International Journal of Informatics and Computation (IJICOM)
Vol. 8, No. 1, 2026
ISSN: 2685-8711, E-ISSN: 2714-5263 DOI: 10.35842/ijicom

Comparison of Machine Learning and Deep Learning Algorithms for Daily Weather Forecasting

Dedy Abdianto Nggego[1], Paskha Marini Thana[2]

Abstract

Global climate change has increased the complexity of weather patterns, particularly in tropical regions such as Merauke Regency. This study evaluates five machine learning models including RF, SVM, Prophet, LSTM, and GRU for daily weather forecasting. We aim to assess the effectiveness of deep learning methods in capturing temporal dependencies in tropical weather systems and compare their performance against conventional machine learning approaches. We apply multiple evaluation metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2), to ensure robust model assessment. The results show that deep learning models consistently outperform traditional methods. GRU achieves the best performance with RMSE = 1.37, MSE = 1.88, and R^2 = 0.87, followed closely by LSTM with RMSE = 1.39 and R^2 = 0.86. In contrast, RF, SVM, and Prophet exhibit higher error rates and lower predictive accuracy. Correlation analysis reveals strong relationships between key meteorological variables, particularly rainfall and humidity, indicating that multi-variable inputs improve forecasting performance. Overall, the findings confirm that GRU is the most effective model for this dataset, while LSTM serves as a strong alternative. This study highlights the superiority of deep learning approaches for modeling complex, nonlinear, and time-dependent weather patterns in tropical regions.

Keywords:
Weather forecasting, GRU, LSTM, Deep Learning

1. Introduction

Climate change continuously increases the complexity and variability of weather patterns across the world. Rising global temperatures, irregular rainfall distribution, storms, floods, droughts, and seasonal shifts directly affect environmental stability and human activities. Tropical countries such as Indonesia experience stronger impacts because they are strongly influenced by monsoon circulation, high humidity, and ocean-atmosphere interactions. Merauke Regency, located in South Papua, faces highly dynamic weather conditions due to its coastal geography and exposure to regional climate systems. These conditions create serious challenges for conventional forecasting approaches that often struggle to capture rapidly changing local weather characteristics. Accurate daily weather forecasting therefore becomes increasingly important for improving environmental monitoring, supporting economic activities, and reducing climate-related risks in vulnerable regions [1], [9], [11].

Weather forecasting plays a critical role in supporting agriculture, fisheries, transportation, and disaster mitigation systems. Farmers depend on weather predictions to determine planting schedules, irrigation management, and harvesting periods. Fishermen require accurate information regarding rainfall, wind speed, and storm potential to ensure operational safety at sea. Transportation sectors, particularly aviation and maritime

Corresponding Author: Dedy Abdianto Nggego (dedyabdianto@unmus.ac.id)
1 Dedy Abdianto Nggego, Teknik Informatika Universitas Musamus, dedyabdianto@unmus.ac.id
2 Paskha Marini Thana, Pendidikan Matematika Universitas Musamus, paskhathana@unmus.ac.id

services, also rely heavily on weather stability to maintain safety and operational efficiency. In disaster-prone regions, weather forecasting functions as an early warning mechanism for floods, storms, and droughts that threaten local communities. However, traditional numerical weather prediction models such as Weather Research and Forecasting (WRF) demand high computational resources and often encounter limitations when modeling localized atmospheric variability. These constraints encourage researchers to explore alternative forecasting approaches that are faster, more adaptive, and capable of learning directly from historical climate patterns [1], [2], [4].

The rapid growth of artificial intelligence and data science significantly transforms modern weather forecasting systems. Machine learning methods enable computers to identify patterns, relationships, and trends from historical weather data without relying entirely on complex physical equations. These approaches become increasingly relevant because meteorological agencies now generate massive datasets from weather stations, satellites, radars, and IoT-based sensors. Machine learning algorithms can process these large datasets efficiently and adaptively, making them suitable for modeling highly nonlinear atmospheric behavior. Researchers also apply preprocessing and feature engineering techniques to improve data quality and enhance predictive performance. As a result, machine learning-based forecasting systems increasingly demonstrate their ability to outperform traditional statistical approaches in handling dynamic and uncertain weather conditions [3], [6], [20], [21].

Several conventional machine learning algorithms have demonstrated promising performance in weather forecasting tasks. Random Forest effectively handles complex and high-dimensional weather datasets by combining multiple decision trees to reduce overfitting and improve prediction stability. Support Vector Machine (SVM) performs well in nonlinear classification and regression problems through kernel-based optimization. Prophet also becomes popular for time-series forecasting because it can capture seasonal and trend-based patterns efficiently. Previous studies show that Random Forest produces reliable weather predictions with strong interpretability, while SVM successfully models nonlinear atmospheric relationships. Gradient boosting approaches such as XGBoost further improve forecasting performance through scalable ensemble learning mechanisms. These machine learning techniques provide practical advantages because they require lower computational costs compared to numerical weather prediction systems while maintaining competitive predictive capability [5], [16], [22], [25], [26], [27], [28].

Deep learning approaches further advance weather forecasting performance by learning temporal dependencies and hidden structures from sequential climate data. Long Short-Term Memory (LSTM) networks address the vanishing gradient problem in recurrent neural networks and effectively capture long-term temporal patterns within weather sequences. Gated Recurrent Unit (GRU) models simplify recurrent architectures while maintaining strong temporal learning capability with lower computational complexity. Recent studies consistently report that LSTM and GRU outperform conventional machine learning models in time-series forecasting applications, including rainfall prediction, temperature forecasting, and wind speed estimation. Temporal Convolutional Networks (TCN) and ConvLSTM architectures also demonstrate strong capability in extracting spatial-temporal weather features. These findings confirm that deep learning methods are highly suitable for modeling dynamic atmospheric behavior characterized by nonlinear and sequential dependencies [2], [7], [12], [13], [14], [17], [29], [31].

Recent advancements in artificial intelligence significantly improve global weather forecasting systems through large-scale deep learning integration. Research by Lam et al. demonstrates that deep learning can achieve skillful medium-range global weather forecasting comparable to physics-based operational systems. Similarly, Rasp and Thuerey show that purely data-driven forecasting approaches can produce competitive results against traditional numerical weather models at similar spatial resolutions. Modern neural network architectures equipped with attention mechanisms and context-matching modules further enhance forecasting precision by capturing complex atmospheric interactions more effectively. These developments indicate that AI-driven forecasting systems no longer function merely as supplementary tools but increasingly become primary alternatives for operational weather prediction. Nevertheless, most advanced forecasting studies focus on regions with extensive meteorological infrastructures and large observational datasets, leaving tropical and remote regions underrepresented in the literature [4], [14], [15], [19], [30].

Tropical regions such as Indonesia present unique forecasting challenges due to the influence of large-scale climate phenomena including El Niño–Southern Oscillation (ENSO), Madden–Julian Oscillation (MJO), and complex monsoon systems. These phenomena create irregular rainfall patterns, extreme seasonal variability, and rapid atmospheric transitions that differ significantly from subtropical climate systems. Merauke Regency experiences additional forecasting difficulties because weather observation infrastructure remains limited and real-time meteorological data availability is relatively low. Models developed using datasets from Europe, North America, or East Asia therefore cannot be directly generalized to tropical regions without local adaptation. Previous studies in Indonesia demonstrate that LSTM and GRU models can effectively capture regional climate variability, but comparative studies involving multiple machine learning and deep learning algorithms remain limited, especially for remote tropical regions such as Merauke [9], [10], [11], [17].

This study addresses these research gaps by comparing multiple machine learning and deep learning algorithms for daily weather forecasting using local tropical climate data from Merauke Regency. This research evaluates the predictive performance of Random Forest, Support Vector Machine (SVM), Prophet, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) models using evaluation metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and coefficient of determination (R^2). This study specifically investigates the ability of deep learning models to capture temporal dependencies and nonlinear weather patterns under highly dynamic tropical climate conditions. Unlike previous studies that mainly focus on subtropical datasets or single-model evaluations, this paper provides a comprehensive comparative analysis using localized weather data from an underrepresented tropical region. The findings are expected to support the development of intelligent weather forecasting systems for smart farming, fisheries management, transportation safety, and disaster mitigation in Indonesia and other tropical regions with similar environmental characteristics [5], [10], [12], [15], [17].

2. Related Works

Recent studies have shown that deep learning models significantly improve weather forecasting performance. Saleem et al. developed an ensemble model combining LSTM and GRU for temperature and humidity forecasting and achieved higher accuracy than single-model approaches [12]. Alfredo and Adytia compared GRU, CNN-GRU, and LSTM models for wave height forecasting and found that hybrid CNN-GRU models captured temporal patterns more effectively [13]. Tekin et al. introduced a CNN-LSTM model integrated with attention and context matcher mechanisms, which improved the ability to model spatial and temporal weather dependencies simultaneously [14]. Lam et al. later demonstrated that deep learning systems such as GraphCast produced global weather forecasts with accuracy comparable to, and sometimes better than, traditional numerical forecasting systems [15].

Several studies in Indonesia also applied machine learning algorithms for local weather prediction. Dwiyanti and Prianto used Random Forest to predict weather conditions in Jakarta and reported satisfactory forecasting accuracy for urban climate data [5]. Fauzi et al. implemented Support Vector Machine (SVM) models for weather prediction and showed that SVM effectively classified weather patterns with high precision [16]. Respaty et al. compared LSTM and GRU models for weather prediction in Jakarta and found that both architectures successfully captured temporal weather dependencies in tropical environments [10]. However, most Indonesian studies focused on urban regions with relatively complete meteorological infrastructure and did not evaluate remote tropical areas with limited observation data.

Research comparing deep learning and traditional statistical models consistently demonstrated the superiority of neural network approaches for time-series forecasting. Sunendar and Rianto compared ARIMA, LSTM, and GRU models and found that LSTM and GRU significantly outperformed ARIMA in handling nonlinear sequential data [18]. Islamy and Wahabi also evaluated LSTM and GRU for wind forecasting and observed that GRU achieved competitive accuracy with lower computational complexity [17]. These findings suggested that recurrent neural network architectures are more suitable for dynamic weather datasets because they can preserve long-term temporal dependencies more effectively than conventional statistical approaches.

Several foundational studies established the theoretical basis of machine learning and deep learning algorithms used in forecasting systems. Breiman introduced Random Forest as an ensemble learning method capable of reducing overfitting while improving predictive stability [22]. Cortes and Vapnik proposed Support Vector Machines (SVM), which became widely used for classification and regression problems due to their strong generalization ability [26]. Hochreiter and Schmidhuber developed the LSTM architecture to solve the vanishing gradient problem in recurrent neural networks [7], [31]. Cho et al. later introduced the GRU architecture as a simpler recurrent model with fewer parameters and faster training performance [29].

Deep learning studies also highlighted the importance of large-scale data and representation learning in weather prediction. Reichstein et al. explained that deep learning enabled Earth system science models to automatically extract hidden patterns from complex environmental datasets [3]. Goodfellow et al. described how deep neural networks could learn hierarchical representations from sequential data and improve prediction quality for nonlinear systems [20]. Chollet further emphasized that modern deep learning frameworks simplified the implementation of advanced forecasting architectures for real-world applications [21]. Géron also showed that machine learning libraries such as Scikit-Learn, TensorFlow, and Keras accelerated practical forecasting system development [24].

Studies on numerical weather prediction demonstrated the limitations of traditional forecasting approaches. Bauer et al. explained that numerical weather prediction systems required extensive computational resources and still struggled to represent highly localized atmospheric phenomena [1]. Rasp and Thuerey proposed a purely data-driven forecasting system and reported that machine learning models achieved prediction performance comparable to physical numerical models at similar spatial resolutions [4]. These findings encouraged researchers to adopt machine learning approaches as alternatives for operational forecasting systems, especially in regions with limited computational infrastructure.

Climate variability research also showed that tropical regions present unique forecasting challenges. McPhaden et al. explained that the El Niño–Southern Oscillation (ENSO) strongly influenced rainfall and temperature variability across tropical regions [9]. Zhang et al. reported that climate variability increased environmental uncertainty and produced unstable weather conditions over time [11]. These phenomena became more complex in Indonesia because monsoon circulation and coastal interactions created highly dynamic local weather patterns. As a result, forecasting models developed for subtropical regions could not be directly applied to tropical environments without adaptation.

Despite rapid progress in machine learning-based forecasting, several limitations remained unresolved. Many studies relied on datasets from developed regions with dense meteorological infrastructure, while tropical regions such as Merauke received limited research attention. Most previous studies also focused on single-model evaluation instead of comprehensive comparisons between machine learning and deep learning algorithms under the same dataset conditions. In addition, researchers often treated machine learning systems as black-box models, making interpretation and operational implementation more difficult [3]. These limitations motivated this study to compare Random Forest, SVM, Prophet, LSTM, and GRU models for daily weather forecasting using tropical climate data from Merauke Regency.

3. Proposed Method

This study proposes a daily weather forecasting framework based on machine learning and deep learning approaches to address the high variability and nonlinear characteristics of tropical weather data in Merauke Regency. This paper applies a comparative-experimental approach by evaluating several forecasting algorithms to identify the most effective model based on predictive accuracy and computational performance. The proposed framework consists of six main stages, namely data collection, data preprocessing, feature engineering, model development, performance evaluation, and model validation. We utilize historical weather datasets to train and test multiple machine learning and deep learning models capable of capturing temporal weather patterns and complex atmospheric relationships. Furthermore, this study adopts a standardized machine learning pipeline for time-series forecasting to ensure systematic model development, reproducibility, and reliable comparative analysis across all evaluated algorithms [6].

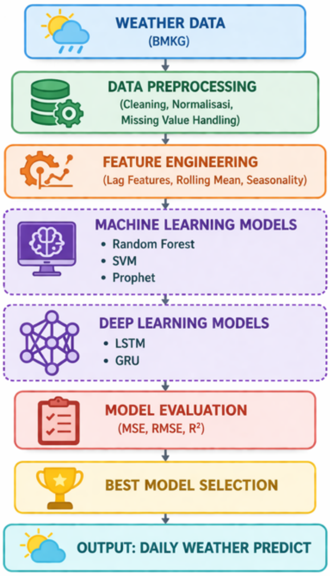

Fig. 1. Research flowchart

This study develops a comparative forecasting model using both conventional machine learning and deep learning algorithms to analyze nonlinear weather patterns with temporal dependencies. We utilize five years of daily weather data obtained from BMKG, including temperature, humidity, rainfall, and wind speed variables. The preprocessing stage includes data cleaning, normalization, and time-series transformation to improve data consistency and model stability. We also apply feature engineering techniques such as lag features, rolling mean, moving averages, and seasonal extraction to strengthen the ability of the models to capture temporal weather patterns.

This paper evaluates RF, SVM, Prophet, LSTM, and GRU models to compare their forecasting performance on tropical weather data. RF effectively handles nonlinear relationships and reduces overfitting through ensemble learning, while SVM performs well on small and medium-sized datasets with complex feature distributions. Prophet efficiently models seasonal and trend components in time-series data. In the deep learning category, LSTM captures long-term temporal dependencies through recurrent memory mechanisms, whereas GRU provides similar forecasting capability with lower computational complexity and faster training time. The selected algorithms represent different modeling approaches and enable comprehensive evaluation of machine learning and deep learning methods for daily weather forecasting.

In this study, we formulate the prediction of the Random Forest (RF) model as the average of all decision tree outputs. The model aggregates the predictions from multiple trees to produce a final estimate, improving stability and accuracy compared to a single decision tree. The RF prediction is defined as:

$$y = \frac{1}{N} \Sigma_{\mathrm{i=1}}^{\mathrm{n}} T_i(x) \tag{1}$$

Where $T_i(x)$ represents the prediction from the i-th decision tree and N denotes the number of trees.

In this study, we model SVM as a method that constructs an optimal hyperplane to minimize prediction error. We define the SVM regression function as:

$$f(x) = w^T x + b \tag{2}$$

Where w is the weight vecror and b is the bias parameter.

LSTM employs memory cells and gating mechanisms to capture long-term dependencies in time series data. The forget gate equation is defined as:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{3}$$

Where f_t denotes the forget gate output, x_t is the input vector, and h_{t-1} is the previous hidden state. LSTM was specifically designed to overcome the vanishing gradient problem commonly encountered in conventional recurrent neural networks when learning long-term temporal dependencies.

GRU simplifies the LSTM architecture using update and reset gates. The update gate can be formulated as:

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \tag{4}$$

Where z_t controls the extent to which previous information is retained.

The next stage involves model evaluation to assess the predictive performance of the developed models in a comprehensive and systematic manner. We evaluate accuracy by comparing predicted values against observed values using three widely adopted regression metrics: MSE, RMSE, and R^2. We use MSE to measure the average magnitude of squared prediction errors, RMSE to provide an interpretable measure of error in the original data scale, and R^2 to quantify how well the model explains the variance in the observed data. Together, these metrics provide a balanced assessment of both error magnitude and explanatory power, ensuring that the model's predictive capability is rigorously validated. The mathematical formulations of these evaluation metrics are presented in the following section.

$$MSE = \left(\frac{1}{n}\right) * \Sigma\left(y_{true} - y_{pred}\right)^2 \tag{5}$$

$$RMSE = \left(\frac{\sqrt{\Sigma\left(y_{true} - y_{pred}\right)^2}}{n}\right) \tag{6}$$

$$R^2 = 1 - \left(\frac{\Sigma\left(y_{true} - y_{pred}\right)^2}{\Sigma(y_{true} - y_{mean})^2}\right) \tag{7}$$

The final stage involves model validation under real-world conditions to ensure that the predictions are not only statistically accurate but also practically reliable when applied outside the training environment. In this stage, we assess how well the model generalizes to unseen data and whether its outputs remain stable under realistic operational scenarios. We further examine the consistency of predictions across different subsets of data and evaluate potential deviations that may arise due to variability in real-world inputs. This process ensures that the model is robust, transferable, and suitable for practical deployment in decision-making contexts.

4. Experimental Setup

In this study, the dataset consists of historical daily weather data obtained from the Meteorology, Climatology, and Geophysics Agency (BMKG) of Merauke Regency for the period 2020–2024. We analyze a tropical climate system characterized by high variability using four main variables: air temperature (°C), humidity (%), rainfall (mm), and wind speed (m/s). We treat the dataset as a time series, where each observation depends on previous time steps, making it suitable for forecasting models that capture temporal dependencies.

We apply standard pre-processing to ensure data quality and consistency before modelling. We configure all models using tuned or controlled parameters for fair comparison. We set RF to 100 trees, SVM with RBF kernel (C = 1.0, gamma = scale), and Prophet with default settings and automatic changepoint detection. We implement DL models (LSTM, GRU) using one hidden layer with 50 neurons, Adam optimizer, 50 epochs, batch size 32, and MSE loss function. We design the evaluation framework using multiple approaches to ensure robust comparison across models. We summarize the dataset structure in Table 1.

Table 1. Dataset and Model Configuration Summary

Component	Description
Data Source	BMKG Merauke Regency
Period	2020–2024
Variables	Temp (°C), Humidity (%), Rainfall (mm), Wind Speed (m/s)
Data Type	Time series (daily)
RF	100 trees
SVM	RBF kernel, C=1.0, gamma=scale
Prophet	Default + changepoint detection
LSTM	1 layer, 50 neurons, Adam, 50 epochs, batch=32
GRU	1 layer, 50 neurons, Adam, 50 epochs, batch=32
Loss Function	MSE
Evaluation Metrics	MSE, RMSE, R^2
Analysis	Model comparison, stability test, correlation analysis

5. Result and Analysis

In this study, we analyze five years of historical weather data from Merauke Regency using five ML models (RF, SVM, Prophet, LSTM, and GRU). We find that DL models outperform traditional models in daily weather prediction. We observe that GRU achieves the best performance with RMSE = 1.37, MSE = 1.88, and R^2 = 0.87, followed closely by LSTM with RMSE = 1.39, MSE = 1.95, and R^2 = 0.86. We also find that RF, SVM, and Prophet produce higher errors and weaker predictive performance compared to DL models.

This paper also examines relationships among weather variables in the dataset. We find a strong positive relationship between rainfall and humidity, while temperature and humidity show a weak negative relationship. We also note that wind speed contributes to weather variability, although its influence is smaller than rainfall and humidity. We conclude that using multiple weather variables improves prediction accuracy. Our results identify GRU as the most effective model, with LSTM serving as a strong alternative due to its consistently high performance.

Table 1. Evaluation Results of Forecasting Models Using MSE, RMSE, and R^2

Model	MSE	RMSE	R^2
Random Forest	2.31	1.52	0.82
SVM	2.85	1.69	0.78
Prophet	2.60	1.61	0.80
LSTM	1.95	1.39	0.86
GRU	**1.88**	**1.37**	**0.87**

We evaluate five forecasting models using MSE, RMSE, and R^2, as shown in Table 1. This study finds that DL models outperform traditional ML models across all metrics. We observe that GRU delivers the best performance with the lowest errors (MSE = 1.88, RMSE = 1.37) and the highest R^2 (0.87). Our results also show that LSTM follows closely with strong accuracy (MSE = 1.95, RMSE = 1.39, R^2 = 0.86). This paper finds that traditional models perform less effectively and observes that RF records MSE = 2.31, RMSE = 1.52, and R^2 = 0.82. We also find that Prophet achieves moderate performance with MSE = 2.60, RMSE = 1.61, and R^2 = 0.80, while SVM shows the weakest results among all models with MSE = 2.85, RMSE = 1.69, and R^2 = 0.78. This study conclude that DL models provide more accurate and stable forecasts than traditional ML methods.

.

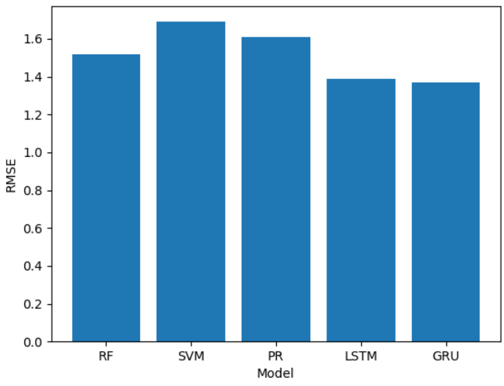

Fig. 3. RMSE Comparison

The R^2 value graph illustrates the ability of each model to explain the variability of the actual data. The GRU model demonstrates the highest R^2 value of 0.87, indicating that it is capable of explaining 87% of the variation in weather data. The LSTM model ranks second with an R^2 value of 0.86, which also reflects very strong performance. RF and Prophet exhibit reasonably good R^2 values; however, they remain below the performance of the deep learning models, while SVM records the lowest R^2 value. These results indicate that the GRU and LSTM models not only produce lower prediction errors but also possess a stronger capability to capture underlying data patterns, making them more stable and accurate for forecasting.

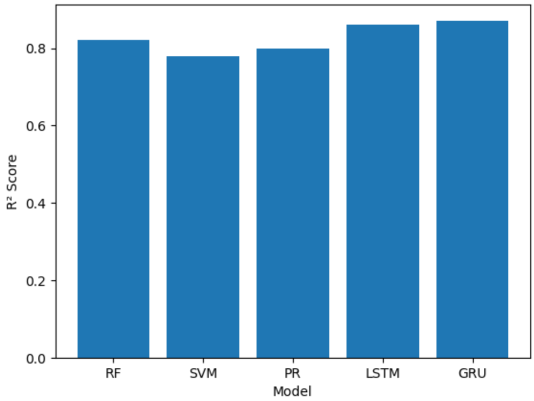

Fig. 4. R^2 Comparison

The superior performance of GRU compared to other algorithms indicates that simplified recurrent architectures are highly effective for weather forecasting tasks involving temporal dependencies. GRU requires fewer parameters than LSTM, resulting in faster convergence and lower computational complexity while maintaining high predictive accuracy. This explains why GRU achieved the lowest RMSE and highest R^2 values in this study.

In contrast, conventional machine learning algorithms such as RF and SVM showed lower predictive capability because these models are not specifically designed to capture sequential temporal dependencies in time-series weather data. Although Prophet performed reasonably well in modelling trend and seasonal components, it was less effective in representing sudden fluctuations and complex non-linear weather patterns commonly observed in tropical regions such as Merauke. The experimental results also demonstrate that deep learning models can better generalise weather patterns under highly dynamic climate conditions. This finding confirms previous studies stating that recurrent neural network architectures are more suitable for time-series forecasting problems involving long-term dependencies and complex atmospheric variability.

The correlation analysis among variables indicates a strong relationship between rainfall and humidity, with a correlation coefficient of 0.85, suggesting that increases in humidity tend to be accompanied by increases in rainfall. In contrast, temperature exhibits a weak negative relationship with humidity, with a correlation value of -0.30, indicating that rising temperatures are generally associated with decreasing humidity. Furthermore, wind speed shows a moderate relationship with extreme weather events, with a correlation coefficient of 0.45. These findings demonstrate that weather variables are interrelated and should be considered collectively within the modelling process.

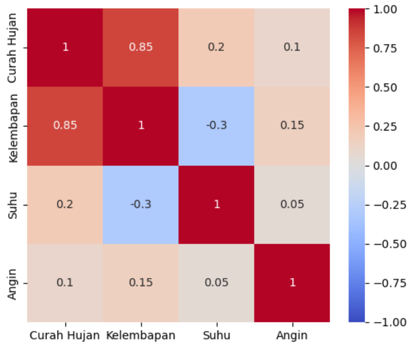

Fig. 5. Correlation Matrix of Variables

We also present the model ranking visualization (Fig. 6) to illustrates the performance order of the algorithms based on the evaluation results. The GRU model ranks first as the best-performing model, followed by LSTM in second place. Random Forest occupies the third position with relatively good performance, while Prophet and SVM rank fourth and fifth, respectively. This ranking indicates that deep learning models consistently outperform other models in terms of prediction accuracy and stability. Therefore, GRU is recommended as the primary model for the implementation of a machine learning-based weather forecasting system.

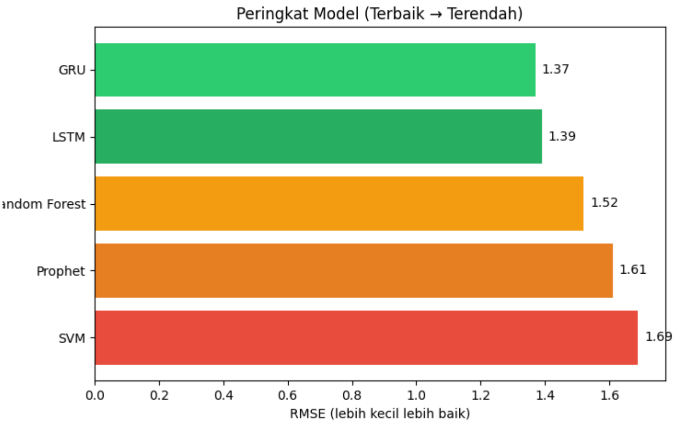

Fig. 6. Model Performance Ranking Visualization

This study finds that DL models such as LSTM and GRU outperform conventional ML models in weather forecasting. We observe that these models better capture long-term dependencies in time series data, which is critical because weather conditions strongly depend on previous states. We show that GRU and LSTM learn sequential patterns more effectively than RF and SVM, which are not designed for temporal dependencies. We also find that RF and SVM still offer advantages in computational efficiency and implementation simplicity, but they struggle to model complex nonlinear and sequential weather patterns, resulting in lower predictive accuracy. This paper reinforces prior studies that report DL models outperform statistical and traditional ML approaches in weather prediction tasks.

We also consider the geographical context of Merauke Regency, a tropical region with high climate variability influenced by monsoon cycles and ENSO events. We observe that these factors create highly dynamic and nonlinear weather patterns that challenge conventional forecasting methods. This study shows that ML-based approaches provide a more adaptive solution for such conditions. We further highlight practical implications across sectors such as agriculture and fisheries. We find that accurate forecasts support farmers in optimizing planting and harvesting schedules, reducing crop failure risks. We also note that fishermen benefit from improved predictions of wind and weather conditions for safer operations. We conclude that ML-based forecasting systems can also support early warning systems for extreme weather events.

6. Conclusion

This study evaluates five ML models (RF, SVM, Prophet, LSTM, and GRU) for daily weather forecasting using five years of historical data from Merauke Regency. We find that DL models consistently outperform conventional ML approaches in both predictive accuracy and stability. We observe that GRU delivers the best performance with the lowest error (RMSE = 1.37, MSE = 1.88) and the highest explanatory power ($R^2 = 0.87$), followed closely by LSTM with similarly strong results (RMSE = 1.39, $R^2 = 0.86$). In contrast, RF, SVM, and Prophet show higher errors and weaker generalization, with SVM performing the worst among all tested models. We conclude that GRU represents the most effective model for capturing temporal dependencies in tropical weather data, while LSTM remains a strong alternative.

This study also confirms that weather variables are strongly interrelated and collectively influence prediction performance. We find a strong positive correlation between rainfall and humidity (0.85), a weak negative relationship between temperature and humidity (−0.30), and a moderate association between wind speed and extreme weather events (0.45). These relationships indicate that multi-variable modeling improves forecasting accuracy. We further show that DL models better capture nonlinear and sequential patterns in highly dynamic tropical climates influenced by monsoon systems and ENSO variability. Based on these findings, we recommend GRU-based forecasting systems for practical deployment, particularly in agriculture, fisheries, and early warning applications, where accurate and stable predictions can support decision-making and reduce climate-related risks.

Acknowledgment

We would like to express our sincere gratitude to Universitas Musamus Merauke for their invaluable support and contributions to this research. We also gratefully acknowledge the financial support provided by the Research and Community Service Institute of Universitas Musamus through the 2025 DIPA grant. Finally, we would like to thank the anonymous reviewers for their constructive comments and valuable suggestions, which have significantly enhanced the quality of this paper.

References

[1] P. Bauer, A. Thorpe, and G. Brunet, “The quiet revolution of numerical weather prediction,” *Nature*, vol. 525, no. 7567, pp. 47–55, 2015, doi: 10.1038/nature14956.
[2] P. Hewage et al., “Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station,” *Soft Computing*, vol. 24, Nov. 2020, doi: 10.1007/s00500-020-04954-0.
[3] M. Reichstein et al., “Deep learning and process understanding for data-driven Earth system science,” *Nature*, vol. 566, no. 7743, pp. 195–204, 2019, doi: 10.1038/s41586-019-0912-1.
[4] S. Rasp and N. Thuerey, “Purely data-driven medium-range weather forecasting achieves comparable skill to physical models at similar resolution,” 2020. doi: 10.48550/arXiv.2008.08626.

[5] Z. A. Dwiyanti and C. Prianto, “Prediksi Cuaca Kota Jakarta Menggunakan Metode Random Forest,” *Jurnal Tekno Insentif*, vol. 17, no. 2, pp. 127–137, Oct. 2023, doi: 10.36787/jti.v17i2.1136.
[6] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. Waltham: Morgan Kaufmann, 2012.
[7] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[8] P. F. Smith, S. Ganesh, and P. Liu, “A comparison of random forest regression and multiple linear regression for prediction in neuroscience,” *Journal of Neuroscience Methods*, vol. 220, no. 1, pp. 85–91, 2013.
[9] M. J. McPhaden, S. E. Zebiak, and M. H. Glantz, “ENSO as an Integrating Concept in Earth Science,” *Science*, vol. 314, no. 5806, pp. 1740–1745, 2006, doi: 10.1126/science.1132588.
[10] W. A. Respaty et al., “Weather Prediction in Jakarta: An Analysis of Climate Data and Regional Influences using LSTM and GRU,” in *Proc. Int. Conf. Data Science and Its Applications (ICoDSA)*, 2023, pp. 408–413, doi: 10.1109/ICoDSA58501.2023.10277097.
[11] Y. Zhang et al., “Climate variability decreases species richness and community stability in a temperate grassland,” *Oecologia*, vol. 188, no. 1, pp. 183–192, 2018.
[12] M. Saleem et al., “An Ensemble Forecasting Method based on optimized LSTM and GRU for Temperature and Humidity Forecasting,” *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18447–18452, 2024, doi: 10.48084/etasr.9047.
[13] C. S. Alfredo and D. A. Adytia, “Time series forecasting of significant wave height using GRU, CNN-GRU, and LSTM,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 5, pp. 776–781, 2022.
[14] S. F. Tekin, A. Fazla, and S. S. Kozat, “Numerical weather forecasting using convolutional-LSTM with attention and context matcher mechanisms,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–13, 2024.
[15] R. Lam et al., “Learning skillful medium-range global weather forecasting,” *Science*, vol. 382, no. 6677, pp. 1416–1421, 2023, doi: 10.1126/science.adi2336.
[16] M. Fauzi et al., “Implementasi Machine Learning Untuk Memprediksi Cuaca Menggunakan Support Vector Machine,” doi: 10.32409/jikstik.23.1.3449.
[17] C. C. Islamy and A. Wahabi, “A Comparative Study of LSTM and GRU Models for Wind Forecasting,” *Sistemasi: Jurnal Sistem Informasi*, vol. 14, no. 6, pp. 2817–2831, 2025.
[18] N. Sunendar and Y. Rianto, “Comparison of ARIMA, LSTM, and GRU Models for Forecasting Sales of Hit Aerosol Products,” *Jurnal Pilar Nusa Mandiri*, vol. 21, no. 2, pp. 153–159, 2025.
[19] A.-C. Akazan et al., “Localized weather prediction using kolmogorov-arnold network-based models and deep rnns,” *arXiv preprint arXiv:2505.22686*, 2025.
[20] Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
[21] F. Chollet, *Deep Learning with Python*. Simon & Schuster, 2021.
[22] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[23] R. Berk, “An introduction to statistical learning from a regression perspective,” in *Handbook of Quantitative Criminology*, Springer, 2009, pp. 725–740.
[24] Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O’Reilly Media, 2022.
[25] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 2013.
[26] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
[27] S. J. Taylor and B. Letham, “Forecasting at scale,” *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.
[28] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
[29] K. Cho et al., “Learning phrase representations using RNN encoder–decoder for statistical machine translation,” in *Proc. EMNLP*, 2014, pp. 1724–1734.
[30] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[31] S. Hochreiter, “The vanishing gradient problem during learning recurrent neural nets and problem solutions,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 6, no. 2, pp. 107–116, 1998.