

# Extreme Flooding in Pakistan: An AI-Powered Framework for Enhanced Urban Flood Management System

M. Yasir Zaheen<sup>1</sup>, Dure Jabeen<sup>2</sup>, Ali Akbar Siddique<sup>2</sup>, M. Aamir<sup>3</sup>

<sup>1</sup>Department of Computer Engineering., Sir Syed University of Engineering & Technology, Karachi, Pakistan

<sup>2</sup>Faculty of Engineering, Sciences and Technology, Iqra University, Pakistan

<sup>3</sup>Department of Telecommunication Engineering, Sir Syed University of Engineering, Karachi, Pakistan

\*Correspondence: [myasir@ssuet.edu.pk](mailto:myasir@ssuet.edu.pk)

**Citation** | Zaheen. Y, Jabeen. D, Siddique. A. A, Aamir. M, “Extreme Flooding in Pakistan: An AI-Powered Framework for Enhanced Urban Flood Management System”, IJIST, Vol. 07 Issue. 03 pp 1584-1596, July 2025

**Received** | June 09, 2025 **Revised** | July 23, 2025 **Accepted** | July 24, 2025 **Published** | July 25, 2025.

Urban flooding poses considerable challenges for metropolitan areas, contributing to rapid urbanization and significant climatic change. This research develops a machine learning-based Urban Flood Management System (UFMS) to predict and manage flood risks, incorporating an enhanced risk warning system for rapidly urbanizing areas. The mitigation of urban flooding parameters, such as rainfall intensity, humidity, temperature, soil moisture, land use, and drainage network capacity, is analyzed in the UFMS. The system employs the artificial intelligence model Support Vector Machine (SVM), in conjunction with ARIMA modeling, achieving a high accuracy rate of 99.99% in flood prediction to forecast flood events. The model undergoes training with two decades of historical meteorological data to augment its predictive prowess and guarantee robust performance. Results show that SVM outperforms other machine learning algorithms in handling complex, multidimensional flood data. This hybrid methodology provides real-time and highly accurate prediction of upcoming floods that leads to actionable insights for urban planners and emergency response teams. Future improvements may involve the utilization of real-time data obtained from Internet of Things (IoT) nodes combined with an advanced deep learning model to improve forecast accuracy, scalability, and reduce response time, ultimately contributing to reduced flood-related damage.

**Keywords:** Flood Prediction, Flash Flood Risk, Natural Disaster Pre-Planning, Flood Management System.



## Introduction:

Flood events inflict extensive damage on infrastructure, agricultural production, and human settlements, leading to severe economic losses and human distress. According to the World Bank, floods account for 43% of all natural disasters globally, affecting over 2 billion people in the last two decades, as shown in Figure 1. Center for Disaster Philanthropy (CDP) reported that from June to August 2024, there were at least 250 deaths in India, 200 in Nepal, 200 in Bangladesh, and 300 deaths in Pakistan, among which half were children [1]. Urban areas are particularly vulnerable due to high population densities and impermeable surfaces, which intensify runoff. Floods have long posed serious challenges for Pakistan, often leaving behind a trail of destruction. One of the worst examples was the 2010 flood, which displaced more than 20 million people and caused an estimated US\$10 billion in damages. Along with economic loss, such disasters bring heartbreaking human consequences, widespread displacement, outbreaks of waterborne diseases, and tragic loss of life.

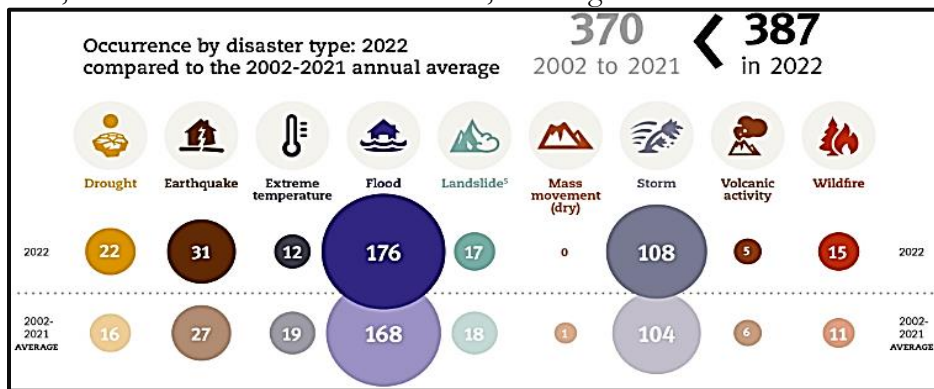


Figure 1. Comparison of disaster occurrences in 2022 with the annual average from the past decade.

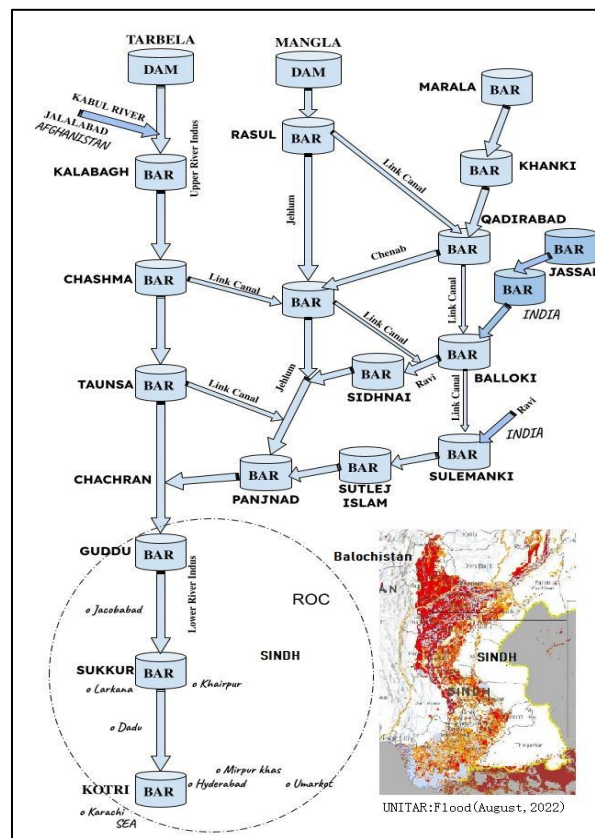


Figure 2. Route guide of Pakistan's river flow, including flood-prone areas.

The situation repeated on an even larger scale in 2022, when heavy monsoon rains led to catastrophic flooding between June and September. According to a report by NASA's Earth Observatory ([earthobservatory.nasa.gov](http://earthobservatory.nasa.gov)), Sindh province alone experienced a staggering 500% increase in rainfall during July and August, contributing to widespread flash floods. Nationwide, the Emergency Events Database (EM-DAT) reported that over 33 million people were affected. 1,739 lives were lost, and damages soared to US\$15 billion. The floods also destroyed more than 1.46 million buildings, killed over 736,000 livestock, and resulted in an additional 1,290 confirmed human casualties [2][3]. Major cities like Khairpur, Larkana, Dadu, Karachi, Mirpur-Khas, and Umer-Kot suffered severe flooding due to prolonged rain, outdated dams, poor drainage mechanisms, and ineffective green infrastructure (GI) as shown in Figure 2 [4]. This study suggests a hybrid AI-driven framework to improve early flood prediction and response because floods are happening more often in Pakistan, and traditional forecasting techniques have their limits. The paper is structured as follows: Section I introduces the background and objectives. Section II presents the proposed UFMS framework. Section III outlines the results and discussion. Section IV concludes the work and presents future directions.

### **Objectives and Novel Contribution:**

The primary objective of this study is to develop an AI-powered Urban Flood Management System (UFMS) capable of accurately predicting flood events in rapidly urbanizing regions of Pakistan. The study also uses worldwide flood management principles to provide a more robust and adaptable solution for local situations.

Designing a hybrid forecasting model that integrates real-time IoT sensor data with Support Vector Machines (SVM) and ARIMA techniques to improve the accuracy of urban flood prediction.

Evaluating the model's performance using key statistical metrics such as accuracy, precision, recall, F1-score, and MAPE, while analyzing its effectiveness under extreme conditions.

Implementing an integrated web and mobile platform that delivers early flood warnings, evacuation routes, real-time monitoring, and supports local authorities in damage estimation and relief planning.

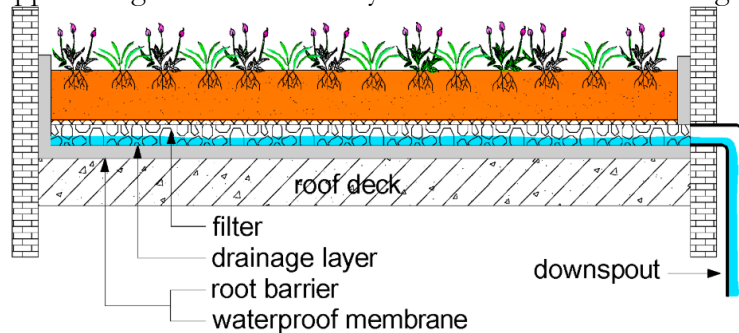
Adapting global flood resilience strategies, such as risk-based land use planning into a localized, intelligent system capable of guiding future improvements through the inclusion of regional hydrological data.

### **Related Work:**

Modern flood management prioritizes a risk-based approach, integrating hydrological, hydraulic, economic, social, and ecological factors to address flood risks effectively. This risk-based method emphasizes tailored solutions for unique floodplain characteristics, with strategies undergoing detailed risk assessments to ensure combined effectiveness. Land use policy is critical for flood mitigation, guiding spatial planning to reduce risks and improve flood-prone areas. Global frameworks, such as the Hyogo Framework and Sendai Framework [5], underscore the importance of sustainable urbanization and adaptive land allocation. By aligning risk-based management with land use policies, the integrated approach enhances resilience, reduces vulnerabilities, and fosters sustainable development while addressing socio-economic and environmental constraints [6][7].

There are many different types of urban floods, including pluvial floods, coastal floods, fluvial floods, flash flooding, sewer flooding, and groundwater flooding [8]. Effective flood management integrates engineered techniques and GI to reduce risks while promoting sustainability. Techniques like rain harvesting, reforestation, soil conservation, and groundwater recharging manage excess water by capturing runoff, enhancing absorption, and replenishing aquifers, as shown in Figure 3. Storage systems, such as reservoirs, and structural

defenses, like dikes and floodwalls, mitigate peak flows and protect urban areas, while flow diversion and river re-profiling improve water conveyance and capacity during floods [9]. Complementing these strategies, GI mimics natural hydrological processes using vegetation, permeable surfaces, and soils to control urban flooding. Beyond flood mitigation, GI delivers co-benefits, including improved air quality, temperature regulation, biodiversity enhancement, and drought resilience [10]. Globally, countries like China and the UK emphasize GI's role in sustainable urbanization and climate adaptation. GI encompasses a range of features, from green roofs and gardens to wetlands and forests, designed to deliver ecosystem services such as flood protection, water purification, and climate resilience. Integrating green infrastructure (GI) into urban planning and adopting proactive, co-designed strategies enables cities to evolve into resilient, multifunctional systems that promote well-being, strengthen the green economy, and support long-term sustainability in the face of climate change [11].



**Figure 3.** Shows structural design for GI-based rainwater harvesting technique.

In a study [12], the authors proposed a unique approach for areas where flooding is a recurring risk. This approach is a part of flood-resilient building design, by elevating the primary living spaces above potential flood levels. With this technique, property damage is highly reduced, and recovery after a flood becomes easier and less costly. Moreover, intense urbanization has intensified groundwater consumption and reduced permeable surfaces, exacerbating flood risks. The authors in paper [13] proposed a solution to mitigate these effects by integrating rooftop Rainwater Harvesting (RWH) along with Managed Aquifer Recharge (MAR) for densely populated areas. The results demonstrated efficient rainwater preservation, minimization of peak flow, and an annual average aquifer recharge rate of 57–255 m<sup>3</sup> from 2004 to 2019 in northeast Brazil.

Artificial Intelligence (AI) algorithms, such as machine learning and data mining techniques, have become indispensable tools for flood prediction. By analyzing historical and real-time meteorological data, these algorithms can identify hidden patterns and trends, enabling accurate rainfall and flood forecasts. This capability is critical for implementing proactive measures to mitigate flood risks, protect infrastructure, and ensure public safety. AI models for flood mitigation utilize diverse inputs categorized into five key types: topographical data (elevation, slope, aspect), meteorological data (rainfall distribution and frequency), geological data (soil properties like lithology and soil type), geographical data (land use, vegetation indices from remote sensing), and anthropogenic data (proximity to artificial structures such as roads). Among these, topographical inputs are the most frequently used random parameters, while factors like slope, land use, aspect, terrain curvature, and distance from rivers are fixed parameters that can vary depending on the location.

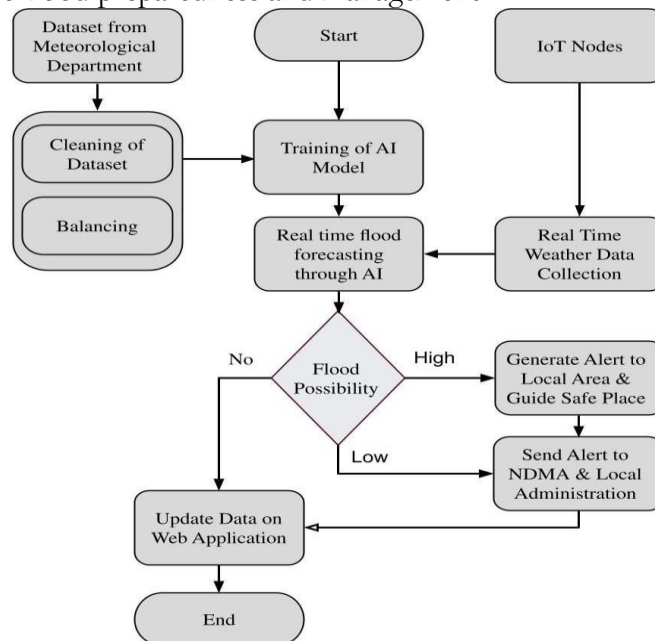
A study in [14] demonstrated the effectiveness of k-Nearest Neighbor (kNN) and Decision Trees in rainfall prediction, achieving high accuracy and recall despite climatic variability. This underscores the potential of AI-driven approaches for improving flood forecasting and risk mitigation. This study [15] investigates flood susceptibility in Ibaraki Prefecture, Japan, using a Random Forest (RF) regression model and GIS. Using eleven environmental variables and data from 224 locations, both flooded and non-flooded. The



model demonstrated exceptional performance, achieving a 99.56% validation accuracy and a high Area Under the Curve (AUC) on the ROC curve. The resulting Flood Susceptibility Map (FSM) categorized approximately one-fifth of the region as high to very high flood risk, providing policymakers with a critical tool for spatial planning and flood mitigation strategies. In paper [16] author highlights an SVM ML algorithm that shows superior performance over benchmarking models like Naïve Bayes and decision tree, incorporating some challenges of over-fitting and parameter selection for flood forecasting. The author investigates that both linear and nonlinear kernels (e.g., RBF) can achieve high accuracy under varying conditions, with SVM exhibiting robust performance even under extreme rainfall inputs. Given the temporal and seasonal complexity of flood-related parameters, hybrid models such as SVM and ARIMA offer complementary strengths in pattern recognition and forecasting accuracy. These models are individually described in the subsections.

### Proposed UFMS Framework:

This section outlines the development process of the Urban Flood Management System (UFMS), including data acquisition, model design, implementation, and system integration. Based on what we learned about how to reduce the danger of flooding around the world and AI-driven methods we looked at in the last part, we suggest a customized UFMS for urban areas in Pakistan. Figure 4 illustrates the flowchart of the proposed UFMS framework. This diagram provides a visual summary of the system's working, starting from real-time sensor data collection to flood prediction using AI models and alert dissemination via web and mobile platforms. The framework integrates IoT-based data acquisition, ARIMA and SVM-based forecasting, threshold-based alert generation, and digital communication channels to enhance flood preparedness and management.



**Figure 4.** Flow chart of proposed UFM framework.

### Data Collection:

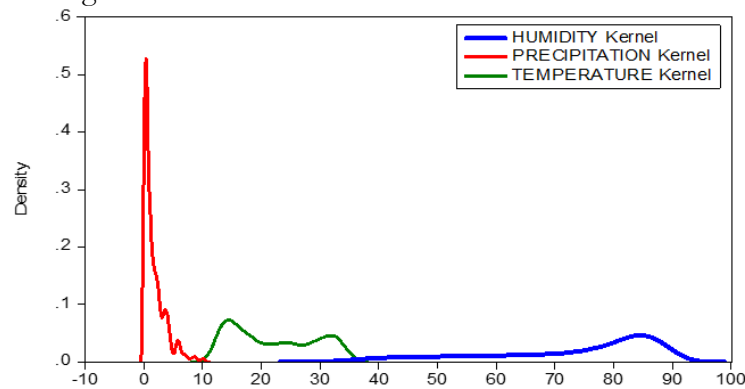
To ensure the safety of citizens in flood-prone areas, UFMS relies on a combination of real-time and historical environmental data. IoT-based sensors are deployed to measure parameters such as water level (ultrasonic sensors), temperature, pH, pressure, wind speed, rainfall intensity, CO<sub>2</sub> levels, and both soil and atmospheric humidity. These devices form IoT nodes that continuously collect environmental data used for flood risk detection. In addition, historical meteorological data is obtained from regional departments for flood-prone areas.

This dataset includes long-term records of weather patterns, flood incidents, and topographical features, enabling robust training of forecasting models.

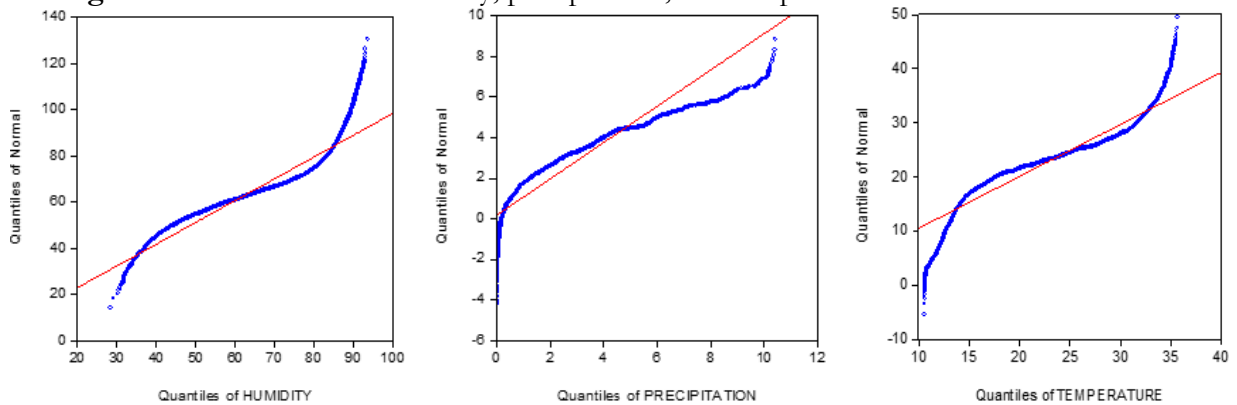
The dataset from the Pakistan Meteorological Department (PMD), though only updated daily, is supplemented by data from similar geographic regions, such as Spain's coastal city of Malaga. Parameters considered include humidity, precipitation, and temperature. The data is averaged every 30 minutes, providing 48 readings per day from 2001 to 2021, and is preprocessed for monthly modeling.

### Data Preprocessing and Feature Engineering:

The data is preprocessed to handle inconsistencies, normalize readings, and convert it into a time-series format suitable for analysis. A large dataset required segmentation and scaling to be manageable for model training. Feature engineering plays a significant role in ensuring model accuracy. Appropriate kernel selection and hyperparameter tuning are essential for SVM, while lag selection and differencing are key for ARIMA. Figures 5 and 6 represent the data distribution and quantile behavior from 2001 to 2021, reflecting trends and variability essential for forecasting.



**Figure 5.** Distribution of humidity, precipitation, and temperature data from 2001 to 2021



**Figure 6.** Data representations for the quantiles from 2001 to 2021.

### Forecasting Model Design:

#### ARIMA Modeling:

ARIMA (AutoRegressive Integrated Moving Average) was used to model the temporal behavior of humidity, precipitation, and temperature data in the historical dataset. The modeling process involved selecting appropriate autoregressive (AR) and moving average (MA) components, denoted as (P, Q), and applying the Dickey-Fuller Test to verify the stationarity of the time series. The statistical evaluation and model selection criteria for ARIMA are detailed in the Results Section Tables 3–5.

#### Support Vector Machines (SVM):

Support Vector Machines (SVM) are used for both classification and regression tasks. They construct a hyperplane decision boundary using support vectors. Kernel functions like

Linear, Polynomial, RBF, and Sigmoid enable SVM to manage non-linearly separable data. The foundational equations of SVM used in this study are:

$$\text{Decision Boundary: } \omega^T x + b = 0 \quad (1)$$

$$\text{Margin: } \frac{1}{\|\omega\|} \quad (2)$$

$$\text{Objective Function: } \min_{\omega, b} \frac{1}{2} \|\omega\|^2 \quad (3)$$

$$\text{Soft Margin Objective Function: } \min_{\omega, b, \xi} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \xi_i \quad (4)$$

Accuracy assesses the overall correctness of forecasts using equation (5), while precision quantifies the percentage of accurately anticipated positive instances among all positive predictions as depicted in equation (6). The study finds recall measures using equation (7), which evaluates the model's capacity to precisely identify positive cases from all real positive instances. Whereas using equation (8), the F1-score offers a single metric for model evaluation by striking a compromise between recall and precision. Support gives the evaluation metrics context by displaying the total number of instances of each class. When combined, these measures provide a thorough grasp of a model's categorization performance and help with decision-making throughout the analytical process. The study used several performance metrics to evaluate the classification performance of our model. These metrics include True-Negative (T-N), True-Positive (T-P), False-Negative (F-N), and False-Positive (F-P). The model's performance is further evaluated in the results section using precision, recall, and F1-score.

$$\text{Accuracy: } \frac{TN + TP}{TN + TP + FN + FP} \quad (5)$$

$$\text{Precision: } \frac{TP}{TP + FP} \quad (6)$$

$$\text{Recall(Sensitivity): } \frac{TP}{TP + FN} \quad (7)$$

$$\text{F1 Score: } \frac{P \times R}{P + R} \quad (8)$$

### System Integration:

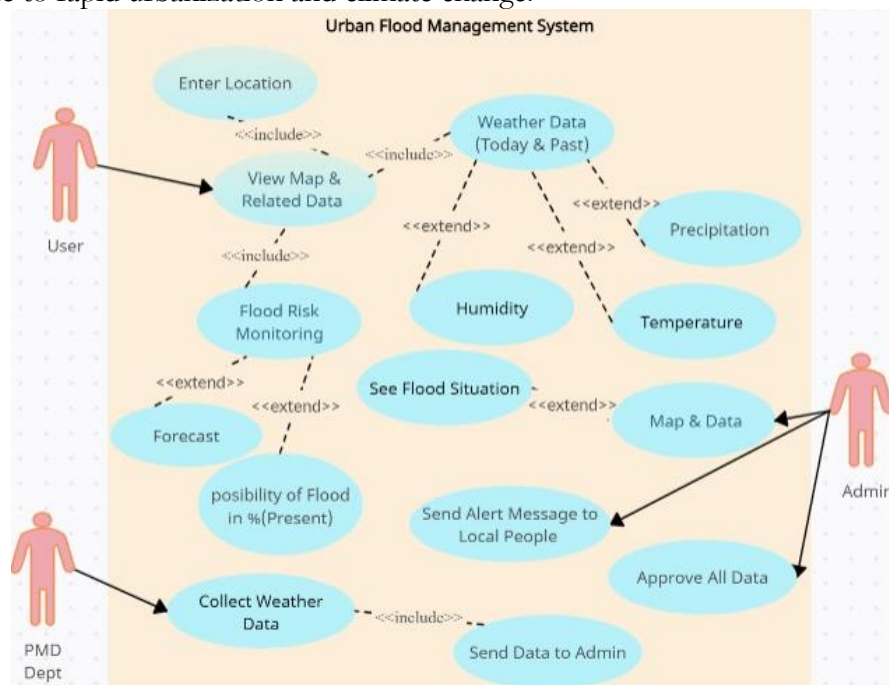
The historical dataset spanning past decades should include multiple critical factors such as rainfall frequency and intensity, water flow and pressure of incoming streams, historical flood records, and the topography of the region. This comprehensive dataset is used to train an optimized AI algorithm. Once trained, the algorithm utilizes real-time data collected from IoT nodes to predict the likelihood of an impending flood. If the probability exceeds 40%, an alert message is automatically generated and disseminated by the National Disaster Management Authority and local administration. Additionally, flood predictions are displayed on a dedicated web application for public awareness and timely response. This application provides weather-related information such as humidity, cloud coverage, and expected rainfall, similar to commercial weather apps like AccuWeather. Additionally, the application displays the predicted flood probability as a percentage. If the probability exceeds 50%, the app suggests nearby safe locations for evacuation based on the user's locality.

### Web and Mobile Application:

The mobile application supports several key functions, including real-time water level monitoring, flood warnings, detection of affected areas, suggested evacuation routes, and timely alert notifications for all users, as illustrated in Figure 7. The corresponding web application offers the same features and extends additional functionality for disaster management, such as estimating the number of victims and calculating relief requirements to support food distribution and recovery planning. Figure 8 presents a use case diagram for the UFMS online portal, highlighting the roles of users, system processes, and data flow. The system administrator is responsible for reviewing and approving weather data submitted by the Pakistan Meteorological Department (PMD). Using this data, the system calculates a risk

The screenshot shows the Urban Flood Management System (UFMS) web application. The interface has a blue header with the UFMS logo and title. A left sidebar contains navigation links: 'Manage Location', 'Weather Data', 'River Data', and 'Flood Risk Monitoring'. The main content area is titled 'Weather Data' and displays three sections: 'Rainfall' with a 'Today' dropdown, 'Temperature' with a 'Max Temp' dropdown, and 'Humidity' with a 'Today' label and a 'Max Temp' dropdown. On the right side, there is a 'Past Record' section with links for 'Monthly Rainfall', 'Monthly Temperature', and 'Monthly Humidity'.

The next step involves collecting historical meteorological data from the relevant regional departments, especially from areas identified as flood-prone. This historical dataset includes past weather patterns, records of flood events, and topographical details of each region. The AI model is trained using this historical data alongside real-time data from the sensors. This enables the system to predict flood events with higher accuracy. Once the flood probability is calculated, alerts are automatically transmitted to the UFMS dashboard and forwarded to local authorities and affected communities. This timely warning gives people a chance to take precautionary actions. The entire system operates in real time, offering continuous monitoring and early warning capabilities for urban areas that face increasing flood threats due to rapid urbanization and climate change.



---

July 2025 | Vol 07 | Issue 03



## Results and Discussions:

To assess the effectiveness of the SVM model in classifying rainfall intensity, the dataset was divided into four categories: No-Rain (N-R), Low-Rain (L-R), Moderate-Rain (M-R), and Heavy-Rain (H-R). There have been 378720 readings in all. The complete dataset is divided into training and testing groups. Training and testing make up 80% and 20% of the total, respectively. The confusion matrix in Table 1 illustrates the model's classification accuracy across the four rainfall categories. NR = 10553, LR = 6988, MR = 23131, and HR = 33072. The trained model correctly predicted NR out of 10553 NR readings with 100% accuracy, LR with 99.92% accuracy, MR and HR with 99.99% and 100% accuracy, respectively. In contrast, Table 2 shows the SVM-based model's performance parameters. These results demonstrate the reliability of the SVM model in accurately classifying rainfall intensities. The high values of Precision and Recall confirm minimal false positives and negatives, which is crucial for flood risk categorization. The high F1-scores further suggest a balanced performance across all categories.

**Table 1.** Confusion matrix of the model trained using SVM.

	N-R	L-R	M-R	H-R
N-R	10553	0	0	0
L-R	5	6983	0	0
M-R	0	0	25130	1
H-R	0	0	0	33072

**Table 2.** Performance Table of SVM Model

S No.	Precision	Accuracy	Recall	F1-Score	Support
N-R	1	0.99992	0.99952	0.99976	10553
L-R	1	0.99993	0.99929	0.99964	6988
M-R	0.99996	0.99998	1.00000	1.00000	25131
H-R	1	0.99998	0.99997	0.99997	33072

Table 3 presents the statistical parameters for the humidity models, derived using the ARIMA model with various (P, Q) values and the Dickey-Fuller Test. This test is a statistical method used to determine whether a given time series is stationary. The idea of stationarity holds great importance in time series analysis. Differentiating data to attain stationarity may be necessary for non-stationary data. Within ARIMA models, the notation (P, Q) denotes the magnitude of the moving average (MA) and autoregressive (AR) components, respectively. It appears that a variety of (P, Q) combinations were explored in the experimentation of the humidity, precipitation, and temperature models. The percentage of the dependent variable's variance that the independent variables in a regression model account for is expressed statistically as adjusted R-squared. The number of predictors in the model is considered by the modified R-squared. In Table 3, the Figures show the percentage of variability described by the models and vary from 0.023 to 0.501. A statistical model's relative quality is gauged by its Akaike's Information Criterion (AIC), and its values range from 5.842 to 5.894. It considers the model's simplicity (in terms of the number of parameters) as well as its good fitness. A better model is indicated by lower AIC values. Schwarz Criterion (SC) or Bayesian information criterion (BIC) is used for model selection in a manner like AIC. SC values range from 5.845 to 5.868. It rewards simpler models that provide a clear explanation of the data and penalizes models with more parameters. The regression standard error represents the regression model's accuracy. It shows the residuals' standard deviation, or the variation between the values that were anticipated and those that were observed. In this case, the values range from 4.489 to 4.606.

The model utilizing (4, 4) as the (P, Q) values achieves the highest adjusted R-squared value of 0.501, indicating that it explains a significant portion of the variability in the data. In

terms of simplicity and goodness of fit, the models with (1, 4), (4, 1), and (4, 4) have comparatively lower.

The AIC and SC values indicate that they might be chosen. The ARIMA (1, 4) model demonstrates the lowest AIC (value) and SC (value) among the tested configurations, indicating a more optimal balance between model complexity and predictive accuracy.

**Table 3.** Humidity statistical parameters with Dickey–Fuller Test Equation

(P, Q)	Adjusted R-squared	AIC	SC	S.E. of regression
(1,1)	0.023	5.894	5.897	4.606
(1,4)	0.072	5.842	5.845	4.489
(4,1)	0.068	5.845	5.845	4.497
(4,4)	0.501	5.865	5.868	4.54

The row highlighted above shows that the best fit equation for the model is (1, 4).

$$(1,4) \rightarrow y(t) = 9.14 \times 10^{-6} + 0.28a - 0.2b(9)$$

In this case, presented in Table 4 for precipitation, the differences in the models' Adjusted R- squared, AIC, SC, and Standard Error of Regression measurements are incredibly small. When the differences between models are minimal, it becomes challenging to definitively determine which one is the best fit based solely on these metrics. However, the model (1, 3) in Equation (9) demonstrates the highest adjusted R-squared value, along with the lowest AIC and SC values, indicating its superior performance among the evaluated models. Furthermore, in comparison to the remaining models, it possesses a slightly elevated Standard Error of Regression.

**Table 4.** Precipitation statistical parameters with Dickey–Fuller Test Equation

(P, Q)	Adjusted R-squared	AIC	SC	S.E. of regression
(1,1)	0.000295	0.388166	0.391647	0.293727
(1,2)	0.000298	0.388169	0.391650	0.294727
(1,3)	0.000319	0.388190	0.391671	0.29731

The row highlighted above shows that the best fit equation for the model is (1, 3).

$$(1,3) \rightarrow y(t) = -1.3 \times 10^{-3} + 0.007a + 0.002b(10)$$

The coefficients (-1.3e-3, 0.007, 0.002) in equation (10) give information on the strength and direction of the relationship between the variables. Table 5 shows that the ARIMA (1, 2) model has a higher adjusted R-squared and an F-statistic of 1.629033, indicating a better goodness of fit. This model also yields the most favorable values for both AIC and SC, further supporting its performance. Based on these criteria, the ARIMA (1, 2) model appears to be the most suitable choice. ARIMA (1, 2) demonstrates a higher explanatory power, lower AIC and SC values, and a Durbin-Watson statistic of 2.000254, which is comparable to other evaluated models.

**Table 5.** Temperature statistical parameters with Dickey–Fuller Test Equation

(P, Q)	Adjusted R-squared	AIC	SC	S.E. of Regression
(1,1)	0.000132	-3.952801	-3.949310	0.033521
(1,2)	0.000236	-3.952905	-3.949414	0.033519
(1,3)	0.000168	-3.952837	-3.949345	0.033520

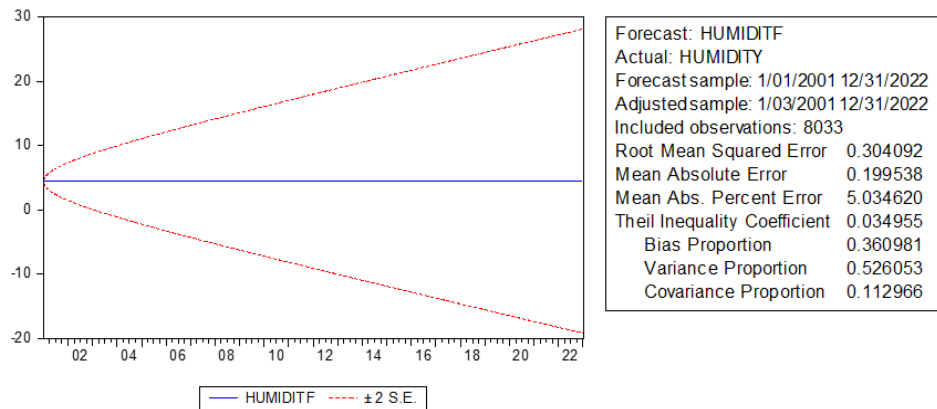
The row highlighted above shows that the best fit equation for the model is (1, 2).

$$(1,2) \rightarrow y(t) = 3.36 \times 10^{-6} + 0.019a + 0.014b(11)$$

Equation (11) represents the precise coefficients (3.36e-6, 0.019, 0.014) signify the magnitude and orientation of the correlation between the variables.

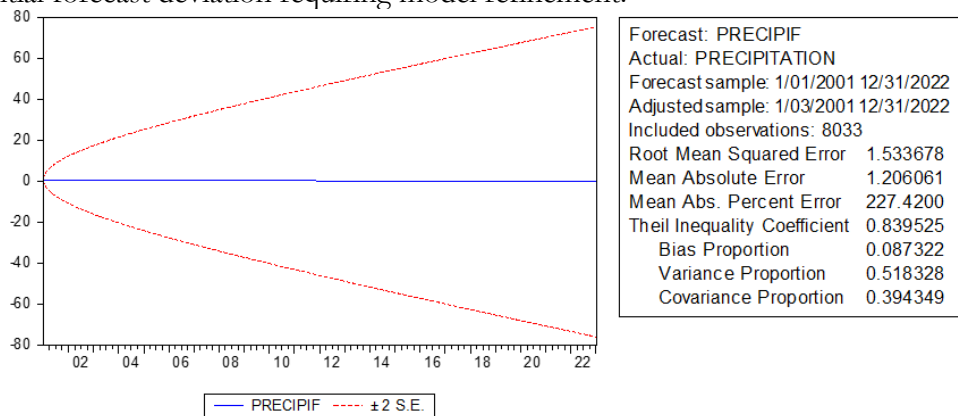
Figure 9 sums up. The ARIMA model (1, 4) that was utilized to predict humidity shows a relatively low Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), which indicates a respectable degree of forecast accuracy. However, the bias ratio suggests the presence of systematic errors, indicating that exploring

the underlying causes of this bias would be beneficial for enhancing the model's performance. Furthermore, the Theil Inequality Coefficient offers a thorough assessment of forecast accuracy by accounting for bias and dispersion, resulting in a full assessment of the model's predictive skills.



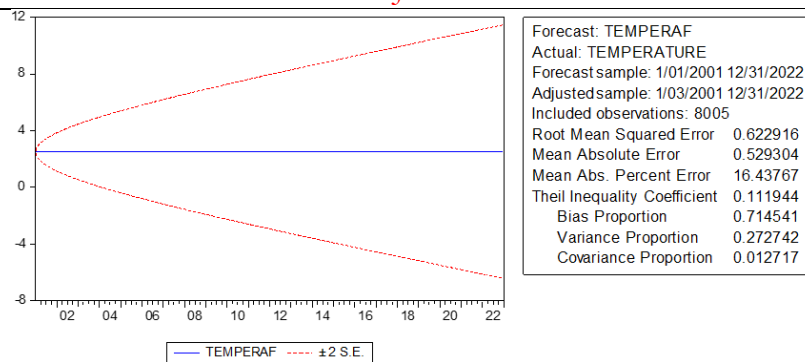
**Figure 9.** Forecast for humidity from 2001 to 2022.

Figure 10 represents the average deviation between the predicted and actual precipitation values, according to the RMSE and MAE, which are 1.206061 units and 1.533678 units, respectively. The anticipated values are, on average, more than twice as high as the actual values in percentage terms, according to the MAPE of 227.4200%, which shows a substantial percentage inaccuracy. Additionally, the Theil Inequality Coefficient is 0.839525, indicating substantial forecast deviation requiring model refinement.



**Figure 10.** Forecast for the precipitation from 2001 to 2022.

Figure 11 presents the RMSE and MAE, indicating that, on average, the predicted temperature values differ from the actual values by 0.622916 and 0.529304 units, respectively. The MAPE value of 16.43767% denotes a moderate percentage error, signifying that the predicted values deviate from the actual values by approximately 16.44% on average. The Theil Inequality Coefficient of 0.111944 suggests a relatively low level of forecast inequality, implying a reasonably accurate forecast. A significant portion of bias is evident in the prediction error breakdown, indicating that systematic errors contribute substantially to the overall forecast inaccuracy. This suggests that the model may encounter challenges in capturing specific systematic patterns in the temperature data. In conclusion, although the model offers predictions, the comparatively large bias ratio suggests the existence of systematic inaccuracies that warrant additional scrutiny or refinement of the model. Exploring alternative models, fine-tuning parameters, or incorporating additional factors may prove beneficial in improving the accuracy of temperature forecasts.



**Figure 11.** Forecast for the temperature from 2001 to 2022.

### Conclusion and Future Work:

The paper presents a novel AI-powered UFMS for real-time flood prediction that can be employed by the flood control departments of several states and countries, particularly in Pakistan's lower Indus region, for pre-flood prediction. The research reveals the potential for better UFMS. The proposed framework comprises extensive literature surveys and different flood prevention techniques, such as AI-based algorithms for the prediction and forecasting of incoming floods. The data is classified into four classes, such as nil, low, moderate, and heavy rainfall for humidity, precipitation, and temperature, respectively. Regression analysis and machine learning techniques are evaluated to understand the stochastic behavior of the unpredictable data. ARIMA modeling is applied to evaluate the best fit model for the flood prediction based on the stochastic parameters such as AIC, SC, and Standard Error squared. The data is also trained using an SVM model to predict floods based on the classification of the data. The SVM algorithm performs with almost 99% accuracy on the training dataset. Future work will focus on deep learning integration and extending this system to more cities with varying topographical challenges.

### Funding Statement:

No funding was received for this study.

### Conflicts of interest:

The authors declare that they have no conflicts of interest to report regarding the present study.

### References:

- [1] I. A. R. Hassam Bin Waseem, "Floods in Pakistan: A state-of-the-art review," *Nat. Hazards Res.*, vol. 3, no. 3, pp. 359–373, 2023, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666592123000641?via%3Dihub>
- [2] Khurram Shehzad, "Extreme flood in Pakistan: Is Pakistan paying the cost of climate change? A short communication," *Sci. Total Environ.*, vol. 880, no. 7, 2023, [Online]. Available: <https://www.bohrium.com/paper-details/extreme-flood-in-pakistan-is-pakistan-paying-the-cost-of-climate-change-a-short-communication/858167261544841216-2971#:~:text=Today%2C%20Pakistan%20is%20facing%20severe,to%20provinces%20and%20major%20cities.>
- [3] S. Ochani, S. I. Aaqil, A. Nazir, F. B. Athar, K. Ochani, and K. Ullah, "Various health-related challenges amidst recent floods in Pakistan; strategies for future prevention and control," *Ann. Med. Surg.*, vol. 82, Oct. 2022, doi: 10.1016/J.AMSU.2022.104667.
- [4] M. N. Hadi Farhadi, Ali Esmacily, "Flood monitoring by integration of Remote Sensing technique and Multi-Criteria Decision Making method," *Comput. Geosci.*, vol. 160, p. 105045, 2022, [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0098300422000127?via%3Dihub>
- [5] R. Mena, "Humanitarianism and the Sendai Framework: A 10-Year Review of Converging and Diverging Paths," *Int. J. Disaster Risk Sci.*, vol. 16, pp. 20–32, 2025, [Online]. Available: <https://link.springer.com/article/10.1007/s13753-024-00595-1>

- [6] I. H. S. P. Wiwandari Handayani, Uchendu Eugene Chigbu, Iwan Rudiarto, "Urbanization and Increasing Flood Risk in the Northern Coast of Central Java—Indonesia: An Assessment towards Better Land Use Policy and Flood Management," *Land*, vol. 9, no. 10, p. 343, 2020, doi: <https://doi.org/10.3390/land9100343>.
- [7] S. T. W. Hafiz Suliman Munawar, Ahmed W.A. Hammad, "A review on flood management technologies related to image processing and machine learning," *Autom. Constr.*, vol. 132, p. 103916, 2021, doi: <https://doi.org/10.1016/j.autcon.2021.103916>.
- [8] N. van de G. Muhammad Atiq Ur Rehman Tariq, Rashid Farooq, "A Critical Review of Flood Risk Management and the Selection of Suitable Measures," *Appl. Sci.*, vol. 10, no. 23, p. 8752, 2020, doi: <https://doi.org/10.3390/app10238752>.
- [9] F. K. S. C. Lei Li, Alexandra M. Collins, Ali Cheshmehzangi, "Identifying enablers and barriers to the implementation of the Green Infrastructure for urban flood management: A comparative analysis of the UK and China," *Urban For. Urban Green.*, vol. 54, p. 126770, 2020, doi: <https://doi.org/10.1016/j.ufug.2020.126770>.
- [10] S. K. Mohit Prakash Mohanty, Sahil Mudgil, "Flood management in India: A focussed review on the current status and future challenges," *Int. J. Disaster Risk Reduct.*, vol. 49, p. 101660, 2020, doi: <https://doi.org/10.1016/j.ijdr.2020.101660>.
- [11] Yinghong Qin, "Urban Flooding Mitigation Techniques: A Systematic Review and Future Studies," *Water*, vol. 12, no. 12, p. 3579, 2020, doi: <https://doi.org/10.3390/w12123579>.
- [12] F. G. Claire Richert, Katrin Erdlenbruch, "The impact of flood management policies on individual adaptation actions: Insights from a French case study," *Ecol. Econ.*, vol. 165, p. 106387, 2019, [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0921800919304914?via%3Dihub>
- [13] C. das N. A. Victor S.G. Baptista, Victor Hugo R. Coelho, Guillaume F. Bertrand, Gustavo B.L. da Silva, Nelson O.L. Caicedo, Suzana Maria G.L. Montenegro, Catalin Stefan, Jana Glass, Ronjon Heim, Anika Conrad, "Rooftop water harvesting for managed aquifer recharge and flood mitigation in tropical cities: Towards a strategy of co-benefit evaluations in João Pessoa, northeast Brazil," *J. Environ. Manage.*, vol. 342, p. 118034, 2023, [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0301479723008228?via%3Dihub>
- [14] S. Biruntha, B. S. Sowmiya, R. Subashri, and M. Vasanth, "Rainfall Prediction using kNN and Decision Tree," *Proc. Int. Conf. Electron. Renew. Syst. ICEARS 2022*, pp. 1757–1763, 2022, doi: 10.1109/ICEARS53579.2022.9752220.
- [15] F. A. Mohamed Wahba, Radwa Essam, Mustafa El-Rawy, Nassir Al-Arifi, "Forecasting of flash flood susceptibility mapping using random forest regression model and geographic information systems," *Heliyon*, vol. 10, no. 13, p. e33982, 2024, [Online]. Available: [https://www.cell.com/heliyon/fulltext/S2405-8440\(24\)10013-8?\\_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS2405844024100138%3Fshowall%3Dtrue](https://www.cell.com/heliyon/fulltext/S2405-8440(24)10013-8?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS2405844024100138%3Fshowall%3Dtrue)
- [16] S. G. Raju, J. Rajesh, V. Vijayaraja, R. Thiagarajan, R. Krishnamoorthy, and S. Arun, "Prediction & Forecasting of Flood Through Rainfall Measurement Using Support Vector Machine," *8th Int. Conf. Smart Struct. Syst. ICSSS 2022*, 2022, doi: 10.1109/ICSSS54381.2022.9782261.



Copyright © by authors and 50Sea. This work is licensed under Creative Commons Attribution 4.0 International License.