1 can be seen in Figure 10 until Figure 14. The graph represents the results of rainfall prediction for Bogor City using an LSTM model with cross-validation across five folds. The evaluation metrics of the cross-validation results show the best values in different folds. The lowest RMSE value of 48.7549 is in Fold 4, the lowest MAE value of 27.4585 is in Fold 5, and the highest correlation value of 0.3764 is in Fold 1. This indicates that using different evaluation metrics can also influence model selection. However, in Folds 3, the model struggles to predict rainfall peaks, resulting in higher RMSE, higher MAE values, and lower correlation. This implies that the model may have difficulty handling extreme or outlier values. In contrast, Folds 1, 4, and 5 show close alignment between the predicted and actual test data because it can be seen that the predicted value can follow the pattern in the actual data, indicating the model's strong predictive capability for those subsets. Cross-validation ensures that the model's performance is tested across various data splits, reducing the risk of overfitting or underfitting. The variation of metrics evaluation across folds highlights how model performance depends on the data distribution in each subset. Overall, the LSTM model with 25 epochs and a batch size of 1 demonstrates the best results, particularly in Fold 5.

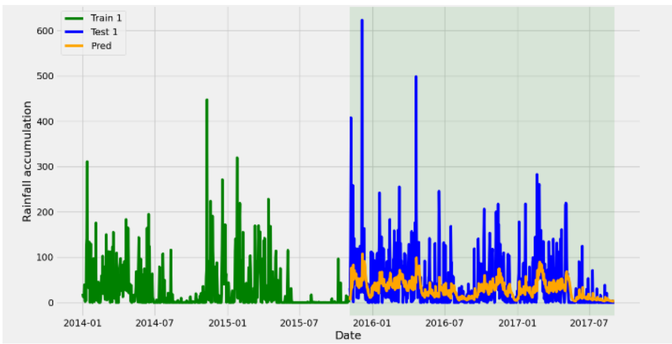

Figure 10. Fold 1 Model 1

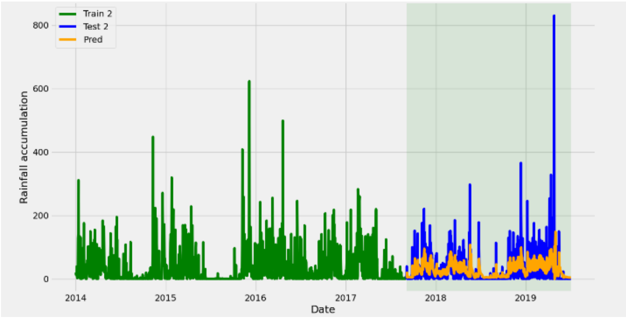

Figure 11. Fold 2 Model 1

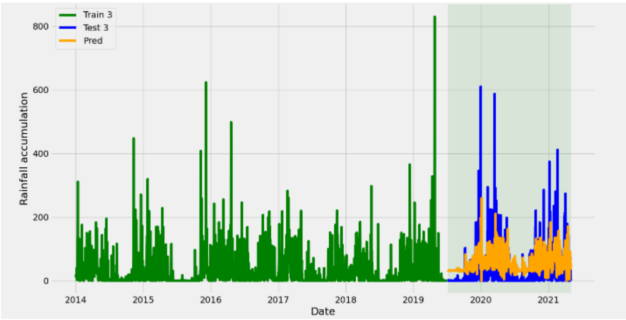

Figure 12. Fold 3 Model 1

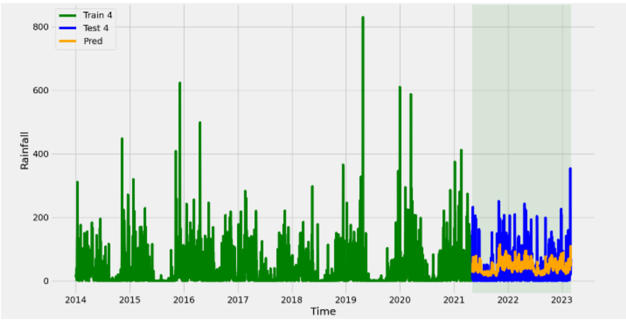

Figure 13. Fold 4 Model 1

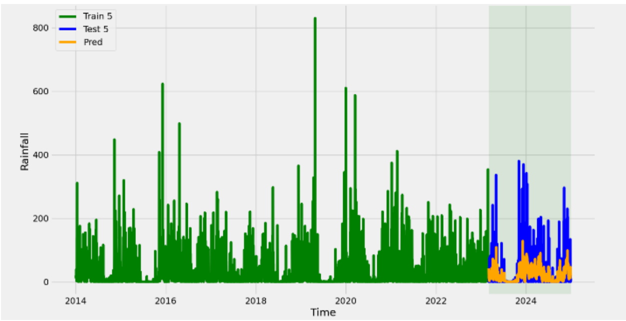

Figure 14. Fold 5 Model 1

E. Forecasting with The Best Model

Model 1 was selected to forecast rainfall for the next 2 years (730 days) based on the evaluation of the time series model using *k*-fold expanding window cross-validation. This model uses an Adam optimizer, the number of epochs is 25, and the batch size is 1. The forecasting process covers the period from January 1, 2025, to December 31, 2026, with all available data being used as training data, which was previously inverse-transformed using Equation 2.

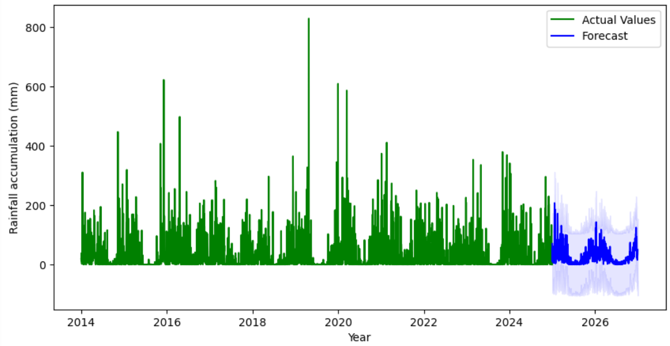

Figure 13. Forecasting with The Best Model

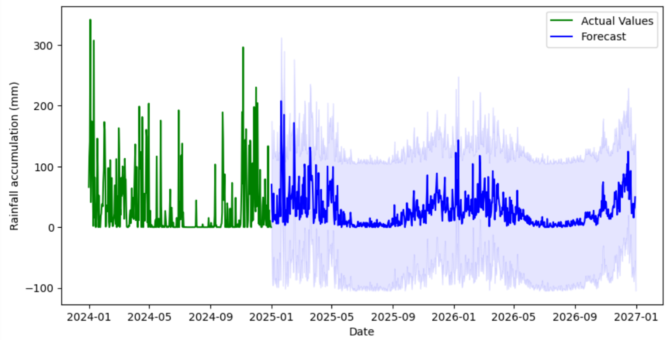

Figure 14. Enlarged Forecasting Results with the Best Model

Figure 13 and Figure 14 shows the results of the rainfall forecasting for the next 2 years with light blue color shows the 95% confidence interval for the forecasting results. The predicted value of rainfall shows quite significant changes and varies from day to day. The highest rainfall prediction will occur on January 20, 2025, with a value of 207.4753 mm, while the lowest rainfall with zero value is expected to occur in mid-June 2025 and early July 2026. However, excluding the zero value, the lowest rainfall prediction is 0.0961 mm on August 5, 2026. In general, these results are in line with [28], where LSTM can capture seasonal patterns, especially in the case of rainfall here in Bogor, where June-July-August has low rainfall and the rest of the year has mixed rainfall.

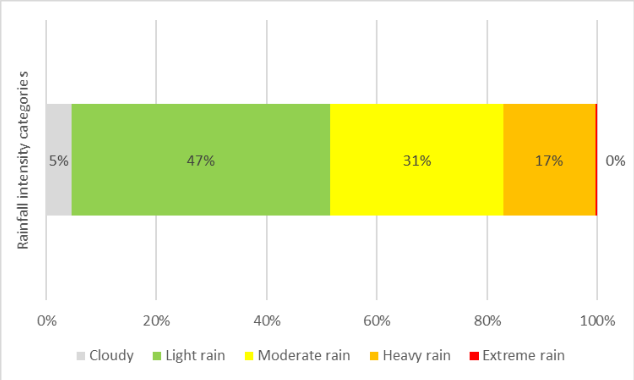

Figure 15. Propose and Labelling with Prediction Data

Figure 15 shows that most precipitation predictions fall into the light rain category, which accounts for 47% of the total data. This category is followed by moderate rain, which accounts for 31%, indicating that the intensity of rainfall generally ranges from light to moderate. Meanwhile, the category of heavy rain has a proportion of 17%, followed by cloudy 5%. The extreme rain category is the least common, accounting for nearly 0%. These results show that light to high-intensity rainfall predominates over other conditions. Therefore, this should be a concern because, according to [29], low rainfall can increase the water crisis, and extreme rainfall can cause natural disasters.

IV. CONCLUSION

The analysis of time series data for predicting rainfall in Bogor City for two years, from January 1, 2025, to December 31, 2026, was conducted using the LSTM model with a robust methodology. Cross-validation show that the best-performing model was achieved with 25 epochs and a batch size of 1, resulting in an average RMSE of 56.3340, MAE of 35.5223, and correlation of 0.3209. This model successfully captured the rainfall data's temporal dependencies and seasonal patterns and provided predictions for the next two years. The forecast indicates a tendency towards light to heavy rainfall during this period.

This research applies univariate forecasting, so it still focuses on modelling. It is still possible to use other hyperparameters to influence the model, such as [27], which varies the batch size. In addition, the addition of variables that are affected/influence rainfall, such as air humidity [30] or temperature [31], so that it becomes a more reliable multivariate analysis for planning and decision-making in agriculture, health, water resource management, and disaster preparedness in the Bogor area.

REFERENCES

[1] K. Yamamoto, T. Sayama, and Apip, “Impact of climate change on flood inundation in a tropical river basin in Indonesia,” *Progress in Earth and Planetary Science*, vol. 8, no. 1, p. 5, Jan. 2021, doi: 10.1186/s40645-020-00386-4.

[2] E. S. Buffa and R. K. Sarin, *Modem Production/ Operations Management*. Los Angeles, CA, USA: John Wiley & Sons, Inc., 1987.

[3] J. Patterson and A. Gibson, *Deep Learning A Practitioners Approach*. California (US): O’Reilly Media, 2017.

[4] O. A. Qasem, M. Akour, and M. Alenezi, “The Influence of Deep Learning Algorithms Factors in Software Fault Prediction,” *IEEE Access*, vol. 8, pp. 63945–63960, 2020, doi: 10.1109/ACCESS.2020.2985290.

[5] J. N. van Rijn and F. Hutter, “Hyperparameter Importance Across Datasets,” in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, in KDD ’18. New York, NY, USA: Association for Computing Machinery, Jul. 2018, pp. 2367–2376. doi: 10.1145/3219819.3220058.

[6] K. Cho and Y. Kim, “Improving streamflow prediction in the WRF-Hydro model with LSTM networks,” *Journal of Hydrology*, vol. 605, p. 127297, Feb. 2022, doi: 10.1016/j.jhydrol.2021.127297.

[7] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[8] A. Chaudhary, K. S. Chouhan, J. Gajrani, and B. Sharma, “Deep Learning With PyTorch,” in *Machine Learning and Deep Learning in Real-Time Applications*, IGI Global Scientific Publishing, 2020, pp. 61–95. doi: 10.4018/978-1-7998-3095-5.ch003.

[9] H. Sak, A. Senior, and F. Beaufays, “Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition,” Feb. 05, 2014, *arXiv*: arXiv:1402.1128. doi: 10.48550/arXiv.1402.1128.

[10] P. Le and W. Zuidema, “Quantifying the vanishing gradient and long distance dependency problem in recursive neural networks and recursive LSTMs,” Mar. 01, 2016, *arXiv*: arXiv:1603.00423. doi: 10.48550/arXiv.1603.00423.

[11] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” Jan. 30, 2017, *arXiv*: arXiv:1412.6980. doi: 10.48550/arXiv.1412.6980.

[12] F. Asferizal, “Analisis Perbandingan Kehandalan Data Hujan GSMaP, TRMM, GPM dan PERSIANN Terhadap Data Obsevasi Dalam Rentang Waktu Penelitian 2020 - 2021,” *Journal of Infrastructure Planning and Design*, vol. 2, no. 1, Art. no. 1, Jul. 2022.

[13] I. Permana and F. N. S. Salisah, “Pengaruh Normalisasi Data Terhadap Performa Hasil Klasifikasi Algoritma Backpropagation: The Effect of Data Normalization on the Performance of the Classification Results of the Backpropagation Algorithm,” *Indonesian Journal of Informatic Research and Software Engineering (IJIRSE)*, vol. 2, no. 1, Art. no. 1, Mar. 2022, doi: 10.57152/ijirse.v2i1.311.

[14] J. Brownlee, *Long Short-Term Memory Networks With Python: Develop Sequence Prediction Models with Deep Learning*. Machine Learning Mastery, 2017.

[15] R. Tanjung, A. Listiani, and F. Lestari, “Prediksi Multivariate Time Series Parameter Cuaca Menggunakan Long Short - Term Memory (LSTM),” *Prosiding Seminar Nasional Sains Data*, vol. 4, no. 1, Art. no. 1, Sep. 2024, doi: 10.33005/senada.v4i1.253.

[16] R. F. Firdaus and I. V. Paputungan, “Prediksi Curah Hujan di Kota Bandung Menggunakan Metode Long Short Term Memory,” *Jurnal Penelitian Inovatif*, vol. 2, no. 3, pp. 453–460, Nov. 2022, doi: 10.54082/jupin.99.

[17] M. Q. Mahmood, X. Wang, F. Aziz, and T. Pang, “Evaluating the sustainability of groundwater abstraction in small watersheds using time series analysis,” *Groundwater for Sustainable Development*, vol. 26, p. 101288, Aug. 2024, doi: 10.1016/j.gsd.2024.101288.

[18] Y. Feng, Y. Zhang, and Y. Wang, “Out-of-sample volatility prediction: Rolling window, expanding window, or both?,” *Journal*

of Forecasting, vol. 43, no. 3, pp. 567–582, 2024, doi: 10.1002/for.3046.

[19] T. Prasetyo, R. A. Putri, D. Ramadhani, Y. Angraini, and K. A. Notodiputro, “Perbandingan Kinerja Metode Arima, Multi-Layer Perceptron, dan Random Forest dalam Peramalan Harga Logam Mulia Berjangka yang Mengandung Pencilan,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 2, Art. no. 2, Aug. 2024, doi: 10.25126/jtiik.20241127392.

[20] CNN Indonesia, “Jadwal Lengkap Awal Musim Hujan di Bogor, Depok, Tangerang, dan Bekasi,” CNN Indonesia. Accessed: Mar. 10, 2025. [Online]. Available: https://www.cnnindonesia.com/teknologi/20241022172439-641-1158456/jadwal-lengkap-awal-musim-hujan-di-bogor-depok-tangerang-dan-bekasi

[21] S. Hidayatulah, “BPBD Prediksi Musim Kemarau di Bogor Hingga Oktober 2023 - Pakuan Raya.” Accessed: Mar. 10, 2025. [Online]. Available: https://pakuanraya.com/bpbd-prediksi-musim-kemarau-di-bogor-hingga-oktober-2023/

[22] A. ADRI, “Kabupaten Bogor Mulai Dilanda Kekeringan,” kompas.id. Accessed: Mar. 10, 2025. [Online]. Available: https://www.kompas.id/baca/metro/2023/06/11/kabupaten-bogor-mulai-dilanda-kekeringan

[23] D. Kasihairani, R. Virgianto, and S. Risnayah, “Dampak El Niño Southern Oscillation Dan Indian Ocean Dipole Mode Terhadap Variabilitas Curah Hujan Musiman Di Indonesia,” Jun. 2014.

[24] M. D. A. Carnegie and C. Chairani, “Perbandingan Long Short Term Memory (LSTM) dan Gated Recurrent Unit (GRU) Untuk Memprediksi Curah Hujan,” *Jurnal Media Informatika Budidarma*, vol. 7, no. 3, Art. no. 3, Jul. 2023, doi: 10.30865/mib.v7i3.6213.

[25] W. li and Z. Liu, “A method of SVM with Normalization in Intrusion Detection,” *Procedia Environmental Sciences*, vol. 11, pp. 256–262, Jan. 2011, doi: 10.1016/j.proenv.2011.12.040.

[26] F. Masri, D. Saepudin, and D. Adytia, “Peramalan Seri Waktu Permukaan Laut Menggunakan Deep Learning Rnn, Lstm, Dan Bilstm, Studi Kasus Di Teluk Jakarta, Indonesia,” *eProceedings of Engineering*, vol. 7, no. 2, Art. no. 2, Aug. 2020, Accessed: Mar. 10, 2025. [Online]. Available: https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/12760

[27] S. S. Chai and K. L. Goh, “Daily Rainfall Forecasting Using Meteorology Data with Long Short-Term Memory (LSTM) Network,” *Journal of Optimization in Industrial Engineering*, vol. 15, no. 1, pp. 187–193, Jun. 2022, doi: 10.22094/joie.2021.1941252.1899.

[28] S. Muzaffar and A. Afshari, “Short-Term Load Forecasts Using LSTM Networks,” *Energy Procedia*, vol. 158, pp. 2922–2927, Feb. 2019, doi: 10.1016/j.egypro.2019.01.952.

[29] N. A. Putri and A. Wibowo, “Rainfall Maps For The Suitability Of Settlement Area In Bogor Raya,” *EnviroScienteae*, vol. 19, no. 2, pp. 123–129, May 2023, doi: 10.20527/es.v19i2.15116.

[30] R. Wulandari, T. Rahmawati, A. Asyary, and F. Nugraha, “Analysis of Climate and Environmental Risk Factors on Dengue Hemorrhagic Fever Incidence in Bogor District,” *Kesmas*, vol. 18, no. 3, pp. 209–214, Aug. 2023, doi: 10.21109/kesmas.v18i3.7351.

[31] E. Khyber, L. Syaufina, and A. Sunkar, “Variability and time series trend analysis of rainfall and temperature in Dramaga Sub-District, Bogor, Indonesia,” *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 771, no. 1, p. 012016, May 2021, doi: 10.1088/1755-1315/771/1/012016.