Journal of Science and Technology in Civil Engineering, HUCE, 2026, 20 (4): 51–67

MACHINE LEARNING-BASED FLOOD SUSCEPTIBILITY ASSESSMENT FOR RAILWAY NETWORKS IN CHIANG MAI: A RANDOM FOREST APPROACH

Sitthikorn Sitthikankun[a], Sattaya Manokeaw[b], Damrongsak Rinchumphu[a,*]

[a]*Department of Civil Engineering, Faculty of Engineering, Chiang Mai University, 239, Huay Kaew Road, Muang District, Chiang Mai, 50200, Thailand*
[b]*Office of Research Administration, Chiang Mai University, Chiang Mai 50200, Thailand*

Article history:
Received 21/7/2026, Revised 16/9/2026, Accepted 16/9/2026, Available online 16/10/2026, Version of Record 25/12/2026

Abstract

Flooding is among the most damaging natural disasters, threatening public safety, causing loss of life and economic disruption, and degrading the resilience of heavy civil infrastructure, including urban railway networks. This study aims to develop a machine-learning-based assessment framework for the flood susceptibility of railway network corridors in the Chiang Mai Comprehensive Plan Area, Thailand. By integrating the Random Forest (RF) algorithm with a Geographic Information System (GIS), the research analyzed five spatial parameters: land use, land cover, elevation, slope, and distance to rivers to identify flood-susceptible sections along the railway network. A probability thresholding technique was applied with a threshold of 0.70 to address prediction errors arising from imbalanced data effectively. The evaluation results demonstrated that the RF model performed effectively, achieving an overall accuracy of 88.2% and an Area Under the Curve (AUC) of 0.863. Feature importance analysis revealed that distance to rivers and land use were the most influential factors driving railway susceptibility. The generated susceptibility map aligns closely with historical flood records, highlighting, from the 2024 event, high-risk clusters in river plains and the city center. Ultimately, these findings can assist engineers and urban planners in flood risk management and the sustainable development and resiliency of railway infrastructure.

Keywords: flood susceptibility; random forest; railway infrastructure; geographic information system.

1. Introduction

Current climate change directly affects the hydrological cycle, causing more intense and frequent heavy rainfall than in the past. This is a major factor contributing to flood crises and severe damage to the economy and urban infrastructure. Thailand is one of the countries most continuously affected by flooding. This susceptibility is primarily caused by its tropical monsoon climate and geographically flood-susceptible river basins, coupled with extensive land-use changes and inadequate drainage infrastructure [1]. Chiang Mai province, in particular, serves as a case study that clearly showcases this issue. Although the province's overall topography consists predominantly of highlands and mountains, the railway corridor exhibits significant physical susceptibility because it lies in the upper Ping River basin and is surrounded by high mountains. This geography makes the area along the railway a natural drainage basin within a catchment [2]. In addition, Chiang Mai is an important economic and cultural tourism center that attracts tourists and generates substantial revenues. Rapid urban expansion requires spatial management to balance preserving the old city with supporting sustainable urban

*Corresponding author. *E-mail address:* damrongsak.r@cmu.ac.th (Rinchumphu, D.)
https://doi.org/10.31814/stce.huce2026-20(4)-04 p-ISSN 1859-2996; e-ISSN 2734-9268
© 2026 The Author(s). Published by Hanoi University of Civil Engineering. Licensed under CC BY-NC-ND 4.0.

growth, given the old city's flood susceptibility, which is exacerbated by altered natural drainage patterns and increased impervious surface coverage [3]. Consequently, the city experiences severe and widespread damage from urban flooding [4].

The railway system is a key infrastructure that drives the regional economy by efficiently, economically, and safely moving large numbers of passengers and goods [5]. However, as an extensive linear infrastructure, it fundamentally modifies existing topographic and hydrological systems, making effective water management and track stability critical challenges [6]. Therefore, the railway route acts as a core link for trade and tourism between the North and other parts of the country. Continuous train operation is vital to maintaining economic balance. If floods disrupt transportation, they directly affect the city's logistics and revenues. Moreover, during flood crises when many roads are cut off, the railway must serve as a reliable alternative for mass transit. Therefore, studying and predicting the susceptibility points of the railway is not just a local flood solution. Rather, this effort enhances the national transportation system's capacity to withstand climate-related flooding and improves overall climate resilience [7].

To design railway infrastructure that avoids flood-susceptible points, it is necessary to have a deep understanding of the surrounding area's physical characteristics. Under land use regulations, the Chiang Mai railway system passes through areas with diverse physical characteristics. It starts by crossing the city center, a commercial, high-density residential zone (red zone). Then, the route extends to the outer ring area, which includes a medium-density residential zone (orange) and a low-density residential zone (yellow). Before reaching the suburbs, most of the route passes through the rural and agricultural zone (green zone), which serves as an open area for natural water infiltration. Because the railway passes through the city center (red and orange zones), which has a high proportion of impermeable surfaces, the water infiltration rate drops significantly. This causes surface runoff to flow quickly into the railway structures during heavy rain [8], as shown in Fig. 1.

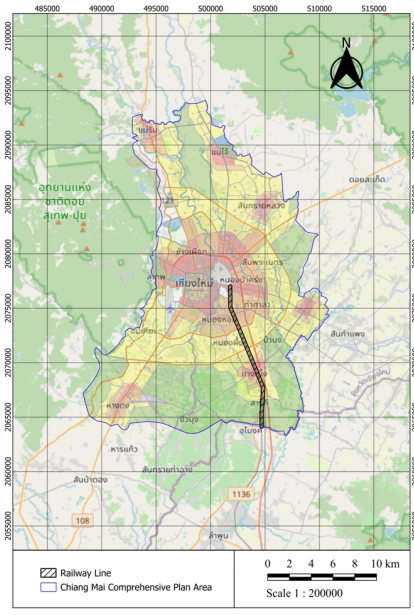


Figure 1. Map of the study area showing the railway alignment across urban land use zones in Chiang Mai. (EPSG:32647)

The rapid urban expansion in Chiang Mai increases the flood risk for railway infrastructure passing through urban areas, as seen in the 2011 and 2024 flood events [9, 10]. This problem is caused by two main factors: the Ping River overflowing its banks and urban waterlogging [10, 11]. Bank overflow occurs when accumulated rainfall exceeds the river's capacity. Meanwhile, waterlogging along the railway tends to concentrate at specific flood-susceptible points, which are localized sections of the track highly susceptible to inundation due to topographic depressions and inadequate drainage infrastructure [12]. This results from changing natural areas that could previously absorb water (pervious surfaces) into roads, buildings, and concrete yards, which are impervious surfaces. Under these conditions, heavy rain increases surface runoff, which quickly floods the railway tracks beyond the drainage system's capacity to handle [11]. Furthermore, if a river overflows during a

heavy rain event, both factors will combine to increase flood severity, significantly affecting train operations. Therefore, rather than analyzing dynamic meteorological variables such as rainfall intensity, assessing railway flood susceptibility requires focusing on the terrain's inherent physical characteristics. Evaluating static spatial parameters such as topography, river proximity, and land use helps identify exactly why a specific section is susceptible to inundation after a rain event. This parsimonious approach ensures the assessment framework remains highly practical and easily deployable for engineering applications.

Assessing railway flood susceptibility is difficult because of complex physical and urban factors. Traditional flood models require too much data for large-scale assessments. Recently, machine learning approaches have been applied to evaluate general flood susceptibility and infrastructure risks in various global contexts [13, 14]. Furthermore, broader studies have assessed the impacts of extreme weather on transportation networks, such as flood-induced road disruptions [15] and large-scale multi-hazard risks to railways [16]. However, despite this growing body of international research, empirical studies focusing specifically on localized, machine-learning-based spatial assessments for railway corridors in Thailand remain limited.

To support flood-resilient infrastructure planning, this study aims to develop a machine-learning-based flood-susceptibility assessment framework for railway networks in Chiang Mai. The Random Forest (RF) algorithm is applied because it excels at handling complex spatial data and is robust to the inherently imbalanced nature of flood datasets, in which non-flooded areas typically significantly outnumber flooded areas [17]. By mitigating majority class bias and preventing overfitting, the RF model ensures reliable predictions. Furthermore, this framework balances predictive performance with engineering interpretability. By coupling feature importance analysis with a spatial susceptibility map, the model provides an empirical tool to support decision-making. Specifically, it identifies critical spatial factors to help new railway alignments avoid highly susceptible zones. Where crossing flood-susceptible urban areas is unavoidable, these spatial insights serve as an essential database for planning targeted preventive measures, such as timely rerouting, elevating railway embankments, and improving sub-track drainage, thereby effectively managing future risks to the railway system.

2. Literature review

This study used a literature review as a conceptual framework to develop a predictive model of flood susceptibility points. The study was divided into three main parts: 1) identifying flood-susceptible areas to prepare the target dataset (dependent variable), 2) analyzing susceptibility indicators to select suitable physical and environmental factors as independent variables, and 3) reviewing approaches for applying the Decision Tree (DT) technique to enhance the model's performance and precision in assessing railway infrastructure in Chiang Mai Province.

2.1. Flood susceptibility identification

Developing a predictive model for flood susceptibility points requires selecting independent and dependent variables that align with the study area's physical characteristics. Historical flood maps can be collected as a dependent variable using the following main approaches.

a. Hydraulic modeling

Modeling is a standard method for analyzing and creating flood maps for use as dependent variable data. Bentivoglio et al. [18] applied this model to analyze flooded areas by return period to train Machine Learning models. However, this method has limitations in its application across areas with different environmental contexts. In addition, Krishnachaithanya et al. [19] stated that this model requires complex and highly detailed input data, which is a major obstacle in data-scarce areas. This

aligns with Alexopoulos et al. [20], who noted that the model's accuracy depends directly on the quality of the initial data. If the spatial data used is inaccurate, the results will not correctly reflect the actual flood situation.

b. Remote sensing

Remote sensing technology is an important tool for creating flood-map databases that can serve as dependent variables. Satellite data is suitable for broad-level overview assessments. Wu et al. [21] applied satellite data and Deep Learning techniques to detect flooded areas in near real time, effectively mitigating the traditional limitation of cloud cover during the rainy season. For cases requiring high resolution, using drones (Unmanned Aerial Vehicles: UAVs) is an effective alternative. Wienhold et al. [22] used high-resolution drone imagery to map and measure water depths.

However, remote sensing technology has limitations to consider. For satellite data, Krishnachaithanya et al. [19] stated that in densely built urban areas, signals are often obscured by high-rise buildings and roofs, making it difficult to identify actual flooded areas. Furthermore, the satellite's limited revisit time may prevent the timely recording of flood events [21]. Meanwhile, regarding drone use, Wienhold et al. [22] identified significant obstacles posed by unfavorable weather conditions, such as heavy rain and strong winds, which often occur during floods. Drones also have flight time limits due to battery capacity and legal barriers that prohibit flying in certain areas, such as near airports or security zones.

c. Ground truth data collection

To mitigate the limitations of hydrodynamic modeling and remote sensing, field surveys of flood marks are an effective tool for data collection because they provide empirical evidence. Wüthrich et al. [23] stated that accurate flood risk assessment requires field visits and interviews with locals to confirm the true extent and depth of the water, which greatly reduces model errors. Similarly, Khosravi et al. [24] confirmed that creating a flood-area database for use as a dependent variable in urban areas requires field survey data combined with Global Positioning System (GPS) coordinates of flooded locations. This method helps reduce errors and provides highly reliable ground truth data for training the model. However, field surveys have significant limitations in terms of cost and resources. Surveying large areas requires significant labor, time, and travel costs. Additionally, there is a time constraint to rapidly collect data immediately after the water recedes to record flood marks before the evidence fades or is cleaned up [25].

d. Classification of flood susceptibility

Once historical flood data have been compiled, the next phase involves categorizing this information to define the dependent variable. This classification scheme is critical to model design, as the method chosen for data grouping significantly impacts the model's predictive accuracy. Existing research typically utilizes two main frameworks: multi-class and binary class classification. 1) Multi-class classification approaches categorize susceptibility into several tiers, which helps in identifying detailed patterns within complex environments. The most prevalent technique uses a five-level scale. For example, Yaseen [26] applied statistical and remote sensing analysis to map flood susceptibility in the Ganga River basin, using five categories ranging from "Very High" to "Very Low" to enable granular assessment. Similar studies, such as Pradhan et al. [27], have integrated Explainable Artificial Intelligence (XAI) with traditional AI models to classify flood risks using this five-tier structure, while Jodhani et al. [28] emphasized that this degree of detail is essential for effective spatial prioritization. Alternative systems include a four-tier scale, in which low-risk categories are often merged to highlight areas of high susceptibility. Chowdhuri et al. [29], for instance, used the Natural Breaks

method to define indices, collapsing the "Very Low" and "Low" categories to reduce interpretative complexity. Others, like Damayanti et al. [30], prefer a three-tier model (High, Moderate, Low), which is often prioritized when the goal is to simplify data presentation for spatial management and policy-making. 2) Binary class Classification simplifies the spatial analysis by separating disaster-prone regions from safe zones. This method assigns values of 1 to areas that experienced flooding and 0 to those that did not. This framework is a standard approach in machine learning, used in research by [31–33] to prepare target datasets for predictive analysis. A common challenge in flood data classification, regardless of the chosen framework, is imbalanced datasets. Because non-flooded areas (Class 0) significantly outnumber flooded areas (Class 1), models often exhibit bias toward the majority class, thereby sacrificing the ability to identify the critical, flood-susceptible areas that are the actual focus of the study [34, 35]. Standard accuracy scores can be misleading in these cases, as a model that labels everything as "safe" can still report high accuracy. Consequently, this study prioritizes evaluation metrics that are more sensitive to class imbalance, including Precision, Recall, F1-Score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) [36, 37]. Ultimately, a binary classification approach (0 and 1) is adopted here to streamline data processing, allowing the model to concentrate specifically on the physical markers of flood susceptibility, which is the core requirement for urban planning and railway safety.

2.2. Factors influencing flood susceptibility

Determining flood susceptibility indicators is an important step in developing the model. This study uses the variable set from Sitthikankun et al. [38], which applied the standard Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) approach for a systematic literature review to extract the main factors influencing urban flooding. The factors considered are as follows.

a. Land use (X1)

Land use reflects the physical changes of an area, which directly affects the flow direction and the water infiltration capacity of the soil surface. Highly developed areas are often covered with impervious surfaces, such as roads, concrete yards, and buildings, which significantly reduce water retention capacity and increase surface runoff [12]. In addition, urban expansion without proper planning, such as encroaching on natural areas or building infrastructure that blocks waterways, will inevitably create more points of susceptibility in those areas [39].

b. Land cover (X2)

Land cover is a physical factor that affects surface water behavior and infiltration rates. Areas covered with vegetation, forests, or grasslands have the potential to retain and slow water flow, thereby reducing surface runoff. Tang et al. [40] stated that changes in land cover, especially the reduction of green areas and their replacement by buildings, will significantly decrease the drainage and infiltration capacity of the soil surface.

c. Elevation (X3)

Elevation, derived from a Digital Elevation Model (DEM), is a variable that reflects topography and plays an important role in identifying risk areas. Low-lying or basin-like areas often experience repeated waterlogging [41]. Furthermore, Rentschler et al. [39] found that current urban expansion often encroaches on low-lying areas, thereby increasing flood susceptibility. Additionally, DEM data serves as the initial database for calculating slope, which directly influences the direction of water flow.

d. Slope (X4)

Slope is a key variable that controls runoff speed and the time to concentration. High-slope areas cause water to flow quickly, reducing infiltration time and increasing the risk of flash floods. Conversely, flat or low-slope areas delay drainage, cause water accumulation, and lead to prolonged waterlogging [41]. Therefore, accurately assessing flood susceptibility points always requires considering the slope factor alongside elevation.

e. Distance to rivers (X5)

Areas near rivers are at high risk of bank overflow, especially when runoff volume exceeds the river's capacity. A. Saleh et al. [42] stated that distance to rivers is a key variable in assessing flood susceptibility because riverbank areas are the first to be affected. Moreover, Yaseen [26] noted that areas near rivers often serve as natural drainage routes. During heavy rain, water from surrounding areas will converge, making these areas more susceptible to waterlogging than those farther away.

When developing Machine Learning models with multiple independent variables, multicollinearity, or the correlation among variables, is a common problem. This can affect precision and cause bias in ranking feature importance. To solve this problem, the Variance Inflation Factor (VIF) analysis is used to test linear relationships. An appropriate VIF value should not exceed 5 or 10 [43]. If the variables meet this criterion, they are considered independent and ready to be used to train the model.

For prediction, this study applied the RF model. The RF algorithm relies on the Bagging (Bootstrap Aggregating) technique, creating multiple decision trees that work together to reduce variance and prevent overfitting. The working process of RF consists of three main steps: 1) Bootstrap Sampling to create data subsets, 2) Tree Building by randomly selecting a group of variables at each node to create diversity, and 3) collecting results through Majority Voting. With its ability to handle large, complex datasets with noise effectively, the RF model is an efficient and widely popular tool [44, 45]

To summarize the literature, although machine learning models are widely applied for general flood prediction, localized susceptibility assessments specific to railway corridors remain limited. This study synthesizes these advanced predictive techniques to address this gap within the Chiang Mai railway system. The proposed framework is an essential engineering tool. By generating highly reliable flood susceptibility maps, it establishes a reliable, data-driven foundation for supporting engineering decision-making. Consequently, this approach not only enhances effective railway management and risk mitigation but also plays a vital role in promoting sustainable railway infrastructure development.

3. Research methodology

To develop a model for assessing flood-susceptibility areas along the railway route using the RF algorithm, this study established an operational framework consisting of three main steps: 1) data management using a GIS, 2) Machine Learning model creation and managing the overfitting problem, and 3) analysis using the RF model and spatial result integration.

This study used Quantum Geographic Information System (QGIS), an open-source GIS, as the primary tool for analyzing and managing spatial data. Key data, including flood extent boundaries, flood susceptibility points, and the railway alignment, were prepared in a polygon vector format. Defining clear coordinates and spatial boundaries helps ensure high accuracy and efficiency when analyzing the data using overlay analysis with other environmental factors [46]. The implementation details are as follows.

3.1. Field survey and GIS data preparation

Preparing a spatial database is a fundamental step in creating a dataset to assess flood susceptibility. In this study, the input data structure was divided into three parts: 1) actual flood extent data from field surveys, 2) flood susceptibility factor data, which serve as independent variables, and 3) 9,450 spatial points generated within a 1-kilometer buffer along the railway corridor, which serve as reference locations for identifying risk levels. Then, all data were processed together in a GIS using overlay analysis to prepare the input data for training the model in the next step.

a. Flood extent data collection

This study relies on empirical data from field surveys of flood marks to determine the actual extent of the 2024 flood. This dataset was selected because it represents the most recent and severe flood and accurately reflects current urban and physical conditions. While validating the model against historical events such as the 2011 flood would be ideal, the methodology fundamentally relies on physical flood marks to obtain highly reliable ground-truth data. Since the 2011 event occurred over a decade ago, those physical traces have completely disappeared, making it impossible to collect reliable field data for independent validation. It integrates data with the urban flood early warning information system project in Chiang Mai province [47]. Furthermore, cooperation was received from the local disaster prevention and mitigation agency during the field visit, which identified waterlogged areas along the railway route and in surrounding areas. This ensures that the data is as accurate and comprehensive as possible. The field data collection process consists of five main steps as follows: 1) Georeferencing: Recording the measurement point code, location details, and geographic coordinates (latitude-longitude) using a GPS receiver. 2) Water Depth Measurement: Measuring the water depth from flood marks appearing on electric poles or nearby structures using a leveling instrument. The road surface level at the base of the pole is used as the initial reference point (+0.000 meters). 3) Benchmarking: Marking the high-water mark at the survey point to serve as physical reference evidence. 4) Database Entry: Compiling the measurement point codes, coordinates, and water depth values, and recording them into a central database system. 5) Flood Extent Mapping: Processing the coordinate and depth data to draw the inundation boundary on a digital map system.

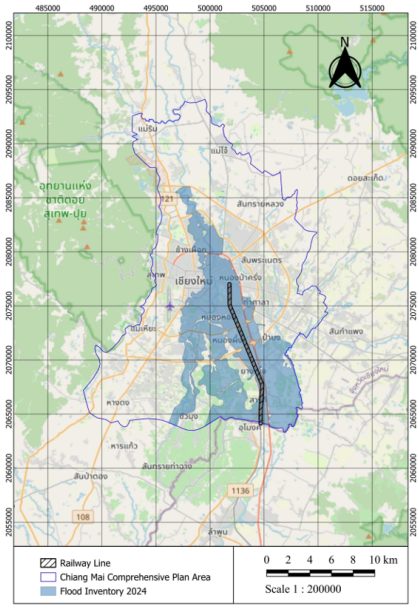


Figure 2. Flood map and railway route in Chiang Mai Province (2024) (EPSG:32647)

After completing this process, all data were exported and converted into a spatial data layer in the GIS. This prepares the target data (dependent variable) for analysis using overlay analysis with other physical susceptibility factors of the railway route in the next step. The resulting spatial distribution of the 2024 flood extent relative to the railway alignment is visually presented in Fig. 2. This figure highlights the specific sections of the railway network that intersect with flooded zones, providing a clear empirical basis for establishing the ground-truth dataset used in the subsequent modeling phase.

b. Flood conditioning factors

To prepare spatial factor data that influence flood susceptibility for use as independent variables, this study collected and created data layers from various sources. This consists of five factors as follows: 1) Land Use involves modifying and reclassifying land use types based on the color-coded zoning regulations of the Chiang Mai Comprehensive Town Planning Law. 2) Land Cover for the year 2024 was extracted from the Google Dynamic World (V1) product accessed via Google Earth Engine (GEE). This globally validated dataset offers a 10-meter spatial resolution derived from Sentinel-2 imagery, with its baseline classification driven by a deep learning Fully Convolutional Neural Network (FCNN) achieving a reported overall agreement of 73.8% [48]. To specifically tailor this dataset for flood susceptibility assessment, the original land cover classes were reclassified into three hydrologically significant categories: Impervious surface, Pervious surface, and Water surface. 3) Elevation was acquired using the NASADEM dataset accessed via GEE. While NASADEM has a native spatial resolution of 30 meters, it was systematically resampled to 10 meters during data extraction to maintain strict spatial uniformity across all analytical inputs. 4) Slope was determined by further analyzing the 10-meter resampled DEM data using spatial analysis tools in QGIS to calculate the percentage of the area's slope. 5) Distance to River includes creating a main river line data layer to analyze the spatial distance from the Ping River.

c. Target dataset generation

To define the study area boundary, a 1-kilometer buffer zone was created on each side of the railway alignment. Within this buffer corridor, a spatial dataset of 9,450 representative points was generated to cover the entire study area. A spatial overlay analysis was then performed by intersecting these points with the 2024 empirical flood extent boundary. Binary target labels were assigned by designating points located inside the flood boundary as flood-susceptible (value of 1) and points located outside the flood boundary but within the 1-km buffer zone as non-susceptible (value of 0). Among the 9,450 data points, 8,200 points were classified as susceptible areas (Class 1) and 1,250 points as non-susceptible areas (Class 0), resulting in a Class 1 to Class 0 ratio of approximately 6.56:1. This exact distribution confirms that, for this specific localized buffer during an extreme event, the flooded areas actually represent the majority class. This spatial sampling process provided a comprehensive and representative dataset for model training and evaluation.

3.2. Machine learning model development

The quality and completeness of the dataset are fundamental factors affecting the model's performance. Therefore, this study systematically split the data into two main sets. This process is used to evaluate the model's ability to predict accurately on unseen data. It also establishes a mechanism to prevent overfitting, as detailed in the steps shown in Fig. 3.

a. Data splitting and validation strategy

To ensure the model's learning process was efficient, this study divided the entire dataset into 80% for training and 20% for testing. This data-splitting approach is a standard practice in spatial modeling, helping to ensure that the training and testing sets adequately represent the underlying data distribution, enabling reliable evaluation [49].

b. Statistical analysis and evaluation metrics

To ensure the reliability of the independent variables prior to model training, a VIF analysis was conducted to detect multicollinearity. A VIF threshold of less than 5 was established to confirm that the spatial variables were independent. Furthermore, to evaluate the RF model's predictive performance, the test set was evaluated using standard statistical metrics. The selected metrics included Accuracy, Precision, Recall, F1-Score, and the AUC-ROC. These specific metrics were chosen to provide a comprehensive evaluation of the model's performance.

Furthermore, because standard Gini impurity-based feature importance is inherently biased toward continuous or high-cardinality variables [50], this study implemented Permutation Feature Importance [51]. This evaluation method ensures that the identified spatial drivers represent true physical signals rather than mathematical artifacts.

c. Model implementation

To implement the RF algorithm for flood susceptibility assessment, this study used Python and the scikit-learn library. To ensure strict reproducibility and enable exact replication of the results, a fixed random seed (SEED = 42) was applied across data partitioning, SMOTE oversampling, and model initialization. The hyperparameters of the RF Classifier were configured with 200 trees (n_estimators = 200) and a 'balanced' class weight (class_weight = 'balanced') to handle class imbalance. In contrast, the remaining hyperparameters (such as maximum tree depth and minimum samples per leaf) were set to the library's default values.

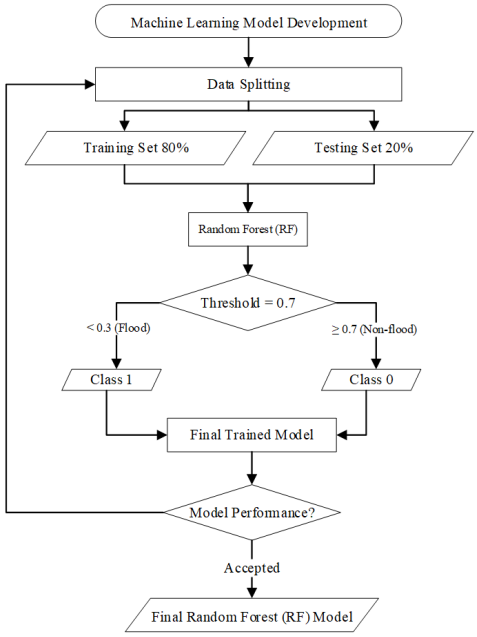


Figure 3. Machine learning framework and overfitting validation

4. Results and discussion

The data analysis and model performance evaluation are presented in accordance with the study's sequential steps. It begins by analyzing the historical flood location database and examining relationships among factors using the DT technique. It also includes creating a model to map flood susceptibility points along the railway alignment and analyzing the physical factors that influence this susceptibility.

4.1. Multicollinearity analysis

Before entering the modeling process, checking the VIF is crucial to prevent data redundancy, which could lead to errors in evaluating feature importance [43, 52]. This study assessed multicollinearity among all five flood risk factors. The verification results, as shown in Table 1, clearly indicate that the VIF values for all factors are below the specified threshold (5). These results confirm that all independent variables are independent of one another, free of multicollinearity, and well-suited for training the model in the next step.

Table 1. VIF values of the flood conditioning factors

Factors	VIF
Elevation (X3)	1.120
Distance to River (X5)	1.068
Slope (X4)	1.040
Land Use (X1)	1.025
Land Cover (X2)	1.016

4.2. Model performance evaluation

The model's capability to predict flood-susceptible areas was assessed by testing it on the previously partitioned test dataset. This evaluation relied on four statistical performance indicators: Accuracy, Precision, Recall, and F1-Score. This ensures that the measurement is comprehensive and robust to imbalanced data. The performance evaluation results of the RF model are detailed in Table 2.

Table 2. Predictive performance of the RF model computed at the 0.70 probability threshold

Class	Support (n)	Precision	Recall	F1-Score
Class 0 (Non-susceptible)	344	0.532	0.552	0.542
Class 1 (Susceptible)	2,378	0.935	0.930	0.932

(Note: Overall Accuracy = 0.882, AUC = 0.863)

Table 3. Confusion matrix of the RF model at the 0.70 probability threshold

Actual / Predicted	Predicted Class 0	Predicted Class 1
Actual Class 0	190 (TN)	154 (FP)
Actual Class 1	167 (FN)	2211 (TP)

a. Performance of models

Evaluating the RF model on the test set using the adjusted 0.70 probability threshold yields an overall accuracy of 0.882 and an AUC of 0.863. The detailed distribution of these predictions is cross-tabulated in the confusion matrix (Table 3). However, considering the classification results at the subgroup level, the limitations due to the imbalanced dataset are clearly evident. For the non-susceptible areas (Class 0, n = 344), the model shows limited performance, with recall and precision values of 0.544 and 0.529, respectively. As explicitly shown in Table 3, this limitation is reflected by the 154 instances where safe areas were misclassified as risk areas (False Positives). Conversely, when considering the study's primary target, the susceptible areas (Class 1), the model demonstrates high efficiency in detecting spatial characteristics. According to Table 3, the model successfully predicted 2,211 True Positives with only 167 False Negatives. It achieves a precision of 0.930, a recall of 0.926, and an F1-score of 0.928.

Therefore, in the context of railway engineering risk management, achieving a recall of 0.926 for detecting susceptible areas successfully meets the study's main objective. In practice, overestimating risk (False Positive) has less impact than neglecting actual susceptible areas (False Negative), which could lead to severe damage to infrastructure and operational stability. Consequently, the RF model is an effective and practical approach for analyzing spatial susceptibility points in subsequent steps.

While this study successfully establishes a foundational predictive framework using a standard 80/20 random data split, an important methodological consideration regarding spatial autocorrelation warrants attention. Because the dataset consists of 9,450 points within a confined 1 km buffer along the railway corridor, neighboring points inherently share similar geographical characteristics. Consequently, a purely random split may introduce spatial data leakage, which could influence the high predictive metrics observed (such as ~0.93 for Class 1). To further refine this approach, future research should implement spatial block or cluster cross-validation strategies [53] to completely mitigate potential spatial leakage and provide an even more rigorous evaluation.

b. Analysis of receiver operating characteristic (ROC) curve

The evaluation of the model's overall classification capability was conducted using the ROC curve and the AUC, as detailed in Fig. 4.

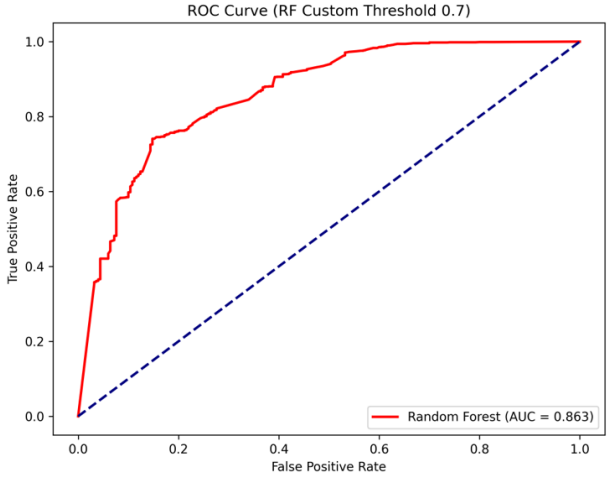


Figure 4. ROC curve

The ROC curve of the RF model clearly points toward the top-left corner of the graph. This reflects a high efficiency in correctly classifying flood-susceptible areas (True Positive Rate) coupled with a low False Positive Rate. Furthermore, the model's AUC of 0.863 confirms the stability and reliability of the analysis results. These statistical data clearly demonstrate that the RF model has sufficient capability and precision to create a map of flood-susceptibility points. This map will serve as a supporting database for engineering planning and risk management along the railway route in Chiang Mai province.

4.3. Railway flood susceptibility map

To address this, the Probability Threshold Moving technique was applied. A Precision-Recall vs. Threshold analysis was conducted, and the dynamic trade-offs among these evaluation metrics are explicitly visualized in Fig. 5. As illustrated in Fig. 5, while mathematically maximizing the standard F1-score points to a lower threshold (0.12), adopting this purely mathematical optimum is impractical for disaster risk management due to the severe penalties of underprediction.

While mathematically maximizing the standard F1-score points to a lower threshold (0.12), adopting this purely mathematical optimum is impractical for disaster risk management due to the severe penalties of underprediction. In railway engineering, neglecting actual susceptible areas (False Negative) results in far more severe consequences than overestimating risk (False Positive).

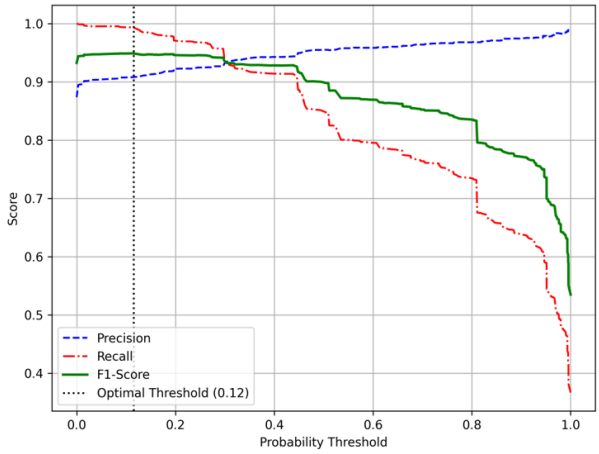


Figure 5. Precision, Recall and F1-Score across various probability thresholds for quantitative threshold selection

Consequently, this study applied a cost-sensitive decision approach, setting a strict 0.70 probability threshold as a conservative engineering safety margin for Class 0. In practice, this means that a spatial segment is classified as 'safe' only if the model is at least 70% confident; otherwise, it is conservatively flagged as susceptible. This optimal adjustment explicitly minimizes critical False Negatives while maintaining overall model stability. The spatial distribution resulting from this calibrated 0.70 threshold is comprehensively depicted in the final flood susceptibility map (Fig. 6). As shown in Fig. 6, the map clearly delineates the continuous high-risk zones along the 1-kilometer railway buffer, directly translating the model's numerical predictions into actionable geographic insights for infrastructure planning.

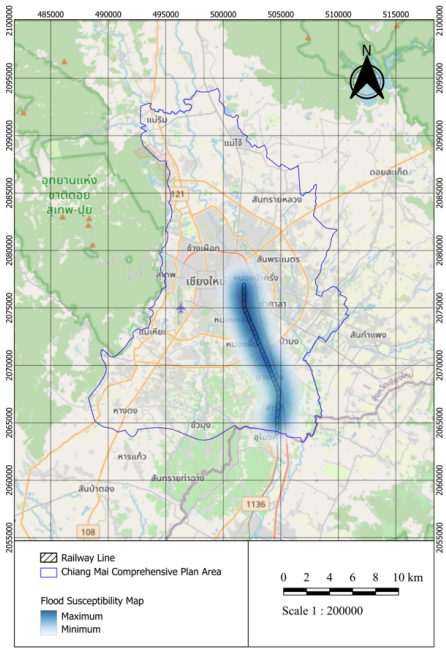


Figure 6. Flood susceptibility map within a 1-kilometer buffer along the railway corridor generated by the RF model (EPSG:32647)

4.4. Feature importance analysis

The initial feature importance analysis derived from the RF model (Table 4) indicated that continuous variables, specifically distance to rivers and elevation, were the most strongly influential factors, with initial weights of 0.310 and 0.293, respectively. However, it is important to note that the standard default method for calculating feature importance in tree-based models relies on impurity reduction, commonly known as Gini importance. Because impurity-based metrics are known to be inherently biased toward continuous or high-cardinality predictors compared to categorical ones (such as land use and land cover), a more rigorous validation was necessary. To ensure that the observed rankings reflect pure physical signals rather than artifacts of the evaluation metric, Permutation Feature Importance was further evaluated on the test set.

The permutation analysis (Table 4) effectively addressed the metric bias and revealed a crucial shift in the variable rankings. Distance to Rivers robustly maintained its position as the most dominant predictive factor (mean accuracy decrease score of 0.1398), confirming it as a true physical driver. Interestingly, Land Use (X1), a categorical variable, emerged as the second most influential factor (score: 0.0316), successfully surpassing elevation (score: 0.0163). This shift confirms that the initial dominance of elevation over land use was partly a metric artifact. Nevertheless, despite the rank swap between the second and third positions, the top three most influential variables, Distance to River, Land Use, and Elevation, ultimately remained consistent, firmly establishing their core significance in the model.

These validated results are physically plausible and highly consistent with the basin-plain setting of the study area. In this relatively flat terrain, the proximity to water bodies (Distance to River) dictates the primary hydrological flood exposure. Concurrently, land use patterns, which govern surface runoff and urban drainage capacities, exert a stronger influence on localized flooding than subtle topographical changes (Elevation). Furthermore, while VIF confirmed the absence of linear multicollinearity, it is vital to conceptually distinguish the specific roles of land use and land cover. Land cover characterizes the immediate physical surface materials, categorized as impervious, pervious, and water surfaces, which directly dictate localized water infiltration rates. In contrast, land use reflects broader socio-economic boundaries and regulatory zoning that govern overall urban infrastructure. Because a single land-use zone can contain a diverse mix of land cover types, retaining both variables ensures that the model effectively captures both microscopic surface hydrology and macroscopic urban planning factors. Consequently, slope (-0.0037) demonstrated a negligible role, which is logical for a flat topography. Ultimately, this interpretable workflow confirms that hydrological proximity and urban structural modifications are the primary spatial variables for screening flood-susceptible sections along the railway corridors.

Table 4. Feature importance analysis comparing standard Gini impurity and Permutation importance of the Random Forest model

Factors	Random Forest	Permutation Importance (Mean ± SD)
Distance to River (X5)	0.310	0.1398 ± 0.0064
Land Use (X1)	0.255	0.0316 ± 0.0049
Elevation (X3)	0.293	0.0163 ± 0.0029
Land Cover (X2)	0.038	-0.0026 ± 0.0022
Slope (X4)	0.104	-0.0037 ± 0.0053

5. Conclusion

The primary objective of this study was to develop a machine-learning-based flood susceptibility assessment framework for railway networks in Chiang Mai, Thailand. It used the RF algorithm as the primary analytical tool to strike an optimal balance between predictive performance and engineering interpretability. The results demonstrate that this model is an effective approach for assessing the physical susceptibility of railway systems, achieving an overall accuracy of 88.2% and a classification efficiency (AUC) of 0.863. Furthermore, the model's ability to analyze feature importance effectively mitigates the computational complexity often associated with advanced black-box algorithms, ensuring that the developed susceptibility map is both practically useful and physically explainable.

The feature importance analysis confirmed that hydrological and urban factors exert the strongest influence on railway route susceptibility. In particular, distance to the river and land use were identi-

fied as the most dominant predictive drivers, followed by elevation. Additionally, the VIF verification confirmed that all independent variables are statistically independent and fall well within standard acceptable criteria. These outcomes indicate that the basin topography, the proximity of the railway alignment to the main river, and the concentration of urban development governed by comprehensive town planning are key variables driving repeated embankment flooding. This aligns seamlessly with the generated susceptibility map, which illustrates clusters of high-risk areas along tracks passing through river plains and the inner city, accurately reflecting historical disaster records that have previously disrupted train operations.

The findings and the developed railway flood susceptibility map serve as empirical tools to support informed decision-making for engineers, the State Railway of Thailand (SRT), and urban planners. Specifically, these findings provide actionable insights for identifying target areas to elevate railway embankments and improve sub-track drainage structures tailored to local hydrological behaviors. For highly susceptible areas in the inner city, land-use regulations under comprehensive town planning must be strictly enforced. This approach should be coupled with promoting green infrastructure to delay and reduce surface runoff entering the railway system. Future studies should consider extending this predictive framework to railway routes in other regions, incorporating climate-change scenario simulations to evaluate long-term risks and to design resilient, sustainable infrastructure.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the author(s) used Claude and Gemini to improve the language or readability of the manuscript. The author(s) reviewed and edited the output and take full responsibility for the accuracy, originality, and integrity of the manuscript.

References

[1] Yang, S., Zhao, B., Yang, D., Wang, T., Yang, Y., Ma, T., Santisirisomboon, J. (2023). Future changes in water resources, floods and droughts under the joint impact of climate and land-use changes in the Chao Phraya basin, Thailand. *Journal of Hydrology*, 620:129454.

[2] Intarat, K., Yoomee, P., Hussadin, A., Lamprom, W. (2024). Assessment of landslide susceptibility in the intermontane basin area of Northern Thailand. *Environment and Natural Resources Journal*, 22(2): 158–170.

[3] Hongnual, K., Leelapattana, W., Thongma, W., Trakansiriwanich, K., Sitthikun, S. (2021). The development of cultural tourism attraction for model creative tourism management in the old city Chiang Mai, Thailand. *Journal of Management Information and Decision Sciences*, 24:1–8.

[4] Liu, S., Liu, R., Tan, N. (2021). A spatial improved-kNN-based flood inundation risk framework for urban tourism under two rainfall scenarios. *Sustainability*, 13(5):2859.

[5] Supanyo, P., Pakdeepinit, P., Katesophit, P., Meeprom, S., Kantasa-ard, A. (2025). Hybrid LSTM-DES models for enhancing the prediction performance of rail tourism: A case study of train passengers in Thailand. In *SMILE 2025*, MDPI, 97(1):1.

[6] Gorbunov, O., Sinsabvarodom, C., Rinchumphu, D. (2026). Hydraulic conductivity estimation for ballasted railway tracks: A systematic review. *Journal of Science and Technology in Civil Engineering (JSTCE) - HUCE*, 20(2):1–19.

[7] Apollonio, C., Iemmolo, G., Di Modugno, M., Apollonio, M., Petroselli, A., Recanatesi, F., Giannetta, D. (2025). A multi-parameter approach to support sustainable hydraulic risk analysis for the protection of transportation infrastructure: The case study of the Gargano Railways (Southern Italy). *Sustainability*, 17(9):4151.

[8] Ministry of Interior (2012). *Ministerial regulation enforcing the comprehensive plan of Chiang Mai City (3rd revision), B.E. 2555*. Government Gazette Vol. 129, Part 44 A.

[9] Choowong, M., Pailoplee, S., Ketthong, C., Udomsak, S., Choowong, N. (2025). Remarkable destruction of the September 2024 heavy rainfall-induced landslide-mudflow and flash flood on major cities in the northern Thailand-Myanmar border after the super typhoon Yagi. *Landslides*, 22(9):2969–2977.

[10] Panyadee, P., Champrasert, P. (2024). Spatiotemporal flood hazard map prediction using machine learning for a flood early warning case study: Chiang Mai Province, Thailand. *Sustainability*, 16(11):4433.
[11] Gu, Z., Phakdimek, S., Nagami, K., Komori, D. (2025). Relationship between urbanization–induced land use changes and flood risk: Case study in Chiang Mai, Thailand. *Water*, 17(3):327.
[12] Dharmarathne, G., Waduge, A. O., Bogahawaththa, M., Rathnayake, U., Meddage, D. P. P. (2024). Adapting cities to the surge: A comprehensive review of climate-induced urban flooding. *Results in Engineering*, 22:102123.
[13] Oliveira de Sousa, F., Ariza Flores, V. A., Cunha, C. S., Oda, S., Xavier Ratton Neto, H. (2025). Multicriteria assessment of flood risk on railroads using a machine learning approach: A case study of railroads in Minas Gerais. *Infrastructures*, 10(1):12.
[14] Sresakoolchai, J., Hamarat, M., Kaewunruen, S. (2023). Automated machine learning recognition to diagnose flood resilience of railway switches and crossings. *Scientific Reports*, 13(1):2106.
[15] Pregnolato, M., Ford, A., Wilkinson, S. M., Dawson, R. J. (2017). The impact of flooding on road transport: A depth-disruption function. *Transportation Research Part D: Transport and Environment*, 55: 67–81.
[16] Koks, E. E., Rozenberg, J., Zorn, C., Tariverdi, M., Vousdoukas, M., Fraser, S. A., Hall, J. W., Hallegatte, S. (2019). A global multi-hazard risk analysis of road and railway infrastructure assets. *Nature Communications*, 10(1):2677.
[17] Al-Aizari, A. R., Alzahrani, H., AlThuwaynee, O. F., Al-Masnay, Y. A., Ullah, K., Park, H.-J., Al-Areeq, N. M., Rahman, M., Hazaea, B. Y., Liu, X. (2024). Uncertainty reduction in flood susceptibility mapping using Random Forest and eXtreme Gradient Boosting algorithms in two tropical desert cities, Shibam and Marib, Yemen. *Remote Sensing*, 16(2):336.
[18] Bentivoglio, R., Isufi, E., Jonkman, S. N., Taormina, R. (2022). Deep learning methods for flood mapping: A review of existing applications and future research directions. *Hydrology and Earth System Sciences*, 26(16):4345–4378.
[19] Krishnachaithanya, N., Bagavathi Sivakumar, P., Deepika, T. (2025). Probabilistic flood risk mapping for Indian megacities using a scalable machine learning framework in Google Earth Engine. *Results in Engineering*, 28:107872.
[20] Alexopoulos, M. J., Iliopoulou, T., Modé, P., Istrati, D., Koutsoyiannis, D., Królewicz, S., Graf, R., Kaczmarek, L., Rączkowski, W. (2025). Flood risk assessment of cultural heritage sites near lakes via advanced hydrodynamic modeling and digital technologies. *Results in Engineering*, 28:107977.
[21] Wu, X., Zhang, Z., Xiong, S., Zhang, W., Tang, J., Li, Z., An, B., Li, R. (2023). A near-real-time flood detection method based on deep learning and SAR images. *Remote Sensing*, 15(8):2046.
[22] Wienhold, K. J., Li, D., Li, W., Fang, Z. N. (2023). Flood inundation and depth mapping using unmanned aerial vehicles combined with high-resolution multispectral imagery. *Hydrology*, 10(8):158.
[23] Wüthrich, D., Korswagen, P. A., Selvam, H., Oetjen, J., Bricker, J., Schüttrumpf, H. (2024). Field survey assessment of flood loads and related building damage from the July 2021 event in the Ahr Valley (Germany). *Journal of Flood Risk Management*, 18(1):e13024.
[24] Khosravi, K., Shahabi, H., Pham, B. T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.-B., Gróf, G., Ho, H. L., Hong, H., Chapi, K., Prakash, I. (2019). A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods. *Journal of Hydrology*, 573:311–323.
[25] Notti, D., Giordan, D., Caló, F., Pepe, A., Zucca, F., Galve, J. P. (2018). Potential and limitations of open satellite data for flood mapping. *Remote Sensing*, 10(11):1673.
[26] Yaseen, Z. M. (2024). Flood hazards and susceptibility detection for Ganga River, Bihar state, India: Employment of remote sensing and statistical approaches. *Results in Engineering*, 21:101665.
[27] Pradhan, B., Lee, S., Dikshit, A., Kim, H. (2023). Spatial flood susceptibility mapping using an explainable artificial intelligence (XAI) model. *Geoscience Frontiers*, 14(6):101625.
[28] Jodhani, K. H., Patel, D., Madhavan, N., Gupta, N., Singh, S. K., Rathnayake, U. (2024). Unravelling flood risk in the Rel River watershed, Gujarat using coupled earth observations, multi criteria decision making and Google Earth Engine. *Results in Engineering*, 24:102836.

[29] Chowdhuri, I., Pal, S. C., Chakrabortty, R. (2020). Flood susceptibility mapping by ensemble evidential belief function and binomial logistic regression model on river basin of eastern India. *Advances in Space Research*, 65(5):1466–1489.

[30] Damayanti, H. N., Wikan, P. A., Annurhutami, F. (2024). GIS-based flood susceptibility mapping in Wawar Watershed, Purworejo Regency. *IOP Conference Series: Earth and Environmental Science*, 1314 (1):012051.

[31] Hong, H., Panahi, M., Shirzadi, A., Ma, T., Liu, J., Zhu, A. X., Chen, W., Kougias, I., Kazakis, N. (2018). Flood susceptibility assessment in Hengfeng area coupling adaptive neuro-fuzzy inference system with genetic algorithm and differential evolution. *Science of The Total Environment*, 621:1124–1141.

[32] Liu, B., Li, Y., Ma, M., Mao, B. (2025). A comprehensive review of machine learning approaches for flood depth estimation. *International Journal of Disaster Risk Science*, 16(3):433–445.

[33] Yu, H., Luo, Z., Wang, L., Ding, X., Wang, S. (2023). Improving the accuracy of flood susceptibility prediction by combining machine learning models and the expanded flood inventory data. *Remote Sensing*, 15(14):3601.

[34] Khalid, R., Khan, U. T. (2024). Flood susceptibility mapping using ANNs: A case study in model generalization and accuracy from Ontario, Canada. *Geocarto International*, 39(1):2316653.

[35] Wang, H., Meng, Y., Xu, H., Wang, H., Guan, X., Liu, Y., Liu, M., Wu, Z. (2024). Prediction of flood risk levels of urban flooded points though using machine learning with unbalanced data. *Journal of Hydrology*, 630:130742.

[36] Ekmekcioğlu, Ö., Koc, K., Özger, M., Işık, Z. (2022). Exploring the additional value of class imbalance distributions on interpretable flash flood susceptibility prediction in the Black Warrior River basin, Alabama, United States. *Journal of Hydrology*, 610:127877.

[37] Rahmati, O., Kornejady, A., Samadi, M., Deo, R. C., Conoscenti, C., Lombardo, L., Dayal, K., Taghizadeh-Mehrjardi, R., Pourghasemi, H. R., Kumar, S., Bui, D. T. (2019). PMT: New analytical framework for automated evaluation of geo-environmental modelling approaches. *Science of The Total Environment*, 664:296–311.

[38] Sitthikankun, S., Chakarin, B., Srivanit, M., Rinchumphu, D. (2024). Identifying flood exposure factors of flooding prediction through spatial analysis and remote sensing. In *IWA 21st International Conference on Navigating Diffuse Pollution Pathways to Carbon Neutrality*, December 11–14, 2024. The Empress Premier Hotel, Chiang Mai, Thailand.

[39] Rentschler, J., Avner, P., Marconcini, M., Su, R., Strano, E., Vousdoukas, M., Hallegatte, S. (2023). Global evidence of rapid urban growth in flood zones since 1985. *Nature*, 622(7981):87–92.

[40] Tang, J., Liu, D., Shang, C., Niu, J. (2024). Impacts of land use change on surface infiltration capacity and urban flood risk in a representative karst mountain city over the last two decades. *Journal of Cleaner Production*, 454:142196.

[41] Gohil, M., Mehta, D., Shaikh, M. (2024). An integration of geospatial and fuzzy-logic techniques for flood-hazard mapping. *Journal of Earth System Science*, 133(2):80.

[42] A. Saleh, R., Al-Areeq, A. M., Al Aghbari, A. A., Ghaleb, M., Benaafi, M., Al-Areeq, N. M., M. Al-Ramadan, B. (2024). A novel voting ensemble model empowered by metaheuristic feature selection for accurate flash flood susceptibility mapping. *Geomatics, Natural Hazards and Risk*, 15(1):2360000.

[43] Abdo, H. G., Richi, S. M., Zeng, T., Alqadhi, S., Prasad, P., Nguyen-Huy, T., Alharbi, M. M., Mallick, J. (2025). Machine learning-based assessment of flood susceptibility in the Eastern Mediterranean: A case study of Baniyas River basin. *Geomatics, Natural Hazards and Risk*, 16(1):2524417.

[44] Myrzatay, A., Rzayeva, L., Bandini, S., Shayea, I., Saoud, B., Çolak, I., Kayisli, K. (2024). Predicting LAN switch failures: An integrated approach with DES and machine learning techniques (RF/LR/DT/SVM). *Results in Engineering*, 23:102356.

[45] Usman-Hamza, F. E., Balogun, A. O., Mamman, H., Capretz, L. F., Basri, S., Oyekunle, R. A., Mojeed, H. A., Akintola, A. G. (2025). Empirical analysis of data sampling-based decision forest classifiers for software defect prediction. *Software*, 4(2):7.

[46] Verma, S., Mehta, D., Pandey, A., Malani, S., Pandey, R. (2024). Flood hazard: A QGIS plugin for assessing flood consequences. *Journal of Water Management Modeling*, 32.

[47] Center of Excellence in Natural Disaster Management (2024). *CM Flood Mark 2024*.
[48] Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., Czerwinski, W., Pasquarella, V. J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., Tait, A. M. (2022). Dynamic World, near real-time global 10 m land use land cover mapping. *Scientific Data*, 9(1):251.
[49] Mulyana, I. S., Suryawan, B., Yamin, M., Juarna, A. (2025). CNN-based prediction of aerodynamic coefficients for airfoils at low Reynolds numbers. *Ingénierie des Systèmes d'Information*, 30(11):3055–3065.
[50] Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8(1):25.
[51] Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.
[52] Rahman, Z. U., Zhang, M., Chen, F., Ullah, S., Ahmad, M., Feroz, A., Danhassan, S. S., Azzam, A. (2025). Flood susceptibility mapping using supervised machine learning models: Insights into predictors' significance and models' performance. *Geomatics, Natural Hazards and Risk*, 16(1):2516728.
[53] Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8):913–929.