GIS Based Hybrid Computational Approaches for Flash Flood Susceptibility Assessment

Binh Thai Pham [1,*], Mohammadtaghi Avand [2,*], Saeid Janizadeh [2], Tran Van Phong [3], Nadhir Al-Ansari [4,*], Lanh Si Ho [5,*], Sumit Das [6], Hiep Van Le [1], Ata Amini [7], Saeid Khosrobeigi Bozchaloei [8], Faeze Jafari [2], and Indra Prakash [9]

[1] University of Transport Technology, Hanoi 100000, Vietnam; hieplv@utt.edu.vn

[2] Department of Watershed Management Engineering, College of Natural Resources, Tarbiat Modares University, Tehran 14115-111, Iran; Janizadeh.saeed@gmail.com (S.J.); Faeze_Jafari86@yahoo.com (F.J.)

[3] Institute of Geological Sciences, Vietnam Academy of Sciences and Technology, 84 Chua Lang Street, Dong da, Hanoi 100000, Vietnam; tphong1617@gmail.com

[4] Department of Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, 971 87 Lulea, Sweden; nadhir.alansari@ltu.se

[5] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

[6] Department of Geography, Savitribai Phule Pune University, Pune 411007, India; sumit.das.earthscience@gmail.com

[7] Kurdistan Agricultural and Natural Resources Research and Education Center, AREEO, Sanandaj 66177-15175, Iran; a.amini@areeo.ac.ir

[8] Department of Watershed Management Engineering, College of Natural Resources, Tehran University, Tehran 1417414418, Iran; saeid.khosro@yahoo.com

[9] Department of Science & Technology, Bhaskarcharya Institute for Space Applications and Geo-Informatics (BISAG), Government of Gujarat, Gandhinagar 382007, India; indra52prakash@gmail.com

* Correspondence: binhpt@utt.edu.vn (B.T.P); mt.avand70@gmail.com (M.A.); nadhir.alansari@ltu.se (N.A.-A.); hosilanh@duytan.edu.vn (L.S.H.).

Abstract: Flash floods are one of the most devastating natural hazards; they occur within a catchment (region) where the response time of the drainage basin is short. Identification of probable flash flood locations and development of accurate flash flood susceptibility maps are important for proper flash flood management of a region. With this objective, we proposed and compared several novel hybrid computational approaches of machine learning methods for flash flood susceptibility mapping, namely AdaBoostM1 based Credal Decision Tree (ABM-CDT); Bagging based Credal Decision Tree (Bag-CDT); Dagging based Credal Decision Tree (Dag-CDT); MultiBoostAB based Credal Decision Tree (MBAB-CDT), and single Credal Decision Tree (CDT). These models were applied at a catchment of Markazi state in Iran. About 320 past flash flood events and nine flash flood influencing factors, namely distance from rivers, aspect, elevation, slope, rainfall, distance from faults, soil, land use, and lithology were considered and analyzed for the development of flash flood susceptibility maps. Correlation based feature selection method was used to validate and select the important factors for modeling of flash floods. Based on this feature selection analysis, only eight factors (distance from rivers, aspect, elevation, slope, rainfall, soil, land use, and lithology) were selected for the modeling, where distance to rivers is the most important factor for modeling of flash flood in this area. Performance of the models was validated and compared by using several robust metrics such as statistical measures and Area Under the Receiver Operating Characteristic (AUC) curve. The results of this study suggested that ABM-CDT (AUC = 0.957) has the best predictive capability in terms of accuracy, followed by Dag-CDT (AUC = 0.947), MBAB-CDT (AUC = 0.933), Bag-CDT (AUC = 0.932), and CDT (0.900), respectively. The proposed methods

presented in this study would help in the development of accurate flash flood susceptible maps of watershed areas not only in Iran but also other parts of the world.

Keywords: machine learning; flash flood; GIS; Iran; decision trees; ensemble techniques

1. Introduction

Flash floods are those events where the rise in water is rapid within a few hours of the heavy rainfall. Flash flood is one of the most common, severely devastating natural hazards, which causes significant damages to the infrastructure and socioeconomy, and most importantly, it brings loss of lives [1–5]. Globally, more than 5000 people die each year due to flash flood events, which is about four times greater than any other category of flood event [6]. The most destructive nature of flood events is generally related to the extreme amount of torrential rainfall within a short duration resulting in high surface runoff [4,7]. Flash floods occur within catchments, where the response time of the drainage basin is short. According to the American Meteorological Society, flash flood events generally do not give advance warning and therefore, they cause significant risk and destruction due to their complex and dynamic environmental settings and nature [8,9].

Flash flood occurrence is affected by various watershed characteristics (type of basin and drainage), anthropogenic activities (land use, deforestation, and civil engineering construction) and meteorological conditions such as amount, intensity, spatial distribution, and time of rainfall. Recently, climate change is altering meteorological conditions which may lead to flash flood condition at one place and drought condition at another place. Therefore, the past may no longer be a reliable guide to the future. Thus, in the planning of flood management, especially of flash flood in urban areas, climate change effect is to be properly considered to avoid future damages to property and loss of life [10,11].

Geomorphological changes due to natural and anthropogenic causes can modify the flood pattern of different areas [12]. Urbanization is one of the important factors in the occurrence of flash floods in cities. Construction of roads and buildings reduces permeable areas and increases sealed surfaces (impermeable areas), thus causing less infiltration and more runoff with the same amount of rainfall causing pluvial flash floods [10]. Therefore, it is essential to identify and map accurately flash flood susceptible areas within a basin considering appropriate factors to develop suitable models for proper planning, management, and mitigation of flash flood events in an area [13].

There are many natural and anthropogenic factors that affect flood occurrence. Among these factors, topography is one of the important elements (land surface slope, river longitudinal profile, river cross section) that affects natural floods [14]. Flood parameters are very sensitive to topography changes. Low areas adjacent to rivers and streams have the highest risk of flooding. However, flash floods can also occur on hill slopes. Digital Elevation Model (DEM) as an indicator of the earth's surface contains information about the elevation of the earth. Flood depth and velocity are the most important parameters used in vulnerability assessment, estimation of casualties, and financial losses based on the land record [14]. Therefore, careful consideration of the topography of the area is desirable to avoid overestimation or underestimation of financial losses, casualties, and thus overall vulnerability assessment of an area [15,16].

Nowadays, multidisciplinary approaches including remote sensing, Geographic Information System (GIS), and machine learning methods are used for effective prediction and management of floods [5,6,12,17–19]. To recognize and delineate flash flood susceptible areas, DEM and other remote sensing satellite images have become popular and useful tools [20,21]. Bui and Hoang [22] reviewed the flash flood studies into three major classes, namely rainfall-runoff models, traditional methods, and pattern classification. In the case of rainfall-runoff models, the methodologies generally focus on establishing the relationship between the rainfall and runoff to determine the spatiotemporal distribution of the floods at a local scale and to carry out such studies in that area [23]. The traditional methods include analysis of long-term time series data and various statistical models [22]. The

problem of predicting flash flood probability by implementing the above methods is the lack of reliable data availability of the long-term time series discharge records. Another method based on the pattern classification is relatively new, which employs monitoring of data at the gauging stations and also preparation of data of flooded and nonflooded group to assess the flash flood probability of a region and to demarcate the area where flash floods can occur [24,25].

Independent simplified decision-making techniques such as Analytical Hierarchy Process (AHP) [5,26–29], Fuzzy Logic (FL) [30,31], and Frequency Ratio (FR) [32,33] are some of the pattern classification-based methods which have been used to generate the flash flood maps around the world. Though these methods are simple, they do not provide a great level of accuracy in flash flood prediction in comparison to modern and advanced machine learning methods such as Support Vector Machine (SVM) [34,35], Artificial Neural Network (ANN) [36–38], Logistic Regression (LR) [39], GARP and QUEST [40], and Random Forest (RF) [41]. In recent years, some hybrid and ensemble machine learning methods such as Hybrid Bayesian Framework [24], Logistic Model Tree with Bagging Ensembles [42], Ensemble Weight-of-Evidence and Support Vector Machines [43], and Neuro-Fuzzy system integrated with Meta-Heuristic Algorithms [44] have been developed which provide better accuracy in comparison to single machine learning methods.

The main objective of the present study is to use GIS Based Hybrid Computational Approaches to develop ensemble models for accurate flash flood susceptibility assessment. In view of this, four hybrid ensemble models for the flash flood prediction were developed with Credal Decision Tree (CDT) as base classifier. These developed ensemble models are: AdaBoostM1 based CDT (ABM-CDT); Bagging based CDT (Bag-CDT); Dagging based CDT (Dag-CDT); and MultiBoostAB based CDT (MBAB-CDT). A small watershed of Tafresh county in the Markazi province of Iran, which experiences many flash floods every year, was selected as a study area for collecting and generating the datasets for the modeling process. To validate and compare performance of the models, various methods such as statistical measures and Area Under the Receiver Operating Characteristic (AUC) curve were used.

2. Materials and Methods

Description of the Research Area

Watershed of Tafresh county is one of the flash flood-prone areas of Markazi province. This county is located in the Markazi province of Iran covering an area of 1605 km^2, between 34°31′ N and 35°5′ N, 49°30′ E to 50°9′ E (Figure 1). Topography of the Tafrash watershed area is hilly with elevation ranging from 1296 to 3101 m. This area experiences cold winters and relatively moderate summers. The average temperature is 19.2 °C in summer and 6.4 °C in winter. Average annual rainfall in this region is 254.3 mm. Major water supply sources in the Tafresh watershed include springs, the perennial GharehChay River, the Ab Kamar seasonal river, and semi-deep wells. The GharehChay river with discharge 3000 ls^{-1} is one of the most important rivers in the area, which provides water for irrigation in Tafresh area, but due to droughts in recent years, discharge has reduced below 2000 ls^{-1}. However, several severe flash floods occur in the Tafrash watershed during winter every year, due to sudden heavy rainfall within a short period.
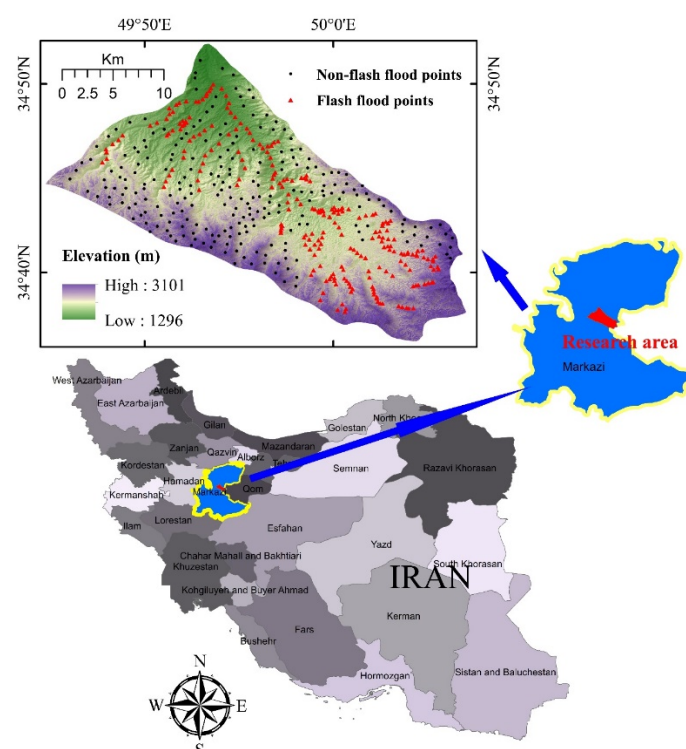
Figure 1. Location of study area.

3. Data Collection and Preparation

3.1. Flash Flood Inventory

Accurate mapping of the past flash flood events has a great impact on the accuracy of developed flash flood susceptibility maps. In order to predict the future flash flood events in a region, it is necessary to have records of the past flash flood events of the area [45]. These events depend on many factors including topography (terrain gradient), meteorology (antecedent rainfall), soil type, vegetative cover, and anthropogenic activities. These factors are considered as important parameters for the preparation of flash flood inventory and for the prediction of future flash flood events. In this research, in total, 320 past flash flood locations (represented on the maps by points) were obtained from the regional water organization of Markazi province (Figures 1 and 2). These flash flood points were divided randomly into 70% data points for training and 30% for validation purposes. In addition, 320 nonflooding points randomly selected from the high-altitude areas with low probability of flooding which were also used to combine with flash flood data for generating the training and testing datasets.


Figure 2. Flash flooding in Tafresh city.

3.2. Flash Flood Conditioning Factors

In this study, nine flash flood affecting parameters, namely distance from river, aspect, elevation, slope, rainfall, distance from faults, soil types, land use, and lithology were considered in the modeling. Thematic maps were generated using ArcGIS 10.1, ENVI 5.1, and SAGA-GIS 2 software (Figure 3). All these maps were converted to raster image (format) of 12.5 m × 12.5 m pixel size, which is up to the resolution of DEM for model studies (Table 1). A detailed description of these factors is given below:

Table 1. Data collection and preparation.

| Row	Primary Input Data	Original Format Sources	Spatial Resolution	Source of Data	Derived Map
1	ALOS-PALSER DEM	Raster	12.5 m	https://search.asf.alaska.edu/	Slope, Aspect, Curvature, Elevation, Distance from river
2	Landsat 8 OLI	Raster	30 m	Department of Natural Resources of Markazi Province	Land use map
3	Meteorological data	Point	-	Markazi County Meteorological Bureau	Rainfall map
4	Geological map	Vector	1:100000	Geological survey and Mineral Exploration of Iran	Lithology and Distance from fault
5	Soil map	Vector	1:100000	Department of Natural Resources of Markazi Province	Soil map
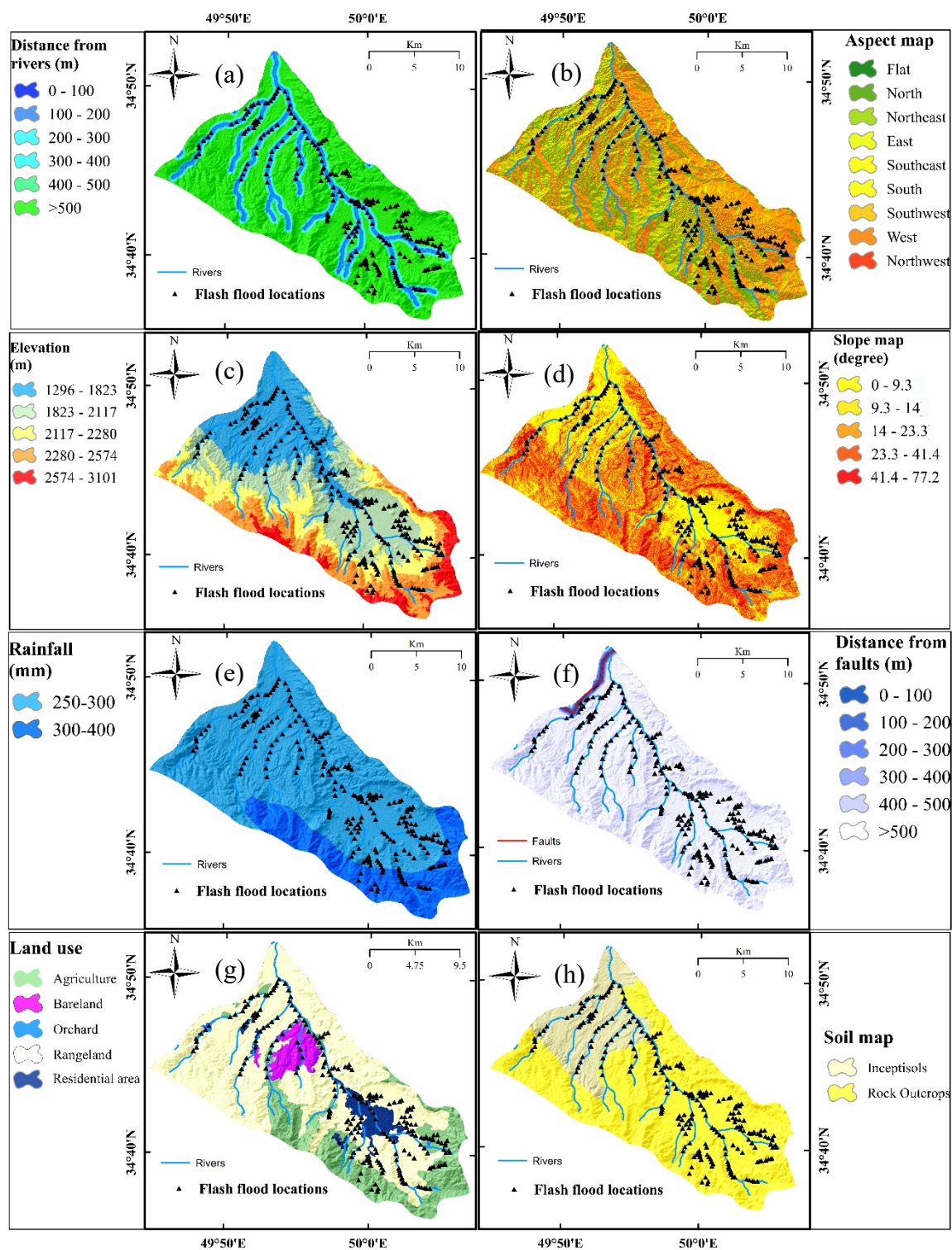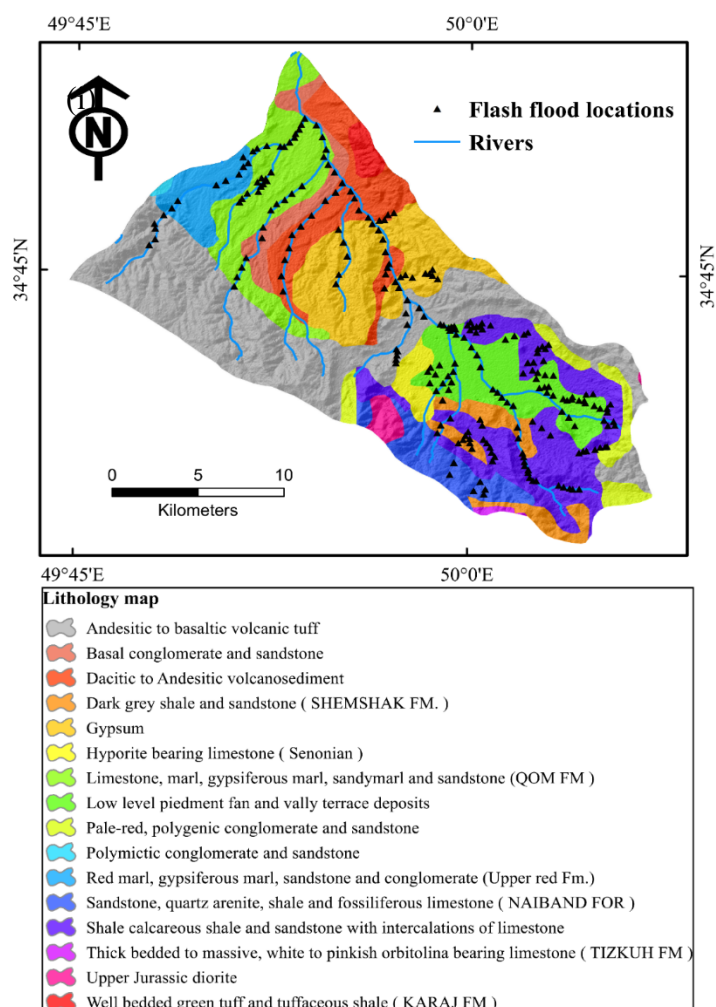
Figure 3. Maps of flash flood conditioning factors: (**a**) distance to rivers, (**b**) aspect, (**c**) elevation, (**d**) slope, (**e**) rainfall, (**f**) distance from faults, (**g**) land use, (**h**) soil, and (**i**) lithology**.**

Distance from rivers:

In general, the area which is close to the rivers is more prone to flooding in both cases of normal flood and flash flood within the river basin as water flows from higher elevation and accumulates at lower elevations. The areas close to other terrestrial water bodies such as ponds, dams, and lakes are also likely to be flooded in the event of heavy rains as the terrain in the vicinity of these water bodies would be almost flat [46]. However, pluvial flash floods may also occur at a distance away from the water bodies depending on the meteorological and topographical conditions. In the present study, six classes of buffer have been developed at buffer distance of 100 m from the river (Figure 3).

Aspect:

An aspect map of a region represents the direction of the surface slope. The direction which a slope faces with respect to the sun (aspect) has a profound influence on microclimate. The aspect map also shows no slope area (flat) where no surface slope is present; this is generally at the base of the hills or near lakes. Regions with low slope or regional flat surface are more vulnerable to the flash flood where water accumulates and rises [6,47]. Therefore, by using this parameter, the flat regions can easily be identified. Besides, flat area flooding also depends on the monsoon wind direction which hits the surface slope (Aspect). In this study, the aspect map was generated from the DEM with nine classes (Figure 3).

Elevation:

Water has a tendency of flowing from high altitude to lower elevation. The continuous flow of the rainwater therefore easily creates a flash flood situation in the low elevation areas [48,49].

However, pluvial flash floods also occur at higher elevation. In this study, an elevation map was generated from the DEM with five classes (Figure 3).

Slope:

Many factors affect catchment hydrologic characteristics, which ultimately influence the production of surface runoff. One of the important factors controlling runoff is surface slope [50]. On the steeper slopes, infiltration will be less and runoff will be more. This excessive runoff will cause flash flooding of the down slope flat areas. Thus, flat areas near and adjacent to high gradient slope generally have high probability of occurrence of flash floods [51]. In the present study, a slope map was created from DEM with five classes (Figure 3).

Rainfall:

Rainfall is the primary source of water for runoff generation over the land surface causing flooding of the low-lying areas. Runoff occurs whenever rain intensity exceeds the infiltration capacity of the ground (soil and jointed weathered rock). Intense short duration rainfall may cause flash floods. Rainfall is the most important factor for flooding of an area [50]. Flooding may also occur due to ice melting. In order to determine the annual rainfall map, the data of four rainfall-gauge stations for a period of 30 years were used. The rainfall map was divided into two classes (Figure 3).

Distance from faults:

Some of the major faults exposed on the surface with wide permeable fault zones may increase infiltration and thus reduce the runoff and can saturate surrounding groundmass causing local flooding. However, fault may also cause failure of levees and earthen dams due to structure failure and may result in flash flooding. In the present study, the distance from fault map was prepared into six classes (Figure 3).

Soil:

Soil is one of the important factors affecting infiltration and runoff and thus has a great impact on flooding. Soils rich in clay are mostly impermeable and cause more runoff and thus cause flooding of the area. In the present study, the soil map was developed from data obtained from the Soil Survey Department of Iran (Figure 3).

Land use:

Land use types affect the degree and frequency of floods in an area [52,53]. Infiltration and runoff depend on the land use pattern as well as other factors. Alterations in the land use configuration can change the flooding pattern of a region. Land-cover change due to anthropogenic activities such as urbanization, deforestation, and cultivation results in increased flash flood frequency and severity. In the present study, the land use map was obtained from the Department of Natural Resources of Markazi Province. Google Earth images and field survey were used to update the map (Figure 3).

Lithology:

Variation of lithology can strongly amplify or reduce the degree of flash flood vulnerability [54,55]. Infiltration and runoff depend on the permeability of lithounits as well as other geo-environmental factors. In this study, the lithology map was prepared from the Geological Survey of Iran data with sixteen groups (Table 2 and Figure 3).

Table 2. Lithology units in the Tafresh watershed and their relative permeability.

Group No	Geo-Units	Description	Permeability
1	Ea.bvt	Andesitic to basaltic volcanic tuff	Low
2	OMc	Basal conglomerate and sandstone	Moderate
3	Ed.avs	Dacitic to andesitic volcanosediment	Moderate
4	TRJs	Dark grey shale and sandstone (SHEMSHAK FM.)	Moderate
5	EKgy	Gypsum	High
6	K2l1	Hyporite bearing limestone (Senonian)	Moderate
7	OMq	Limestone, marl, gypsiferous marl. Sandymarl and sandstone (QOM FM)	Low
8	Qft2	Low level piedment fan and valley terrace deposit	High
9	Plc	Polymictic conglomerate and sandstone	Moderate
10	Mur	Red marl, gypsiferous marl, sandstone and conglomerate (upper red Fm.)	High
11	TRn	Sandstone, quartze arenite, shale and fossiliferous limestone (NAIBAND for)	Moderate
12	K2shm	Sale calcareous shale and sandstone with intercalations of limestone	Moderate

13	Ktzl	Thick bedded to massive, white to pinkish orbitolina bearing limestone (TIZKUh FM)	Moderate
14	Judi	Upper Jurassic diorite	Low
15	EK	Well bedded green tuff and tuffaceousshle (KARAJ FM)	Moderate

4. Methods Used

4.1. Frequency Ratio

Frequency Ratio (FR) determines the quantitative relationship between a flash flood event and its various variables [32,56]. In order to determine the FR, the ratio of flash flood events in each class of influencing factors is calculated relative to the total flash flood events. The ratio of the area of each class to the total area is also determined. Finally, by dividing the percentage of flash flood events in each class by the percentage of the area of each class relative to the entire research area, the FR of the classes of each factor is calculated. FR for each class of factors affecting the flash flood are calculated using the following equation [32,33]:

$$FR = \left(\frac{\frac{A}{B}}{\frac{C}{D}} \right) = \frac{E}{F} \tag{1}$$

where A: number of flash flood pixels per class, B: total flash flood pixels of the entire area, C: number of pixels per subclass of effective flash flood factors, D: total number of pixels in a region. E: percentage of flash flood occurrence in each class of effective factors, F: relative percentage of area of each class of total area.

4.2. Correlation Based Feature Selection

Irrelevant and redundant factors must be removed to improve data quality for modeling [57]. According to Pham et al. [58], working with a large number of factors reduces the speed of model execution, low modeling accuracy, and overfitting due to the large number of irrelevant factors as model inputs. There are many factors influencing the flood phenomenon, but the factors with higher correlation coefficients are more relevant in modeling and vice versa [58]. In this study, correlation based feature selection was selected to evaluate the importance of the factors used for better modeling of landslide susceptibility. This method is based on the assumption that features/factors are relevant if their values vary systematically with category membership [57,59]. In other words, a feature is useful if it is correlated with or predictive of the class; otherwise it is irrelevant [57,59]. In correlation based feature selection, the score of the evaluation is defined as Average Merit (AM) which is expressed as the following equation [57]:

$$AM_i = \frac{AC_i}{AI_i} \tag{2}$$

where AM_i is the score of factor *ith*, AC_i is the average correlation between the subset *i*[th] with the dependent variable, and AI_i is the average intercorrelation within the subset *i*[th].

4.3. AdaBoostM1

AdaBoostM1 is a popular adaptive boosting algorithm proposed by Freund and Schapire [60]. AdaBoostM1 enhanced the predictive ability of the classifier. This method is employed to solve the classification problem, which contains a complicated dataset generated from previous classifiers. Firstly, the weight values are allocated to occurrences in learning dataset. After that, the weights are substituted in iterations of training process according to the performance of the previous base classifier. The training process will be terminated when the optimal weights have been given specifically to achieve the best performance of the base classifier [61].

4.4. Bagging

Bagging is known as one of the earliest ensemble methods which was proposed by Breiman [62] to improve the algorithm accuracy of machine learning methods [63]. In this method, bootstrap sample technique is used to produce numerous samples for creating a training classifier. Each generated training set is then employed to establish a decision tree. After that, these subsets are combined with the output in the final model [61]. This method not only enhances the capacity of generalization but also decreases the error of classification [64,65]. The optimum result of classification can be drawn using the following equation:

$$L'(x) = \underset{y \in Y}{\arg\max} \sum_{i=1}^{t} 1\left(C_i(x) = y\right) \tag{3}$$

where L′ (x) expresses a combination of classifier and Ci(x) denotes an indicator function.

4.5. Dagging

Dagging was initially introduced by Ting and Witten [66]. This method is recognized as one of the famous ensemble techniques. Aim of Dagging method is to improve accuracy in prediction of the classifier by combining varied samples of the training set [67]. A number of disjointed samples are employed rather than bootstrap samples to achieve the base classifier [66,68]. This method is a powerful technique for a single classifier, which has a poor time of complexity; thus, the outputs of algorithms with weak training are linked via the popular voting rule [67].

4.6. MultiBoostAB

MultiBoostAB is a combination ensemble learning algorithm, which is established on the basis of AdaBoostM1 and Wagging methods in order to hinder overfitting problem [69]. Wagging is a variable of Bagging, which exploits training cases using various weights that can reduce remarkably the bias of AdaBoostM1 technique [70]. Combinations of Wagging and AdaBoostM1 produce a framework that can transform a weak training classifier to a robust one. As MultiBoostAB is able to perform parallel processing, it is considered as a potential and computational method that has more advantages in comparison to Wagging and AdaBoostM1 methods [69]. MultiBoostAB method involves three main steps: (1) selection of a subset randomly from the original learning data and then to use it to produce fundamental classifier-based models; (2) the weights of occurrence are adjusted based on the predictive competence of the models; and (3) finally, new subsets from the occurrence weighting are chosen for training newer models [71].

4.7. Credal Decision Tree

Abellan and Moral originally proposed Credal Decision Tree (CDT) using an original split criterion that was built based on uncertainty measures as well as inaccurate probabilities [70]. CDT is used to tackle classification problems by employing credal sets [64,72,73]. In order to reduce generating a complicated decision tree in the building process of CDT, an exclusive criterion was introduced in case of the summation of uncertainties raising due to splitting, the construction process will stop [64,74]. In order to quantitatively evaluate the entire uncertainty of credal sets, an updated method was recommended based on the theory of Dempster and Shafer [75,76]. The function applied in measuring the total uncertainty is expressed in the following equation [77]:

$$EU(\chi) = NG(\chi) + RG(\chi) \tag{4}$$

where EU expresses a value of entire uncertainty (i.e., total uncertainty), NG denotes a general nonspecificity function, and RG is a general randomness function for a credal set that represents a credal set. The successes and conclusions on the measurement of the summation of uncertainty were derived in previous literature of Abellan and Moral [78]. Besides, the detailed procedure for computing and properties of this measurement on EU were clearly described in previous studies [72,78]. To analyze

probability intervals of individual variables, the inaccurate probability model was adopted [79,80]. Supposing that the Z is known as a variable that has values which are expressed by zj; then p(zj) is considered as the probability distribution, which reflects that each value of zj is determined as per the following formula [73,81]:

$$p\left(z_j\right) \in \left[\frac{m_{zj}}{M+r}, \frac{m_{zj}+r}{M+r}\right], \mathrm{j} = 1,\dots, \mathrm{k};$$

(5)

where M and m_{zj} express the sample size and the event frequency (Z = z_j), respectively; and r is called the hyperparameter, which has values of 1 or 2, as stated by Walley [80].

4.8. Validation of the Models

Validation is important to determine the accuracy of the flash flood susceptibility models. To verify the prediction capability of models, it is desirable to assess as well as compare both learning and validating datasets [17,25,42]. In the present study, various validation criteria were adopted, namely Area Under the Receiver Operating Characteristic (ROC) curve and statistical measures.

4.8.1. Receiver Operating Characteristic (ROC) Curve

ROC curve is considered as a good tool for analyzing landslide and flood susceptibility models [17,42,82,83]. The *x*-axis of the ROC curve graph shows the specificity whereas the *y*-axis presents the sensitivity [84–88]. The area located under the ROC curve which is called the AUC is commonly employed to evaluate the prediction capacity of models [89–93]. Normally, the value of AUC has a range of 0.5–1.0 [94–96]. Higher value of AUC indicates better prediction capacity of the models [97–100]. The value of AUC is calculated by the following equation:

$$AUC = \frac{\left(\sum EC + \sum IC\right)}{\left(a+b\right)}$$

(6)

where EC indicates the number of the accurately classified flash flood events, IC denotes the number of the inaccurately classified flash flood events, a is single flash flood event, and b is denotes the total number of flash flood events.

4.8.2. Statistical Measures

In the present study, seven popular statistical measures, namely Positive Predictive Value (PPV), Negative Predictive Value (NPV), Root Mean Square Error (RSME), Accuracy (ACC), Sensitivity (SST), Specificity (SPF), and Kappa index (k) were employed for assessing performance of the flash flood prediction models. The description of these indexes is summarized in Table 3.

Table 3. List of statistical measures employed in this research [101–105].

Statistical Measures	Formula
PPV (%)	$PPV = \dfrac{A}{A + B}$
NPV (%)	$NPV = \dfrac{C}{C + D}$
ACC (%)	$ACC = \dfrac{A + C}{A + C + B + D}$
SST (%)	$SST = \dfrac{A}{A + D}$
SPF (%)	$SPF = \dfrac{C}{C + B}$
k	$k = \dfrac{P_a - P_{est}}{1 - P_{est}}$ $P_a = (A + C)$ $P_{est} = (A + D) \times (A + D) + (B + C) \times (D + C)$

where, A (true positive) and C (true negative) denotes the number of pixels of flash flood event classified correctly, whereas B (false positive) and D (false negative) are the numbers of pixels of nonflash flood event classified incorrectly. P_a and P_{est} are the measured and expected agreements, respectively.

RMSE is defined as the squared difference error between the model simulated and measured values. This method is popularly employed to assess flash flood susceptibility maps [17,42]. The smaller values of RMSE means the prediction capacity of the model is better. Determination of RMSE is calculated as follows [59,106–108]:

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{L} \left(X_{model} - X_{act} \right)^2} \qquad (7)$$

where X_{model} and X_{act} denote the model simulated and actual (i.e., measured) value, respectively; L stands for the summation of samples.

5. Methodology

Methodology of the study is presented below in several main steps: (1) Data collection and preparation; (2) Generating training and testing datasets; (3) Building the flash flood models; (4) Validation of the models; and (5) Generation of flash flood susceptibility maps (Figure 4). A more detailed description of these steps is given below:
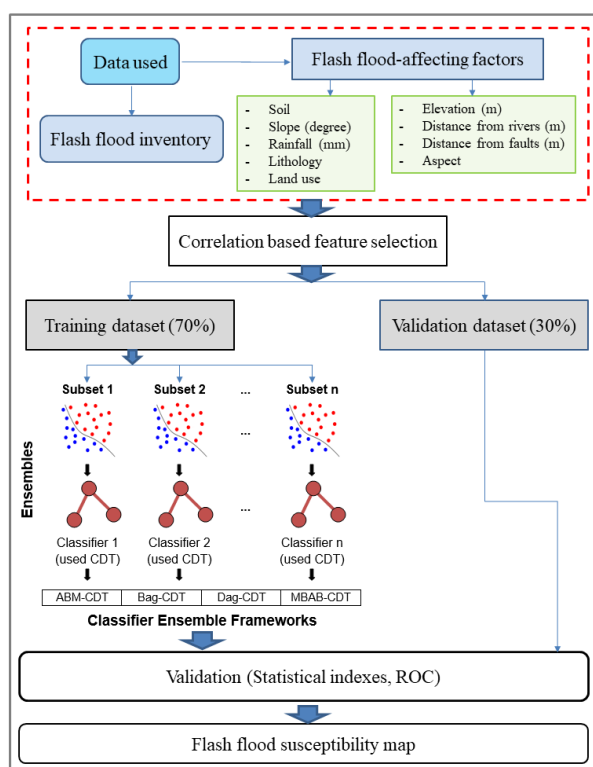
Figure 4. Methodological flowchart of the flash flood susceptibility mapping used in the study**.**

5.1. Data Collection and Preparation

Flash flood inventory map and the conditioning factor maps were generated in the raster format with 12.5 m pixel size. Thereafter, the inventory map was overlaid with the conditioning factor maps to calculate the FR values of each class of the conditioning factor using FR method. These FR values were then used as the weights of the class of the factors. In addition, correlation-based feature selection was used to validate and select the important factors and also to asses relative importance of these factors for modeling of flash floods.

5.2. Generating Training and Testing Datasets

Flash flood inventory was randomly divided into two parts with the ratio of 70/30. Out of these parts, 70% of inventory was used to sample with the conditioning factors assigned the weights for generating the training dataset, whereas the 30% remaining was used to sample with the conditioning factors assigned the weights for generating testing dataset. Selection of ratio for division of training and testing inventory might affect performance of the models. In this study, the ratio of 70/30 was used as it is a common ratio used in modeling [109–111]. This step was carried out in ArcGIS application.

5.3. Building the Flash Flood Models

Different hybrid models, namely ABM-CDT, Bag-CDT, Dag-CDT, MBAB-CDT, and a single classifier CDT were developed in this step using training dataset. Out of these methods, ABM-CDT is a combination of AdaBoostM1 ensemble and CDT classifier, Bag-CDT is a combination of Bagging ensemble and CDT, Dag-CDT is a combination of Dagging and CDT, and MBAB-CDT is a combination of MultiBoostAB and CDT. In these hybrid models, ensemble techniques were used to optimize the training dataset which was then used as input data in CDT classifier for flash flood susceptibility assessment. To construct these models, internal parameters should be selected and optimized to get the best performance of the models. More specifically, in CDT, initial parameters such as batch size, initial count, maximum of depth, minimum total weight of instances in a leaf, minimum proportion of variance, number of folds and seed were selected as 100, 0.0, −1, 2.0, 0.001, 3,

1, respectively. In ABM-CDT, initial parameters such as batch size, number of iterations, seed and weight of threshold were selected as 100, 10, 1, and 100, respectively. In Bag-CDT, initial parameters such as batch size, number of execution slots, number of iterations, and seed were selected as 100, 1, 15, and 1, respectively. In Dag-CDT, initial parameters such as batch size, number of folds, and seed were selected 100, 10, and 1, respectively. In MBAB-CDT, initial parameters such as batch size, number of iterations, number of subcommittees, seed, and weight of threshold were selected 100, 20, 3, 1, and 100, respectively. The values of these initial parameters of the models were determined by the trial-error process. This step was carried out using the packages and codes included in the Weka software.

5.4. Validation of the Models

Validation of the models was carried out on both training and testing datasets using various criteria such as PPV, NPV, SST, SPF, ACC, Kappa, RMSE, and AUC. While validation using training dataset shows the goodness-of-fit of the models, validation using testing datasets shows predictive capability of the models. This step was carried out using the packages and codes included in the Weka software.

5.5. Generation of Flash Flood Susceptibility Maps

In this step, flash flood susceptibility maps of Tafresh watershed were prepared based on ABM-CDT, Bag-CDT, Dag-CDT, MBAB-CDT hybrid machine learning models and CDT model in ArcGIS software. To construct the flash flood susceptibility maps, flash flood susceptibility indexes generated from the construction of the models were used to assign all pixels of the study area. Thereafter, these indexes were classified into five classes of flash flood susceptibility, namely very low, low, moderate, high, and very high to construct final maps using geometric interval classification method available in GIS software.

6. Results and Discussion

6.1. Impact Weight of each Class of Variables Affecting Flash Flood Susceptibility by FR Method

The impact weight of each class of variables was determined based on the comparative analyses of relationships between the location of past floods with the topographical and geo-environmental variables affecting flash flood occurrences (Figure 5). Analysis indicated that the highest weight in the variable of altitude classes belongs to the elevation class of 1296–1823 m. In the slope percentage of the surface slope, the weight of 0–9.3 degrees was the highest weight. In the slope direction variable, the northwest slope direction has a higher weight than the other aspects. In variable distance from the fault class of 400–500 m, weight has more influence than other classes. Examination of the variable distance from river showed that most of the flood-related weight was located at 0–100 m class. In the rainfall variable, the rainfall class 250–300 mm has higher weight than the other class. This means this class of rainfall belongs to threshold value for the occurrence of flash flood. Higher rainfall above this value can also cause flash flood depending on the duration in combination with other factors. Land use classes of the orchard and residential, which are in proximity to the main river and at gentle slopes, had the highest weighting factor compared to other land uses. Soil analysis indicates that the weight of the inceptisols soil is higher than that of the rocky outcrops. The lithology in this area indicates that Qom formation (OMq) has higher weight than other classes.
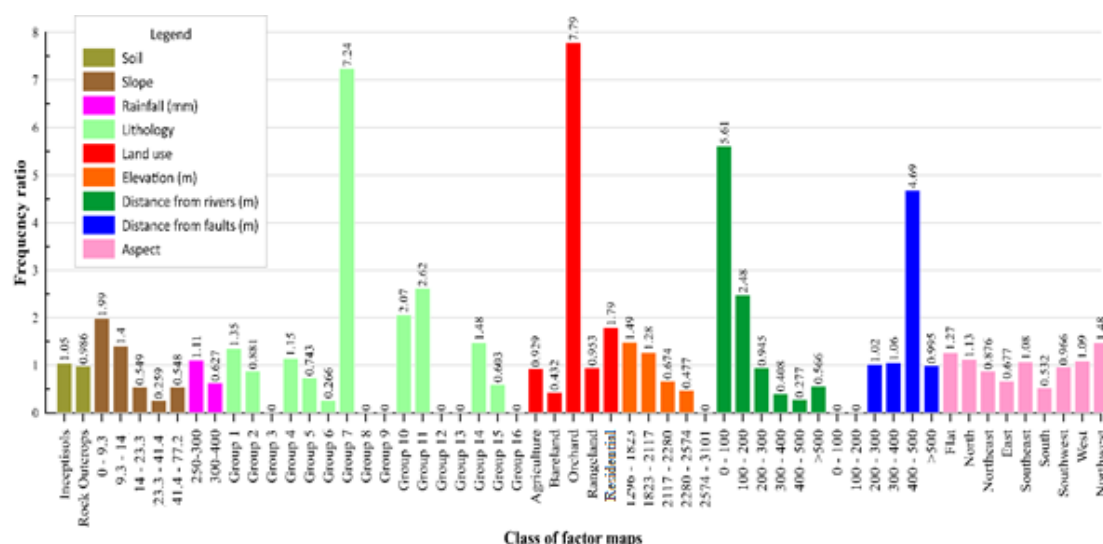
Figure 5. Frequency analysis of flash flood occurrence on the factor maps.

*6.2. Importance of Factors Using Correlation-Based Feature Selection*

Relative importance analysis of factors affecting flash floods was carried out using correlation-based feature selection method as shown in Table 4. It can be seen that distance from rivers is the most important factor for flash floods as the value of *AM* (0.608) is the highest compared with other factors. Following factors are slope (*AM* = 0.484), elevation (*AM* = 0.337), lithology (*AM* = 0.125), soil (*AM* = 0.099), rainfall (*AM* = 0.049), land use (*AM* = 0.024), aspect (*AM* = 0.022), and distance from faults (*AM* = 0.007), respectively. The feature selection results are reasonable as the areas close to the river are more likely to be affected by floods. This is true for normal river floods and also for flash floods in case of torrential rains within short period in this area [112,113]. Slope is also important as it influences surface runoff, volume, and velocity of flow. In the study area (Tafresh), there is more accumulation than outflow due to gentle topography (slope factor) resulting in the rapid rise of the flood water level within short time during torrential rain. Therefore, slope factor is the second most influential factor in the flood modeling (Table 4), which is consistent with many studies [114,115]. At higher elevation, slope factor is important resulting in higher velocity and runoff thus draining the water rapidly towards lower levels [116–118]. Other factors, namely lithology, soil, rainfall, land use, and aspect are also important factors for modeling of flash floods though their *AM* value varies as mentioned in Table 4. Here, we would like to mention that though *AM* of rainfall factor is only 0.049, it is the main and also triggering factor on which flash flood depends, especially in this area. However, the feature selection results show that distance to faults is the least important factor to flash flood occurrence and modeling (*AM* = 0.007), and thus this factor has a very small contribution to the performance of the models, and it should be removed from the datasets for further analysis of the models. Therefore, out of nine factors, only eight factors (distance from river, aspect, elevation, slope, rainfall, soil types, land use, and lithology) were reasonably selected for modeling of flash floods in this study.

Table 4. Importance of factors using correlation based feature selection.

Ranked	Class	Average Merit (*AM*)
1	Distance from rivers	0.608
2	Slope	0.484
3	Elevation	0.337
4	Lithology	0.125
5	Soil	0.099
6	Rainfall	0.049
7	Land use	0.024
8	Aspect	0.022
9	Distance from faults	0.007

6.3. Validation of Different Models

Performance of the machine learning models was validated using various criteria on both training and testing datasets (Figures 6–9). Validation of all the models was done by the ROC method (Figure 6). Results indicated very high AUC value during both training (ABM-CDT = 0.995; Bag-CDT = 0.972; Dag-CDT = 0.947; MBAB-CDT = 0.986; and CDT= 0.933) and testing phase (ABM-CDT = 0.96; Bag-CDT = 0.93; Dag-CDT = 0.47; MBAB-CDT = 0.933; and CDT = 0.90). Among all five models, ABM-CDT shows the maximum level of AUC compared with other models. All the models indicate a very low value of RMSE, both on the training dataset (ABM-CDT = 0.168; Bag-CDT = 0.245; Dag-CDT = 0.316; MBAB-CDT = 0.206; and CDT = 0.279) and testing dataset (ABM-CDT = 0.291; Bag-CDT = 0.307; Dag-CDT = 0.329; MBAB-CDT = 0.31; and CDT = 0.323) period, which clearly indicate high reliability of the proposed models (Figure 7). However, the ABM-CDT model indicates the best performance in comparison to other models, and it has the lowest RMSE value.
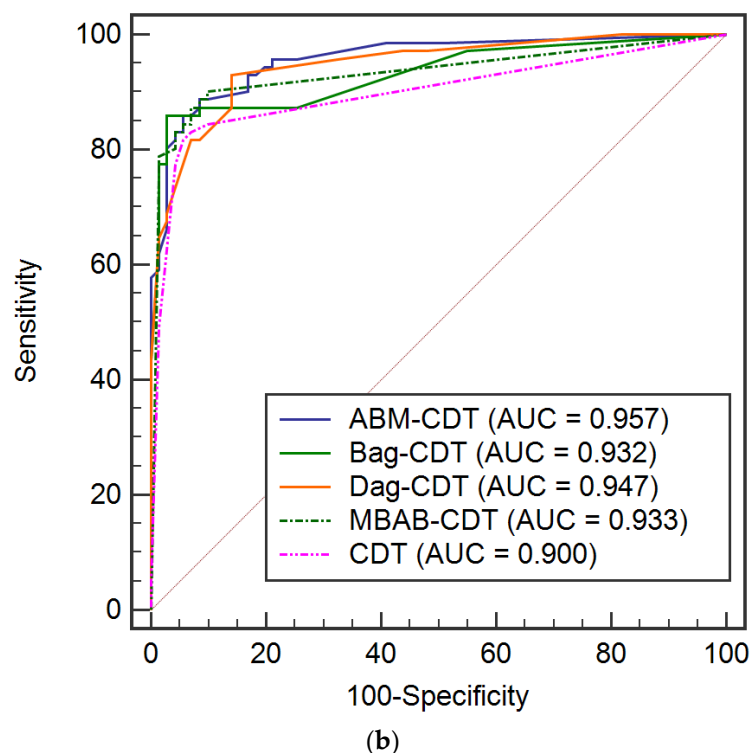

(**a**)

(**b**)
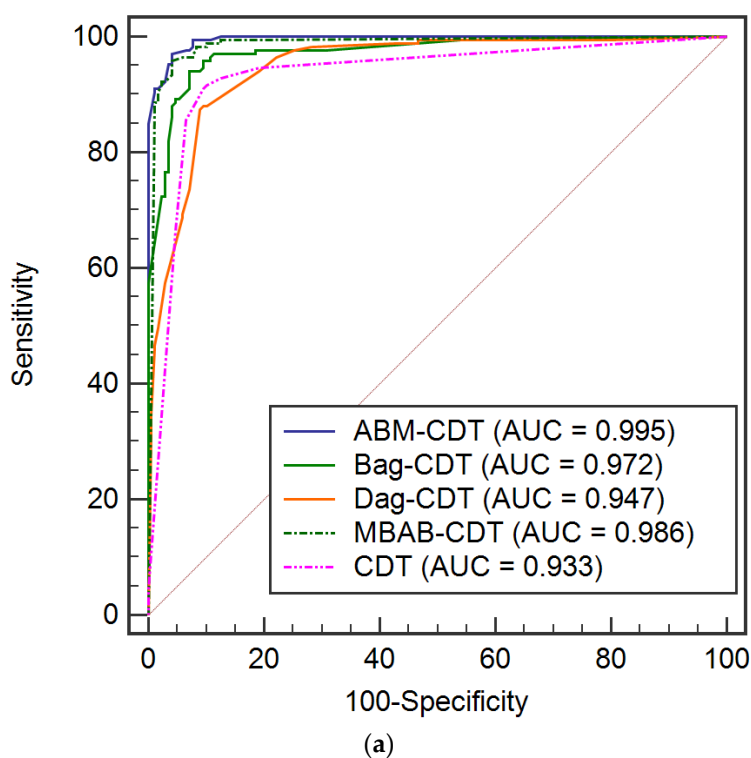
Figure 6. Analysis of Receiver Operating Characteristic (ROC) of the models: (**a**) training dataset and (**b**) validating dataset.
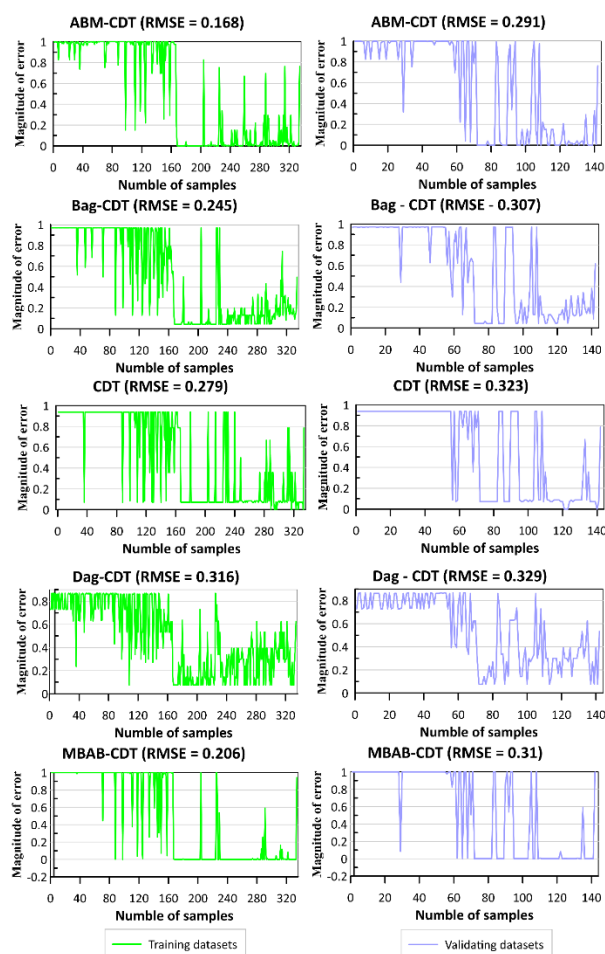

Figure 7. Analysis of RMSE of models.

Figure 8 indicates performance of the models using other validation criteria. It can be observed that all models have good performance with high values of PPV, NPV, SST, SPF, and ACC. Out of these, the ABM-CDT model has high values of PPV (95.81% for training and 94.37% for testing), NPV (96.41% for training and 85.92% for testing), SST (96.39% for training and 87.01% for testing), SPF (95.83% for training and 93.85% for testing), and ACC (96.11% for training and 90.14% for testing), the Bag-CDT model has values of PPV (88.62% for training and 94.37% for testing), NPV (97.01% for training and 85.92% for testing), SST (96.73% for training and 87.01% for testing), SPF (89.5% for training and 93.85% for testing), and ACC (92.81% for training and 90.14% for testing), the Dag-CDT model has values of PPV (89.82% for training and 91.55% for testing), NPV (88.02% for training and 81.69% for testing), SST (88.24% for training and 83.33% for testing), SPF (89.63% for training and 90.63% for testing), and ACC (88.92% for training and 86.62% for testing), the MBAB-CDT model has values of PPV (92.22% for training and 94.37% for testing), NPV (96.41% for training and 84.51% for testing), SST (96.25% for training and 85.9% for testing), SPF (92.53% for training and 93.75% for testing), and ACC (94.31% for training and 89.44% for testing) and the CDT model has values of PPV (90.42% for training and 94.37% for testing), NPV (91.02% for training and 81.69% for testing), SST (90.96% for training and 83.75% for testing), SPF (90.48% for training and 93.55% for testing), and ACC (90.72% for training and 88.03% for testing). Kappa statistics also show a satisfactory accuracy in both the case of training (ABM-CDT = 0.922; Bag-CDT = 0.856; Dag-CDT = 0.788; MBAB-CDT = 0.898; and CDT = 0.814) and testing (ABM-CDT = 0.803; Bag-CDT = 0.803; Dag-CDT = 0.732; MBAB-CDT = 0.789; and CDT = 0.761) (Figure 9).
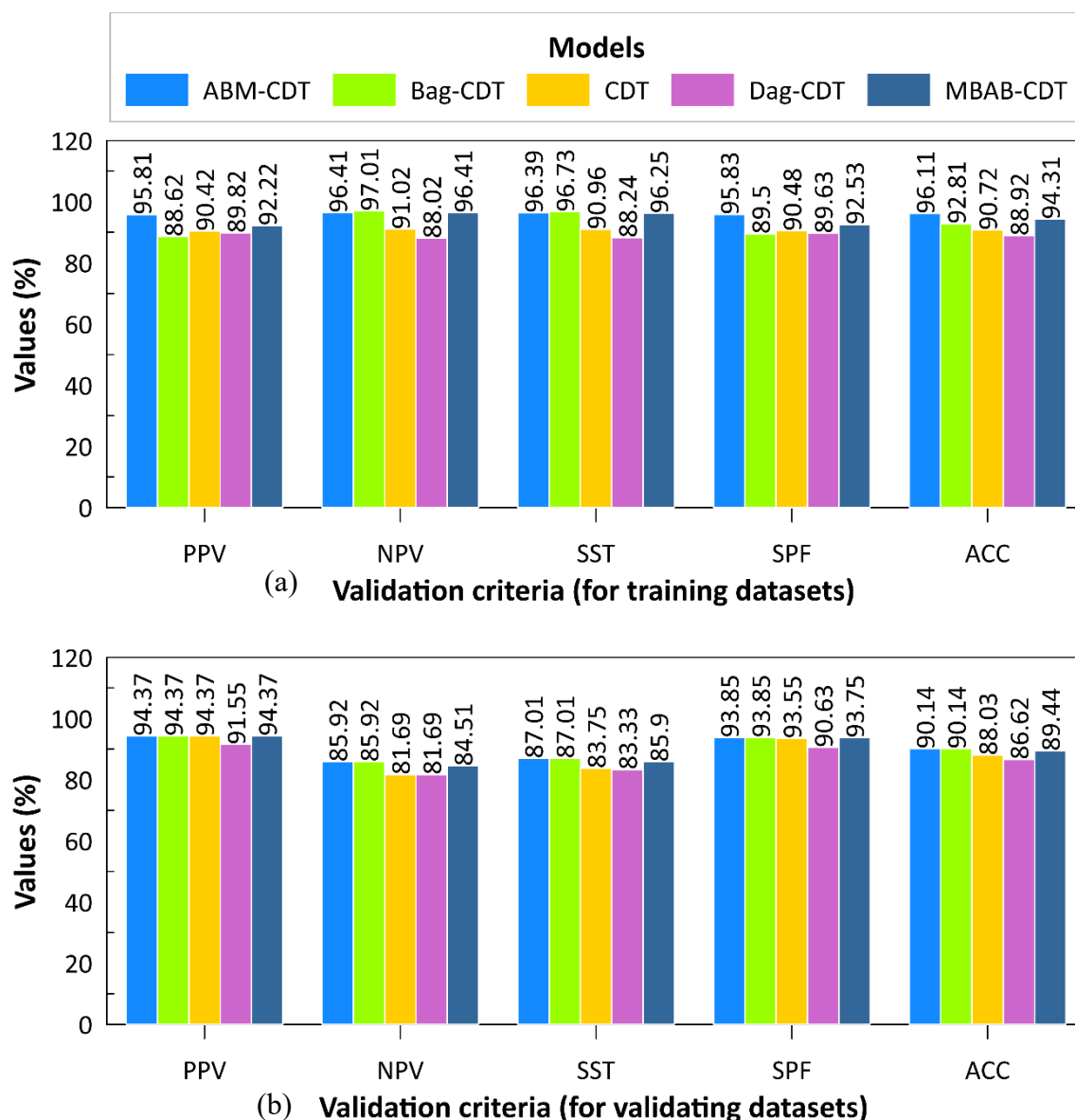
Figure 8. Analysis of accuracy of the models using: (**a**) training dataset and (**b**) validating dataset.
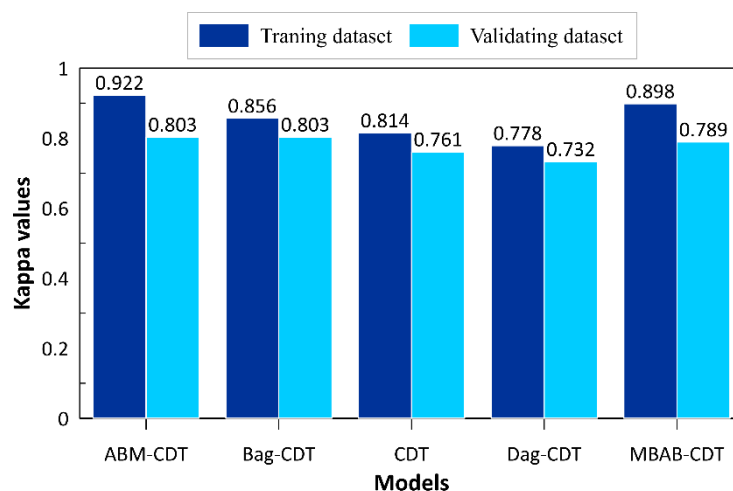

Figure 9. Kappa values for the models.

Considering analysis of the above results, it can be stated that all these developed and applied models performed well for flash flood susceptibility mapping in this study. In particular, the

prediction capability of the CDT model has been enhanced by more than 5% with AdaBoost, about 3% with Bagging and MultiBoostAB, and 5% with Dagging. In general, CDT algorithm is one of the good data mining models built on the decision tree and uses IDM and general uncertainty measures [69]. However, it has a low accuracy as the built tree decides to categorize a new sample of data, especially with incomplete or missing values of the data. Therefore, the use of ensemble frameworks like AdaBoostM1, Bagging, Dagging, and MultiboostAB is a great help in improving performance of the CDT as these techniques have the capability to condense the bias as well as the variance and avoid the problem of overfitting [119]. Comparison results of different ensemble frameworks used in this study (ABM-CDT, Bag-CDT, Dag-CDT, and MBAB-CDT) showed that ABM-CDT outperforms other ensemble frameworks (Bag-CDT, Dag-CDT, and MBAB-CDT). Thus, it can be stated that AdaBoostM1 is more effective than other ensemble techniques (Dagging, Bagging, and MultiBoostAB) in improving performance of the CDT for flash flood susceptibility assessment of this study. This result is reasonable as AdaBoostM1 can be considered to make a classification of the binary classes and enhance the prediction accuracy [120,121]. It is a very well-known fact that among all these ensembles, AdaBoostM1 is an interpretable and highly robust algorithm that prevents noise in order to make significant improvement in classifying error in comparison to the base decision tree classifier [122]. Our results are comparable to the previous ensemble model-based studies, which report that the ensemble models lead to a boost in the performance of a standalone model [123–125].

6.4. Development of Flash Flood Susceptibility Maps

Flash flood susceptibility maps of the research area were produced using ABM-CDT, Bag-CDT, Dag-CDT, MBAB-CDT, and CDT models (Figure 10). Figure 11 shows the comparison of results of all the models of flash flood susceptibility classes and their percentage of class pixels and flash flood pixels. All the models indicated that more than 50% of past flash floods were observed on very high susceptibility class of the maps (ABM-CDT = 51.3%; Bag-CDT = 53.8%; CDT = 61.8%; Dag-CDT = 69.7%; and MBAB-CDT = 86.1%). Evaluation of the frequency ratio data of the historical flash flood locations and the generated flash flood maps for the very high susceptible pixel class was done. The maximum FR was observed for ABM-CDT (3.46) followed by Bag-CDT (3.44); Dag-CDT (3.4); CDT (2.88), and MBAB-CDT (2.66), which clearly indicated higher degree of reliability of ABM-CDT and Bag-CDT algorithms.
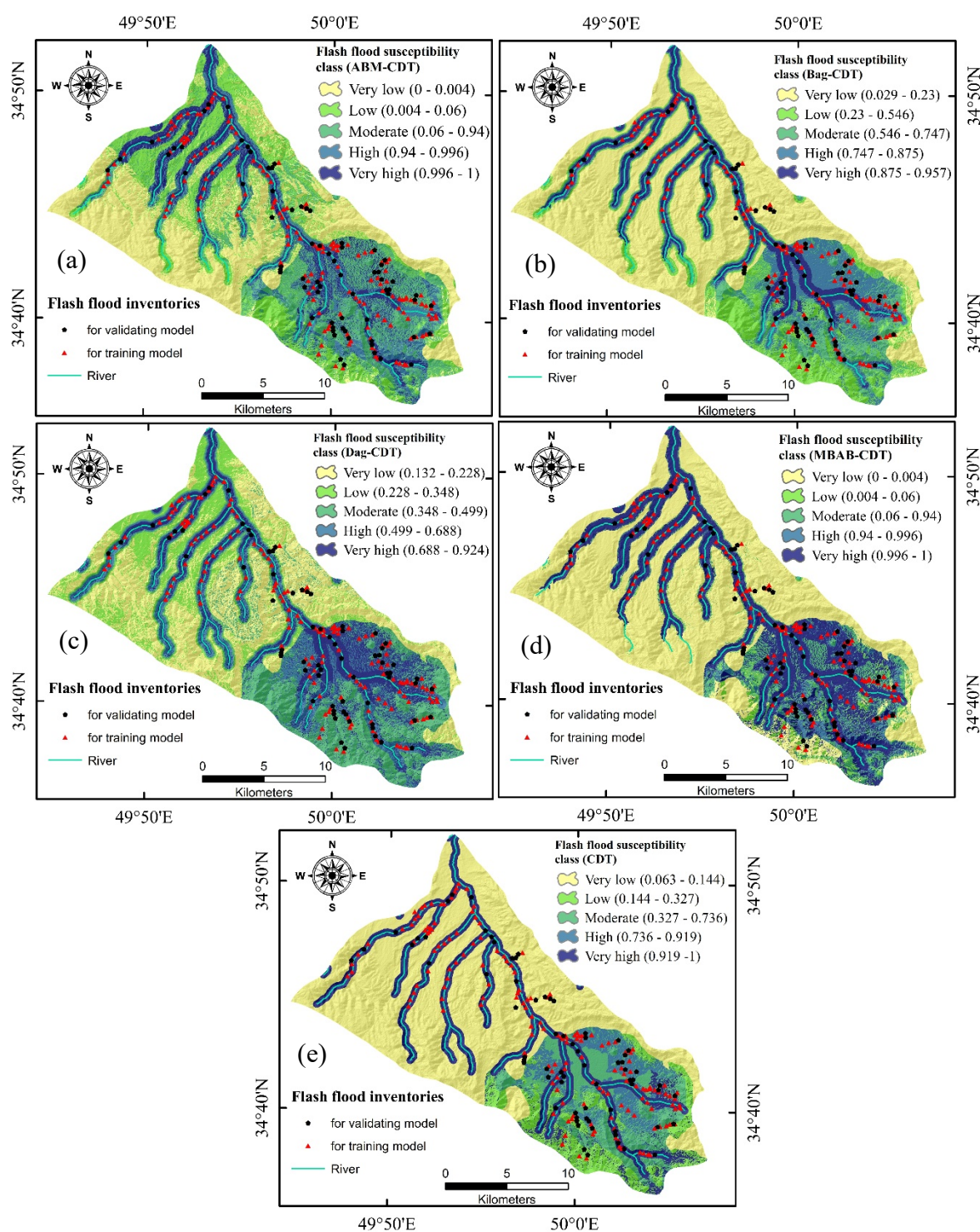
Figure 10. Flash flood susceptibility maps of the models: (**a**) ABM-CDT, (**b**) Bag-CDT, (**c**) Dag-CDT, (**d**) MBAB-CDT, and (**e**) CDT.
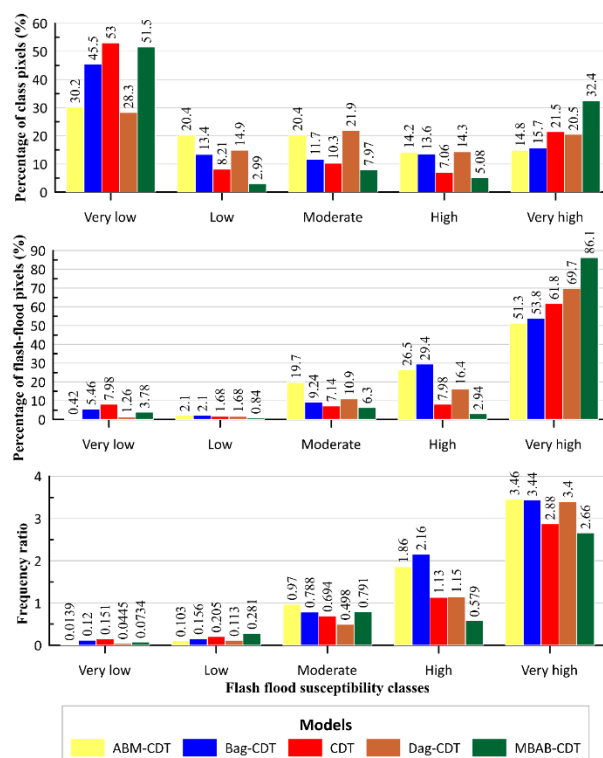
Figure 11. Analysis of performance of flash flood susceptibility maps using different models.

Analysis of the results of flash flood susceptibility maps shows that the Tafresh city area, which is located in the Tafresh watershed, belongs to very high susceptibility class. This is due to rapid development and expansion of the city area by encroaching topographically vulnerable areas to flash floods. Moreover, construction of buildings and roads in urban areas resulted in the increase of surface areas of impermeable structures and thus less infiltration and more runoff, causing flash floods in the event of intense rainfall during short periods [40,126,127]. The results of flood susceptibility zoning in Tafresh watershed showed that the southeastern parts have high to very high susceptibility to flash floods. The most important causes of flood susceptibility in these areas are related with anthropogenic activities causing drastic changes in catchment morphology, such as leveling of the land, altering the natural drainage, and increasing the impervious surfaces in the city. This has exacerbated the risk of floods and flooding of the infrastructure facilities thus increasing the potential threat to life and financial losses.

7. Concluding Remarks

In flash flood management studies, it is required to use accurate flash flood susceptibility maps by governing bodies and the policy makers for better flash flood mitigation and systematic development of the area. Since recent decades, a large number of methodologies have been developed to improve the accuracy of such maps. In this study, we proposed five new hybrid machine learning computational approaches to predict the possibility of flash flood occurrences in a studied catchment of Iran, where devastating flash flood events are frequent. The proposed methods are four hybrid models: ABM-CDT, Bag-CDT, Dag-CDT, MBAB-CDT, and single classifier: CDT. To construct the flash flood map, in total nine flash flood conditioning factors were taken into consideration to train and test the proposed models. Correlation based feature selection method was used to validate and select the important factors and also to asses relative importance of these factors for modeling of flash floods. Analysis shows that the lowest AM value (0.007) is of distance to fault and the highest AM value (0.608) is of distance to rivers. Distance to faults was then removed from the datasets for the flash flood modeling. Therefore, in the present study, we have considered only eight factors (distance from river, aspect, elevation, slope, rainfall, soil types, land use, and lithology) in the modeling.

The results show that performance of all the studied models in terms of accuracy was good as these models show very low RMSE values and a high percentage of AUC. Results indicate very high AUC value during both training phase (ABM-CDT = 0.995; Bag-CDT = 0.972; Dag-CDT = 0.958; MBAB-CDT = 0.983; and CDT = 0.933) and testing phase (ABM-CDT = 0.96; Bag-CDT = 0.93; Dag-CDT = 0.95; MBAB-CDT = 0.933; and CDT = 0.90). Among all five models, ABM-CDT shows the maximum level of accuracy compared with other models. Evaluation of the FR data of the historical flash flood locations and generated flash flood maps was done for the very high susceptible pixel class. The maximum frequency ratio was observed for ABM-CDT (3.46), followed by Bag-CDT (3.44); Dag-CDT (3.5); CDT (2.88), and MBAB-CDT (2.65) which clearly indicated higher degree of reliability of ABM-CDT and Bag-CDT algorithms. The models, as an outcome of the study, would also help in the development of accurate flash flood susceptible maps in other watersheds of Iran. However, in the model studies, physical link between cause and effect is to be maintained considering local geo-environmental and hydrological factors for better flash flood prediction and management.

In this study, we performed a systematic analysis using multisource geospatial data; a significant number of limitations still exist in this study about data configuration. We have used 12.5 m spatial resolution ALOS-PALSER DEM which is freely available; a higher resolution DEM can provide a more reliable flood map which may be more useful for the practical use of flood mitigation. In addition, feature selection method such as Information Gain should be applied to evaluate the importance of input factors used for better investigation and application of the machine learning models. Furthermore, despite employing robust methodologies, our study area is local in nature. Therefore, this study is required to be extended to other places for the evaluation of its practical application in different terrains and environments.

In this study we did not consider dynamic changes which may be induced by human activities in the form of land use changes, topography alteration, infrastructure development, as well as climate change. These changes may affect the natural hydrological cycle and thus the pattern of floods, in particular of flash flood in urban areas impacting the life and property of communities affected. Another limitation of the model study is the lack of dynamic consideration of changing parameters related with physical changes, flow levels, direction, erosion, sedimentation, blocking of the drainage system, etc. on flood simulation and its causative effect on land development and flood management.

However, there is a great scope for further research related with the assessment, prediction, and mapping of flash floods by applying other combinations of hybrid artificial intelligence models in different areas using high resolution geo-spatial data for better production of flash flood susceptibility maps.

References

1. Douben, K.-J. Characteristics of river floods and flooding: A global overview, 1985–2003. *Irrig. Drain. J. Int. Comm. Irrig. Drain.* **2006**, *55*, S9–S21.
2. Anagnostou, M.N.; Kalogiros, J.; Anagnostou, E.N.; Tarolli, M.; Papadopoulos, A.; Borga, M. Performance evaluation of high-resolution rainfall estimation by X-band dual-polarization radar for flash flood applications in mountainous basins. *J. Hydrol.* **2010**, *394*, 4–16.
3. Javelle, P.; Fouchier, C.; Arnaud, P.; Lavabre, J. Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations. *J. Hydrol.* **2010**, *394*, 267–274.

4. Modrick, T.M.; Georgakakos, K.P. The character and causes of flash flood occurrence changes in mountainous small basins of Southern California under projected climatic change. *J. Hydrol. Reg. Stud.* **2015**, *3*, 312–336.

5. Das, S. Geospatial mapping of flood susceptibility and hydro-geomorphic response to the floods in Ulhas Basin, India. *Remote Sens. Appl. Soc. Environ.* **2019**, *14*, 60–74.

6. Bui, D.T.; Tsangaratos, P.; Ngo, P.-T.T.; Pham, T.D.; Pham, B.T. Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods. *Sci. Total Environ.* **2019**, *668*, 1038–1054.

7. Georgakakos, K.P.; Hudlow, M.D. Quantitative precipitation forecast techniques for use in hydrologic forecasting. *Bull. Am. Meteorol. Soc.* **1984**, *65*, 1186–1200.

8. Georgakakos, K.P. On the design of national, real-time warning systems with capability for site-specific, flash-flood forecasts. *Bull. Am. Meteorol. Soc.* **1986**, *67*, 1233–1239.

9. Collier, C.G. Flash flood forecasting: What are the limits of predictability? *Q. J. R. Meteorol. Soc. J. Atmos. Sci. Appl. Meteorol. Phys. Oceanogr.* **2007**, *133*, 3–23.

10. Recanatesi, F.; Petroselli, A.; Ripa, M.N.; Leone, A. Assessment of stormwater runoff management practices and BMPs under soil sealing: A study case in a peri-urban watershed of the metropolitan area of Rome (Italy). *J. Environ. Manag.* **2017**, *201*, 6–18.

11. Szewrański, S.; Kazak, J.; Szkaradkiewicz, M.; Sasik, J. Flood risk factors in suburban area in the context of climate change adaptation policies—Case study of Wroclaw, Poland. *J. Ecol. Eng.* **2015**, *16*, 13–18.

12. Tien Bui, D.; Khosravi, K.; Shahabi, H.; Daggupati, P.; Adamowski, J.F.; Melesse, A.M.; Pham, B.T.; Pourghasemi, H.R.; Mahmoudi, M.; Bahrami, S.; et al. Flood spatial modeling in northern Iran using remote sensing and gis: A comparison between evidential belief functions and its ensemble with a multivariate logistic regression model. *Remote Sens.* **2019**, *11*, 1589.

13. Hammond, M.J.; Chen, A.S.; Djordjević, S.; Butler, D.; Mark, O. Urban flood impact assessment: A state-of-the-art review. *Urban Water J.* **2015**, *12*, 14–29.

14. Saksena, S.; Merwade, V. Incorporating the effect of DEM resolution and accuracy for improved flood inundation mapping. *J. Hydrol.* **2015**, *530*, 180–194.

15. Komolafe, A.A.; Herath, S.; Avtar, R. Sensitivity of flood damage estimation to spatial resolution. *J. Flood Risk Manag.* **2018**, *11*, 370–381, doi:10.1111/jfr3.12224.

16. Annis, A.; Nardi, F.; Morrison, R.R.; Castelli, F. Investigating hydrogeomorphic floodplain mapping performance with varying DTM resolution and stream order. *Hydrol. Sci. J.* **2019**, *64*, 525–538.

17. Khosravi, K.; Pham, B.T.; Chapi, K.; Shirzadi, A.; Shahabi, H.; Revhaug, I.; Prakash, I.; Bui, D.T. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz Watershed, Northern Iran. *Sci. Total Environ.* **2018**, *627*, 744–755.

18. Nikoo, M.; Ramezani, F.; Hadzima-Nyarko, M.; Nyarko, E.K.; Nikoo, M. Flood-routing modeling with neural network optimized by social-based algorithm. *Nat. Hazards* **2016**, *82*, 1–24.

19. Pradhan, B.; Shafiee, M.; Pirasteh, S. Maximum flash flood prone area mapping using RADARSAT images and GIS: Kelantan river basin. *Int. J. Geoinform.* **2009**, *5*, 11–23.

20. Noman, N.S.; Nelson, E.J.; Zundel, A.K. Review of automated floodplain delineation from digital terrain models. *J. Water Resour. Plan. Manag.* **2001**, *127*, 394–402.

21. Papaioannou, G.; Vasiliades, L.; Loukas, A. Multi-criteria analysis framework for potential flash flood prone areas mapping. *Water Resour. Manag.* **2015**, *29*, 399–418.

22. Bui, D.T.; Hoang, N.-D. 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.* **2017**, *10*, 3391.

23. Brunner, G.W. *HEC-RAS River Analysis System. Hydraulic Reference Manual, Version 1.0*; Hydrologic Engineering Center: Davis, CA, USA, 1995.

24. Bui, D.T.; Ngo, P.-T.T.; Pham, T.D.; Jaafari, A.; Minh, N.Q.; Hoa, P.V.; Samui, P. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. *Catena* **2019**, *179*, 184–196.

25. Bui, D.T.; Pradhan, B.; Nampak, H.; Bui, Q.-T.; Tran, Q.-A.; Nguyen, Q.-P. Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flash flood susceptibilitgy modeling in a high-frequency tropical cyclone area using GIS. *J. Hydrol.* **2016**, *540*, 317–330.

26.	Chen, Y.-R.; Yeh, C.-H.; Yu, B. Integrated application of the analytic hierarchy process and the geographic information system for flash flood risk assessment and flash flood plain management in Taiwan. *Nat. Hazards* **2011**, *59*, 1261–1276.

27.	Das, S. Geographic information system and AHP-based flood hazard zonation of Vaitarna Basin, Maharashtra, India. *Arab. J. Geosci.* **2018**, *11*, 576.

28.	Radwan, F.; Alazba, A.A.; Mossad, A. Flash flood risk assessment and mapping using AHP in arid and semiarid regions. *Acta Geophys.* **2019**, *67*, 215–229.

29.	Souissi, D.; Zouhri, L.; Hammami, S.; Msaddek, M.H.; Zghibi, A.; Dlala, M. GIS-based MCDM-AHP modeling for flash flood susceptibility mapping of arid areas, Southeastern Tunisia. *Geocarto Int.* **2019**, No. just-accepted, 1–25.

30.	Pierdicca, N.; Pulvirenti, L.; Chini, M.; Guerriero, L.; Ferrazzoli, P. A fuzzy-logic-based approach for flash flood detection from cosmo-skymed data. In Proceedings of the IEEE International Geoscience & Remote Sensing Symposium, IGARSS 2010, Honolulu, HI, USA, 25–30 July 2010; pp. 4796–4798.

31.	Zou, Q.; Zhou, J.; Zhou, C.; Song, L.; Guo, J. Comprehensive flash flood risk assessment based on set pair analysis-variable fuzzy sets model and fuzzy AHP. *Stoch. Environ. Res. Risk Assess.* **2013**, *27*, 525–546.

32.	Lee, M.-J.; Kang, J.; Jeon, S. Application of frequency ratio model and validation for predictive flooded area susceptibility mapping using GIS. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 895–898.

33.	Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Flash flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 1149–1165.

34.	Yan, J.; Jin, J.; Chen, F.; Yu, G.; Yin, H.; Wang, W. Urban flash flood forecast using support vector machine and numerical simulation. *J. Hydro.* **2018**, *20*, 221-231.

35.	Tehrany, M.S.; Pradhan, B.; Mansor, S.; Ahmad, N. Flash flood susceptibility assessment using GIS-based support vector machine model with different kernel types. *Catena* **2015**, *125*, 91–101.

36.	Sahoo, G.B.; Ray, C.; De Carlo, E.H. Use of neural network to predict flash flood and attendant water qualities of a mountainous stream on Oahu, Hawaii. *J. Hydrol.* **2006**, *327*, 525–538.

37.	Youssef, A.M.; Pradhan, B.; Hassan, A.M. Flash flash flood risk estimation along the St. Katherine Road, Southern Sinai, Egypt using GIS based morphometry and satellite imagery. *Environ. Earth Sci.* **2011**, *62*, 611–623, doi:10.1007/s12665-010-0551-1.

38.	Kia, M.B.; Pirasteh, S.; Pradhan, B.; Mahmud, A.R.; Sulaiman, W.N.A.; Moradi, A. An artificial neural network model for flash flood simulation using GIS: Johor River Basin, Malaysia. *Environ. Earth Sci.* **2012**, *67*, 251–264.

39.	Nandi, A.; Mandal, A.; Wilson, M.; Smith, D. Flash flood hazard mapping in Jamaica using principal component analysis and logistic regression. *Environ. Earth Sci.* **2016**, *75*, 465.

40.	Darabi, H.; Choubin, B.; Rahmati, O.; Torabi Haghighi, A.; Pradhan, B.; Kløve, B. Urban flash flood risk mapping using the GARP and QUEST models: A comparative study of machine learning techniques. *J. Hydrol.* **2019**, *569*, 142–154, doi:10.1016/j.jhydrol.2018.12.002.

41.	Lee, S.; Kim, J.-C.; Jung, H.-S.; Lee, M.J.; Lee, S. Spatial prediction of flash flood susceptibility using random-forest and boosted-tree models in Seoul Metropolitan City, Korea. *Geomat. Nat. Hazards Risk* **2017**, *8*, 1185–1203.

42.	Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Bui, D.T.; Pham, B.T.; Khosravi, K. A novel hybrid artificial intelligence approach for flash flood susceptibility assessment. *Environ. Model. Softw.* **2017**, *95*, 229–245.

43.	Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Flash flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J. Hydrol.* **2014**, *512*, 332–343, doi:10.1016/j.jhydrol.2014.03.008.

44.	Bui, D.T.; Panahi, M.; Shahabi, H.; Singh, V.P.; Shirzadi, A.; Chapi, K.; Khosravi, K.; Chen, W.; Panahi, S.; Li, S.; et al. Novel hybrid evolutionary algorithms for spatial prediction of floods. *Sci. Rep.* **2018**, *8*, 15364.

45.	Choubin, B.; Moradi, E.; Golshan, M.; Adamowski, J. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Sci. Total Environ.* **2019**, *651*, 2087–2096, doi:10.1016/j.scitotenv.2018.10.064.

46.	Reager, J.T.; Thomas, B.F.; Famiglietti, J.S. River basin flash flood potential inferred using grace gravity observations at several months lead time. *Nat. Geosci.* **2014**, *7*, 588.

47. Hoang, L.P.; Biesbroek, R.; Tri, V.P.D.; Kummu, M.; Van Vliet, M.T.H.; Leemans, R.; Kabat, P.; Ludwig, F. Managing flash flood risks in the mekong delta: How to address emerging challenges under climate change and socioeconomic developments. *Ambio* **2018**, *47*, 635–649, doi:10.1007/s13280-017-1009-4.

48. Fernández, D.S.; Lutz, M.A. Urban flash flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis. *Eng. Geol.* **2010**, *111*, 90–98.

49. Dahri, N.; Abida, H. Monte carlo simulation-aided Analytical Hierarchy Process (AHP) for flash flood susceptibility mapping in Gabes Basin (Southeastern Tunisia). *Environ. Earth Sci.* **2017**, *76*, 302.

50. Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Spatial prediction of flash flood susceptible areas using rule based Decision Tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *J. Hydrol.* **2013**, *504*, 69–79.

51. Li, K.; Wu, S.; Dai, E.; Xu, Z. Flash flood loss analysis and quantitative risk assessment in China. *Nat. Hazards* **2012**, *63*, 737–760.

52. Garcia-Ruiz, J.M.; Regüés, D.; Alvera, B.; Lana-Renault, N.; Serrano-Muela, P.; Nadal-Romero, E.; Navas, A.; Latron, J.; Marti-Bono, C.; Arnáez, J. Flash flood generation and sediment transport in experimental catchments affected by land use changes in the central pyrenees. *J. Hydrol.* **2008**, *356*, 245–260.

53. Benito, G.; Rico, M.; Sánchez-Moya, Y.; Sopeña, A.; Thorndycraft, V.R.; Barriendos, M. The impact of late holocene climatic variability and land use change on the flash flood hydrology of the Guadalentin River, Southeast Spain. *Glob. Planet. Chang.* **2010**, *70*, 53–63.

54. Xu, Y.; Chung, S.-L.; Jahn, B.; Wu, G. Petrologic and geochemical constraints on the petrogenesis of Permian-Triassic Emeishan flash flood basalts in Southwestern China. *Lithos* **2001**, *58*, 145–168.

55. Kazakis, N.; Kougias, I.; Patsialis, T. Assessment of flash flood hazard areas at a regional scale using an index-based approach and analytical hierarchy process: Application in Rhodope-Evros Region, Greece. *Sci. Total Environ.* **2015**, *538*, 555–563.

56. Rahmati, O.; Pourghasemi, H.R.; Zeinivand, H. Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. *Geocarto Int.* **2016**, *31*, 42–70.

57. Hall, M.A. Correlation-based feature selection of discrete and numeric class machine learning. In Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000), Stanford, CA, USA, 29 June–2 July 2000.

58. Pham, B.T.; Pradhan, B.; Bui, D.T; Prakash, I.; Dholakia, M.B. A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India). *Environ. Model. Softw.* **2016**, *84*, 240–250.

59. Duma, M.; Twala, B.; Nelwamondo, F.V.; Marwala, T. Partial imputation to improve predictive modelling in insurance risk classification using a hybrid positive selection algorithm and correlation-based feature selection. *Curr. Sci.* **2012**, *103*, 697–705.

60. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139.

61. Pham, B.T.; Bui, D.T.; Prakash, I.; Dholakia, M.B. Hybrid integration of multilayer perceptron neural networks and machine learning ensembles for landslide susceptibility assessment at Himalayan Area (India) using GIS. *Catena* **2017**, *149*, 52–63.

62. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140.

63. Piao, Y.; Piao, M.; Jin, C.H.; Shon, H.S.; Chung, J.-M.; Hwang, B.; Ryu, K.H. A new ensemble method with feature space partitioning for high-dimensional data classification. *Math. Probl. Eng.* **2015**, *2015*, 1–12.

64. He, Q.; Xu, Z.; Li, S.; Li, R.; Zhang, S.; Wang, N.; Pham, B.T.; Chen, W. Novel entropy and rotation forest-based credal decision tree classifier for landslide susceptibility modeling. *Entropy* **2019**, *21*, 106.

65. Khosravi, K.; Cooper, J.R.; Daggupati, P.; Pham, B.T.; Bui, D.T. Bedload transport rate prediction: Application of novel hybrid data mining techniques. *J. Hydrology* **2020**, 124774.

66. Ting, K.M.; Witten, I.H. Stacking bagged and dagged models. In Proceedings of the 14th International Conference on Machine Learning, San Francisco, CA, USA, July 1997.

67. Onan, A.; Korukouglu, S.; Bulut, H. Ensemble of keyword extraction methods and classifiers in text classification. *Expert Syst. Appl.* **2016**, *57*, 232–247.

68. Thai, B.; Dieu, P.; Bui, T.; Prakash, I. Landslide susceptibility assessment using bagging ensemble based alternating decision trees, logistic regression and J48 decision trees methods: A comparative study. *Geotech. Geol. Eng.* **2017**, *35*, 2597–2611, doi:10.1007/s10706-017-0264-2.

69. Webb, G.I. Multiboosting: A technique for combining boosting and wagging. *Mach. Learn.* **2000**, *40*, 159–196.

70. Kotti, M.; Benetos, E.; Kotropoulos, C.; Pitas, I. A neural network approach to audio-assisted movie dialogue detection. *Neurocomputing* **2007**, *71*, 157–166.

71. Bui, D.T.; Ho, T.-C.; Pradhan, B.; Pham, B.-T.; Nhu, V.-H.; Revhaug, I. GIS-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with adaboost, bagging, and multiboost ensemble frameworks. *Environ. Earth Sci.* **2016**, *75*, 1101.

72. Abellán, J.; Moral, S. Building classification trees using the total uncertainty criterion. *Int. J. Intell. Syst.* **2003**, *18*, 1215–1225

73. Mantas, C.J.; Abellán, J. Credal-C4.5: Decision tree based on imprecise probabilities to classify noisy data. *Expert Syst. Appl.* **2014**, *41*, 4625–4637, doi:10.1016/j.eswa.2014.01.017.

74. Abellán, J.; Masegosa, A.R. Combining decision trees based on imprecise probabilities and uncertainty measures. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*; Springer, Berlin, Heidelberg, 2007; pp. 512–523..

75. Dempster, A.P. Upper and lower probabilities induced by a multivalued mapping. In *Classic Works of the Dempster-Shafer Theory of Belief Functions*; Springer: New York, NY, USA, 2008; pp. 57–72.

76. Shafer, G. *A Mathematical Theory of Evidence*; Princeton University Press: Princeton, NJ, USA, 1976; p. 42.

77. Abellan, J.; Moral, S. Completing a total uncertainty measure in the dempster-shafer theory. *Int. J. Gen. Syst.* **1999**, *28*, 299–314.

78. Abellan, J.; Moral, S. A non-specificity measure for convex sets of probability distributions. *Int. J. Uncertain. Fuzziness Knowl. Based Syst.* **2000**, *8*, 357–367.

79. Mantas, C.J.; Abellán, J.; Castellano, J.G. Analysis of credal-C4. 5 for classification in noisy domains. *Expert Syst. Appl.* **2016**, *61*, 314–326.

80. Walley, P. Inferences from multinomial data: Learning about a bag of marbles. *J. R. Stat. Soc. Ser. B* **1996**, *58*, 3–34.

81. Mantas, C.J.; Abellán, J. Analysis and extension of decision trees based on imprecise probabilities: Application on noisy data. *Expert Syst. Appl.* **2014**, *41*, 2514–2525.

82. Hong, H.; Panahi, M.; Shirzadi, A.; Ma, T.; Liu, J.; Zhu, A.-X.; Chen, W.; Kougias, I.; Kazakis, N. Flash flood susceptibility assessment in hengfeng area coupling adaptive neuro-fuzzy inference system with genetic algorithm and differential evolution. *Sci. Total Environ.* **2018**, *621*, 1124–1141.

83. Pham, B.T.; Bui, D.T.; Dholakia, M.B.; Prakash, I.; Pham, H.V. A comparative study of least square support vector machines and multiclass alternating decision trees for spatial prediction of rainfall-induced landslides in a tropical cyclones area. *Geotech. Geol. Eng.* **2016**, *34*, 1807–1824, doi:10.1007/s10706-016-9990-0.

84. Ayalew, L.; Yamagishi, H.; Ugawa, N. Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa Area of Agano River, Niigata Prefecture, Japan. *Landslides* **2004**, *1*, 73–81.

85. Van Dao, D.; Jaafari, A.; Bayat, M.; Mafi-Gholami, D.; Qi, C.; Moayedi, H.; Van Phong, T.; Ly, H.-B.; Le, T.-T.; Trinh, P.T. A spatially explicit deep learning neural network model for the prediction of landslide susceptibility. *Catena* **2020**, *188*, 104451.

86. Termeh, S.V.R.; Khosravi, K.; Sartaj, M.; Keesstra, S.D.; Tsai, F.T.C.; Dijksma, R.; Pham, B.T. Optimization of an adaptive neuro-fuzzy inference system for groundwater potential mapping. *Hydrogeol. J.* **2019**, *27*, 2511–2534.

87. Pham, B.T.; Jaafari, A.; Prakash, I.; Singh, S.K.; Quoc, N.K.; Bui, D.T. Hybrid computational intelligence models for groundwater potential mapping. *Catena* **2019**, *182*, 104101.

88. Tien Bui, D.; Shirzadi, A.; Chapi, K.; Shahabi, H.; Pradhan, B.; Pham, B.T.; Singh, V.P.; Chen, W.; Khosravi, K.; Ahmad, B.B.; et al. A hybrid computational intelligence approach to groundwater spring potential mapping. *Water* **2019**, *11*, 2013.

89. Phong, T.V.; Phan, T.T.; Prakash, I.; Singh, S.K.; Shirzadi, A.; Chapi, K.; Ly, H.B.; Ho, L.S.; Quoc, N.K.; Pham, B.T. Landslide susceptibility modeling using different artificial intelligence methods: A case study at Muong Lay district, Vietnam. *Geocarto Int.* **2019**, doi:10.1080/10106049.2019.1665715.

90. Tien Bui, D.; Shirzadi, A.; Shahabi, H.; Geertsema, M.; Omidvar, E.; Clague, J.J.; Thai Pham, B.; Dou, J.; Talebpoor, D.; Lee, S.; et al. New ensemble models for shallow landslide susceptibility modeling in a semi-arid watershed. *Forests* **2019**, *10*, 743.

91. Pham, B.T.; Prakash, I.; Singh, S.K.; Shirzadi, A.; Shahabi, H.; Bui, D.T. Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learning approaches. *Catena* **2019**, *175*, 203–218.

92. Pham, B.T.; Bui, D.T.; Pham, H.V.; Le, H.Q.; Prakash, I.; Dholakia, M.B. Landslide hazard assessment using random subspace fuzzy rules based classifier ensemble and probability analysis of rainfall data: A case study at Mu Cang Chai District, Yen Bai Province (Viet Nam). *J. Indian Soc. Remote Sens.* **2017**, *45*, 673–683.

93. Jaafari, A.; Zenner, E.K.; Pham, B.T. Wildfire spatial pattern analysis in the Zagros Mountains, Iran: A comparative study of decision tree based classifiers. *Ecol. Inform.* **2018**, *43*, 200–211.

94. Khosravi, K.; Shahabi, H.; Pham, B.T.; Adamowski, J.; Shirzadi, A.; Pradhan, B.; Dou, J.; Ly, H.; Grof, G.; Ho, H.L.; et al. A comparative assessment of flood susceptibility modeling using multi-criteria decision making analysis and machine learning methods. *J. Hydrol.* **2019**, *573*, 311–323.

95. Khosravi, K.; Sartaj, M.; Tsai, F.T.; Singh, V.P.; Kazakis, N.; Melesse, A.M.; Prakash, I.; Bui, D.T.; Pham, B.T. A comparison study of drastic methods with various objective methods for groundwater vulnerability assessment. *Sci. Total Environ.* **2018**, *642*, 1032–1049.

96. Miraki, S.; Zanganeh, S.H.; Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Pham, B.T. Mapping groundwater potential using a novel hybrid intelligence approach. *Water Resour. Manag.* **2019**, *33*, 281–302.

97. Dou, J.; Yunus, A.P.;Bui, D.T.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.; Khosravi, K.; Yang, Y.; Pham, B.T. 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.* **2019**, *662*, 332–346.

98. Abedini, M.; Ghasemian, B.; Shirzadi, A.; Shahabi, H.; Chapi, K.; Pham, B.T.; Ahmad, B.B.; Tien Bui, D. A novel hybrid approach of bayesian logistic regression and its ensembles for landslide susceptibility assessment. *Geocarto Int.* **2018**, doi:10.1080/10106049.2018.1499820.

99. Chang, K.T.; Merghadi, A.; Yunus, A.P.; Pham, B.T.; Dou, J. Evaluating scale effects of topographic variables in landslide susceptibility models using GIS-based machine learning techniques. *Sci. Rep.* **2019**, *9*, 1–21.

100. Nohani, E.; Moharrami, M.; Sharafi, S.; Khosravi, K.; Pradhan, B.; Pham, B.T.; Lee, S.; Melesse, A.M. Landslide susceptibility mapping using different GIS-based bivariate models. *Water* **2019**, *11*, 1402.

101. Pham, B.T.; Nguyen, M.D.; Bui, K.T.; Prakash, I.; Chapi, K.; Bui, D.T. A novel artificial intelligence approach based on multi-layer perceptron neural network and biogeography based optimization for predicting coefficient of consolidation of soil. *Catena* **2019**, *173*, 302–311.

102. Nguyen, V.V.; Pham, B.T.; Vu, B.T.; Prakash, I.; Jha, S.; Shahabi, H.; Shirzadi, A.; Ba, D.N.; Kumar, R.; Chatterjee, J.M.; et al. Hybrid machine learning approaches for landslide susceptibility modelling. *Forests* **2019**, *10*, 157.

103. Pham, B.T.; Prakash, I.; Jaafari, A.; Bui, D.T. Spatial prediction of rainfall-induced landslides using aggregating one-dependence estimators classifier. *J. Indian Soc. Remote Sens.* **2018**, *46*, 1457–1470.

104. Pham, B.T.; Prakash, I. A novel hybrid model of bagging-based naïve bayes trees for landslide susceptibility. *Bull. Eng. Geol. Environ.* **2019**, *78*, 1911–1925.

105. Pham, B.T. A novel classifier based on composite hyper-cubes on iterated random projections for assessment of landslide susceptibility. *J. Geol. Soc. India* **2018**, *91*, 355–362.

106. Dou, J.; Yunus, A.P.; Xu, Y.; Zhu, Z.; Chen, C.W.; Sahana, M.; Yang, Y.; Khosravi, K.; Pham, B.T. Torrential rainfall-triggered shallow landslide characteristics and susceptibility assessment using ensemble data-driven models in the Dongjiang Reservoir Watershed, China. *Nat. Hazards* **2019**, *97*, 579–609.

107. Pham, B.T.; Prakash, I.; Dou, J.; Singh, S.K.; Trinh, P.T.; Tran, H.T.; Le, T.M.; Phong, T.V.; Khoi, D.K.; Shirzadi, A.; et al. A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. *Geocarto Int.* **2019**, 1–25.

108. Peng, Y.; Shi, Y.; Yan, H.; Chen, K.; Zhang, J. Coincidence risk analysis of floods using multivariate copulas: Case study of Jinsha River and Min River, China. *J. Hydrol. Eng.* **2018**, *24*, 05018030.

109. Le, L.M.; Ly, H.B.; Pham, B.T.; Le, V.M.; Pham, T.A.; Nguyen, D.H.; Tran, X.T.; Le, T.T. Hybrid artificial intelligence approaches for predicting buckling damage of steel columns under axial compression. *Materials* **2019**, *12*, 1670.

110. Ly, H.B.; Desceliers, C.; Le, L.M.; Le, T.T.; Pham, B.T.; Nguyen-Ngoc, L.; Doan, V.T.; Le, M. Quantification of uncertainties on the critical buckling load of columns under axial compression with uncertain random materials. *Materials* **2019**, *12*, 1828.

111. Shahabi, H.; Jarihani, B.; Tavakkoli Piralilou, S.; Chittleborough, D.; Avand, M.; Ghorbanzadeh, O. A semi-automated object-based gully networks detection using different machine learning models: A case study of Bowen Catchment, Queensland, Australia. *Sensors* **2019**, *19*, 4893.

112. Jalayer, F.; De Risi, R.; De Paola, F.; Giugni, M.; Manfredi, G.; Gasparini, P.; Topa, M.E.; Yonas, N.; Yeshitela, K.; Nebebe, A.; et al. Probabilistic GIS-based method for delineation of urban flooding risk hotspots. *Nat. Hazards* **2014**, *73*, 975–1001, doi:10.1007/s11069-014-1119-2.

113. Chapman, L. Increasing vulnerability to floods in new development areas : Evidence from Ho Chi Minh City. *Int. J. Clim. Chang. Strateg. Manag.* **2018**, doi:0.1108/IJCCSM-12-2016-0169.

114. Dano, U.L.; Balogun, A.L.; Matori, A.N.; Wan Yusouf, K.; Rimi Abubakar, I.; Mohamed, S.; Aina, Y.A.; Pradhan, B. Flood susceptibility mapping using GIS-based analytic network process: A case study of Perlis, Malaysia. *Water* **2019**, *11*, 615.

115. Zhao, G.; Pang, B.; Xu, Z.; Peng, D.; Xu, L. Assessment of urban flood susceptibility using semi-supervised machine learning model. *Sci. Total Environ.* **2019**, *659*, 940–949.

116. Khosravi, K.; Melesse, A.M.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Hong, H. Flood susceptibility mapping at Ningdu catchment, China using bivariate and data mining techniques. In *Extreme Hydrology and Climate Variability*; Elsevier: London, UK, 2019; pp. 419–434.

117. Termeh, S.V. R.; Kornejady, A.; Pourghasemi, H.R.; Keesstra, S. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. *Sci. Total Environ.* **2018**, *615*, 438–451.

118. Ahmadlou, M.; Karimi, M.; Alizadeh, S.; Shirzadi, A.; Parvinnejhad, D.; Shahabi, H.; Panahi, M. Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). *Geocarto Int.* **2019**, *34*, 1252–1272.

119. Thai Pham, B.; Tien Bui, D.; Prakash, I. Landslide susceptibility modelling using different advanced decision trees methods. *Civil Eng. Environ. Syst.* **2018**, *35*, 139–157.

120. Li, H.; Ouyang, J.; Li, F.; Xie, X. Study on safety evaluation model of small and medium-sized earth-rock dam based on BP-AdaBoost algorithm. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2019; p. 032024.

121. Avand, M.; Janizadeh, S.; Tien Bui, D.; Pham, V. H.; Ngo, P. T. T.; Nhu, V. H. A tree-based intelligence ensemble approach for spatial prediction of potential groundwater. *International Journal of Digital Earth*, **2020**. p. 1-22.

122. Kuncheva, L. *Combining Pattern Classifiers Methods and Algorithms*; John Wiley&Sons. Inc. Publication: Hoboken, NI, USA, 2014.

123. Thai Pham, B.; Shirzadi, A.; Shahabi, H.; Omidvar, E.; Singh, S.K.; Sahana, M.; Asl, D.T.; Ahmad, B.B.; Quoc, N.K.; Lee, S. Landslide susceptibility assessment by novel hybrid machine learning algorithms. *Sustainability* **2019**, *11*, 4386.

124. Dou, J.; Yunus, A.P.; Bui, D.T.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.; Han, Z.; Pham, B.T. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides* **2019**, doi:10.1007/s10346-019-01286-5.

125. Merghadi, A.; Abderrahmane, B.; Tien Bui, D. Landslide susceptibility assessment at Mila Basin (Algeria): A comparative assessment of prediction capability of advanced machine learning methods. *ISPRS Int. J. Geo-Inf.* **2019**, *7*, 268.

126. Gautam, D.; Dong, Y. Multi-hazard vulnerability of structures and lifelines due to the 2015 Gorkha earthquake and 2017 central Nepal flash flood. *J. Build. Eng.* **2018**, *17*, 196–201.

127. Eem, S.-h.; Yang, B.-j.; Jeon, H. Simplified methodology for urban flood damage assessment at building scale using open data. *J. Coast. Res.* **2018**, *85*, 1396–1400.