Integrating hybrid deep learning and CSI for multi-interval hydrological data in enhanced flood prediction

Indrastanti Ratna Widiasari, Eko Sediyono, Rissal Efendi
Department of Informatics Engineering, Faculty of Information Technology, Satya Wacana Christian University, Salatiga, Indonesia

Article Info

Article history:

Received Aug 25, 2025
Revised Mar 26, 2026
Accepted Apr 19, 2026

Keywords:

Cubic spline interpolation
Flood prediction
Hybrid deep learning
Hydrological modeling
Long short-term memory-convolutional neural network

ABSTRACT

Flood prediction accuracy is often constrained by heterogeneous and asynchronous hydrological data collected at different time intervals. This study proposes a hybrid deep learning–based flood prediction framework that integrates long short-term memory (LSTM), convolutional neural network (CNN), and cubic spline interpolation (CSI) to address these challenges. Rainfall, river discharge, and water level data representing upstream, midstream, and downstream conditions of the Bengawan Solo watershed were utilized. CSI was applied as a preprocessing step to harmonize multi-interval data, reduce noise, and recover missing observations, thereby improving data consistency. The experimental results show that the proposed hybrid LSTM–CNN model enhanced with CSI outperforms baseline LSTM and non-interpolated hybrid models, achieving a mean absolute percentage error (MAPE) of 5.84%, root mean square error (RMSE) of 0.125 m, mean absolute error (MAE) of 0.082 m, and R^2 of 0.948. The integration of spatio-temporal feature learning with data harmonization enables more accurate flood level prediction and supports timely flood early warning systems. The proposed approach demonstrates strong potential for improving flood risk management and disaster preparedness in flood-prone regions.

Corresponding Author:

Indrastanti Ratna Widiasari
Department of Informatics Engineering, Faculty of Information Technology
Satya Wacana Christian University
Diponegoro 52-60, Salatiga, Indonesia
Email: indrastanti@uksw.edu

1. INTRODUCTION

Flooding is a recurring natural disaster in Indonesia, particularly in areas intersected by large river systems and complex irrigation networks [1], [2]. This phenomenon poses serious threats to human safety, damages infrastructure, and disrupts agricultural activities. As climate variability increases, the frequency and intensity of floods continue to rise, especially in lowland and river-adjacent regions. An effective and accurate flood prediction system is essential to enable communities and authorities to take early and appropriate mitigation measures [3], [4].

Bojonegoro Regency is one of the most vulnerable to floods which is located along the Bengawan Solo River and its tributaries. Water levels at dams such as Gerak, Gongseng, and Pacal are often elevated as a result of high rainfall and increased upstream water discharge, endangering downstream communities. Although these essential hydraulic, current early warning systems often lack the accuracy and timeliness needed for effective disaster response. A significant disparity can be observed between available hydrological data and the predictive frameworks being utilized, in particular, in terms of integrating spatial and temporal variables in real time [5]–[7]. Another challenge is that hydrological variables are often recorded at different time intervals, leading to inconsistencies and irregularities in the datasets. Such variations complicate data integration and forecasting, since models must align and process asynchronous time series.

Various computational approaches to flood forecasting have been investigated in recent studies, for example conventional machine learning algorithms, such as support vector machines, decision trees, and regression models [8]–[10]. However, these models are inadequate in modelling nonlinear dependencies, complex temporal-spatial interactions, and are limited in their scalability for real-world applications [11]–[13]. In contrast, deep learning approaches, particularly long short-term memory (LSTM) networks and convolutional neural networks (CNN) have shown better potential [14]–[16]. LSTM is effective in capturing long-term temporal sequences [17], [18], whereas CNN demonstrates processing spatially structured data [19]. Nonetheless, most current models have not yet achieved optimal integration of both temporal and spatial features in a unified predictive framework. Hydrological variables are often recorded at different time intervals, leading to inconsistencies and irregularities in the datasets. Such variations complicate data integration and forecasting, since models must align and process asynchronous time series. To address this challenge, this study employs cubic spline interpolation (CSI) to harmonize multi-interval data. This study introduces a hybrid approach that integrates LSTM, CNN, and CSI to improve the accuracy and responsiveness of flood prediction [20], [21]. LSTM, CNN, and CSI play complementary roles: LSTM captures long-term temporal dependencies and CNN extracts spatial features, and CSI is utilized for the smoothing and reconstructive hydrological multi-interval time-series data, reducing noise and improving the quality of input fed into the deep learning models [22], [23].

Despite the extensive application of machine learning and deep learning models for flood prediction, several critical limitations remain in existing studies. Most previous works focus primarily on temporal modeling using LSTM or spatial feature extraction using CNN independently, without effectively integrating both capabilities within a unified prediction framework. In addition, hydrological monitoring data are often collected at heterogeneous and asynchronous time intervals, yet many prior studies assume uniform temporal resolution or apply simple interpolation techniques that do not adequately address noise and missing values. These limitations reduce model robustness and may lead to unstable predictions, particularly during rapidly changing hydrological conditions. Therefore, a clear research gap exists in developing an integrated flood prediction framework that simultaneously handles spatio-temporal feature learning and harmonizes multi-interval hydrological datasets. This study addresses this gap by integrating a hybrid LSTM–CNN architecture with CSI to improve data consistency and predictive accuracy.

The novelty of this study lies in the explicit integration of a hybrid LSTM–CNN architecture with CSI to address both spatio-temporal feature learning and multi-interval hydrological data harmonization within a single flood prediction framework. This study uses a hybrid model that integrates LSTM, CNN, and CSI architectures to improve the accuracy and responsiveness of flood prediction. In contrast to previous models that process either temporal or spatial features separately, the proposed model combines real-time rainfall [24], [25], water discharge [26], [27], and water level data [28] from multiple upstream dams into a unified prediction model. Quantitatively, the proposed approach reduces prediction error from 7.85% (baseline LSTM) to 5.84% mean absolute percentage error (MAPE) and improves the coefficient of determination from 0.921 to 0.948, demonstrating the measurable contribution of the proposed framework.

This study focuses on three main objectives: developing a deep learning-based flood prediction system capable of handling spatio-temporal hydrological data, enhancing the speed and precision of water level forecasts for earlier flood warnings and employing CSI reconstruct time-series data with differing intervals, reducing noise and improving the quality of inputs for the predictive model. The expected benefits include more accurate flood risk assessments, improved disaster preparedness, and more timely mitigation actions for at-risk communities. This research addresses a critical need for more intelligent and integrated flood prediction systems in vulnerable regions like Bojonegoro and Tuban, East Java, Indonesia. The combination of deep learning methods with real-time monitoring data offers a practical solution to bridge existing gaps in early warning systems. The contribution is both scientific and societal: it advances the field of hydrological modelling through a novel hybrid architecture and supports disaster mitigation efforts through accurate and actionable flood forecasts. As extreme weather events become more frequent, such predictive systems will be increasingly vital for climate resilience and public safety.

2. METHOD

This study aims to develop a flood prediction model using a hybrid approach that combines deep learning techniques with CSI through four main stages. The first stage involves the collection and preprocessing of hydrological data, including rainfall, river discharge, and water level. Data were obtained from Bengawan Solo Central River Basin Authority, the Water Resources Management Agency, and the Regional Disaster Management Agency, as well as through field studies conducted in the Bengawan Solo River Basin and the Gerak Dam in Bojonegoro Regency, East Java. The data, recorded at varying time intervals (rainfall every 1 hour, discharge every 15 minutes, and water level every 10 minutes), were used to analyze hydrological dynamics and the influence of rainfall on discharge and water level fluctuations. To synchronize data with differing temporal resolutions, CSI was applied to enhance the accuracy and consistency of inputs for the hydrological model. The second stage focuses on the design and development of a flood prediction model using a hybrid configuration based on LSTM and CNN architectures, with optimized hyperparameters such as the number of epochs, batch size, and learning rate. The third stage involves training and validating the model using the interpolated dataset, followed by performance evaluation using metrics such as MAPE, R^2, mean absolute error (MAE), and root mean square error (RMSE). The fourth stage includes result interpretation and model analysis to support decision-making in water resource management and flood mitigation efforts in the study area. Figure 1 presents the workflow of the proposed hybrid flood forecasting model, outlining each stage from data acquisition to model deployment, including preprocessing, spline interpolation, and deep learning integration.

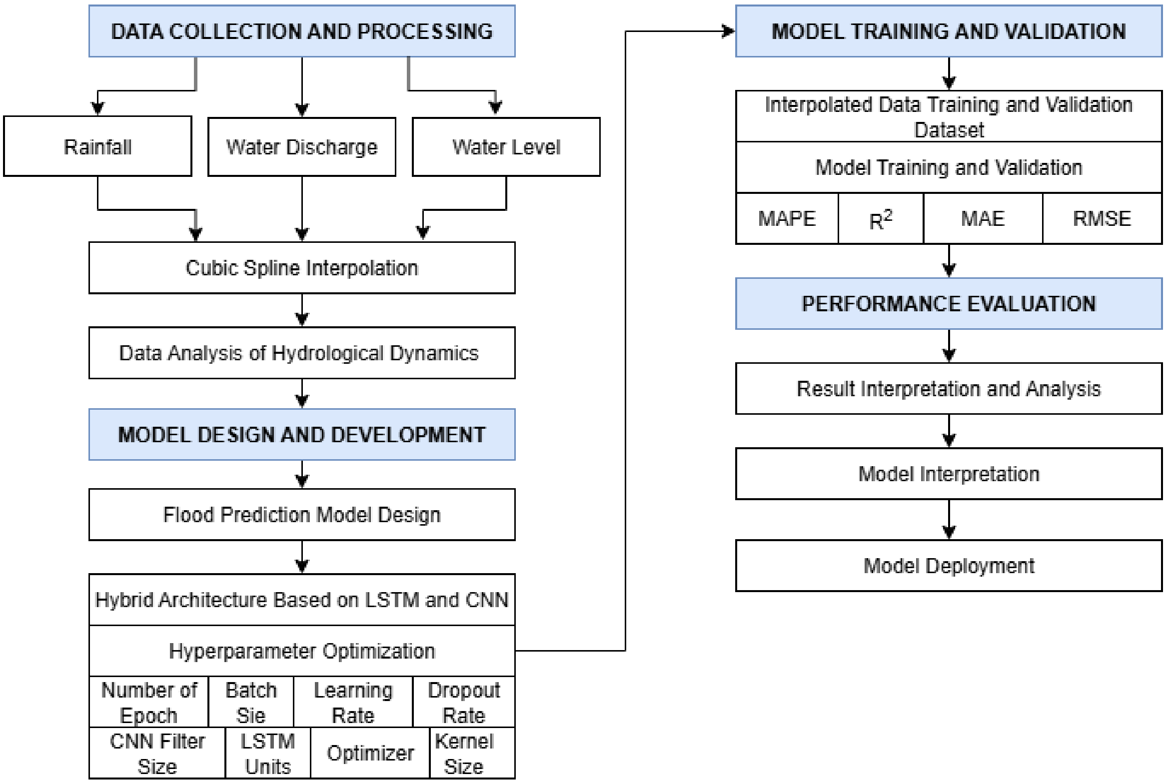


Figure 1. Architecture of the proposed flood prediction system

2.1. Data collection and preprocessing

Accurate flood forecasting requires high-resolution hydrological data and consistent temporal alignment across multiple parameters. In this study, the hydrological dataset comprises multivariate time-series observations with heterogeneous temporal intervals: hourly rainfall (R_t), 15-minute discharge (Q_t), and 10-minute water level measurements (H_t). These variables originate from diverse measurement systems and require uniform sampling rates for effective temporal modeling. Rainfall data were obtained from the Bojonegoro, Pacal Dam, and Padangan Rainfall Posts, representing the city, upstream, and middle reaches of the Bengawan Solo Watershed. Water discharge and water level data were collected from the Malo Bridge and Sumberejo water Level Posts, the primary monitoring points in the river segment around Bojonegoro. We chose these monitoring posts because the locations represent the contribution of local rainfall and flooding from upstream. This study employed rainfall, discharge, and water level data in combination to analyze the relationship between rainfall intensity, increased flow volume, and potential overflow in the flood-prone areas of Bojonegoro. Table 1 presents the hydrometeorological monitoring posts, including rainfall, water discharge, and water level stations, that are relevant for analyzing flood risk in Bojonegoro and Tuban, East Java, Indonesia.

Table 1. Monitoring posts for rainfall, water discharge and water level related to flood risk

Variables	Monitoring posts	Role in flood prediction	Key rationale
Rainfall	Pacal Dam Rainfall Post	Upstream primary predictor	Provides the earliest signal of heavy upstream rainfall that drives discharge toward Bojonegoro (longest lead time).
Water discharge	Karangnongko Water Discharge Post	Early upstream discharge predictor	Indicates inflow from the upstream area before reaching Bojonegoro, critical for early warning.
Water level	Karangnongko Water Level Post	Early upstream water-level signal	Detects initial rise in water level upstream, which can propagate downstream and affect Bojonegoro.
Water discharge	Malo Bridge Water Discharge Post	Confirmation of inflow to the city	Represents the volume of flow that directly affects Bojonegoro City.
Rainfall	Bojonegoro Rainfall Post	Local rainfall predictor	Captures intense local rainfall that may trigger urban/flash flooding.
Water level	Sumberejo Water Level Post	Target/flood trigger	Direct measurement of potential river overflow; used as the final indicator of flooding.

As a first step in the development of a flood prediction system, rainfall and water level data obtained from environmental sensors often experience irregularities, noise, or missing data. To overcome this, the CSI approach is used, which is a numerical interpolation method for reconstructing continuous data from discrete observations. The cubic spline function is arranged in the form of a third-degree polynomial segment for each interval between two consecutive data points, as shown in (1). To harmonize the time scales, CSI was utilized. This technique constructs a piecewise third-degree polynomial between data points, ensuring continuity in the first and second derivatives across intervals. Let the observed data points be defined as $(x_o, y_0).(x_1, y_1), \dots, (x_n, y_n)$, where x_i denotes the time and y_i the observed value (e.g., rainfall or discharge). The goal is to estimate intermediate values $y(x)$ at arbitrary points $x \in [x_i, x_{i+1}]$.

$$S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3 \quad (1)$$

The CSI function $S_i(x)$, shown in (1) is used to generate a smooth curve based on available rainfall or water level data. Each polynomial segment constructed in the CSI process ensures continuity and smoothness in the overall hydrological time series. In flood prediction modeling, abrupt variations in rainfall or water level data can significantly compromise model accuracy. As represented in (2), the spline function is required to pass through all original data points, preserving the integrity of the temporal hydrological dynamics.

$$S_i(x) = y_i, \quad S_i(x_{i+1}) = y_{i+1} \quad (2)$$

The CSI applies continuity constraints at the segment connection points of each spline segment, maintaining smoothness throughout the interpolated hydrological time series. This continuity constraints reinforce the model's suitability for reliable flood prediction as expressed in (3) and (4):

$$S_i'(x_{i+1}) = S_{i+1}'(x_{i+1}) \quad (3)$$

$$S_i''((x_{i+1}) = S_{i+1}''(x_{i+1}) \quad (4)$$

The spline condition ensures interpolation stability at the edges and prevents unrealistic extrapolation in rainfall or water level modeling, as represented in (5):

$$S_0''(x_0) = 0, \quad S_{n-1}''(x_n)=0 \quad (5)$$

The construction of the spline requires establishing the interval between successive time points. The width of interval is critical in determining the cubic spline coefficients, which ensures the accuracy of temporal interpolation for hydrological variables such as rainfall and water level. The interval width h_i between each pair of data points is mathematically defined in (6):

$$h_i = x_{i+1} - x_i \quad (6)$$

CSI provides a mathematical approach to capture nonlinear variations in rainfall and water level data by estimating intermediate values [29]–[31]. The spline coefficients, preserve the continuity and smoothness of the curve throughout each sub interval. The coefficient a_i, representing the function value at the beginning of an interval, is defined in (7):

$$a_i = y_i \quad (7)$$

where y_i is the actual data value at position x_i.

The coefficient b_i represents the local gradient of the interpolation function. The coefficient is determined by considering the variation in function values across points, as well as the influence of the adjacent spline curves, as described in (8):

$$b_i = \frac{y_{i+1} - y_i}{h_i} - \frac{h_i}{3} - (2c_i + c_{i+1}) \quad (8)$$

The coefficient d_i represents the curvature or the cubic contribution of the spline segment and is essential for shaping the smoothness of the interpolated curve. It is computed using the difference in second derivatives across consecutive intervals, as shown in (9):

$$d_i = \frac{c_{i+1} - c_i}{3h_i} \quad (9)$$

After the interpolated data is obtained and recovered from missing values and noise, the next step is to normalize hydrological features such as rainfall, water level, and discharge. This normalization is important so that the deep learning algorithm is not biased towards features with large scales. The normalization method used is min-max scaling, as formulated in (10):

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (10)$$

This normalization produces features in the range [0, 1] which are then used as input in a hybrid deep learning architecture based on LSTM-CNN, to capture temporal patterns and long-term dependencies of hydrological data that have been smoothed through spline interpolation.

2.2. Model design and development

The hybrid flood prediction model in this study combines LSTM and CNN to effectively extract both spatial and temporal features from multivariate hydrological time-series data. The integration of CNN for local feature learning and LSTM for temporal sequence modeling produces a robust prediction model capable of capturing complex spatio-temporal patterns.

2.2.1. Long short-term memory architecture design

A hybrid LSTM–CNN model is employed for spatio-temporal feature extraction. Initially, sequential patterns from the rainfall, discharge, and water level data are captured by an LSTM network, which models long-term temporal dependencies. The features extracted by the LSTM layer are then processed by a CNN module, enabling the model to learn spatial representations from the sequential data, which identifies localized spatial features and short-term fluctuations. This hybrid architecture ensures that both global temporal trends and local spatial variations are effectively learned.

To effectively represent temporal dynamics in the hydrological data, an LSTM network is applied to model sequential dependencies, such as rainfall, discharge, and water level. Through its memory and gated mechanisms, the LSTM network effectively models long-term patterns and dependencies that play a pivotal role in flood dynamics. The mathematical formulation of an LSTM cell at time step t is described as follows:

- Forget gate

The forget gate decides which past information from the cell state should be discarded. This mechanism ensures that only relevant hydrological features (e.g., rainfall patterns or sudden discharge changes) are retained for prediction, as represented in (11):

$$f_t = \sigma(W_f . [h_{t-1}, x_t] + b_f \quad (11)$$

Where:
f_t: forget gate vector (same size as the cell state)
σ: sigmoid activation function, output in range [0, 1]
h_{t-1}: hidden state from the previous time step
x_t: input at the current time step (e.g., rainfall or water level at time t)
W_f, b_f: weight matrix and bias for forget gate

- Input gate

The input gate regulates what new information is added to the cell state, allowing the model to incorporate new events such as sudden rainfall spikes or water surges. The input gate is represented as (12):

$$i_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \quad (12)$$

Where:
i_t: input gate vector controlling update strength.
W_f, b_f: weight matrix and bias for input gate.
– Candidate cell state
In (13) calculates the candidate values to be potentially added to the cell state, representing new knowledge learned from the current input.

$$\tilde{C}_t = tanh(W_c.[h_{t-1}, x_t] + b_{fc} \quad (13)$$

Where:
$\tilde{C}_t$: candidate cell state.
$tanh$: hyperbolic tangent activation to regulate value range.
W_c, b_c: weight matrix and bias for cell state update.
– Memory cell state update
The cell state C_t is updated by blending the retained memory and new candidate values. In flood prediction, this reflects both past patterns and current anomalies in hydrological parameters. Memory cell state update is represented as (14):

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (14)$$

Where:
C_t: updated cell state.
C_{t-1}: previous cell state.
– Output gate
The output gate filters the current cell state to produce the hidden state h_t, which captures the relevant context used for the final output or further CNN processing.

$$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \quad (15)$$

$$h_t = o_t \times tanh(C_t) \quad (16)$$

Where:
o_t: output gate vector.
h_t: hidden state at time ttt, passed to the next LSTM step or CNN.
W_o, b_o: weight matrix and bias for output gate.
The LSTM-based model was implemented using the Keras library with TensorFlow backend. The architecture consisted of two stacked LSTM layers, each with 128 units and a rectified linear unit (ReLU) activation function, followed by a dropout layer with a rate of 0.2 to reduce overfitting. A dense output layer with linear activation was used to predict continuous water level values. The model was trained using the Adam optimizer with a learning rate of 0.001 for 100 epochs and a batch size of 64. The LSTM architecture was selected for its proven ability to capture and preserve long-term temporal dependencies in sequential datasets, thus improving prediction performance and robustness.

2.2.2. Convolutional neural network architecture design

The LSTM–CNN model architecture is designed to leverage the strengths of both recurrent and convolutional networks for flood prediction. The LSTM layers first process sequential hydrological data to capture long-term temporal dependencies and generate hidden representations. These LSTM outputs are then fed into the CNN layers, which apply 1D convolution to extract local temporal features. Table 2 summarizes the structure of the CNN component used to refine the LSTM-derived features.

Table 2. CNN architecture for flood prediction

Layer	Type	Output shape	Filters/units	Kernel size	Activation
Input	-	(None, timesteps, 1)	-	-	-
Convolutional layer 1	Conv1D	(None, t−2, 32)	32	3	ReLU
Convolutional layer 2	Conv1D	(None, t−4, 64)	64	3	ReLU
Convolutional layer 3	Conv1D	(None, t−6, 128)	128	3	ReLU
Fully connected	Dense	(None, units)	-	-	ReLU
Output	Dense	(None, 1)	-	-	Linear

To effectively process spatio-temporal patterns in time-series data such as rainfall, discharge, and water level, the CNN model applies a one-dimensional (1D) convolution operation. This operation allows the model to recognize local patterns by shifting the kernel along the time series. Mathematically, this process is described by (17):

$$F_t = \sum_{i=0}^{L-1} X_{t+1} \cdot K_i \quad (17)$$

Where F_t is the feature map at time step t, X is the input sequence (such as rainfall or water level), and K is the convolutional kernel of length L. This equation illustrates how CNN captures local temporal patterns using a sliding window mechanism. Each convolutional layer progressively extracts more abstract features, starting from simple spikes to complex rainfall and discharge combinations, with the ReLU activation enabling non-linear learning. The dense layers then convert these features into continuous flood risk predictions, making CNN valuable for early warning systems.

2.2.3. Implementation details and experiments setup

The hyperparameters of the deep learning models were determined through empirical tuning. The LSTM–CNN model was trained for 100 epochs using a batch size of 32 and the Adam optimizer with a learning rate of 0.001. Early stopping was applied to prevent overfitting by monitoring the validation loss. The same training configuration was used for the baseline LSTM and CNN models to ensure a fair comparison.

All experiments were implemented using Python with the TensorFlow and Keras deep learning libraries. Model training and evaluation were conducted on a workstation with GPU acceleration to improve computational efficiency. This setup ensured efficient training while maintaining reproducibility of the results.

2.2.4. Algorithm description

This subsection describes the algorithmic workflow of the proposed hybrid LSTM–CNN flood prediction method. CSI is employed to harmonize heterogeneous hydrological time-series data prior to deep learning-based feature extraction and temporal modeling. The complete sequence of processing steps is presented in Algorithm 1.

Algorithm 1. Hybrid LSTM–CNN flood prediction algorithm with CSI
Data: Multi-interval hydrological time-series data (rainfall, discharge, water level)
Result: Predicted flood water level, MAPE, RMSE, MAE, R²

```
X ← collect hydrological data from multiple monitoring stations
Xraw ← merge rainfall, discharge, and water level datasets
if time intervals of Xraw are inconsistent then
    Xc ← apply Cubic Spline Interpolation (Xraw)
else
    Xc ← Xraw
end if
Xn ← normalize Xc

Set hyperparameters:
    epochs ← E
    batch_size ← B
    learning_rate ← α
    dropout_rate ← d

for each time window t do
    Fc ← extract features using CNN layers from Xn(t)
    Ft ← model temporal dependencies using LSTM layers from Fc
    Apply dropout with rate d
end for

Y ← predict flood water level using fused CNN–LSTM features

Split dataset into training and testing sets
Train hybrid LSTM–CNN model using training data with (E, B, α)
```

26: Validate model using testing data
27:
28: Evaluate prediction performance using MAPE, RMSE, MAE, and R^2

The proposed method begins by collecting multi-source hydrological data and harmonizing heterogeneous time intervals using CSI. The normalized data are then processed through CNN layers to extract representative features, followed by LSTM layers to capture long-term temporal dependencies. A hybrid CNN–LSTM architecture is employed to generate flood level predictions based on fused spatio-temporal features. Finally, the model performance is evaluated using standard error metrics, including MAPE, RMSE, MAE, and R^2.

3. RESULTS AND DISCUSSION

The hybrid LSTM–CNN flood prediction model, enhanced with CSI, was tested through a series of experimental scenarios. These scenarios were designed to evaluate the model's performance across different configurations and data preprocessing conditions, allowing for a clear comparison of predictive accuracy and robustness. The following subsections present the results for each scenario, along with an in-depth discussion of their implications for real-world flood forecasting.

3.1. Scenario 1: baseline model performance

The baseline scenario was designed to establish a performance reference for subsequent experiments. In this stage, the standalone LSTM and CNN architectures were trained and tested using raw hydrological data without any interpolation or smoothing. As shown in Table 3, the LSTM model achieved a MAPE of 7.85% and RMSE of 0.152 m, outperforming the CNN model, which recorded a MAPE of 8.42% and RMSE of 0.167 m. These results indicate that LSTM is more effective at capturing temporal dependencies inherent in sequential hydrological datasets, while CNN’s reliance on local feature extraction may limit its ability to represent long-term patterns.

Table 3. Performance results of baseline flood prediction models

Model	MAPE (%)	RMSE (m)	MAE (m)	R^2
LSTM	7.85	0.152	0.098	0.921
CNN	8.42	0.167	0.105	0.908

Although the LSTM model showed better performance than CNN, both models exhibited moderate error rates, which could be attributed to noise, missing values, and heterogeneous temporal resolutions in the dataset. The R^2 values (0.921 for LSTM and 0.908 for CNN) suggest that neither model could fully capture the variance in observed water levels. These findings emphasize the need for more advanced modeling techniques and preprocessing strategies to improve forecasting accuracy.

3.2. Hybrid model without interpolation

In the second scenario, the LSTM–CNN hybrid architecture was applied to raw data without any interpolation. The purpose was to determine whether integrating temporal sequence modeling (via LSTM) with spatial pattern extraction (via CNN) could outperform individual models. As presented in Table 4, the hybrid model achieved a MAPE of 6.95% and RMSE of 0.138 m, showing a significant improvement over both baseline models. This performance gain demonstrates that the combination of LSTM and CNN enables better representation of both long-term dependencies and short-term variations in hydrological time series.

Table 4. Performance results of the hybrid LSTM–CNN flood prediction model without CSI

Model	MAPE (%)	RMSE (m)	MAE (m)	R^2
LSTM–CNN	6.95	0.138	0.09	0.936

The improvement over the standalone models is consistent across all metrics, with the hybrid approach delivering an R^2 of 0.936 compared to 0.921 for LSTM and 0.908 for CNN. This confirms that the hybrid architecture benefits from the complementary strengths of its components. However, residual prediction errors during high-variability periods indicate that irregular sampling and noise still affect model performance. Therefore, further enhancements, particularly in preprocessing, are expected to unlock the full potential of this architecture.

3.3. Effect of cubic spline interpolation

The third scenario introduced CSI as a preprocessing step before feeding the data into the hybrid LSTM–CNN model. This interpolation method was applied to address missing data points, reduce measurement noise, and harmonize temporal resolutions across rainfall, discharge, and water level datasets. As reported in Table 5, the proposed hybrid LSTM–CNN model enhanced with CSI consistently outperformed the baseline models across all evaluation metrics. This significant improvement underscores the importance of clean and temporally consistent data for deep learning model performance.

Table 5. Performance comparison of the hybrid LSTM–CNN+CSI model

Model	MAPE (%)	RMSE (m)	MAE (m)	R^2
LSTM–CNN+CSI	5.84	0.125	0.082	0.948

The observed performance gains can be attributed to the CSI's ability to preserve smooth transitions in hydrological variables while maintaining the integrity of their temporal dynamics. By eliminating abrupt jumps between measurements and filling in missing values, CSI effectively reduces the uncertainty in model inputs. This leads to more stable learning during training and reduces the risk of overfitting to noisy observations. The results highlight that preprocessing using interpolation is not just a data preparation step but a critical factor that directly influences model predictive capability, particularly in domains like hydrology where measurement inconsistencies are common.

3.4. Comparative analysis

To provide a comprehensive performance comparison, all three configurations—baseline models, hybrid without interpolation, and hybrid with CSI—were evaluated side-by-side. As shown in Table 6, the hybrid model with CSI achieved the lowest error rates (MAPE: 5.84%, RMSE: 0.125 m, and MAE: 0.082 m) and the highest R^2 value of 0.948. In contrast, the standalone CNN recorded the weakest performance across all metrics, reaffirming the benefits of temporal modelling in hydrological forecasting. This comparison clearly demonstrates the cumulative advantages of architectural integration and data preprocessing.

Table 6. Hybrid model performance with CSI

Scenario	MAPE (%)	RMSE (m)	MAE (m)	R^2
LSTM (baseline)	7.85	0.152	0.098	0.921
CNN (baseline)	8.42	0.167	0.105	0.908
Hybrid LSTM–CNN	6.95	0.138	0.09	0.936
Hybrid LSTM–CNN+CSI	5.84	0.125	0.082	0.948

From a practical perspective, these results imply that operational flood forecasting systems can benefit from both model design improvements and robust data preparation pipelines. The reduction in error from the baseline to the best-performing configuration translates to more accurate flood warnings and potentially more lead time for disaster response. The consistency of performance improvements across multiple metrics further reinforces the reliability of the hybrid LSTM–CNN+CSI configuration. Such findings have implications not only for flood forecasting but also for other spatio-temporal prediction tasks in environmental and climate modelling.

Figure 2 illustrates the prediction error differences (predicted minus observed water level) for the four evaluated models. The results reveal a clear pattern in which the error distribution becomes progressively narrower as the model architecture becomes more integrated. The standalone CNN and LSTM models show larger fluctuations around zero, indicating higher prediction variance and less stable forecasting behavior during dynamic hydrological conditions. The hybrid LSTM–CNN model reduces this variability, demonstrating that combining temporal sequence modeling with local feature extraction improves prediction stability. The most consistent pattern is observed in the LSTM–CNN+CSI model, where error values remain tightly concentrated around zero across the time series. This indicates that the interpolation-based preprocessing effectively harmonizes multi-interval data and reduces noise in the input signals. The reduced error dispersion confirms that the proposed model provides more reliable flood level predictions. These findings support the research objective of improving prediction stability and accuracy when handling heterogeneous hydrological datasets. Overall, this comparison highlights that integrating CSI into the LSTM–CNN framework effectively reduces prediction differences and improves model accuracy.

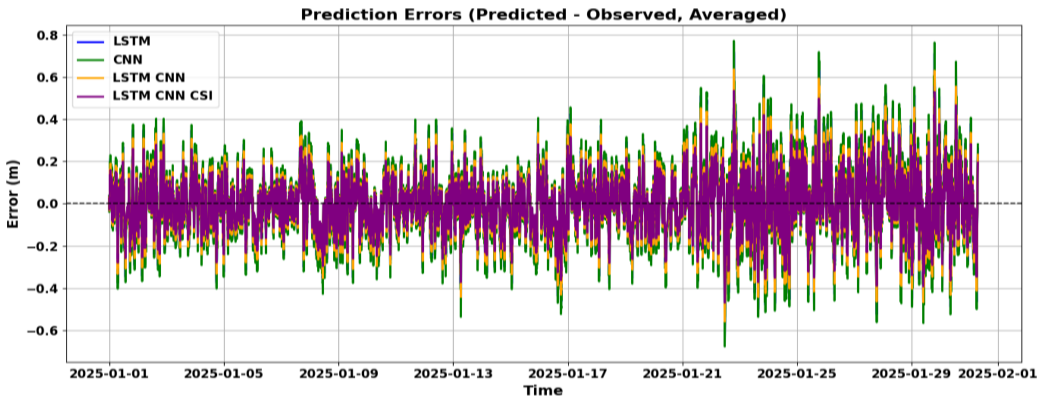


Figure 2. Time-series comparison of prediction errors (predicted minus observed water level) for four models: LSTM, CNN, LSTM–CNN, and LSTM–CNN with CSI

Figure 3 presents scatter plots comparing observed and predicted water levels for the four models. A clear improvement trend can be observed from the standalone models to the proposed hybrid approach. The CNN and LSTM models exhibit wider dispersion around the one-to-one reference line, indicating prediction bias and reduced accuracy, particularly at higher water levels where flood events are more critical. The hybrid LSTM–CNN model improves alignment with the ideal diagonal line, suggesting that the integration of temporal and spatial feature extraction enhances the model's ability to capture complex hydrological relationships. The tightest clustering around the diagonal line is achieved by the LSTM–CNN+CSI model. This pattern demonstrates that interpolated and harmonized data enable the model to learn more representative spatio-temporal patterns. Consequently, prediction errors decrease significantly across the full range of observed water levels. This result highlights the effectiveness of combining hybrid deep learning with data interpolation in improving flood prediction reliability.

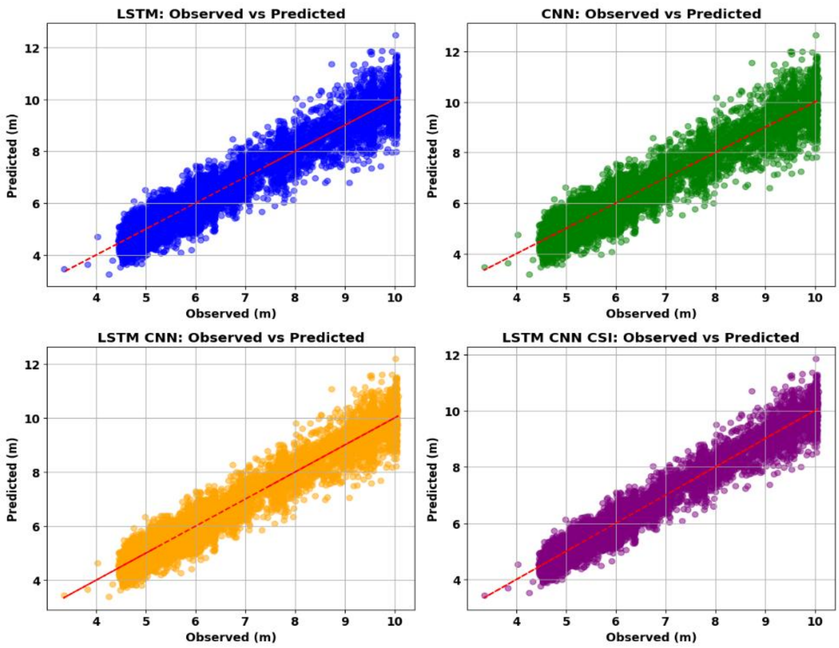


Figure 3. Scatter plots comparing observed and predicted water levels for four models: LSTM, CNN, LSTM–CNN, and LSTM–CNN with CSI

Figure 4 compares the predictive performance of four model configurations using MAPE, RMSE, MAE, and R^2 metrics. A consistent performance improvement pattern is observed as the modeling approach evolves from standalone models to the hybrid architecture with interpolation. The CNN model records the highest prediction errors, reflecting its limitations in capturing long-term temporal dependencies. The LSTM model performs better due to its ability to learn sequential patterns in hydrological data. However, the hybrid LSTM–CNN model further reduces prediction errors, indicating that combining temporal modeling with convolutional feature extraction improves representation of complex flood dynamics. The best performance is achieved by the hybrid LSTM–CNN model with CSI preprocessing, which shows the lowest error values and the highest R^2 score. This improvement demonstrates that harmonizing heterogeneous hydrological data significantly enhances the effectiveness of deep learning models. These results confirm that the integration of hybrid architectures and interpolation-based preprocessing provides a more robust framework for flood prediction systems.

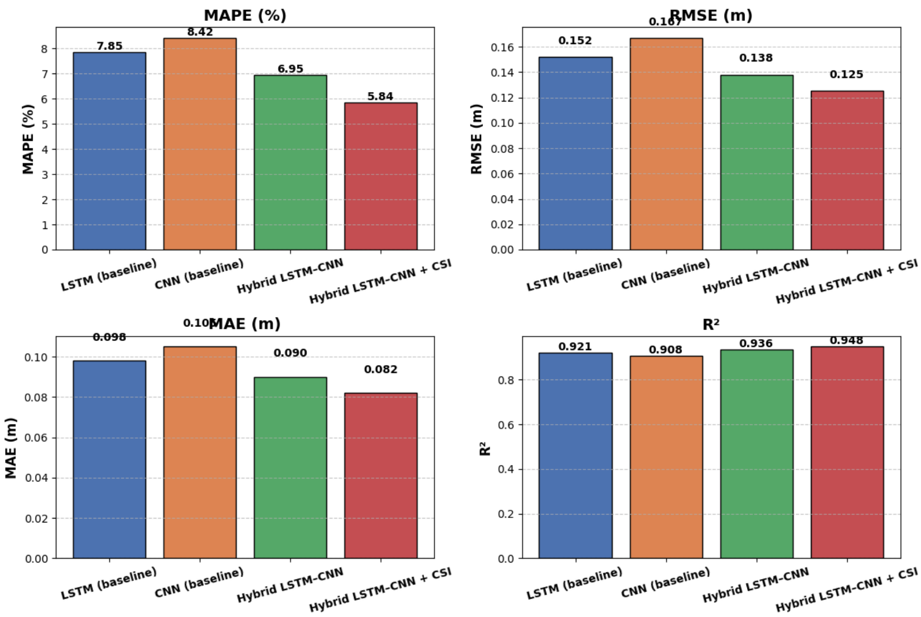


Figure 4. Performance comparison of baseline and hybrid models using MAPE, RMSE, MAE, and R^2

The use of CSI has been shown to significantly improve prediction accuracy compared to other interpolation methods. Compared to linear interpolation, CSI produces smoother and more realistic transitions because it maintains continuity up to the second derivative, allowing for representation of water level rise and fall patterns without sharp distortion. Unlike high-order polynomial interpolation, which often results in extreme oscillations (Runge's phenomenon), CSI utilizes a piecewise polynomial approach that is stable over non-uniform measurement intervals, making it more suitable for hydrological data. Furthermore, compared to simple imputations such as the mean or median, which ignore temporal dynamics, CSI maintains the shape of the temporal curve, allowing for representation of discharge spikes and water level rises. Empirically, the application of CSI in this study reduced prediction error (MAPE from 6.95% to 5.84%) and increased R^2 (from 0.936 to 0.948). This demonstrates that CSI is not simply a data cleaning technique, but rather a strategic component that strengthens the ability of deep learning models to more accurately capture hydrological temporal-spatial patterns.

3.5. Discussion

The comparative studies summarized in Table 7 illustrate the advancement of deep learning methods for flood forecasting across different spatial and temporal contexts. Dehghani *et al.* [32] comparatively evaluated LSTM, CNN, and ConvLSTM models for hourly short-term streamflow forecasting in Malaysia using rainfall, discharge, and streamflow data. Their results showed that ConvLSTM achieved the best performance (RMSE≈0.21, R^2≈0.95), highlighting the importance of spatio-temporal deep learning architectures for real-time hydrological forecasting. Building upon this direction, Wang *et al.* [33] proposed a dual-stage attention-based LSTM (encoder–decoder) model for multi-step flood forecasting in China (6–24 hours ahead), demonstrating that attention mechanisms significantly improve forecasting reliability (R^2 up to 0.92) by emphasizing the most relevant temporal features.

At the daily forecasting scale, Cho *et al.* [34] integrated rainfall, discharge, water level, temperature, and humidity variables within a hybrid LSTM–GRU framework in Korea, achieving strong predictive performance (NSE 0.942, MAE 2.22), thereby confirming the benefits of combining multiple meteorological inputs with recurrent architectures. Extending to long-term forecasting, Kordani *et al.* [35] proposed an ensemble E-LSTM–GRU model enhanced with an attention mechanism for monthly and seasonal prediction in Illinois, USA, achieving a high R^2 of approximately 0.98 and demonstrating the effectiveness of ensemble-based deep learning strategies in capturing complex hydrological dynamics.

Furthermore, Liu *et al.* [36] combined physically based hydrological modeling with LSTM networks in a hybrid framework applied to mixed catchments, showing that integrating process-driven models with data-driven approaches improves NSE and reduces RMSE, particularly when handling interpolated daily-to-monthly datasets.

Table 7. Comparative studies on flood forecasting using deep learning models

Study/year	Model	Input data	Domain/case study (interval)	Key metrics	Results (key values)	Main findings
Dehghani *et al.* (2023) [32]	LSTM, CNN, and ConvLSTM	Rainfall, discharge, streamflow	Malaysia – short-term forecasting (hourly)	RMSE, MAPE, R^2	RMSE≈0.21, R^2≈0.95 (best ConvLSTM)	Highlighted the importance of spatio-temporal deep learning approaches for real-time streamflow forecasting.
Wang *et al.* (2024) [33]	Dual-stage attention LSTM (encoder–decoder)	Rainfall, discharge, river flow	China – multi-step forecasting (6–24 hours ahead)	RMSE, R^2	RMSE ↓ 0.35, R^2 ↑ 0.92	Showed that attention mechanisms significantly improve the reliability of multi-step flood forecasting.
Cho *et al.* (2022) [34]	Hybrid LSTM–GRU	Rainfall, discharge, water level, temp, humidity	Korea (Yeojubo) – daily prediction	MSE, NSE, MAE	MSE 3.92, NSE 0.942, MAE 2.22	Combining meteorological features with a hybrid LSTM–GRU model yielded superior performance across multiple metrics.
Kordani *et al.* (2024) [35]	Ensemble E-LSTM–GRU+attention	Rainfall, discharge, hydro-met data	Illinois, USA – long-term (monthly/seasonal)	R^2	R^2≈0.98	The ensemble framework achieved exceptionally high accuracy for long-term flood forecasting.
Liu *et al.* (2024) [36]	Hybrid physical-based+LSTM	Rainfall, streamflow (observed & simulated)	Mixed catchments – daily to monthly (interpolated)	NSE, RMSE	Improved NSE and reduced RMSE	Integration of physical models with LSTM improved predictive performance and effectively handled interpolated datasets.
This study (2025)	Hybrid LSTM–CNN+CSI	Hourly rainfall (R_t), 15-min discharge (Q_t), 10-min water level (H_t)	Bojonegoro, East Java – real-time, multi-interval forecasting	RMSE, MAPE, MAE, R^2	MAPE 5.84%, RMSE 0.125, MAE 0.082, R^2 0.948	The hybrid model with CSI successfully captured multi-interval data, enhancing high-frequency flood forecasting accuracy in Bojonegoro.

The superior performance of the proposed hybrid LSTM–CNN model enhanced with CSI can be attributed to the complementary strengths of its architectural components and data preprocessing strategy. The LSTM layer effectively captures long-term temporal dependencies in hydrological time series, such as delayed river discharge responses to upstream rainfall, while the CNN component extracts localized temporal patterns and short-term fluctuations that are often associated with sudden water level rises. The integration of CSI further enhances model performance by harmonizing multi-interval hydrological data, reducing noise, and recovering missing observations. This combination allows the model to learn from smoother and temporally consistent inputs, leading to more stable training and improved generalization, particularly during high-variability flood events.

From a computational perspective, the standalone CNN model exhibits the lowest training cost due to its limited sequential processing, whereas the LSTM-based models require higher computational resources because of their recurrent structure and dependence on time-step-wise processing. The hybrid LSTM–CNN architecture introduces additional computational overhead compared to individual models; however, this increase remains manageable within the experimental setup and is justified by the substantial improvement in predictive accuracy. Moreover, the use of CSI as a preprocessing step does not significantly increase runtime during model training and inference, as interpolation is performed offline prior to learning. Overall, the proposed hybrid model achieves a favorable trade-off between computational complexity and forecasting accuracy, making it suitable for practical flood early warning applications.

In line with these advancements, this study (2025) proposes a hybrid LSTM–CNN model enhanced with CSI for real-time, multi-interval flood forecasting in Bojonegoro, East Java. By harmonizing heterogeneous temporal inputs—hourly rainfall (R_t), 15-minute discharge (Q_t), and 10-minute water level

(H_t)—the proposed approach achieved MAPE 5.84%, RMSE 0.125 m, MAE 0.082 m, and R^2 0.948, demonstrating its effectiveness in capturing high-frequency hydrological dynamics while maintaining strong predictive accuracy. Compared with prior studies, this work not only confirms the effectiveness of hybrid deep learning models but also highlights the importance of interpolation as a preprocessing step, bridging the gap between irregular monitoring intervals and the demands of real-time flood prediction.

4. CONCLUSION

This study demonstrates that integrating a hybrid LSTM–CNN architecture with CSI significantly improves flood prediction accuracy when handling heterogeneous, multi-interval hydrological data. The proposed approach effectively captures long-term temporal dependencies and localized hydrological patterns while enhancing data continuity through interpolation, achieving the lowest prediction errors among all evaluated models (MAPE 5.84%, RMSE 0.125 m, MAE 0.082 m, and R^2 0.948). Despite these promising results, this study is limited to a single watershed and relies on a finite number of monitoring stations, which may affect model generalizability under different hydrological and climatic conditions. Future research will focus on extending the model to multiple river basins, incorporating additional meteorological variables and real-time IoT sensor data, and exploring advanced deep learning mechanisms such as attention-based architectures to further enhance prediction robustness and scalability.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Balai Besar Wilayah Sungai Bengawan Solo (Bengawan Solo River Basin Authority) for providing the hydrological and meteorological data used in this study and Badan Penanggulangan Bencana Daerah (Regional Disaster Management Agency) of Bojonegoro for their valuable guidance and information.

FUNDING INFORMATION

This research was funded by the Vice Rector for Research, Innovation, and Entrepreneurship, Satya Wacana Christian University (Contract No. 146/SPK-PF/RIK/08/2025).

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Indrastanti Ratna Widiasari	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Eko Sediyono		✓				✓		✓	✓	✓	✓	✓		
Rissal Efendi	✓		✓	✓			✓			✓	✓		✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this manuscript

DATA AVAILABILITY

Water level, rainfall, and discharge data were obtained from the Bengawan Solo River Basin Authority (BBWS Bengawan Solo) via its Hydrology portal (https://hidrologi.bbws-bsolo.net/) for the period 1 January–30 April 2025. Data are available on the portal subject to its terms of use.

REFERENCES

[1] H. P. Rahayu, K. I. Zulfa, D. Nurhasanah, R. Haigh, D. Amaratunga, and I. I. Wahdiny, "Unveiling transboundary challenges in river flood risk management: learning from the Ciliwung River basin," *Natural Hazards and Earth System Sciences*, vol. 24, no. 6, pp. 2045–2064, 2024, doi: 10.5194/nhess-24-2045-2024.

[2] A. Sigit and M. Harada, "River-Basin Classification for Flood Risk Assessment in Indonesia," in *Geo-Sustainnovation for Resilient Society*, Springer Nature Singapore, 2024, pp. 309–319, doi: 10.1007/978-981-99-9219-5_29.

[3] R. Ramadhan, M. Marzuki, W. Suryanto, S. Sholihun, H. Yusnaini, and R. Muharsyah, "Rainfall variability in Indonesia new capital associated with the Madden-Julian Oscillation and its contribution to flood events," *Quaternary Science Advances*, vol. 13, p. 100163, Jan. 2024, doi: 10.1016/j.qsa.2024.100163.

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

[5] T. Niedzielski and M. Halicki, "Improving Linear Interpolation of Missing Hydrological Data by Applying Integrated Autoregressive Models," *Water Resources Management*, vol. 37, no. 14, pp. 5707–5724, Oct. 2023, doi: 10.1007/s11269-023-03625-7.

[6] A. Akkala, S. F. Boubrahimi, S. M. Hamdi, P. Hosseinzadeh, and A. Nassar, "Spatio-Temporal Graph Neural Networks for Streamflow Prediction in the Upper Colorado Basin," *Hydrology*, vol. 12, no. 3, p. 60, Mar. 2025, doi: 10.3390/hydrology12030060.

[7] S. Zhu, J. Wei, H. Zhang, Y. Xu, and H. Qin, "Spatiotemporal deep learning rainfall-runoff forecasting combined with remote sensing precipitation products in large scale basins," *Journal of Hydrology*, vol. 616, p. 128727, Jan. 2023, doi: 10.1016/j.jhydrol.2022.128727.

[8] M. Jahanbani, M. H. Vahidnia, H. Aghamohammadi, and Z. Azizi, "Flood susceptibility mapping through geoinformatics and ensemble learning methods, with an emphasis on the AdaBoost-Decision Tree algorithm, in Mazandaran, Iran," *Earth Science Informatics*, vol. 17, no. 2, pp. 1433–1457, Jan. 2024, doi: 10.1007/s12145-023-01213-2.

[9] A. R. M. T. Islam *et al.*, "Flood susceptibility modelling using advanced ensemble machine learning models," *Geoscience Frontiers*, vol. 12, no. 3, p. 101075, May 2021, doi: 10.1016/j.gsf.2020.09.006.

[10] A. Habibi, M. R. Delavar, B. Nazari, S. Pirasteh, and M. S. Sadeghian, "A novel approach for flood hazard assessment using hybridized ensemble models and feature selection algorithms," *International Journal of Applied Earth Observation and Geoinformation*, vol. 122, p. 103443, Aug. 2023, doi: 10.1016/j.jag.2023.103443.

[11] J. Zhou, L. Chen, T. Hu, H. Lu, Y. Shi, and L. Chen, "The comparative study of machine learning agent models in flood forecasting for tidal river reaches," *Scientific Reports*, vol. 15, no. 1, May 2025, doi: 10.1038/s41598-025-04633-w.

[12] V. Kumar, K. Sharma, T. Caloiero, D. Mehta, and K. Singh, "Comprehensive Overview of Flood Modeling Approaches: A Review of Recent Advances," *Hydrology*, vol. 10, no. 7, p. 141, 2023, doi: 10.3390/hydrology10070141.

[13] V. Kumar, H. M. Azamathulla, K. V. Sharma, D. J. Mehta, and K. T. Maharaj, "The State of the Art in Deep Learning Applications, Challenges, and Future Prospects: A Comprehensive Review of Flood Forecasting and Management," *Sustainability*, vol. 15, no. 13, p. 10543, 2023, doi: 10.3390/su151310543.

[14] A. Halbouni, T. S. Gunawan, M. H. Habaebi, M. Halbouni, M. Kartiwi, and R. Ahmad, "CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System," *IEEE Access*, vol. 10, pp. 99837–99849, 2022, doi: 10.1109/access.2022.3206425.

[15] A. Wan, Q. Chang, K. AL-Bukhaiti, and J. He, "Short-term power load forecasting for combined heat and power using CNN-LSTM enhanced by attention mechanism," *Energy*, vol. 282, p. 128274, Nov. 2023, doi: 10.1016/j.energy.2023.128274.

[16] A. Agga, A. Abbou, M. Labbadi, Y. El Houm, and I. H. Ou Ali, "CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production," *Electric Power Systems Research*, vol. 208, p. 107908, 2022, doi: 10.1016/j.epsr.2022.107908.

[17] E. Heidari, V. Samadi, and A. A. Khan, "Leveraging Recurrent Neural Networks for Flood Prediction and Assessment," *Hydrology*, vol. 12, no. 4, p. 90, Apr. 2025, doi: 10.3390/hydrology12040090.

[18] X. Pan *et al.*, "LSTM Model-Based Rapid Prediction Method of Urban Inundation with Rainfall Time Series," *Water Resources Management*, vol. 39, no. 2, pp. 661–688, Oct. 2024, doi: 10.1007/s11269-024-03972-z.

[19] H. Malik, J. Feng, P. Shao, and Z. A. Abduljabbar, "Improving flood forecasting using time-distributed CNN-LSTM model: a time-distributed spatiotemporal method," *Earth Science Informatics*, vol. 17, no. 4, pp. 3455–3474, 2024, doi: 10.1007/s12145-024-01354-y.

[20] L. Zhong, R. Liu, X. Miao, Y. Chen, S. Li, and H. Ji, "Compressor Performance Prediction Based on the Interpolation Method and Support Vector Machine," *Aerospace*, vol. 10, no. 6, p. 558, 2023, doi: 10.3390/aerospace10060558.

[21] P. Amirzehni, S. Samadianfard, A. Nazemi, and A. Sadraddini, "Evaluating capabilities of the spline and cubic spline interpolation functions in reference evapotranspiration estimation implementing satellite image data," *Earth Science Informatics*, vol. 16, no. 4, pp. 3779–3795, Oct. 2023, doi: 10.1007/s12145-023-01127-z.

[22] C. Skoulikaris, P. Venetsanou, G. Lazoglou, C. Anagnostopoulou, and K. Voudouris, "Spatio-Temporal Interpolation and Bias Correction Ordering Analysis for Hydrological Simulations: An Assessment on a Mountainous River Basin," *Water*, vol. 14, no. 4, p. 660, Feb. 2022, doi: 10.3390/w14040660.

[23] E. Abdi, M. Ali, C. A. G. Santos, A. Olusola, and M. A. Ghorbani, "Enhancing groundwater level prediction accuracy using interpolation techniques in deep learning models," *Groundwater for Sustainable Development*, vol. 26, p. 101213, Aug. 2024, doi: 10.1016/j.gsd.2024.101213.

[24] P. Li, J. Zhang, and P. Krebs, "Prediction of Flow Based on a CNN-LSTM Combined Deep Learning Approach," *Water*, vol. 14, no. 6, p. 993, Mar. 2022, doi: 10.3390/w14060993.

[25] I. Ebtehaj and H. Bonakdari, "CNN vs. LSTM: A Comparative Study of Hourly Precipitation Intensity Prediction as a Key Factor in Flood Forecasting Frameworks," *Atmosphere*, vol. 15, no. 9, p. 1082, 2024, doi: 10.3390/atmos15091082.

[26] P. J. Depetris, "The Importance of Monitoring River Water Discharge," *Frontiers in Water*, vol. 3, Oct. 2021, doi: 10.3389/frwa.2021.745912.

[27] C. Prakash, A. Barthwal, and D. Acharya, "FLOODWALL: A Real-Time Flash Flood Monitoring and Forecasting System Using IoT," *IEEE Sensors Journal*, vol. 23, no. 1, pp. 787–799, Jan. 2023, doi: 10.1109/jsen.2022.3223671.

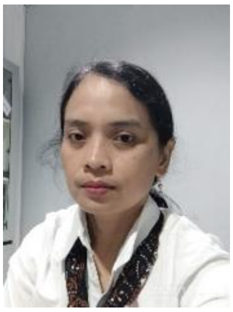
[28] W. Almikaeel, A. Šoltész, L. Čubanová, and D. Baroková, "Hydro-informer: a deep learning model for accurate water level and flood predictions," *Natural Hazards*, vol. 121, no. 4, pp. 3959–3979, Oct. 2024, doi: 10.1007/s11069-024-06949-8.

[29] M. Meggiorin, G. Passadore, S. Bertoldo, A. Sottani, and A. Rinaldo, "Comparison of Three Imputation Methods for Groundwater Level Timeseries," *Water*, vol. 15, no. 4, p. 801, Feb. 2023, doi: 10.3390/w15040801.

[30] M. R. Zaghiyan, S. Eslamian, A. Gohari, and M. S. Ebrahimi, "Temporal correction of irregular observed intervals of

groundwater level series using interpolation techniques," *Theoretical and Applied Climatology*, vol. 145, no. 3–4, pp. 1027–1037, 2021, doi: 10.1007/s00704-021-03666-1.
[31] W. Peng, D. Tao, Q. Ma, Q. Xie, and J. Wang, "Clipped seismic record recovery analysis based on the cubic spline interpolation algorithm," *Journal of Seismology*, vol. 28, no. 3, pp. 843–857, Apr. 2024, doi: 10.1007/s10950-024-10207-6.
[32] A. Dehghani *et al.*, "Comparative evaluation of LSTM, CNN, and ConvLSTM for hourly short-term streamflow forecasting using deep learning approaches," *Ecological Informatics*, vol. 75, p. 102119, 2023, doi: 10.1016/j.ecoinf.2023.102119.
[33] F. Wang, W. Wang, W. Bi, W. Lin, and D. Zhang, "Dual-Stage Attention-Based LSTM Network for Multiple Time Steps Flood Forecasting," *Proceedings of IAHS*, vol. 386, pp. 141–146, Apr. 2024, doi: 10.5194/piahs-386-141-2024.
[34] M. Cho, C. Kim, K. Jung, and H. Jung, "Water Level Prediction Model Applying a Long Short-Term Memory (LSTM)–Gated Recurrent Unit (GRU) Method for Flood Prediction," *Water*, vol. 14, no. 14, p. 2221, 2022, doi: 10.3390/w14142221.
[35] M. Kordani, M. R. Nikoo, M. Fooladi, I. Ahmadianfar, R. Nazari, and A. H. Gandomi, "Improving Long-Term Flood Forecasting Accuracy Using Ensemble Deep Learning Models and an Attention Mechanism," *Journal of Hydrologic Engineering*, vol. 29, no. 6, Dec. 2024, doi: 10.1061/jhyeff.heeng-6262.
[36] J. Liu, J. Koch, S. Stisen, L. Troldborg, and R. J. M. Schneider, "A national-scale hybrid model for enhanced streamflow estimation – consolidating a physically based hydrological model with long short-term memory (LSTM) networks," *Hydrology and Earth System Sciences*, vol. 28, no. 13, pp. 2871–2893, 2024, doi: 10.5194/hess-28-2871-2024.

BIOGRAPHIES OF AUTHORS



Indrastanti Ratna Widiasari was born in Sukoharjo, Central Java, Indonesia in 1978. She earned her Doctoral Degree in Electrical Engineering study program Universitas Gadjah Mada Yogyakarta in 2018. She is a lecturer in Informatics Engineering Satya Wacana Christian University. Her research interests are in the field of computer networks, wireless sensor networks, smart technology, deep learning, and pervasive computing. She can be contacted at email: indrastanti@uksw.edu.

Eko Sediyono is a Professor of Information Technology at Satya Wacana Christian University and currently serves as Vice Rector for Research, Innovation, and Entrepreneurship. Graduate Degree from Bogor Agricultural Institute and the University of Indonesia (Master and Doctoral Degree), his research focuses on spatial information systems, smart cities, and artificial intelligence. With a Google Scholar h-index of 16 and Scopus h-index of 5, he has published extensively and developed several software-related intellectual property rights. He can be contacted at email: eko@staff.uksw.edu or eko@uksw.edu.

Rissal Efendi earned his Master of Information System from Diponegoro University. Since 2018, he has been a lecturer in Education of Informatics Engineering and Computer Science Study Program, Satya Wacana Christian University. He has working experiences as a computer network engineer in some private companies. His research interests are in the field of computer networks, cyber security, machine learning, deep learning, and pervasive computing. Now he focusses on conducting IoT-based flood disaster prediction research in Semarang, Indonesia. He can be contacted at email: rissal.efendi@uksw.edu.