

## Enhancing flood risk assessment through integration of ensemble learning approaches and physical-based hydrological modeling

Mohamed Saber, Tayeb Boulmaiz, Mawloud Guermoui, Karim I. Abdrabo, Sameh A. Kantoush, Tetsuya Sumi, Hamouda Boutaghane, Tomoharu Hori, Doan Van Binh, Binh Quang Nguyen, Thao T. P. Bui, Ngoc Duong Vo, Emad Habib & Emad Mabrouk

To cite this article: Mohamed Saber, Tayeb Boulmaiz, Mawloud Guermoui, Karim I. Abdrabo, Sameh A. Kantoush, Tetsuya Sumi, Hamouda Boutaghane, Tomoharu Hori, Doan Van Binh, Binh Quang Nguyen, Thao T. P. Bui, Ngoc Duong Vo, Emad Habib & Emad Mabrouk (2023) Enhancing flood risk assessment through integration of ensemble learning approaches and physical-based hydrological modeling, *Geomatics, Natural Hazards and Risk*, 14:1, 2203798, DOI: [10.1080/19475705.2023.2203798](https://doi.org/10.1080/19475705.2023.2203798)

To link to this article: <https://doi.org/10.1080/19475705.2023.2203798>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 04 May 2023.



Submit your article to this journal [↗](#)



Article views: 1498



View related articles [↗](#)



View Crossmark data [↗](#)

## Enhancing flood risk assessment through integration of ensemble learning approaches and physical-based hydrological modeling

Mohamed Saber<sup>a</sup>, Tayeb Boulmaiz<sup>b</sup>, Mawloud Guermoui<sup>c</sup>, Karim I. Abdrabo<sup>d,e</sup>, Sameh A. Kantoush<sup>a</sup>, Tetsuya Sumi<sup>a</sup>, Hamouda Boutaghane<sup>f</sup>, Tomoharu Hori<sup>a</sup>, Doan Van Binh<sup>g</sup>, Binh Quang Nguyen<sup>a,h</sup>, Thao T. P. Bui<sup>a</sup>, Ngoc Duong Vo<sup>h</sup>, Emad Habib<sup>i</sup> and Emad Mabrouk<sup>j,k</sup>

<sup>a</sup>Disaster Prevention Research Institute (DPRI), Kyoto University, Kyoto, Japan; <sup>b</sup>Materials, Energy Systems Technology and Environment Laboratory, Université de Ghardaia, Ghardaia, Algeria; <sup>c</sup>Unité de Recherche Appliquée en Energies Renouvelables, URAER, Centre de Développement des Energies Renouvelables, CDER, Ghardaia, Algeria; <sup>d</sup>Faculty of Urban and Regional Planning, Cairo University, Giza, Egypt; <sup>e</sup>Department of Urban Management, Graduate School of Engineering, Kyoto University, Kyoto, Japan; <sup>f</sup>Hydraulic Department, Badji Mokhtar-Annaba University, Annaba, Algeria; <sup>g</sup>Master Program in Water Technology, Reuse and Management, Faculty of Engineering, Vietnamese German University, Binh Duong, Vietnam; <sup>h</sup>The University of Danang-University of Science and Technology, Danang, Vietnam; <sup>i</sup>Department of Civil Engineering, University of Louisiana at Lafayette, Lafayette, LA, USA; <sup>j</sup>College of Engineering and Technology, American University of the Middle East, Egaila, Kuwait; <sup>k</sup>Department of Mathematics, Faculty of Science, Assiut University, Assiut, Egypt

### ABSTRACT

This study aims to examine three machine learning (ML) techniques, namely random forest (RF), LightGBM, and CatBoost for flooding susceptibility maps (FSMs) in the Vietnamese Vu Gia-Thu Bon (VGTB). The results of ML are compared with those of the rainfall-runoff model, and different training dataset sizes are utilized in the performance assessment. Ten independent factors are assessed. An inventory map with approximately 850 flooding sites is based on several post-flood surveys. The inventory dataset is randomly split between training (70%) and testing (30%). The AUC-ROC results are 97.9%, 99.5%, and 99.5% for CatBoost, LightGBM, and RF, respectively. The FSMs developed by the ML methods show good agreement in terms of an extension with flood inundation maps developed using the rainfall-runoff model. The models' FSMs showed 10–13% of the total area to be highly susceptible to flooding, consistent with RRI's flood map. The FSMs show that downstream areas (both urbanized and agricultural) are under high and very high levels of susceptibility. Additionally, different sizes of the input datasets are tested to determine the least number of data points having acceptable reliability. The results demonstrate that the ML methods can realistically predict FSMs, regardless of the number of training samples.

### ARTICLE HISTORY

Received 26 December 2022  
Accepted 12 April 2023

### KEYWORDS

Machine learning; random forest; LightGBM; CatBoost; flood susceptibility mapping; rainfall-runoff inundation model

CONTACT Mohamed Saber  [mohamedmd.saber.3u@kyoto-u.ac.jp](mailto:mohamedmd.saber.3u@kyoto-u.ac.jp)

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

## 1. Introduction

Floods are the greatest catastrophic natural disaster on a global scale. Because of their short lag times, flash floods are more devastating than other types of flooding (Vinet 2008; Bui, Ngo, et al. 2019; Abdrabo, Kantosh, et al. 2022). Flash floods have the highest mortality rates per event and are the leading cause of flood-related deaths in developed countries, due to their high-speed flow and limited warning time, (Jonkman and Kelman 2005; Ashley and Ashley 2008; Bisht et al. 2018; Esmaili et al. 2022; Abdrabo et al. 2023). However, floods are more destructive in developing countries like Vietnam. Extreme fluctuations in storm patterns and global climate change are the leading causes of the reported rise in flash floods (Hirabayashi et al. 2013; IPCC 2014; Abdrabo, Saber, et al. 2022; Saber et al. 2022). Typhoons, tropical cyclones, extended coastal areas, and dense river networks are the primary causes of severe flooding in Vietnam. It is also highly vulnerable to floods caused by extreme storms. Vietnam is rated eighth among the top ten countries in weather events (Thao et al. 2020), where densely populated areas are more vulnerable to floods. Consequently, continuous risks in human life and assets will always exist (Luu et al. 2021). Vietnam is susceptible to natural disasters, with over 13,000 deaths and 1% of GDP lost annually in the last two decades. More than half of the country's land area and population are at risk of being affected by tropical cyclones and floods (World Bank report, 2010). Vietnam's geography and location make it prone to the effects of climate change (IPCC, 2007; Wang et al. 2010). For instance, In 2020, Central Vietnam was hit by severe natural disasters that caused significant loss of life and damage to property. 357 people died or went missing, 876 were injured, and over half a million houses were submerged or damaged. The floods and storms also damaged infrastructure and hampered aid distribution. The estimated loss was around VND 35,180,997,000, making it the worst disaster to hit Central Vietnam in the past century (The International Federation of Red Cross Red Crescent Societies 2022). Additionally, The country's agriculture heavily relies on fertile, low-lying regions that benefit from normal flooding, but also renders them vulnerable to severe flooding and crop damage (IPCC, 2007; Wang et al. 2010). Flash flood mitigation for risk reduction and management requires efficient monitoring measures (Arora et al. 2021). Food susceptibility mapping is critical for scientists and governments worldwide to keep cities and human settlements safe and resilient (Ali et al. 2020).

Several studies have been performed to forecast the likelihood of flooding events. These studies can be divided into rainfall-runoff analysis, conventional analysis, and pattern categorization (Tien Bui and Hoang 2017). The traditional analysis uses time-series data an extended period obtained from rainfall stations to produce regression models. The rainfall-runoff models (e.g. MIKE, PCSWMM 2D, HEC-RAS, etc.) determine the correlation between runoff and rainfall to calculate temporal and spatial floods (Nguyen et al. 2015). In general, this task is complicated because of difficulties in accessing affected areas, especially in developing countries; as a result, the hydrological models' performance may be compromised, and comprehensive observational datasets are needed for the calibration and validation of models (Abushandi and Merkel 2011; Abdrabo et al. 2020). Both groups have a significant deficiency: the lack of required data frequently limits their applications and incurs substantial costs for

data collection (Fenicia et al. 2014). On the other hand, the last group (pattern classification), uses machine learning (ML) models that utilize historical geological, environmental, and flood data. Accordingly, flood-prone areas are defined as flood and non-flood classes (Bui, Ngo, et al. 2019). However, comparative studies and integration between these groups are lacking (Hsu et al. 1995; Demirel et al. 2009; Humphrey et al. 2016; Yang et al. 2020).

Over the last 20 years, the application of ML methods for flood susceptibility forecasting has been extensively evaluated globally. As a result, the recent advancement of ML methods has significantly improved flood modeling. Because of the ability of ML techniques to capture information without making predetermined assumptions, process complex datasets, and promptly provide high accuracy and reliable results, such practices have become widespread (Arabameri et al. 2020; Costache, Popa, et al. 2020). Several articles have employed GIS techniques and remote sensing to develop reliable flooding susceptibility maps (FSMs). ML models are currently associated with GIS to address various hydrological and environmental issues (Akay and Taş 2020). Logistic regression (LR), support vector machines (SVMs), Artificial neural networks (ANNs), adaptive neuro-fuzzy inference system (ANFIS), and random forest (RF) models are the most utilized in ML for FSM (Hong et al. 2018; Choubin et al. 2019; Darabi et al. 2019; Costache, Hong, et al. 2020; Dodangeh et al. 2020; Shirzadi et al. 2020; Arora et al. 2021; Shahabi et al. 2021; Gharakhanlou and Perez 2023).

Ensemble and hybrid ML models have recently appeared, outperforming single models' accuracy predictions (Zenggang et al. 2021). Several ensemble ML techniques, such as the alternating decision tree, bagging, dagging, reduced-error pruning tree, naïve Bayes tree, logistic model tree, AdaBoost, J48 decision tree, and random subspace ensembles have been applied to enhance the predictive accuracy of the FSM (Luu et al. 2021; Pham, Jaafari, et al. 2021; Tuyen et al. 2021). CatBoost, LightGBM, and RF (Random Forest) models are popular machine-learning algorithms used in various applications, including flood susceptibility modeling. While their effectiveness has been demonstrated in other fields, their wide applicability in flood susceptibility modeling remains limited (Saber et al. 2021; Aydin and Iban 2022; Seydi et al. 2022). CatBoost and LightGBM are gradient-boosting algorithms designed to handle categorical variables efficiently. It is known for its superior accuracy in classification tasks and has been used successfully in flood susceptibility modeling (Saber et al. 2021), especially in their simplicity of implementation and fast training speed. In the prediction of flood susceptibility, LightGBM was found to have higher accuracy than other algorithms such as random forest and gradient boosting (Aydin and Iban 2022). Similarly, CatBoost has been reported to outperform other methods like Logistic Regression and SVM (Support Vector Machine) (Seydi et al. 2022). RF is a popular ensemble learning algorithm that uses decision trees to make predictions. Rivals of these models are the Deep Learning techniques (CNN, LSTM, BI-LSTM, GRU, Transformer, and their hybridization techniques). However, the major drawbacks of these deep learning techniques include the requirement for longer historical data and their complex architecture with multiple hyper-parameters to tune. The advantage of the chosen classifier methods lies in their ease of implementation, fast training, and high accuracy (ranging from 95.5% to 99.5%). Additionally, these methods do not

require extensive historical data to select the optimal model parameters. Several studies have developed flooding susceptibility maps in Vietnam using ML, which can be classified into three groups. The first evaluates the utilization of new ML models and their ability to detect areas prone to floods. For instance, the AdaBoost, bagging, and random subspace ensemble learning methods were combined with the Partial Decision Tree (PART) classifier to develop new GIS-based ensemble methods for FSM in the province of Quang Binh (Luu et al. 2021). (Nguyen et al. 2022) used the hybridization of the relevance vector machine and coyote optimization algorithm to generate FSM of the Gianh River watershed (Central Vietnam). The second group attempts to overcome the limitations in the study's numbers that utilize remote sensing data to generate input variables for FSM despite the merits of using such available data (Pham et al. 2019, 2010–2018). As such, (Dhara et al. 2020; Nguyen et al. 2020; Nhu et al. 2020; Ngo et al. 2021) suggested a hybrid approach using remotely sensed data with ML models for flooding susceptibility. The third group introduced a novel deep learning neural network (DLNN) algorithm for FSM (Tien Bui et al. 2020), integrating particle swarm optimization (PSO) and extreme learning machines (ELMs) (Bui, Ngo, et al. 2019; Bui et al. 2020) along with a comparison between ML and deep learning techniques (Pham, Luu, et al. 2021) for the same study area. ML techniques have been utilized in previous studies in Vietnam for flood susceptibility mapping. However, there has been no exploration of the potential of two ML models (CatBoost and LightGBM) in predicting flooding susceptibility in humid environments as in Vietnam.

Several ensemble methods to predict FFS have been used (Shahabi et al. 2021). ML methods consist of multiple stages (Arora et al. 2021), including the preparation of the inventory and influencing factors, as well as the assessment of the accuracy of the ML model. Despite many FSM studies using ML techniques, most focus on the model's accuracy. While a depth analysis of the used approach was not well explored. Some limitations in this regard should be considered, such as the limited availability of high-quality and comprehensive data, such as elevation data, hydrologic and hydraulic data, land use and land cover data, and historical flood data. Additionally, the diversity and complexity of flood mechanisms and the dynamic nature of floods pose challenges to developing robust and accurate machine-learning models (Nguyen et al. 2022). Another limitation is the lack of standardization and comparison across different studies, which makes it difficult to compare and generalize results.

Moreover, they were not considered return periods in spatial modeling of the flood because the selection of training and validation points has not been based on the return periods due to a lack of hazard maps for each return period (Choubin et al. 2023). Additional to the previous limitations, two crucial aspects must be addressed; 1) the inventory flood/non-flood database, an essential part of developing the ML model, must be better defined. A larger inventory dataset typically provides more information about the environment and its susceptibility to flooding, which can improve the accuracy of the machine learning models. However, there may be diminishing returns with a vast inventory dataset, where additional data points do not provide further information or improve the model performance. In general, the optimal inventory dataset size will depend on several factors, such as the model's complexity,

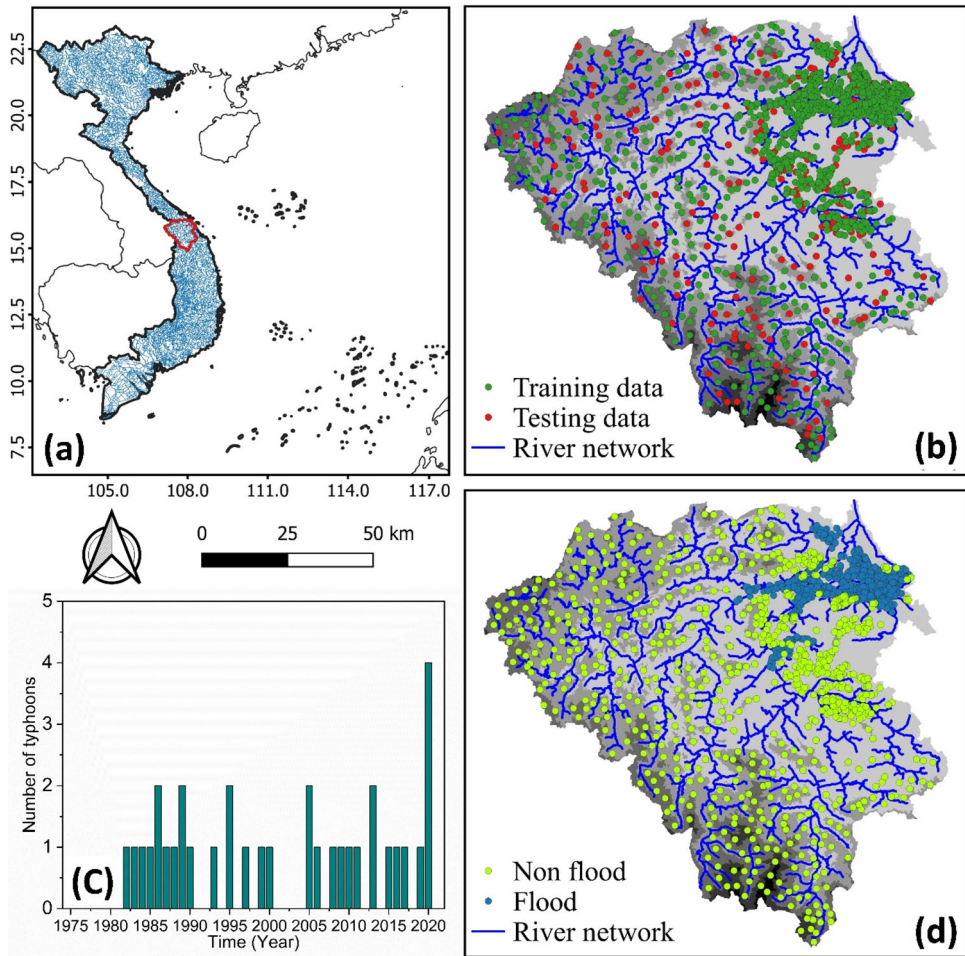
the data's quality, the spatial resolution of the inventory data, and the environmental characteristics of the study area. Therefore, it is crucial to evaluate the performance of the machine learning models for different inventory dataset sizes to determine the optimal dataset size for the specific use case. To the best of our knowledge, no previous studies have addressed or analyzed the impact of inventory dataset size on the accuracy of FSM results, except for one study that focused solely on predicting water levels (Tiwari and Chatterjee 2010, 2) the generated FSM using the recently developed system with ML should be validated using the conventional method (hydrological and hydraulic modeling).

In the present study, we examined two ML models, the light gradient boosting machine (LightGBM) and categorical boosting (CatBoost), for FSM for the first time in humid regions after successful application in arid areas (Saber et al. 2021). Previously, both methods had been applied to some applications. LightGBM, for example, has been employed in some previous studies due to its accuracy in predictions, short computational time, and exceptional prevention of overfitting problems. Accordingly, our primary objectives are (1) to evaluate how practical the two ML approaches (CatBoost and LightGBM) are for predicting flooding susceptibility in humid environments (Vu Gia-Thu Bon basin in Vietnam); (2) to compare the results of the two models used with those of the conventional RF method; (3) to test the effect of the inventory datasets (number of points) on the accuracy of the results in the study area; and (4) to compare the rainfall-runoff inundation (RRI) 2D hydrological model with the proposed ML integrated models in terms of flood extent.

## 2. Study area

The River Basin of Vu Gia-Thu Bon (VGTB) (Figure 1) is one of the major river basins in Vietnam, with a surface area of 10,350 km<sup>2</sup> (RETA 2011). The land use types in the basin are forest (47%), agriculture (26%), and pasture (20%) (Avitabile et al. 2016). The climate in this basin is tropical monsoon, with two seasons: dry summer (January-August) and wet winter (September-December). The basin's topographic features are hilly mountainous areas, with approximately 60% of the basin having an elevation of over 552 m. The average annual rainfall varies significantly, from 2000 mm in the downstream regions to more than 4000 mm in mountainous areas. There are seasonal differences, with 65% to 80% of the annual rainfall between September and December (RETA 2011). The rain in the eight months of the dry season is only 20–35% of the annual rainfall (Nauditt et al. 2017). Due to rainfall's spatial distribution, the VGTB basin's runoff varies substantially across seasons. River flow in this period accounts for around 62.5–69.2% of the total annual flow. The impacts of both heavy rainfall and steep terrain usually lead to flash flooding. Approximately 4–8 floods occur annually. Due to meteorological patterns such as tropical depressions, typhoons, and cold air, the highest flood peak occurs in October and November (Vu et al. 2011). According to the Quang Nam Province Commanding Committee for Disaster Prevention, Search and Rescue report, the number of fatalities and property losses caused by floods and storms has been growing, particularly in 2020 (Figure 2).

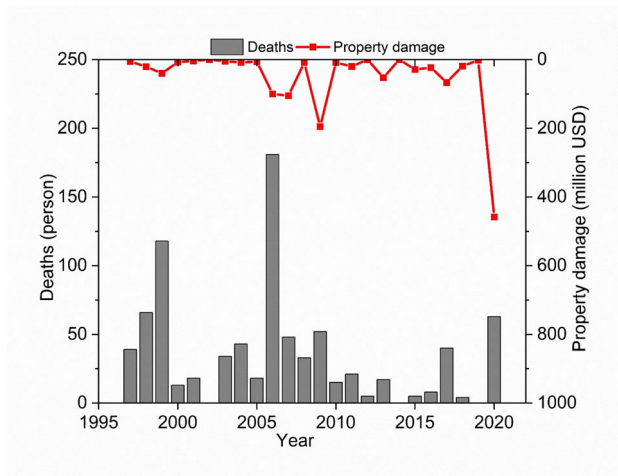




**Figure 1.** Location of the river basin of VGTB: (a) Vietnam Map, (b) flood inventory dataset map for training and validation, (c) total annual precipitation of the entire basin, and (d) flooded and non-flooded locations.

There are two main sub-basins in the VGTB river system: Thu Bon Basin and the Vu Gia Basin. The Quang Hue River connects both Rivers. The Vu Gia River starts from the western slope of Kon Tum and flows towards the province of Quang Nam and the city of Danang. It connects with the sea at the Cua Han estuary. The length of the main river from the source to the Cua Han estuary is 204 km. At the same time, the Thu Bon River originates from a mountain of 1500 m in Kon Tum province. The river length from the source to the Cua Dai estuary is 198 km.

The River Basin of VGTB has been chosen for a flood susceptibility mapping due to its location, flood history, environmental factors, and socioeconomic factors. The VGTB covers approximately  $10,350 \text{ km}^2$  in central Vietnam and is an essential water source for the region. However, it is also one of the most flood-prone areas in the country, experiencing major floods almost every decade. For example, in 2000 and 2009, the region experienced catastrophic floods that resulted in significant loss of life and property damage (Luu et al. 2018). Steep mountain slopes, narrow valleys,



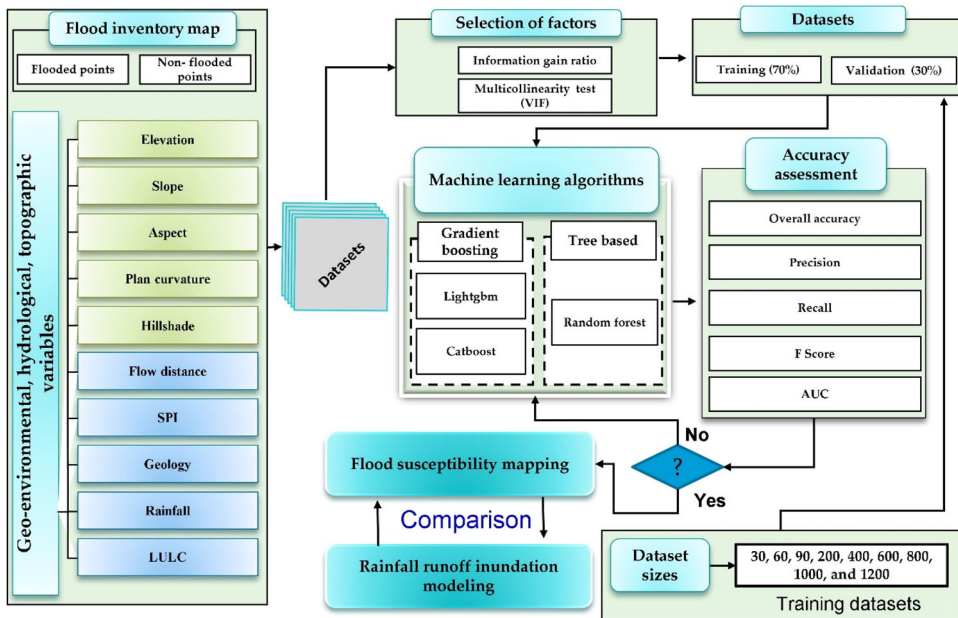
**Figure 2.** Number of deaths and property damage caused by storms and floods from 1997 to 2020 in the River basin of VGTB (source: Commanding Committee for Natural Disaster Prevention and Control, Search and Rescue in Quang Nam Province).

typhoons, and extreme weather events exacerbate flooding in the region. The VGTB river basin includes both Quang Nam province (1.5 million in 2012 (UN Habitat 2013)) and Da Nang city (1.1 million in 2018 (United Nations Economic and Social Commission for Asia and the Pacific 2018)) is home to approximately 3 million people, many of whom live in low-lying areas vulnerable to flooding. The region is also an important agricultural center. The comprehensive flood susceptibility mapping in the VGTB will help not only researchers but also decision makers understand the situation of the region under the threat of flooding. The results will provide a guidance for proposing effective measures to mitigate the impacts of flooding on both human populations and the environment, in order to reduce the risks associated with flooding in the region.

### 3. Methodology

This study's methodology comprises multiple phases, as shown in the flowchart in Figure 3. This methodology has two main parts. First, a flood inventory map is created using 850 flooded spots. These spots were determined primarily through post-flood assessments conducted after typhoons in 1999, 2006, 2007, 2009, 2013, and 2020. Non-flooded points (850) over the catchment were randomly selected using GIS tools. Additionally, Ten commonly used independent flood susceptibility variables (FSFs) covering hydrological, topographical, geological, and landform characteristics were considered for modeling. The flooding susceptibility influencing factors, namely elevation, aspect, slope, hillshade, horizontal flow distance, plan curvature, stream power index (SPI), geology, land use/land cover, and rainfall, were used to define the linear relationship with other variables. In subsequent phases, the data was divided into two sets using a random selection scheme: (70%) for training and (30%) for testing. ArcGIS was used to create spatial maps for each flooding susceptibility





**Figure 3.** Methodology flowchart for flood susceptibility mapping.

factor while keeping spatial resolution consistency in mind. Following that, two approaches, the variance inflation factor (VIF) and the information gain ratio (IGR) were used to investigate the significance of the influencing factors in flooding susceptibility. The ML algorithms RF, CatBoost, and LightGBM were then implemented. The accuracy of the ML models' final results was assessed using various statistical processes, including the most dominant area under the curve (AUC). Moreover, as we have very high-quality observational flood locations, we tested the models to check the different sizes of the training datasets (Figure 3). The final FSM developed by the ML models was then compared with the flood inundation maps from the 2D physical hydrological model regarding the flood extent.

### 3.1. Datasets

#### 3.1.1. Flooding inventory datasets

The initial stage of flood susceptibility mapping identifies flood locations (points) based on prior flood records using several sources, such as field surveys, remote sensing data, and flood forecasting records (Tehrany et al. 2014; Wang et al. 2019; Band et al. 2020; Esfandiari et al. 2020). The locations of future hazardous events might be forecasted using previous information (Devkota et al. 2013; Tehrany and Kumar 2018). As a result, the fundamental phase of a flood susceptibility study is an examination of prior historical flood occurrences and their contributing elements (Masood and Takeuchi 2012). The accuracy in selecting flood points is reflected in the model accuracy for FSM (Tehrany et al. 2013; Arora et al. 2019). In this study, 1700 ground control points (Figure 1) were identified for flooded (850) and non-flooded points (850). Approximately 1250 were used for training and 450 for testing the models.

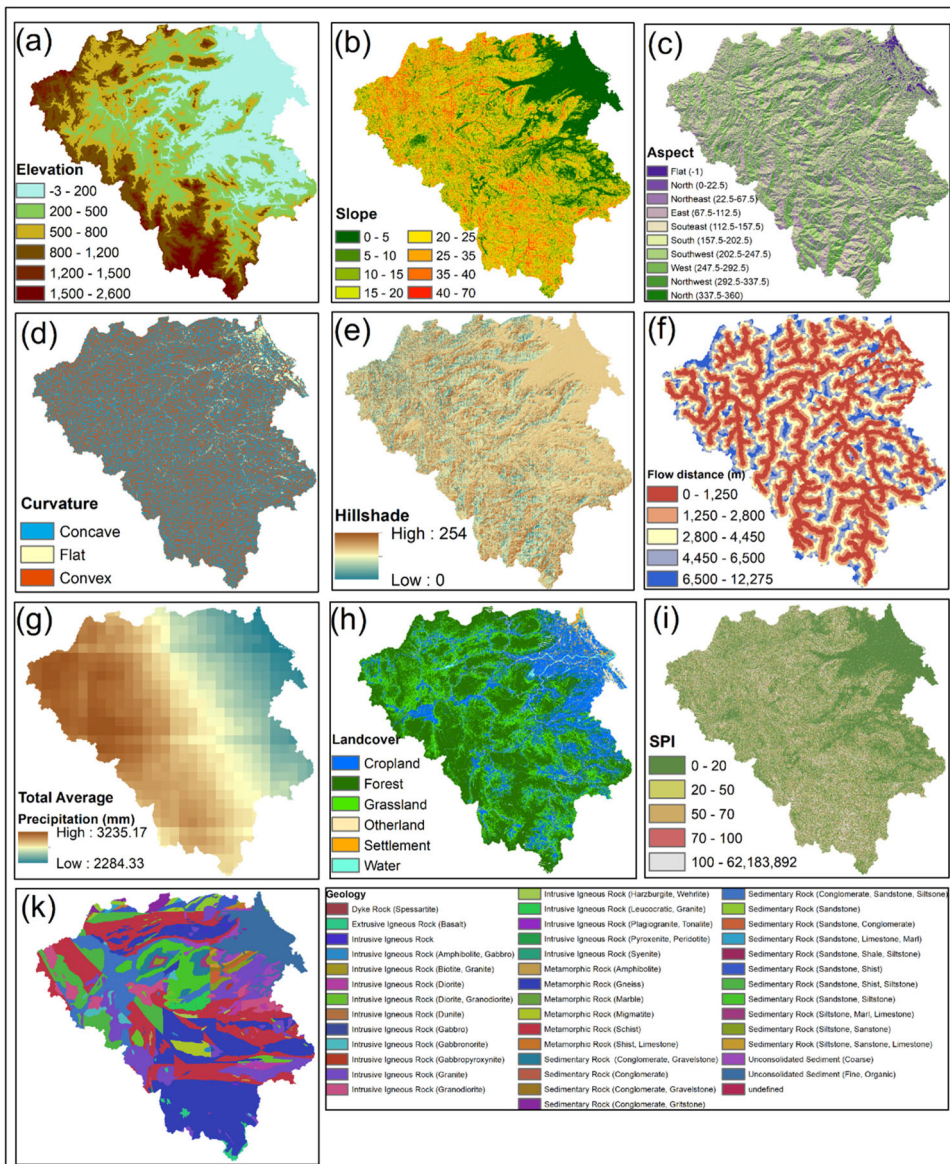
The flooded locations were compiled from historical flood records and post-flood field surveys in 1999, 2006, 2007, 2009, 2013, and 2020 (Figure 1). Flood and non-flood locations were assigned values 1 and 0, respectively. Using the random selection approach, the points were divided into 70% for training to create the flooding prediction model and 30% for testing the model's performance and generalization abilities.

### 3.1.2. Spatial datasets (flood controlling parameters)

Identifying flood governing parameters for flooding susceptibility mapping is critical and influences model accuracy (Kia et al. 2012). Runoff in a drainage system is influenced by the watershed features, terrain, catchment area, land use types, and land cover during floods (Hölting and Coldewey 2019). Generally, there are no uniform and standard selection criteria for FSM controlling factors. The selection of flood-controlling parameters depends on various factors such as the area's location, topography, hydrology, and human activities. Here are some common parameters used for flood control, along with the justification behind their selection (Rahman, Chen, Elbeltagi, et al. 2021; Rahman, Chen, Islam, et al. 2021b): (1) Watershed characteristics: such as its size, shape, and slope, can affect the amount and speed of water runoff, which can, in turn, affect flood risk. (2) River channel characteristics: such as shape, width, depth, and roughness of a river channel can all affect how water flows through it. (3) Topography data: such as elevation maps and terrain models, can help identify areas more prone to flooding. (4) Land use and land cover: Human activities such as urbanization, deforestation, and agriculture can alter the natural landscape and affect flood risk. For example, urbanization can increase the number of impervious surfaces, leading to more runoff and higher flood risk. Land use and land cover analysis can help identify areas where land use changes can be made to reduce flood risk.

Depending on previous research and the features of the studied area (Rahman, Chen, Elbeltagi et al. 2021; Rahman, Chen, Islam, et al. 2021), as well as the availability of data, we were able to develop fifteenth flood governing indicators that include topographic, geological, hydrological, and landform factors. The fifteenth indicators are plan curvature, elevation, slope, aspect, horizontal and vertical distance from streams, flow direction and accumulations, hillshade, SPI, rainfall, geology, NDVI, and land use/land cover. After feature selection of such parameters (See section 3.2), only ten parameters were considered as influencing features in this study. Using ArcGIS, the data were developed in raster formats (Figure 4). All topographic factors were constructed based on the spatial analysis of the MERIT digital elevation model (Yamazaki et al. 2017). The terrain's elevation had a spatial resolution of 3 s (90 m at the equator). It was created by removing the incorrect components from existing digital elevation models (DEMs), such as Aw3D-30m v1 and SRTM3 v2.1. These data are freely available and accessible at [http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT\\_DEM/](http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/). Below are the details of this study's considered flood-influencing parameters.

**Elevation:** According to (Tehrany et al. 2013), there is a clear correlation between elevation and flooding, which makes lowland surfaces more susceptible to flooding than higher ones (Khosravi et al. 2016). This implies that the likelihood of flooding decreases with increasing topographic elevation (Youssef et al. 2016). The research



**Figure 4.** Flood influencing factors: (a) elevation, (b) slope, (c) aspect, (d) plan curvature, (e) hillshade, (f) horizontal flow distance, (g) rainfall, (h) land use/land cover, (i) SPI, and (j) geology.

area contains complicated topography characteristics, including very high elevations up to 2600 m and low altitudes ranging from 3 m to 200 m in the downstream section of the basin and the coastline area, primarily residential and agricultural areas (Figure 4a).

**Slope:** This is a significant factor influencing flooding (Khosravi et al. 2016; Tien Bui et al. 2016; Meraj et al. 2018) because of its effect on water velocity and surface flow (Torabi Haghghi et al. 2018). The study area's slope varied from  $0^{\circ}$  to  $70^{\circ}$  (Figure 4b).

**Aspect:** As stated by (Choubin et al. 2019), this aspect influences the hydrological parameters. There is an indirect relationship between the aspects of floods owing to their control over several geo-environmental factors, such as rainfall, vegetation, and soils (Rahmati et al. 2016). When aspects receive a low intensity of sunlight, which means more soil moisture, the moist slope will likely increase runoff, leading to increased flooding risk (Yariyan et al. 2020). The aspect raster map was categorized into ten classes from flat to the northwest (Figure 4c).

**Plan curvature:** Many researchers consider this an essential flood-controlling factor (Hong et al. 2018) and affect heterogeneity and hyporheic (Cardenas et al. 2004). The different values of curvatures differentiate the areas of faster runoff from those with slower runoff. While negative values cause an increase in runoff, positive values decrease it. The surface runoff is affected by the shape of the slope, as zero curvature (flat) and negative curvature (concave) have more potential for flooding than the convex form (positive) (Tehrany et al. 2015, 2014; Shahabi et al. 2021). Concave slopes, for instance, slow surface flow and increase filtering losses, while convex slopes do precisely the opposite of concave slopes (Cao et al. 2016). The curvature map was developed based on the DEM with three forms (concave, flat, and convex), and the flat class was more dominant in the downstream area, as shown in Figure 4d.

**Hillshade:** A hill's length and shadow are intertwined with its hillshade or topshade, which may affect where the surface flow converges (Aryal et al. 2003). Prior research has shown minimal interest in topshade (Bui, Ngo, et al. 2019). Predicting flooding vulnerability requires it after slope and elevation (Bui, Hoang, et al. 2019). Figure 4e shows that toposhade was chosen as a flood influencing factor.

**Flow distance:** Any area's likelihood of flooding is influenced mainly by distance from major rivers or streams (Glenn et al. 2012). Typically, nearby streams are more vulnerable to floods (Chapi et al. 2017). The farther away from rivers, the greater the chance of floods. Floods are common in places near rivers, which have been stressed as a primary influencing factor for flooding (Predick and Turner 2008; Bui et al. 2018; Darabi et al. 2019). According to Gigović et al. (2017) and González-Arqueros et al. (2018), the distance from streams is the primary conduit for surface flow. ArcGIS calculated the horizontal flow distance for the current investigation using the flow direction, flow accumulation, and DEM (Figure 4f.).

**Rainfall:** Precipitation is one of the triggering factors for flooding, as no rainfall indicates a lack of flooding. The average rainfall was estimated between 2001 and 2019 using the PERSIANN Dynamic Infrared–Rain rate model (PDIR). Estimating precipitation was done using remotely sensed information that utilizes ANNs (Nguyen et al. 2020). It is a real-time global dataset with a high resolution of approximately  $0.04^\circ \times 0.04^\circ$ , or  $4 \text{ km} \times 4 \text{ km}$ , at (<https://chrsdata.eng.uci.edu/>). According to the geographical maps, the average annual precipitation in the upstream and mountainous parts is 3284 mm, while 2235 mm in the downstream.

**Land use/land cover:** The influence of this factor was confirmed using the global cover map developed by the geospatial Japan information authority ([https://www.gsi.go.jp/kankyochiri/gm\\_global\\_e.html](https://www.gsi.go.jp/kankyochiri/gm_global_e.html)) and mainly from this website (<https://global-maps.github.io/glcnm.html>). Land use and land cover types were also considered as controlling factors due to their influence on filtration and runoff velocity. The study

area has approximately six classes (Figure 4h), including cropland, forest, grassland, other lands, settlement, and water. The forest is the dominant type of land cover in the mountainous area, especially upstream of the basins, and agricultural land and urban areas are located in the downstream region.

**Stream power index:** This parameter indicates the power of erosion and discharges within a specific area of the river system (Poudyal et al. 2010). Several researchers have considered the SPI a flooding contributor because it indicates surface flow. The highest values of SPI imply a fast flow of downstream water, which reveals lower flooding susceptibility, and low values imply slow flow leading to more inundation (Tehrany and Kumar 2018). SPI was calculated based on a method derived from Jebur et al. (2014). The present study area classified SPI into five classes (Figure 4i).

**Geology:** In terms of infiltration and flow velocity, this is a critical parameter. Lithology data was given by the Land Use and Climate Change Interaction in Central Vietnam (LUCCI) (Nauditt and Ribbe 2017). In terms of geological classification, it has been subdivided into several different kinds with high variation in sedimentary, igneous, and metamorphic rock types (Figure 4k).

### 3.2. Selection of flood influencing factors

The selection of controlling factors is an important stage in ML modeling for FSM. the estimated capabilities of the model may be impacted by an inaccurate selection of the hyperparameter values or redundancy (Öztürk and Akdeniz 2000). As a result, the feature selection procedure was based on Spearman's rank correlation, the IGR, and the multicollinearity test to identify irrelevant features.

#### 3.2.1. Spearman's correlation coefficient

The nonparametric Spearman rank correlation coefficient is used to show the strength of the monotonic association between two variables, X and Y. From  $-1$  to  $1$ , and the coefficient indicates more significant and weaker correlations. As the coefficient value approaches  $0$ , the relationship between the two variables, X and Y, becomes weaker. Correlation coefficient values above  $0.7$  imply considerable collinearity (Tien Bui et al. 2016). According to this formula, the correlation coefficient is calculated:

$$r(x, y) = 1 - \frac{6 \sum (x - y)^2}{n(n^2 - 1)} \quad (1)$$

where  $r$  refers to the correlation coefficient,  $x$ , and  $y$  are defined as the two variables, and  $n$  is the length of each variable.

#### 3.2.2. Multicollinearity test

Using Spearman's coefficient, multicollinearity was examined between all contributing elements and the correlations between two characteristics. The VIF was used in this study's multicollinearity analysis to identify any existing interrelatedness between variables. This element is frequently utilized in investigations of flood susceptibility (Bui et al. 2019; Khosravi et al. 2019; Rahman et al. 2019), suggesting a threshold  $> 5$  to



consider multicollinearity. The relevant predictors are, however, deemed collinear in other research if the VIF value is more than 10; hence it is advised to leave them out of the models (Dou et al. 2019; Wang et al. 2019). Thus, we considered a value of five as the threshold for selection. The independent predictors are specified as  $X = \{X_1, X_2, \dots, X_n\}$  and  $R_j^2$ , and refer to the determination coefficient when the  $j$ th independent predictor  $X_j$  is regressed on the other predictors. The following equation is used to determine VIF:

$$VIF = \frac{1}{1 - R_j^2} \quad (2)$$

### 3.2.3. Information gain ratio

The IGR test assessed conditioning factors to determine their relative relevance in floods (Quinlan 1986; Xu et al. 2013). The latter is one of the feature selection techniques considered by many previous studies (Shahabi et al. 2021). When an input has zero IGR, there is no correlation between the input and the output. This circumstance suggests that including such input in the model will not provide any information; rather, it will create noise, reducing the model's capacity for prediction. Therefore, it is strongly advised that these elements be eliminated from the inputs. Eq. (3) is used to compute the IGR.

$$IGR(x, Z) = \frac{Entropy(Z) - \sum_{i=1}^n \frac{|Z_i|}{|Z|} Entropy(Z_i)}{\sum_{i=1}^n \frac{|Z_i|}{|Z|} \log \frac{|Z_i|}{|Z|}} \quad (3)$$

### 3.3. Machine learning methods

ML approaches are the basic concept of employing algorithms to analyze and learn from the data to produce forecasting or classification systems. These techniques can be learned from previous experience or a given historical database. These methods can generalize the learning examples provided in the training phase to identify the main tasks that must be performed.

Several ML algorithms have been developed. These techniques can be classified according to their learning mechanisms (i.e. supervised, unsupervised, and semi-supervised learning). Choosing a suitable ML model and training method depends on the problems to be solved or the available data and its types. The current study focused on using supervised ML techniques for flooding susceptibility assessment. According to previous research, various ML techniques have been proposed recently to deal with flooding susceptibility assessment (i.e. SVMs, ELMs, ANNs, Gaussian process regression (GPR), and classification and regression trees (CART)). In addition, few studies have addressed flooding susceptibility using ensemble-learning approaches based on decision trees. These algorithms are based on boosting techniques concentrating on misclassified data during the training phase. In this respect,



this study aims to assess the performance of two new modeling techniques, CatBoost and LightGBM, benchmarked against the conventional RF approach.

### 3.3.1. *Random forest*

RF models have proven efficiency when dealing with prediction and classification problems (Esfandiari et al. 2020; Schoppa et al. 2020). RF is an ensemble learning approach based on a decision tree model. It was developed by Breiman (2001), who combined bagging (Breiman 2001) and random subspace (Ho 1998) techniques. This ML algorithm has proven reliable in many fields (Zahedi et al. 2018; Izquierdo-Verdiguier and Zurita-Milla 2020; Pourghasemi et al. 2020). This study aimed to predict flood or non-flood regions according to several conditioning factors; therefore, the RF model was used as a classifier method.

The weakness of decision trees is their sensitivity to training data, which may result in very different tree structures. In the RF method, the original training set is used to randomly generate several training sets, thereby allowing the creation of different trees (bagging method). The inputs of the decision trees have the same amount of data as the initial training, and because the data are randomly generated, the samples may be repeated two or more. In addition, each tree in the RF is trained with a subset of features that allows the development of diversified trees that are not correlated. The final result (classification) was obtained using a majority voting method on each decision tree result (Pal 2005).

Decision tree models are simple to use and easy to interpret; however, their performance is not always better than other classification methods (Malekipirbazari and Aksakalli 2015). On the other hand, RF outperformed other ML algorithms, such as ANNs (Bachmair et al. 2017).

### 3.3.2. *Light gradient boosting machine*

Microsoft created the gradient boosting decision tree (GBDT) variation called LightGBM. (Ke et al. 2017). It uses a combination of weak learners to generate a robust model. The new variant includes algorithms such as histograms, leaf tree growth, gradient-based one-side sampling (GOSS), and exclusive feature bundling (EFB)).

In GBDT models, the presorted algorithm is commonly used for split operations. All possible split points are tested based on the information gained, which is time-consuming to determine the optimal split. A new histogram algorithm was adopted in the LightGBM method. To reduce the time and complexity of the operation, the data are grouped into a histogram, and the split point is chosen based on it (Figure 5).

In LightGBM, the decision tree growth strategy was changed by replacing the level-wise approach with the leaf-wise tree growth approach. When finding the best node to split, the former approach of the GBDT splits one level down, forming symmetric trees (Figure 6). In LightGBM, only the leaves that reduced the maximum error were split (Figure 6). Ge et al. (2019) recommended defining a maximum leaf-wise depth to avoid deep growth of trees and prevent overfitting of the model.

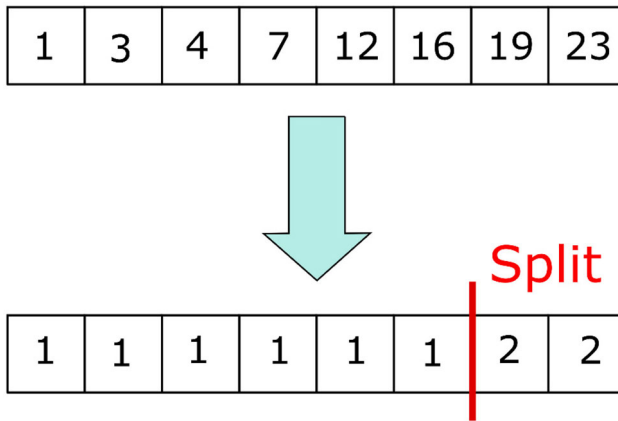


Figure 5. Split operation example based on histogram algorithm.

Level-wise tree growth

Leafwise tree growth

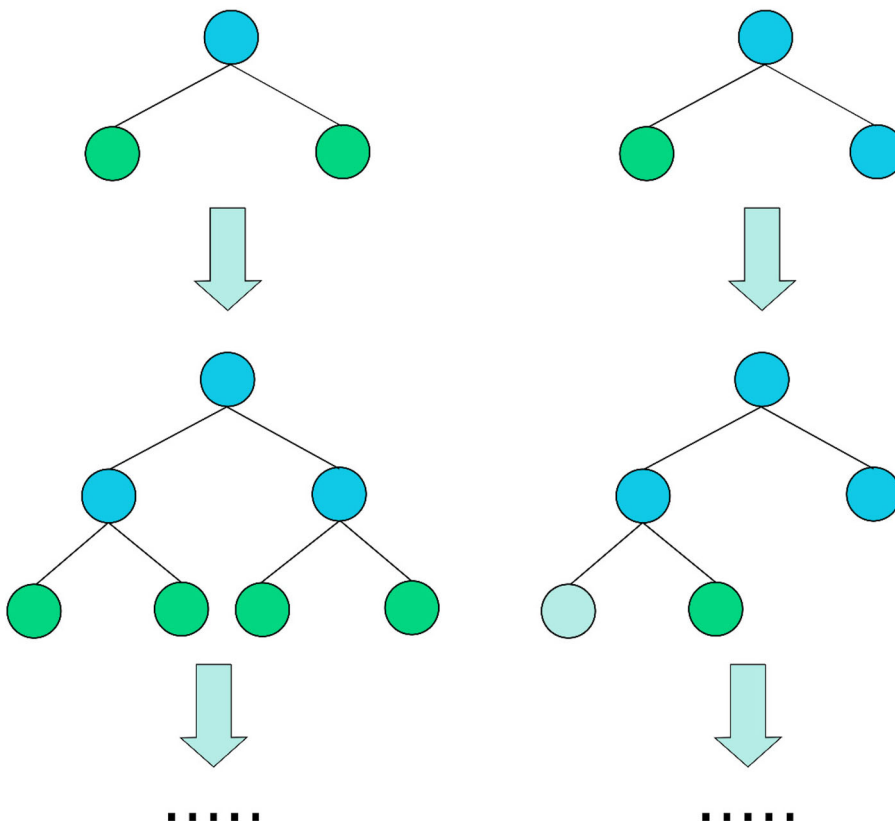


Figure 6. Level-wise and leafwise tree growth strategies.

The LightGBM model also uses two algorithms (GOSS and EFB), making it faster than GBDT models while maintaining a high performance (Saber et al. 2021).

### 3.3.3. Categorical boosting

The CatBoost model is another enhanced boosting decision-tree learning technique proposed by (Dorogush et al. 2018). It employs a gradient-boosting scheme to construct a regression model through adjusted estimation. Furthermore, various refinements were performed to minimize the overfitting of the model. The gradient boosting model is a useful ML tool that has yielded accurate results in many disciplines, including environmental parameter estimation, geospatial ecosystem factor dispersion, and meteorological forecasting. The CatBoost model operates well in terms of categorical attributes. Typically, the absence of definite characteristics increases the accuracy of the model. It is primarily dependent on the use of gradient boosting, which employs a binary-tree classification scheme. The following points outline the differences between CatBoost and the other boosting techniques.

- A sophisticated method was incorporated to convert category characteristics into numerical information. As mentioned by (Prokhorenkova et al. 2018), target statistics are very effective for dealing with categorical attributes with minimal information errors.
- CatBoost combines categorical variables to take advantage of the existing relationship between different parameters.
- To reduce the overfitting problem and improve the classification performance, a symmetrical tree strategy is used.

Let us suppose we have a dataset:

$$D = \{(X_J, Y_J)\} \quad J = 1, \dots, m \quad (4)$$

where  $X_J = (x_j^1, x_j^2, \dots, x_j^n)$  is a combination of attributes, and  $Y_J \in R$ , denotes the desired target. Input-output datasets are dispersed identically and independently depending on an unknown function  $\rho(\cdot, \cdot)$ . The target of the learning techniques is to train and examine a function  $H : R^n \rightarrow R$  that can decrease information loss, that is,  $L(H) := EL(y, H(X))$ , where  $L$  is the smoothness error function and  $(X, y)$  denotes the testing samples from  $D$ . The gradient boosting approach builds a greedy series of approximations  $H_t: R^n \rightarrow R$ ,  $t = 0, 1, 2, \dots$ ,  $H_t = H((t-1)) + g_t$  is the final function produced from prior approximation using an additive process  $H_t = H((t-1)) + g_t$ .

$$g^t = \arg \min_{g \in G} L(H^{t-1} + g) = \arg \min_{g \in G} E L(y, H^{t-1}(X)) \quad (5)$$

In general, greedy techniques, such as Newton's method, employing a second-order approach of  $L(H(t-1) + g)$  at  $H(t-1)$  or adopting (negative) gradient stages, are used to address the optimization issue.

### 3.4. Rainfall-runoff inundation model (RRI)

Japan's International Center for Water Hazard and Risk Management has developed the RRI model. It is a 2D distributed hydrological model capable of simultaneously simulating flow discharge and flood inundation (Sayama et al. 2012). The model has been applied in many previous studies worldwide (Perera et al. 2017; Abdel-Fattah et al. 2018; Tam et al. 2019; Saber et al. 2020; Try et al. 2020). In this study, the model was calibrated and validated based on the typhoon of 2020, showing acceptable results with the actual flood discharge and good agreement with flood inundation maps. The final flooding inundation map developed was used for comparison with the ML FSMs for the flood extent mapping.

### 3.5. Evaluation of the model's performance validation

The receiver operating characteristic (ROC) curve measure is a commonly used and validated strategy for assessing the reliability of a model in geospatial research (Tehrany et al. 2013; Chen et al. 2020). The most popular method for evaluating flood vulnerability and landslide approaches is the ROC curve. The classification performance of a given technique was evaluated using the AUC in several studies (Bui et al. 2012; Youssef et al. 2016; Youssef and Hegab 2019). A high classification efficiency for a given classification model should have an AUC-ROC value of 0.– 1, and the model's performance is enhanced by boosting the AUC-ROC scores. When the AUC-ROC value was close to 1.0, the models offered the best rate of precision and consistency. This demonstrates the model's ability to forecast disasters without bias (Bui et al. 2012). In this study, the ROC score was determined using the following formula (Chang et al. 2018):

Other quantitative metrics (accuracy, recall, precision, and F1-score) were employed to check the model performance and compare its classification ability with its counterpart models in the literature. Accuracy is the ratio of correctly classified data to total observations [Eq. (6)]; precision can be defined by the ratio of properly positive classified data to total positive data [Eq. (7)]. Recall, is known as sensitivity and is defined by the ratio of positive to the total observations [Eq. (8)]. F1-score uses weighted averaging for both the precision and recall [Eq. (9)].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$F1 \text{ score} = \frac{2 (Recall * Precision)}{Recall + Precision} \quad (9)$$

**Table 1.** Spearman's correlation coefficients for flooding susceptibility mapping.

	Aspect	Rainfall	Curvature	DEM	NDVI	Flow acc.	Flow dir.	H. Dist.	V. Dist.	Geology	Hillshade	Land use	Slope	SPI	TWI
Aspect	1.00														
Rainfall	-0.03	1.00													
Curvature	0.08	0.01	1.00												
DEM	0.03	0.20	0.27	1.00											
NDVI	0.04	-0.03	0.20	0.72	1.00										
Flow acc.	-0.07	0.01	-0.48	0.05	0.06	1.00									
Flow dir.	0.06	-0.07	0.11	0.06	0.06	-0.11	1.00								
H. Dist.	-0.04	0.03	0.02	0.49	0.39	0.08	0.02	1.00							
V. Dist.	0.04	0.17	0.28	0.97	0.72	0.00	0.06	0.52	1.00						
Geology	0.00	0.29	-0.02	0.08	-0.02	-0.01	-0.03	-0.09	0.06	1.00					
Hillshade	0.38	-0.05	0.08	0.03	0.02	-0.05	0.51	-0.06	-0.03	-0.06	1.00				
Land use	-0.03	-0.10	-0.11	0.59	-0.53	-0.13	-0.05	-0.25	-0.58	-0.07	0.00	1.00			
Slope	0.10	0.07	0.28	0.85	0.69	0.03	0.08	0.38	0.86	0.06	0.01	-0.53	1.00		
SPI	0.04	0.03	-0.41	0.32	0.27	0.86	-0.06	0.21	0.30	-0.01	-0.05	-0.27	0.38	1.00	
TWI	-0.10	-0.07	-0.49	0.76	-0.61	0.35	-0.12	-0.31	-0.79	-0.07	-0.04	0.45	-0.90	0.00	1.00

where true positive (TP) represents a properly categorized flooded pixel, true negative (TN) represents a correctly categorized non-flood pixel, false positive (FP) indicates the number of pixels miscategorized as flood pixels, and false negative (FN) refers to the number of pixels miscategorized as non-flood pixels.

## 4. Results and discussion

### 4.1. Multicollinearity assessment and feature selection

According to Chen et al. (2020), a value greater than 0.7 indicates a strong correlation between variables. This value was adopted in this study to detect the existence of a correlation between the flood-influencing factors. Ten conditioning factors (DEM, NDVI, flow accumulation, vertical distance from the river, and slope) were identified as correlated (Table 1). The VIF of the vertical distance from the river (= 12), DEM (= 10.5), SPI (= 7.7), and flow accumulation (= 7.4) factors were more significant than the threshold value ( $> 5$ ), which indicates a problem of multicollinearity (Figure 7a).

To formulate an opinion on the importance of influencing factors concerning flood generation, the IGR scores were computed and illustrated in Figure 7b. According to the results; most factors had an IGR greater than 0.05. Only four had an inferior IGR, that is, flow accumulation, flow direction, rainfall, and aspect.

The selection of conditioning factors was performed as follows:

1. Based on multicollinearity analysis, the vertical distance from the river, DEM, SPI, and flow accumulation factors were removed from the selection list.
2. Using the IGR as a selection criterion, flow direction, rainfall, and aspect were removed because their IGR was almost equal to zero.
3. After removing the abovementioned factors, only the slope and the topographic wetness index (TWI) remained correlated. By comparing the IGR (Figure 7), we find that the slope factor is more critical than the TWI concerning flood generation. Therefore, the slope factor was selected for flood prediction based on the normalized difference vegetation index (NDVI), land use, curvature, geology, hillshade, and horizontal distance from the river.

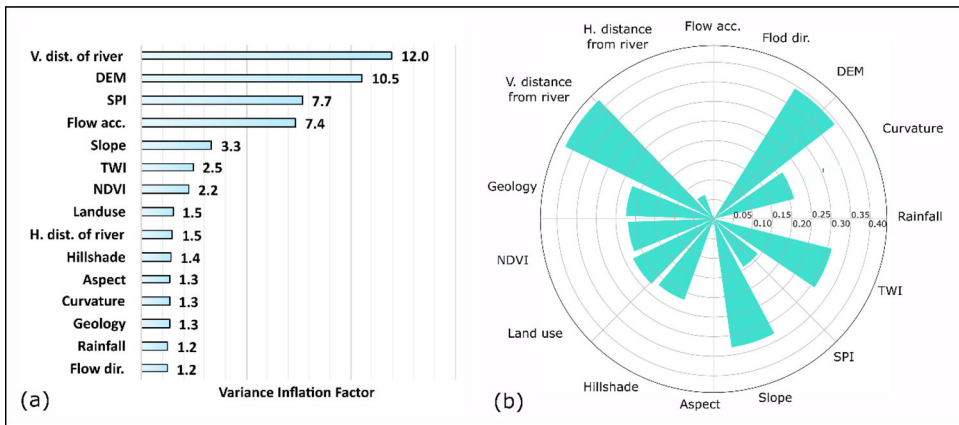


Figure 7. Analysis of influencing factors: (a) VIF and (b) IGR for flood susceptibility.

Table 2. Parameter values of random forest, CatBoost, and LightGBM models.

Method	Hyperparameter	Grid search values	Selected value
RF	Max depth	[2:2:40]	22
	Number of trees	[100:100:2000]	800
LightGBM	Learning rate	[0.01:0.01:0.1]	0.09
	Max depth	[2:2:40]	32
	Number of leaves	[20:20:300]	220
CatBoost	Min data	[5:5:100]	25
	Learning rate	[0.01:0.01:0.1]	0.07
	Max depth	[2:2:40]	6
	Leaf estimation iterations	[1:1:10]	8
	l2 leaf reg	Log [10 (-21),10 (-8)]	Log (10 (-18))

#### 4.2. Evaluation of the models

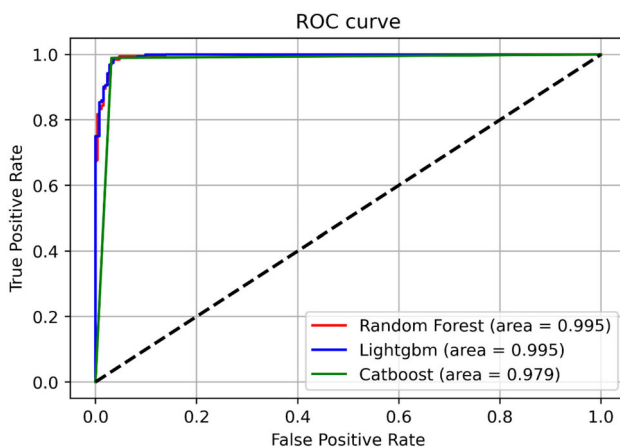
This section offers a thorough analysis and comparison of all models created for this research concerning several categorization criteria. K-fold cross-validation was used throughout the learning phase. A learning set (60%) was created from the reviewed data, and the remaining data were used to gauge accuracy. The learning datasets were divided into two groups: validation data, which was used for hyperparameter tuning, and training data (80%), which was used to modify and reduce classification mistakes and model weights. The relevant hyperparameters for each classification method were selected using the grid search method. A broad range of hyperparameter values was evaluated during the process. The best designs for each classifier are listed in Table 2.

The accuracy rates of all the studied models are listed in Table 3. As can be seen, all developed classification techniques achieved approximately identical results in terms of statistical metrics. The LightGBM model slightly outperformed the others regarding speed convergence and classification metrics. The ROC curve of the generated models on the test ensembles is displayed in Figure 8, which reveals that the three suggested boosting strategies have similar qualities and provide significant accuracy. The maximum AUC was reached by LightGBM and RF models with the same score (99.5%), and CatBoost was ranked as the worst model with an AUC of 97.9%.



**Table 3.** Statistical parameters used for the model performance evaluation.

Criteria	Random Forest		LightGBM		CatBoost	
	Train	Test	Train	Test	Train	Test
Accuracy	0.999	0.955	0.999	0.973	0.999	0.978
Precision	0.998	0.962	0.998	0.950	0.998	0.960
Recall	1.000	0.932	1.000	0.990	1.000	0.990
F1_score	0.999	0.947	0.999	0.969	0.999	0.974
AUC	1.000	0.995	1.000	0.995	0.999	0.979

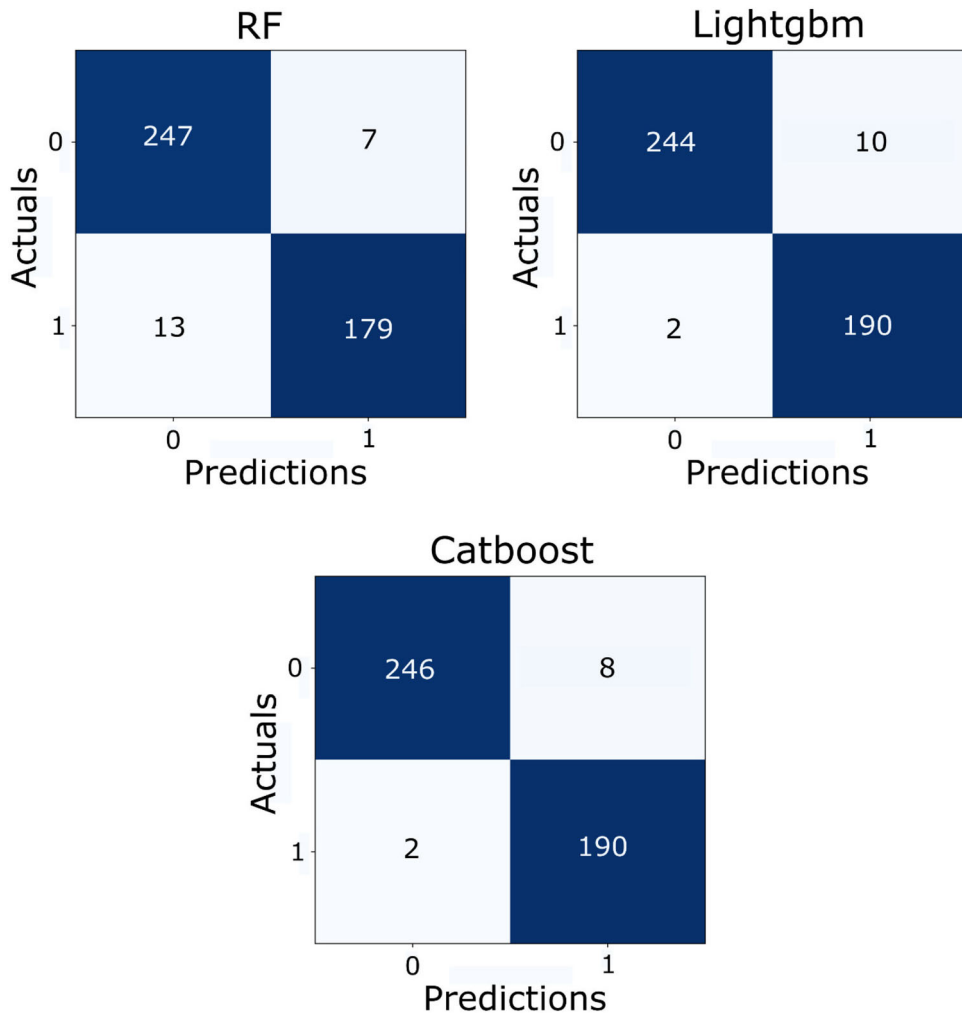
**Figure 8.** Performance of random forest, CatBoost, and LightGBM models based on AUC-ROC curves.

Furthermore, CatBoost scored the first rank in terms of accuracy, with performance accuracy equal to 97.8% and precision equal to 96%, accompanied by the LightGBM classifier with an accuracy of 97.3% and precision of 95%. Finally, the RF model was ranked as the last classifier model with an accuracy equal to 95.5% and a precision of 96.2%. In comparison with previous studies, RF in this study outperformed many of the previous applications, including (e.g. AUC = 0.925, Chen et al. (2020); AUC = 0.886, Tang et al. (2020); AUC = 0.7878, Lee et al. (2017); AUC = 0.972, Achour and Pourghasemi et al. (2020)).

The confusion matrix in Figure 9. shows the performance of the used models in the study area, where CatBoost shows better prediction followed by LightGBM and, finally, the RF methods; however, all of them display an acceptable prediction.

This study examined two novel boosting classification models for flooding susceptibility assessment in the VGTB River Basin. From the evaluation statistics, we can conclude that the LightGBM and CatBoost models proved their performance for flooding susceptibility and can be used as essential tools for real-time application compared to their counterpart models because of their high performance and speed convergence.

This work examined two novel boosting classification approaches for predicting flood vulnerability in the VGTB. This is the first work investigating CatBoost and LightGBM for flood classification in humid environments against the frequently used



**Figure 9.** Confusion matrix showing the performance of the used models in VGTB River Basin.

RF models. The results revealed that LightGBM outperformed its counterpart ML models, especially regarding processing time and classification metrics. This agrees with the findings of Saber et al. (2021) that LightGBM has proven its efficiency in flash flood prediction and outperforms the other two methods in classification and processing time. In addition, it was stated that LightGBM outperformed other methods, such as the RF, M5Tree, and empirical models for estimating daily evapotranspiration in China as a humid subtropical area (Fan et al. 2019).

Similarly, it was also found that LightGBM performed better than the others in terms of AUC (99.5%). The accuracy of CatBoost (97.9%) was also high compared to the previous studies in other fields. Among other methods, CatBoost, SVM, and RF have been applied to evapotranspiration modeling in China (Huang et al. 2019). They stated that CatBoost presented higher accuracy and lowered computational cost than the other approaches (RF and SVM).

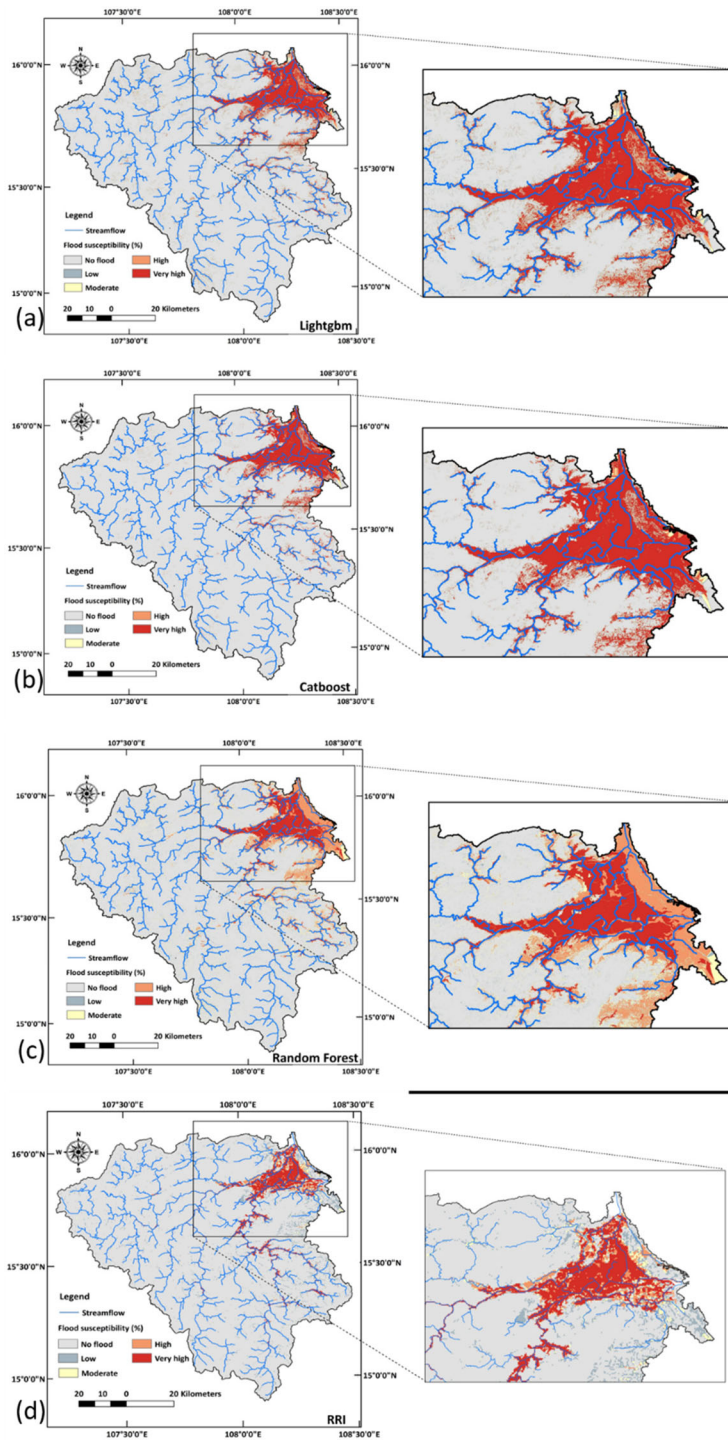
### 4.3. Flood susceptibility modeling

The newly applied boosting techniques (CatBoost and LightGBM) and RF demonstrated their high performance in predicting flooding in a humid climate environment. The flood susceptibility maps for the whole VGTB river basin were thus estimated using these approaches. Then, the three FSMs developed using the three models were compared with the flood inundation map of the RRI model regarding the flood extent, as shown in [Figure 10](#). The flooding susceptibility values were then mapped under five levels of susceptibility classes: no flood, low, moderate, high, and very high.

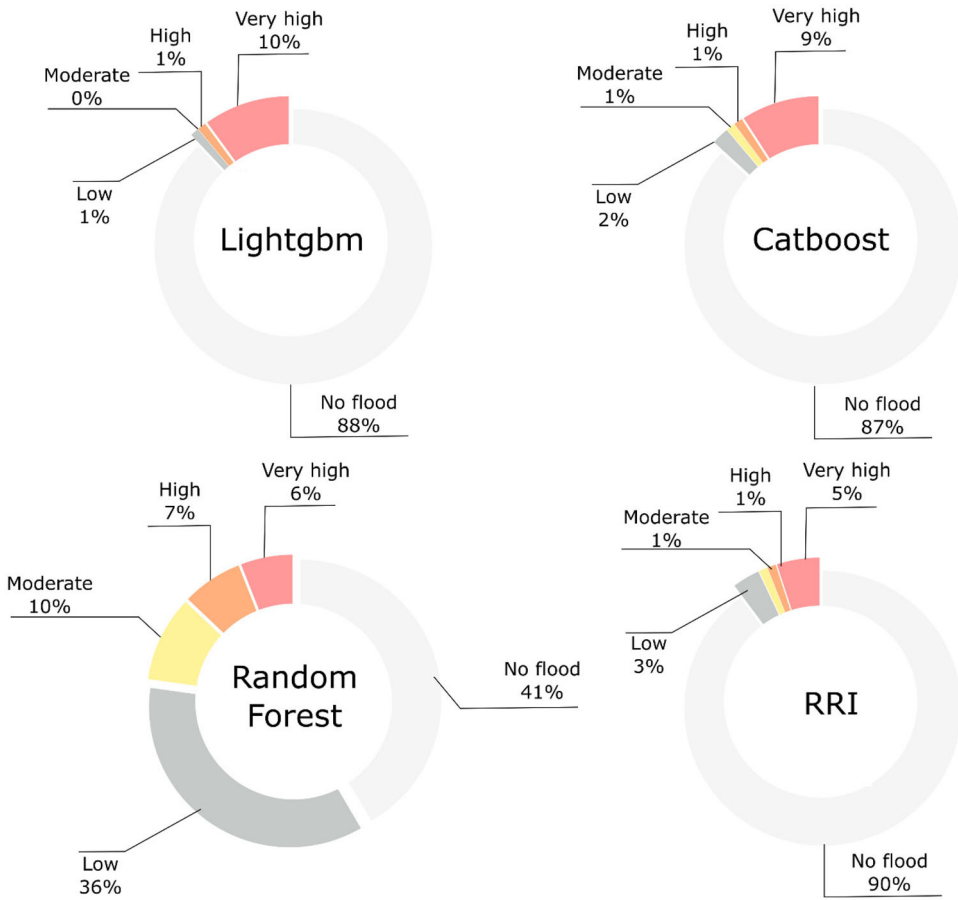
The FSMs by the employed models showed that the areas of high and very high levels of susceptibility to flooding to be 13% (RF), 11% (LightGBM), and 10% (CatBoost) of the total area, which agrees with the flood inundation map developed by RRI at approximately 11%. This level of susceptibility is predominant in the coastal and plain areas along the Vu Gia and Thu Bon Rivers ([Figures 10 and 11](#)). The spatial distributions of the high and very high levels were similar in all the maps produced by the ML and RRI models. The areas affected by a moderate level of susceptibility to flooding ([Figure 11](#)) were estimated at 10% (RF), 0% (LightGBM), and 1% (CatBoost), indicating that both LightGBM and CatBoost are more similar to the RRI model which shows a value of approximately 1%. The areas affected by the low level of susceptibility to flooding ([Figure 11](#)) were estimated at 36% (RF), 1% (LightGBM), and 2% (CatBoost), which also revealed that both LightGBM and CatBoost performed better, with good agreement with the RRI model showing the value of approximately 3%. It was also found that the areas that were not subjected to the flooding were approximately 42% (RF), 83% (LightGBM), and 87% (CatBoost) of the total study area ([Figure 11](#)), showing good agreement with RRI model that shows approximately 90%. However, the performances of the employed models are almost the same. The two new methods of LightGBM and CatBoost outperform RF in terms of the spatial coverage of the flood susceptibility levels compared with the RRI model. The RF overestimated the low flood susceptibility in the study area. The spatial distribution of FSM is consistent across utilized ML models, emphasizing that most of the residential and agricultural sectors are concentrated in coastal regions prone to flooding.

### 4.4. Testing different sizes of the datasets

In this section, we tested different sizes of the datasets, including flooded and non-flooded points (1250, 1000, 800, 600, 400, 90, 60, and 30) of the training model ([Figure 12](#)). The training datasets were classified as 50% and 50% for flooded and non-flooded points, respectively; however, the testing datasets were the same during the simulation ([Figure 12](#)). We found that accuracy scores for all the models and all the tested cases were greater than 90% ([Figure 13](#)), except for dataset sizes of 60 and 30 points in LightGBM. The accuracy score slightly decreases with the decrease in the datasets in both the LightGBM and CatBoost models but is inconsistent in the RF model. This implies that the ML approaches employed in this study can effectively work with very limited training datasets with a slight decrease in accuracy, which will



**Figure 10.** Flood susceptibility maps by LightGBM (a), CatBoost (b), RF (c), and RRI (d), respectively, from top to bottom.



**Figure 11.** Affected area of the flood susceptibility levels for the three applied ML methods and flood inundation map of RRI model.

be applicable for ungauged regions with deficient monitoring and observations of flooding occurrences and impacts. The FSMs developed based on different training datasets show that most spatial maps are acceptable as overall spatial coverage; however, there are some small spatial differences in the susceptibility flooding levels (Figures 14 and 15). For instance, all datasets (1250, 1000, 800, 600, 400, 90, 60, and 30) except for the dataset of 200 had almost the same percentage of impacted regions (Figures 14 and 15) in the category of extremely high flood susceptibility. On the other hand, the affected areas by the high flood susceptibility level also vary. Still, the highest percentage was 9% for the 200 and 60 datasets, and the lowest was 6% for the 30, 90, 600, and 800 datasets. The variation in moderate flood occurrence ranged from 17% to 9%. The dataset size of 30 was the highest among the others, about 17%. The range of the low flood susceptibility category was highly variable, from 20% to 41%, the lowest was for the dataset of 30, and the highest was for the dataset of 800. The reasons for such variation are probably the random selection of the flooded samples, which in some cases are not representative of all the influencing factors. We noticed that the spatial coverage was not extremely different, but some differences were observed based

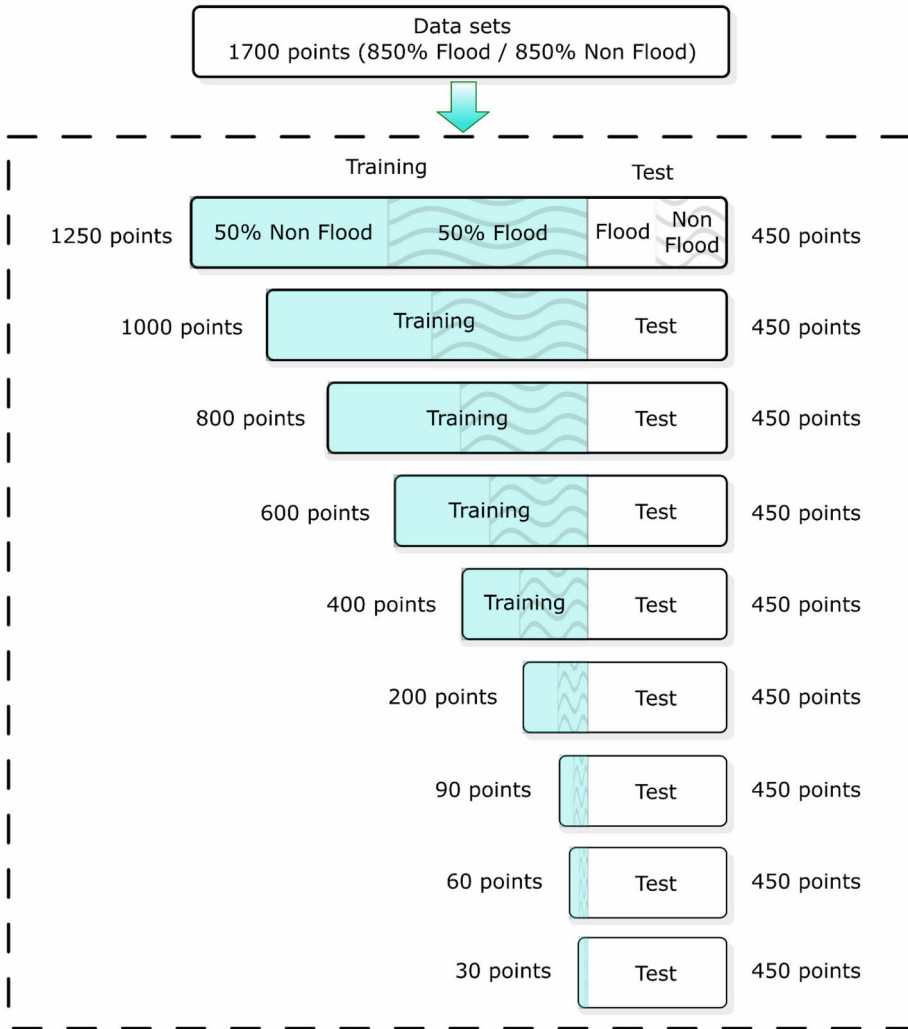


Figure 12. Datasets used in the training and testing of the ML models.

on the categories. The areas with no flood levels are also changeable by about 42%, 42%, 38%, 45%, 44%, 45%, 42%, 44%, and 51% for the datasets of 1250, 1000, 800, 600, 400, 200, 90, 60, and 30, respectively. Interestingly, the highest percentage was recorded by the dataset of 30 points and the lowest by dataset 800. The analysis of different data sizes for ML training show that ML can effectively predict the flood susceptibility maps in the study area regardless of the number of samples, with the condition of the used data being observational flooded sites.

#### 4.5. Discussion and comparison with results of the RRI model

The flood risk assessment scientific community is endeavoring to develop much more logical and mathematical methods for FSM forecasting at different catchment scales



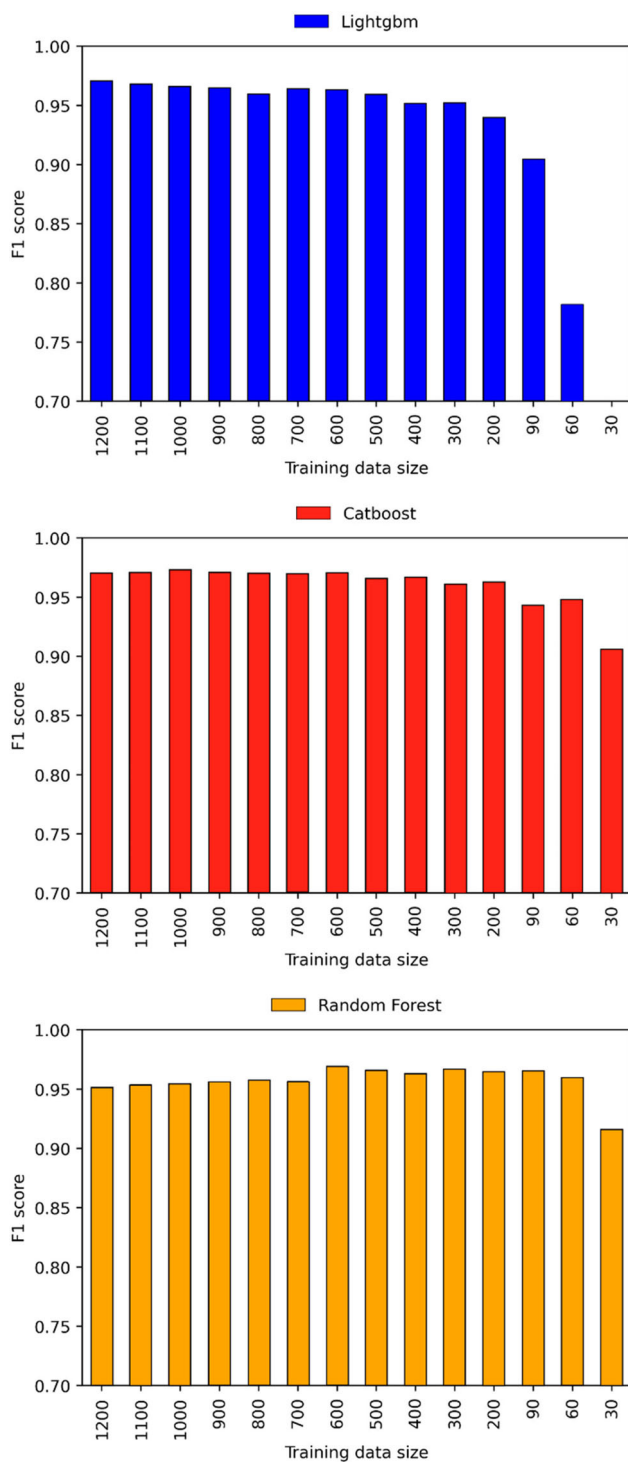


Figure 13. Accuracy of the models based on different training datasets.

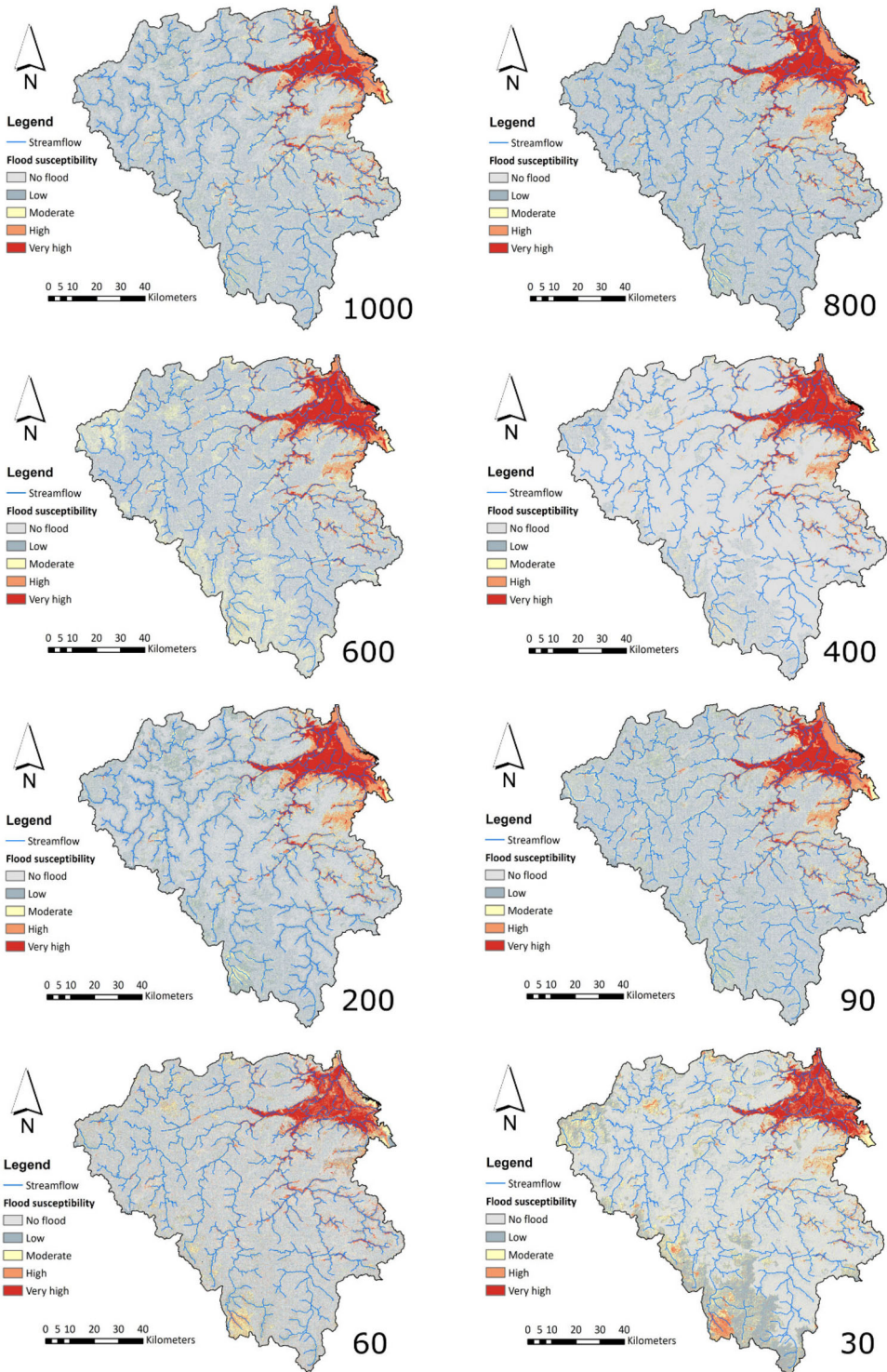
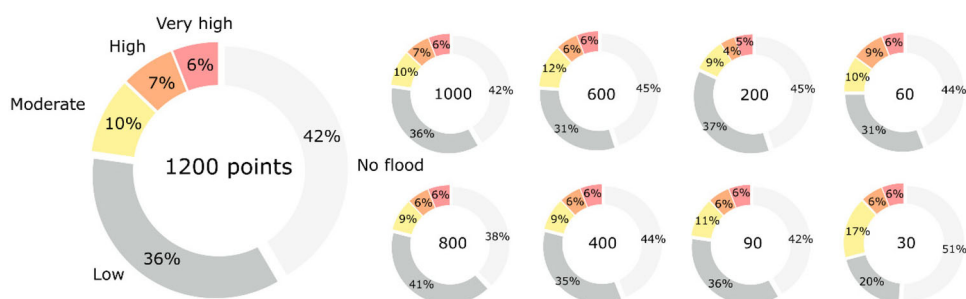


Figure 14. Impact of data size on flood susceptibility map.



**Figure 15.** Percentage of the affected areas under different flood susceptibility classes using different dataset sizes (RF method).

(Arora et al. 2021). Some previous studies on flood susceptibility mapping use ML approaches and deep learning in the study area. Testing many models is therefore strongly advised, especially in areas with little data and complex hydrological models. This study applies three ML methods: LightGBM, CatBoost, and RF. The LightGBM and CatBoost techniques were put to the test for the first time for mapping flood susceptibility in this humid area with a high frequency of typhoon occurrences. Compared to the commonly used RF approach, the findings of the flooding susceptibility maps show that the two methods can forecast flood-prone regions with respectable accuracy. AUC = 78% (Band et al. 2020), 99.3% (Li et al. 2019), 94.5% (Talukdar et al. 2020), 93.8% (Park and Lee 2020), and 89.4% (Talukdar et al. 2020) use RFs in several additional related research with varying degrees of accuracy (Nguyen et al. 2018). Compared to most earlier research, the AUC = 99% for RF in this study was greater.

Additionally, the newly applied methods of LightGBM and CatBoost showed almost the same accuracy of 99% and 98%, respectively, revealing better performance than most previous studies. These three methods have been tested in Hurghada, Egypt (Saber et al. 2021), stating that LightGBM has the advantage of better classification metrics and fast processing time and outperforms other methodologies such as CatBoost and RF. In addition, their results showed that LightGBM and CatBoost had proven their efficiency in flash flood prediction in arid regions (Saber et al. 2021).

The three techniques also outperformed the 90% average performance of previously used methods for mapping flood susceptibility, which was based on an average of about 140 prior applications from more than 30 papers that have been analyzed. Based on AUC, the effectiveness of the prior techniques used for FSM ranges from 64 percent (Shafizadeh-Moghadam et al. 2018) to 99.3% (Li et al. 2019). CatBoost was also applied in Germany, with better performance than other methods, showing good accuracy with an AUC of 0.816 (Kaiser (2021)).

The maps of flood susceptibility developed using ML techniques (Figure 10) showed an acceptable fit with the generated flood inundation map by the RRI model, showing that the ML approaches are promising for flood prediction and can be used without detailed observations and challenges of model calibrations as alternative tools for hydrological models. The results of LightGBM and CatBoost are more comparable to the flood inundation map developed by the physical RRI model, indicating that

they are more acceptable than RF, which overestimates the low flood susceptibility level in the study area.

The flood susceptibility maps developed using these models showed that high and very high levels of susceptibility to flooding covered 13% (RF), 11% (LightGBM), and 10% (CatBoost) of the total area, respectively. These levels of susceptibility were mainly concentrated in the coastal and plain areas along the Vu Gia and Thu Bon Rivers. The moderate level of susceptibility to flooding covered 10% (RF), 0% (LightGBM), and 1% (CatBoost) of the total area, while the low level of susceptibility covered 36% (RF), 1% (LightGBM), and 2% (CatBoost). The areas not susceptible to flooding were 42% (RF), 83% (LightGBM), and 87% (CatBoost) of the total area. The spatial distribution of FSMs is consistent across utilized ML models, emphasizing that most of the residential and agricultural sectors are concentrated in coastal regions prone to flooding. The two new methods, LightGBM and CatBoost, outperform RF in terms of the spatial coverage of the flood susceptibility levels compared to the RRI model.

Furthermore, we tested different datasets for training the three ML models, concluding that datasets with more than 90 points can be sufficiently accurate for reasonable prediction of the FSM. LightGBM and CatBoost showed a slightly declining trend in the accuracy of the results based on the dataset sizes; however, RF did not show such a trend. These results are precious for applying ML to ungauged basins with limited datasets.

## 5. Conclusions

Flooding resulting from typhoons is one of the most threatening disasters in Asian countries and worldwide. Therefore, the present study introduced three ML methods to accurately predict flooding susceptibility in humid Vietnamese areas. The first method is RF, which is well known and widely applicable in many applications, including FSM, and the other methods of LightGBM and CatBoost were examined for the first time for FSM in this humid region. On the basis of a flood inventory map and ten flood-influencing factors, the models were trained and validated. Owing to the availability of high-quality observations, we also tested different datasets for the training (i.e. 30, 60, 90, 200, 400, 600, 800, 1000, and 1250 data points) to determine the minor data points that provided acceptable reliability, as well as to understand the differences in the spatial FSMs in the study area.

Interestingly, we found that the accuracy of results based on all the tested datasets was higher than 90%, indicating that a limited number of observations can be used efficiently in model accuracy. However, the final FSMs differed spatially from one susceptibility level to the others. This finding is significant to demonstrate that ML methods can work efficiently with an acceptable level of accuracy within a small number of actual training datasets. The conclusions of this study can be summarized as follows:

- We applied three ML models—RF, LightGBM, and CatBoost—to predict flood susceptibility in humid areas that experienced successive extreme typhoons.
- The LightGBM and CatBoost models were tested for the first time in this specific climatic region and showed high performance compared to the RF method.

- The results of the ML methods showed good agreement with the rainfall-runoff model for flood inundation mapping, especially the LightGBM and CatBoost models, in terms of coverage areas of the flood susceptibility levels.
- Different training datasets were examined to determine ML's lowest acceptable number of observations for flooding susceptibility.
- The FSMs demonstrated that downstream areas with high residential and agricultural activity are highly susceptible to flooding.
- These results might be utilized as a guide and reference for flood risk reduction and management in this region, assisting managers, decision-makers, and planners in successfully managing and reducing floods in high-risk flood zones.

The study's conclusion highlights the effectiveness of the ensemble learning approach in accurately predicting flood susceptibility, with a high level of agreement with hydrological models in flood mapping. The results of this study demonstrate the potential of ensemble learning algorithms, such as CatBoost, LightGBM, and RF, to provide valuable insights into the assessment of flood risk and support the development of effective flood management strategies. However, it is important to note that while these models provide a powerful tool for flood prediction, they are not without limitations. One of the main limitations of this study, as the majority of the studies utilized machine learning (ML) models, is that the presented susceptibility map has not considered return periods in spatial modeling of the flood because the selection of training and validation points has not been based on the return periods due to lack of hazard maps for each return period. Further research is needed to address these limitations and improve the accuracy and reliability of these models in others regions of the world.

In light of the findings, an ongoing extension of this research aims to utilize the power of machine learning algorithms, combined with physically-based hydrological models, to predict flood depth. This effort seeks to develop a more comprehensive and accurate understanding of flood risk and provide decision-makers with valuable information to support evidence-based decision-making for flood mitigation and adaptation. Overall, the study provides promising results for the application of machine learning in flood susceptibility mapping and the prediction of flood depth, and it opens the door for future research in this field.

## Acknowledgment

This study was supported by the JSPS Core-to-Core Program (Grant number: JPJSCCB20220004); JSPS Aid for Scientific Research (KAKENHI) Program 2023: Grant number: 23K04328. Data sources include the Geospatial Information Authority of Japan at Chiba University and collaborating organizations. The Ministry of Higher Education of the Arab Republic of Egypt for providing a full scholarship to the author Karim Abdrabo

## Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work presented in this study.



## Funding

This work was funded by the Asia-Pacific Network for Global Change Research (APN) under project reference number CRRP2020-09MYKantoush (Funder ID: <https://doi.org/10.13039/100005536>).

## Data availability statement

The data used in this study are available from the corresponding author, Mohamed Saber, upon reasonable request.

## References

- Abdel-Fattah M, Kantoush SA, Saber M, Sumi T. 2018. Rainfall-runoff modeling for extreme flash floods in wadi samail, oman. *J JSCE, Ser B1*. 74(5):I\_691–I\_696.
- Abdrabo KI, Kantosh SA, Saber M, Sumi T, Elleithy D, Habiba OM, Alboshy B. 2022. The Role of Urban Planning and Landscape Tools Concerning Flash flood risk reduction within arid and semiarid regions. In: *Wadi Flash Floods*. Springer, p. 283–316.
- Abdrabo KI, Kantoush SA, Esmail A, Saber M, Sumi T, Almamari M, Elboshy B, Ghoniem S. 2023. An integrated indicator-based approach for constructing an urban flood vulnerability index as an urban decision-making tool using the PCA and AHP techniques: a case study of Alexandria, Egypt. *Urban Clim*. 48:101426.
- Abdrabo KI, Kantoush SA, Saber M, Sumi T, Habiba OM, Elleithy D, Elboshy B. 2020. Integrated methodology for urban flood risk mapping at the microscale in ungauged regions: a case study of Hurghada, Egypt. *Remote Sens*. 12(21):3548.
- Abdrabo KI, Saber M, Kantoush SA, ElGharbawi T, Sumi T, Elboshy B. 2022. Applications of remote sensing for flood inundation mapping at urban areas in MENA Region: case studies of five Egyptian cities. In: *applications of space techniques on the natural hazards in the MENA region*. Cham: Springer International Publishing, p. 307–330.
- Abushandi EH, Merkel BJ. 2011. Application of IHACRES rainfall-runoff model to the Wadi Dhuliel arid catchment, Jordan. *J Water Clim Change*. 2(1):56–71.
- Akay AE, Taş İ. 2020. Mapping the risk of winter storm damage using GIS-based fuzzy logic. *J Res*. 31(3):729–742.
- Ali SA, Parvin F, Pham QB, Vojtek M, Vojteková J, Costache R, Linh NTT, Nguyen HQ, Ahmad A, Ghorbani MA. 2020. GIS-based comparative assessment of flood susceptibility mapping using hybrid multi-criteria decision-making approach, naïve Bayes tree, bivariate statistics and logistic regression: a case of Topľa basin, Slovakia. *Ecol Indic*. 117:106620.
- Arabameri A, Saha S, Mukherjee K, Blaschke T, Chen W, Ngo PTT, Band SS. 2020. Modeling spatial flood using novel ensemble artificial intelligence approaches in Northern Iran. *Remote Sens*. 12(20):3423.
- Arora A, Arabameri A, Pandey M, Siddiqui MA, Shukla UK, Bui DT, Mishra VN, 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.
- Arora A, Pandey M, Siddiqui MA, Hong H, Mishra VN. 2019. Spatial flood susceptibility prediction in Middle Ganga Plain: comparison of frequency ratio and Shannon's entropy models. *Geocarto Int*. 36(18):2085–2116.
- Aryal SK, Mein RG, O'Loughlin EM. 2003. The concept of effective length in hillslopes: assessing the influence of climate and topography on the contributing areas of catchments. *Hydrol Process*. 17(1):131–151.,
- Ashley ST, Ashley WS. 2008. Flood fatalities in the United States. *J Appl Meteorol Climatol*. 47(3):805–818.



- Avitabile V, Schultz M, Herold N, De Bruin S, Pratihast AK, Manh CP, Quang HV, Herold M. 2016. Carbon emissions from land cover change in Central Vietnam. *Carbon Manag.* 7(5-6):333–346.
- Aydin HE, Iban MC. 2022. Predicting and analyzing flood susceptibility using boosting-based ensemble machine learning algorithms with SHapley Additive exPlanations. *Nat Hazards.* 1–35.
- Bachmair S, Svensson C, Prosdocimi I, Hannaford J, Stahl K. 2017. Developing drought impact functions for drought risk management. *Nat Hazards Earth Syst Sci.* 17(11):1947–1960.
- Band SS, Janizadeh S, Chandra Pal S, Saha A, Chakraborty R, Melesse AM, Mosavi A. 2020. Flash flood susceptibility modeling using new approaches of hybrid and ensemble tree-based machine learning algorithms. *Remote Sens.* 12(21):3568.
- Bisht S, Chaudhry S, Sharma S, Soni S. 2018. Assessment of flash flood vulnerability zonation through Geospatial technique in high altitude Himalayan watershed, Himachal Pradesh India. *Remote Sens Appl Soc Environ.* 12:35–47.
- Breiman L. 2001. Random forests. *Mach Learn.* 45(1):5–32.
- Bui DT, Hoang N-D, Pham T-D, Ngo P-TT, Hoa PV, Minh NQ, 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.
- Bui DT, Ngo P-TT, Pham TD, Jaafari A, Minh NQ, Hoa PV, Samui P. 2019. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. *CATENA.* 179:184–196.
- Bui DT, Ngo P-TT, Pham TD, Jaafari A, Minh NQ, Hoa PV, Samui P. 2019b. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. *Catena* 179:184–196.
- Bui DT, Panahi M, Shahabi H, Singh VP, Shirzadi A, Chapi K, Khosravi K, Chen W, Panahi S, Li S, et al. 2018. Novel hybrid evolutionary algorithms for spatial prediction of floods. *Sci Rep.* 8(1):14.
- Bui DT, Pradhan B, Lofman O, Revhaug I, Dick OB. 2012. Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): a comparative assessment of the efficacy of evidential belief functions and fuzzy logic models. *Catena.* 96:28–40.
- Bui Q-T, Nguyen Q-H, Nguyen XL, Pham VD, Nguyen HD, 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.
- Cao Y, Nishihara R, Qian ZR, Song M, Mima K, Inamura K, Nowak JA, et al. Regular aspirin use associates with lower risk of colorectal cancers with low numbers of tumor-infiltrating lymphocytes. *Gastroenterology.* 151(5):879–892.
- Cardenas MB, Wilson J, Zlotnik VA. 2004. Impact of heterogeneity, bed forms, and stream curvature on subchannel hyporheic exchange. *Water Resour Res.* 40(8):1–13.
- Chang M-J, Chang H-K, Chen Y-C, Lin G-F, Chen P-A, Lai J-S, Tan Y-C. 2018. A support vector machine forecasting model for typhoon flood inundation mapping and early flood warning systems. *Water* 10(12):1734.
- Chapi K, Singh VP, Shirzadi A, Shahabi H, Bui DT, Pham BT, Khosravi K. 2017. A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environ Model Softw.* 95:229–245.
- Chen T-HK, Qiu C, Schmitt M, Zhu XX, Sabel CE, Prishchepov AV. 2020. Mapping horizontal and vertical urban densification in Denmark with Landsat time-series from 1985 to 2018: a semantic segmentation solution. *Remote Sens Environ.* 251:112096.
- Choubin B, Hosseini FS, Rahmati O, Youshanloei MM. 2023. A step toward considering the return period in flood spatial modeling. *Nat Hazards.* 115(1):431–460.
- Choubin B, Moradi E, Golshan M, Adamowski J, Sajedi-Hosseini F, Mosavi A. 2019. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Sci Total Environ.* 651(Pt 2): 2087–2096.

- Costache R, Hong H, Pham QB. 2020. Comparative assessment of the flash-flood potential within small mountain catchments using bivariate statistics and their novel hybrid integration with machine learning models. *Sci Total Environ.* 711:134514.
- Costache R, Popa MC, Tien Bui D, Diaconu DC, Ciubotaru N, Minea G, Pham QB. 2020. Spatial predicting of flood potential areas using novel hybridizations of fuzzy decision-making, bivariate statistics, and machine learning. *J Hydrol.* 585:124808.
- Darabi H, Choubin B, Rahmati O, Haghghi AT, Pradhan B, Kløve B. 2019. Urban flood risk mapping using the GARP and QUEST models: a comparative study of machine learning techniques. *J Hydrol.* 569:142–154.
- Demirel MC, Venancio A, Kahya E. 2009. Flow forecast by SWAT model and ANN in Pracana basin, Portugal. *Adv Eng Softw.* 40(7):467–473.
- Devkota KC, Regmi AD, Pourghasemi HR, Yoshida K, Pradhan B, Ryu IC, Dhital MR, Althuwaynee OF. 2013. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. *Nat Hazards.* 65(1):135–165.
- Dhara S, Dang T, Parial K, Lu XX. 2020. Accounting for uncertainty and reconstruction of flooding patterns based on multi-satellite imagery and support vector machine technique: a case study of Can Tho City, Vietnam. *Water.* 12(6):1543.
- Dodangeh E, Choubin B, Eigdir AN, 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.
- Dorogush AV, Ershov V, Gulin A. 2018. CatBoost: gradient boosting with categorical features support. *ArXiv Prepr. ArXiv181011363.* 1–7.
- Dou J, Yunus AP, Bui DT, Merghadi A, Sahana M, Zhu Z, Chen C-W, Khosravi K, Yang Y, Pham BT. 2019. Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. *Sci Total Environ.* 662:332–346.
- Esfandiari M, Jabari S, McGrath H, Coleman D. 2020. flood mapping using random forest and identifying the essential conditioning factors; a case study in Fredericton, New Brunswick, Canada. *ISPRS AnnPhotogramm Remote Sens Spat Inf Sci.* 5:609–615.
- Esmail A, Abdrabo KI, Saber M, Sliuzas RV, Atun F, Kantoush SA, Sumi T. 2022. Integration of flood risk assessment and spatial planning for disaster management in Egypt. *Prog. Disaster Sci.* 15:100245.
- Fan J, Wu L, Zhang F, Cai H, Ma X, Bai H. 2019. Evaluation and development of empirical models for estimating daily and monthly mean daily diffuse horizontal solar radiation for different climatic regions of China. *Renew Sust Energ Rev.* 105:168–186.
- Fenicia F, Kavetski D, Savenije HH, Clark MP, Schoups G, Pfister L, Freer J. 2014. Catchment properties, function, and conceptual model representation: is there a correspondence? *Hydrol Process.* 28(4):2451–2467.
- Ge F, Zhu S, Peng T, Zhao Y, Sielmann F, Fraedrich K, Zhi X, Liu X, Tang W, Ji L. 2019. Risks of precipitation extremes over Southeast Asia: does 1.5° C or 2° C global warming make a difference?. *Environ Res Lett.* 14(4):044015.
- Gharakhanlou NM, Perez L. 2023. Flood susceptible prediction through the use of geospatial variables and machine learning methods. *J Hydrol.* 617:129121.
- Gigović L, Pamučar D, Bajić Z, Drobnjak S. 2017. Application of GIS-interval rough AHP methodology for flood hazard mapping in urban areas. *Water* 9(6):360.
- Glenn EP, Morino K, Nagler PL, Murray RS, Pearlstein S, Hultine KR. 2012. Roles of saltcedar (*Tamarix* spp.) and capillary rise in salinizing a non-flooding terrace on a flow-regulated desert river. *J Arid Environ.* 79:56–65.
- González-Arqueros ML, Mendoza ME, Bocco G, Castillo BS. 2018. Flood susceptibility in rural settlements in remote zones: the case of a mountainous basin in the Sierra-Costa region of Michoacán, Mexico. *J Environ Manage.* 223:685–693.
- Hirabayashi Y, Mahendran R, Koirala S, Konoshima L, Yamazaki D, Watanabe S, Kim H, Kanae S. 2013. Global flood risk under climate change. *Nature Clim Change.* 3(9):816–821.

- Hörling B, Coldewey WG. 2019. Surface water infiltration. *Hydrogeology*. Springer, p. 33–37.
- Hong H, Tsangaratos P, Ilia I, Liu J, Zhu A-X, Chen W. 2018. Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China. *Sci Total Environ*. 625:575–588.
- Ho TK. 1998. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell*. 20(8):832–844.
- Hsu K, Gupta HV, Sorooshian S. 1995. Artificial neural network modeling of the rainfall-run-off process. *Water Resour Res*. 31(10):2517–2530.
- Huang G, Wu L, Ma X, Zhang W, Fan J, Yu X, Zeng W, Zhou H. 2019. Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. *J Hydrol*. 574:1029–1041.
- Humphrey GB, Gibbs MS, Dandy GC, Maier HR. 2016. A hybrid approach to monthly streamflow forecasting: integrating hydrological model outputs into a Bayesian artificial neural network. *J Hydrol*. 540:623–640.
- IPCC. 2007. The physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Camb Univ Press Camb U KN Y NY USA. 996:113–119.
- Izquierdo-Verdiguier E, Zurita-Milla R. 2020. An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing. *Int J Appl Earth Obs Geoinformation*. 88:102051.
- Jebur MN, Pradhan B, Tehrany MS. 2014. Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. *Remote Sens Environ*. 152:150–165.
- Jonkman SN, Kelman I. 2005. An analysis of the causes and circumstances of flood disaster deaths. *Disasters*. 29(1):75–97.
- Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu T-Y. 2017. Lightgbm: a highly efficient gradient boosting decision tree. *Adv Neural Inf Process Syst*. 30:3146–3154.
- Khosravi K, Nohani E, Maroufinia E, Pourghasemi HR. 2016. A GIS-based flood susceptibility assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision-making technique. *Nat Hazards*. 83(2):947–987.
- Khosravi K, Shahabi H, Pham BT, Adamowski J, Shirzadi A, Pradhan B, Dou J, Ly H-B, Gróf G, Ho HL, et al. 2019. A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods. *J Hydrol*. 573:311–323.
- Kia MB, Pirasteh S, Pradhan B, Mahmud AR, Sulaiman WNA, Moradi A. 2012. An artificial neural network model for flood simulation using GIS: johor River Basin, Malaysia. *Environ Earth Sci*. 67(1):251–264.
- Lee S, Kim JC, Jung HS, Lee MJ, Lee S. 2017. Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. *Geomat Nat Hazards Risk*. 8(2), 1185–1203.
- Li X, Yan D, Wang K, Weng B, Qin T, Liu S. 2019. Flood risk assessment of global watersheds based on multiple machine learning models. *Water* 11(8):1654.
- Luu C, Von Meding J, Kanjanabutra S. 2018. Assessing flood hazard using flood marks and analytic hierarchy process approach: A case study for the 2013 flood event in Quang Nam. Vietnam *Nat Hazards*. 90:1031–1050.
- Luu C, Pham BT, Phong TV, Costache R, Nguyen HD, Amiri M, Bui QD, Nguyen LT, Le HV, Prakash I, et al. 2021. GIS-based ensemble computational models for flood susceptibility prediction in the Quang Binh Province, Vietnam. *J Hydrol*. 599:126500.
- Malekipirbazari M, Aksakalli V. 2015. Risk assessment in social lending via random forests. *Expert Syst Appl*. 42(10):4621–4631.
- Masood M, Takeuchi K. 2012. Assessment of flood hazard, vulnerability and risk of mid-eastern Dhaka using DEM and 1D hydrodynamic model. *Nat Hazards*. 61(2):757–770.

- Meraj G, Khan T, Romshoo SA, Farooq M, Rohitashw K, Sheikh BA. 2018. An integrated geoinformatics and hydrological modelling-based approach for effective flood management in the Jhelum Basin, NW Himalaya. *Multidiscip Digit Publ Inst Proc.* 7:8.
- Nauditt A, Firoz ABM, Trinh VQ, Fink M, Stolpe H, Ribbe L. 2017. Hydrological drought risk assessment in an anthropogenically impacted tropical catchment, Central Vietnam. In: A. Nauditt, L. Ribbe editors, *Land use and climate change interactions in central Vietnam*. Singapore: Springer, p. 223–239.
- Nauditt A, Ribbe L. 2017. *Land use and climate change interactions in central Vietnam*. Berlin: Springer Nature.
- Ngo P-TT, Pham TD, Hoang N-D, Tran DA, Amiri M, Le TT, Hoa PV, Bui PV, Nhu V-H, Bui DT. 2021. A new hybrid equilibrium optimized SysFor based geospatial data mining for tropical storm-induced flash flood susceptible mapping. *J Environ Manage.* 280:111858.
- Nguyen HD, Quang-Thanh B, Nguyen Q-H, Nguyen TG, Pham LT, Nguyen XL, Vu PL, Thanh Nguyen TH, Nguyen AT, Petrisor A-I. 2022. A novel hybrid approach to flood susceptibility assessment based on machine learning and land use change. Case study: a river watershed in Vietnam. *Hydrol Sci J.* 67(7):1065–1083.
- Nguyen HQ, Degener J, Kappas M. 2015. Flash flood prediction by coupling KINEROS2 and HEC-RAS models for tropical regions of Northern Vietnam. *Hydrology* 2(4):242–265.
- Nguyen P, Ombadi M, Goroooh VA, Shearer EJ, Sadeghi M, Sorooshian S, Hsu K, Bolvin D, Ralph MF. 2020. PERSIANN dynamic infrared–rain rate (PDIR-Now): a near-real-time, quasi-global satellite precipitation dataset. *J Hydrometeorol.* 21(12):2893–2906.
- Nguyen V-N, Tien Bui D, Ngo P-TT, Nguyen Q-P, Nguyen VC, Long NQ, Revhaug I. 2018. An Integration of least squares support vector machines and firefly optimization algorithm for flood susceptible modeling using GIS. In: D. Tien Bui, A. Ngoc Do, H.-B. Bui, N.-D. Hoang editors, *Advances and applications in geospatial technology and earth resources*. Cham: Springer International Publishing, p. 52–64.
- Nguyen V-N, Yariyan P, Amiri M, Dang Tran A, Pham TD, Do MP, Thi Ngo PT, Nhu V-H, Quoc Long N, Tien Bui D. 2020. A new modeling approach for spatial prediction of flash flood with biogeography optimized CHAID tree ensemble and remote sensing data. *Remote Sens.* 12(9):1373.
- Nhu V-H, Thi Ngo P-T, Pham TD, Dou J, Song X, Hoang N-D, Tran DA, Cao DP, Aydilek İB, Amiri M, et al. 2020. A new hybrid firefly–PSO optimized random subspace tree intelligence for torrential rainfall-induced flash flood susceptible mapping. *Remote Sens.* 12(17): 2688.
- Öztürk F, Akdeniz F. 2000. Ill-conditioning and multicollinearity. *Linear Algebra Appl.* 321(1–3):295–305.
- IPCC. 2014. *Climate Change 2014: Synthesis Report*. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. Geneva, Switzerland: IPCC.
- Pal M. 2005. Random forest classifier for remote sensing classification. *Int J Remote Sens.* 26(1):217–222.
- Park S-J, Lee D-K. 2020. Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms. *Environ Res Lett.* 15(9):094052.
- Perera EDP, Sayama T, Magome J, Hasegawa A, Iwami Y. 2017. RCP8. 5-based future flood hazard analysis for the lower Mekong river basin. *Hydrology* 4(4):55.
- Pham BT, Jaafari A, Nguyen-Thoi T, Van Phong T, Nguyen HD, Satyam N, Masroor M, Rehman S, Sajjad H, Sahana M, et al. 2021. Ensemble machine learning models based on reduced error pruning tree for prediction of rainfall-induced landslides. *Int J Digit Earth.* 14(5):575–596.
- Pham BT, Luu C, Phong TV, Trinh PT, Shirzadi A, Renoud S, Asadi S, Le HV, von Meding J, Clague JJ. 2021. Can deep learning algorithms outperform benchmark machine learning algorithms in flood susceptibility modeling? *J Hydrol.* 592:125615.

- Pham TD, Xia J, Ha NT, Bui DT, Le NN, Tekeuchi W. 2019. A review of remote sensing approaches for monitoring blue carbon ecosystems: mangroves, seagrasses and salt marshes during 2010–2018. *Sensors* 19(8):1933.
- Poudyal CP, Chang C, Oh HJ, Lee S. 2010. Landslide susceptibility maps comparing frequency ratio and artificial neural networks: a case study from the Nepal Himalaya. *Environ Earth Sci.* 61:1049–1064.
- Pourghasemi HR, Kariminejad N, Amiri M, Edalat M, Zarafshar M, Blaschke T, Cerda A. 2020. Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. *Sci Rep.* 10(1):11.
- Predick KI, Turner MG. 2008. Landscape configuration and flood frequency influence invasive shrubs in floodplain forests of the Wisconsin River (USA). *J Ecol.* 96:91–102.
- Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. 2018. CatBoost: unbiased boosting with categorical features. *Adv Neural Inf Process Syst.* 31.
- Quinlan JR. 1986. Induction of decision trees. *Mach Learn.* 1(1):81–106.
- Rahman M, Chen N, Elbeltagi A, Islam MM, Alam M, Pourghasemi HR, Tao W, Zhang J, Shufeng T, Faiz H, et al. 2021. Application of stacking hybrid machine learning algorithms in delineating multi-type flooding in Bangladesh. *J Environ Manage.* 295:113086.,
- Rahman M, Chen N, Islam MM, Mahmud GI, Pourghasemi HR, Alam M, Rahim MA, Baig MA, Bhattacharjee A, Dewan A. 2021b. Development of flood hazard map and emergency relief operation system using hydrodynamic modeling and machine learning algorithm. *J Clean Prod.* 311:127594.
- Rahman M, Ningsheng C, Islam MM, Dewan A, Iqbal J, Washakh RMA, Shufeng T. 2019. Flood susceptibility assessment in Bangladesh using machine learning and multi-criteria decision analysis. *Earth Syst Environ.* 3(3):585–601.
- Rahmati O, Pourghasemi HR, Zeinivand H. 2016. Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golistan Province, Iran. *Geocarto Int.* 31(1):42–70.
- RETA. 2011. Investment, Managing water in Asia's river basins: charting progress and facilitating – The Vu Gia-Thu Bon Basin.
- Saber M, Abdrabo KI, Habiba OM, Kantosh SA, Sumi T. 2020. Impacts of triple factors on flash flood vulnerability in Egypt: urban growth, extreme climate, and mismanagement. *Geosciences* 10(1):24.
- Saber M, Boulmaiz T, Guermoui M, Abdrabo KI, Kantoush SA, Sumi T, Boutaghane H, Nohara D, Mabrouk E. 2021. Examining LightGBM and CatBoost models for wadi flash flood susceptibility prediction. *Geocarto Int.* 37(25):7462–7487.
- Saber M, Kantoush SA, Abdel-Fattah M, Sumi T, Moya JA, Abdrabo K. 2022. Flash flood modeling and mitigation in arid and semiarid basins: case studies from Oman and Brazil. In: *Wadi Flash Floods*. Singapore: Springer, , p. 355–381.
- Sayama T, Ozawa G, Kawakami T, Nabesaka S, Fukami K. 2012. Rainfall–runoff–inundation analysis of the 2010 Pakistan flood in the Kabul River basin. *Hydrol Sci J.* 57(2):298–312.
- Schoppa L, Disse M, Bachmair S. 2020. Evaluating the performance of random forest for large-scale flood discharge simulation. *J Hydrol.* 590:125531.
- Seydi ST, Kanani-Sadat Y, Hasanlou M, Sahraei R, Chanussot J, Amani M. 2022. Comparison of Machine Learning Algorithms for Flood Susceptibility Mapping. *Remote Sens.* 15:192.
- Shafizadeh-Moghadam 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.
- Shahabi H, Shirzadi A, Ronoud S, Asadi S, Pham BT, Mansouripour F, Geertsema M, Clague JJ, Bui DT. 2021. Flash flood susceptibility mapping using a novel deep learning model based on deep belief network, back propagation and genetic algorithm. *Geosci Front.* 12(3): 101100.
- Shirzadi A, Asadi S, Shahabi H, Ronoud S, Clague JJ, Khosravi K, Pham BT, Ahmad BB, Bui DT. 2020. A novel ensemble learning based on Bayesian Belief Network coupled with an



- extreme learning machine for flash flood susceptibility mapping. *Eng Appl Artif Intell.* 96: 103971.
- Talukdar S, Ghose B, Salam R, Mahato S, Pham QB, Linh NTT, Costache R, Avand, M, Shahfahad. 2020. Flood susceptibility modeling in Teesta River basin, Bangladesh using novel ensembles of bagging algorithms. *Stoch Environ Res Risk Assess.* 34(12):2277–2300.
- Tam TH, Abd Rahman MZ, Harun S, Hanapi MN, Kaoje IU. 2019. Application of Satellite rainfall products for flood inundation modelling in Kelantan River Basin, Malaysia. *Hydrology* 6(4):95.
- Tang X, Li J, Liu M, Liu W, Hong H. 2020. Flood susceptibility assessment based on a novel random Naïve Bayes method: A comparison between different factor discretization methods. *Catena.* 190:104536.
- Tehrany MS, Kumar L. 2018. The application of a Dempster–Shafer-based evidential belief function in flood susceptibility mapping and comparison with frequency ratio and logistic regression methods. *Environ Earth Sci.* 77:1–24.
- Tehrany MS, Pradhan B, Jebur MN. 2015. Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. *Stoch Environ Res Risk Assess.* 29(4):1149–1165.
- Tehrany MS, Pradhan B, Jebur MN. 2014. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J Hydrol.* 512:332–343.
- Tehrany MS, Pradhan B, Jebur MN. 2013. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *J Hydrol.* 504:69–79.
- Thao NTP, Linh TT, Ha NTT, Vinh PQ, Linh NT. 2020. Mapping flood inundation areas over the lower part of the Con River basin using Sentinel 1A imagery. *Vietnam J Earth Sci.* 42: 288–297.
- Tien Bui D, Hoang N-D. 2017. A Bayesian framework based on a Gaussian mixture model and radial-basis-function Fisher discriminant analysis (BayGmmKda V1. 1) for spatial prediction of floods. *Geosci Model Dev.* 10(9):3391–3409.
- Tien Bui D, Hoang N-D, Martínez-Álvarez F, Ngo P-TT, Hoa PV, Pham TD, Samui P, Costache R. 2020. A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area. *Sci Total Environ.* 701: 134413.
- Tien Bui D, Pradhan B, Nampak H, Bui Q-T, Tran Q-A, Nguyen Q-P. 2016. Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibility modeling in a high-frequency tropical cyclone area using GIS. *J Hydrol.* 540:317–330.
- Tiwari MK, Chatterjee C. 2010. Uncertainty assessment and ensemble flood forecasting using bootstrap based artificial neural networks (BANNs). *J Hydrol.* 382(1-4):20–33.
- Torabi Haghighi A, Menberu MW, Darabi H, Akanegbu J, Kløve B. 2018. Use of remote sensing to analyse peatland changes after drainage for peat extraction. *Land Degrad Develop.* 29(10):3479–3488.
- Try S, Tanaka S, Tanaka K, Sayama T, Oeurng C, Uk S, Takara K, Hu M, Han D. 2020. Comparison of gridded precipitation datasets for rainfall-runoff and inundation modeling in the Mekong River Basin. *PLoS One.* 15(1):e0226814.
- Tuyen TT, Jaafari A, Yen HPH, Nguyen-Thoi T, Phong TV, Nguyen HD, Van Le H, Phuong TTM, Nguyen SH, Prakash I, et al. 2021. Mapping forest fire susceptibility using spatially explicit ensemble models based on the locally weighted learning algorithm. *Ecol. Inform.* 63: 101292.
- UN Habitat. 2013. Quang Nam Provincial Socio-Economic Development: orientation To 2020 And Vision To 2025.
- United Nations Economic and Social Commission for Asia and the Pacific. 2018. Da Nang City, Viet Nam.
- Vinet F. 2008. Geographical analysis of damage due to flash floods in southern France: the cases of 12–13 November 1999 and 8–9 September 2002. *Appl. Geogr.* 28(4):323–336.

- Vu TTL, Nguyen LD, Hoang TS, Bui TA, Nguyen MT, Nguyen TH. 2011. Solutions for flood and drought prevention and mitigation in Quang Nam.
- Wang X, Mahul O, Stutley C. 2010. Weathering the storm: options for disaster risk financing in Vietnam. *The World Bank, Washington DC.*
- Wang Y, Hong H, Chen W, Li S, Panahi M, Khosravi K, Shirzadi A, Shahabi H, Panahi S, Costache R. 2019. Flood susceptibility mapping in Dingnan County (China) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm. *J Environ Manage.* 247:712–729.
- Xu Y, Dai Y, Dong ZY, Zhang R, Meng K. 2013. Extreme learning machine-based predictor for real-time frequency stability assessment of electric power systems. *Neural Comput Applic.* 22(3-4):501–508.
- Yamazaki D, Ikeshima D, Tawatari R, Yamaguchi T, O'Loughlin F, Neal JC, Sampson CC, Kanae S, Bates PD. 2017. A high-accuracy map of global terrain elevations. *Geophys Res Lett.* 44(11):5844–5853.
- Yang S, Yang D, Chen J, Santisirisomboon J, Lu W, Zhao B. 2020. A physical process and machine learning combined hydrological model for daily streamflow simulations of large watersheds with limited observation data. *J Hydrol.* 590:125206.
- Yariyan P, Janizadeh S, Van Phong T, Nguyen HD, Costache R, Van Le H, Pham BT, Pradhan B, Tiefenbacher JP. 2020. Improvement of best first decision trees using bagging and dagging ensembles for flood probability mapping. *Water Resour Manage.* 34(9):3037–3053.
- Youssef AM, Hegab MA. 2019. Flood-hazard assessment modeling using multicriteria analysis and GIS: a case study—Ras Gharib area, Egypt. *Spatial modeling in GIS and R for Earth and environmental sciences.* Elsevier, p. 229–257.
- Youssef AM, Pradhan B, Sefry SA. 2016. Flash flood susceptibility assessment in Jeddah city (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models. *Environ Earth Sci.* 75:12.
- Zahedi P, Parvande S, Asgharpour A, McLaury BS, Shirazi SA, McKinney BA. 2018. Random forest regression prediction of solid particle Erosion in elbows. *Powder Technol.* 338:983–992.
- Zenggang X, Zhiwen T, Xiaowen C, Xue-Min Z, Kaibin Z, Conghuan Y. 2021. Research on image retrieval algorithm based on combination of color and shape features. *J Sign Process Syst.* 93(2-3):139–146.