Prediction of flood-affected areas based on geographic information system data using machine learning

Amrul Faruq[1], Lailis Syafaah[1], Muhammad Irfan[1], Shahrum Shah Abdullah[2], Shamsul Faisal Mohd Hussein[3], Fitri Yakub[2]

[1]Department of Electrical Engineering, Faculty of Engineering, Universitas Muhammadiyah Malang, Malang, Indonesia
[2]Department of Electronic Systems Engineering, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
[3]Fakulti Teknologi dan Kejuruteraan Mekanikal, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

Article Info

Article history:

Received Feb 2, 2025
Revised Aug 18, 2025
Accepted Nov 8, 2025

Keywords:

Disaster risk reduction
Flood forecasting
Geographic information systems
Machine learning
Susceptibility mapping

ABSTRACT

Flood disasters have become more frequent and severe due to climate variability, posing significant threats to human lives, agriculture, and infrastructure. Effective disaster management and mitigation require accurate identification of flood-prone areas. This study develops an intelligent flood prediction system by integrating machine learning algorithms with geographic information systems (GIS) data to enhance flood risk assessment. The proposed system utilizes two machine learning models including random forest (RF) and support vector machine (SVM) to predict flood-susceptible areas. The models are trained on historical flood data and GIS-derived features, including elevation, slope, topographic wetness index (TWI), aspect, and curvature. The dataset undergoes preprocessing, including normalization and feature selection, before being divided into training, validation, and test sets. The models are then trained and evaluated based on their predictive performance. Evaluation metrics, particularly the area under the curve (AUC), demonstrate that RF outperforms SVM in predicting flood-prone areas. RF achieves an accuracy of 82%, while SVM records a lower accuracy of 68%. The superior performance of RF is attributed to its ability to handle complex, nonlinear relationships in flood prediction. These results highlight the effectiveness of machine learning algorithms in flood susceptibility modeling and support the integration of data-driven techniques into flood and disaster risk reduction management strategies.

Corresponding Author:

Shamsul Faisal Mohd Hussein
Fakulti Teknologi dan Kejuruteraan Mekanikal, Universiti Teknikal Malaysia Melaka
Hang Tuah Jaya, Durian Tunggal 76100, Melaka, Malaysia
Email: shamsul.faisal@utem.edu.my

1. INTRODUCTION

There has been a drastic increase in climate-related disasters in recent times [1]. The majority of disasters caused by climate change are influenced by changes in land use, population density, geological conditions, and geographical location. Among all natural disasters driven by climate, floods account for approximately 50% of global fatalities, with estimated annual losses of US$100 billion worldwide, impacting human lives and damaging agricultural land and existing infrastructure [2]. These remarkable changes in meteorological and socio-economic dynamics have increased the frequency of flood events over the years, prompting disaster management officials and policymakers to develop measures for identifying flood-prone areas by implementing technology-based flood prevention strategies.

The increasing frequency of flood disasters, as observed in recent studies, highlights how climatic variability has led to a rise in annual rainfall volume and increased surface runoff from a hydrological perspective. Data from the National Disaster Management Agency (BNPB) recorded that in November 2023 alone, floods occurred at several locations in Malang City, East Java, causing damage to infrastructure. Affected areas included the Sigura-gura Residence housing complex in Karangbesuki Village, Sukun District, as well as several other locations in Klojen District and Lowokwaru District [3]. This phenomenon has intensified the pressure to develop accurate flood risk maps that ensure sustainable flood risk mitigation and protect communities and infrastructure from hazardous threats.

A common approach to flood management involves the creation of flood hazard maps, which help identify areas at risk or prone to flooding, enabling the development and allocation of appropriate measures through either structural defenses or land-use planning [4]. Taking the Malang region as an example, the lack of comprehensive flood data and the continuous expansion of settlements into flood-prone zones have significantly increased the estimated annual losses due to flood disasters. Therefore, it is now crucial to conduct assessments of areas vulnerable to flooding by developing flood vulnerability maps that highlight and rank the likelihood of flooding at varying scales. Such maps can aid in ensuring the proper prioritization of areas in urgent need of intervention and attention from local governments.

In previous studies, the prediction of flood-prone areas has involved various hydrological or statistical modeling frameworks. For instance, rainfall-runoff hydrological models are among the most common methods used to estimate flood-vulnerable regions [5], [6]. The use of accurate flood prediction models can significantly contribute to disaster management strategies, policy formulation, and the prioritization of mitigation measures for existing hazards. Recent studies on flood prediction predominantly employ specific data-driven models that incorporate various simplified assumptions [7]. These models can include physical, data-based, and machine learning approaches. Therefore, the research problem addressed in this study is how to design an intelligent system for predicting flood-affected areas based on geographic information system (GIS) data using a machine learning algorithm approach.

Research on the implementation and development of machine learning models for flood disaster prediction began in early 2018, with the first publications highlighting the use of artificial neural networks [8]–[10]. In this initial phase, artificial neural networks were employed to model and predict flood events based on various relevant variables, such as rainfall, soil moisture, and river conditions. For advance, the research evolved to integrate multi-model and ensemble machine learning approaches to enhance the accuracy and robustness of predictions. These multi-model and ensemble techniques involved combining several different machine learning models to produce more reliable predictions [7], [11]. Additionally, the research began incorporating data from GIS as input for machine learning models. GIS data provides more detailed and specific geographic information, such as topography, land use, and water flow patterns, which are crucial for more accurately mapping flood-affected areas. By leveraging GIS data, machine learning models can generate more detailed and informative flood risk maps, which are essential for effective disaster management and mitigation planning [12].

The research aims to achieve two primary objectives. First, to develop a predictive model for flood-affected areas by leveraging machine learning algorithms integrated with spatial data based on GIS, and second, to build a web-based information management system that provides predictions of flood-affected areas using machine learning algorithms and GIS data. This system is designed to offer a reliable and accessible tool for authorities and communities, enabling more effective flood risk management and mitigation strategies.

2. METHOD

Figure 1 illustrates the core methodology for developing this system is centered on machine learning, specifically utilizing random forest (RF) and support vector machine (SVM) algorithms for predicting flood-affected areas. In the data preprocessing stage, historical flood data is exported into shapefile format and combined with non-flood data. These shapefiles are then assigned values derived from GIS data, such as digital elevation model (DEM), aspect, curvature, topographic wetness index (TWI), and slope, based on the coordinates of flood and non-flood points. During the training phase, the preprocessed data is divided into three sets: training data, validation data, and test data [13]. Once the training is complete, the RF and SVM algorithms are employed to predict flood impacts, particularly for the Malang City, using historical flood data and relevant geographic information as inputs. RF is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes (for classification) or mean prediction (for regression) of the individual trees [14], [15]. It is robust to overfitting and can handle large datasets with high dimensionality, making it suitable for integrating diverse GIS data inputs. SVM is a supervised learning algorithm that finds the optimal hyperplane to separate data into classes [16]. It is particularly effective in high-dimensional spaces and is well-suited for binary classification tasks, such as distinguishing between flood and non-flood areas [17].

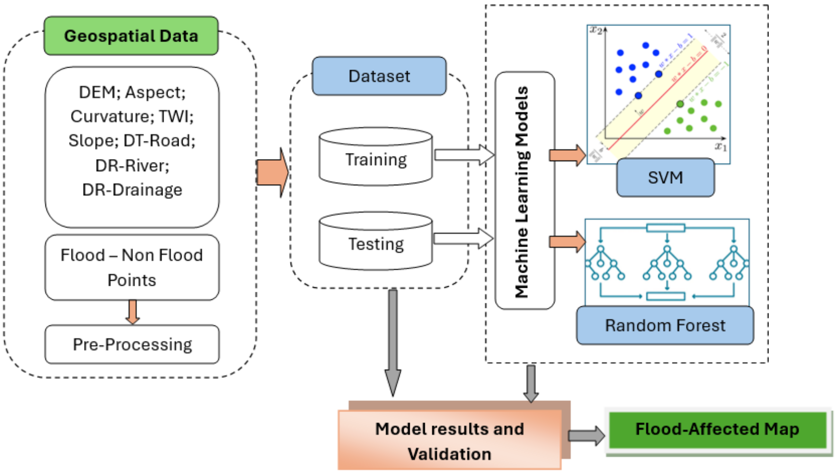


Figure 1. Machine learning-based model for flood-affected map forecasting

Historical flood data collection and preparation form the foundational step in developing the flood prediction system. Data is gathered from reliable sources, such as government agencies like the BNPB, Malang City as well as meteorological department to ensure accuracy and relevance [18]. This historical flood data, which includes records of past flood events, is exported into shapefile format, a standard GIS data format that stores geographic features (e.g., points, lines, and polygons) and their associated attributes. To create a balanced dataset, non-flood data areas with no recorded flood events is also collected. The shapefiles are then enriched with additional geographic attributes, such as coordinates, to facilitate integration with GIS data layers. This comprehensive dataset serves as the input for training machine learning models, enabling them to learn patterns and relationships between flood occurrences and environmental factors [19]. Proper preparation of this data, including handling missing values and ensuring consistency, is critical to the accuracy and reliability of the flood prediction system.

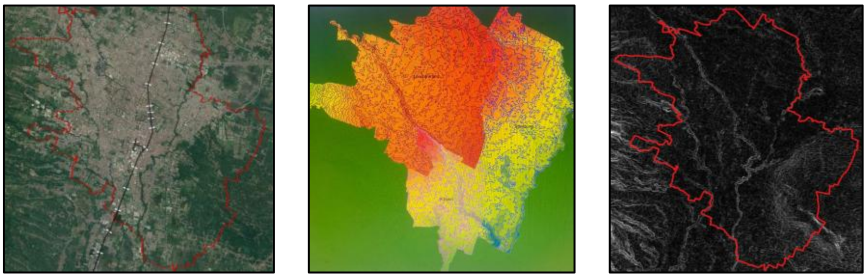
Integration with GIS data enhances the flood prediction model by incorporating detailed spatial information. Key GIS layers, such as DEM, slope, TWI, aspect, and curvature, are extracted and linked to historical flood and non-flood points based on their coordinates. DEM provides elevation data, slope indicates terrain steepness, TWI measures water accumulation potential, while aspect and curvature describe terrain orientation and shape. These attributes capture environmental and topographical factors critical to flood dynamics, enriching the dataset and improving the accuracy of machine learning models in predicting flood-prone areas [20].

The preprocessed dataset is divided into three distinct subsets to ensure effective training, validation, and testing of the machine learning models. The largest portion, training data (70-80%), is used to train the RF and SVM models, allowing them to learn the relationships between input features (e.g., GIS data) and the target variable (flood or non-flood). A smaller portion, validation data (10-15%), is reserved for tuning hyperparameters and optimizing model performance, ensuring the models are neither overfitting nor underfitting the data. Finally, the test data (10-15%) is used to evaluate the final model's accuracy and generalization ability, providing an unbiased assessment of how well the model performs on unseen data.

The evaluation method used for classification in this study is the area under the curve (AUC) of the support vector classifier (SVC), which is a robust metric for assessing the performance of binary classification models, such as flood prediction (flood vs. non-flood) [21]. The AUC is a performance metric derived from the receiver operating characteristic (ROC) curve, which plots the true positive rate (TPR) against the false positive rate (FPR) at various classification thresholds. The AUC provides a single scalar value that summarizes the model's ability to distinguish between the two classes (flood and non-flood). The SVC is a supervised machine learning algorithm used for binary or multi-class classification tasks. In this study, SVC is employed to classify areas as either flood-prone or non-flood-prone based on input features derived from GIS data and historical flood records.

3. RESULTS AND DISCUSSION

The utilization of satellite imagery data for mapping the Malang City area is carried out through the integration of SAS.Planet and ArcGIS applications. SAS.Planet is used as a tool to download high-resolution satellite imagery, which is then imported into ArcGIS for further analysis and visualization. The satellite imagery of Malang City obtained from SAS.Planet is processed in ArcGIS to generate a topographic map that represents the area's elevation based on meters above sea level (MASL). This approach has been widely used in geospatial research, as stated by [22], [23] to support spatial data management and disaster-prone area analysis. The implementation of the Malang City area map is illustrated in Figure 2. Figure 2(a) shows the satellite image, Figure 2(b) presents DEM, and Figure 2(c) displays the slope map.



(a) (b) (c)

Figure 2. The geospatial information data of (a) satellite image, (b) DEM, and (c) slope

The raw data is converted into point data using QGIS and stored in a shapefile format. This process involves extracting and utilizing geographic coordinates to spatially visualize flood events in QGIS. The shapefile format is chosen for its high compatibility with other geospatial mapping software, such as ArcGIS, and its ability to store vector data, including attribute information and location geometry. A preprocessing step is applied to the previously generated point data to prepare training data for the machine learning model. Mapped flood event locations are assigned the label ID_1, indicating flood-affected areas. As a comparison, additional randomly selected points are labeled ID_0, representing non-affected areas. These datasets are combined into a single shapefile for easier management and compatibility with geospatial software. The processed data is then used for training the machine learning model. This approach aligns with spatial data-based geospatial studies by [24] which emphasize the importance of labeling and classification for improving model prediction accuracy.

The GIS data used in this study includes the DEM, topographic aspect, curvature, slope, TWI, distance to road (DTRoad), distance to river (DTRiver), and distance to drainage (DTDrainage). All datasets are merged into a single shapefile for further processing in machine learning modeling. GIS data preprocessing is conducted using Python-based programs. In this stage, DEM is utilized to generate coordinate points covering the entire Malang City area. These coordinate points serve as references for extracting attribute values from each GIS layer. Afterward, these configurations can be treated as training and testing data as illustrated in Figure 3. This process ensures that each coordinate point contains relevant attributes for spatial analysis.

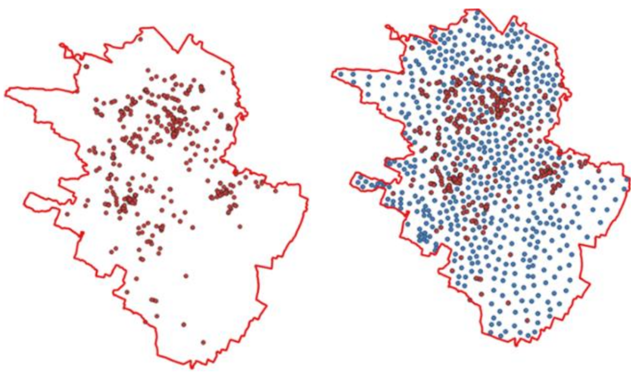
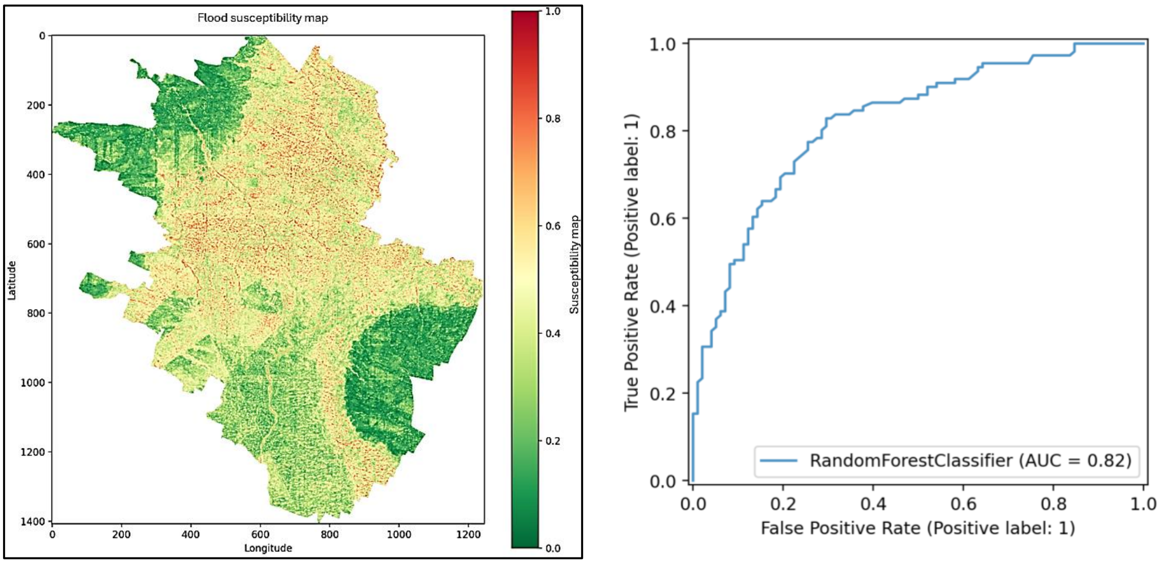


Figure 3. Flood-point data before and after augmentation for model training and testing

3.1. Random forest model evaluation

Across 15 test iterations with different data splits and tree counts, the RF model's accuracy ranged between 79% and 82%, demonstrating consistent performance despite variations in parameter settings. This consistency suggests that the model effectively generalizes flood risk patterns without significant overfitting or underfitting. Figure 4 shows the generated flood susceptibility map visually represents risk levels across the study area as showed by Figure 4(a). High-risk flood zones are marked in red, indicating areas with a significant likelihood of flooding. Moderate-risk zones appear in yellow, signifying regions with a balanced probability of flood occurrence. Meanwhile, low-risk zones are shaded in green, highlighting areas with minimal flood susceptibility. The spatial distribution of these flood-prone areas aligns with known geographic and hydrological characteristics, such as proximity to rivers, drainage channels, and low-lying regions. The results further validate the integration of GIS and machine learning in flood prediction, supporting its application in disaster risk management and urban planning [25] as depicted in Figure 4(b).



(a) (b)

Figure 4. RF model result: (a) flood-affected area and (b) model performance

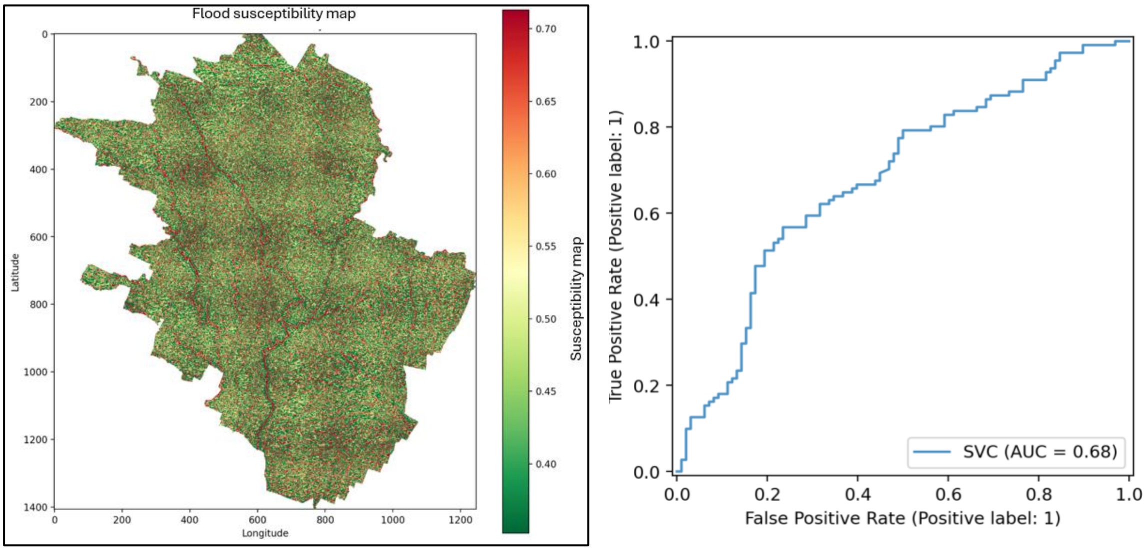
3.2. Support vector machine model evaluation

The SVM model was tested using three different kernel types: radial basis function (RBF), sigmoid, and polynomial. The accuracy varied across kernels, with RBF and polynomial achieving similar accuracy levels between 64% and 70%, while the sigmoid kernel performed significantly worse, ranging from 49-58%. The flood-affected area and model performance as shown in Figure 5. Upon evaluating the results, the RBF kernel produced the clearest flood susceptibility map compared to the other kernels. It effectively displayed the designated three-class flood risk zones, with distinct red (high risk), yellow (moderate risk), and green (low risk) areas, it is indicated in Figure 5(a). In contrast, both the sigmoid and polynomial kernels generated maps dominated by yellow, indicating an overgeneralization of moderate flood risk and a lack of clear classification boundaries. These findings suggest that the SVM model using RBF kernel is the most suitable for flood susceptibility mapping in this study, as it maintains both accuracy and interpretability [26]. These performances as depicted in Figure 5(b).

3.3. Model's performance discussion

The comparative analysis of the flood susceptibility maps generated using the RF and SVM models reveals key differences in predictive accuracy and spatial representation of flood-prone areas. The RF model demonstrated superior classification performance, achieving an accuracy of 82%, whereas the SVM model, depending on the kernel type used, exhibited lower and more variable accuracy, with the RBF and polynomial kernels ranging from 64-70% and the sigmoid kernel performing the worst at 49-58%. These discrepancies in predictive capability directly influenced the spatial delineation of flood-prone regions.

The RF-based flood susceptibility map exhibited a well-defined classification of flood risk zones, effectively capturing the high-risk (red), moderate-risk (yellow), and low-risk (green) areas with clear spatial boundaries. This outcome aligns with the model's ability to handle complex, nonlinear relationships within high-dimensional datasets, ensuring that the predictive mapping reflects real-world flood distribution more accurately. Moreover, RF's algorithm approach reduces overfitting and increases generalizability, making it a robust choice for geospatial flood modeling. Conversely, the SVM-generated flood maps varied in interpretability depending on the kernel applied. The RBF kernel produced a clearer susceptibility distribution compared to the polynomial and sigmoid kernels, yet it still lacked the distinct zonal separation achieved by RF. Notably, maps produced by the polynomial and sigmoid kernels displayed an overgeneralized classification, with an excessive dominance of moderate-risk (yellow) areas, suggesting the models struggled to define clear spatial boundaries. This result may be attributed to the sensitivity of SVM to class imbalances and its reliance on kernel-based transformations, which, in flood mapping contexts, might not fully capture the intricate spatial variability of hydrological and topographical factors.



(a) (b)

Figure 5. SVM model result of (a) flood-affected area and (b) model performance

4. CONCLUSION

This study successfully developed a flood prediction model by integrating machine learning algorithms with GIS-based spatial data. The model aims to provide rapid and accurate flood susceptibility assessments, offering valuable support for disaster mitigation and risk management. By leveraging advanced computational techniques, the study enhances the capability of predicting flood-prone areas, which is crucial for effective planning and decision-making. The experimental results indicate that the RF model outperforms the SVM model in both predictive accuracy and stability. RF achieved the highest accuracy of 82% with a data split of 70% training, 10% validation, and 20% testing using 200 decision trees. Across different configurations, RF maintained a consistently high accuracy ranging between 79% and 82%, demonstrating its robustness in flood susceptibility modeling. Meanwhile, SVM's performance varied significantly depending on the kernel type used. Among the tested kernels, the RBF kernel yielded the best accuracy at 68%, while the sigmoid kernel had the lowest accuracy at 49%. A key advantage of RF lies in its ability to handle complex, nonlinear relationships within geospatial data while maintaining high stability across different parameter settings. Its ensemble learning approach minimizes overfitting and enhances the reliability of flood susceptibility predictions. On the other hand, SVM showed greater sensitivity to kernel selection, leading to inconsistencies in classification results. Additionally, the flood susceptibility maps produced by RF exhibited well-defined spatial classifications, effectively distinguishing between high-risk (red), moderate-risk (yellow), and low-risk (green) areas. In contrast, maps generated by the SVM model, particularly those using polynomial and sigmoid kernels, displayed excessive dominance of moderate-risk zones, suggesting limitations in capturing spatial variability accurately. These findings highlight the significant potential of machine learning, particularly RF, in flood risk assessment. The study demonstrates that integrating machine learning with GIS can enhance predictive accuracy and provide a data-driven approach for disaster preparedness. Future research could explore further improvements, such as incorporating additional hydrological and meteorological parameters, optimizing hyperparameters, and testing deep learning models for enhanced performance. Moreover, integrating real-time flood monitoring data and early warning systems could further strengthen the practical application of machine learning-based flood prediction models in disaster management. Future improvements could involve refining input features, incorporate real-time hydrological data, and test deep learning approaches for enhanced predictive accuracy.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to Universiti Teknikal Malaysia Melaka (UTeM) for supporting this project. Also, to the Faculty of Engineering, Universitas Muhammadiyah Malang (UMM) for the support and resources provided throughout this work. We deeply appreciate the institution's commitment to advancing scientific research and innovation. To the MJIIT-UTM and UTeM Malaysia as our research's collaborator, the research methodology and discussion analysis are highly appreciated.

FUNDING INFORMATION

This study was funded under UMM's Research Grant No. E.6.1/95.09/RPK-UMM/2024. The work is for data acquisition and analysis. For the publication, this work was funded by Universiti Teknikal Malaysia Melaka (UTeM).

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
Amrul Faruq	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓
Lailis Syafaah		✓				✓		✓	✓	✓	✓	✓		
Muhammad Irfan					✓		✓			✓		✓	✓	
Shahrum Shah Abdullah	✓		✓	✓			✓			✓	✓	✓	✓	
Shamsul Faisal Mohd Hussein		✓	✓	✓	✓	✓	✓	✓	✓	✓				✓
Fitri Yakub					✓		✓			✓		✓		

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

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author, [SFMH]. The data, which contain information that could compromise the privacy of research participants, are not publicly available due to certain restrictions.

REFERENCES

[1] D. Partini and A. N. Hidayaht, "Disaster risk reduction efforts through education in Indonesia: a literature review," *IOP Conference Series: Earth and Environmental Science*, vol. 1314, no. 1, 2024, doi: 10.1088/1755-1315/1314/1/012049.

[2] A. Gaviglio, M. E. Marescotti, E. Demartini, and A. Corradini, "Flood damage on dairy farms: a what-if analysis to assess economic losses," *Water*, vol. 14, no. 24, 2022, doi: 10.3390/w14243984.

[3] R. Ridhoi, "Natural hazard of Southern Malang: Sitiardjo flash floods, 1932–1939," in *Embracing New Perspectives in History, Social Sciences, and Education*, London, United Kingdom: Routledge, 2022, pp. 44–48, doi: 10.1201/9781003295273-9.

[4] N. H. M. Ghazali and S. Osman, "Flood hazard mapping in Malaysia: case study Sg. Kelantan River Basin," *Catalogue of Hydrologic Analysis: Flood Hazard Mapping*, vol. 1, pp. 1–30, 2019.

[5] C. Hu, Q. Wu, H. Li, S. Jian, N. Li, and Z. Lou, "Deep learning with a long short-term memory networks approach for rainfall-runoff simulation," *Water*, vol. 10, no. 11, 2018, doi: 10.3390/w10111543.

[6] J. Tian, J. Liu, D. Yan, L. Ding, and C. Li, "Ensemble flood forecasting based on a coupled atmospheric-hydrological modeling system with data assimilation," *Atmospheric Research*, vol. 224, pp. 127–137, 2019, doi: 10.1016/j.atmosres.2019.03.029.

[7] A. Faruq, A. Marto, and S. S. Abdullah, "Flood forecasting of Malaysia Kelantan River using support vector regression technique," *Computer Systems Science and Engineering*, vol. 39, no. 3, pp. 297–306, 2021, doi: 10.32604/csse.2021.017468.

[8] M. S. M. Sabre, S. S. Abdullah, and A. Faruq, "Flood warning and monitoring system utilizing internet of things technology," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, vol. 4, no. 4, pp. 287–296, 2019, doi: 10.22219/kinetik.v4i4.898.

[9] A. Faruq, S. S. Abdullah, A. Marto, M. A. A. Bakar, S. F. M. Hussein, and C. M. C. Razali, "The use of radial basis function and non-linear autoregressive exogenous neural networks to forecast multi-step ahead of time flood water level," *International Journal of Advances in Intelligent Informatics*, vol. 5, no. 1, pp. 1–10, 2018, doi: 10.26555/ijain.v5i1.280.

[10] A. Faruq, H. P. Arsa, S. F. M. Hussein, C. M. C. Razali, A. Marto, and S. S. Abdullah, "Deep learning-based forecast and warning of floods in Klang River, Malaysia," *Ingénierie des systèmes d information*, vol. 25, no. 3, pp. 365–370, 2020, doi: 10.18280/isi.250311.

[11] A. Faruq, S. F. M. Hussein, A. Marto, and S. S. Abdullah, "Flood river water level forecasting using ensemble machine learning for early warning systems," *IOP Conference Series: Earth and Environmental Science*, vol. 1091, 2022, doi: 10.1088/1755-1315/1091/1/012041.

[12] M. Wahba, R. Essam, M. El-Rawy, N. Al-Arifi, F. Abdalla, and W. M. Elsadek, "Forecasting of flash flood susceptibility mapping using random forest regression model and geographic information systems," *Heliyon*, vol. 10, no. 13, 2024, doi: 10.1016/j.heliyon.2024.e33982.

[13] Z. Demissie, P. Rimal, W. M. Seyoum, A. Dutta, and G. Rimmington, "Flood susceptibility mapping: Integrating machine learning and GIS for enhanced risk assessment," *Applied Computing and Geosciences*, vol. 23, 2024, doi: 10.1016/j.acags.2024.100183.

[14] Z. Wang, C. Lai, X. Chen, B. Yang, S. Zhao, and X. Bai, "Flood hazard risk assessment model based on random forest," *Journal of Hydrology*, vol. 527, pp. 1130–1141, 2015, doi: 10.1016/j.jhydrol.2015.06.008.

[15] J. Chen, Q. Li, H. Wang, and M. Deng, "A machine learning ensemble approach based on random forest and radial basis function neural network for risk evaluation of regional flood disaster: a case study of the yangtze river delta, China," *International Journal of Environmental Research and Public Health*, vol. 17, no. 1, 2020, doi: 10.3390/ijerph17010049.

[16] R. Kashef, "A boosted SVM classifier trained by incremental learning and decremental unlearning approach," *Expert Systems with Applications*, vol. 167, 2021, doi: 10.1016/j.eswa.2020.114154.

[17] J. Wu, H. Liu, G. Wei, T. Song, C. Zhang, and H. Zhou, "Flash flood forecasting using support vector regression model in a small mountainous catchment," *Water*, vol. 11, no. 7, 2019, doi: 10.3390/w11071327.

[18] Perdinan *et al.*, "Initiative collaboration tool of early warning systems for early action to mitigate flood disaster impacts in Indonesia," *IOP Conference Series: Earth and Environmental Science*, vol. 1359, no. 1, 2024, doi: 10.1088/1755-1315/1359/1/012035.

[19] D. I. Putra and M. Matsuyuki, "The disaster-management capabilities of local governments: a case study in Indonesia," *Journal of Disaster Research*, vol. 15, no. 4, pp. 471–480, 2020, doi: 10.20965/jdr.2020.p0471.

[20] V. Isazade, A. B. Qasimi, A. Al Kafy, P. Dong, and M. Mohammadi, "Simulation of flood-prone areas using machine learning and GIS techniques in Samangan Province, Afghanistan," *Geodesy and cartography*, vol. 50, no. 1, pp. 20–29, 2024, doi: 10.3846/gac.2024.18555.

[21] N. Mohamadiazar, A. Ebrahimian, and H. Hosseiny, "Integrating deep learning, satellite image processing, and spatial-temporal analysis for urban flood prediction," *Journal of Hydrology*, vol. 639, 2024, doi: 10.1016/j.jhydrol.2024.131508.

[22] A. B. Adeyemi and A. A. Komolafe, "Flood hazard zones prediction using machine-learning-based geospatial approach in lower Niger River basin, Nigeria," *Natural Hazards Research*, vol. 5, no. 2, pp. 399–412, 2025, doi: 10.1016/j.nhres.2025.01.002.

[23] N. Lamichhane and S. Sharma, "Development of flood warning system and flood inundation mapping using field survey and LiDAR data for the grand river near the City of Painesville, Ohio," *Hydrology*, vol. 4, no. 2, 2017, doi: 10.3390/hydrology4020024.

[24] N. Khoirunisa, C.-Y. Ku, and C.-Y. Liu, "A GIS-based artificial neural network model for flood susceptibility assessment," *International Journal of Environmental Research and Public Health*, vol. 18, no. 3, 2021, doi: 10.3390/ijerph18031072.

[25] E. H. Ighile, H. Shirakawa, and H. Tanikawa, "A study on the application of GIS and machine learning to predict flood areas in Nigeria," *Sustainability*, vol. 14, no. 9, 2022, doi: 10.3390/su14095039.

[26] N. I. Saikh and P. Mondal, "GIS-based machine learning algorithm for flood susceptibility analysis in the Pagla river basin, Eastern India," *Natural Hazards Research*, vol. 3, no. 3, pp. 420–436, 2023, doi: 10.1016/j.nhres.2023.05.004.

BIOGRAPHIES OF AUTHORS

Amrul Faruq is an Electrical Engineer and Computer Science Engineer. He obtained bachelor's and master's degree in Electrical Engineering in 2009 and 2013, from Universitas Muhammadiyah Malang and Universiti Teknologi Malaysia, respectively. His Ph.D. obtained from the Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia, Kuala Lumpur, in 2022 recently. His research interests about computational data science and optimization algorithms. He can be contacted at email: faruq@umm.ac.id.

Lailis Syafaah received her Ph.D. in 2014 at Brawijaya University, Malang City with specializes in electrical and electronics engineering, biomedical engineering, control system, and artificial intelligence. Currently she is an Associate Professor at the Universitas Muhammadiyah Malang. She can be contacted at email: lailis@umm.ac.id.

Muhammad Irfan graduated in 1991 with a Bachelor of Engineering degree, from the Department of Electrical Engineering, Brawijaya University Malang, and a Master of Engineering in 2000 from the Department of Informatics, Sepuluh Nopember Institute of Technology (ITS), Surabaya. Currently, he is a senior lecturer at the University of Muhammadiyah Malang (UMM) and is active in research and community service. He is currently pursuing his doctoral research program at the Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia. His research interests are in renewable energy and computer engineering. He can be contacted at email: irfan@umm.ac.id.

Shahrum Shah Abdullah is an Associate Professor at Department of Electronic Systems Engineering in Malaysia Japan International Institute of Technology, Universiti Teknologi Malaysia. He receives B.Eng. (Electrical) (McGill), M.Sc. (Control Systems) (Sheffield), Ph.D. in Control Systems (Imperial College London). His expertise including electronics, control system, artificial intelligence and optimization. He can be contacted at email: shahrum@fke.utm.my or shahrum@utm.my.

Shamsul Faisal Mohd Hussein specializes in electronics engineering, with a focus on mechatronics engineering, control systems engineering, and artificial intelligence. He earned a bachelor's degree and a master's degree from the Faculty of Electrical Engineering at Universiti Teknologi Malaysia (UTM) Skudai campus in 2006 and 2011, respectively. He subsequently earned a Ph.D. degree from the Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur campus in 2022. He has worked for several years in the manufacturing and oil and gas sectors before becoming a lecturer at Universiti Teknikal Malaysia Melaka (UTeM). His current research interests encompass system identification, time series prediction and forecasting, as well as modelling and simulation. He can be contacted at email: shamsul.faisal@utem.edu.my.

Fitri Yakub received his Dip.Eng. and B.Eng. degrees in Mechatronics Engineering and Electronics Engineering from University of Technology Malaysia in 2001 and 2006 respectively. He obtained M.Sc. in Mechatronics Engineering from International Islamic University Malaysia in 2011. He received doctorate in Automatic Control Laboratory, Tokyo Metropolitan University in 2015. He is now with the Malaysia-Japan International Institute of Technology since 2012. He was attached to Alcon Johor (Ciba Vision Sdn Bhd) under MOHE CEO faculty programme from Feb 2020- Aug 2020. He is a senior member of IEEE, charted engineer from IET, member of SAE. He was recipient of an Asian Human Resource Fund by Tokyo Metropolitan Government from 2012 until 2015. His field of research interest includes intelligent control, automatic and robust control, and motion control, which related to applications of positioning systems, vehicle dynamics system, and vibration and control systems. Now expanding to IoT and machine learning application. He can be contacted at email: mfitri.kl@utm.my.