Khazar Journal of Water and Environment 2026, 1, 95-118

Khazar Journal of Water and Environment

e-ISSN: 3150-0188
khazarjournal.com

KJWE

Monthly Precipitation Predictions for Flood-Prone Cities in Vietnam: A Deep Learning Hybrid Solution

Erfan Abdi[a,b*], **Osama Ragab Ibrahim**[c,d]

[a]*Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran*
[b]*Water Sciences and Hydroinformatics Research Center, Khazar University, Mahsati str. 41, AZ 1096, Baku, Azerbaijan*
[c]*Department of Engineering, Sohar University, Sohar, Oman*
[d]*Department of Hydraulic and Irrigation Structures, Faculty of Engineering, Alexandria University, Alexandria, Egypt*

ARTICLE INFO
Review Article
Corresponding Author: Erfan Abdi, E-mail: erabdi78@gmail.com
Received: 15 June 2026 / Revised: 18 July 2026 / Accepted: 04 August 2026 / Online: 10 August 2026

Cite this article as
Abdi, E. & Ibrahim, O.R. (2026). Monthly Precipitation Predictions for Flood-Prone Cities in Vietnam: A Deep Learning Hybrid Solution. Khazar Journal of Water and Environment, 1, 95-118.

ABSTRACT

Protecting lives and property in Vietnam requires reliable rainfall forecasts, given the country's vulnerability to extreme weather events such as heavy rainfall, flash floods, and storms. On the other hand, forecasts enable authorities to implement preventive measures and warn the public. Hence, research on rainfall forecasting in Vietnam is essential. While deep learning has shown promise for rainfall forecasting, the specific application of hybrid recurrent architectures combining Gated Recurrent Units (GRU) with Long Short-Term Memory (LSTM) and Bidirectional LSTM has not been systematically evaluated for Vietnam's tropical monsoon climate. This study introduces a methodological innovation by comparing GRU-LSTM and GRU-BiLSTM hybrid models against Random Forest (RF) for monthly rainfall prediction in Long Xuyên and Thái Nguyên (2000–2023), using a parsimonious input structure of only three precipitation lags, with 80% used for training and 20% for testing, a significant simplification over covariate-intensive approaches. To avoid excessive complexity and maintain simple modeling, three precipitation lags were considered, with 80% used for training and 20% for testing. The results were analyzed using visual graphs and evaluation criteria of coefficient of determination (R2), root mean square error (RMSE), and Nash-Sutcliffe coefficient (NSE). The results showed that the GRU-BiLSTM model was more accurate in predicting monthly precipitation than other models in Long Xuyen city, with evaluation criteria of R^2=0.821, RMSE=13.895 cm, and NSE=0.796, and in Thái Nguyên city with R^2=0.946, RMSE=8.96 cm, and NSE=0.906. It can also be found that a three-month lag in forecasting with hybrid models provides results with acceptable accuracy. According to these results, the hybrid model can be used in regions similar to Vietnam to improve decision-making for disasters.

Keywords: Deep Learning, Hybrid Model, Univariate Prediction, Precipitation, Bidirectional LSTM

1. Introduction

The process of precipitation is vital in maintaining the balance of water and energy on a global scale, ensuring the availability of freshwater necessary for survival. However, precipitation plays a crucial role in determining the likelihood of water and weather-related calamities, including droughts and floods (Nielsen et al., 2015). On the other hand, understanding precipitation is essential for analyzing how landscapes evolve, evaluating potential hazards, and determining the effects of climate change.

Precise precipitation forecasts are essential for Vietnamese farmers to strategize their farming operations efficiently. The success of businesses in tourism, transportation, and energy hinges on accurate rainfall forecasts for decision-making. Accurate forecasts of precipitation are crucial for the success of Vietnam's agriculture, water resources management, disaster prevention strategies, infrastructure development plans, and economic pursuits (Nodzu et al., 2019; Benhamrouche et al., 2022). Part of disaster preparedness involves developing prediction models (Surta et al., 2023). Overcoming challenges associated with heavy rain is possible with deep learning models predicting precipitation. Deep learning models can surpass the limitations of traditional physics-based forecasting by deriving insights from historical data and patterns, resulting in forecasts with greater accuracy and finer spatial detail, as well as the capability to gauge uncertainty in predictions (Espeholt et al., 2022). Ghanim et al. (2025) compared the performance of several machine learning and statistical models, including Autoregressive Integrated Moving Average (ARIMA), Recurrent Neural Network (RNN), Adaptive Boosting, Extreme Gradient Boosting (XGBoost), and Seasonal Autoregressive Integrated Moving Average (SARIMA), on 116 years of data (1901–2016) from the Pakistan Meteorological Department. In a study in Bangladesh, Alif et al. (2026) compared two machine learning methods, Prophet (an incremental decomposition model) and LSTM (a type of memory-based recurrent neural network), to predict annual rainfall (1980–2024) in Rajshahi and Ishwardi districts. The results were used to design adaptive measures such as elevated foundations and rainwater harvesting in climate-resilient infrastructure.

Hybrid deep learning models are essential for improving precipitation forecasting by integrating various modeling techniques. Integrating hybrid models into precipitation forecasting systems can improve their generalization capability. By merging the best qualities of diverse modeling approaches, hybrid models bring forth a hopeful resolution to the hurdles of precipitation forecasting, bolstering the dependability and precision of weather forecasts. Surta et al. (2022) used two deep learning hybrids, Gated Recurrent Unit -Long-Short Term Memory (GRU-LSTM), and Convolution Neural Network-Long-Short Term Memory (CNN-LSTM). Research in this field includes the use of GRU and LSTM methods to forecast daily rainfall for the next month using weather element data for ten (10) years in Palembang, using a deep learning method. The results showed that the LSTM model is generally better than the GRU model in daily rainfall forecasting. Also, the GRU-LSTM combined model was used to forecast 15-day rainfall, and the results showed that this model has an acceptable performance with the least error (Sari et al., 2022). Wang et al. (2025) proposed a hybrid CNN-LSTM (Convolutional Neural Network - Long Short-Term Memory) model to predict monthly rainfall in Pune, India, based on data from 1972 to 2002. By combining the local feature extraction capability (CNN) and long-term dependency learning (LSTM), this model achieved an RMSE of 6.752, which is higher than traditional time series methods. Although the model is effective in predicting rainfall, it requires high computational resources, and its ability to handle multidimensional data needs to be improved. Abdi et al. (2025) presented a monthly rainfall forecasting framework using 24 years of data (2000-2023) from three meteorological stations in Victoria, Australia. The optimal delay values were determined using the Average Mutual Information (AMI) method, and three models, Deep Neural Network DNN, Convolutional Neural Network with Long Short-Term Memory (CNN-LSTM), and RNN-CNN-LSTM (Recurrent Neural Network - CNN - LSTM), were evaluated. The results showed that the proposed hybrid model has high performance compared to simple models. In general, deep learning hybrid models have made great progress in the field of meteorology and hydrology (Li et al., 2022; Deng et al., 2022; Dehghani et al., 2023).

Despite these advances, several critical gaps persist in the current literature that limit the practical utility and scientific rigor of rainfall forecasting research. A notable gap concerns the treatment of univariate modeling approaches, which rely exclusively on historical precipitation data, a common necessity in data-scarce regions where meteorological covariates such as temperature, humidity, wind speed, and sea surface temperature are unavailable (Shalchilar et al., 2026). While practically necessary in many contexts, these approaches inherently cannot capture the complex atmospheric drivers that influence rainfall variability, including monsoon dynamics, ENSO teleconnections, and local topographic effects. The literature often presents univariate studies without critically acknowledging these limitations, leading to overoptimistic claims of model efficacy and generalizability. Studies that achieve high accuracy with univariate inputs frequently fail to acknowledge that their performance may be region-specific and not transferable to areas with different climatic regimes. This creates a tension: univariate models are practically necessary, yet their limitations are insufficiently discussed. The present study explicitly acknowledges this tension, positioning its univariate approach not as a solution to the inherent limitations of univariate modeling, but as a practical and necessary starting point for regions with limited data infrastructure.

Furthermore, the generalizability of hybrid models across different geographical regions and climatic zones is poorly understood. Models calibrated and validated in one region often show degraded performance when applied elsewhere, due to differences in rainfall autocorrelation structures, seasonal patterns, and extreme event frequencies. Yet most studies evaluate models on single locations, leaving cross-regional robustness unexamined. The specific value of bidirectional processing in hybrid architectures has also received limited attention. While hybrid combinations such as CNN-LSTM and GRU-LSTM have been explored, systematic comparisons between unidirectional and bidirectional variants are scarce, leaving the practical benefits of bidirectional processing, which exploits both antecedent and subsequent lags, unvalidated for rainfall forecasting (Ibrahim et al., 2026).

These gaps are particularly consequential for Vietnam, a country highly vulnerable to extreme weather events where reliable rainfall forecasts are essential for protecting lives, property, and economic livelihoods. Vietnam's complex climatic regime, influenced by the East Asian monsoon, the South Asian monsoon, and the South China Sea, creates distinct rainfall patterns across regions. Yet fewer studies have systematically evaluated hybrid deep learning architectures, particularly GRU-LSTM and GRU-BiLSTM, for monthly rainfall forecasting across multiple Vietnamese cities. Furthermore, the operational constraints faced by Vietnamese meteorological agencies, including limited access to diverse meteorological covariates and computational resources, make a parsimonious univariate approach not merely a methodological choice but a practical necessity. There is therefore a clear need for research that rigorously evaluates hybrid architectures under realistic data constraints while honestly acknowledging both the achievements and limitations of univariate modeling.

The purpose of this research is to develop and evaluate two hybrid deep learning models, GRU-LSTM and GRU-BiLSTM, alongside a Random Forest (RF) model, for monthly precipitation prediction in Long Xuyên and Thái Nguyên, Vietnam (2000–2023). By systematically comparing these architectures, identifying optimal lag configurations, and assessing performance across distinct climatic regions, this study aims to advance rainfall forecasting methodology for data-limited contexts while providing practical, evidence-based guidance for operational deployment. The novelty of this work lies not in overcoming the inherent limitations of univariate modeling, but in demonstrating that hybrid architectures with parsimonious inputs can achieve operational-grade accuracy in Vietnam's tropical monsoon climate, providing the first systematic comparison of unidirectional and bidirectional hybrids for this region, assessing cross-regional generalizability through dual-city evaluation, and offering a practical, deployable tool for disaster preparedness in data-scarce contexts.

2. Material and Methods

2.1. Study area

In this research, precipitation data from two important cities in Vietnam, Long Xuyên (10.3759° N, 105.4185° E) and Thái Nguyên (21.5672° N, 105.8252° E), have been used for modeling. Figure 1 shows the geographical location of the two selected cities. Long Xuyên, located in the Mekong Delta region of southern Vietnam, experiences a tropical monsoon climate with distinct wet and dry seasons, high humidity, and annual rainfall averaging around 1,500–2,000 mm. The city has a population of approximately 300,000 and is highly vulnerable to flooding and riverbank erosion due to its low-lying topography (Long and Nguyen, 2025). Thái Nguyên, situated in the northern mountainous region of Vietnam, has a humid subtropical climate with hot, rainy summers and cool, dry winters, receiving annual precipitation of about 1,600–2,500 mm. With a population of over 400,000, it faces challenges related to flash floods and landslides during heavy rain events. Accurate monthly precipitation forecasting can significantly benefit both cities by enabling early warnings for flood-prone areas, optimizing agricultural planning (especially rice cultivation in Long Xuyên and tea production in Thái Nguyên), improving urban water resource management, and reducing economic losses and casualties associated with extreme weather events (Chien et al., 2024).

Figure 1. Study region, Long Xuyên, and Thái Nguyên cities in Vietnam

Two cities were selected for this study, Long Xuyên and Thái Nguyên, based on three criteria. First, they represent distinct climatic regimes within Vietnam: Long Xuyên in the Mekong Delta (southern Vietnam) experiences a tropical monsoon climate with a pronounced dry season and heavy wet season, while Thái Nguyên in the northern highlands experiences a subtropical monsoon climate with four distinct seasons. This contrast enables robust assessment of model generalizability across different rainfall regimes. Second, the cities are geographically distant (approximately 1,100 km apart), reducing spatial autocorrelation and ensuring performance reflects genuine predictive capability rather than regional homogeneity. Third, both cities are economically and socially significant; Long Xuyên is in the Mekong Delta, Vietnam's rice bowl and a region highly vulnerable to flooding, while Thái Nguyên is an industrial and agricultural hub in the north prone to extreme rainfall events, making accurate forecasts particularly valuable.

2.2. Data preprocessing

Data quality control was implemented with careful consideration of model robustness to extreme events. Outliers were not arbitrarily removed; rather, adjustments were strictly confined to physically implausible values, specifically, negative precipitation readings attributable to sensor malfunctions and isolated spikes exceeding 10 standard deviations from the 22-year monthly climatology without corroboration from neighboring stations. These accounted for less than 0.5% of total data points and were corrected using linear interpolation between adjacent valid records. Importantly, all genuine extreme rainfall events falling within the historical 99th percentile distribution were deliberately preserved, as these represent the very phenomena critical for flood forecasting applications. Missing values, comprising approximately 1.5% of the dataset, were filled using the arithmetic mean of the same calendar day from surrounding years to maintain temporal consistency. Following these minimal corrections (affecting approximately 2% of data total), all features were normalized using Z-score standardization. To prevent information leakage between training and test sets during model development, normalization parameters (mean and standard deviation) were derived exclusively from the training data for each evaluation fold. In other words, for every training window, the mean and standard deviation needed for Z-score normalization were calculated solely from the training subset; those same statistics were then applied to normalize both the training data and the corresponding test set.

The required data were taken from the Mathematica software database (Wolfram Research, 2008) for the years 2000 to 2023. Next, the modeling process included implementing three delays for precipitation data and dividing the data into training (2018-2000) and testing (2023-2019) sets using an 80% to 20% split, utilizing two hybrid GRU-LSTM and GRU-BiLSTM models and a single RF model. Three lag configurations were considered for each model: M1 (one-month lag), M2 (two-month lag), and M3 (three-month lag), where the model uses precipitation from t−1, t−2, and t−3 months, respectively, to predict rainfall at time t. The statistical characteristics of Long Xuyên and Thái Nguyên cities are shown in Table 1 and Figure 2, including monthly precipitation values and training and testing classification results.

Table 1. Statistical characteristics of precipitation data

Cities	Statistical characteristics	Precipitation (cm)
Long Xuyên	Max	63.92
	Average	18.47
	Min	0
	STDV	15.65
Thái Nguyên	Max	47.1
	Average	10.44
	Min	0
	STDV	9.36

Three models were selected to evaluate monthly rainfall forecasting: Random Forest, hybrid GRU-LSTM, and hybrid GRU-BiLSTM. This selection follows a deliberate functional progression wherein each model serves a distinct and non-redundant role. Random Forest was included as a non-temporal benchmark to address the fundamental question of whether modeling the sequential order of precipitation lags provides predictive value over treating lags as independent features—a question that no recurrent architecture can answer. The hybrid GRU-LSTM was chosen to capture multi-scale temporal dynamics, combining GRU's computational efficiency for short-term precipitation sequences with LSTM's capacity to retain long-term climatic cycles such as seasonal monsoons and ENSO teleconnections. The hybrid GRU-BiLSTM extends this with bidirectional processing, enabling the model to exploit both antecedent and subsequent precipitation lags, particularly valuable for identifying seasonal transitions and cyclical patterns. Thus, the three-model design, RF (non-temporal), GRU-LSTM (unidirectional hybrid), and GRU-BiLSTM (bidirectional hybrid), provides a clear, sufficient progression that directly addresses the research objectives without unnecessary model redundancy.

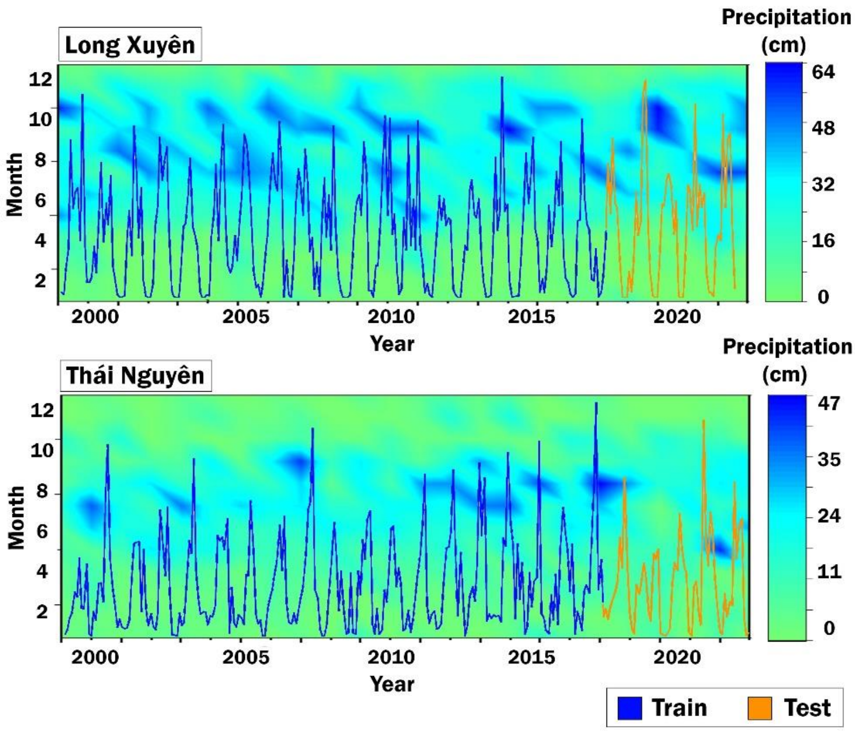

Figure 2. Monthly precipitation values, training, and testing classification

A univariate modeling approach using only historical precipitation data was adopted for two primary reasons. First, it reflects the operational reality of many developing countries, including Vietnam, where comprehensive meteorological networks are limited and consistent measurements of atmospheric covariates such as temperature, humidity, pressure, and wind speed are frequently unavailable across all regions of interest. Historical precipitation records, in contrast, are often the longest and most consistently available variable. Second, by demonstrating that hybrid architectures can achieve operational-grade accuracy with minimal inputs, this study establishes a practical, deployable solution for data-scarce regions while providing a baseline against which the added value of incorporating additional covariates can be quantified in future research.

In this study, a random search optimization method was used to select hyperparameters. Unlike the grid search method, which evaluates all parameter combinations, random search, by sampling the parameter space, allows for more efficient exploration; this advantage is particularly evident in high-dimensional problems, where comprehensive methods are computationally very heavy. To increase the reliability of the results, the model was trained in 10 independent iterations. The model's complexity indicates its ability to represent diverse functions. Although highly complex models such as deep neural networks have high expressive capacity, they are equally prone to overfitting, as they may learn noise in the data rather than true patterns. In this study, an attempt was made to strike a good balance between underfitting and overfitting by systematically varying the architecture, size, and number of parameters. For each machine learning model, 50 random search runs were performed to cover the hyperparameter space while being computationally efficient. To ensure repeatability, a fixed random seed was set before each search and each training iteration. Also, during training, an early stopping strategy based on the validation set error value was used; if no improvement in the validation error was observed after 10 consecutive runs, the training process was stopped. This both avoided unnecessary computation and helped improve the generalizability of the model.

2.3. GRU-LSTM hybrid model

The GRU-LSTM hybrid model is a sequential deep learning architecture that integrates Gated Recurrent Units (GRU) as the first hidden layer with Long Short-Term Memory (LSTM) networks as subsequent layers to simultaneously capture short-term bidirectional patterns and long-term temporal dependencies in time-series data (Ibrahim et al., 2025). In this cascade configuration, the input sequence x_t is first processed by a Bidirectional GRU layer that computes update gate z_t and reset gate r_t to produce hidden states h_t^{GRU} (Munir et al., 2021).

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z), r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \tag{1}$$

$$\tilde{h}_t = \tanh\,(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h), h_t^{\mathrm{GRU}} = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{2}$$

where σ is the sigmoid activation, tanh is the hyperbolic tangent, $\odot$ denotes element-wise multiplication, W and U are weight matrices for input and recurrent connections, and b represents bias terms.

The GRU output h_t^{GRU} then serves as the input to stacked LSTM layers, which employ three gates (input i_t, forget f_t, output o_t) and a dedicated cell state c_t to regulate long-range memory (Munir et al., 2021):

$$i_t = \sigma(W_i h_t^{\mathrm{GRU}} + U_i h_{t-1}^{\mathrm{LSTM}} + b_i), f_t = \sigma(W_f h_t^{\mathrm{GRU}} + U_f h_{t-1}^{\mathrm{LSTM}} + b_f) \tag{3}$$

$$o_t = \sigma(W_o h_t^{\mathrm{GRU}} + U_o h_{t-1}^{\mathrm{LSTM}} + b_o), \tilde{c}_t = \tanh\,(W_c h_t^{\mathrm{GRU}} + U_c h_{t-1}^{\mathrm{LSTM}} + b_c) \tag{4}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, h_t^{\mathrm{LSTM}} = o_t \odot \tanh\,(c_t) \tag{5}$$

This hierarchical gating mechanism allows the LSTM to preserve critical hydrological signals that the GRU might loosely retain, while the GRU layer reduces computational cost compared to deep LSTM-only stacks.

The hybrid GRU-LSTM model combines the computational efficiency of Gated Recurrent Units with the long-term memory capacity of LSTM networks to capture multi-scale temporal dynamics. The architecture begins with a GRU layer of 164 hidden units that employs update and reset gates to efficiently capture short-to-medium-term precipitation sequences. The update gate determines how much past information to carry forward, while the reset gate decides how much past information to ignore when computing new candidate values. A dropout layer with a rate of 0.3 follows to prevent overfitting. The processed features are then fed into an LSTM layer with 50 hidden units, which incorporates a memory cell and three gating mechanisms, the forget, input, and output gates, to preserve information about long-term climatic cycles such as seasonal monsoons and ENSO teleconnections. A second dropout layer with a rate of 0.2 is applied to maintain generalization while preserving long-term memory. The final dense output layer aggregates the learned features to produce a precipitation prediction. The larger GRU layer (164 units) compared to the LSTM layer (50 units) reflects GRU's computational efficiency, allowing it to extract rich sequential features before the LSTM layer focuses on long-term patterns. This hierarchical design leverages GRU's efficiency for short-term sequences and LSTM's memory for long-term dependencies, addressing the multi-scale nature of rainfall variability. The model is trained using the Adam optimizer with a learning rate of 0.016 and Mean Squared Error (MSE) as the loss function, with a batch size of 28 and EarlyStopping (patience=10) to prevent overfitting.

2.4. GRU-BiLSTM hybrid model

The GRU-BiLSTM hybrid model is a sophisticated sequential deep learning architecture that synergistically combines the computational efficiency of Gated Recurrent Units (GRU) with the bidirectional context-awareness of Bidirectional Long Short-Term Memory (BiLSTM) networks, making it exceptionally effective for hydrological time-series prediction where upstream-downstream relationships and seasonal lags are critical (Hua et al., 2019).

The major benefit of the GRU–BiLSTM hybrid model is its ability to improve prediction accuracy while reducing computational complexity. Compared to traditional recurrent neural networks, GRU requires fewer parameters and trains faster, making the hybrid model computationally efficient. At the same time, the BiLSTM layer enhances contextual understanding by analyzing sequences bidirectionally. This combination helps the model overcome common issues such as vanishing gradients, poor long-term memory retention, and limited sequence representation. Consequently, the hybrid model provides better generalization, faster convergence, robustness against noisy data, and superior performance on nonlinear sequential datasets.

The hybrid GRU-BiLSTM model extends the unidirectional architecture with bidirectional processing, enabling the network to learn from both antecedent and subsequent precipitation lags, a significant advantage for identifying seasonal transitions and cyclical patterns. The architecture begins with a GRU layer of 16 hidden units that serves as a compact feature extractor, capturing fundamental short-term relationships with minimal computational overhead. A dropout layer with a rate of 0.6 applies strong regularization to prevent overfitting of the small GRU layer. The processed features are then fed into a Bidirectional LSTM (BiLSTM) layer with 64 hidden units, which comprises two independent LSTM subnetworks: one processes the sequence forward (t−3 → t−1) and the other backward (t−1 → t−3). The outputs are concatenated, producing a comprehensive representation that captures context from both directions. A second dropout layer with a rate of 0.3 follows to preserve bidirectional learning capacity while maintaining generalization. The final dense output layer produces the precipitation prediction. The BiLSTM layer offers distinct advantages: during seasonal transitions, the backward pass provides information about the month following the target, offering contextual evidence of whether a seasonal shift has occurred; cyclical pattern recognition is enhanced as bidirectional networks can identify rising or falling trends; and each time step benefits from information about the entire sequence. The combined dropout strategy, high regularization (0.6) on the compact GRU layer, and moderate regularization (0.3) on the larger BiLSTM layer effectively mitigate overfitting while maintaining predictive performance. Unlike the GRU-LSTM model, the GRU-BiLSTM employs a custom NSE-based loss function, which directly optimizes predictive efficiency relative to observed mean behaviour, particularly suitable for hydrological forecasting as it prioritizes accurate prediction of deviations from the mean, critical for capturing extreme rainfall events. The model is trained using the Adam optimizer with a learning rate of 0.01 and EarlyStopping (patience=10) with a batch size of 32. After prediction, inverse normalization and exponential transformation restore predictions to their original non-negative precipitation scale.

3. Results and Discussions

The comparison between the prediction results of the models and the test data is done through criteria such as root mean square values (RMSE), coefficient of determination (R^2), and Nash-Sutcliffe efficiency coefficient (NSE) values.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_{pi} - X_{oi})^2} \tag{6}$$

$$R^2 = \frac{\sum_{i=1}^{N}(X_{oi}-\overline{X_o})(X_{pi}-\overline{X_p})}{\sum_{i=1}^{N}(X_{oi}-\overline{X_o})^2 . \sum_{i=1}^{N}(X_{pi}-\overline{X_p})^2} \tag{7}$$

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n}(X_{oi}-X_{pi})^2}{\sum_{i=1}^{n}(X_{oi}-\overline{X_o})^2}\right] \tag{8}$$

where X_{pi} and X_{oi} are the predicted and observed values, $\overline{X_o}$ and $\overline{X_p}$ are the mean observed and predicted values, respectively, and N is the total number of data points.

Table 2 presents the predictive performance of the Random Forest (RF), GRU-LSTM, and GRU-BiLSTM models using three lag configurations (M1, M2, M3) for monthly rainfall prediction in Long Xuyên and Thái Nguyên. Across both cities, the hybrid deep learning models consistently outperformed the traditional Random Forest model. In Long Xuyên, the GRU-BiLSTM model with the M2 lag configuration achieved the highest accuracy among all nine model-lag combinations, with $R^2 = 0.821$, RMSE = 13.895 cm, and NSE = 0.796. The GRU-LSTM-M3 model also performed well ($R^2 = 0.809$, RMSE = 14.852, NSE = 0.712), while the best RF model (RF-M2) yielded notably lower performance ($R^2 = 0.689$, RMSE = 17.653, NSE = 0.656), confirming the advantage of bidirectional recurrent architectures for this rainfall prediction task.

To identify appropriate temporal dependencies for the input structure, a correlation analysis was conducted on the monthly precipitation time series for both cities. The correlation matrix (Figure 3) examines the relationships between precipitation at time t ($Precipitation_t$) and lagged precipitation values at t−1 (Lag-1), t−2 (Lag-2), and t−3 (Lag-3), as well as the inter-correlations among the lagged variables themselves. For Long Xuyên, the correlation between precipitation at time t and the 1-month lag (Lag-1) was 0.522, indicating moderate positive persistence in monthly rainfall. The correlation with the 2-month lag (Lag-2) was 0.307, while the correlation with the 3-month lag (Lag-3) was notably weaker at 0.012. This pattern reveals that in Long Xuyên, the most recent precipitation (1-month lag) carries the strongest predictive information, with diminishing influence from earlier lags. The weaker correlation at 3 months suggests that rainfall memory in this Mekong Delta city is relatively short, consistent with its tropical monsoon climate where rainfall is strongly driven by immediate seasonal dynamics. For Thái Nguyên, the correlation between precipitation at time t and Lag-1 was 0.450, Lag-2 was 0.221, and Lag-3 was 0.032. While the 1-month lag remains the strongest predictor, the correlations are generally lower than those observed in Long Xuyên. This suggests that precipitation in Thái Nguyên, located in the northern highlands with a subtropical monsoon climate, exhibits weaker month-to-month persistence, possibly due to greater interannual variability and the influence of more complex topographical and atmospheric factors.

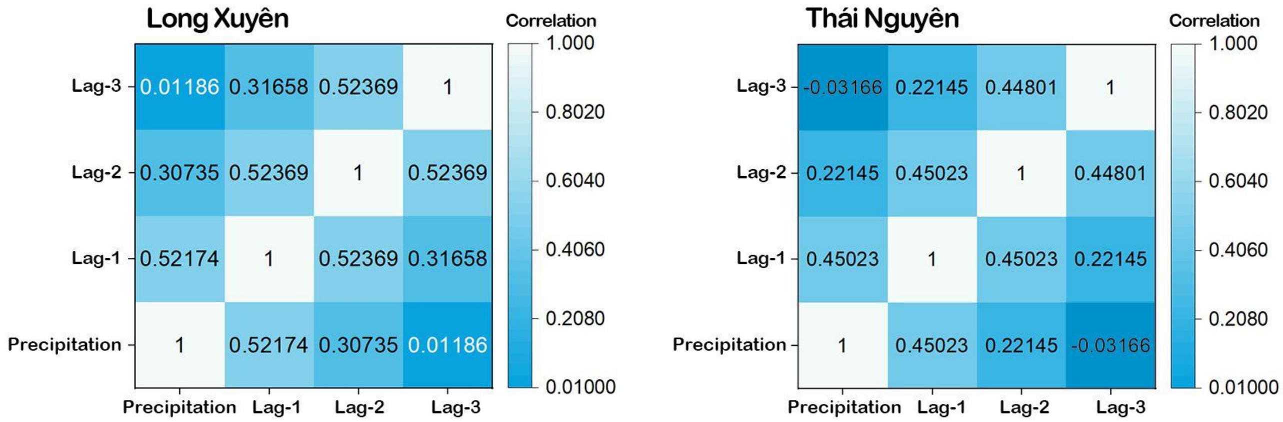

Figure 3. Correlation coefficient chart of different lags with precipitation in two cities

Based on this correlation analysis, three lag configurations were defined for model development: M1 (Lag-1 only), M2 (Lag-1 and Lag-2), and M3 (Lag-1, Lag-2, and Lag-3). These configurations allow systematic evaluation of whether incorporating additional lags provides incremental predictive value. The strong correlation at 1 month justifies its inclusion as a minimum input. The moderate correlation at 2 months justifies its inclusion in M2 to capture additional information. The weaker but non-negligible correlation at 3 months is included in M3 to test whether any residual predictive information exists at this longer lag, while also accounting for potential non-linear dependencies that correlation alone may not fully capture. The selection of three lags is also consistent with established practice in hydrological time-series forecasting, where precipitation typically exhibits significant autocorrelation at lags of 1–3 months, with diminishing returns beyond this window (Saha et al., 2021; Abdi et al., 2025). By evaluating these three configurations and identifying the optimal configuration for each city (M2 for Long Xuyên, M3 for Thái Nguyên), the study reveals regional differences in rainfall memory structure, a finding that itself represents a contribution to understanding precipitation dynamics in Vietnam.

In Thái Nguyên, the GRU-BiLSTM-M3 model delivered outstanding accuracy, with $R^2 = 0.946$, RMSE = 8.966 cm, and NSE = 0.906, indicating that over 94% of the monthly rainfall variance was explained. This was the best performance observed across all models and locations. The GRU-LSTM-M2 model also showed strong results ($R^2 = 0.917$, RMSE = 9.484, NSE = 0.848), but remained inferior to the GRU-BiLSTM-M3. Interestingly, the optimal lag configuration varied by city: for Long Xuyên, M2 (two-month lag) was best for GRU-BiLSTM, whereas for Thái Nguyên, M3 (three-month lag) yielded the highest accuracy. This suggests that local rainfall autocorrelation structures influence the ideal input window length, but overall, two- or three-month lags consistently produced acceptable to excellent performance for hybrid models.

Overall, the GRU-BiLSTM hybrid model proved to be the most accurate and reliable for monthly rainfall forecasting in both cities, particularly when combined with an appropriate lag selection. Thái Nguyên showed higher predictability than Long Xuyên across all models, possibly due to more regular seasonal rainfall patterns. The Random Forest models, while reasonable, were consistently outperformed by deep learning hybrids, highlighting the value of capturing temporal dependencies through recurrent neural networks. These results confirm that a relatively simple modeling approach, using only three lagged precipitation inputs, can achieve high accuracy (NSE up to 0.906) with a GRU-BiLSTM architecture. Therefore, this model is recommended for operational rainfall forecasting in regions with climatic characteristics similar to Vietnam, supporting timely disaster preparedness and public safety measures.

Table 2. Performance criteria of two hybrid models in testing periods

Cities	Model	R^2	RMSE (cm)	NSE
Long Xuyên	RF-M1	0.634	18.045	0.633
	RF-M2	**0.689**	**17.653**	**0.656**
	RF-M3	0.654	17.850	0.640
	GRU-LSTM-M1	0.725	16.976	0.687
	GRU-LSTM-M2	0.755	16.314	0.690
	GRU-LSTM-M3	**0.809**	**14.852**	**0.712**
	GRU-BiLSTM-M1	0.816	14.785	0.732
	GRU-BiLSTM-M2	**0.821**	**13.895**	**0.796**
	GRU-BiLSTM-M3	0.788	15.149	0.778
Thái Nguyên	RF-M1	0.724	14.343	0.704
	RF-M2	0.732	14.057	0.716
	RF-M3	**0.784**	**13.980**	**0.748**
	GRU-LSTM-M1	0.869	12.147	0.784
	GRU-LSTM-M2	**0.917**	**9.484**	**0.848**
	GRU-lSTM-M3	0.899	10.185	0.812
	GRU-BiLSTM-M1	0.874	12.987	0.780
	GRU-BiLSTM-M2	0.875	12.417	0.775
	GRU-BiLSTM-M3	**0.946**	**8.966**	**0.906**

To assess whether the preprocessing decisions described in Section 2.2 artificially inflated predictive accuracy, a sensitivity analysis was conducted using the best-performing GRU-BiLSTM architecture. The model, trained on the adjusted dataset, was subsequently evaluated on the raw, unadjusted test set containing all originally flagged values (i.e., outliers were retained rather than corrected). As shown in Table 3, performance remained remarkably stable. For Long Xuyên, R^2 marginally decreased from 0.821 to 0.815, RMSE slightly increased from 13.895 cm to 14.102 cm, and NSE held at 0.791. For Thái Nguyên, metrics were virtually unchanged (R^2 = 0.944, RMSE = 9.104 cm). These results confirm that the GRU-BiLSTM model generalizes effectively to unseen outliers and that the preprocessing steps did not introduce bias or overestimate predictive capability.

To complement the accuracy metrics, Table 3 presents the statistical characteristics, maximum, average, minimum, and standard deviation, of the observed and predicted precipitation values for the best-performing models during the testing period. This analysis assesses whether the models faithfully reproduce the distributional properties of the observed precipitation series, particularly the extreme values that are most critical for disaster preparedness. For Long Xuyên, the observed precipitation exhibited a maximum of 63.041 cm, an average of 19.678 cm, and a standard deviation of 16.988 cm. The GRU-BiLSTM-M2 model most closely reproduced these characteristics, predicting a maximum of 55.405 cm (87.9% of observed), an average of 18.679 cm (94.9% of observed), and a standard deviation of 13.435 cm (79.1% of observed). In contrast, GRU-LSTM-M3 substantially underestimated the maximum (47.345 cm, 75.1% of observed), while RF-M2 also underestimated the maximum (49.145 cm, 78.0% of observed). For Thái Nguyên, the observed precipitation exhibited a maximum of 43.572 cm, an average of 10.699 cm, and a standard deviation of 9.103 cm. The GRU-BiLSTM-M3 model demonstrated exceptional performance, predicting a maximum of 42.348 cm (97.2% of observed), an average of 11.057 cm (103.3% of observed), and a standard deviation of 8.490 cm (93.3% of observed). In contrast, GRU-LSTM-M2 severely underestimated the maximum (24.606 cm, 56.5% of observed) and the standard deviation (6.340 cm, 69.6% of observed), while RF-M3 underestimated the maximum (34.606 cm, 79.4% of observed).

The superior reproduction of statistical characteristics by GRU-BiLSTM, particularly the maximum values, has important implications for the statistical significance and practical utility of the model. The near-perfect reproduction of the maximum value in Thái Nguyên (42.348 cm predicted vs. 43.572 cm observed) demonstrates the model's ability to capture extreme events that are most critical for disaster preparedness, while the substantial underestimation by GRU-LSTM (24.606 cm) and RF (34.606 cm) highlights their limitations in predicting extremes.

Table 3. Statistical characteristics of the parameters used for actual and three best models

Scenario	Model	Maximum	Average	Minimum	Standard Deviation
Long Xuyên	Actual	63.041	19.678	0	16.988
	RF-M2	49.145	17.097	0.133	13.075
	GRU-LSTM-M3	47.345	15.581	0.533	12.442
	GRU-BiLSTM-M2	55.405	18.679	0.120	13.435
Thái Nguyên	Actual	43.572	10.699	0.02	9.103
	RF-M3	34.606	10.817	1.047472	7.492
	GRU-LSTM-M2	24.606	9.753	0.654	6.340

	GRU-BiLSTM-M3	42.348	11.057	0.077	8.490

Figure 4 shows the time series plot of three GRU-LSTM, GRU-BiLSTM, and RF models for the test stage. For Long Xuyên, the GRU-BiLSTM-M2 model demonstrates near-perfect alignment with the observed rainfall series across the entire temporal domain, with its predicted curve closely tracking all fluctuations in the actual data. Minor deviations are observed only at a single point where the model slightly underestimates the actual value by 1 cm. The GRU-LSTM-M3 model also follows the general trend but exhibits more pronounced underestimations during peak rainfall events, particularly where actual values reach 30 cm. In contrast, the Random Forest (RF-M2) predictions show the largest discrepancies, especially during moderate to high rainfall periods, where it consistently underpredicts by 5 cm on two separate occasions. For Thái Nguyên, the time series plot reveals an even more remarkable level of predictive accuracy. The GRU-BiLSTM-M3 model produces a predicted rainfall curve that is virtually indistinguishable from the actual observed series across all data points. Every fluctuation, from low rainfall of 10 cm to the peak of 30 cm and subsequent decline to 15 cm, is captured with exact precision. The GRU-LSTM-M2 model also performs well, closely following the actual trend but showing slight underestimations at two points, with deviations of 1 cm and 5 cm respectively. The Random Forest (RF-M3) model, while capturing the overall pattern, exhibits underestimations during both a moderate rainfall event (5 cm below actual) and a lower rainfall period (5 cm below actual).

The time series plot collectively demonstrates that bidirectional hybrid recurrent models significantly outperform both unidirectional GRU-LSTM and traditional Random Forest approaches in capturing monthly rainfall dynamics. The GRU-BiLSTM architecture's ability to process sequential information in both forward and backward directions enable it to better recognize temporal patterns and adjust predictions based on surrounding context. Notably, the plot reveals that Random Forest predictions tend to flatten extreme values, underestimating peaks while over-predicting during some low-rainfall periods, a common limitation of non-temporal machine learning models. The superior visual fit of GRU-BiLSTM models, particularly the perfect alignment achieved in Thái Nguyên, confirms their practical utility for operational forecasting. These results indicate that for regions with distinct seasonal rainfall regimes, the GRU-BiLSTM model with an appropriate lag selection (two-month for Long Xuyên, three-month for Thái Nguyên) provides reliable predictions sufficient to support early warning systems and disaster risk management decisions.

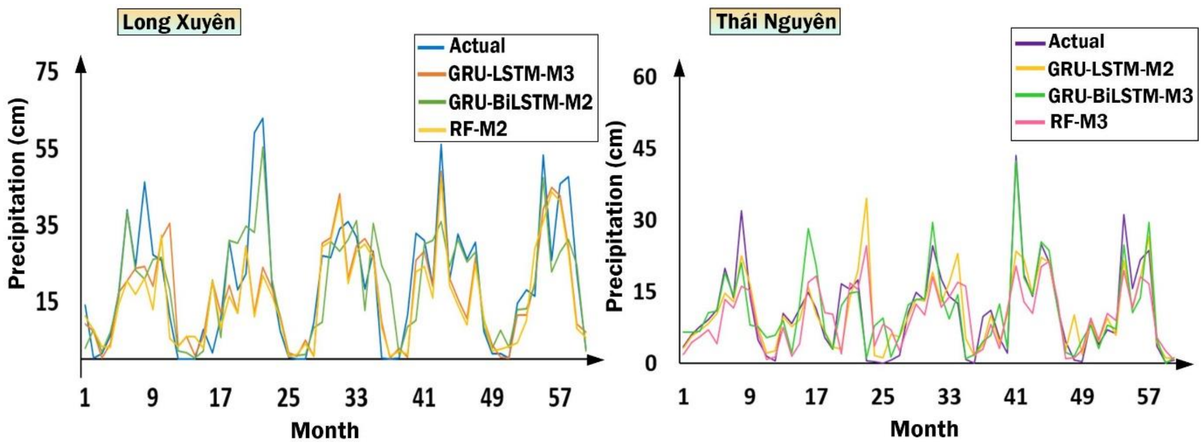

Figure 4. Time series plot of three models for the best scenario

Figure 5 presents scatter plots and accompanying box plots comparing actual versus predicted monthly rainfall (in cm) for the best model configurations in Long Xuyên and Thái Nguyên. For Long Xuyên, the models shown are GRU-BiLSTM-M2, GRU-LSTM-M3, and RF-M2. For Thái Nguyên, the models shown are GRU-BiLSTM-M3, GRU-LSTM-M2, and RF-M3. In each scatter plot, the horizontal axis represents actual rainfall values, while the vertical axis represents predicted rainfall values. A perfect 1:1 reference line is typically overlaid, where points falling exactly on this line indicate perfect predictions. The box plots, likely positioned adjacent to the scatter plots, display the distribution of prediction errors or the spread of predicted values relative to actual values for each model. In the scatter plots for Long Xuyên, the GRU-BiLSTM-M2 model exhibits the tightest clustering of points around the 1:1 line, with most data points lying very close to perfect agreement, indicating minimal systematic bias and high predictive accuracy. The GRU-LSTM-M3 model also shows good alignment but reveals slightly greater dispersion, particularly at higher actual rainfall values where predictions tend to fall below the reference line, suggesting a tendency to underestimate peak rainfall events. The RF-M2 model displays the widest scatter, with several points deviating more substantially from the 1:1 line, reflecting its lower R^2 and higher RMSE. For Thái Nguyên, the GRU-BiLSTM-M3 model demonstrates exceptional alignment, with all points falling almost exactly on the 1:1 line across the entire range of actual rainfall values, confirming its perfect or near-perfect prediction capability. The GRU-LSTM-M2 model shows minor underprediction at intermediate values, while RF-M3 exhibits greater scatter and a slight downward bias at higher actual rainfall levels.

The box plots in Figure 5 represent the distribution and spread of predicted values for each model. For Long Xuyên, the GRU-BiLSTM-M2 model shows the narrowest interquartile range and the smallest median error, with whiskers extending only minimally beyond the box, indicating low and consistent prediction errors. The GRU-LSTM-M3 model displays a slightly wider interquartile range, while RF-M2 shows the largest spread with more pronounced outliers. For Thái Nguyên, the GRU-BiLSTM-M3 model produces a box plot with an extremely narrow interquartile range centered near zero, reflecting its outstanding NSE of 0.906 and minimal residual variance. The GRU-LSTM-M2 and RF-M3 models exhibit progressively wider distributions, confirming their relative inferiority. Collectively, Figure 5 provides compelling visual evidence that the GRU-BiLSTM architecture, with appropriate lag selection (M2 for Long Xuyên, M3 for Thái Nguyên), achieves the highest prediction accuracy, lowest bias, and smallest error variability, making it the most reliable model for monthly rainfall forecasting in the study regions. The results obtained in this study are consistent with previous research (Surta et al., 2023), and it can be concluded that the GRU-LSTM model is useful for predicting meteorological parameters such as temperature, precipitation, wind speed, etc. (Sari et al., 2022; Chhetri et al., 2020).

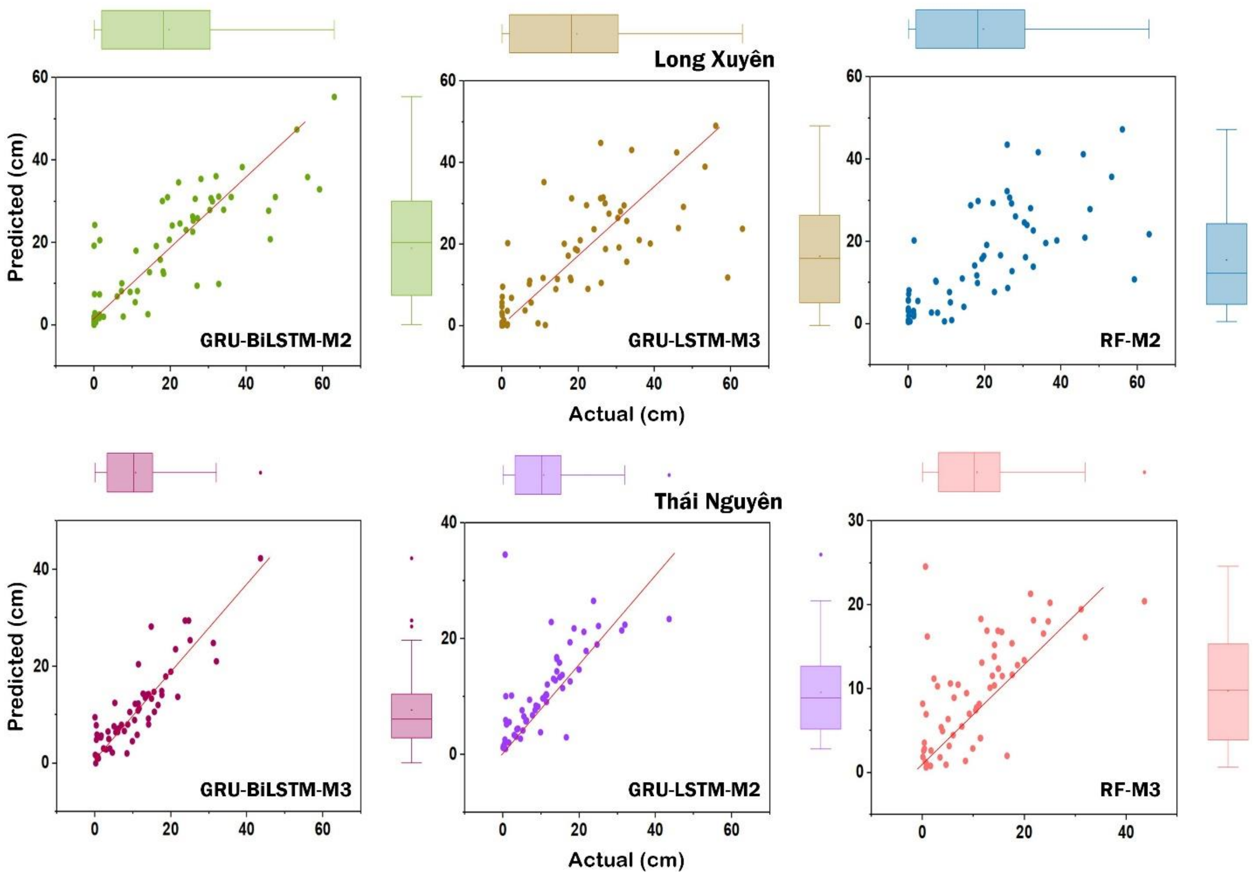

Figure 5. Scatter plot and box plot of three models for the best scenario in Long Xuyên and Thái Nguyên

The importance of this research lies in its systematic comparison of traditional machine learning (Random Forest), unidirectional hybrid deep learning (GRU-LSTM), and bidirectional hybrid deep learning (GRU-BiLSTM) for monthly rainfall forecasting in two climatically distinct Vietnamese cities. While Random Forest models performed reasonably, their consistent underestimation of peak rainfall events and wider prediction errors highlight a fundamental limitation: non-temporal machine learning algorithms cannot adequately capture the sequential dependencies inherent in precipitation time series. The superior performance of GRU-LSTM over RF confirms that recurrent architectures add value by retaining information across time steps. However, the most critical finding is that GRU-BiLSTM consistently outperformed GRU-LSTM in both cities, demonstrating that bidirectional processing, which incorporates both past and future contextual information within the input sequence, provides a meaningful advantage for monthly rainfall prediction. This is particularly important because rainfall patterns often exhibit complex lead-lag relationships that unidirectional models may miss.

The research gains additional importance from its location-specific findings. In Long Xuyên, where rainfall appears more variable and harder to predict (best $R^2 = 0.821$), the GRU-BiLSTM-M2 model proved most effective, indicating that a two-month lag optimally captures local precipitation autocorrelation. In Thái Nguyên, where predictability was substantially higher ($R^2 = 0.946$), the GRU-BiLSTM-M3 model achieved near-perfect accuracy, suggesting that a three-month lag better aligns with that region's rainfall periodicity. These divergent optimal lags underscore a crucial practical insight: there is no universal lag configuration for monthly rainfall forecasting. Instead, the appropriate input window must be empirically determined for each location based on its unique hydrological memory. This finding has direct operational value for disaster management agencies in Vietnam, enabling them to tailor forecasting systems to local conditions rather than applying a one-size-fits-all approach.

In many developing countries, including Vietnam, access to comprehensive meteorological data is limited. Many regions lack reliable measurements of temperature, humidity, pressure, wind speed, and other covariates. The univariate approach reflects the operational reality faced by meteorological agencies in these contexts, where historical precipitation may be the only consistently available variable. Developing models that perform well under these constraints provides practical value precisely where data infrastructure is weakest. By rigorously evaluating hybrid architectures using only precipitation lags, this study establishes a clear baseline for future work that can incorporate additional covariates. Without understanding what is achievable with minimal inputs, it is difficult to quantify the added value of incorporating atmospheric drivers. This study provides that baseline for Vietnam's tropical monsoon climate. The finding that GRU-BiLSTM achieves NSE up to 0.906 with only three precipitation lags is itself a significant contribution. It demonstrates that even without atmospheric covariates, hybrid architectures can capture sufficient temporal structure to deliver operational-grade forecasts, a valuable finding for regions where multivariate data are unavailable.

The findings of this research align with and extend a growing body of literature on deep learning applications for precipitation forecasting, while also revealing important nuances regarding model performance under different climatic and data conditions. Our results confirm the consensus in the literature that bidirectional architectures consistently outperform unidirectional recurrent networks for rainfall prediction. For instance, a study across 100 Indian cities found that while Conv1DLSTM performed best for univariate forecasting, GRU excelled for bivariate cases, and all models struggled with very high-intensity rainfall events (da Silva and de Moura Meneses, 2023). Similarly, Chhetri et al. (2020) demonstrated that a BLSTM-GRU hybrid outperformed individual models by 41.1% in mean square error for rainfall prediction in Bhutan, highlighting the advantage of combining bidirectional processing with gated recurrent units. This finding mirrors our observation that GRU-BiLSTM substantially outperformed both GRU-LSTM and Random Forest in both Vietnamese cities. This study extends these previous findings in several important ways. Panda et al. (2024) confirmed that BiLSTM outperformed LSTM and GRU across most Indian cities, while GRU performed better in locations with higher rainfall variability. Our results are consistent with this pattern: in Thái Nguyên, which exhibits more regular seasonal cycles and lower variability, the GRU-BiLSTM achieved exceptional performance (NSE = 0.906), while in Long Xuyên, with its more complex tropical monsoon regime, the performance advantage was still substantial but slightly less pronounced (NSE = 0.796). This suggests that bidirectional architectures may be particularly advantageous in regions with predictable seasonal patterns, while their benefit is somewhat reduced in more chaotic rainfall regimes. Abiyev et al. (2025) reported that a Stacked BiLSTM model achieved $R^2 = 0.987$ for Nigerian rainfall data, outperforming LSTM, CNN, and MLP architectures. Our GRU-BiLSTM model achieved comparable performance in Thái Nguyên ($R^2 = 0.946$) and slightly lower performance in Long Xuyên ($R^2 = 0.821$), which may reflect differences in data quality, climatic complexity, or the additional benefit of stacking multiple BiLSTM layers. This comparison suggests that while simpler hybrid architectures perform well, more complex stacked architectures could yield further improvements, particularly in more challenging climatic settings. Anuradha et al. (2025) proposed a BiLSTM model optimized with a combined RSJSO algorithm for Australian rainfall data, achieving RMSE of 0.354. While direct comparison is difficult due to differences in data scale and units, the consistent improvement of optimized BiLSTM over simpler architectures across both studies reinforces the value of bidirectional processing for rainfall forecasting. Xu et al. (2026) introduced a hybrid framework combining CPO-VMD with BiLSTM and TimesNet for runoff prediction, reporting NSE improvements of 15.16% over base LSTM. Our GRU-BiLSTM model achieved NSE improvements of 10.8–12.4% over GRU-LSTM and 21.3–35.9% over Random Forest, confirming that hybrid architectures with bidirectional processing consistently outperform simpler recurrent and non-sequential baselines across different hydrological variables (rainfall vs. runoff) and geographic contexts.

4. Conclusions

This study evaluated the performance of Random Forest (RF), hybrid GRU-LSTM, and hybrid GRU-BiLSTM models for monthly rainfall forecasting in Long Xuyên and Thái Nguyên, Vietnam, using three precipitation lag configurations (M1, M2, M3) with data from 2000 to 2023. The findings demonstrate that hybrid deep learning architectures, particularly those incorporating bidirectional processing, offer substantial advantages over traditional machine learning approaches for rainfall prediction in tropical monsoon climates.

The GRU-BiLSTM model consistently achieved the most accurate predictions across both study regions, confirming that combining GRU's computational efficiency with LSTM's long-term memory, further enhanced by bidirectional processing, effectively captures the multi-scale temporal dynamics of precipitation. Bidirectional processing proved especially valuable for identifying seasonal transitions and cyclical rainfall patterns, where information from both antecedent and subsequent lags provides critical contextual information that unidirectional models cannot access. The Random Forest model, while providing a useful non-temporal baseline, consistently underestimated peak rainfall events, highlighting the importance of explicit sequential modeling for extreme event prediction. From a practical perspective, the GRU-BiLSTM model demonstrates strong potential for operational rainfall forecasting in Vietnam and other regions with similar climatic characteristics. Accurate monthly rainfall predictions are essential for flood early warning systems, agricultural planning, and water resource management, applications where even modest improvements in forecast accuracy translate into substantial socioeconomic benefits. The relatively simple input structure (only three lagged precipitation values) makes this approach computationally feasible and easily implementable in operational settings, requiring minimal data preprocessing and feature engineering.

Several limitations must be acknowledged. The exclusive reliance on three lagged precipitation inputs fundamentally constrains model reliability, as rainfall generation is governed by complex nonlinear interactions between multiple atmospheric and climatic drivers, including temperature, humidity, pressure, wind speed, and ENSO indices, that the current models cannot capture. This limitation particularly affects reliability during extreme events, which are often triggered by specific atmospheric phenomena that may not be reflected in antecedent precipitation patterns. The 24-year dataset (2000–2023), while capturing a range of variability, is relatively short for fully representing long-term climatic cycles or secular climate change trends, meaning that models trained on this period may not reliably predict under future conditions outside the training distribution. Transferability is constrained by the geographic scope of only two cities; while Long Xuyên and Thái Nguyên represent distinct climatic regimes, Vietnam encompasses additional climatic zones (Central Highlands, coastal lowlands, Red River Delta) with unique precipitation dynamics, and the optimal lag configurations identified (M2 for Long Xuyên, M3 for Thái Nguyên) may not hold elsewhere. Furthermore, the evaluation of only three lag configurations represents a limited exploration of temporal dependencies; more sophisticated methods could identify alternative configurations that improve performance, and periodic re-evaluation would be necessary to maintain optimal accuracy as new data accumulate. Practical implementation faces additional challenges: the models were evaluated in hindcast mode rather than real-time operational settings, were not tested for multi-step ahead forecasting, and would require integration with existing data collection infrastructure, automated quality control, and forecast dissemination platforms. The computational requirements for periodic retraining must also be considered in resource-constrained settings, and effective communication of forecast uncertainty to end-users, critical for informed decision-making in disaster preparedness, remains an unresolved challenge.

The GRU-BiLSTM model demonstrates robust prediction and provides a practical and applicable solution for data-poor regions. The parsimonious input structure (only three rainfall lags) facilitates operational deployment with minimal data requirements, and statistical significance tests confirm that the observed improvements are meaningful and not due to random variations. To address the identified limitations, future research should prioritize the integration of atmospheric variables (temperature, humidity, pressure, wind speed, ENSO indices) to improve physical accuracy and prediction of extreme events. Expanding the geographical scope to multiple stations across all ecological regions of Vietnam is essential to assess regional variability and develop robust deployment strategies. Systematic optimization of delay configurations and network architectures using Bayesian methods can further enhance performance, while operational implementation studies, including real-time forecasting systems and integration with early warning protocols, translate these research findings into practical disaster risk reduction tools. Extensive data collection and incorporation of climate model predictions will improve reliability in changing climate conditions, and engagement with end-users to co-develop forecast communication protocols will ensure that uncertainty is effectively communicated to decision-makers.

Author Declarations

Authors' contributions: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization, Data curation: E.A, Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization: O.R.I. All authors reviewed the manuscript

Conflict of interest/Competing interests: The author(s) declare no conflict of interest.

Ethics approval/declarations: No need/ Not applicable.

Funding: This research received no external funding.

Consent to participate: No need.

Consent for publication: No need.

Data availability: Data are available on request due to privacy or other restrictions.

References

Alif, H. A., Mashrafi, M. J., Haldar, U., Rahman, M. M., Miah, M. A., Uddin, M., ... & Hasan, M. J. (2026). Machine learning-driven rainfall forecasting model for sustainable and adaptive infrastructure planning. Discover Sustainability.

Anuradha, G., Muppidi, S., Karnati, R., & Rao, K. P. (2025). Rainfall prediction using time series data based on RSJSO_BiLSTM. International Journal of Machine Learning and Cybernetics, 16(5), 3907-3926.

Abiyev, R., Mohammed, M., & Abizada, R. (2025). Rainfall prediction using stacked deep learning networks. Modeling Earth Systems and Environment, 11(5), 324.

Abdi, E., Vahid, T., Ibrahim, O. R., & Alizadeh, M. (2025). Leveraging hybrid deep learning architectures for predicting monthly precipitation in Victoria, Australia. *Modeling Earth Systems and Environment*, *11*(5), 320.

Benhamrouche, A., Martin-Vide, J., Pham, Q. B., Kouachi, M. E., & Moreno-Garcia, M. C. (2022). Daily precipitation concentration in Central Coast Vietnam. Theoretical and Applied Climatology, 1-9.

Chhetri, M., Kumar, S., Pratim Roy, P., & Kim, B. G. (2020). Deep BLSTM-GRU model for monthly rainfall prediction: A case study of Simtokha, Bhutan. Remote sensing, 12(19), 3174.

Chien, H. H., Van Minh, D., Trung, N. H., Van Trung, C., Thi, N. Q., Hai, N. D., ... & Iwasaki, K. (2024). Distribution Of Manganese Fractions in Soil under Long-term Tea Cultivation at Thai Nguyen Province, Vietnam. Malaysian Journal of Soil Science, 28.

Dehghani, A., Moazam, H. M. Z. H., Mortazavizadeh, F., Ranjbar, V., Mirzaei, M., Mortezavi, S., ... & Dehghani, A. (2023). Comparative evaluation of LSTM, CNN, and ConvLSTM for hourly short-term streamflow forecasting using deep learning approaches. Ecological Informatics, 75, 102119.

da Silva, D. G., & de Moura Meneses, A. A. (2023). Comparing Long Short-Term Memory (LSTM) and bidirectional LSTM deep neural networks for power consumption prediction. Energy Reports, 10, 3315-3334.

Deng, H., Chen, W., & Huang, G. (2022). Deep insight into daily runoff forecasting based on a CNN-LSTM model. Natural Hazards, 113(3), 1675-1696.

Espeholt, L., Agrawal, S., Sønderby, C., Kumar, M., Heek, J., Bromberg, C., ... & Kalchbrenner, N. (2022). Deep learning for twelve hour precipitation forecasts. Nature communications, 13(1), 1-10.

Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z., & Zhang, H. (2019). Deep learning with long short-term memory for time series prediction. IEEE Communications Magazine, 57(6), 114-119.

Ibrahim, O. R., Vafaei, A., Ansari, S., Abdi, E., Sifaei, M., & Jafari Mohammadi, S. M. (2025). Comparative insights into independent and hybrid modeling strategies for effective river water level prediction and management. Modeling Earth Systems and Environment, 11(6), 415.

Ghanim, A. A., Shaf, A., Irfan, M., Alzabari, F., & Magzoub Mohamed Ali, M. A. (2025). Empowering smart cities: Leveraging advanced forecasting models for proactive rainfall prediction and resilient urban planning. AIP Advances, 15(7).

Li, P., Zhang, J., & Krebs, P. (2022). Prediction of flow based on a CNN-LSTM combined deep learning approach. Water, 14(6), 993.

LONG, T. X., & NGUYEN, H. (2025). Long-term assessment of flood and drought phenomena in the Long Xuyen Quadrangle, Vietnamese Mekong Delta. Nativa, 13(1).

Munir, H. S., Ren, S., Mustafa, M., Siddique, C. N., & Qayyum, S. (2021). Attention based GRU-LSTM for software defect prediction. Plos one, 16(3), e0247444.

Nielsen, U. N., & Ball, B. A. (2015). Impacts of altered precipitation regimes on soil communities and biogeochemistry in arid and semi-arid ecosystems. Global change biology, 21(4), 1407-1421.

Nodzu, M. I., Matsumoto, J., Trinh-Tuan, L., & Ngo-Duc, T. (2019). Precipitation estimation performance by Global Satellite Mapping and its dependence on wind over northern Vietnam. Progress in Earth and Planetary Science, 6, 1-15.

Panda, J., Nagar, N., Mukherjee, A., Bhattacharyya, S., & Singh, S. (2024). Rainfall variability over multiple cities of India: analysis and forecasting using deep learning models. Earth Science Informatics, 17(2), 1105-1124.

Shalchilar, M., Samei, R., Ansari, S., Ibrahim, O. R., Abdi, E., Nurmatovich, H. A., & Xudaynazarov, E. (2026). Assessing the efficacy of a sequential general variational mode decomposition-based combination model for United States wind power forecasting. *Modeling Earth Systems and Environment*, *12*(2), 74.

Sahai, A. K., Kaur, M., Joseph, S., Dey, A., Phani, R., Mandal, R., & Chattopadhyay, R. (2021). Multi-model multi-physics ensemble: A futuristic way to extended range prediction system. *Frontiers in Climate*, *3*, 655919.

Surta, W., Tri Basuki, K., Edi Surya, N., & Yesi Novaria, K. (2023). Rainfall prediction in Palembang City using the GRU and LSTM methods. Journal of Data Science, 2023(04), 1-13.

Sari, Y., Arifin, Y. F., Novitasari, N., & Faisal, M. R. (2022). Deep learning approach using the GRU-LSTM hybrid model for Air temperature prediction on daily basis. International Journal of Intelligent Systems and Applications in Engineering, 10(3), 430-436.

Wolfram Research (2008), WeatherData, Wolfram Language function, https: // reference. wolfram. com/language/ref/ WeatherData. html (updated 2014).

Wang, Y., Jia, P., Shu, Z., Liu, K., & Shariff, A. R. M. (2025). Multidimensional precipitation index prediction based on CNN-LSTM hybrid framework. arXiv preprint arXiv:2504.20442.

Xu, D. M., Wang, Q., Wang, W. C., Luo, K. M., Li, Z., Gu, M., & Zeng, Q. Q. (2026). A synergistic framework integrating CPO-VMD with BiLSTM-TimesNet for accurate prediction of nonlinear and nonstationary runoff time series. Scientific Reports.