PARAMETER
JURNAL MATEMATIKA, STATISTIKA DAN TERAPANNYA

e-ISSN. 2829-6303
Vol 04 No 3, December 2025
Page 591-601

Enhancing Rainfall Forecasting Performance in Bandung City Using Bi-LSTM with Grid Search Optimization on Gregorian and Lunar Calendar Data

Mahdayani Putri Yunizar [1*], Andrew Hosea Talakua [2], Gumgum Darmawan [3]

[1,2,3] Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran
Jl. Raya Bandung–Sumedang KM 21, Hegarmanah, Jatinangor, Sumedang 45363, West Java, Indonesia

E-mail Correspondence Author: *mahdayani25001@mail.unpad.ac.id

Abstract

Rainfall is a climatic factor that strongly influences human activities and plays a crucial role in decision making related to water resources, mobility, and disaster preparedness. High rainfall intensity may escalate into hydrometeorological hazards, underscoring the importance of accurate rainfall forecasting to support early warning and mitigation efforts. This study aims to compare the forecasting accuracy of monthly rainfall predictions between the Gregorian and lunar calendars using the Bidirectional Long Short-Term Memory (Bi-LSTM) model optimized through a grid search approach. The method is designed to capture temporal patterns arising from the distinct structures of two asynchronous calendars. Daily rainfall data from Bandung City, Indonesia, covering the period from 2000 to 2025, were converted into monthly series in both calendar systems. The results reveal that the Gregorian calendar provides significantly better forecasting performance, achieving the lowest MAPE value of 11.60 percent at the three-month horizon. In contrast, the lunar calendar shows higher variability and reaches its best MAPE of 31.43 percent at the same horizon. These findings indicate that the Gregorian calendar offers a more stable temporal representation for rainfall forecasting in Bandung and supports improved predictive modeling for climate-related decision making.

Keywords: *Bi-LSTM, calendar conversion, rainfall forecasting, time series, tropical climatology*

: https://doi.org/10.30598/parameterv4i1pp55-62

1. INTRODUCTION

Rainfall constitutes a fundamental component of hydrological, agricultural, and urban environmental systems, particularly in tropical regions such as Indonesia. The capacity to generate accurate rainfall forecasts is essential to support water resource management, disaster risk reduction, and climate-sensitive urban planning [1]. Bandung City, situated in the highland area of West Java, exhibits complex spatiotemporal rainfall dynamics due to its basin morphology and surrounding orographic structures. These characteristics generate high variability in precipitation, thereby reducing the effectiveness of traditional statistical forecasting approaches. The intensification of climate variability further underscores the need for more adaptive and data-driven forecasting techniques in regional climatology [2].

Rainfall forecasting has conventionally relied on the Gregorian calendar as the primary temporal reference. However, several atmospheric studies have reported that lunar phases and gravitational atmospheric tides influence cloud formation and precipitation, indicating that the lunar calendar may reveal cyclic rainfall variations that are not captured by the Gregorian system [3]. Incorporating lunar time structures has been shown to enhance the identification of periodic climatic signals, resulting in improved forecast accuracy in several recent studies [4]. Consequently, analytical approaches that integrate both Gregorian and lunar calendars have emerged as a promising direction in contemporary climate modeling research [5].

Advancements in artificial intelligence have further strengthened rainfall forecasting methodologies. Deep learning models, particularly Long Short-Term Memory (LSTM) and Bidirectional LSTM (Bi LSTM), demonstrate superior capability in capturing nonlinear temporal dependencies and long-range interactions within precipitation time series [6]. Despite these advantages, the predictive effectiveness of such models is greatly influenced by the selection of hyperparameters. Systematic optimization methods such as grid search are therefore required to ensure robust model performance and to minimize structural bias in the learning process [7].

Although Bi LSTM combined with hyperparameter optimization has been employed in various meteorological studies, the application of an integrated Gregorian and lunar calendar framework remains limited, particularly in West Java. A recent investigation conducted in Bogor City revealed that rainfall modeling based on the lunar calendar achieved higher accuracy than that based solely on the Gregorian calendar, thereby highlighting the potential benefits of dual temporal structures [8]. Bandung City, despite possessing distinct climatological and topographical characteristics, has not yet been examined using a dual calendar and a deep learning optimized forecasting framework. This gap indicates the need for more comprehensive research that integrates alternative temporal representations with advanced predictive modeling techniques.

To address this research gap, the present study proposes an enhanced rainfall forecasting framework for Bandung City by integrating Gregorian and lunar calendar time series into a Bi LSTM model optimized through grid search. The scientific contribution of this work lies in the formulation of a dual calendar based temporal transformation combined with systematic hyperparameter optimization, which collectively aims to improve forecasting accuracy beyond that achieved by conventional single calendar deep learning models. The outcomes of this research are expected to advance regional climate forecasting methodologies and provide robust analytical support for climate resilient urban planning in Bandung City [9].

2. METHOD

2.1. Data Acquisition

Daily rainfall data for Bandung City were obtained from the official portal of the Indonesian Meteorology, Climatology, and Geophysics Agency (BMKG). The daily rainfall series is represented as $D = \{r_t \mid t = 1,2,\dots,T\}$, where $r_t$denotes the rainfall on day t. Long-term datasets are essential because extended temporal records allow richer exploration of seasonal and periodic features, which is critical for time-series modeling using deep learning methods such as Long Short-Term Memory (LSTM). Similar approaches have been adopted in studies of rainfall prediction in tropical regions, including work in Jambi City, Indonesia, using daily rainfall data from 2016–2024 [10].

Before analysis, all records were validated and preprocessed to handle missing values, ensure continuity and consistency, and normalize the data. This process ensures that the rainfall

series meets quality standards necessary for predictive modeling, reducing biases caused by incomplete or erroneous observations [11]. Thus, the daily series Dcan be considered a reliable representation of historical rainfall complete, continuous, and of sufficient quality suitable as univariate input for LSTM-based predictive modeling.

2.2. Data Preprocessing

Preprocessing included missing value handling, normalization, and temporal aggregation.

2.2.1. Misssing Value Treatment

Missing values were imputed using mean imputation:

$$r_t^* = \begin{cases} r_t, & if\ r_t\ is\ observed \\ \bar{r}, & if\ r_t\ is\ missing \end{cases}$$

Where $\bar{r}$ is the mean of available observations. This method is widely used in hydrometeorological data processing and preserves statistical characteristics of the dataset [12].

2.2.2. Normalization

Min Max Scaling was applied to stabilize model convergence:

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

Denormalization after prediction used:

$$x_{denorm} = x_{norm}(x_{max} - x_{min}) + x_{min} \quad (2)$$

Normalization is essential in deep learning to ensure numerical stability during optimization [11].

2.3. Gregorian to Lunar Calendar Transformation

Transformation into the lunar calendar followed a structured temporal segmentation approach that reconstructs monthly rainfall totals according to lunar month boundaries.
Monthly rainfall in calendar system C is defined as:

$$R_m^C = \sum_{t \in M_C(m)} r_t \quad (3)$$

Where:

1. R_m^C: rainfall in month $m$for calendar C
2. $M_C(m)$: set of days in month $m$for Gregorian or lunar system

The dual calendar representation is:

$$R_{dual}(m) = \{R_{greg}(m), R_{lunar}(m)\} \quad (4)$$

The rationale for adopting this dual-calendar approach is supported by empirical studies that found statistically significant correlations between lunar phases (or lunar cycles) and rainfall variability. For instance, a recent study covering 42 years of rainfall data for a major city in Pakistan reported that incorporating lunar-phase information alongside meteorological factors improved the performance of rainfall/extreme-rainfall prediction models [13]. Another study across a very large network of weather stations in Mexico showed that historical cumulative rainfall exhibited systematic variation with lunar phases: rainfall was most frequently highest during the New Moon phase and lowest during Waning-Crescent (and related) phases [14].

2.4. Bidirectional Long Short-Term Memory (Bi-LSTM)

Long Short-Term Memory (LSTM) networks were originally introduced to overcome the vanishing gradient limitation commonly encountered in standard recurrent neural networks, particularly when modeling long temporal sequences. Each LSTM cell is equipped with three gating mechanisms—namely the forget gate, input gate, and output gate—which regulate how information is retained, updated, or discarded throughout the learning process. Owing to this structure, LSTM has been successfully applied in various domains including sentiment classification [15], analysis of public responses to COVID-19 vaccination [16], and human activity recognition using smartphone sensor data [17]. Despite these advantages, LSTM architectures often involve high computational costs due to their extensive internal operations, motivating the exploration of more efficient variants with comparable predictive power. One such extension is the Bidirectional LSTM (Bi-LSTM).

Bi-LSTM was introduced by Graves and Schmidhuber as an enhancement to both traditional recurrent neural networks and the unidirectional LSTM architecture, which propagate information solely forward in time. In such models, the hidden representation at time step t depends only on

prior temporal information. In contrast, Bi-LSTM incorporates two separate LSTM layers: one processing the input sequence in the forward direction and another processing it in reverse. This bidirectional structure enables the model to extract contextual information from both past and future observations within a given sequence, thereby providing a richer temporal representation [18]. The additional feature extraction performed by these dual layers has been shown to substantially improve model performance across a variety of prediction tasks [19].

By learning temporal dependencies in two directions simultaneously, Bi-LSTM is able to preserve information from long-range sequences more effectively than its unidirectional counterpart. Its architectural design prevents the model from prematurely forgetting earlier or later observations during training, thus offering improved stability in long sequence modeling [20]. As a result, Bi-LSTM often demonstrates superior predictive accuracy compared to conventional LSTM networks [28]. The general architecture of the Bi-LSTM model is illustrated in Figure 1 [21].

In this architecture, the forward LSTM layer operates in the same order as a standard LSTM, producing hidden states at t–1, t, and t+1. Meanwhile, the backward layer processes the same sequence in reverse order, generating hidden states from t+1 backward to t–1. These two hidden states, denoted by the forward component $h_t^{\rightarrow}$ and the backward component $h_t^{\leftarrow}$, are jointly utilized to form a more comprehensive output representation. According to [22], the hidden state updates for the two directions can be expressed as:

$$h_t^{\rightarrow} = \mathrm{LSTM}(x_t, h_{t-1}^{\rightarrow}), \qquad h_t^{\leftarrow} = \mathrm{LSTM}(x_t, h_{t+1}^{\leftarrow}) \tag{5}$$

The final output at time t is then computed by combining the contributions of both directional hidden states as formulated in [23]:

$$y_t = U_y h_t^{\rightarrow} + W_y h_t^{\leftarrow} + b_y \tag{6}$$

where U_y and W_y represent the weight matrices associated with the forward and backward outputs, respectively, and b_y denotes the output bias.

The application of Bi-LSTM has been widely documented in multiple fields. Prior studies have utilized this architecture for forecasting wastewater flow rates [6], predicting the development of tropical cyclones [24], estimating soil and groundwater moisture content [25], and modeling streamflow variations in major river basins [26]. These studies collectively demonstrate the capability of Bi-LSTM to deliver enhanced accuracy in time series prediction tasks that exhibit complex temporal dependencies.

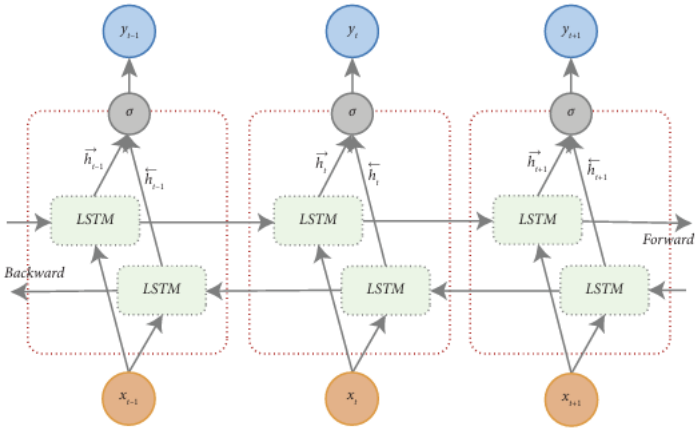

Figure 1. Architecture of Bidirectonal LSTM (Bi-LSTM)

2.5. Hyperparameter Optimization Using Grid Search

In constructing machine learning models such as Bi-LSTM, selecting an appropriate set of hyperparameters is essential because these parameters directly influence model performance, stability, and learning efficiency. Hyperparameter tuning enables the model to operate under its optimal configuration by systematically evaluating predefined parameter combinations [7]. In this

study, the Bi-LSTM architecture is enhanced through an integrated grid search procedure, as illustrated in Figure 3, to ensure that the resulting forecasting model achieves the highest possible accuracy.

The Bi-LSTM–grid search architecture consists of several key components: an input layer, a Bi-LSTM processing layer, a dropout layer, a dense layer, an output layer, and a hyperparameter search module. The input layer receives the sequential rainfall data, while the Bi-LSTM layer performs dual-direction temporal feature extraction. To mitigate overfitting, dropout regularization is inserted between layers, reducing the co-adaptation of neurons during training [5]. The dense layer acts as the final transformation layer that maps the extracted features into numerical predictions, and the output layer generates the final rainfall forecast.

The grid search procedure systematically explores the predefined hyperparameter space: number of hidden neurons (hidden units), batch size, number of epochs, dropout rate; and where applicable learning rate and optimizer type. For each candidate combination, the model is trained and evaluated using k-fold cross-validation. In this study, a 5-fold cross-validation scheme is adopted to balance computational cost and validation robustness. The configuration that achieves the lowest average validation error is selected as the optimal hyperparameter set.

The use of cross-validation in hyperparameter search ensures robust evaluation: the dataset is partitioned into k subsets (folds); in each iteration, the model is trained on $k-1$ subsets and validated on the remaining subset. This process repeats k times, each fold serving once as validation, and the performance is averaged across folds thus reducing variance in performance estimates and mitigating overfitting risks.

The effectiveness of hyperparameter tuning via grid search (or similar systematic search) for (Bi-)LSTM-based time-series forecasting has been demonstrated in recent literature. For instance, in the context of photovoltaic power forecasting, Sutarna et al. (2024) showed that Bi-LSTM with optimized hyperparameters (optimizer, learning rate, activation function) significantly improved forecasting accuracy [27]. In renewable-energy time series forecasting (wind power), a hybrid RNN-LSTM optimized with Grid Search + cross-validation also achieved high accuracy[28]. In a broader review of rainfall forecasting using AI techniques, hyperparameter tuning was reported to play an important role in minimizing forecasting errors in a substantial fraction of studies [29].

Therefore, employing Grid Search combined with cross-validation in this study provides a methodologically sound and empirically justified approach to finding optimal hyperparameter configurations for the Bi-LSTM rainfall forecasting model, improving both accuracy and generalization capability.

Table 1. Parameter Description

Parameter	Value
Neuron Hidden	5, 10, 15, 20
Batch	4, 8,16, 32
Epoch	50, 100, 150. 200
Dropout	0.1, 0.2

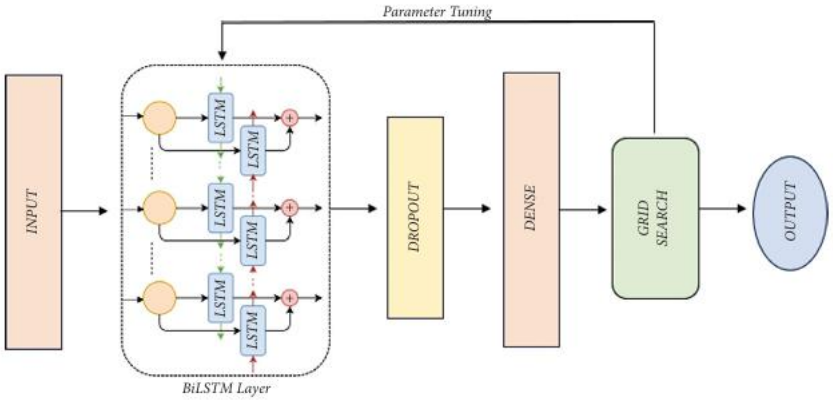

Figure 2. Proposed model Bi-LSTM-Grid Search

2.6. Model Evaluation

Two primary evaluation metrics were used: Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE).

MAPE

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (10)$$

RMSE

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (11)$$

Where y_i denotes the observed rainfall and $\hat{y}_i$ denotes the model's forecast at time i.

MAPE gives the average percentage error relative to the observed values, making it easy to interpret how large the error is in relative (percentage) terms, which is often meaningful in rainfall forecasting. RMSE gives a measure of the magnitude of the forecast error in the same units as the target variable (e.g. mm of rainfall), and penalizes larger errors more heavily, which is useful for capturing large deviations.

Several recent studies on rainfall forecasting use these metrics for model evaluation. For example, in a monthly rainfall forecasting study in Chuping, the authors used MAE, RMSE and MAPE to evaluate model performance in both temperature and rainfall forecasting [30].
Another recent research applying a hybrid ANN–Fuzzy model for annual rainfall prediction in Indonesia reports RMSE (alongside MAE and R^2) as one of primary evaluation metrics [31].

Therefore, in this study, using MAPE and RMSE as primary evaluation metrics is aligned with common practice in rainfall and hydrometeorological forecasting literature, and provides complementary perspectives on forecast accuracy: percentage-based relative error (MAPE) and absolute error magnitude (RMSE)

3. RESULTS AND DISCUSSION

This section presents the forecasting performance of the Bi-LSTM model developed under two calendar systems Gregorian and lunar. The analysis includes hyperparameter optimization, model evaluation using mean absolute percentage error (MAPE), and an examination of performance variation across forecasting horizons. All tables and figures are inserted following standard scientific reporting conventions.

3.1. Calendar Conversion

The data used in this study consist of daily rainfall observations recorded from 1 January 2000 to 31 August 2025 in Bandung City. These observations were obtained from the Meteorological, Climatological, and Geophysical Agency. The daily rainfall values in the Gregorian calendar were first aggregated into monthly totals to form the Gregorian-based monthly time series. The

conversion process produced 273 monthly observations over the twenty-five-year period, as presented in Table 2.

Subsequently, the same daily rainfall dataset was transformed into a lunar-based monthly series using the calendar conversion procedure described earlier. Due to differences in month length and calendar segmentation between the lunar and Gregorian systems, the lunar-based aggregation resulted in a total of 286 months, as shown in Table 3. This increase in the number of months reflects the shorter average duration of lunar months and the distinct structural properties of the lunar calendar. The conversion outcome confirms that the two temporal frameworks produce different time series representations, which are further analyzed in the subsequent forecasting evaluation.

Table 1. Monthly Rainfall Data (mm) Based on The Georgian Calendar

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2000	261.4	140.7	135.7	259	240.1	47.4	80.2	19.8	44.8	152.4	317.1	70.6
2001	219	248.9	208	244.3	82.9	87.5	187.2	52.3	107	409.7	526.4	75.5
2002	364.8	81.4	344.1	183.5	55	54.1	121.9	37.8	10.3	20.8	195.5	457.2
2003	69.1	265.6	359	136	111.7	37.4	40.5	74.7	76.3	320.2	185.6	203.2
2004	195.6	191.2	240.8	301.8	286.5	76.2	34.4	11.4	84.7	83.5	184.4	238.9
2005	168.2	416.7	307.7	213.5	190.6	201	76.3	64.2	145.3	114.9	225.8	204.7
2006	299.9	282.3	53.4	232.6	89.7	32.2	45	0	0.3	57.1	109.3	499.8
2007	127.5	405.7	105.4	472	88.6	164.1	11	0	44.1	98.4	301.7	359.7
2008	240.9	116.5	242.4	297.1	165.4	65.3	3.6	58.6	41.5	137	277.3	332.8
2009	208.5	200.5	365.7	165.6	183.8	101	24.2	0.5	24	234.5	318.2	271.1
2010	353.3	557.1	531	93	345	131.9	220.8	106.1	424.4	292.2	401.4	237.5
2011	63	76.7	89.4	381.5	193.4	117.6	77.2	3.1	102.8	103.6	321.4	259
2012	82.9	303.7	155.5	290.8	257.1	60.5	34.2	0	27	125	537	636.9
2013	216.9	249.6	304.8	285.8	170.9	231.5	159.1	74.3	171.7	35.8	64.1	325.6
2014	272.7	83.4	260.6	195.1	176.7	173	164.8	119.8	0.6	60.8	246.8	246.8
2015	167.3	179.7	248	231	208.1	56.4	0.3	50.4	43.2	34.5	419.4	301.8
2016	391.5	387.5	376.2	523	317.8	139.3	182.3	128.7	286.2	362.3	442.5	62.1
2017	68.3	196.3	396.5	210.8	222.3	68.4	7.9	45.7	90.8	345.3	442	129.9
2018	190.8	239.3	292	297.5	123.9	33.4	0.3	38.9	40.8	124.8	483.2	322.9
2019	231.4	269.3	223.3	298.9	243	36.9	13.4	0.2	55	84.2	270.9	315.5
2020	207.6	336.6	290.8	271.4	292.3	30.3	63.7	41.6	87.7	327.3	207.3	261.8
2021	148.6	153.9	309	177.2	238.9	92.4	33.2	91.8	73	218.4	454.3	198.5
2022	59.5	117.1	238.9	336.2	146.9	150.6	82.2	29.9	182.2	366.7	307.2	277.7
2023	68.3	111.1	199.8	275.5	268.5	89.8	24	29.7	18.2	71.2	239.3	366.4
2024	328.5	263	275	213.4	69.8	156.8	17	53.2	135.5	62.8	477.6	89.7
2025	225.5	131.4	274.1	233.1	319	146.5	112.9	211.9				

Table 3. Conversion of Monthly Rainfall (mm) Data Based on The Lunar Calendar.

Year	Muh	saf	rawal	rakhir	juwal	jukhir	rajab	syaban	ram	syawal	dzq	Dzh
1420									65.9	265.8	73.9	191.2
1421	273.8	166.3	52.4	75.2	13.3	51.3	117.5	342.7	58.4	119.5	313.2	170
1422	323.7	81.7	86.7	179.8	76.7	77	179.6	484.4	335.6	114.2	315.9	204.1
1423	238.2	173.4	54.1	63.5	83.4	37.8	28.1	9.5	192.2	454.7	91.1	259.5
1424	342.4	137.4	110.3	37.4	40.5	19.9	131.1	252.4	185.6	258.1	127.1	215.6
1425	251	247.2	291	167.9	39.4	0.2	62.4	117.2	64.6	155.7	221.9	246.2
1426	411.2	362.8	142.1	148.9	179.6	117	11.5	160.5	107.1	222	201.1	284.7

1427	297.5	51.2	228.6	68.6	59.5	41.1	3.9	0.3	6.1	279.3	235.6	162.3
1428	366.9	247.1	241.2	333.3	97.7	83	2.2	3.7	41.7	279.3	235.6	301
1429	199.9	139.7	258.2	276	151.4	44.9	10.8	51.4	41.5	136.8	266.8	343.5
1430	177.7	226.5	328.8	164.6	196.5	127	26.6	2.1	10	174.8	147.4	347.9
1431	410.3	503.4	373.5	399.7	166	266.2	173.6	149.8	185	319	312.9	476.5
1432	126.9	84.9	56.9	91.9	403.1	210.2	87	63.4	3.1	102.8	71.2	296
1433	273.5	102.3	234.8	166.1	343.3	265.7	80.2	8.3	26.2	8	96	148
1434	473.3	255.1	178.1	396.6	191.5	137.5	207.4	385.5	11.5	171.1	171.7	36.8
1435	64.1	325.6	272.7	84.4	259.6	195.1	176.7	173	164.7	68.1	52.4	4.8
1436	263.5	246.8	67.5	317.3	97.6	388	215.9	70.5	6	50.7	0	76.8
1437	147.5	371	358.8	371.3	353.6	400.9	433.1	302.3	156.6	159.2	79	335.7
1438	329	472.3	65.8	52	52	146.8	638.7	223	101.6	8.4	16.4	32
1439	222.6	489.8	218.4	234.6	132.1	344.4	173.5	254	79.4	33.4	8.7	50.4
1440	28.9	198.3	424	314.9	223.3	358	193	273.6	209	31.6	13.4	0.2
1441	55	14.6	303.4	291.7	192.6	369.5	192.8	381.5	307	39.5	53.3	58
1442	27.1	98.8	356.9	199.4	257.4	191	172.4	232.8	161.8	218.4	110.5	12.3
1443	90.3	82.4	352	318.6	192.1	64.5	112.5	252.7	329.7	139.2	150.6	82.2
1444	28.4	183.7	361.9	193.6	354.2	96.1	86.9	201	174.8	410.1	49.6	56.9
1445	42.2	0.2	28.1	163.8	281.8	450.4	138.2	299.5	230.9	188.8	149.4	73.1
1446	58.8	10.3	134	94.3	446.1	89.7	223.4	133.5	274.1	228.6	308.3	148
1447	120	176.6	41.9									

3.2. Hyperparameter Optimization Results

Hyperparameter tuning was performed using grid search for four forecasting horizons: 3, 6, 12, and 24 months. The optimal configurations for the Gregorian calendar are presented in Table 4, while the results for the lunar calendar are given in Table 5.

Table 4. Results of Bi LSTM Grid Search Tuning Parameters on the Gregorian Calendar

Forecasting length	Neuron	Batch	Epoch	Dropout	MSE
3	20	4	200	0.1	0.02889
6	20	4	200	0.1	0.03849
12	15	4	200	0.1	0.03054
18	20	4	200	0.1	0.02378
24	15	4	200	0.1	0.02343

The Gregorian-based model shows relatively stable optimal neuron counts between 15 and 20 across all forecasting horizons. The lowest MSE values are obtained at 18 and 24 months (0.02378 and 0.02343, respectively), indicating that the Gregorian calendar provides more reliable forecasting performance at longer prediction intervals. In contrast, short-term forecasting horizons (3 and 6 months) exhibit slightly higher error levels, suggesting that short-range rainfall fluctuations are less effectively captured using Gregorian temporal segmentation alone.

Table 5. Results of Bi LSTM Grid Search Tuning Parameters on the Lunar Calendar

Forecasting length	Neuron	Batch	Epoch	Dropout	MSE
3	20	4	200	0.1	0.02432
6	20	4	200	0.1	0.01876
12	15	4	200	0.1	0.03478
18	20	4	200	0.2	0.03528
24	15	4	200	0.1	0.02379

The lunar-based model achieves its lowest MSE at 6 months (0.01876), indicating strong sensitivity to semi-seasonal rainfall cycles. Short-term forecasting (3–6 months) benefits substantially from the lunar temporal structure, while performance decreases at longer horizons (12–18 months), where MSE values rise above 0.03. Interestingly, the 24-month forecasting horizon again shows improved performance (0.02379), comparable to the Gregorian model, suggesting that certain long-term rainfall oscillations are also detectable within lunar-aligned time segmentation.

3.3. Comparative Evaluation of Forecasting Accuracy

The forecasting accuracy of both models was evaluated using MAPE. The results are shown in Table 6.

Table 6. MAPE Value in the Gregorian and Lunar Calendars

Forecasting Length	MAPE Gregorian (%)	MAPE Lunar (%)
3	11.60	31.43
6	33.98	58.57
12	33.40	65.47
18	77.70	84.24
24	77.15	54.57

The comparative analysis based on MAPE values demonstrates that the Bi-LSTM model employing the Gregorian calendar provides superior forecasting performance when compared with the model structured under the lunar calendar. Across almost all forecasting horizons, the Gregorian model consistently produces lower error rates. For the short-term horizons of three and six months, the Gregorian calendar achieves MAPE values of 11.60 percent and 33.98 percent, whereas the lunar calendar shows substantially higher errors of 31.43 percent and 58.57 percent. A similar pattern is observed in the twelve-month horizon, with the Gregorian model reaching 33.40 percent and the lunar calendar exceeding 65 percent. These results indicate that Gregorian-based temporal segmentation is more effective in capturing rainfall variability for short to medium forecasting periods.

At the eighteen-month horizon, the performance gap becomes more evident. The lunar model records a MAPE value of 84.24 percent, while the Gregorian model achieves 77.70 percent. This outcome further confirms the greater stability of the Gregorian structure in modeling intermediate-range rainfall dynamics.

Although the lunar model attains a lower MAPE value than the Gregorian model at the twenty-four-month horizon, this single instance does not alter the overall conclusion. The lunar calendar performs less accurately in four out of five forecasting horizons and displays greater variability across the evaluation period. In contrast, the Gregorian model demonstrates more consistent and predictable behavior, avoiding the large fluctuations in error that characterize the lunar-based forecasts.

Considering all forecasting horizons collectively, the Gregorian calendar provides a more coherent and stable temporal foundation for rainfall prediction in Bandung City. The isolated improvement observed in the longest horizon of the lunar model does not outweigh its broader pattern of underperformance. Thus, the Gregorian calendar remains the more reliable choice for practical and operational rainfall forecasting applications.

4. CONCLUSION (11 PT)

This study examined the forecasting performance of Bi-LSTM models using Gregorian and lunar calendar structures for monthly rainfall prediction in Bandung City. The dual-calendar approach was developed to evaluate whether alternative temporal segmentation could enhance predictive accuracy. The results demonstrate clear compatibility between what was expected in the Introduction and the empirical findings presented in the Results and Discussion.

The analysis shows that the Gregorian calendar provides more accurate and stable forecasting performance across nearly all prediction horizons. The MAPE evaluation confirms

that the Gregorian-based model consistently outperforms the lunar model in four out of five forecasting lengths, particularly within short and medium horizons where rainfall variability is strongly governed by seasonal and monsoonal cycles. Although the lunar model achieves a lower error at the longest forecast horizon, this single result does not outweigh its overall higher error levels and instability in other periods. Therefore, the Gregorian calendar remains the more reliable temporal structure for rainfall forecasting in this region.

The findings of this study highlight the importance of selecting an appropriate temporal framework when modeling hydrometeorological variables. Future research may expand this approach by incorporating additional climate indicators, integrating attention-based deep learning architectures, or applying hybrid calendar models that combine both solar and lunar temporal cues. Further application of this methodology to flood early warning systems or climate adaptive planning tools may also provide valuable operational benefits for decision-making in Bandung City and similar tropical regions.

REFERENCES

[1] P. Feng *et al.*, "Machine learning-based integration of large-scale climate drivers can improve the forecast of seasonal rainfall probability in Australia," *Environ. Res. Lett.*, vol. 15, no. 8, p. 84051, 2020.

[2] V. Masson-Delmotte *et al.*, "Climate change 2021: the physical science basis," *Contrib. Work. Gr. I to sixth Assess. Rep. Intergov. panel Clim. Chang.*, vol. 2, no. 1, p. 2391, 2021.

[3] T. Kohyama and J. M. Wallace, "Rainfall variations induced by the lunar gravitational atmospheric tide and their implications for the relationship between tropical rainfall and humidity," *Geophys. Res. Lett.*, vol. 43, no. 2, pp. 918–923, 2016.

[4] S. Sen Roy, "Impact of lunar cycle on the precipitation in India," *Geophys. Res. Lett.*, vol. 33, no. 1, 2006.

[5] G. Darmawan, B. Handoko, D. Y. Faidah, and D. Islamiaty, "Improving the Forecasting Accuracy Based on the Lunar Calendar in Modeling Rainfall Levels Using the Bi-LSTM Method Through the Grid Search Approach," *Sci. World J.*, vol. 2023, no. 1, p. 1863346, 2023.

[6] H. Kang, S. Yang, J. Huang, and J. Oh, "Time series prediction of wastewater flow rate by bidirectional LSTM deep learning," *Int. J. Control. Autom. Syst.*, vol. 18, no. 12, pp. 3023–3030, 2020.

[7] I. Priyadarshini and C. Cotton, "A novel LSTM–CNN–grid search-based deep neural network for sentiment analysis," *J. Supercomput.*, vol. 77, no. 12, pp. 13911–13932, 2021.

[8] D. Endalie, G. Haile, and W. Taye, "Deep learning model for daily rainfall prediction: case study of Jimma, Ethiopia," *Water Supply*, vol. 22, no. 3, pp. 3448–3461, 2022.

[9] Y. Zhang and A. Ye, "Machine learning for precipitation forecasts postprocessing: Multimodel comparison and experimental investigation," *J. Hydrometeorol.*, vol. 22, no. 11, pp. 3065–3085, 2021.

[10] A. Indryani, U. Khaira, and M. F. Putri, "Rainfall Prediction Using Long Short-Term Memory Method (Case Study: Jambi City)," *Publ. Elektron. Pengemb. Apl. Digit. Untuk Negeri*, vol. 6, no. 1, pp. 57–70, 2025, doi: https://doi.org/10.23960/pepadun.v6i1.256.

[11] R. Firdaus and I. Paputungan, "Prediksi Curah Hujan di Kota Bandung Menggunakan Metode Long Short Term Memory," *J. Penelit. Inov.*, vol. 2, pp. 453–460, Nov. 2022, doi: 10.54082/jupin.99.

[12] A. Wangwongchai, M. Waqas, P. Dechpichai, P. T. Hlaing, S. Ahmad, and U. W. Humphries, "Imputation of missing daily rainfall data; A comparison between artificial intelligence and statistical techniques," *MethodsX*, vol. 11, p. 102459, 2023.

[13] U. Rasool, X. Yin, Z. Xu, M. A. Rasool, M. Hussain, and F. Iftikhar, "Influence of lunar phases and meteorological factors on rainfall in Karachi City, Pakistan," *J. Hydrol.*, vol. 629, p. 130628, 2024.

[14] J. I. Avila-Carrazco *et al.*, "Correlation between Lunar Phases and Rainfall Patterns in Mexico," *Atmosphere (Basel).*, vol. 15, no. 7, p. 746, 2024.

[15] A. Ligthart, C. Catal, and B. Tekinerdogan, "Systematic reviews in sentiment analysis: a tertiary study," *Artif. Intell. Rev.*, vol. 54, no. 7, pp. 4997–5053, 2021.

[16] F. Rustam, M. Khalid, W. Aslam, V. Rupapara, A. Mehmood, and G. S. Choi, "A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis," *PLoS One*, vol. 16, no. 2, p. e0245909, 2021.

[17] S. Mekruksavanich and A. Jitpattanakul, "Lstm networks using smartphone data for sensor-based human activity recognition in smart homes," *Sensors*, vol. 21, no. 5, p. 1636, 2021.

[18] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment analysis of comment texts based on

BiLSTM," *Ieee Access*, vol. 7, pp. 51522–51532, 2019.
[19] S. Liang, D. Wang, J. Wu, R. Wang, and R. Wang, "Method of Bidirectional LSTM Modelling for the Atmospheric Temperature.," *Intell. Autom. Soft Comput.*, vol. 30, no. 2, 2021.
[20] C. Cai, Y. Tao, T. Zhu, and Z. Deng, "Short-term load forecasting based on deep learning bidirectional lstm neural network," *Appl. Sci.*, vol. 11, no. 17, p. 8129, 2021.
[21] Z. Cui, R. Ke, Z. Pu, and Y. Wang, "Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting network-wide traffic state with missing values," *Transp. Res. Part C Emerg. Technol.*, vol. 118, p. 102674, 2020.
[22] Z. Hameed and B. Garcia-Zapirain, "Sentiment classification using a single-layered BiLSTM model," *Ieee Access*, vol. 8, pp. 73992–74001, 2020.
[23] J. Kim and N. Moon, "BiLSTM model based on multivariate time series data in multiple field for forecasting trading area," *J. Ambient Intell. Humaniz. Comput.*, pp. 1–10, 2019.
[24] P. Varalakshmi, N. Vasumathi, and R. Venkatesan, "Tropical Cyclone prediction based on multi-model fusion across Indian coastal region," *Prog. Oceanogr.*, vol. 193, p. 102557, 2021.
[25] F. Granata, F. Di Nunno, M. Najafzadeh, and I. Demir, "A stacked machine learning algorithm for multi-step ahead prediction of soil moisture," *Hydrology*, vol. 10, no. 1, p. 1, 2022.
[26] R. K. Vogeti, B. R. Mishra, and K. S. Raju, "Machine learning algorithms for streamflow forecasting of Lower Godavari Basin," *H2Open J.*, vol. 5, no. 4, pp. 670–685, 2022.
[27] N. Sutarna, C. Tjahyadi, P. Oktivasari, M. Dwiyaniti, and T. Tohazen, "Hyperparameter Tuning Impact on Deep Learning Bi-LSTM for Photovoltaic Power Forecasting," *J Robot Control*, vol. 4, no. 3, pp. 677–693, 2024, doi: https://doi.org/10.18196/jrc.v5i3.21120.
[28] A. G. AbdElkader, H. ZainEldin, and M. M. Saafan, "Optimizing wind power forecasting with RNN-LSTM models through grid search cross-validation," *Sustain. Comput. Informatics Syst.*, vol. 45, p. 101054, 2025, doi: https://doi.org/10.1016/j.suscom.2024.101054.
[29] M. A. Saleh, H. M. Rasel, and B. Ray, "A comprehensive review towards resilient rainfall forecasting models using artificial intelligence techniques," *Green Technol. Sustain.*, vol. 2, no. 3, p. 100104, 2024, doi: https://doi.org/10.1016/j.grets.2024.100104.
[30] L. S. Ee, M. Misiran, H. Sapiri, and Z. M. Yusof, "Forecasting the Monthly Temperature and Rainfall in Chuping Perlis," *Educ. J. Sci. Math. Technol.*, vol. 11, no. 2, pp. 20–33, 2024, doi: 10.37134/ejsmt.vol11.2.3.2024.
[31] S. Asiah, W. Riana, D. C. Purba, M. I. Azharsum, and P. V. A. Elyakim, "Annual Rainfall Prediction in Indonesia Using A Hybrid Artificial Neural Network and Fuzzy Algorithm Model," *JOMLAI J. Mach. Learn. Artif. Intell.*, vol. 4, no. 2, pp. 99–105, 2025, doi: https://doi.org/10.55123/jomlai.v4i2.5964.