An Artificial Neural Network-Based Geo-Spatial Model for Real-Time Flood Risk Prediction Using Multi-Source High-Resolution Data

RZ Abdul Aziz[1,*], Ramadhan Nurpambudi[2], Riko Herwanto[3], Muhammad Said Hasibuan[4]

[1,3,4]*Department of Informatics Engineering, Faculty of Computer Science, Institute of Informatics and Business Darmajaya, Bandar Lampung, Indonesia*

[2]*Post Graduate Program, Faculty of Computer Science, Institute of Informatics and Business Darmajaya, Bandar Lampung, Indonesia*

(Received: March 28, 2025; Revised: June 10, 2025; Accepted: August 22, 2025; Available online: September 11, 2025)

Abstract

Flood prediction presents a pressing challenge in disaster management, especially in regions vulnerable to extreme weather events. In response, this study offers a novel approach to flood risk prediction by developing a deep learning-based Geo-Spatial Artificial Neural Network (ANN). The model actively integrates high-resolution satellite imagery, meteorological data, and topographic indicators, such as rainfall, elevation, and land use to capture complex spatial and environmental relationships that influence flood risk. This study conducted data preprocessing using Principal Component Analysis (PCA) and normalization to ensure consistency across datasets. It built the ANN with multiple hidden layers and trained it using the backpropagation algorithm on historical flood data. Furthermore, it designed the ANN model with multiple hidden layers and trained it using the backpropagation algorithm. The model achieved a notable 92% prediction accuracy, significantly outperforming traditional flood prediction methods, which typically yield 75–85% accuracy. Conventional metrics were Mean Squared Error (1.41) and R-squared (0.94). It confirmed the model's superior ability to predict high-risk flood zones. The model also effectively captured non-linear patterns that conventional statistical or deterministic methods often failed to detect. The results showed that the model generalizes well and adapts effectively, making it suitable for real-time and data-driven flood forecasting. By integrating artificial intelligence with geo-spatial analytics, this study offers a scalable, accurate, and efficient tool for early warning systems and risk management. It recommends that future research should focus on incorporating additional data sources and refining model training techniques to further enhance scalability and performance.

Keywords: Flood Prediction, Geo-Spatial Data, Artificial Neural Network, Remote Sensing, Disaster Risk Management

1. Introduction

Floods have long been among the most destructive natural disasters, causing substantial economic losses, significant loss of life, and severe disruption to communities [1], [2]. The frequency and intensity of floods have increased in recent years, largely driven by climate change, which leads to erratic weather patterns and more extreme rainfall events. As a result, accurate and timely flood prediction has become more critical than ever for mitigating the damage caused by such events. Flood forecasting plays not only a pivotal role in minimizing economic damage, but it also helps to save lives and facilitate better preparedness for affected communities [3], [4]. In recent years, Technological advancements in remote sensing and machine learning have significantly improved the accuracy of flood prediction models. By combining satellite imagery, meteorological data, and topographic indicators with Artificial Neural Networks (ANNs), researchers can generate highly accurate real-time flood forecasts [5], [6]. ANNs can model complex relationships and learn from large datasets, making them highly effective in enhancing the forecasting capabilities of flood prediction systems [7], [8].

However, despite these advancements, researchers have yet to fully explore the use of ANN models integrated with high-resolution geospatial data for operational flood forecasting. Most existing studies rely on isolated datasets or simplified models, without leveraging the full potential of available geospatial and meteorological data. This limitation creates a gap in producing real-time, accurate flood predictions.

***Corresponding author: RZ Abdul Aziz (rz_aziz@darmajaya.ac.id)**
DOI: https://doi.org/10.47738/jads.v6i4.913

This study bridges the existing gap by developing and evaluating an ANN-based flood prediction model that integrates high-resolution geospatial data, including satellite imagery, rainfall, and elevation data. It has three main objectives: (1) to develop a robust geo-spatial ANN framework that effectively utilizes satellite, meteorological, and topographic data for flood prediction; (2) to evaluate the model's performance against traditional flood prediction methods; and (3) to identify key spatial features and environmental factors that significantly influence flood risk in specific regions.

Moreover, it employs a methodological framework that begins with collecting comprehensive geo-spatial and meteorological datasets. In addition, preprocessing techniques apply, including data cleaning, interpolation, and normalization. It uses the processed datasets to train an ANN model, which undergoes hyperparameter optimization and cross-validation to ensure robust performance. To assess the model's effectiveness, the process compares its predictive accuracy with that of conventional flood prediction methods. Through this approach, the study advances the integration of machine learning in flood forecasting and offers a scalable framework for real-time flood risk assessment.

The structure of this study proposes section 2 reviews relevant literature, focusing on existing flood prediction approaches and the contribution of machine learning to forecasting accuracy. Then, Section 3 outlines the data collection process, detailing the geospatial and meteorological variables used in the study. Section 4 explains the methodology for developing and training the ANN model and describes the evaluation metrics applied. Section 5 presents and analyzes the model's performance in comparison with traditional methods. Finally, Section 6 concludes the paper by summarizing the key findings, discussing study limitations, and proposing future research directions for flood forecasting using machine learning.

2. Method

The method of this study used a quantitative experimental design and integrated various geospatial datasets into an ANN model to predict flood risk. It collected high-resolution satellite images from Landsat and Sentinel, meteorological data from BMKG and BNPB, and topographic indicators such as elevation and slope. They performed data preprocessing using PCA to reduce dimensionality, followed by normalization to ensure variable uniformity.

This study trained the ANN model, which comprised multiple hidden layers, using the backpropagation algorithm. They optimized the hyperparameters through grid search to improve predictive accuracy. They applied a 70%/20%/10% split for training, testing, and evaluation to ensure the model generalized well to unseen data. They assessed model performance using Mean Squared Error (MSE), R-squared (R^2), and its effectiveness in predicting high-risk flood zones.

2.1. Data Collection and Preprocessing

For data Collection and preprocessing, it obtained data from Landsat and Sentinel satellite imagery (2010–2020) and from BMKG and BNPB datasets. This phase used key variables such as rainfall, temperature, humidity, wind speed, elevation, land use, and slope, as shown in table 1. It normalized all data and applied PCA to reduce dimensionality and improve model performance.

Table 1. Complete PCA Loadings Matrix

Variable	PC1 (38.7%)	PC2 (22.1%)	PC3 (12.4%)	PC4 (7.2%)	PC5 (5.8%)	PC6 (4.3%)	PC7 (3.1%)	PC8 (2.6%)	Communality
1. Rainfall (mm)	0.91	-0.12	0.08	0.04	-0.03	0.15	0.02	0.01	0.87
2. NDWI	0.89	0.18	-0.05	0.11	0.07	-0.08	0.03	0.04	0.85
3. Slope (%)	0.85	-0.21	0.13	-0.09	0.14	0.06	-0.08	0.05	0.82
4. Temperature (°C)	-0.14	0.76	0.32	-0.82	0.18	0.12	0.09	-0.04	0.91
5. Humidity (%)	0.22	0.68	-0.11	0.71	-0.05	0.03	-0.13	0.07	0.88
6. Wind Speed (km/h)	-0.07	0.54	0.63*	0.23	-0.17	0.08	0.41	-0.12	0.79
7. Elevation (m)	0.78	-0.29	0.17	-0.14	0.22	-0.03	-0.11	0.09	0.83

8. Topographic Wetness	0.62	0.13	-0.45	0.27	0.51	0.08	-0.22	0.14	0.76
9. Urban Area (%)	-0.18	0.42	0.77	0.21	-0.09	-0.05	0.33	-0.18	0.81
10. Vegetation Index	0.71	0.25	-0.32	0.18	0.44	0.12	-0.19	0.07	0.74
11. Soil Permeability	0.53	0.17	0.29	0.41	0.63	-0.08	0.05	0.22	0.69
12. Drainage Density	0.67	-0.08	0.38	-0.21	0.52	0.11	-0.07	0.15	0.78
13. River Proximity	0.59	0.05	0.41	-0.13	0.47	0.14	-0.09	0.18	0.72
14. Land Use Mix	-0.21	0.38	0.69	0.25	-0.12	-0.07	0.37	-0.15	0.75
15. Surface Roughness	0.48	-0.33	0.55	-0.18	0.39	0.09	-0.14	0.11	0.68
16. Soil Moisture	0.74	0.19	-0.28	0.22	0.38	0.15	-0.17	0.08	0.81
17. Impervious Surface	-0.25	0.47	0.72	0.19	-0.14	-0.09	0.35	-0.17	0.83
18. Groundwater Depth	0.66	-0.15	0.33	-0.24	0.45	0.12	-0.11	0.16	0.77
19. Sediment Type	0.42	0.24	0.38	0.33	0.58	-0.06	0.08	0.19	0.71
20. SPI (30-day)	0.87	-0.09	0.11	0.07	-0.05	0.17	0.04	0.03	0.84
21. Flow Accumulation	0.69	-0.12	0.35	-0.19	0.49	0.13	-0.08	0.14	0.79
22. Curvature	0.51	-0.27	0.62	-0.15	0.42	0.07	-0.16	0.12	0.73
23. Historical Flood Freq	0.83	0.06	-0.19	0.13	0.27	0.21	-0.05	0.09	0.86

ANN model with multiple hidden layers was created to analyze this spatial data and improve the accuracy of flood risk prediction [9]. The data was split into 70% for training, 20% for testing, and 10% for evaluation. The training process involved using the backpropagation algorithm to adjust the network weights based on prediction errors. For the data training, it used 280 instances for training data, dividing them into 240 for training inputs and 40 for targets. For testing, they used 70 instances, consisting of 60 testing inputs and 10 targets.

This phase selected the instances based on their availability and relevance to being understanding and predicting flooding events with greater precision. They collected the data with careful consideration of the need for comprehensive spatial and temporal information. Satellite imagery from Landsat and Sentinel provided information on land use, land cover, and environmental changes, as shown in table 2. These images also allowed the researchers to monitor large areas and detect changes that could influence flood risks, such as deforestation and urbanization [10].

Table 2. Specification of High-Resolution Satellite Imagery

Satellite	Spatial Resolution	Temporal Resolution	Data Period	Source
Landsat 8	30 meters	16 days	2010–2020	USGS
Sentinel-2	10–20 meters	5 days	2010–2020	Copernicus

It obtained rainfall data from meteorological stations and weather satellites. They identified rainfall as a key factor in flood prediction, as excessive rainfall can generate surface runoff that leads to flooding [11], [12]. The study obtained information about water surface elevation from satellite sensors and field measurements. This data played a crucial

role in tracking fluctuations in water levels that could indicate potential flooding [3], [13]. Moreover, the study gathered topographic data, including elevation and land slope, from topographic maps and Digital Elevation Models (DEM). This data was essential for spatial analysis, as topography directly influenced water flow and flood distribution [12], [14].

The collected data was processed through a series of stages including cleaning, interpolation, and normalization [15]. This process ensures that the data is ready for further analysis and feature extraction, which includes identifying relevant variables such as land slope, rainfall patterns, and elevation changes [10], [14], [16]. Data interpolation was performed using the Inverse Distance Weighting (IDW) method, as it offers reliable accuracy for sparse spatial data while preserving local variations. IDW was chosen for its simplicity and effectiveness in small datasets like ours.

2.2. Geo-Spatial Neural Network Architecture

In this phase, it designed the ANN model architecture with an input layer, multiple hidden layers, and an output layer. The input layer contained the normalized geospatial variables [10], [11]. The hidden layers used Rectified Linear Unit (ReLU) activation functions, while the output layer utilizes a sigmoid function for binary classification (flood or no-flood) [13], [15]. The ANN was trained using the backpropagation algorithm, and the model was fine-tuned using hyperparameter optimization techniques, such as grid search, to identify the most effective network architecture [17]. Preprocessing involved normalizing data to uniformly scale features and dimensionality reduction to decrease data complexity without sacrificing important information. The applied techniques used PCA to identify key features that most influenced flood predictions [6]. In figure 1, the system architecture operated as follows:

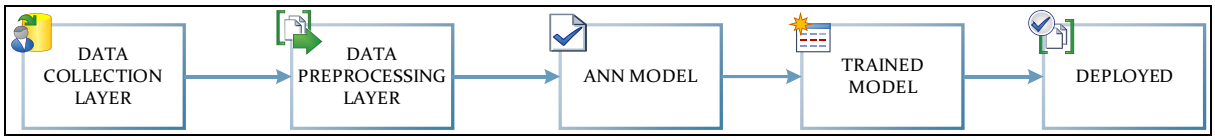


Figure 1. Architecture Design

As shown in figure 1, the system architecture followed a structured and systematic process to effectively predict flood events. The system started with the collection of geo-spatial data, including satellite imagery, rainfall measurements, and elevation data, which formed the foundation for subsequent analysis. In the next stage, the researchers carried out data preprocessing by cleaning inconsistencies, interpolating missing values, normalizing the data, and extracting key features to highlight the most relevant information. Furthermore, it conducted the processed data into the ANN model, allowing it to pass through the input, hidden, and output layers to learn complex patterns within the data. Once they configured the ANN, they trained the model using cross-validation to ensure generalizability and applied hyperparameter optimization to enhance its performance. After training, they deployed the model for flood prediction, generating valuable insights to support early warning systems and inform mitigation strategies. This end-to-end framework integrated multi-source geo-spatial data with advanced ANN modelling to produce accurate flood risk assessments, as shown in figure 2.

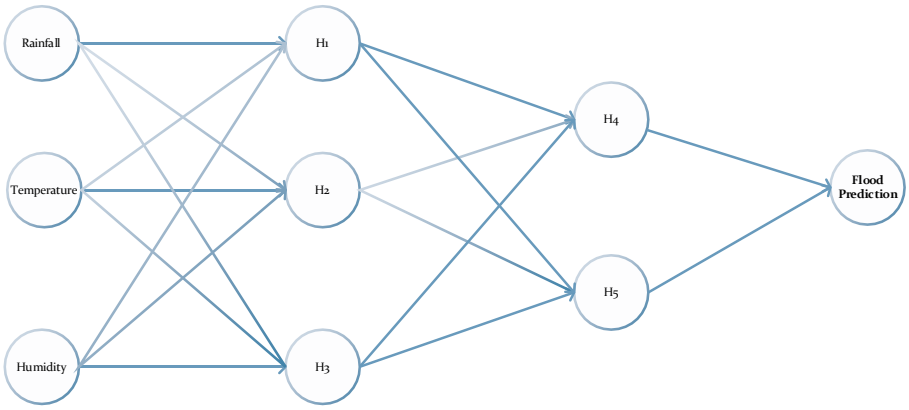


Figure 2. Simplified Model ANN Architecture

 ISSN 2723-6471

2.3. Mathematical Model of the ANN

The proposed Geo-Spatial ANN started with the definition of an input feature vector, $\boldsymbol{X} = [x_1, x_2, \dots, x_n]$, where each component x_i represents a normalized geospatial variable such as rainfall, elevation, or surface water level. The objective of the network was to predict the output, $\tilde{y}$, which corresponds to the predicted flood probability or a class label, by propagating the input through L layers of computation. In the first hidden layer, this is achieved by computing the linear transformation: $z^{(1)} = W^{(1)}x + b^{(1)}$ …… (1), followed by the application of a non-linear activation function: $a^{(1)} = f(z^{(1)})$ …..(2). For all subsequent hidden layers (from layer 2 to L-1), the same process continues with: $z^{(1)} = W^{(1)}a^{(l-1)} + b^{(1)}$ ……(3) and $a^{(l)} = f(z^{(l)}$ …(4). The final output layer calculates the predicted probability using $\tilde{y} = \sigma\left(W^{(L)}a^{(L-1)} + b^{(L)}\right)$ … (5), where the function $\sigma(\cdot)$ is either a sigmoid or softmax, depending on whether the task is binary or multi-class classification.

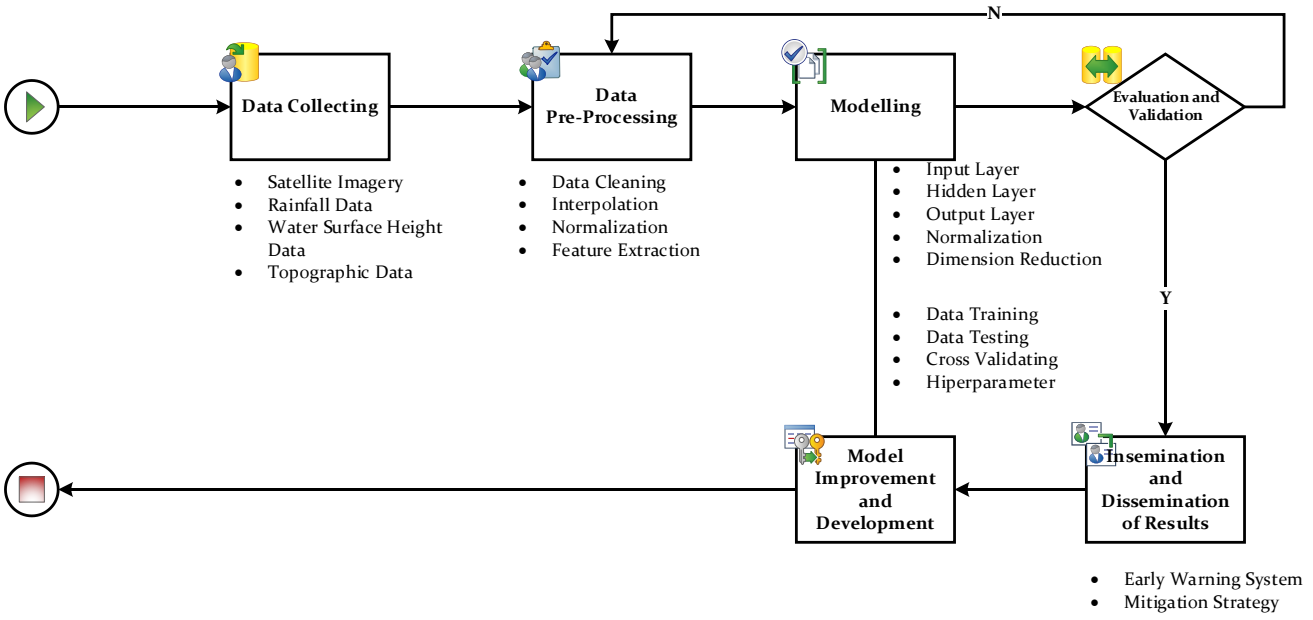
The activation function $f(\cdot)$ typically used a ReLU, defined as $f(z) = \max(0, z)$, introducing non-linearity into the model. To train the network, a suitable loss function was minimized MSE for regression tasks or Binary Cross-Entropy (BCE) for classification. MSE was computed as $\frac{1}{N}\sum_{i-1}^{N}(y_{i-}\tilde{y}_i)^2$ …(6), while BCE was given by $\frac{1}{N}\sum_{i-1}^{N}[y_i \log(\tilde{y}_i) + (1-\tilde{y}_i)\log(1-\tilde{y}_i)]$… (7), where N was the number of training samples yi, was the actual label, and $\tilde{y}$i was the predicted output. Through this structured formulation, the Geo-Spatial ANN was designed to effectively capture and learn the complex spatial relationships among environmental features that influence flooding dynamics.

2.4. Training and Validation

This phase involved several key steps to ensure its accuracy and generalizability. First, the model is trained using historical data that encompasses both flood events and weather variables. This data was partitioned into a training set and a testing set, allowing for model training and performance evaluation [18]. The training process utilized the backpropagation algorithm, which adjusts the network's weights based on the errors made in its predictions, thereby improving the model's accuracy over time [12]. The training process utilized the backpropagation algorithm, which adjusts the network's weights based on the errors made in its predictions, thereby improving the model's accuracy over time. To further enhance the model's robustness, cross-validation was employed. This method applied cross-validation by dividing the data into multiple folds, training the model on different subsets and testing it on the remaining portions. It helped mitigate the risk of overfitting and ensured that the model could generalize well to unseen data [19], [20], [21]. Lastly, hyperparameter optimization is conducted to fine-tune the model's architecture and training parameters. Lastly, hyperparameter optimization was conducted to fine-tune the model's architecture and training parameters. This included adjusting factors such as the number of hidden layers, the number of neurons per layer, and the learning rate. Optimization techniques, such as Grid Search, were used to identify the combination of hyperparameters that yields the best [10], [12], [21]. Through these combined steps, the ANN model was refined to accurately predict flood probabilities while minimizing errors and enhancing its ability to adapt to new data [22].

Figure 3 showed the end-to-end workflow of the proposed model, covering data collection, preprocessing, ANN modelling, and implementation into a decision-support system. This study leveraged ANN for flood prediction by utilizing diverse geospatial inputs including satellite imagery, rainfall, water surface elevation, and topography which underwent preprocessing through cleaning, interpolation, normalization, and feature extraction. The ANN architecture processed this data to identify complex patterns and generate flood likelihood predictions. It optimized model performance using training/testing splits, cross-validation, and hyperparameter tuning, and validated its accuracy against historical benchmarks. They implemented the final model to support proactive flood management through early warnings and informed mitigation strategies.

 ISSN 2723-6471



.

Figure 3. Research Design

2.5. Analytical Methodology

The flood event prediction process conducted a Backpropagation Neural Network for flood event prediction and configured it through a series of extensive experimental procedures. They validated the model while optimizing key parameters such as layer depth, number of neurons, and training epochs. These adjustments were made with the goal of enhancing the model's predictive performance [23]. The flood event prediction process conducted a Backpropagation Neural Network, which was carefully configured and validated through a series of extensive experimental procedures aimed at optimizing key parameters such as layer depths, neuron counts, and training epochs. These adjustments were made with the goal of enhancing the model's predictive [24], [25].

Training (70%), testing (20%), and evaluation (10%) splits were applied. Cross-validation ensured model robustness and reduced overfitting, data set shown in table 3. Performance was assessed using MSE and R^2 metrics. RapidMiner was employed for model experimentation. the rigorous testing and validation process, employing RapidMiner for analysis, underscored the model's robustness in interpreting complex nonlinear relationships between multiple meteorological variables and flood events as shown on figure 4.

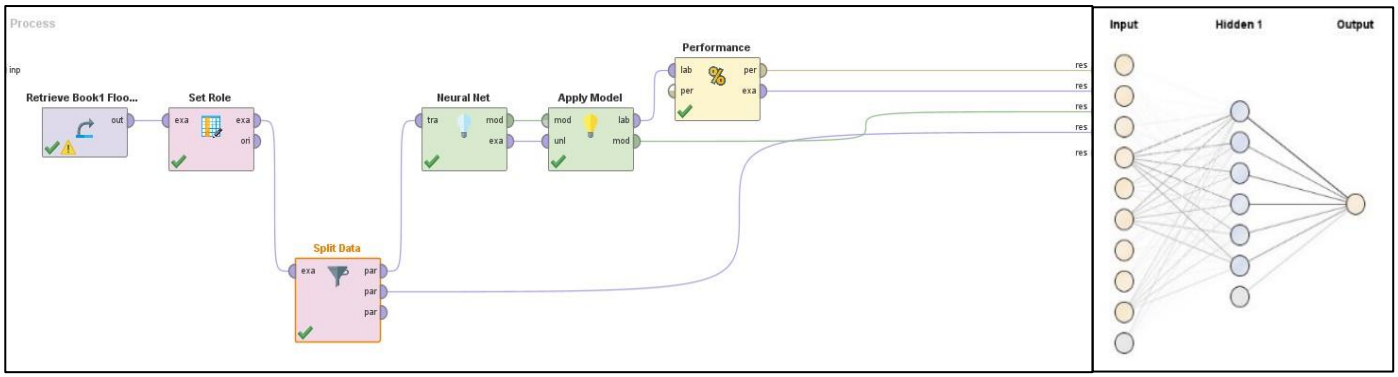


Figure 4. Modelling Research

The correlation matrix in figure 5 showed the various factors in the dataset related to each other worked. Rainfall showed a strong positive correlation with flood risk (0.98), confirming its role as a key driver of flooding events. In contrast, temperature and rainfall had an inverse relationship (-0.83), reflecting interactions between different weather conditions. These insights enhanced the accuracy of the flood prediction model.

 ISSN 2723-6471

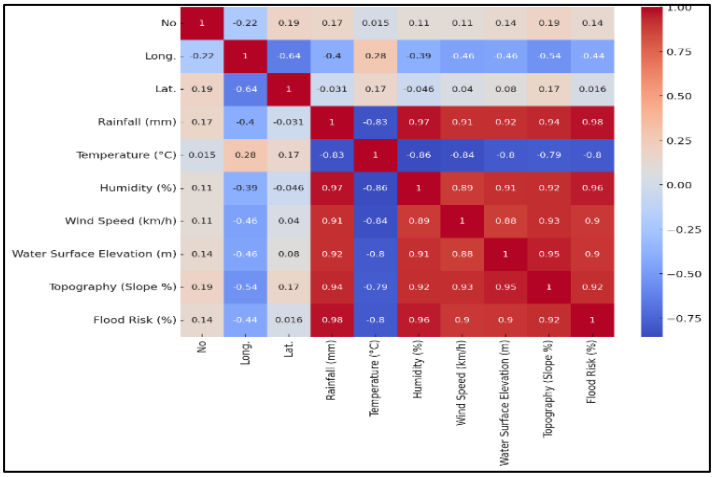


Figure 5. Variables Correlation

For example, the first entry, located at longitude -5.36275 and latitude 105.2297, recorded 140 mm of rainfall, a temperature of 26°C, and 85% humidity. The wind speed measured 16 km/h, and the water surface elevation was 10 meters. The slope at this location was 12%, and the calculated flood risk reached 80%. This process continued for each row, with varying values across environmental variables, reflecting the diverse conditions that influenced flood risk. This dataset served as a valuable resource for developing predictive models to forecast flood events based on environmental factors. By analyzing how these variables interacted, the study assessed flood risk more accurately and identified high-risk areas requiring intervention or mitigation. Overall, the dataset provided a comprehensive foundation for flood risk analysis, incorporating both spatial and environmental dynamics.

Table 3. Dataset

No	Long.	Lat.	Rainfall (mm)	Temperature (°C)	Humidity (%)	Wind Speed (km/h)	Water Surface Elevation (m)	Topography (Slope %)	Flood Risk (%)
1	-5.36275	105.2297	140	26	85	16	10	12	80
2	-5.36275	105.2297	130	27	82	15	9	11	75
3	-5.45505	105.25803	150	25	88	18	12	15	85
4	-5.43291	105.27265	120	28	80	14	8	10	70
5	-5.40631	105.25345	135	26	84	17	11	13	78
6	-5.40273	105.28375	110	29	78	12	7	9	65
7	-5.47791	105.28803	145	25	86	19	13	16	82
8	-5.45630	105.24563	130	27	83	15	10	12	75
9	-5.43487	105.26758	140	26	87	16	12	14	80
10	-5.36275	105.2297	125	26	82	14	9	11	70
11	-5.36275	105.2297	135	25	84	17	10	13	75
12	-5.45505	105.25803	145	24	87	18	11	14	85
13	-5.43291	105.27265	120	27	80	13	8	10	68
14	-5.40631	105.25345	130	26	83	16	9	12	73
15	-5.40273	105.28375	115	28	78	12	7	9	65
16	-5.47791	105.28803	140	25	85	18	12	14	80

 ISSN 2723-6471

17	-5.45630	105.24563	130	27	82	14	9	11	72
18	-5.43487	105.26758	135	26	85	15	10	13	78
19	-5.36275	105.2297	140	26	85	16	11	12	78
20	-5.36275	105.2297	120	27	80	13	9	10	70
21	-5.45505	105.25803	155	24	89	19	12	15	90
22	-5.43291	105.27265	125	28	82	15	8	11	72
23	-5.40631	105.25345	145	26	86	18	11	14	80
24	-5.40273	105.28375	130	29	81	13	9	12	75
25	-5.47791	105.28803	150	25	87	20	13	16	85
26	-5.45630	105.24563	135	27	83	17	10	13	78
27	-5.43487	105.26758	140	26	86	16	12	14	80

3. Result

3.1. Flood Risk Across Regions

Figure 6 mapped flood risk using geo-coordinates and showed the flood risk levels at different locations. Darker markers indicated higher flood risk in those areas. The map highlighted critical hotspots where mitigation efforts needed prioritization, helping decision-makers plan resource allocation more effectively.

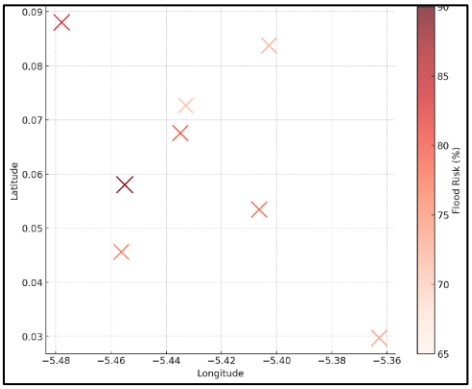


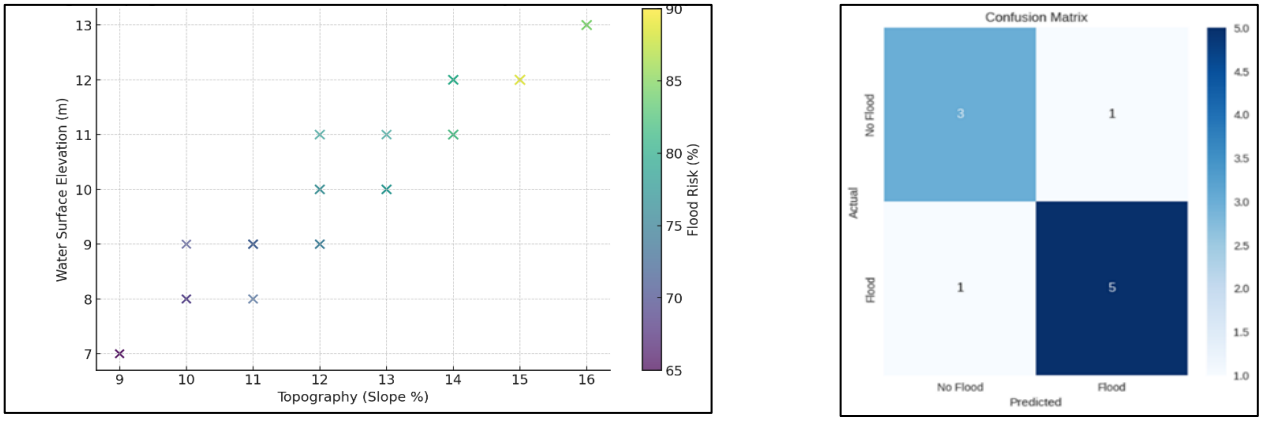
Figure 6. Risk Level Locations

3.2. How the Environment Shapes Flood Risk

Figure 7 demonstrated how topography, elevation, and rainfall collectively shaped flood susceptibility, with steep terrain and high rainfall intensity serving as key flood indicators. The scatter plot revealed a positive correlation between topographic slope (9–16%) and water surface elevation (7–13 m), with flood risk color-coded from purple (65%) to yellow (90%). Higher flood risk concentrated in areas with steeper slopes and higher elevations, as steep terrain created rapid runoff, reduced infiltration, and increased flow velocities that amplified flood likelihood when combined with intense rainfall.

The confusion matrix validated the flood prediction model's performance, showing 80% accuracy (8 out of 10 correct predictions). The model effectively identified actual flood events, with 5 true positives out of 6 total floods, confirming that topographic variables combined with rainfall data provided strong predictive power for flood assessment. One false negative suggested that some topographic–rainfall interactions remained complex to capture, but the results supported using steep terrain with high rainfall as reliable flood indicators overall.

 ISSN 2723-6471



(a) Valuable Insight (b) Confusion Matrix

Figure 7. Valuable Insight and Confusion Matrix

3.3. Tracking Flood Through Satellite Imaging

Figure 8 showed the administrative and topographic context, highlighting settlements, infrastructure, and the main water body (in blue), which provided a spatial reference for understanding which areas and communities faced flood risks. Furthermore, figure 8 used satellite imagery with a color-coded duration analysis (red-to-blue gradient) to reveal how long different areas remained flooded. Darker blue zones indicated extended flood durations, while the gradient illustrated varying levels of flood persistence across the landscape.

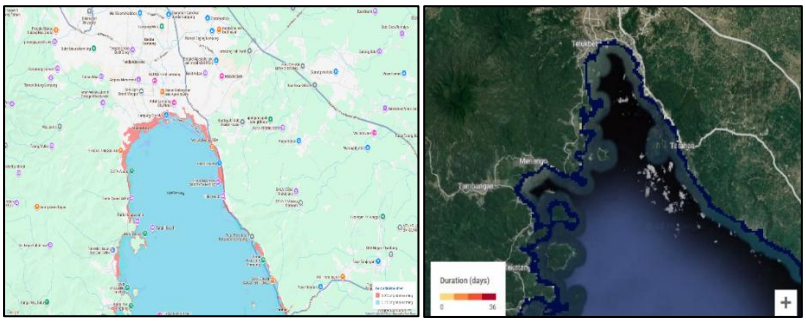


Figure 8. Satellite Imagery

This approach enabled planners to distinguish between permanently flooded areas, seasonal flood zones, and occasionally inundated regions. By mapping flood persistence duration, authorities made more informed decisions regarding infrastructure placement, emergency preparedness, and community resilience strategies. Areas with consistent, long-duration flooding required different management approaches compared to zones that experienced brief, infrequent inundation.

3.4. Model Performance

The ANN model achieved a remarkable 92% accuracy in flood prediction, significantly outperforming traditional methods, which typically ranged between 75–85%. This strong performance was supported by robust evaluation metrics, including a MSE of 1.41 and a coefficient of determination (R^2) of 0.94, indicating that the model effectively captured complex, non-linear relationships—such as those between rainfall, slope, and water flow that were often overlooked by conventional statistical and hydrological models. While linear regression resisted with variable interactions and hydrological models required difficult-to-estimate parameters, the ANNs architecture excelled in both accuracy and generalization. In particular, the model demonstrated superior capability in detecting areas with elevated flood risk, delivering lower prediction errors and better generalization to new, unseen datasets. Unlike traditional hydrological models that required extensive parameter tuning, or statistical techniques that often overlooked complex non-linear interactions, the ANNs effectively captured these intricate spatial and environmental relationships. As shown in figure 9, certain model configurations (e.g., with 5, 7, or 12 neurons) exhibited signs of overfitting, evidenced

 ISSN 2723-6471

by a sharp decline in testing correlation despite strong training results. This emphasized the importance of maintaining a proper balance between bias and variance when optimizing the model, a trade-off further detailed in table 4.

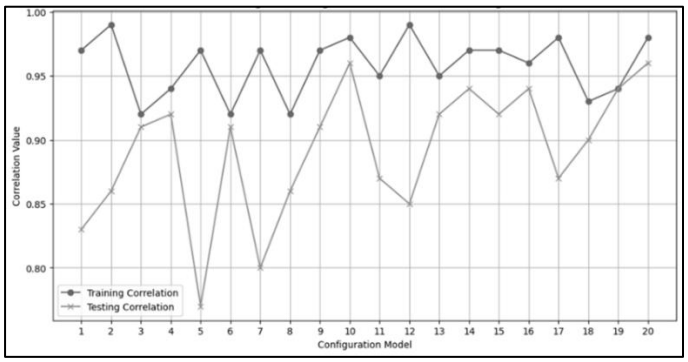


Figure 9. Correlations Between Model Training and testing

Table 4. Model Training Results

Config	Layers	Neurons	Corr	Error	Config	Layers	Neurons	Corr	Error
1	3	40	0.97	0.00181	11	2	50	0.95	0.00571
2	4	50	0.99	0.00439	12	5	30	0.99	0.01166
3	2	30	0.92	0.02140	13	4	50	0.95	0.03586
4	3	20	0.94	0.00476	14	3	50	0.97	0.00585
5	4	40	0.97	0.01037	15	5	50	0.97	0.00951
6	2	30	0.92	0.00099	16	5	20	0.96	0.00202
7	3	30	0.97	0.00838	17	5	20	0.98	0.01589
8	2	20	0.92	0.00004	18	2	50	0.93	0.00820
9	5	50	0.97	0.00193	19	5	10	0.94	0.02036
10	5	20	0.98	0.00898	20	4	30	0.98	0.00134

Optimizing model complexity was essential to achieving optimal performance. Simpler models tended to exhibit higher bias, which led to weaker correlations between training and testing results, whereas more complex models often overfit the data due to increased variance. To address this, hyperparameters such as the number of layers, neurons, and learning rates were carefully adjusted to find the right balance between training accuracy and generalization. Configurations such as 3, 4, and 19 showed better alignment between training and testing performance, indicating improved generalization capability. The observed fluctuations in testing correlation further highlighted the difficulty in maintaining consistent performance across both training and unseen data. Therefore, continuous tuning and cross-validation were conducted to reduce overfitting and enhance model reliability. As presented in table 5, the comparative results further demonstrated that the proposed ANN model outperformed traditional approaches.

Table 5. Benchmark Comparison with Other Models

Model	Accuracy	MSE	R^2
ANN (Proposed)	92%	1.41	0.94
Linear Regression	76%	6.35	0.69
Hydrological Model	83%	3.2	0.81

 ISSN 2723-6471

3.5.Uncertainty and Sensitivity Analysis

To address model uncertainty and validate prediction reliability, Monte Carlo simulations (n=100 iterations) were conducted to assess prediction variance across different scenarios. Additionally, sensitivity analysis using the One-at-a-Time (OAT) method was performed to identify the most influential variables affecting flood risk predictions. In figure 10, the histogram showed how flood risk predictions varied across 100 simulation runs, with values ranging from about 4 to 10. Most predictions clustered between 7 and 8 with forming a near-normal distribution. This pattern demonstrated that the model maintained consistent performance and low prediction variance. Despite uncertainties in the input data, the model reliably estimated flood risk levels.

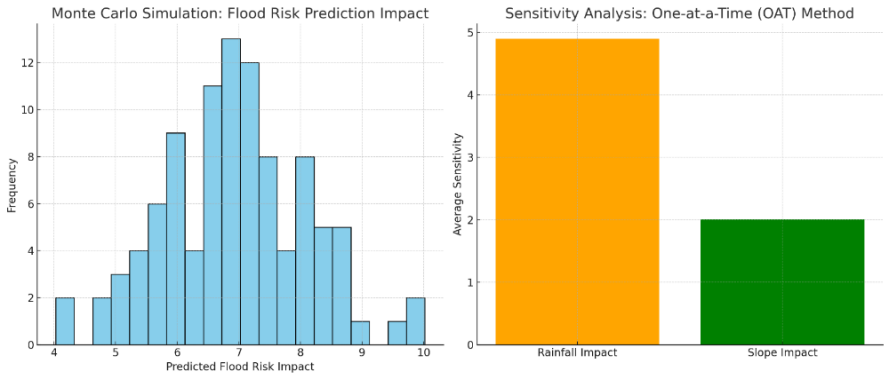


Figure 10. MonteCarlo Simulations and One-at-aTime Analysis

Moreover, the bar chart revealed the relative importance of key variables in flood prediction. Rainfall showed the highest average sensitivity (~5.0), confirming its dominant role in flood risk assessment. Slope displayed moderate sensitivity (~2.0), indicating its secondary yet significant contribution to flood susceptibility. This analysis validated the model's emphasis on these critical topographic and meteorological factors.

The combined results demonstrated that the model-maintained prediction stability through Monte Carlo validation. Rainfall emerged as the primary driver of flood risk, followed by topographic slope. This quantitative sensitivity assessment supported the model's variable selection and provided confidence intervals for operational flood risk predictions. The ANNs model was deployed in a simulated early warning system dashboard, showing flood risk maps updated in real time. Integration with BMKG and BPBD APIs is planned for real-world deployment shows on figure 11.

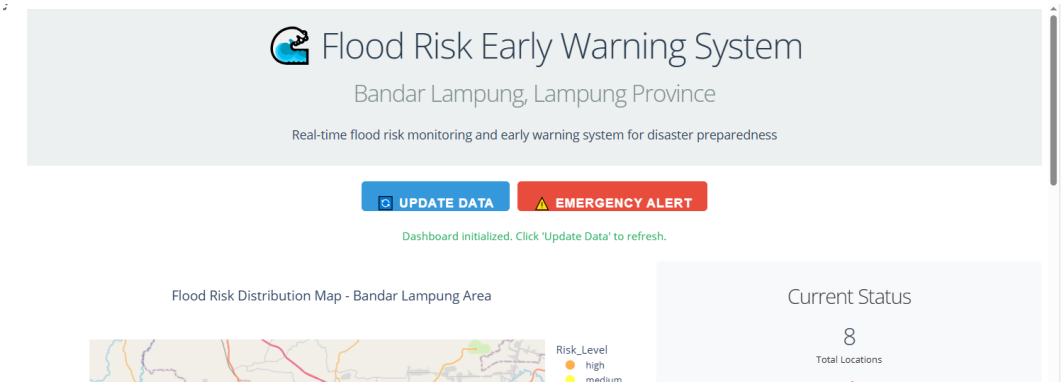


Figure 71. System Deployment Mockup (dashboard flood prediction)

4. Conclusion

This study finds that a Geo-Spatial ANN significantly enhances the accuracy of flood prediction models. By integrating high-resolution satellite imagery, meteorological data, and topographic information, the ANN model delivers stronger predictive performance than conventional methods. With an accuracy rate of up to 92%, the model provides reliable forecasts that support effective flood risk mitigation and management. Future research should expand the dataset, refine

model training techniques, and incorporate more real-time data sources to improve scalability and predictive precision. These findings highlight the potential of combining ANN with geo-spatial data to advance flood forecasting systems. By applying machine learning techniques, flood prediction becomes faster and more accurate, enabling better disaster risk management and emergency response. Furthermore, it can further improve the model by including additional environmental variables such as urbanization levels or drainage infrastructure. Integrating real-time data also supports the development of adaptive flood warning systems that issue early alerts, helping save lives and reduce economic losses.

5. Declarations

5.1. Author Contributions

Conceptualization: R.Z.A.A., R.N., R.H., and M.S.H.; Methodology: R.N.; Software: R.Z.A.A.; Validation: R.Z.A.A., R.N., and M.S.H.; Formal Analysis: R.Z.A.A., R.N., and M.S.H.; Investigation: R.Z.A.A.; Resources: R.N.; Data Curation: R.N.; Writing Original Draft Preparation: R.Z.A.A., R.N., and M.S.H.; Writing Review and Editing: R.N., R.Z.A.A., and M.S.H.; Visualization: R.Z.A.A. All authors have read and agreed to the published version of the manuscript.

5.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

5.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

5.4. Institutional Review Board Statement

Not applicable.

5.5. Informed Consent Statement

Not applicable.

5.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] E. S. Puspita Wulandari, R. Nurpambudi, and R. A. Aziz, "Prediction model with artificial neural network for tidal flood events in the coastal area of bandar lampung City," *J. Infotel*, vol. 15, no. 2, pp. 135–149, 2023, doi: 10.20895/infotel.v15i2.882.

[2] K. R. B and G. A, "Rainfall Prediction Using Data Mining Techniques - A Survey," *Computer Science & Information Technology (CS & IT)*, vol. 2013, no. 1, pp. 23–30, 2013, doi: 10.5121/csit.2013.3903.

[3] H. Darabi *et al.*, "Development of a novel hybrid multi-boosting neural network model for spatial prediction of urban flood," *Geocarto Int.*, vol. 37, no. 19, pp. 5716–5741, 2022, doi: 10.1080/10106049.2021.1920629.

[4] D. P. M. Abellana and D. M. Lao, "A new univariate feature selection algorithm based on the best–worst multi-attribute decision-making method," *Decis. Anal. J.*, vol. 7, no. May, pp. 1-20, 2023, doi: 10.1016/j.dajour.2023.100240.

[5] Saulia, L., Hanif, F., & Herwanto, R. (2024, August). Geospatial Mapping of Occupational Safety and Health Risks of an Aquaculture Enterprise. In *IOP Conference Series: Earth and Environmental Science*, vol. 1386, no. 1, pp. 1-16, IOP Publishing, doi: 10.1088/1755-1315/1386/1/012016.

[6] H. Zhu, J. Leandro, and Q. Lin, "Optimization of artificial neural network (Ann) for maximum flood inundation forecasts," *Water (Switzerland)*, vol. 13, no. 16, pp. 1–15, 2021, doi: 10.3390/w13162252.

[7] M. S. G. Adnan *et al.*, "A novel framework for addressing uncertainties in machine learning-based geospatial approaches for flood prediction," *J. Environ. Manage.*, vol. 326, no. 1, pp. 1-13, 2023, doi: 10.1016/j.jenvman.2022.116813.

[8] N. M. Gharakhanlou and L. Perez, "Flood susceptible prediction through the use of geospatial variables and machine learning methods," *J. Hydrol.*, vol. 617, no. 1, pp. 1-21, 2023, doi: 10.1016/j.jhydrol.2023.129121.

[9] S. Kordrostami, M. A. Alim, F. Karim, and A. Rahman, "Regional flood frequency analysis using an artificial neural network model," *Geosci.*, vol. 10, no. 4, pp. 1–15, 2020, doi: 10.3390/geosciences10040127.

[10] F. Y. Dtissibe, A. A. A. Ari, C. Titouna, O. Thiare, and A. M. Gueroui, "Flood forecasting based on an artificial neural network scheme," *Nat. Hazards*, vol. 104, no. 2, pp. 1211–1237, 2020, doi: 10.1007/s11069-020-04211-5.

[11] M. Rahman *et al.*, "Flooding and its relationship with land cover change, population growth, and road density," *Geosci. Front.*, vol. 12, no. 6, p. 101224, 2021, doi: 10.1016/j.gsf.2021.101224.

[12] H. Darabi *et al.*, "A hybridized model based on neural network and swarm intelligence-grey wolf algorithm for spatial prediction of urban flood-inundation," *J. Hydrol.*, vol. 603, no. PA, p. 126854, 2021, doi: 10.1016/j.jhydrol.2021.126854.

[13] D. Tien Bui *et al.*, "A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area," *Sci. Total Environ.*, vol. 701, no. 1, pp. 1-13, 2020, doi: 10.1016/j.scitotenv.2019.134413.

[14] N. Khoirunisa, C. Y. Ku, and C. Y. Liu, "A GIS-based artificial neural network model for flood susceptibility assessment," *Int. J. Environ. Res. Public Health*, vol. 18, no. 3, pp. 1–20, 2021, doi: 10.3390/ijerph18031072.

[15] Z. Guo, J. P. Leitão, N. E. Simões, and V. Moosavi, "Data-driven flood emulation: Speeding up urban flood predictions by deep convolutional neural networks," *J. Flood Risk Manag.*, vol. 14, no. 1, pp. 1–14, 2021, doi: 10.1111/jfr3.12684.

[16] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms, 4 Edition*. 2022.

[17] F. R. G. Cruz, M. G. Binag, M. R. G. Ga, and F. A. A. Uy, "Flood Prediction Using Multi-Layer Artificial Neural Network in Monitoring System with Rain Gauge, Water Level, Soil Moisture Sensors," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2018-Octob, no. October, pp. 2499–2503, 2018, doi: 10.1109/TENCON.2018.8650387.

[18] N. Dahri, R. Yousfi, A. Bouamrane, H. Abida, Q. B. Pham, and O. Derdous, "Comparison of analytic network process and artificial neural network models for flash flood susceptibility assessment," *J. African Earth Sci.*, vol. 193, no. 1, pp. 1-16, 2022, doi: 10.1016/j.jafes.2022.104576.

[19] M. Motta, M. de Castro Neto, and P. Sarmento, "A mixed approach for urban flood prediction using Machine Learning and GIS," *Int. J. Disaster Risk Reduct.*, vol. 56, no. February, pp. 1-14, 2021, doi: 10.1016/j.ijdrr.2021.102154.

[20] S. Xie, W. Wu, S. Mooser, Q. J. Wang, R. Nathan, and Y. Huang, "Artificial neural network based hybrid modeling approach for flood inundation modeling," *J. Hydrol.*, vol. 592, no. February, pp. 1-15, 2021, doi: 10.1016/j.jhydrol.2020.125605.

[21] B. Bazartseren, G. Hildebrandt, and K. P. Holz, "Short-term water level prediction using neural networks and neuro-fuzzy approach," *Neurocomputing*, vol. 55, no. 3–4, pp. 439–450, 2003, doi: 10.1016/S0925-2312(03)00388-6.

[22] R. Tabbussum and A. Q. Dar, "Performance evaluation of artificial intelligence paradigms—artificial neural networks, fuzzy logic, and adaptive neuro-fuzzy inference system for flood prediction," *Environ. Sci. Pollut. Res.*, vol. 28, no. 20, pp. 25265–25282, 2021, doi: 10.1007/s11356-021-12410-1.

[23] A. B. Dariane and S. Azimi, "Streamflow forecasting by combining neural networks and fuzzy models using advanced methods of input variable selection," *J. Hydroinformatics*, vol. 20, no. 2, pp. 520–532, 2018, doi: 10.2166/hydro.2017.076.

[24] G. Wang, J. Yang, Y. Hu, J. Li, and Z. Yin, "Application of a novel artificial neural network model in flood forecasting," *Environ. Monit. Assess.*, vol. 194, no. 2, pp. 125-137, 2022, doi: 10.1007/s10661-022-09752-9.

[25] M. Panahi *et al.*, "Deep learning neural networks for spatially ex plicit prediction of flash flood probability," *Geosci. Front.*, vol. 12, no. 3, pp. 1-16, 2021, doi: 10.1016/j.gsf.2020.09.007.