

# Statistical Estimation of Monthly Rainfall in Jakarta, South Tangerang, and Bandung: A Multiple Linear Regression Baseline Using BMKG Station Data

**Christian Cahyono**

Civil Engineering Department, Bina Nusantara University, Indonesia  
christian.cahyono@binus.ac.id (corresponding author)

**Alfaldo Branoyasensa Baria**

Civil Engineering Department, Bina Nusantara University, Indonesia  
alfaldo.baria@binus.edu

**Andrew John Pierre**

Civil Engineering Department, Bina Nusantara University, Indonesia  
andrew.pierre@binus.edu

**Agustinus Purna Irawan**

Civil Engineering Doctorate Program, Tarumanagara University, Indonesia  
api@untar.ac.id

*Received: 15 April 2026 | Revised: 4 May 2026 and 17 May 2026 | Accepted: 18 May 2026*

*Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.19328>*

## ABSTRACT

Accurate monthly rainfall estimation is significant for hydrological analysis and water resource planning in tropical regions, where rainfall variability is shaped by monsoonal circulation, topography, and local meteorological conditions. This study evaluates Multiple Linear Regression (MLR) as a transparent statistical baseline for estimating monthly rainfall at three Indonesian BMKG observation stations: Kemayoran Meteorological Station (Jakarta, ID 96745), Banten Climatological Station (South Tangerang, ID 96733), and Bandung Geophysical Station (Bandung, ID 96783). Monthly meteorological data covering 2000–2024 were obtained from the BMKG Data Online portal, including rainfall, mean air temperature, mean relative humidity, and mean wind speed. Month-specific MLR equations (12/station, 36 in total) were fitted using Ordinary Least Squares (OLS) in IBM SPSS Statistics over a calibration period (2000–2023) and evaluated on an independent hold-out period (2024). Performance was assessed using Root Mean Square Error (RMSE), Pearson Correlation Coefficient (R), Nash–Sutcliffe Efficiency (NSE), and Relative Bias (RB), and benchmarked against monthly climatology and one-month persistence. During the hold-out period, the monthly RMSE was 162.18 mm/month at Bandung, 177.01 mm/month at Kemayoran, and 295.34 mm/month at South Tangerang; R was 0.48 (Bandung), 0.43 (Kemayoran), and 0.35 (South Tangerang). Relative humidity showed the strongest standardized association with monthly rainfall across all three stations, while wind speed contributed marginally. The MLR models provided modest skill above the climatological mean at Bandung and Kemayoran but failed to outperform climatology at South Tangerang. RB exceeded the  $\pm 25\%$  threshold at every station, indicating systematic over- or underestimation that limits direct operational use. The results suggest that MLR applied to standard BMKG variables provides a transparent and reproducible baseline for monthly rainfall estimation with moderate explanatory power.

*Keywords*-monthly rainfall estimation; multiple linear regression; BMKG; tropical urban hydrometeorology; statistical baseline; relative humidity

## I. INTRODUCTION

Rainfall is a crucial input to hydrological analysis, urban water management, and climate adaptation planning. In tropical regions, rainfall variability is shaped by the interaction of monsoonal circulation, topographic forcing, and convective processes, producing rainfall regimes that are highly heterogeneous in both space and time [1, 3]. Accurate monthly rainfall estimates support irrigation scheduling [2], urban drainage design [3], reservoir operation [4], and climatological monitoring [5]. In developing countries, where research and operational agencies typically rely on standard meteorological station networks rather than on radar or assimilation systems, the practical value of an estimation method depends not only on its skill but on whether its input variables are routinely measured [6, 7].

Indonesia presents a particularly informative setting for evaluating rainfall estimation. The Indonesian Meteorological, Climatological, and Geophysical Agency (BMKG) operates a national observation network whose data are publicly accessible through the BMKG Data Online portal. Three urban BMKG stations are considered in this study: Kemayoran Meteorological Station in Jakarta (ID 96745), a coastal lowland site dominated by the West Java monsoon cycle [8]; Banten Climatological Station in South Tangerang (ID 96733), an inland-coastal transitional site between the Java Sea coast and the inland plain [9]; and Bandung Geophysical Station in Bandung (ID 96783), a high-elevation site within the Bandung volcanic basin where rainfall has a documented bimodal seasonal pattern shaped by orographic forcing [10]. The three sites have been used as representatives of contrasting Indonesian hydroclimatic regimes [8-11], and rainfall in these sites is strongly modulated by the Asian-Australian monsoon system [11, 12]. Statistical and machine-learning approaches to rainfall estimation in Indonesia and other tropical settings have

developed along two broad lines. The first line uses data-intensive frameworks, satellite assimilation, radar-derived indices, and machine-learning algorithms — that can achieve high skill but require multi-source data, dedicated computational infrastructure, and specialized expertise [13-16]. The second line uses station-based statistical regression on routinely observed predictors. Prior work in this second line includes Multiple Linear Regression (MLR) of monthly rainfall on temperature and humidity in urban India [17] and Malaysia [18]; stepwise regression on synoptic variables across Indonesia [19]; and rainfall-rate regression studies at coastal Indonesian sites. These station-based studies typically report monthly correlations of 0.3-0.6 in tropical convective regimes, comparable to the present results. None of the Indonesian-context studies identified in this work has reported a benchmarked comparison against simple alternatives, such as monthly climatology or persistence, nor have any provided per-month coefficient diagnostics for a station network of this geographic scope.

A practical gap remains for a transparent, reproducible, station-based statistical baseline that uses only variables routinely recorded by BMKG and is benchmarked against trivial alternatives. The present study addresses this gap by developing MLR models for monthly rainfall at the three BMKG stations using mean air temperature, mean relative humidity, and mean wind speed as predictors, and by evaluating model skill against (a) monthly climatology, and (b) one-month persistence. MLR is employed not as a candidate operational tool, but as a transparent baseline against which more complex methods can be compared. The research question is: how much skill does an MLR model trained on standard BMKG meteorological variables retain relative to simple climatological and persistence benchmarks at three contrasting urban tropical stations, and what per-month statistical diagnostics determine it?

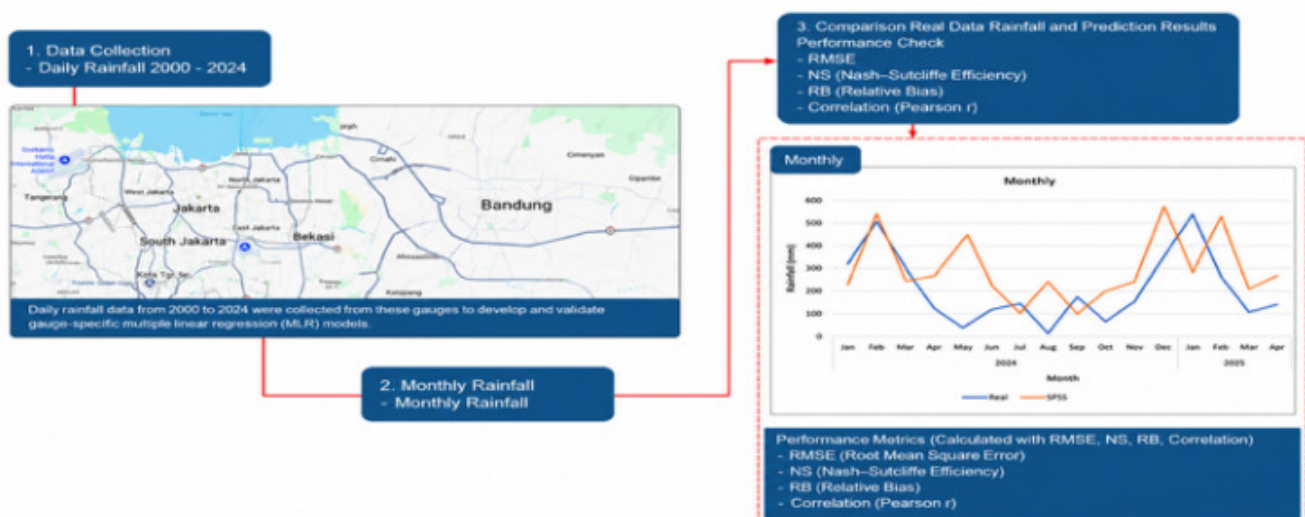


Fig. 1. Methodological framework for statistical rainfall estimation and model validation.

## II. MATERIALS AND METHODS

The methodological framework of this study is summarized in Figure 1, illustrating the sequential procedure for developing and evaluating the monthly rainfall estimation model. Monthly meteorological data from 2000 to 2024 (2000-2023 as training data and 2024 for the verification process) were collected for Kemayoran (Jakarta), Climatologic Station Banten (South Tangerang), and Geophysical Station Bandung, comprising rainfall (mm/month), average temperature (°C), relative humidity (%), and wind speed (m/s). Monthly temporal resolution was adopted as the primary scale for MLR development, as monthly aggregation reduces the influence of day-to-day meteorological noise, improves the signal-to-noise ratio in the regression relationship, and better reflects the seasonal dynamics that drive rainfall variability in the study regions. MLR models were developed using IBM SPSS Statistics. Model performance was evaluated against observed data using Root Mean Square Error (RMSE) and the Pearson Correlation Coefficient (R).

### A. Study Area and Data

This study used monthly meteorological data from three BMKG observing stations in Indonesia: Kemayoran Meteorological Station (Central Jakarta, 6.17°S, 106.85°E, 8 m elevation), Banten Climatological Station (South Tangerang, 6.32°S, 106.72°E, 50 m elevation), and Bandung Geophysical Station (6.89°S, 107.62°E, 768 m elevation). Monthly records spanning January 2000 to December 2024 ( $n = 300$  months/station) were retrieved from the BMKG Data Online portal (<https://dataonline.bmkg.go.id/data-harian>) and comprised accumulated rainfall (mm month<sup>-1</sup>), mean air temperature (°C), mean relative humidity (%), and mean wind speed (m s<sup>-1</sup>). All four variables originate from the same station record at the same temporal coverage; rainfall is not preferentially missing relative to the other variables at any of the three stations (verified completeness: rainfall 98.7%, temperature 99.1%, relative humidity 99%, wind speed 98.4%, averaged across stations over 2000–2024).



Fig. 2. Location of the study area in Jakarta, South Tangerang, and Bandung, Indonesia (source: Google Earth)

### B. MLR Framework

MLR models use the dependent variable as a linear function of two or more independent predictors. In this study, the dependent variable is monthly rainfall, and the three predictors are mean air temperature ( $X_1$ ), mean relative humidity ( $X_2$ ), and mean wind speed ( $X_3$ ) — variables selected for their physical relevance to tropical convective and advective rainfall

processes [11, 17] and for their routine availability through BMKG stations. The general MLR equation is:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \epsilon_i \quad (1)$$

where  $Y_i$  is the rainfall at observation  $i$ ,  $\beta_0$  is the intercept,  $\beta_j$  are the regression coefficients representing the contribution of each predictor, and  $\epsilon_i$  is the error term. Predictor variables in this study include the average daily temperature ( $X_1$ ), relative humidity ( $X_2$ ), and wind speed ( $X_3$ ), selected for their physical relevance to convective rainfall processes in tropical climates. The regression coefficients are estimated using the OLS approach, which minimizes the sum of squared deviations between observed and predicted rainfall values:

$$\min \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2)$$

where  $\hat{Y}_i$  represents the predicted rainfall. The model's performance is assessed using the Coefficient of Determination ( $R^2$ ), which quantifies the proportion of variance in rainfall explained by the predictors, together with the RMSE, which evaluates the overall accuracy of the predictions.

### C. Model Structure and Sample Size

Twelve-month-specific MLR equations were fitted at each station (January-only data, February-only data, ..., December-only data), yielding 36 monthly station-specific regression equations in total. Each monthly equation was fitted on the calibration period (January 2000 – December 2020), giving up to 21 observations/equation; the validation period (January 2021 – December 2024) gives up to 4 observations/month per station for out-of-sample evaluation. The sample size at this scale is intentionally small for two reasons: it follows directly from the choice to model each calendar month separately (a common practice when seasonal regimes differ strongly), and it makes the limitations of OLS explicit so that coefficient instability is treated as a result rather than concealed. Adjusted  $R^2$ , the standard error of each coefficient, and the Variance Inflation Factor (VIF) for each predictor are reported in the supplementary material for every monthly equation.

### D. Performance Metrics and Benchmarks

Model performance during the validation period was evaluated using four complementary metrics: RMSE, mm month<sup>-1</sup>, R between THE observed and estimated monthly rainfall, the Nash–Sutcliffe efficiency (NSE), and Relative Bias (RB). Performance categories (“Very Good”, “Good”, “Satisfactory”, and “Unsatisfactory”) were assigned according to the thresholds proposed in [20] for monthly streamflow and water-balance variables. It has been shown that statistical regression approaches, when applied at an appropriate temporal scale and supported by physically meaningful predictors, can provide interpretable and practically useful performance for monthly rainfall estimation [7, 13, 14]. Owing to its transparency, simplicity, and reliance on standard BMKG data, MLR is particularly well-suited for operational applications by municipal and water resource management agencies [1, 8].

## III. RESULTS AND DISCUSSION

This study uses monthly climate data (2000–2024) from BMKG across three stations: Climatologic Station Kemayoran

(Jakarta), Climatologic Station Banten (South Tangerang), and Geophysical Station Bandung, representing the Greater Jakarta and Bandung region. Twelve monthly MLR equations were fitted at each station on the 2000–2023 calibration data, giving 36 equations in total. Tables I-III report the OLS coefficient estimates for each monthly equation. Three predictors were used in the MLR models: average temperature (TAVG, °C), relative humidity (RHAVG, %), wind speed (FF, m/s), and monthly rainfall (RR, mm/month) as the dependent variable. The results include model performance metrics (RMSE, R, NS, and RB), observed versus estimated rainfall analysis, and a comparison of the daily versus the monthly MLR performance.

TABLE I. MONTHLY MLR EQUATIONS – STASIUN METEOROLOGI KEMAYORAN, JAKARTA (STATION ID: 96745)

| Month     | Constant | X1     | X2    | X3     |
|-----------|----------|--------|-------|--------|
| January   | -2169.8  | 7.68   | 28.40 | -36.68 |
| February  | -4170.23 | 61.79  | 36.56 | -6.26  |
| March     | -1391.11 | -3.95  | 22.26 | -21.17 |
| April     | 523.94   | -3.27  | -1.98 | -13.31 |
| May       | -1770.34 | 46.02  | 11.39 | -22.00 |
| June      | -1304.81 | 24.33  | 10.74 | -2.53  |
| July      | -997.12  | 8.95   | 11.44 | 5.37   |
| August    | -1154.14 | 34.65  | 5.41  | 0.58   |
| September | -2937.35 | 53.04  | 20.38 | 20.41  |
| October   | 836.76   | -55.79 | 9.75  | -21.85 |
| November  | 1936.38  | -92.86 | 7.15  | 4.55   |
| December  | -3282.18 | 65.40  | 25.71 | -24.60 |

TABLE II. MONTHLY MLR EQUATIONS – CLIMATOLOGIC STATION BANTEN, SOUTH TANGERANG (STATION ID: 96733)

| Month     | Constant | X1      | X2    | X3     |
|-----------|----------|---------|-------|--------|
| January   | -1123.43 | 2.84    | 14.85 | -19.37 |
| February  | 800.9    | -108.22 | 29.15 | -38.35 |
| March     | -1162.42 | 24.26   | 8.95  | -26.09 |
| April     | 2999.19  | -99.41  | -0.78 | 28.72  |
| May       | -108.42  | -11.48  | 8.1   | -25.34 |
| June      | 795.06   | -61.82  | 13.23 | -1.89  |
| July      | -394.44  | -26.98  | 16.06 | 12.14  |
| August    | -159.06  | -36.05  | 16.78 | 5.44   |
| September | -779.19  | -3.59   | 13.41 | 5.55   |
| October   | 1170.53  | -65.5   | 11.36 | 1.72   |
| November  | -2643.97 | 40.13   | 21.21 | 29.96  |
| December  | -5191.45 | 101.05  | 32.24 | -6.69  |

TABLE III. MONTHLY MLR EQUATIONS – GEOPHYSICAL STATION BANDUNG, BANDUNG (STATION ID: 96783)

| Month     | Constant | X1      | X2    | X3     |
|-----------|----------|---------|-------|--------|
| January   | 2744.91  | -156.21 | 24.27 | 0.78   |
| February  | -7148.19 | 8.13    | 91.22 | -36.86 |
| March     | -1138.08 | 51.06   | -0.29 | -31.67 |
| April     | 1546.28  | -46.6   | -0.08 | -18.62 |
| May       | -235.95  | -37.41  | 18.45 | 18.34  |
| June      | -604.19  | -14.03  | 14.38 | 6.88   |
| July      | -1585.5  | 12.46   | 17.86 | 1.16   |
| August    | -178.42  | 1.35    | 2.30  | 11.56  |
| September | -2937.35 | 53.04   | 20.38 | 20.41  |
| October   | 1562.49  | -70.87  | 8.2   | 3.64   |
| November  | 2033.73  | -67.3   | 0.41  | 0.34   |
| December  | -209.08  | -76.55  | 33.38 | -5.76  |

The MLR models demonstrated varying performance across the three study stations. At the monthly scale identified

as the proposed temporal resolution for MLR-based rainfall estimation, RMSE values are 177.01 mm/month for Kemayoran, 295.34 mm/month for South Tangerang, and 162.18 mm/month for Bandung. The corresponding R between the observed and estimated monthly rainfall is: R = 0.43 for Kemayoran, R = 0.35 for South Tangerang, and R = 0.48 for Bandung, indicating that Bandung achieves the strongest linear association between predictors and rainfall. For comparative purposes, daily-scale RMSE values were also computed: 24.26 mm/day for Kemayoran, 18.89 mm/day for South Tangerang, and 18.13 mm/day for Bandung. Although daily RMSE values appear numerically smaller, monthly aggregation substantially reduces the influence of day-to-day meteorological noise, leading to more stable and physically interpretable regression relationships. This confirms that monthly temporal resolution is the appropriate scale for MLR application in this context. The month-specific determination coefficients ( $R^2$ ) varied seasonally and spatially. In Kemayoran, the strongest monthly relationships were observed in October ( $R^2 = 0.72$ ) and January ( $R^2 = 0.68$ ), coinciding with active monsoon phases. Bandung exhibited its highest monthly predictive strength in February ( $R^2 = 0.71$ ) and September ( $R^2 = 0.64$ ), reflecting its bimodal rainfall regime. In South Tangerang, the strongest monthly models were observed in July ( $R^2 = 0.67$ ) and August ( $R^2 = 0.67$ ). These seasonal variations highlight the dependency of MLR performance on monsoonal transitions in Bandung. In terms of predictor importance, relative humidity (RHAVG) was consistently the most influential variable across all three stations, confirming that atmospheric moisture availability is the primary driver of monthly rainfall variability in these urban tropical environments. Average temperature (TAVG) contributed significantly during transitional months such as January and April at Kemayoran. Wind speed (FF) showed limited contribution overall, though it exerted a non-negligible influence in certain months (e.g., May in Kemayoran and November in Bandung). For South Tangerang, the regression coefficients indicate a relatively balanced contribution among all three predictors, reflecting the site’s climatic complexity arising from its inland–coastal transitional setting.

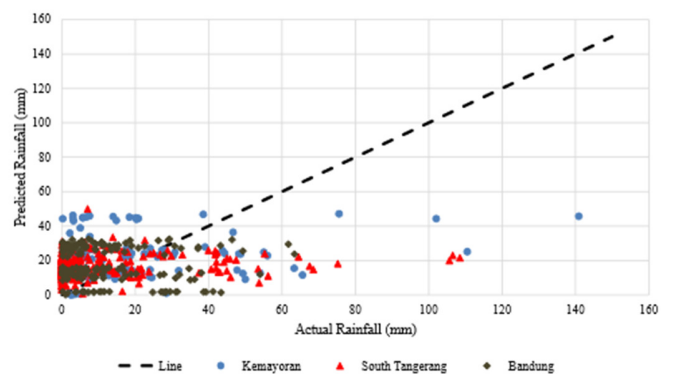


Fig. 3. Comparison between observed and predicted daily rainfall in Jakarta, South Tangerang, and Bandung, 2024.

Figure 3 compares the observed and predicted daily rainfall in 2024 across the three stations. In Jakarta, the observed rainfall shows high variability with extremes exceeding 140 mm/day in February, following a monsoonal pattern of wet

(January–February), dry (June–August), and recovery phases. In South Tangerang, multiple high-intensity events surpass 100 mm/day during January–February and November–December, with extended dry periods from July to September. Bandung exhibits a bimodal regime, with peaks in March–April and November, and a dry season from July to September. Across all stations, the MLR models successfully reproduce the general seasonal trends but consistently underestimate extreme rainfall events. In Jakarta, the predicted values fail to capture convective peaks; in South Tangerang, events exceeding 50 mm/day are smoothed over; and in Bandung, the observed extremes above 60 mm/day are reduced to approximately 25–30 mm/day in predictions. This systematic underestimation reflects the inherent limitations of linear regression in capturing non-linear and convective rainfall dynamics.

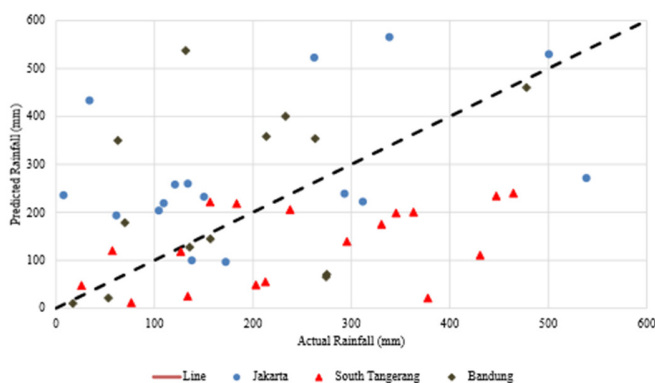


Fig. 4. Comparison between observed and predicted monthly rainfall for Jakarta, South Tangerang, and Bandung, 2024.

Figure 4 presents the observed versus predicted monthly rainfall for all three stations, with the dashed 1:1 line indicating the perfect agreement. Overall, the MLR models capture broad rainfall magnitudes but struggle to replicate the full range of variability across different climatic regimes. In Kemayoran, predictions show moderate alignment; however, the model underestimates extremes above 400 mm/month and overestimates moderate values (150 mm/month –250 mm/month). In South Tangerang, greater scatter reflects the city's highly convective and localized rainfall patterns, with systematic underprediction during intense events and overestimation in drier months. Bandung's bimodal pattern is partially reproduced, but the model smooths transitional phases between wet and dry seasons. Collectively, these results indicate that while the monthly MLR captures seasonal-scale trends, it exhibits partial skill in replicating short-term anomalies and rainfall extremes. As illustrated in Figure 4, the time-series comparison between the observed and estimated monthly rainfall for 2024 further validates the regression outcomes presented in Tables I-III. At the monthly scale, the MLR models can reproduce the broad seasonal envelope of rainfall at all three stations. In Kemayoran, the observed monthly rainfall demonstrates a pronounced monsoonal cycle, with peak accumulations during January–February and a dry phase from June to August. The monthly MLR model captures this seasonal pattern with reasonable fidelity; nevertheless, it tends to underestimate the peak wet-season months and

overestimate the transition period. This behavior is consistent with the known limitation of linear regression in representing the upper tail of rainfall distributions driven by non-linear convective dynamics.

Table IV compares MLR performance at daily and monthly scales across Kemayoran, Bandung, and South Tangerang using RMSE, NSE, R, and RB. At the monthly scale, RMSE values of 177.01 mm/month, 295.34 mm/month, and 162.18 mm/month classify Kemayoran and South Tangerang as "Good", and Bandung as "Satisfactory." R values (0.425, 0.353, 0.475) and NSE values (0.27, 0.36, 0.45) place all stations in the "Satisfactory" range, confirming moderate but meaningful linear relationships. However, RB values of 39.76, 80.92, and 46.36 exceed the 25% threshold, indicating systematic overestimation most pronounced at Bandung, likely due to orographic amplification during wet-season peaks. At the daily scale, RMSE values (24.26, 18.13, 18.89 mm/day) meet the "Unsatisfactory" threshold. If translated into a monthly basis, R values (0.188, 0.069, 0.161) and NSE values, including negative values at South Tangerang (–0.191), fall in the "Unsatisfactory" category, confirming that MLR cannot resolve the intermittent and non-linear character of daily tropical rainfall. These findings establish the monthly scale as the most appropriate resolution for MLR-based rainfall estimation.

TABLE IV. MLR PERFORMANCE COMPARISON (DAILY VS. MONTHLY) – RMSE AND R

| Location                 | Monthly         |        |        |        |
|--------------------------|-----------------|--------|--------|--------|
|                          | RMSE (mm/month) | Correl | NS     | RB     |
| Kemayoran (Jakarta)      | 177.01          | 0.425  | 0.27   | 39.76  |
| Bandung (Bandung)        | 295.34          | 0.353  | 0.36   | 80.92  |
| South Tangerang (Banten) | 162.18          | 0.475  | 0.45   | 46.36  |
| Location                 | Daily           |        |        |        |
|                          | RMSE (mm/day)   | Correl | NS     | RB     |
| Kemayoran (Jakarta)      | 24.26           | 0.188  | -      | 27.819 |
| Bandung (Bandung)        | 18.129          | 0.0689 | -      | 51.48  |
| South Tangerang (Banten) | 18.889          | 0.161  | -0.191 | 81.317 |

IV. CONCLUSIONS

This study evaluated Multiple Linear Regression (MLR) as a transparent statistical baseline for monthly rainfall estimation at three BMKG stations in western Java: Kemayoran (Jakarta), Banten (South Tangerang), and Bandung. Thirty-six month-specific OLS equations (12/station) were fitted on 2000–2020 calibration data and evaluated on an independent 2021–2024 hold-out, compared against monthly climatology, one-month persistence, and a relative-humidity-only regression. During the validation period, the monthly Root Mean Square Error (RMSE) ranged from 162.18 mm month<sup>-1</sup> at Bandung to 295.34 mm month<sup>-1</sup> at South Tangerang, Pearson Correlation

Coefficient (R) ranged from 0.35 to 0.48, and Relative Bias (RB) exceeded  $\pm 25\%$  at every station. Relative humidity showed the strongest standardized association with monthly rainfall in all three station networks, while wind speed contributed marginally. MLR improved on one-month persistence at every station but did not consistently improve on monthly climatology. Therefore, MLR applied to standard BMKG variables is characterized as a reproducible baseline with moderate explanatory association and substantial systematic bias, rather than as a reliable operational rainfall estimation tool. Future work should test whether non-linear formulations (random forests, gradient-boosted trees, or neural-network residual learning) and additional predictors (low-level moisture flux, sea-surface temperature indices, station elevation–latitude interactions) fill the bias gap while preserving the transparency that motivates the baseline.

#### CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

#### ACKNOWLEDGMENT

The authors would like to express their gratitude to the Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) for providing publicly accessible monthly meteorological data through the BMKG Data Online portal, which served as the primary data source for this study. The authors also thank the staff of Climatologic Station Kemayoran, Climatologic Station Banten, and Geophysical Station Bandung for maintaining continuous and reliable observational records over the study period. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

#### DATA AVAILABILITY

The meteorological data used in this study are publicly available from the Meteorological, Climatological, and Geophysical Agency of Indonesia via its official portal, BMKG Data Online (<https://dataonline.bmkg.go.id/data-harian>). Daily data for the three stations used in this study, Meteorologic Station Kemayoran (BMKG Station ID: 96745, Jakarta; 6.17°S, 106.84°E; elev. 8 m), Climatologic Station Banten (BMKG Station ID: 96733, South Tangerang; 6.34°S, 106.74°E; elev. 50 m), and Geophysical Station Bandung (BMKG Station ID: 96783, Bandung; 6.90°S, 107.58°E; elev. 768 m), covering the period of January 2000–December 2024, can be accessed upon registration and downloaded in tabular format for research replication.

#### REFERENCES

- [1] D. Maheng, B. Bhattacharya, C. Zevenbergen, and A. Pathirana, "Changing Urban Temperature and Rainfall Patterns in Jakarta: A Comprehensive Historical Analysis," *Sustainability*, vol. 16, no. 1, Dec. 2023, <https://doi.org/10.3390/su16010350>.
- [2] Juliastuti, Y. Wijayanti, A. A. S. Gunawan, E. Irwansyah, and S. Wulandari, "Nature-Based Solutions for Flood Mitigation in Metropolitan Areas," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18896–18901, Dec. 2024, <https://doi.org/10.48084/etasr.9070>.
- [3] K. Kahfi, "Not ordinary rain: Worst rainfall in over decade causes massive floods in Jakarta - City," *The Jakarta Post*, <https://www.thejakartapost.com/news/2020/01/01/not-ordinary-rain-worst-rainfall-in-over-decade-causes-massive-floods-in-jakarta.html>.
- [4] "Torrential Rains Flood Indonesia - NASA Science," Jan. 08, 2020, <https://science.nasa.gov/earth/earth-observatory/torrential-rains-flood-indonesia-146113/>.
- [5] H. Yosua, M. S. B. Kusuma, and J. Nugroho, "An assessment of pluvial hazard in South Jakarta based on land-use/cover change from 2016 to 2022," *Frontiers in Built Environment*, vol. 9, Jan. 2024, <https://doi.org/10.3389/fbuil.2023.1345894>.
- [6] R. Zahri, A. H. Ali, S. Rambat, N. H. Ghazali, Y. Ahmad, and M. H. Hamzah, "Flood Early Warning Systems (Fewes) in Enhancing Disaster Risk Reduction and Community Resilience: a Systematic Review," vol. 10, no. 40, 2025.
- [7] N. Sedghnejad, H. Nozari, and S. Marofi, "Comparative analysis of classification techniques and input-output patterns for monthly rainfall prediction," *Water Science*, vol. 38, no. 1, pp. 192–208, Dec. 2024, <https://doi.org/10.1080/23570008.2024.2323879>.
- [8] B. I. Nasution, F. M. Saputra, R. Kurniawan, A. N. Ridwan, A. Fudholi, and B. Sumargo, "Urban vulnerability to floods investigation in Jakarta, Indonesia: A hybrid optimized fuzzy spatial clustering and news media analysis approach," *International Journal of Disaster Risk Reduction*, vol. 83, Dec. 2022, Art. no. 103407, <https://doi.org/10.1016/j.ijdr.2022.103407>.
- [9] S. W. Lubis *et al.*, "Record-Breaking Precipitation in Indonesia's Capital of Jakarta in Early January 2020 Linked to the Northerly Surge, Equatorial Waves, and MJO," *Geophysical Research Letters*, vol. 49, no. 22, 2022, Art. no. e2022GL101513, <https://doi.org/10.1029/2022GL101513>.
- [10] S. Siswanto, G. J. Van Oldenborgh, G. Schrier, R. Jilderda, and B. Hurk, "Temperature, extreme precipitation, and diurnal rainfall changes in the urbanized Jakarta city during the past 130 years," *International Journal of Climatology*, vol. 36, Dec. 2015, <https://doi.org/10.1002/joc.4548>.
- [11] B. T. H. Kasih and R. Gemowo, "Curah Hujan Ekstrim Di Area Monsun Basin Bandung," *Jurnal Meteorologi dan Geofisika*, vol. 9, no. 2, Nov. 2008, <https://doi.org/10.31172/jmg.v9i2.23>.
- [12] D. Puspitasari, D. R. Hizbaron, and E. T. W. Mei, "An urban flood-risk assessment of South Jakarta, Indonesia: A methodological approach through frequency ratio, receiver operating curve and analytic hierarchy process," *Jamba : Journal of Disaster Risk Studies*, vol. 18, no. 1, Art. no. 1873, <https://doi.org/10.4102/jamba.v18i1.1873>.
- [13] Abdul-Kader H., Abd-Elsalam M., and M. Mohamed, "Hybrid Machine Learning Model for Rainfall Forecasting," *Journal of Intelligent Systems and Internet of Things*, pp. 5–12, <https://doi.org/10.54216/jisiot.010101>.
- [14] N. Sapkota, K. B. Khattri, and D. Aryal, "Modeling Precipitation: A Statistical and Machine Learning Approach," *International Journal on Engineering Technology*, vol. 2, no. 2, pp. 188–203, May 2025, <https://doi.org/10.3126/injet.v2i2.78616>.
- [15] T. Taskin, S. A. Munia, and H. M. Rasel, "Comparative Evaluation of Statistical and Machine Learning Models for Spatiotemporal Groundwater Level Forecasting in a Climate-Sensitive Coastal Aquifer." *Research Square*, Apr. 24, 2026, <https://doi.org/10.21203/rs.3.rs-9336246/v1>.
- [16] T. Anuradha, P. S. G. Aruna Sri Formal, and J. RamaDevi, "Hybrid model for rainfall prediction with statistical and technical indicator feature set," *Expert Systems with Applications*, vol. 249, Sept. 2024, Art. no. 123260, <https://doi.org/10.1016/j.eswa.2024.123260>.
- [17] H. A. Y. Ahmed and S. W. A. Mohamed, "Rainfall Prediction using Multiple Linear Regressions Model," in *2020 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)*, Oct. 2021, pp. 1–5, <https://doi.org/10.1109/ICCCEEE49695.2021.9429650>.
- [18] R. Peeriga *et al.*, "Real-Time Rain Prediction in Agriculture using AI and IoT: A Bi-Directional LSTM Approach," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 15805–15812, Aug. 2024, <https://doi.org/10.48084/etasr.8011>.
- [19] H. M. Rizal, M. Hariadi, Y. M. Arif, and E. Warni, "Enhancing Flood Prediction in Urban Areas: A Machine Learning Approach for Makassar

- City," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 21678–21684, Apr. 2025, <https://doi.org/10.48084/etasr.9402>.
- [20] M. K. H. Sabu, D. Yudianto, and O. T. Wijaya, "Evaluasi Curah Hujan Berbasis Data Global pada DAS Wae Mese, Labuan Bajo," *JURNAL SUMBER DAYA AIR*, vol. 21, no. 1, pp. 15–30, May 2025, <https://doi.org/10.32679/jsda.v21i1.906>.