https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Discussions

Elevating Flash Flood Prediction Accuracy: A Synergistic Approach with PSO and GA Optimization

Duc-Vinh Hoang[1] and Yuei-An Liou[2*]

[1]The National Key Laboratory of River and Coastal Engineering, Vietnam Academy for Water Resources, 171 Tay Son Street, Dongda District, Hanoi 100000, Vietnam
[2]Center for Space and Remote Sensing Research, National Central University, 300 Jhongda Rd., Jhongli District, Taoyuan City 320317, Taiwan

Correspondence to: Yuei-An Liou (yueian@csrsr.ncu.edu.tw)

Abstract. Flash floods are frequent and devastating natural disasters in small mountainous river basins worldwide, causing significant harm to people, infrastructure, and property. Flash flood susceptibility mapping is a crucial tool for damage prevention and reduction. This study is focused on the creation of flash flood susceptibility maps in a mountainous region in northern Vietnam. We enhanced the performance of robust machine learning models, including Support Vector Machines (SVM), Random Forests (RF), and XGBoost (XGB), by applying advanced optimization techniques such as Particle Swarm Optimization (PSO) and Genetic Algorithms (GA). These models were developed based on 14 key factors, including elevation, slope, aspect, curvature, topographic wetness index (TWI), stream power index (SPI), flow accumulation, river density, distance to the river, NDVI, land use/land cover (LULC), rainfall, soil type, geology, and 412 flood inventory points. Nine models were tested, including three standalone ML algorithms (SVM, RF, XGB), three ensemble models optimized with PSO (PSO-SVM, PSO-RF, PSO-XGB), and three optimized with GA (GA-SVM, GA-RF, GA-XGB). The results indicated that ensemble models outperformed standalone ones, with the PSO-XGB, GA-XGB, and GA-RF models exhibiting outstanding performance, achieving accuracy rates of 0.939, 0.927, and 0.933, along with remarkable AUC-ROC scores of 0.957, 0.968, and 0.977, respectively. This innovative study introduces a novel set of associative models, contributing significantly to the advancement of flood prediction techniques. The methodology holds applicability for various regions characterized by similar topographical and climatic attributes. Furthermore, enhancing the precision of flood forecasting contributes to the formulation of mitigation strategies by municipal authorities to mitigate prospective flood-related impacts.

Keywords: Flash flood; optimize ML models; PSO, GA, Vietnam

1 Introduction

A flash flood is defined as a short, abrupt, and high-volume local flood, occurring within a limited timeframe, typically less than six hours, and precipitated by intense rainfall, rapid snowmelt due to temperature fluctuations or rain on snow, as well as the abrupt discharge of water resulting from dam breaches, levee collapses, or obstructions caused by ice jams (WMO and GWP, 2012). A flashflood involves a rapid surge in river and stream water levels, often accompanied by high flow speeds and

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

significant debris accumulation, with the force of the flood capable of dislodging boulders, uprooting trees, and demolishing structures in its course (WMO, 2007). Flash floods induce erosional processes, leaching, and landslides, leading to rocky mudflows characterized by exceptionally high velocities, causing substantial destruction within mountainous regions and valley environments. These events are primarily prevalent in elevated terrains characterized by steep slopes, significantly harming human populations and infrastructure (Sˇpitalar et al., 2014). Being among the most typhoon-prone nations in Asia, Vietnam has become especially susceptible to the threats of typhoon-driven flooding (Nguyen et al., 2019). According to the report of the Department of Natural Disaster Prevention and Control of Vietnam (DNDPC, 2019), flash floods have been documented across most of the country's mountainous and midland regions. Over the period spanning from 2000 to the present, an annual average of 15 to 16 flash flood events has been recorded, resulting in substantial harm to human lives, residential structures, critical infrastructure, and agricultural lands.

Forecasting and early warning of flash floods is a key component of risk reduction strategies used in flood management. However, despite recent developments in measurement, forecasting, and nowcasting methodologies, hydrologists and meteorologists claim that it is now possible to accurately predict the location and timing of flood events just one hour before they occur (Braud et al., 2018; Collier, 2007). This suggests that the current predictive methods might not account for all types of flash floods, particularly those brought on by anthropogenic or natural infrastructure failures, making it difficult to provide accurate and timely warnings in all circumstances. Moreover, flash floods are common in modest mountainous catchments characterized by a deficiency in hydrological measurements and observational data. This paucity of information presents difficulties in conducting precise hydraulic simulations for flash flood prediction (Braud et al., 2018). Consequently, the creation of flash flood susceptibility maps has emerged as a viable and supplementary strategy to inform decision-making procedures, encompassing zoning regulations, infrastructure expansion, and evacuation strategies, aiming to ameliorate flash flood consequences and reinforce resilience in regions prone to flooding.

In contemporary times, there has been a notable surge in the utilization of Remote Sensing (RS) and Geographic Information Systems (GIS) for the development of hazard and flood susceptibility maps (Chapi et al., 2017; Dao and Liou, 2015; Wang et al., 2022). These maps are constructed based on forecasting flash flood likelihood, which is accomplished by identifying the relationship between historical flood occurrences and influential variables. Typically, this relationship is elucidated through a combination of subjective and numerical methodologies (Tien Bui et al., 2019). Subjective approaches normally depend on the knowledge and insights of domain specialists, like Analytic Hierarchy Process (AHP) (Liou et al., 2017; Nachappa et al., 2020) or Fuzzy logic (Costache et al., 2020). Numerical methodologies include statistical approaches like the Frequency Ratio (FR) technique (Cao et al., 2016) or Weight of Evidence (WoE) (Costache et al., 2020). In recent years, Machine Learning (ML) techniques have drawn a lot of interest and broader utilization in the field of flood prediction.

Recent studies have underscored the effectiveness of ML models in delineating regions prone to flooding. Well-established machine learning algorithms have been widely employed and have demonstrated encouraging outcomes, including Support Vector Machines (SVM) (Madhuri et al., 2021; Roy et al., 2020), Naïve Bayes (Chen et al., 2020), Random Forests (RF) (Vinh and Liou, 2024), Adaptive Neuro-Fuzzy Inference Systems (ANFIS) (Arora et al., 2021; Nguyen, 2023), and XGBoost (Abedi

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

et al., 2021; Zhang et al., 2019). Comparative evaluations have revealed that ML-based techniques surpass subjective methods, like AHP and Fuzzy Logic, as well as statistical approaches, including FR and WoE, in terms of flood prediction accuracy. Floods and flash floods represent intricate hydrological phenomena influenced by multiple natural, environmental, and human factors. Consequently, enhancing prediction accuracy necessitates the utilization of nonlinear and dynamic predictive models. Striving towards applying ML and hybrid ML models has yielded promising outcomes. One commonly employed technique for hybridizing ML involves optimizing algorithms to fine-tune the model's hyperparameters. Prominent optimization strategies, including Genetic Algorithms (GA) (Linh et al., 2022), Particle Swarm Optimization (PSO) (Bui et al., 2020; Razavi et al., 2018), Gradient-Based Optimizers (GBO) (Nguyen, 2023), and Differential Evolution (DE) (Nguyen, 2023), have been incorporated into these hybrid frameworks, yielding heightened accuracy in the domain of flood prediction.

The prevailing optimization techniques frequently employed are PSO and GA. As of October 2020, PSO and GA collectively constitute 78.77% of the nature-inspired optimization algorithms (NIOAs) utilized in research articles accessible within the WOS and Scopus databases (Wang et al., 2021). Overall, their global search capabilities, adaptability, versatility, and suitability for non-linear and high-dimensional optimization problems make PSO and GA popular choices for optimizing ML algorithms. Researchers often select these methods when seeking to enhance the performance of ML models in various applications.

Numerous endeavors in hydrology have focused on hybridizing ML methodologies, particularly concerning flood prediction. Nevertheless, the prevailing approach in these investigations is either employing a single optimization algorithm across various ML algorithms or utilizing multiple optimization algorithms for a single ML algorithm. In this research, our primary objective is to synergistically integrate PSO and GA with three robust ML algorithms, specifically SVM, RF, and XGB. The aim is to engineer hybrid models characterized by exceptional performance and an elevated capacity to forecast the onset of flash floods in the Song Ma district of the mountainous northern region of Vietnam. It is important to note that this study represents a pioneering effort in introducing a comprehensive suite of hybrid models, thus serving as a valuable complement to existing flood forecasting methodologies. By harnessing the combined strengths of these optimization techniques and ML algorithms, we aspire to significantly advance the accuracy and reliability of flash flood prediction within this geographically vulnerable area.

2. Study region and dataset

2.1 Study region

The study was undertaken within the locale of Song Ma, situated in the mountainous region of Vietnam northwestern sector (Figure 1). Encompassing an area of about 1640 square kilometres, the study site exhibits an elevation range from 300 meters to 1300 meters above sea level, characterized by an intricate topography featuring elevated mountain chains intertwined with valleys, rivers, and streams. An annual recurrence of floods and flash floods transpires from June to August, yielding considerable detriment to infrastructure, assets, and human lives. A particularly noteworthy occurrence transpired in September

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Discussions

1975 as a historic flash flood, leading to the substantial devastation of road and bridge networks, agricultural land being subjected to landslides and inundation, and swept of numerous residential structures. The resultant traffic disruption reached a near-complete standstill, necessitating aircraft deployment to distribute relief provisions to the local populace. Given the recurrent prevalence of flash floods, the Song Ma district emerges as a suitable and rational research domain for investigating the susceptibility to flash flood occurrences.

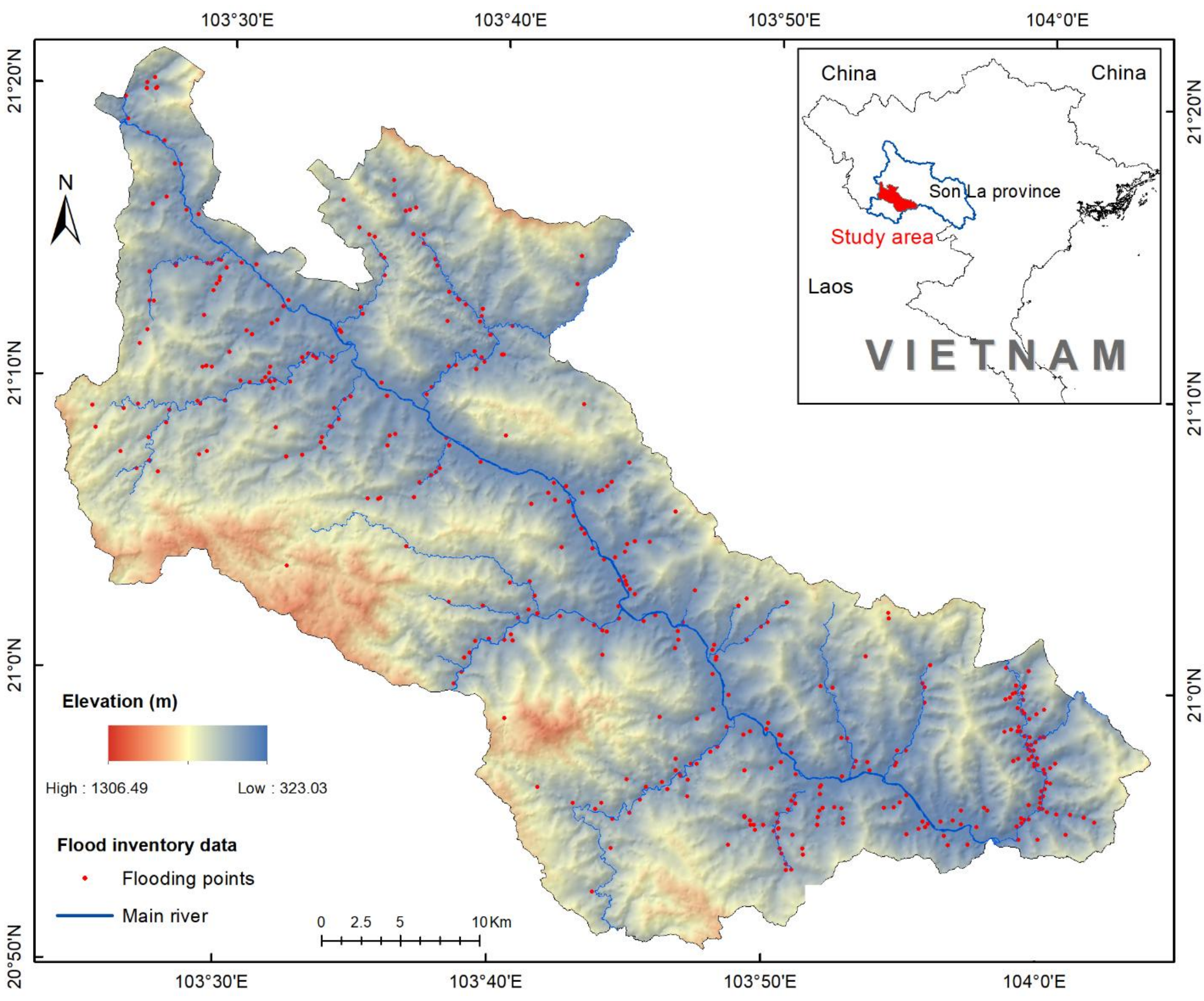


Figure 1: Location of the study area and flood inventory data

2.2 Inventory data on flash floods

The study used flash flood history data acquired from the Vietnam Disaster Management Authority (VDMA), complemented by field research conducted by the Vietnam Academy for Water Resources (VAWR). Encompassing the study region are a cumulative 412 sites associated with flash floods that transpired from 2001 to 2018. An equivalent set of 412 locations devoid

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

of flash floods was randomly designated to establish a binary classification model for flood events, ensuring methodological consistency and precision. Within this context, the non-flood locations were ascribed a numerical value of 0, whereas the flood locations were designated a value of 1. Afterwards, the 824 flood history sites were partitioned into a 70% subset for training and a 30% subset for testing, thereby serving as the input data for the proposed machine learning algorithms (Figure 1).

2.3 Factors influencing flash floods

The identification of flood-influencing factors constitutes a fundamental prerequisite for the creation of flood susceptibility maps. In research within a designated study area, determining influential factors is frequently based on prior inquiries, where significant variables are discerned owing to the unique geographical and environmental attributes that exhibit variation among different regions (Tehrany et al., 2019). These factors in our study were elevation, aspect, slope, curvature, stream density, distance to river, flow accumulation, topographic wetness index (TWI), stream power index (SPI), rainfall, LULC, and Normalized Difference Vegetation Index (NDVI). We used ArcGIS 10.5 to prepare the factors from different sources (Figure 1 and Table 1).

Physical factors related to the topographic surface, including elevation, slope, aspect, curvature, TWI, SPI, stream density, and distance to river, directly affect flood flows. Elevation exerts a significant influence on water retention capacity, and regions situated at lower elevations are inherently more susceptible to experiencing flood events. Slope dictates water movement from higher to lower elevations, thereby playing a pivotal role in the genesis of flooded areas. Aspect affects soil moisture, infiltration, and local climate conditions. The curvature assesses how the water is distributed across the surface. The SPI is a measure of the stability of the terrain and the possibility for stream erosion. Meanwhile, TWI indicates the gravity downward movement of water and makes it possible to estimate where water will accumulate (Moore et al., 1991). Stream density influences the capacity for runoff, the storage within floodplains, and the susceptibility to channel obstructions, thereby ultimately affecting the probability and level of flooding at a certain location. Flow accumulation is indicative of the aggregation of water from numerous contributing cells or regions within a watershed. Greater flow accumulation corresponds to a larger potential volume of water capable of contributing to flooding events (Santos et al., 2019). The distance to the river plays a crucial role in the propagation of floods. Close proximity to watercourses is frequently linked to heightened vulnerability to flooding, as water exceeds the natural boundaries of the riverbanks. These nine factors were extracted from the digital elevation model (DEM) in ArcGIS. We obtained a 10 * 10 m resolution DEM from a contour line map of 1:10,000 scale provided by the Ministry of Natural Resources and Environment of Vietnam (MONRE).

Rainfall, specifically intense or extended rain, assumes a pivotal role in the susceptibility of floods, as it can lead to heightened runoff, elevated water levels in river systems, and, consequently, an augmented probability of flooding occurrence within vulnerable regions. Here, we generated a spatial map representing rainfall distribution for the three peak months of the year, namely June, July, and August. The rainfall dataset was collected from a network of 11 rain gauge stations spanning 53 years from 1970 to 2022.

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

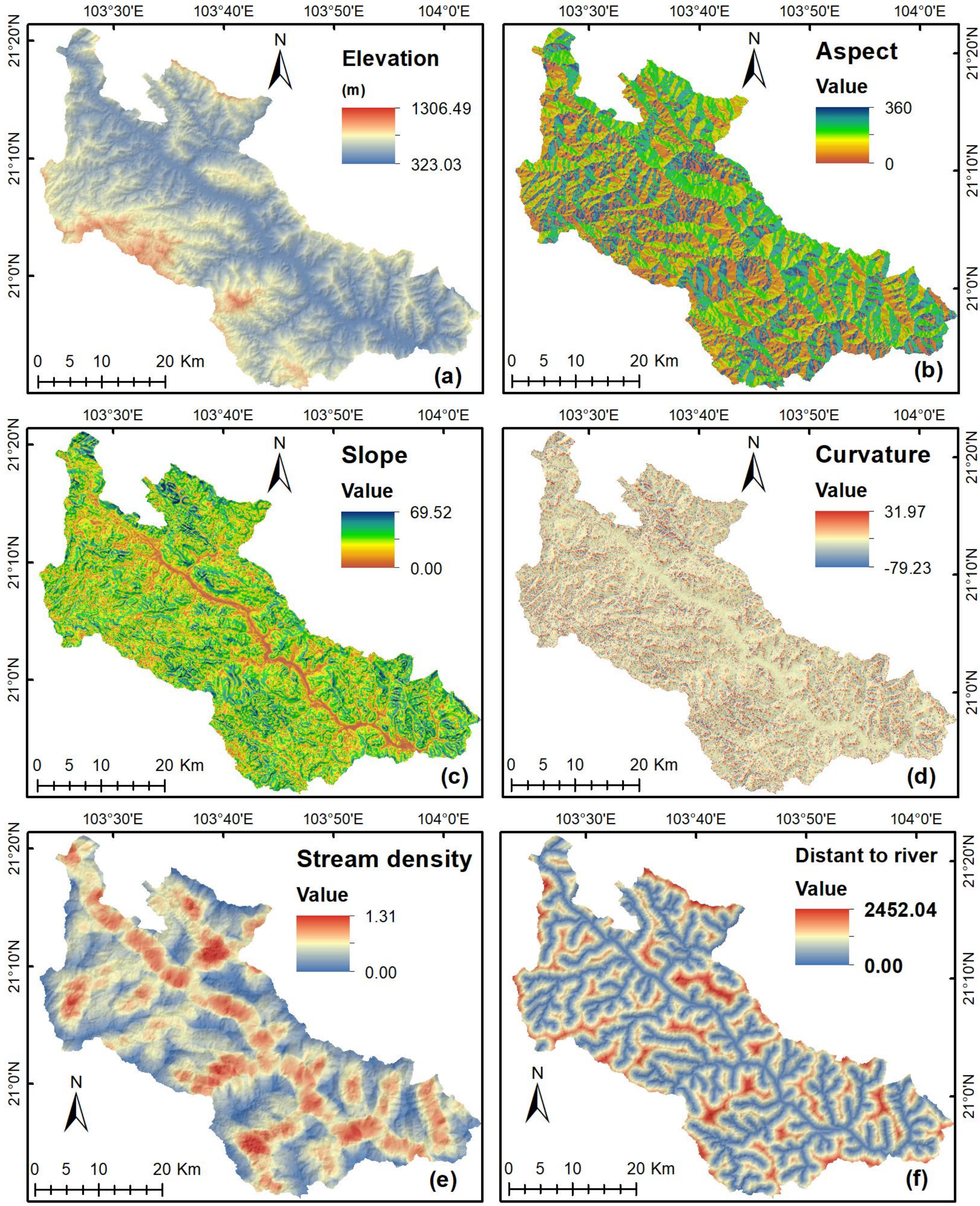
Soil type directly correlates with vulnerability to flooding, as it governs both the capacity for water infiltration and the terrain's permeability (Costache et al., 2020). Soil type with infiltration capacities have enhanced potential to absorb and retain water, thereby reducing the probability of flooding. Similarly, geology characteristics influence flood risk by influencing surface runoff patterns, permeability, and groundwater storage (Panahi et al., 2021). Places where the geology is impermeable are more prone to heightened flood susceptibility, whereas lithology with greater permeability can facilitate efficient water drainage, reducing flooding risks. Data on soil type and geology were taken from the digital maps at a scale of 1:100,000 provided by the Ministry of Agriculture and Rural Development of Vietnam (MARD) and the MONRE, respectively.

LULC significantly influences flood susceptibility through its impact on various hydrological factors (Talukdar et al., 2020). The nature of LULC in a region affects surface runoff, infiltration rates, drainage patterns, and the capacity for water storage (Razavi et al., 2018). Natural land covers, such as forests and wetlands, often have better infiltration rates and act as buffers against excessive runoff. In contrast, urban areas with impervious surfaces increase surface runoff, potentially overwhelming drainage systems and leading to localized flooding. Moreover, changes to natural drainage patterns and modifications in river courses can disrupt the natural flow of water, further contributing to flood susceptibility. The type and arrangement of land use directly shape the hydrological response of an area and play a crucial role in determining the likelihood and severity of flooding events. Here, LULC (2017) was extracted from the 1:10,000 scale digital maps provided by MONRE.

The NDVI exhibits a robust correlation with the occurrence of flash floods (Dodangeh et al., 2020). High NDVI values consistently signify a greater quantity of vegetation cover, facilitating improved soil infiltration capacities and reduced surface runoff, consequently lowering the likelihood of flash floods (Shafizadeh et al., 2018). Conversely, low NDVI values signify sparse vegetation or unvegetated land, resulting in diminished water absorption capabilities and increased susceptibility to surface runoff, heightening the risk of flash flood incidents. In this study, NDVI was extracted from Sentinel 2 satellite images. Sentinel 2 images with a cloud ratio lower than 5% were composited, and the median value for the whole year of 2017 and calculated the NDVI on the GEE platform.

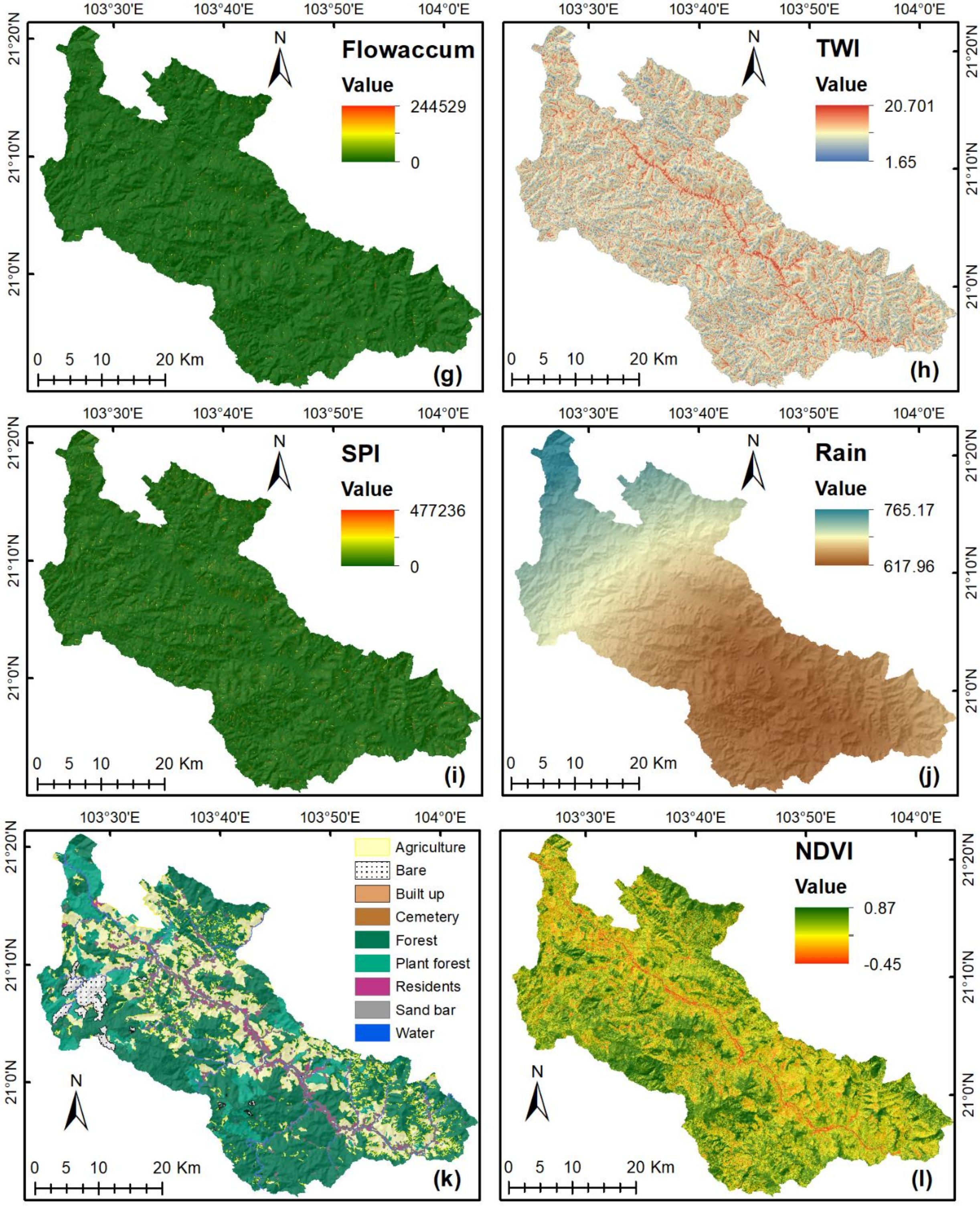
https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Discussions



https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions



165

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Discussions

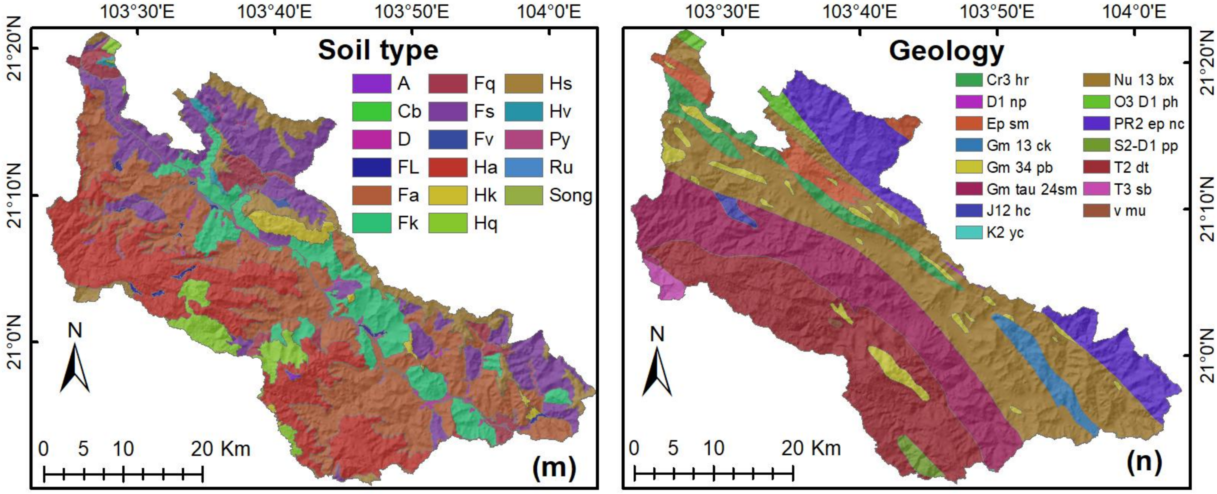


Figure 2: Flash flood influencing factors: (a) Elevation, (b) Aspect, (c) Slope, (d) Curvature, (e) Stream density, (f) Distance to river, (g) Flow accumulation, (h) TWI, (i) SPI, (j) Rainfall, (k) LULC, (l) NDVI, (m) Soil type, and (n) Geology

Table 1. Sources and type of influencing factors

Factors	**Type**	**Derived from**	**Resolution /scale**	**Sources**
Elevation	Raster	DEM	10m	MONRE
Slope	Raster	DEM	10m	MONRE
Aspect	Raster	DEM	10m	MONRE
Curvature	Raster	DEM	10m	MONRE
TWI	Raster	DEM	10m	MONRE
SPI	Raster	DEM	10m	MONRE
Stream density	Raster	DEM	10m	MONRE
Distance to river	Raster	DEM	10m	MONRE
Flow accumulation	Raster	DEM	10m	MONRE
Rainfall	Raster	Rainfall gauges	10m	CHIRPS
Soil type	Polygon	Digital soil map	1:100,000	MARD
Geology	Polygon	Digital lithology map	1:100,000	MONRE
LULC	Polygon	Digital land cover map	1:10,000	MONRE
NDVI	Raster	Sentinel 2	10m	Copernicus

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

3. Methodology

Our approach encompassed four steps: (i) input data preprocessing, (ii) Machine learning modeling, (iii) model evaluation, and (iv) flash flood susceptibility mapping (0)

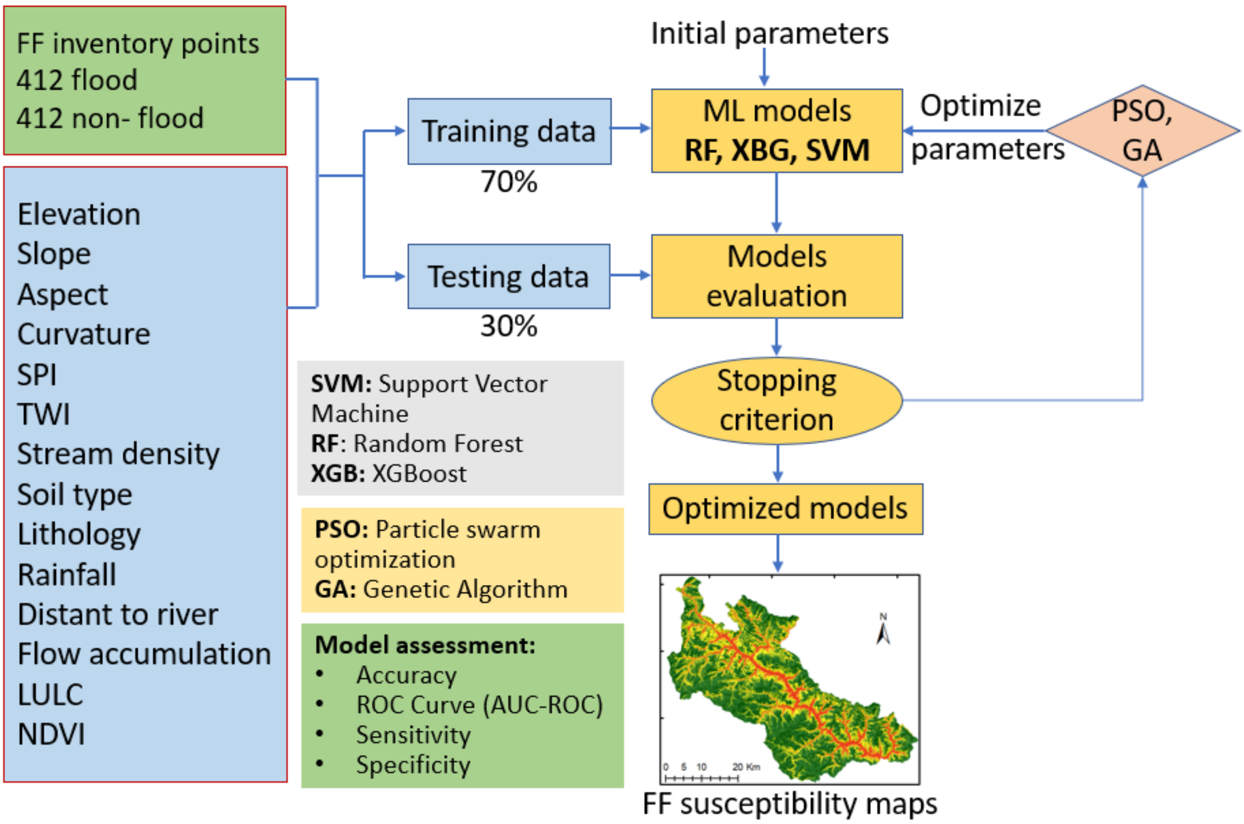


Figure 3: Methodology flowchart

3.1. Input data preprocessing

We acquired and treated unprocessed data, comprising Digital Elevation Model (DEM), land cover details, soil categorization, geological mappings, precipitation records, Sentinel 2 imagery, and historical flash flood archives. These datasets produced an inventory map and identified fourteen factors impacting flash floods. Before model development, the dataset underwent preprocessing steps, including assessing multicollinearity, normalization of variables, and division into training (70%) and testing (30%) subsets.

3.2 Machine learning modeling

In the study, three well-established machine learning techniques known for their strong capability in predicting flash flood susceptibility, namely Support Vector Machine (SVM) (Roy et al., 2020), Random Forest (RF) (Abedi et al., 2021; Chen et

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

al., 2020), and Extreme Gradient Boosting (XGB) (Linh et al., 2022), were employed. To enhance the predictive power of each model, advanced optimization algorithms, including PSO and GA, were utilized to fine-tune the hyperparameters specific to each machine learning approach. GA and PSO were chosen because they are the most widely used optimization algorithms in scientific research, with a total annual publication of 2051 papers for GA and 1346 papers for PSO (Wang et al., 2021). The underlying operational principles of these methods are elaborated upon in the subsequent sections.

3.2.1. Support Vector Machine (SVM)

SVM represents a machine learning technique designed to discover the most suitable hyperplane for to separate data elements that belong to different classes. The process involves transforming the input data into a feature space with higher dimensions, where the SVM determines a hyperplane which optimally splits these categories. SVM is a flexible approach, proficient in addressing linear and non-linear classification challenges, facilitated by its utilization of diverse kernel functions (Boser et al., 1992). The significance of SVM in flood susceptibility evaluation originates from its ability to manage non-linear associations and mitigate the risk of overfitting, rendering it an invaluable resource for studying and projecting flood-prone areas (Roy et al., 2020).

3.2.2. Random Forest (RF)

RF is a ML approach that utilizes a collection of decision trees to make predictions. It excels in handling large datasets, capturing complex relationships, and mitigating the risk of overfitting. The utilization of RF in evaluating and predicting regions prone to flooding has become increasingly popular (Abedi et al., 2021; Chen et al., 2020; Roy et al., 2020). Its effectiveness is notable due to its ability to incorporate various influential factors, handle non-linear relationships, and produce accurate and reliable results. By conducting thorough analyses of the inputs, RF allows for the recognition of critical parameters that significantly contribute to the vulnerability to floods.

3.2.3. Extreme Gradient Boosting (XGB)

XGB is a potent ML algorithm extensively employed for for classifying and flood susceptibility applications. Its operation involves the sequential construction of a chain of decision trees, with every successive tree rectifying the errors introduced by its predecessors (Chen and Guestrin, 2016). XGB seeks to optimize a designated loss function to find the best divides that reduce the total forecast mistake. To reduce the risk of overfitting, XGB uses regularization approaches such shrinking through the learning rate and feature subsampling. Moreover, it employs gradient boosting, meaning every tree receives instruction to adjust any residuals left by previous trees, so addressing disparities between anticipated and real value.

The iterative procedure continues until an exact amount of trees is met, leading to a remarkably precise and dependable classification model. In terms of flood susceptibility assessments, XGB adeptly captures intricate and non-linear relationships among influencing variables and flood incidents. Its application in many research endeavors has consistently yielded precise, robust, and trustworthy outcomes when forecasting regions prone to flooding (Abedi et al., 2021; Linh et al., 2022).

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

3.2.4. Particle Swarm Optimization (PSO)

PSO is a powerful optimization technique that can be applied to fine-tune the hyperparameters of machine learning models in flood susceptibility prediction (Razavi et al., 2018). A fundamental characteristic of PSO lies in its emulation of the behaviors exhibited by biological organisms, such as birds, fish, and ants. PSO operates by simulating the behavior of a swarm of particles moving through a multidimensional parameter space, searching for the optimal set of hyperparameters that minimizes a predefined objective function, typically related to the model's performance (Kennedy and Russell, 1995).

For SVM, PSO can explore the hyperparameter space of SVM, including the selection of kernel, regularization parameters, and other settings. By adjusting these hyperparameters, PSO aims to find the optimal configuration that maximizes the separation of flood and non-flood locations. This results in a more accurate SVM model for flood susceptibility assessment.

For RF, PSO can tune parameters like the number of trees in the forest, the depth of individual trees, and other settings that influence the model's performance. By optimizing these parameters, PSO helps RF better capture the complex relationships between various factors and flood occurrences, improving accuracy in identifying regions susceptible to flooding.

For XGB, PSO can fine-tune the hyperparameters of XGB, such as the learning rate, the number of boosting rounds, the depth of trees, and other relevant settings. By optimizing these parameters, PSO enhances XGB's ability to handle non-linear associations, correct errors sequentially, and effectively capture the complexity of flood-related factors. This leads to more precise and reliable predictions of flood-prone areas.

3.2.5. Genetic Algorithms (GA)

The GA concept, drawn from nature, has been extensively applied to addressing optimization challenges and learning procedures (Sheta, 2006). In the natural world, improved combinations of chromosomes lead to more advanced generations, with occasional mutations enhancing subsequent iterations. The genetic algorithm employs a similar approach to tackle optimization problems, leveraging the idea of refining solutions through successive generations, including beneficial mutations, for improved outcomes. They initialize a population of potential solutions, evaluate their fitness using a defined objective function, select the best-performing solutions for reproduction (crossover), introduce small changes (mutation) to create diverse offspring, and then replace the old generation with the new. This iterative process continues until a termination criterion is met, yielding an optimal or near-optimal solution. GAs effectively explore complex solutions, making them valuable for solving optimization problems in various fields by adapting and refining candidate solutions through generations, ultimately improving their quality based on fitness evaluation (Sivanandam and Deepa, 2008).

GA can optimize the hyperparameters of ML algorithms for flash flood susceptibility prediction through an evolutionary approach. By representing hyperparameters as individuals in a population, GA performs selection, crossover, and mutation operations to evolve and refine the parameter combinations, aiming to improve the models' performance in identifying flood-prone regions. This iterative process allows GA to navigate the hyperparameter space effectively, leading to more accurate

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Discussions

and reliable models by adjusting key parameters that influence the models' ability to handle non-linear relationships, overfitting, and other factors relevant to flood susceptibility assessment.

3.3. Model evaluation

We subjected all models to fitting and cross-validation processes. They were assessed by comparing the performance of each algorithm using four common metrics: Accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC-ROC) (equations (1), (2), (3), (4)). Accuracy pertains to the proportion of occurrences accurately categorized as flash floods or non-flash floods. The sensitivity indicates the portion of occurrences correctly identified as flash floods, while specificity indicates the percentage of incidents accurately identified as non-flash floods. AUC-ROC refers to the overall performance of the predictive model in distinguishing between flood or non-flood. A higher AUC-ROC value implies that the model is better at correctly identifying locations likely to experience flash floods than areas less prone to this type of natural disaster.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{FP + TN} \tag{3}$$

$$AUC = \sum \mathrm{TP} + \sum \frac{TN}{P} + N \tag{4}$$

where TP and TN represent true positives and true negatives, respectively; FP and FN denote false positives and false negatives; and P and N refer to the quantities of flood and non-flood pixels.

4. Results

4.1. The significance of flash flood influencing factors

Before creating a flash flood susceptibility map, it is necessary to conduct a sensitivity analysis to ensure the model balance and emphasize the importance of influencing factors (Nguyen, 2023). Figure 4 depicts the outcomes of the sensitivity analysis and the importance of influencing factors using the RF algorithm. The findings indicated that elevation (0.216) and distance to the river (0.207) were most strongly correlated with flood incidence. The significance of LULC (0.104), stream density (0.088), slope (0.082), NDVI (0.069), and curvature (0.049) fell within a moderate range. While TWI (0.039), rainfall (0.032), aspect (0.028), soil type (0.023), flow accumulation (0.023), SPI (0.022), and geology (0.016) had the least importance.

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

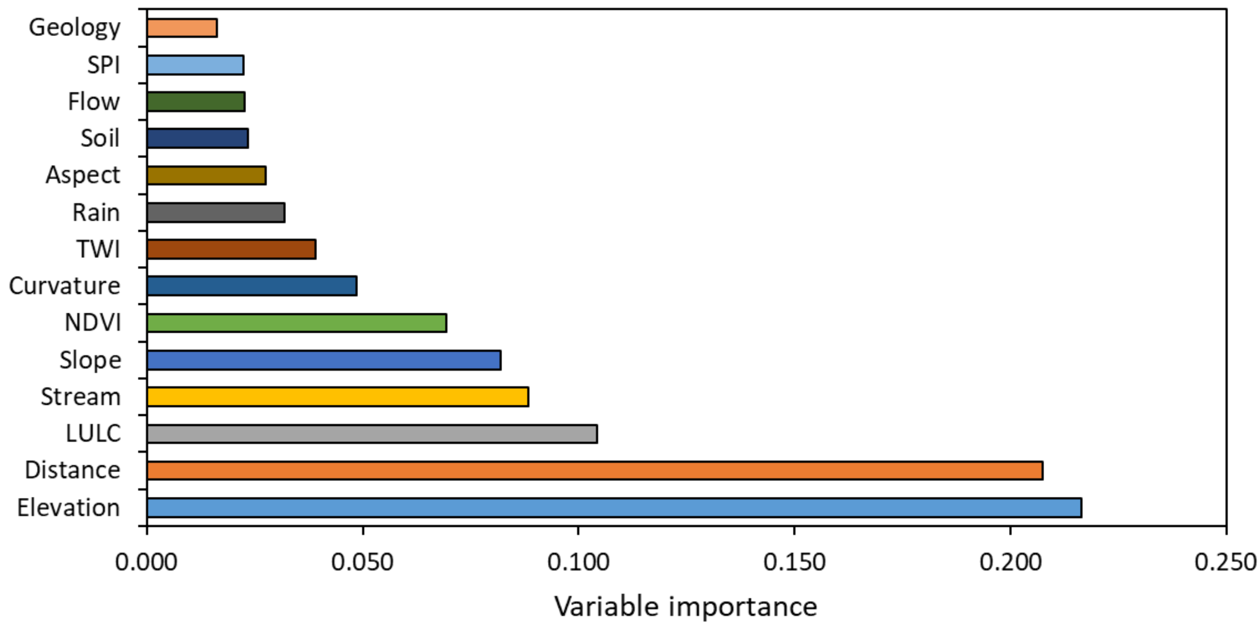


Figure 4: The importance of influencing factors using the RF model

4.2. Models performance

A total of nine models were formed by combining three machine learning algorithms (SVM, RF, XGB) with two optimization algorithms (PSO, GA). These nine models were then evaluated using accuracy, sensitivity, specificity, and AUC-ROC metrics on both training and testing datasets. The performance outcomes of these models are depicted in Figure 5 and Figure 6, as well as summarized in Table 2.

Accuracy: The lowest accuracy of 0.861 was observed in the SVM model, while the highest accuracy of 0.939 was achieved by the PSO-XGB model.

AUC-ROC: The lowest AUC-ROC value of 0.932 was seen in the SVM model, while the highest AUC-ROC value of 0.977 was attained by the Genetic Algorithm-optimized Random Forest (GA-RF) model.

Sensitivity: The lowest sensitivity score of 0.833 was observed in the Random Forest (RF) model, while the highest sensitivity score of 0.964 was achieved by the Genetic Algorithm-optimized Random Forest (GA-RF) model.

Specificity: The lowest specificity score of 0.851 was seen in the SVM model, while the highest specificity score of 0.920 was attained by the Genetic Algorithm-optimized Support Vector Machine (GA-SVM) and Genetic Algorithm-optimized Extreme Gradient Boosting (GA-XGB) models.

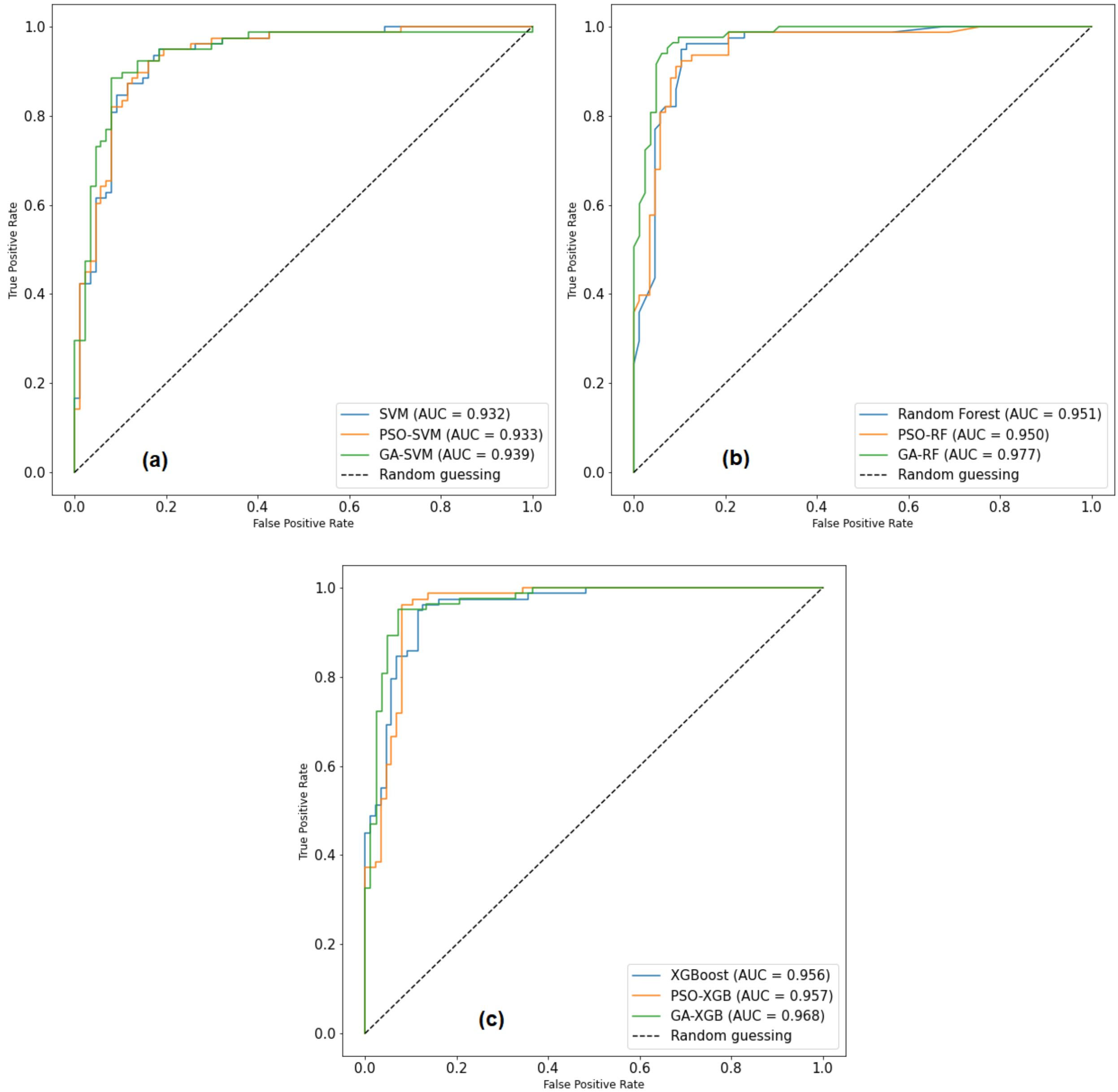


Figure 5: The AUC-ROC of models in testing dataset. (a) SVM and optimized SVM models; (b) RF and optimized RF models; (c) XGB and optimized XGB models

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Discussions

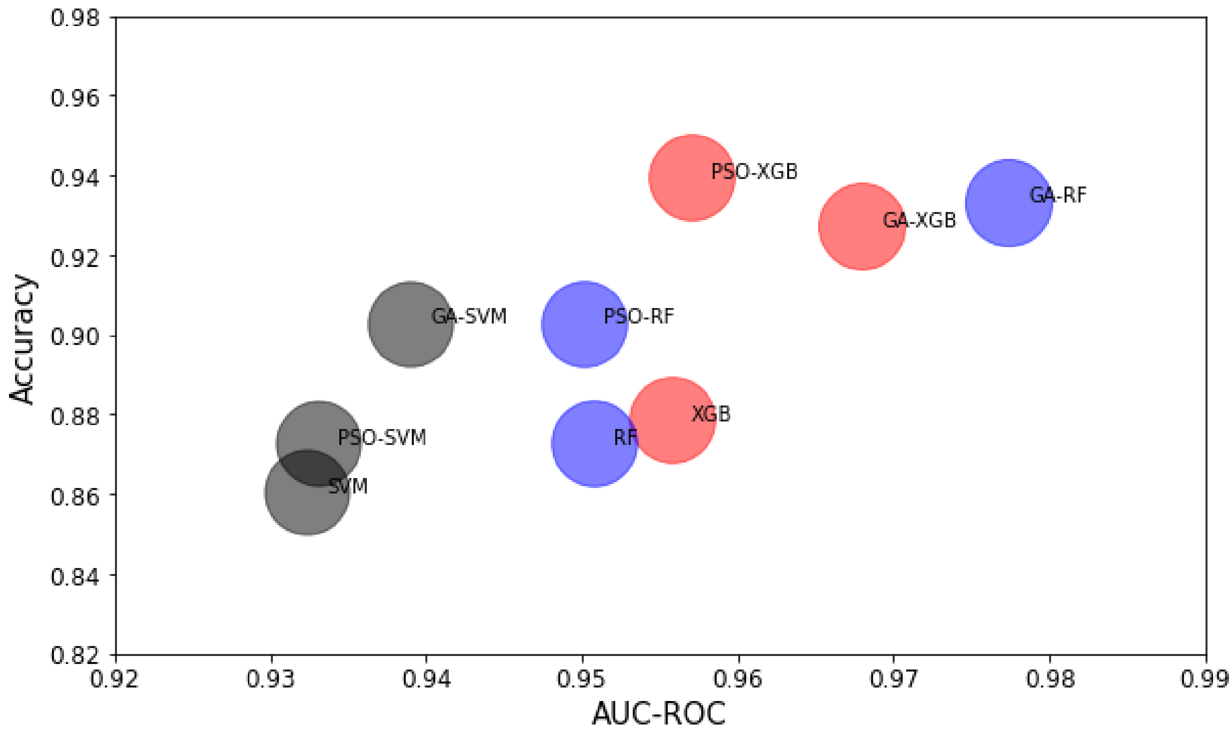


Figure 6: AUC-ROC and Accuracy scatter plot of difference models

290 **Table 2. The performance of each model**

Models	Accuracy	AUC-ROC	Sensitivity	Specificity
SVM	0.861	0.932	0.872	0.851
PSO-SVM	0.873	0.933	0.872	0.874
GA-SVM	0.903	0.939	0.885	0.920
RF	0.873	0.951	0.833	0.908
PSO-RF	0.903	0.950	0.885	0.920
GA-RF	0.933	0.977	0.964	0.902
XGB	0.879	0.956	0.872	0.885
PSO-XGB	0.939	0.957	0.962	0.920
GA-XGB	0.927	0.968	0.928	0.927

In summary, the evaluation of different models revealed that the GA-RF model consistently achieved the highest performance in terms of AUC-ROC, and sensitivity, while the SVM model exhibited the lowest performance across these metrics. However, the Genetic Algorithm-optimized SVM (GA-SVM) and XGB (GA-XGB) models consistently achieved the highest specificity scores among the evaluated models. These disparities highlight the varying effectiveness of different models and optimization
295 techniques in flood susceptibility prediction across other key evaluation metrics.

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

4.3. Flash flood susceptibility maps

The study employed nine models, SVM, SPO-SVM, GA-SVM, RF, SPO-RF, GA-RF, XGB, SPO-XGB, and GA-XGB to assess the probability of flash flood occurrence in each pixel of the Song Ma district. We employed the probability map produced by the GA-RF model, which exhibited the highest accuracy, AUC-ROC, and sensitivity. The Natural Break approach was then used to reclassify this map into five categories: very low, low, moderate, high, and very high flash flood susceptibility (Feizizadeh and Blaschke, 2013). Subsequently, the outcomes from the remaining models were reclassified based on the thresholds derived from GA-RF. Figure 7 visually presents the flood susceptibility maps generated by the nine proposed models. The distribution of each level of flood susceptibility was displayed in Figure 8.

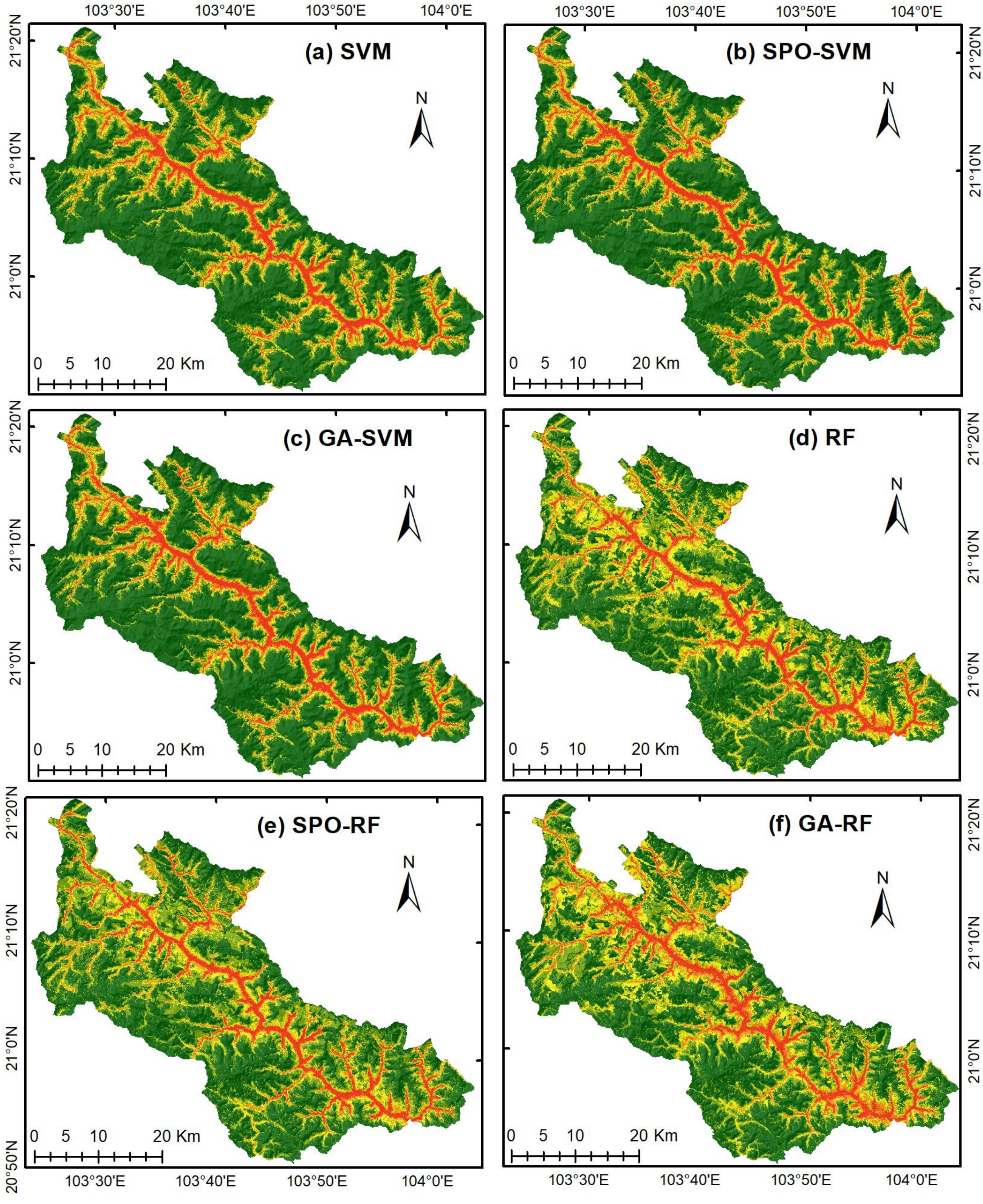
Among the different susceptibility levels, the "Very low" category dominates all models, constituting the largest portion of predictions. Notably, the XGB model exhibits the highest proportion in the "Very low" susceptibility class with 61.33%. Both the SVM and XGB models, in conjunction with their enhanced iterations utilizing PSO and GA, yield predominant outcomes in the "Very low" susceptibility classification, encompassing a range from 50.92% to 61.33%. In contrast, the RF algorithm and its optimized versions only contain a narrower range of 43.02% to 45.04%.

When analyzing the "High" susceptibility level, the outcomes portray a spectrum of predictions. Models such as PSO-XGB and GA-RF, empowered by optimization algorithms, exhibit a noteworthy proportion in this category, with 12.44% and 10.56%, respectively. In contrast, XGB and GA-XGB showed the lowest "High susceptibility" area, at 5.89% and 6.42%, respectively.

Meanwhile, at the "Very High" susceptibility level, the XGB algorithm and its optimized versions, particularly GA-XGB, stand out as remarkable performers in this critical category with 20.51%. At the same time, the residual six algorithms yield relatively similar outcomes, ranging between 7.11% and 7.74%.

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions



https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

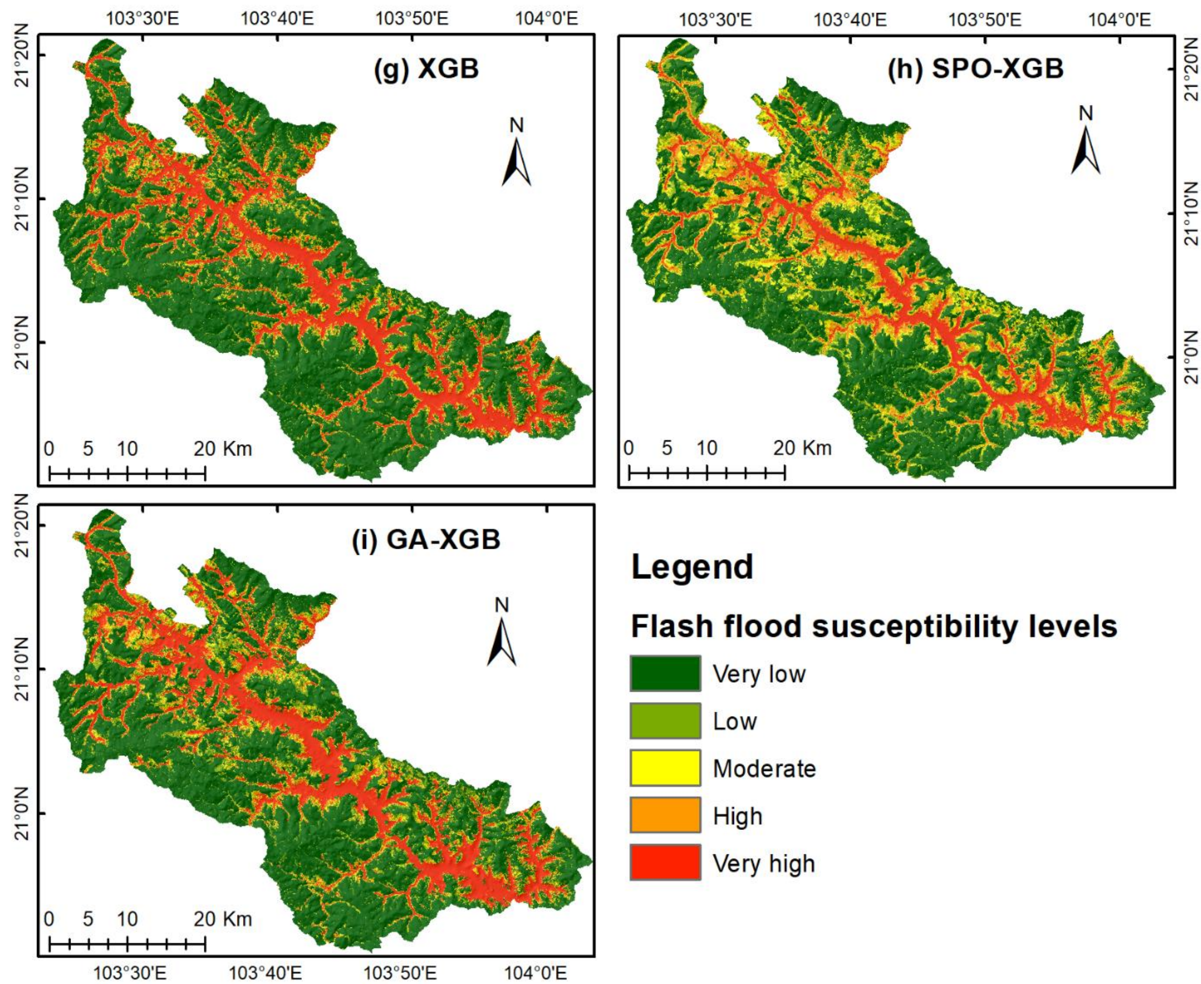


Figure 7:The flash flood susceptibility maps of nine proposed models

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Discussions

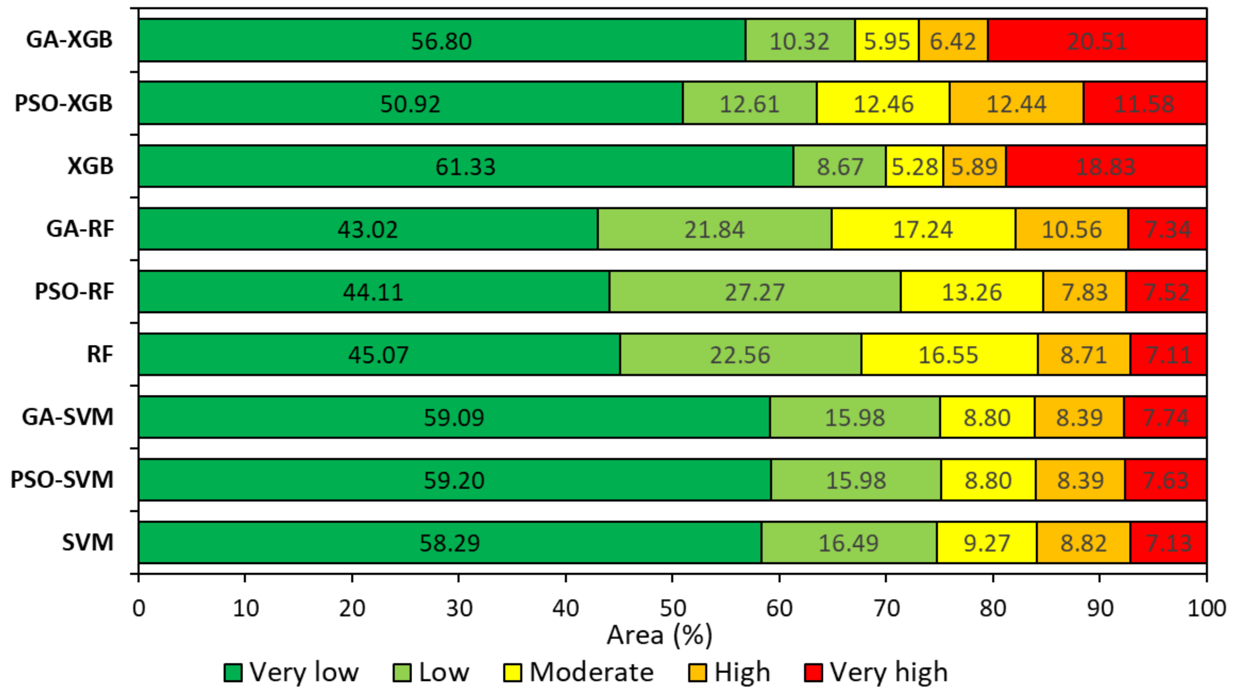


Figure 8:Flash flood susceptibility levels of different models

5. Discussion

Precisely delineating flood susceptibility within a specified geographical zone is essential for protecting residents and developing effective mitigation strategies. As robust novel methodologies are being formulated, the enhancement of outcome accuracy is evident, better assisting those in charge of managing flood risk (Nguyen, 2023). This study aimed to generate flash flood susceptibility maps by using powerful machine learning models (SVM, RF, and XGB) and optimization algorithms (PSO, GA) to create hybrid ML versions to enhance the precision of flash flood occurrence probability prediction. We found that integrating PSO and GA to optimize the SVM, RF, and XGB models has yielded substantial enhancements in their performance across various evaluation metrics.

The study found that, elevation, distance to the river, LULC, stream density, and slope were the most important of the 14 factors that affect flash flood prediction. The significance of elements relies on the particular features of the studied area, such as geographical in nature ecological, climatological, hydrological, and anthropological properties. In addition, the selection and execution of the method approach have a significant influence on determining the relative prominence of these variables. The findings are consistent with previous study on flash floods in several worldwide zones, including Vietnam. Using RF modeling, Chen et al. (2020) underscored the importance of land use, NDVI, slope, distance to rivers, and elevation among 13 inputs, yielding consistent findings for flood susceptibility model in Quannan, China. Bui et al., (2019) used the Learning

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences Open Access EGU Discussions

Vector Quantization method. They showed that slope, TWI, and elevation factors are the most important factors in predicting flash floods in the northern mountainous area of Vietnam. A similar concurrence was observed in Nguyen, (2023) study conducted in Ha Tinh Province, Vietnam, utilizing the RF model, wherein land use, elevation, and slope were identified as primary contributors to flooding susceptibility.

Elevation, distance to the river, LULC, stream density, and slope are critical variables in predicting flash flood probability due to their fundamental roles in shaping a region's hydrological and geomorphological characteristics. These factors directly influence the pathways and processes through which precipitation and runoff interact with the landscape, leading to flash flood events. For elevation, areas at lower elevations are more prone to flooding because they tend to accumulate water from higher ground during intense rainfall, increasing the potential for flash flooding (Abedi et al., 2021). Areas close to rivers or streams are more susceptible because they are in the immediate path of rapidly rising water levels, especially during heavy rainfall. The composition of LULC in a region affects how water is absorbed, infiltrated, or runoff (Nguyen, 2023). The density of streams or drainage networks in an area affects how water is channeled and transported. High stream density can facilitate rapid runoff accumulation during heavy rainfall, contributing to water levels' rapid rise and flash flood likelihood (El-Haddad et al., 2021). Slope or the gradient of the terrain influences the speed at which water flows. Steeper slopes promote faster runoff, which can swiftly increase water levels in downstream areas, potentially leading to flash floods (El-Haddad et al., 2021). Collectively, these variables interact in complex ways to determine the response of a landscape to intense rainfall. The interplay between elevation, proximity to water bodies, land cover characteristics, drainage patterns, and terrain steepness influences the potential for runoff to accumulate quickly and trigger flash flood events.

The findings of the study further demonstrate enhancements in the performance metrics of the optimal algorithms associated with each distinct ML model. The employment of PSO-SVM and GA-SVM ensemble methodologies resulted in improvements for SVM accuracy metrics, elevating them from initial values of 0.861 to 0.873 and 0.903, accompanied by parallel progress in AUC-ROC values, advancing from 0.932 to 0.933 and 0.939, correspondingly. In a similar vein, the utilization of PSO-RF and GA-RF strategies yielded augmented accuracy outcomes for the RF model, ameliorating its accuracy from 0.873 to 0.903 and 0.933, in conjunction with corresponding advancements in AUC-ROC measures, progressing from 0.951 to 0.950 and 0.977, respectively. Analogously, the XGB model exhibited enhanced accuracy through optimization via PSO and GA techniques, resulting in accuracy elevations from 0.879 to 0.939 and 0.927, while simultaneously experiencing improvements in AUC-ROC statistics, ascending from 0.956 to 0.957 and 0.968, respectively. These advancements collectively underline the effectiveness of PSO and GA in fine-tuning the hyperparameters of the SVM, RF, and XGB models. These findings align harmoniously with prior investigations concerning flood forecasting. For instance, a study conducted by (Razavi et al., 2018), focusing on the flood susceptibility within the Jahrom basin of Iran, employed PSO and GA techniques to fine-tune the Adaptive Neuro-Fuzzy Inference System (ANFIS) model. This endeavor yielded noteworthy enhancements in the AUC-ROC values, demonstrating increments of up to 0.945 and 0.926, respectively. Moreover, a study in the year 2022 pursued a similar approach by employing GA to optimize the XGB model, specifically in the context of flash flood prediction within the Tafresh watershed in Iran (Linh et al., 2022). The outcomes of this undertaking manifested a considerable advancement in AUC-ROC

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

statistics, surging from an initial value of 0.85 to a heightened level of 0.87. Additionally, an investigation involving the optimization of the SVM model through the application of GA also demonstrated significant improvements in AUC-ROC performance. The AUC-ROC value experienced an elevation from 0.839 to 0.886 in this study (Arabameri et al., 2022). The optimization process of PSO and GA enhanced the models' ability to capture intricate relationships among influential factors, thereby refining their predictive accuracy and robustness in flood susceptibility analysis.

The nine proposed models' generated flash flood susceptibility maps displayed distinct levels of susceptibility zoning. Across the spectrum of machine learning models analyzed a consistent trend emerges when examining the "Very Low" flood susceptibility level. This class, which designates areas with the lowest probability of encountering flash floods, claims a substantial proportion across all models. Notably, the PSO-XGB distinguishes itself with the highest representation in this category, showcasing its effectiveness in identifying regions with minimal risk. Other models, including SVM and GA-RF, exhibit noteworthy shares in the "Very Low" susceptibility tier. This unified pattern underscores the collective capacity of these models to accurately pinpoint locales with low vulnerability to flash floods, providing a foundational basis for comparison against more vulnerable regions. Conversely, when focusing on the "Very High" flood susceptibility level, Genetic Algorithm-optimized models, specifically GA-XGB, exhibit exceptional proficiency in predicting this category, indicating their prowess in identifying areas at the highest risk. With notably elevated proportions compared to other models, GA-XGB excels in recognizing regions susceptible to severe flash flooding. The diversity in model performance is also evident in the varying ratios across the other susceptibility classes. The "Moderate" and "High" susceptibility levels demonstrate fluctuations in distribution, indicating the models' capacity to distinguish areas with moderate and elevated risks of flash floods. These findings underscore the significance of optimization techniques in refining model outcomes and bolstering their sensitivity to different levels of flood susceptibility. The distribution trends in each category collectively highlight the strengths and nuances of each model in predicting flash flood risk, providing valuable insights for decision-makers and stakeholders engaged in proactive disaster preparedness and response planning.

6. Conclusions

The accurate delineation of an area's susceptibility to flash floods is of utmost importance, serving both to protect residents and to inform the development of effective mitigation strategies. As innovative and robust methodologies continue to advance, there is a noticeable improvement in the precision of our results. This improvement is particularly beneficial for stakeholders responsible for flood risk management. The primary objective of this study was to generate susceptibility maps to flash floods, achieved through the utilization of advanced machine learning models, including SVM, RF, and XGB, combined with optimization algorithms such as PSO and GA. This synergistic approach resulted in hybrid machine learning frameworks that significantly enhanced the accuracy of predicting the probability of flash flood occurrences. The outcomes underscore that the amalgamation of PSO and GA in optimizing the SVM, RF, and XGB models led to marked improvements in their performance across a spectrum of evaluation metrics. The PSO-XGB and GA-RF hybrid approaches exhibit more favorable outcomes among the proposed models. Consequently, we recommend applying these specific models for predicting flash flood susceptibility in the Song Ma district and analogous regions characterized by comparable environmental circumstances within Vietnam and across diverse global locations.

Author Contributions

Yuei-An Liou performed conceptualization, funding acquisition, project administration, supervision, writing – review & editing.

Duc-Vinh Hoang performed conceptualization, methodology, formal analysis, data curation, writing – original draft, visualization.

All authors read and approved the final manuscript.

Acknowledgments

This research is funded by Taiwan National Science and Technology Council (NSTC) under project codes: 112-2111-M-008 -027, 112-2121-M-008 -002 and 111-2634-F-008 -001. The authors would like to thank the Vietnam Academy for Water Resources (VAWR), Vietnam Disaster Management Authority (VDMA), and Vietnam Institute of Geosciences and Mineral Resources (VIGMR) for providing the flood inventory data.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Availability of data and materials

Data will be made available on request.

References

Abedi, R., Costache, R., Shafizadeh-Moghadam, H., Pham, Q.B., 2021. Flash-flood susceptibility mapping based on XGBoost, random forest and boosted regression trees. Geocarto Int. 37, 5479–5496. https://doi.org/10.1080/10106049.2021.1920636

Arabameri, A., Seyed Danesh, A., Santosh, M., Cerda, A., Chandra Pal, S., Ghorbanzadeh, O., Roy, P., Chowdhuri, I., 2022. Flood susceptibility mapping using meta-heuristic algorithms. Geomatics, Nat. Hazards Risk 13, 949–974. https://doi.org/10.1080/19475705.2022.2060138

Arora, A., Arabameri, A., Pandey, M., Siddiqui, M.A., Shukla, U.K., Bui, D.T., Mishra, V.N., Bhardwaj, A., 2021. Optimization of state-of-the-art fuzzy-metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India. Sci. Total Environ. 750, 141565. https://doi.org/10.1016/j.scitotenv.2020.141565

Boser, B., Guyon, I., 5th, V.V.-P. of the, 2003, U., 1992. A training algorithm for optimal margin classifiers. Gautampendse.Com 144–152.

Braud, I., Vincendon, B., Anquetin, S., Ducrocq, V., Creutin, J.D., 2018. The challenges of flash flood forecasting. Mobil.

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

Face Extrem. Hydrometeorol. Events 1 Defin. Relev. Scales Anal. 63–88. https://doi.org/10.1016/B978-1-78548-289-2.50003-3

Bui, D.T., Tsangaratos, P., Ngo, P.T.T., Pham, T.D., Pham, B.T., 2019. Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods. Sci. Total Environ. 668, 1038–1054. https://doi.org/10.1016/j.scitotenv.2019.02.422

Bui, Q.T., Nguyen, Q.H., Nguyen, X.L., Pham, V.D., Nguyen, H.D., Pham, V.M., 2020. Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. J. Hydrol. 581, 124379. https://doi.org/10.1016/j.jhydrol.2019.124379

Cao, C., Xu, P., Wang, Y., Chen, J., Zheng, L., Niu, C., 2016. Flash flood hazard susceptibility mapping using frequency ratio and statistical index methods in coalmine subsidence areas. Sustain. 8. https://doi.org/10.3390/su8090948

Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Bui, D.T., Pham, B.T., Khosravi, K., 2017. A novel hybrid artificial intelligence approach for flood susceptibility assessment. Environ. Model. Softw. 95, 229–245. https://doi.org/10.1016/j.envsoft.2017.06.012

Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. 13-17-Augu, 785–794. https://doi.org/10.1145/2939672.2939785

Chen, W., Li, Y., Xue, W., Shahabi, H., Li, S., Hong, H., Wang, X., Bian, H., Zhang, S., Pradhan, B., Ahmad, B. Bin, 2020. Modeling flood susceptibility using data-driven approaches of naïve Bayes tree, alternating decision tree, and random forest methods. Sci. Total Environ. 701, 134979. https://doi.org/10.1016/j.scitotenv.2019.134979

Collier, C.G., 2007. Flash flood forecasting: What are the limits of predictability? Q. J. R. Meteorol. Soc. 133, 3–23. https://doi.org/10.1002/qj.29

Costache, R., Țîncu, R., Elkhrachy, I., Pham, Q.B., Popa, M.C., Diaconu, D.C., Avand, M., Costache, I., Arabameri, A., Bui, D.T., 2020. New neural fuzzy-based machine learning ensemble for enhancing the prediction accuracy of flood susceptibility mapping. Hydrol. Sci. J. 65, 2816–2837. https://doi.org/10.1080/02626667.2020.1842412

Dao, P.D., Liou, Y.A., 2015. Object-based flood mapping and affected rice field estimation with landsat 8 OLI and MODIS data. Remote Sens. 7, 5077–5097. https://doi.org/10.3390/rs70505077

Dodangeh, E., Choubin, B., Eigdir, A.N., Nabipour, N., Panahi, M., Shamshirband, S., Mosavi, A., 2020. Integrated machine learning methods with resampling algorithms for flood susceptibility prediction. Sci. Total Environ. 705, 135983. https://doi.org/10.1016/j.scitotenv.2019.135983

El-Haddad, B.A., Youssef, A.M., Pourghasemi, H.R., Pradhan, B., El-Shater, A.H., El-Khashab, M.H., 2021. Flood susceptibility prediction using four machine learning techniques and comparison of their performance at Wadi Qena Basin, Egypt, Natural Hazards. https://doi.org/10.1007/s11069-020-04296-y

Feizizadeh, B., Blaschke, T., 2013. GIS-multicriteria decision analysis for landslide susceptibility mapping: Comparing three methods for the Urmia lake basin, Iran. Nat. Hazards 65, 2105–2128. https://doi.org/10.1007/s11069-012-0463-3

Kennedy, J., Russell, E., 1995. Particle Swarm Optimization. TIn Proc. IEEE Int. Conf. neural networks 1942–1948.

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

https://doi.org/10.1007/978-3-319-46173-1_2

Linh, N.T.T., Pandey, M., Janizadeh, S., Sankar Bhunia, G., Norouzi, A., Ali, S., Bao Pham, Q., Tran Anh, D., Ahmadi, K., 2022. Flood susceptibility modeling based on new hybrid intelligence model: Optimization of XGboost model using GA metaheuristic algorithm. Adv. Sp. Res. 69, 3301–3318. https://doi.org/10.1016/j.asr.2022.02.027

Liou, Y.A., Nguyen, A.K., Li, M.H., 2017. Assessing spatiotemporal eco-environmental vulnerability by Landsat data. Ecol. Indic. 80, 52–65. https://doi.org/10.1016/j.ecolind.2017.04.055

Madhuri, R., Sistla, S., Srinivasa Raju, K., 2021. Application of machine learning algorithms for flood susceptibility assessment and risk management. J. Water Clim. Chang. 12, 2608–2623. https://doi.org/10.2166/wcc.2021.051

Moore et al., 1991. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. Hydrol. Process. 5, 3–30. https://doi.org/10.1002/hyp.3360050103

Nachappa, T., Tavakkoli, P.S., Gholamnia, K., Ghorbanzadeh, O., Rahmati, O., Blaschke, T., 2020. Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory. J. Hydrol. 590, 125275. https://doi.org/10.1016/j.jhydrol.2020.125275

Nguyen, H.D., 2023. Spatial modeling of flood hazard using machine learning and GIS in Ha Tinh province, Vietnam. J. Water Clim. Chang. 14, 200–222. https://doi.org/10.2166/wcc.2022.257

Nguyen, K.A., Liou, Y.A., Terry, J.P., 2019. Vulnerability of Vietnam to typhoons: A spatial assessment based on hazards, exposure and adaptive capacity. Sci. Total Environ. 682, 31–46. https://doi.org/10.1016/j.scitotenv.2019.04.069

Panahi, M., Dodangeh, E., Rezaie, F., Khosravi, K., Van Le, H., Lee, M.J., Lee, S., Thai Pham, B., 2021. Flood spatial prediction modeling using a hybrid of meta-optimization and support vector regression modeling. Catena 199, 105114. https://doi.org/10.1016/j.catena.2020.105114

Razavi, T.S.V., Kornejady, A., Pourghasemi, H.R., Keesstra, S., 2018. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. Sci. Total Environ. 615, 438–451. https://doi.org/10.1016/j.scitotenv.2017.09.262

Roy, P., Chandra Pal, S., Chakrabortty, R., Chowdhuri, I., Malik, S., Das, B., 2020. Threats of climate and land use change on future flood susceptibility. J. Clean. Prod. 272, 122757. https://doi.org/10.1016/j.jclepro.2020.122757

Špitalar, M., Gourley, J.J., Lutoff, C., Kirstetter, P.E., Brilly, M., Carr, N., 2014. Analysis of flash flood parameters and human impacts in the US from 2006 to 2012. J. Hydrol. 519, 863–870. https://doi.org/10.1016/j.jhydrol.2014.07.004

Santos, P.P., Reis, E., Pereira, S., Santos, M., 2019. A flood susceptibility model at the national scale based on multicriteria analysis. Sci. Total Environ. 667, 325–337. https://doi.org/10.1016/j.scitotenv.2019.02.328

Shafizadeh, M.H., Valavi, R., Shahabi, H., Chapi, K., Shirzadi, A., 2018. Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. J. Environ. Manage. 217, 1–11. https://doi.org/10.1016/j.jenvman.2018.03.089

Sheta, A., 2006. A Comparsion between Genetic Algorithms and Sequential Quadratic Programming in Solving Constrained Optimization Problems. AIML J. 6, 67–74.

https://doi.org/10.5194/nhess-2024-215
Preprint. Discussion started: 12 December 2024

Sivanandam, S.N., Deepa, S.N., 2008. Introduction to genetic algorithms, Introduction to Genetic Algorithms. https://doi.org/10.1007/978-3-540-73190-0

Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y., A., R., 2020. Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations — A Review. Remote Sens. 12(7):1135.

Tehrany, M.S., Kumar, L., Shabani, F., 2019. A novel GIS-based ensemble technique for flood susceptibility mapping using evidential belief function and support vector machine: Brisbane, Australia. PeerJ 2019. https://doi.org/10.7717/peerj.7653

Tien Bui, D., Hoang, N.D., Pham, T.D., Ngo, P.T.T., Hoa, P.V., Minh, N.Q., Tran, X.T., Samui, P., 2019. A new intelligence approach based on GIS-based Multivariate Adaptive Regression Splines and metaheuristic optimization for predicting flash flood susceptible areas at high-frequency tropical typhoon area. J. Hydrol. 575, 314–326. https://doi.org/10.1016/j.jhydrol.2019.05.046

Vinh, H.D., Liou, Y.-A., 2024. Assessing the influence of human activities on flash flood susceptibility in mountainous regions of Vietnam. Ecol. Indic. 158, 111417. https://doi.org/10.1016/j.ecolind.2023.111417

Wang, L.C., Hoang, D.V., Liou, Y.A., 2022. Quantifying the Impacts of the 2020 Flood on Crop Production and Food Security in the Middle Reaches of the Yangtze River, China. Remote Sens. 14. https://doi.org/10.3390/rs14133140

Wang, Z., Qin, C., Wan, B., Song, W.W., 2021. A comparative study of common nature-inspired algorithms for continuous function optimization. Entropy 23, 1–40. https://doi.org/10.3390/e23070874

WMO, 2007. Guidance on Flash Flood Management. Recent experiences from Central and Eastern Europe. Flood Manag. Tools Ser. APFM, Gene, 65.

WMO, GWP, 2012. Management of Flash Floods. Integr. Flood Manag. Tools Ser. Manag. flash flood 44.

Zhang, Y., Ge, T., Tian, W., Liou, Y.-A., 2019. Debris Flow Susceptibility Mapping Using Machine-Learning Techniques in Shigatse Area, China. Remote Sens. 11, 1–26. https://doi.org/doi:10.3390/rs11232801