Journal of Hydrology 592 (2021) 125815

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Flood risk assessment using hybrid artificial intelligence models integrated with multi-criteria decision analysis in Quang Nam Province, Vietnam

Binh Thai Pham[a,*], Chinh Luu[b,*], Tran Van Phong[c], Huu Duy Nguyen[d,*], Hiep Van Le[e,*], Thai Quoc Tran[f], Huong Thu Ta[g], Indra Prakash[h]

[a] University of Transport Technology, Hanoi 100000, Viet Nam
[b] Faculty of Hydraulic Engineering, National University of Civil Engineering, Hanoi 100000, Viet Nam
[c] Institute of Geological Sciences, Vietnam Academy of Sciences and Technology, 84 Chua Lang Street, Dong da, Hanoi 100000, Viet Nam
[d] Faculty of Geography, VNU University of Science, Vietnam National University, 334 Nguyen Trai, Hanoi 100000, Viet Nam
[e] Institute of Research and Development, Duy Tan University, Da Nang 550000, Viet Nam
[f] Department of Urban Planning, National University of Civil Engineering, Hanoi 100000, Viet Nam
[g] Vietnam Academy for Water Resources, Hanoi 100000, Viet Nam
[h] Dy. Director General (R), Geological Survey of India, Gandhinagar, (Guj.) 382002, India

ARTICLE INFO

This manuscript was handled by Emmanouil Anagnostou, Editor-in-Chief, with the assistance of YYY, Associate Editor

Keywords:
Flood risk assessment
Machine learning
Multi-criteria decision analysis
Quang Nam, Vietnam

ABSTRACT

Flood risk assessment is an important task for disaster management activities in flood-prone areas. Therefore, it is crucial to develop accurate flood risk assessment maps. In this study, we proposed a flood risk assessment framework which combines flood susceptibility assessment and flood consequences (human health and financial impact) for developing a final flood risk assessment map using Multi-Criteria Decision Analysis (MCDA) method. Two hybrid Artificial Intelligence (AI) models, namely ABMDT (AdaBoost-DT) and BDT (Bagging-DT) were developed with Decision Table (DT) as a base classifier for creating a flood susceptibility map. We used 847 flood locations of major flooding events in the years 2007, 2009 and 2013 in Quang Nam province of Vietnam; and 14 flood influencing factors of topography, geology, hydrology and environment to construct and validate the hybrid AI models. Various statistical measures were used to validate the models, including the Area Under Receiver Operating Characteristic (ROC) Curve called AUC. Results show that all the proposed models performed well, but the performance of the BDT model (AUC = 0.96) is the best in comparison to other models ABMDT (AUC = 0.953) and single DT (AUC = 0.929). Therefore, the flood susceptibility map produced by the BDT model was used to combine with a flood consequences map to develop a reliable flood risk assessment map for the study area. The final flood risk map can provide a useful source for better flood hazard management of the study area, and the proposed framework and models can be applied to other flood-prone areas.

1. Introduction

Flood is one of the most devastating natural hazards affecting life, damaging properties, disrupting communication and causing death all over the world (Khosravi, 2018; Khosravi et al., 2016; Luu et al., 2018; Nardi et al., 2006; Rahmati et al., 2016). Floods affect the socio-economic condition of the country (Degiorgis, 2013, 2012b; Tran et al., 2009; Van Aalst and Burton, 2002). Global climate change has recently increased the frequency and duration of precipitation and thus the incidence of flooding in the river basins and low-lying areas (Bui et al., 2019a; Nguyen et al., 2018). Climate change effects have also resulted in increasing flood incidents due to the increase in rainfall intensity and frequency (Bouwer et al., 2010). In addition, anthropogenic activities such as deforestation and land-use pattern change have contributed to severe flooding incidents (Bubeck et al., 2012). Flood risk has been increasing in coastal and low-lying areas due to migration of population and climate change effects (Nicholls et al., 1999).

Nowadays, flood risk assessment is gaining importance for proper flood management of an area (Manfreda and Samela, 2019; Merz et al., 2010). Flood risk can be determined as a product of the probability and potential consequences of a flood event (Sayers et al., 2002). Flood risk can also be considered the probability of losses and includes three

* Corresponding authors.
E-mail addresses: binhpt@utt.edu.vn (B.T. Pham), luuthidieuchinh@nuce.edu.vn (C. Luu), nguyenhuuduy@hus.edu.vn (H.D. Nguyen), levanhiep2@duytan.edu.vn (H.V. Le).

https://doi.org/10.1016/j.jhydrol.2020.125815
Received 11 September 2020; Accepted 25 November 2020
Available online 30 November 2020

components of hazard, exposure and vulnerability (Ronco, 2014; Winsemius et al., 2013). The probability of a flood occurring at one location depends on geo-morphological factors and geo-environmental factors (Khosravi et al., 2016). The potential consequences concern humans and exposed elements (Jato-Espino et al., 2019). Currently, flood risk studies have also been developed at the interface between nature and society (Brown and Damery, 2002).

Many methods can be used to incorporate different indicators into an integrated tool for flood risk assessment and management. In this regard, Multi-Criteria Decision Analysis (MCDA) methods have been widely used for combining, integrating and evaluating flood risk-related factors. Ozturk and Batuk (2011) applied Analytic Hierarchy Process (AHP), which is one of the popular MCDA methods for identifying the flood vulnerability of South Marmara Basin, Turkey and then integrated the risk assessment into a Geographic Information System (GIS) framework. Kappes et al. (2012) presented an indicator-based framework for assessing vulnerability for natural hazards using MCDA approach. Papathoma-Köhle (2016) used this approach to determine the debris-flow hazard for South Tyrol, in which MCDA approach was applied to assign the weights of vulnerability criteria. Toosi et al. (2019) used the SWAT model to generate a runoff coefficient map and then combined flood hazard indexes using AHP method to create a flood hazard assessment map for Mashhad Plain Basin, Iran. However, when there are a large amount of data, and a large number of criteria need to be are analyzed for flood studies, it is difficult to estimate which criteria are relevant or influential in decision-making processes using MCDM methods. This is one of the limiations of MCDA approach where complex decisions are needed for the flood management (Ishizaka and Labib, 2009). Therefore, the Artificial Intelligence (AI) or Machine Learning (ML) methods can be integrated with MCDA for the automatic estimation of weights of each criterion to avoid human bias.

In recent years, many ML methods have been used to predict natural hazards in general and flood hazards in particular. These methods include Artificial Neural Networks (ANN) (Falah, 2019; Lekkas et al., 2004; Rang et al., 2006), Support Vector Machines (SVM) (Bafitlhile and Li, 2019; Choubin, 2019; Shrifi Garmdareh et al., 2019; Xiong et al., 2019), Random Forest (RF) (Naghibi et al., 2019; Victoriano et al., 2019), Logistic Regression (LR) (Nandi et al., 2016; Pradhan, 2009). The performance of hybrid ML models is better to compare with individual models in many cases (Jaafari et al., 2018; Tien Bui, 2016). The development of new hybrid ML models is intended to further improve the predictive performance of the models (Bui et al., 2019c; Bui and Pham, 2018; Tien Bui, 2018).

In this study, the main objective was to develop a novel approach which is a combination of hybrid ML models and MCDA technique for creating a flood risk assessment map for Quang Nam province which is one of the flood-prone areas of Vietnam. With this aim, we developed two hybrid ML models, namely AdaBoost-DT and Bagging-DT with Decision Table (DT) as a base classifier for flood susceptibility assessment and mapping. One of the popular MCDA methods named AHP was employed to create a flood consequences map, which was then used to combine with the flood susceptibility map to make the final flood risk assessment map. The main difference of this study with previous studies is that this is the first time the AI or ML models integrated with MCDA for developing the flood risk framework and assessment. In addition, this integrated approach can provide a useful tool for flood risk assessment mapping, which can be applied to decision-making processes in flood risk management. The model performance was evaluated with the standard statistical indices measures, including Area Under the Receiver Operating Characteristic (ROC) Curve called AUC. Weka and ArcGIS programs were used for data analysis, and model development and visualization.

2. Study area

Quang Nam province ($14^0 57'10''$ to $16003'50''$ Northern latitude; $107012'40''$ to $108044'20''$ Eastern longitude) is located in the central coastal region of Vietnam covering about 1,057,474 ha area (Fig. 1). This is a key economic region of Central Vietnam, bordered in the North by Danang City, and in the South by Dung Quat Economic Zone. The topography of the province is complex varying from high mountains in the west; midland in the middle; delta and coastal areas in the east. Two river basins of Vu Gia and Thu Bon flow through the province and form valleys and flood plains.

The climate in the region is a moderate tropical type with a dry season and a rainy season. Total annual rainfall ranges from 2000 mm to 4000 mm (Pham et al., 2019c). Rainfall during September to December period accounts for 70–75% annual rainfall. Rainfall in mountainous areas is much higher than in the plains (Fig. 3(b)). The rainy season coincides with the typhoon season, so storms and tropical depressions often cause landslides and flash floods in mountainous districts of Nam Tra My, Bac Tra My, Tay Giang, Dong Giang and Nam Giang and flooding in plain districts of Dai Loc, Dien Ban, Duy Xuyen and Hoi An. Most of the area of the province is covered by forest (Fig. 3(m)). However, the forest area has been reduced in recent years due to agricultural expansion, dam construction and commercial development (Meyfroidt et al., 2013). Anthropogenic activities and climate change effect have increased the risk of flooding in this province (Luu and von Meding, 2018).

3. Material and methods

3.1. Conceptual framework

Flood risk can be defined as the combination of the probability of occurrence of a flood event and its associated negative consequences which may cover different types of impacts (Schanze, 2006). The likelihood or susceptibility of a flood can be integrated with consequences in risk estimation (Woodruff, 2005). The flood risk can be managed by reducing the consequences it may cause (de Moel, 2015). Therefore, in this study, we assessed flood risk by combining the flood susceptibility and its negative consequences, as indicated in the following equation:

$$Floodrisk = Susceptibility^{*}Consequences \tag{1}$$

The potential impacts of flooding are often considered along with flood susceptibility in assessing flood risk (Merz, 2014). The consequence indexes considered in this study include human health and financial implications. The indicators are selected based on the critical analysis of past and present flood-related data. The description of flood consequences is depicted in Fig. 2.

3.2. Geospatial database

3.2.1. Flood influencing factors

Physical, anthropogenic and meteorological factors affect the occurrence of floods in any area. In this study, 14 flood influencing factors were selected based on the analysis of the topography, geo-environment conditions, meteorology and anthropogenic activities considering the nature of flood occurrences of the study area; and other published works (Costache, 2019; Hosseini, 2019; Wang, 2019). Topography and hydrology factors (Slope, Curvature, Curvature plan, Curvature profile, Flow accumulation, Elevation, Topographic Wetness Index (TWI), Sediment Transport Index (STI), Stream Power Index (SPI), River density and Distance from rivers) were extracted from Aster Digital Elevation Model (https://earthexplorer.usgs.gov). Lithology map was obtained from the Geology Institute, Vietnam. Land cover map was developed from Google Earth images and data of the Department of Natural Resources and Environment of Quang Nam province. Rainfall map was developed from the data of Meteorology Institute, Vietnam. Thematic maps of the flood influencing factors are presented in Fig. 4. In addition, we carried out the frequency analysis to evaluate the probability of past and present flood occurrences on each class of the flood

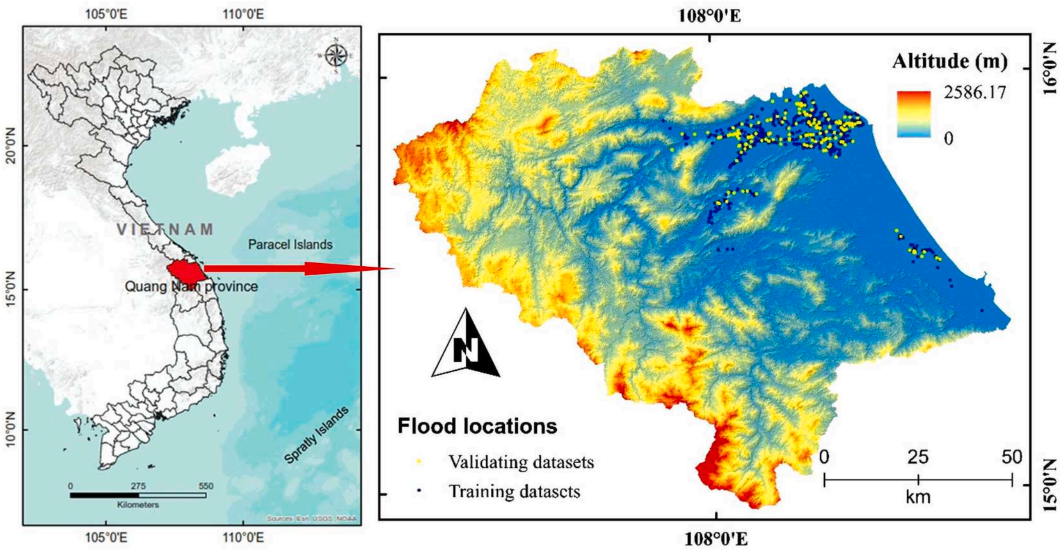

Fig. 1. Location of the study area, Quang Nam province.

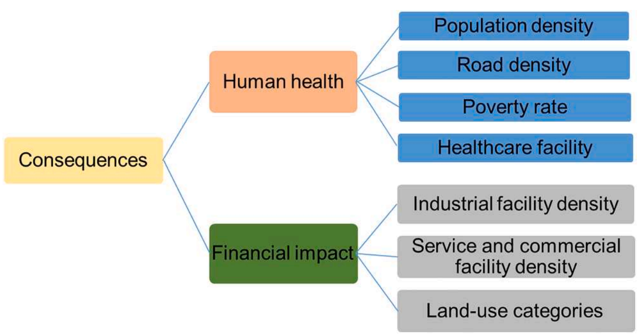

Fig. 2. Conceptual model of the flood consequences indicators.

Fig. 3. An example of a flood mark for the 2013 flood event constructed by Quang Nam Provincial Committee of Natural Disaster Prevention and Control.

conditioning factors. In frequency analysis, the frequency ratio was calculated by dividing the percentage of past and present flood pixels by the percentage of the class pixels on each class of each factor map (Rahmati et al., 2016). In each factor map, the classes with higher frequency ratio values have a higher probability of flood occurrences than those with lower frequency ratio values (Siahkamari et al., 2018). The results of the frequency analysis of these factors are presented in Fig. 5.

3.2.2. *Flood consequences factors*

Floods have many positive and negative consequences for the communities and natural environment (Albano et al., 2014). Positive consequences include replenishing nutrients in flood plain areas, filling of the reservoirs, recharging of the groundwater and removal of contaminants/pollutants. The negative consequences of flooding include disruption of life, fatalities and health problems (Hajat, 2005), disruption of communication, destruction of crops and adverse financial impacts on societies (Thieken et al., 2014). Flood consequence indicators in the flood zone area vary from place to place depending on local conditions. In this study, we analyzed adverse flood consequences taking into account human health (Fig. 6) and financial aspects (Fig. 7) to assess flood risk.

3.2.3. *Flood consequences on human health*

The flood consequences on human health are based on the population density, available infrastructures and the medical capacity of the area (Scheuer et al., 2011). Population density is the most important indicator of flood consequences analysis since it directly involves humans (Grothmann and Reusswig, 2006). Thus, flood consequences would be greater for a large number of people at risk. The poverty rate and flood risk are interrelated in Vietnam (Tran et al., 2009). A higher poverty rate receives a higher weight for flood consequences. The mountainous areas often have a high poverty rate (>40%). In the present study, the socio-economic indicators including population density, poverty rate, and number of medical staff available in the study area were synthesized from the 2015 statistical yearbooks of 18 districts in Quang Nam province. The road density (indicator) is derived from the intersecting transportation network and the commune boundary. The well-connected areas provide additional response support to community (Gain and Hoque, 2013; Masuya et al., 2015). The number of medical staff criterion affects the response activities during flood events (Scheuer et al., 2011).

3.2.4. *Flood consequences on financial impact*

The financial impact can be estimated based on the concentration of economic activities such as the density of industrial facilities and services in the research area (Jato-Espino et al., 2019). The industrial facility density, and service and commercial facility density data were synthesized from the 2015 statistical yearbooks of 18 districts in Quang Nam province. In addition, the land-use categories indicator was also

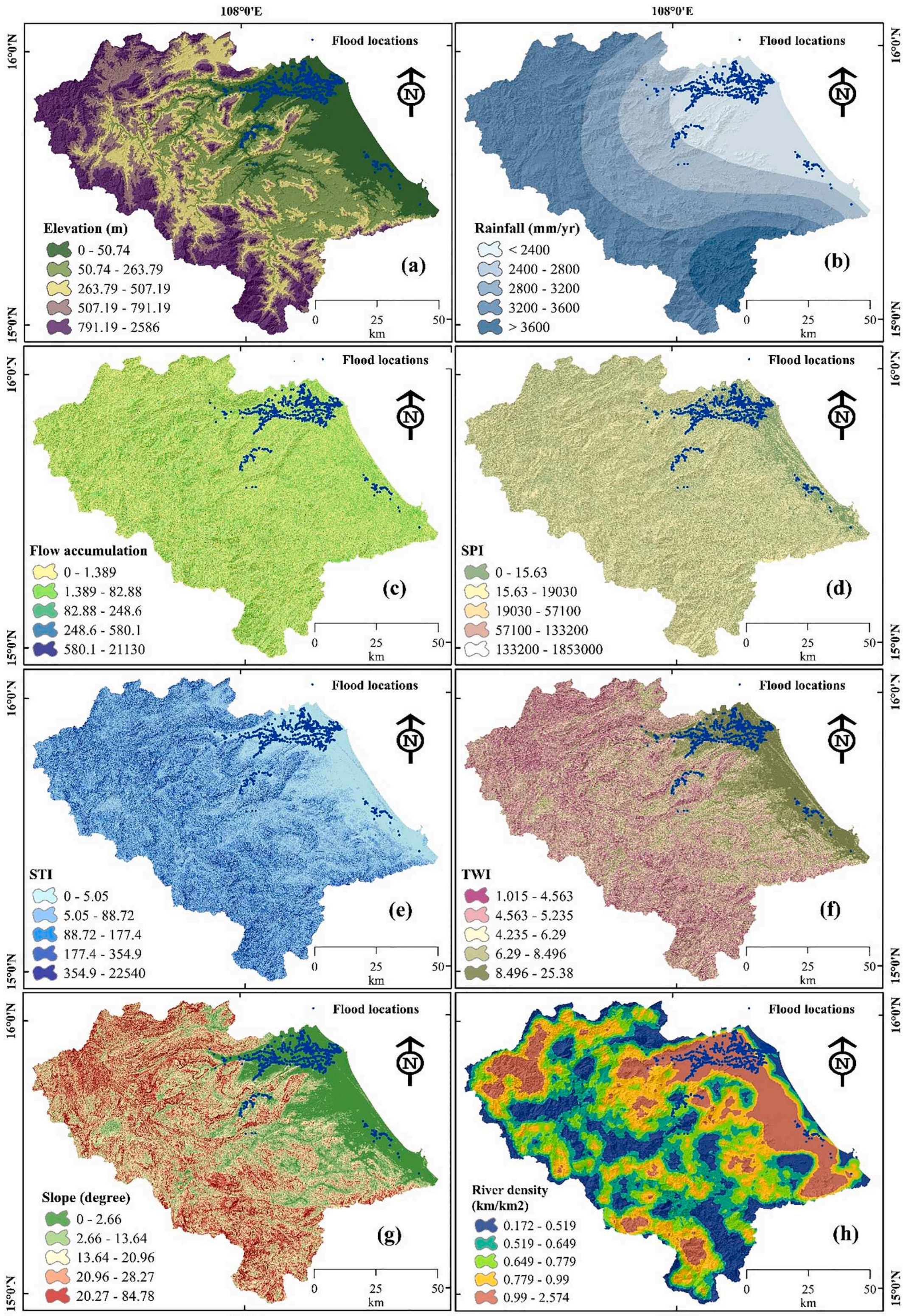

Fig. 4. Maps of flood influencing factors: (a) elevation, (b) rainfall, (c) flow accumulation, (d) SPI, (e) STI, (f) TWI, (g) slope, (h) river density, (i) distance from rivers, (j) plan curvature, (k) profile curvature, (l) curvature, (m) land cover, and (n) lithology.

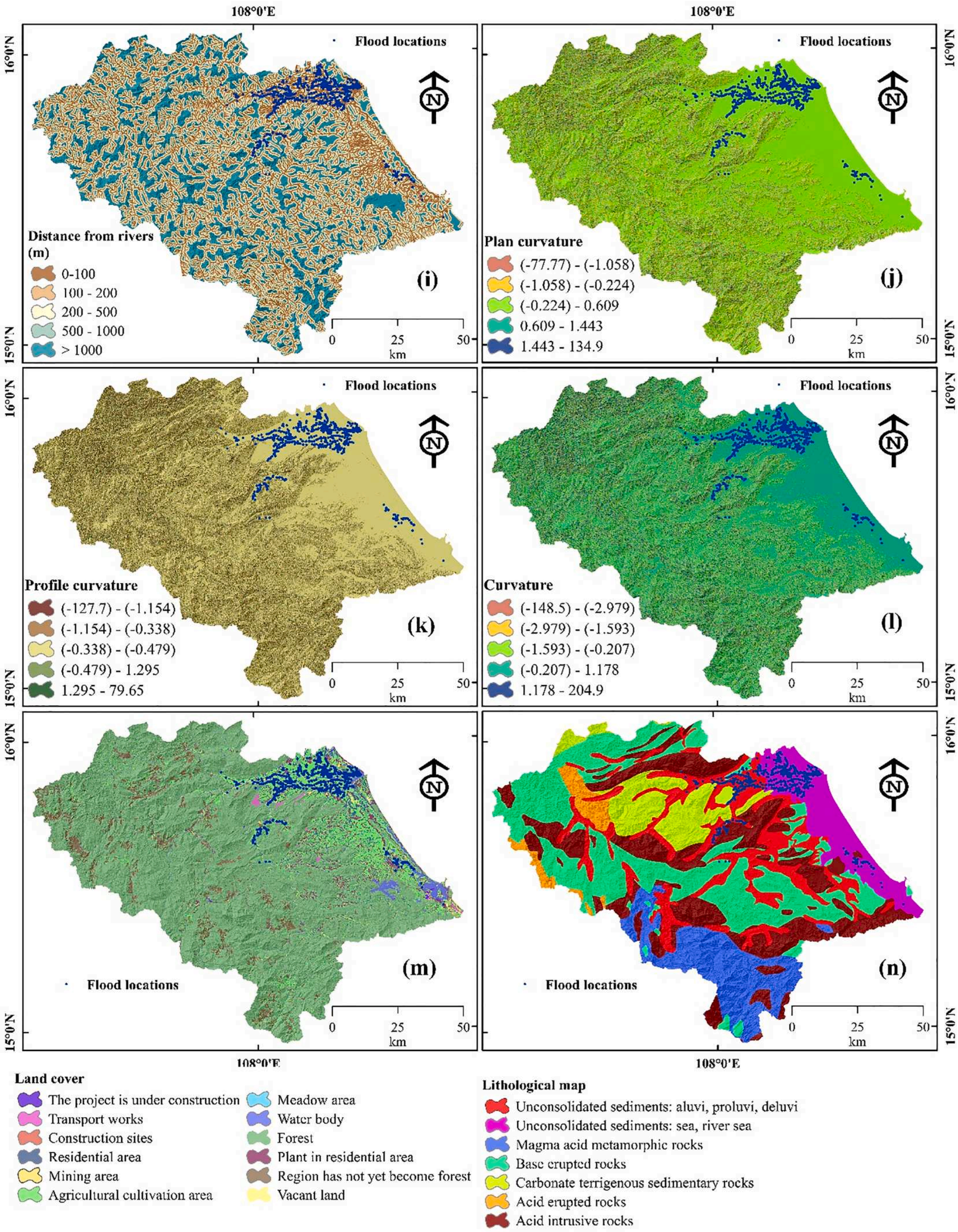

Fig. 4. (*continued*).

considered in flood consequences assessment (Jongman et al., 2012). In the study area, the potential impact of flooding is highest for homestead and built-up areas as it is directly related to people. The agricultural land indicator has the second-highest impact because it affects the residents' livelihood.

3.3. *Methodology flowchart for flood risk assessment*

The methodology of the present study involves three main steps: (i) flood susceptibility assessment, (ii) flood consequences assessment, and (iii) flood risk assessment. The methodology flowchart is shown in Fig. 8, and the details of the three steps are presented in the following subsections:

3.3.1. *Flood susceptibility assessment*

First, preprocessing of input data to select relevant features for the modelling was done using Relief-F feature selection method, which is a simple and effective approach of feature weight estimation. A spatial geodatabase of flood susceptibility factors and flood locations was used to generate the training (70% flood locations) and testing (30% flood locations) datasets. Frequency ratio analysis was done to create the weights of the classes of the flood conditioning factors used in the modelling. Second, training dataset was used to construct the susceptibility maps using hybrid AI techniques, namely ensemble ABMDT and BDT; and single DT. In the ABMDT method, AdaBoost ensemble was first used to generate the optimal training dataset, which was then used for training the base classifier DT. In BDT method, Bagging ensemble was first used to generate the optimal training dataset, which was then used

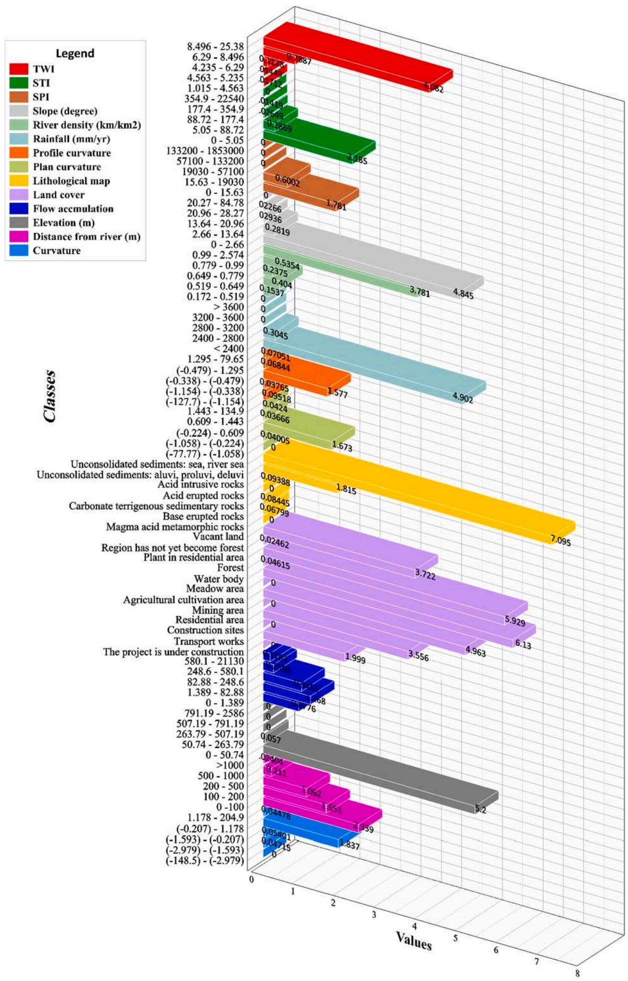

Fig. 5. Frequency analysis of past and present flood occurrences on the factor maps.

for training the base classifier DT. Third, the validation of these models was done on both the training and testing datasets using various standard quantitative indices, including AUC. Fourth, the results of these indices were evaluated for selecting the best method. Finally, flood susceptibility maps were constructed using these models, and the best flood susceptibility assessment map was then used to generate a final flood risk assessment map.

3.3.2. Flood inventory map

Flood inventory represents the spatial relationships between the history of flood and factors of physical and human geography such as geomorphology, climate, hydrology and human activities (Bui et al., 2019b). It plays an important role in modelling and generating flood susceptibility and risk maps. The study area of Quang Nam province is affected by several flood events each year due to its geographic location. The drainage basins of the area are steep slopes and short rivers, so the runoff is substantial in the mountainous regions. This causes rapidly flooding in low-lying areas.

In this study, the flood inventory map consists of 847 historical flooded locations which occurred in the year 2007, 2009 and 2013 (Fig. 1). The flooded locations were identified from the historical flood marks (Fig. 3). At a flood mark location, the information of code, coordinates (longitude and latitude), address, and flood depth were recorded. The flood inventory data were collected from the Quang Nam Provincial Committee of Natural Disaster Prevention and Control. The data was used for the development of flood susceptibility maps. We also referenced the inundation maps data of previous studies by Luu et al. (2018), Chau et al. (2013) and Ho and Umitsu (2011); and generated the same number of non-flood locations.

3.3.3. Flood consequences assessment

In this step, the maps of flood consequences related to human health and financial impact were combined to create the final flood consequences assessment map. The weights of the layers were generated using the AHP method.

3.3.4. Flood risk assessment

In the final step, the best flood susceptibility assessment map in the first step was combined with the consequences assessment map to generate the final flood risk assessment map for the study area. We used the *Weighted Sum* tool to integrate layers in GIS application.

3.4. Methods used

3.4.1. Decision Table (DT)

Decision tables are classification models used for prediction (Kohavi, 1997). It generates several training variables which made up the labelled instances. The DT is invested for the majorities of class with a rules mapping. It is made up of two components: (1) the set of functionalities designated by the diagram and (2) the set of the majority of labelled instances (Kohavi, 1997). Each instance has a value for a feature and is labelled in the schema (Kohavi, 1997). With an instance is not labelled *I*, the labelled attributes to the instance classified by DT is calculated as follows:

If τ is the entire instance that is labelled in the corresponding DT with the given instance *I*, where the instances in the schema request a match and all other functionality is ignored. If $\tau = 0$, return the majority classes in DT, otherwise return the majority classes in τ, the values do not determine which calls the distinct values in the process. If err (h, f) presents the hypothesis errors *h* for the function *f*, we estimate errors when using independent test data τ as in the following equation:

$$\widehat{err}(h,\tau) = \frac{1}{|\tau|}\sum_{(x_i;y_i)} L(h(x_i),(y_i)) \quad (2)$$

In this study, the DT was trained with hyper-parameters provided in Table 1.

3.4.2. AdaBoost

AdaBoost algorithm is considered to be one of the popular methods that are used to solve binary classification math. Its algorithm is replaced in the multiclass without dividing it to several problems with two classes (Fan and Wang, 2011). In addition, this algorithm combines with a weak classifier to build a powerful hybrid algorithm. The weak classifications are assigned weights to improve the classifications of the previous incorrectly classified sampling. The construction model is the decision tree model, which is supported by the model coefficients. In this study, the AdaBoost was trained and constructed with hyper-parameters provided in Table 1.

3.4.3. Bagging

Bagging is considered to be one of the smooth machine learning methods proposed by Breiman in 1996 (Breiman, 1996). It is proposed to reduce variance without increasing too much bias error. Bagging method is a valuable tool for evaluating the natural disasters in general and the flood in particular. It is related to the model prediction improvement capacity due to the capacity for sensitivity in small changes to training data. In Bagging, the bootstrap technique is used to select the variables randomly with the replacement capacity to construct several samples to create the training data (Galar et al., 2011). Each of the subsets is prepared to build the decision tree, which is synthesized in the final models. Class prediction of the new element comes with the majority voting strategy derived from the class prediction of the basic

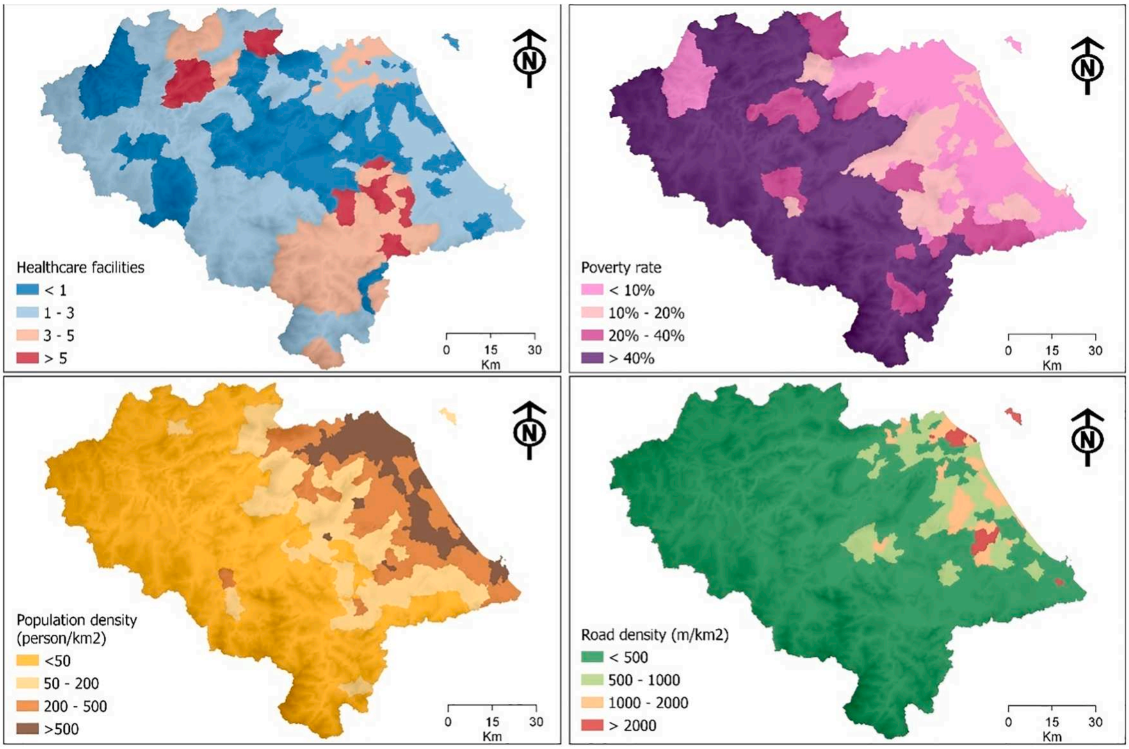

Fig. 6. Maps of flood consequences on human health.

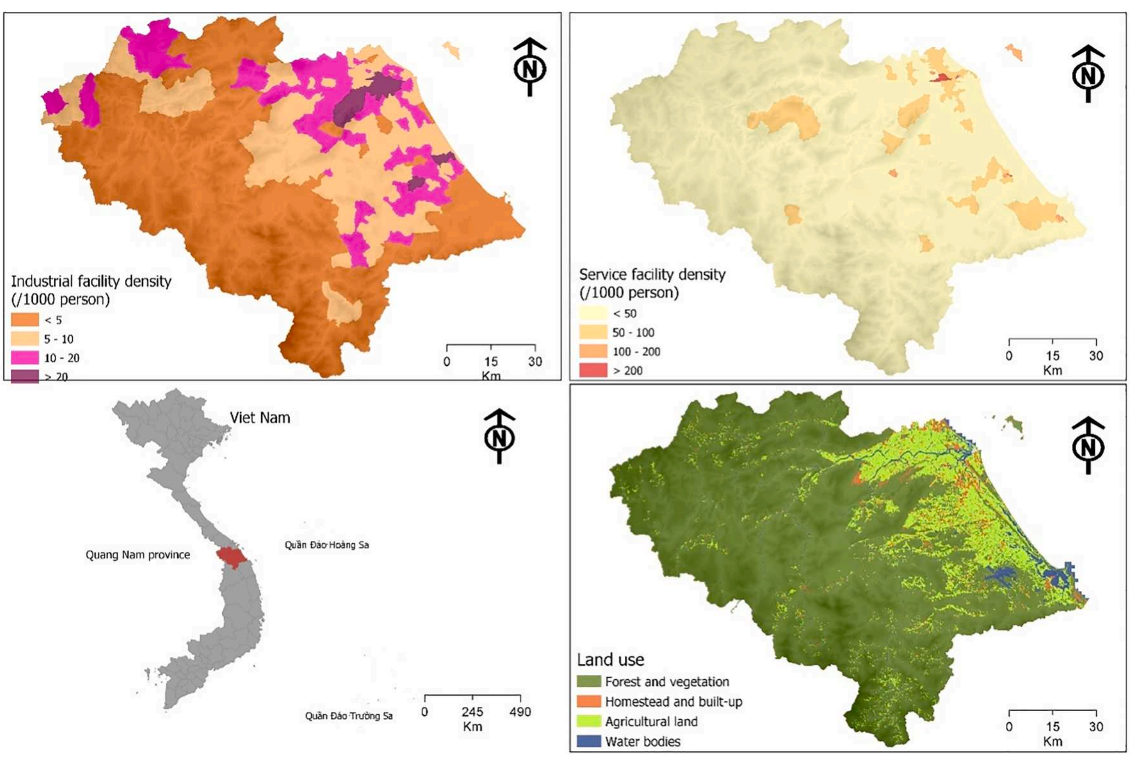

Fig. 7. Maps of flood consequences on financial impact.

model. In this study, the Bagging was trained and constructed with hyper-parameters provided in Table 1.

3.4.4. *Relief-F feature selection method*

The selection of influencing factors plays an essential role in the process of constructing the flood susceptibility model because all the features (factors) may not be relevant in modelling (Hong, 2018; Wang et al., 2016; Yariyan, 2020a). Feature selection is considered to be one of the most critical issues in the sample determination process, and this supported by machine learning focuses the algorithm on the factors most significant influence on the classification prediction (Van Dao, 2020). The Relief-F method is considered to be one of the most popular methods

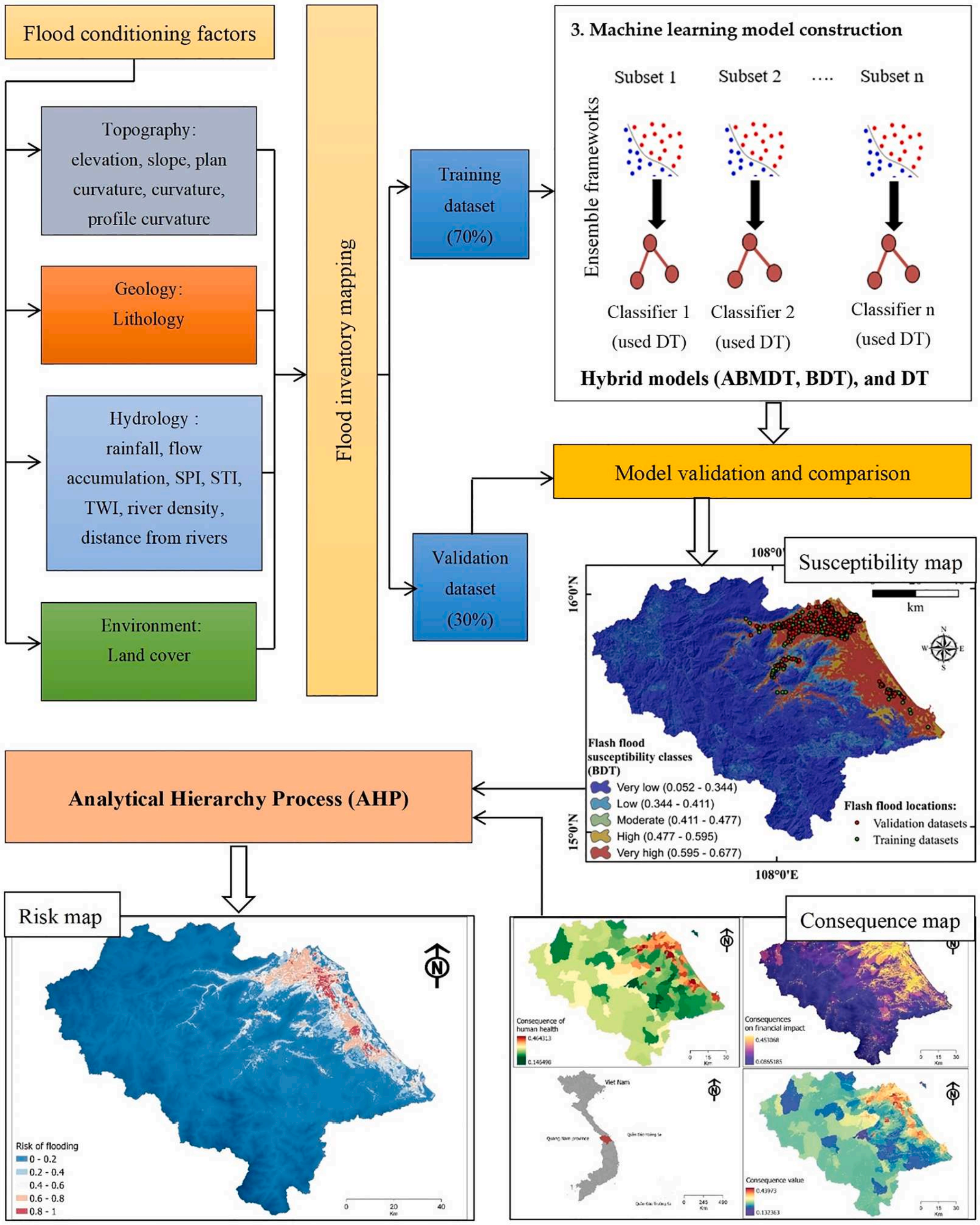

Fig. 8. . The methodology of flood risk assessment used in this study.

Table 1
Hyper-parameters of models used in this study.

No	Parameters	Models		
		DT	ABMDT	BDT
1	Batch Size	100	100	100
2	Cross validation	1	–	–
3	Number of decimal places	2	2	2
4	Search	Best Frist	–	–
5	Number of interations	–	10	10
6	Seed	–	1	1
7	Weight Threshold	–	100	–
8	Number of execution slots	–	–	1
9	Base classifier	–	DT	DT

for performing this task. This algorithm has the ability to not only support binary class problems but also to solve binary class problems (Urbanowicz et al., 2018). This algorithm is based on the summaries of the weightings of the factors according to their suitability for the objective (Saha et al., 2020).

3.4.5. Validation methods

Statistical measures used to validate the models include True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), Positive predictive value (PPV), Negative predictive value (NPV), Sensitivity (SST), Specificity (SPF), Accuracy (ACC), Kappa (K), Root Mean Squared Error (RMSE) and Reciever Operating Characteristic (ROC) curve. These methods were used to evaluate the performance of models (Pham et al., 2019b) for the development of reliable flood susceptibility maps.

In this study, statistical measures PPV and NPV are the proportion of pixels that are correctly classified as flood and non-flood, respectively. SST and SPF are the rates of the number of flood and non-flood pixels. ACC is one of the most popular metrics that is used to assess the performance of models (Bennett, 2013; Janizadeh, 2019). This measure is represented by the proportion of the number of pixel flood and non-flood pixel. Cohen's kappa coefficient (k) is an efficient statistics measure that helps to measure random consensus among classification components. It calculates the consensus between the two evaluations variables used to classify N objects into C mutually exclusive sets. The value of kappa varies from 0 to 1, where value 1 indicates perfect model performance (Pham et al., 2016, 2018). RMSE indicates the level of dispersion of predicted values from actual values. The model is considered perfect when the RMSE value is small. ROC is one of the most used measures by researchers to assess the predictive capacity and classification quality of the model (Gorsevski et al., 2006; Mohammady et al., 2012). Area Under the ROC curve (AUC) is often used to quantitatively evaluate the predictive capability of the models. Value of AUC equals to 1 indicates perfect model performance.

In addition, the Wilcoxon signed-rank test was used to check the statistical differences between the flood models (Pham et al., 2019a). This algorithm has been applied widely to verify the statistical significance between models (Khosravi, 2018). To apply this method, two values (p and Z) were used with the null hypothesis that the models were not statistically different at the level of 0.05, which was constructed. If the Z values are higher than the upper limit value (1.96) and lower limit value (−1.96), as well as $p < 0.05$. This assumption was rejected; therefore, the performance of the models is different in terms of statistics (Dietterich, 1998).

3.4.6. Multi-Criteria decision analysis (MCDA)

MCDA methods allow to work with quantitative variables and are applied to decision-making processes (Huang et al., 2011). Several popular MCDA methods used in flood risk assessment are Analytical Hierarchy Process (AHP) (Saaty, 2003), Analytic Network Process (ANP) (Saaty and Vargas, 2006), and Ordered Weighted Averaging (OWA) (Makropoulos and Butler, 2006). The integration of MCDA methods with GIS analytical techniques are used for spatial analysis of various assessments and decision problems (Malczewski, 2006). Weighted Linear Combination (WLC) method is often used to integrate MCDA methods in GIS environment (Malczewski, 2000). In this study, we used one of the most popular MCDA methods, namely AHP, to generate weights for the indicators. The WLC method was then used to incorporate the weighted indicator layers into a GIS framework.

Table 2
Importance of the flood factors using Relief-F feature selection.

Rank	Weights (W)	Factor conditions
1	0.7235	Elevation
2	0.6729	Rainfall
3	0.6069	Slope
4	0.5982	Land cover
5	0.5803	Lithology
6	0.4384	TWI
7	0.2949	STI
8	0.2743	River density
9	0.2587	Curvature
10	0.2435	Plan curvature
11	0.1929	Distance from rivers
12	0.1856	Profile curvature
13	0.0819	SPI
14	0.0152	Flow accumulation

4. Results and analysis

4.1. Flood susceptibility assessment map

4.1.1. Important factors used for the flood models

In the susceptibility modelling, it is desirable to remove unnecessary or unimportant factors which might affect the accuracy of the prediction models. Therefore, feature selection techniques are being used to validate and select the important factors for developing machine learning models (Hoa, 2019). In this study, Relief-F, which is a popular feature selection method (Yang et al., 2011), was applied in the model study (Table 2). The results suggest that elevation (W = 0.7235), rainfall (W = 0.6729), slope (W = 0.6069) are the most important influencing factors (W > 6) for the flood modelling in the study area. This is consistent with other similar flood studies (Degiorgis, 2012a; Garrote and Bras, 1995; Hosseini, 2020; Shrestha, 2014).

4.1.2. Validation of the models

Validation of the studied susceptibility models (ABMDT, BDT, and DT) was carried out using various standard quantitative indices on both training and testing datasets (Table 3 and Fig. 9). In terms of training data, the results show that the BDT model has a better value of NPV (99.115%) and SST (99.032%) while the ABMDT has a better value of PPV (92.183%), SPF (92.639), and ACC (95.280) compared to the other models. For validation data, the BDT model has the best value of PPV (86.391%), NPV (99.408%), SST (99.320%), SPF (87.958%), and ACC (92.899) compared to other models. The best k value of 0.906 is for the ABMDT model in terms of training data, whereas the best k value for the BDT model is 0.858 in terms of testing data.

The ABMDT model has the lowest value of RMSE on training dataset (0.197), whereas the BDT model has the lowest RMSE value on testing dataset (0.245) (Table 3). The results of the ROC analysis show that the models of BDT and ABMDT have better performance with AUC (0.977), then the DT model (0.961) in terms of training data. For validation data, the BDT model has better performance with AUC (0.96), followed by the ABMDT with AUC (0.953) and DT model with AUC (0.929), respectively (Fig. 9).

Wilcoxon signed-rank test was carried out to evaluate the significance of the statistical differences of the model performance (Tables 4 and 5). It can be observed that on both training and testing datasets, the Z-values of the comparative pairs are higher than the upper limit value (1.96) and lower limit value (−1.96), and the p-values are smaller than the threshold value of 0.05. Therefore, it can be stated that the differences in the performance of the models are statistical significance in this study.

4.1.3. Flood susceptibility mapping

Flood susceptibility indexes of the study area were generated during the training phase of the models. These indexes were used for the construction and classification of flood susceptibility maps into five classes: very low susceptibility, low susceptibility, moderate susceptibility, high susceptibility, and very high susceptibility using natural break classification method (Fig. 10). It can be observed that for the BDT model,

Table 3
Validation of the models using quantitative indices.

Parameters	Training dataset			Validating dataset		
	BDT	ABMDT	DT	BDT	ABMDT	DT
PPV (%)	90.560	92.183	90.560	86.391	86.391	85.799
NPV (%)	99.115	98.378	98.378	99.408	98.225	98.817
SST (%)	99.032	98.270	98.240	99.320	97.987	98.639
SPF (%)	91.304	92.639	91.245	87.958	87.831	87.435
ACC (%)	94.838	95.280	94.469	92.899	92.308	92.308
k	0.897	0.906	0.889	0.858	0.846	0.846
RMSE	0.228	0.197	0.219	0.245	0.249	0.262

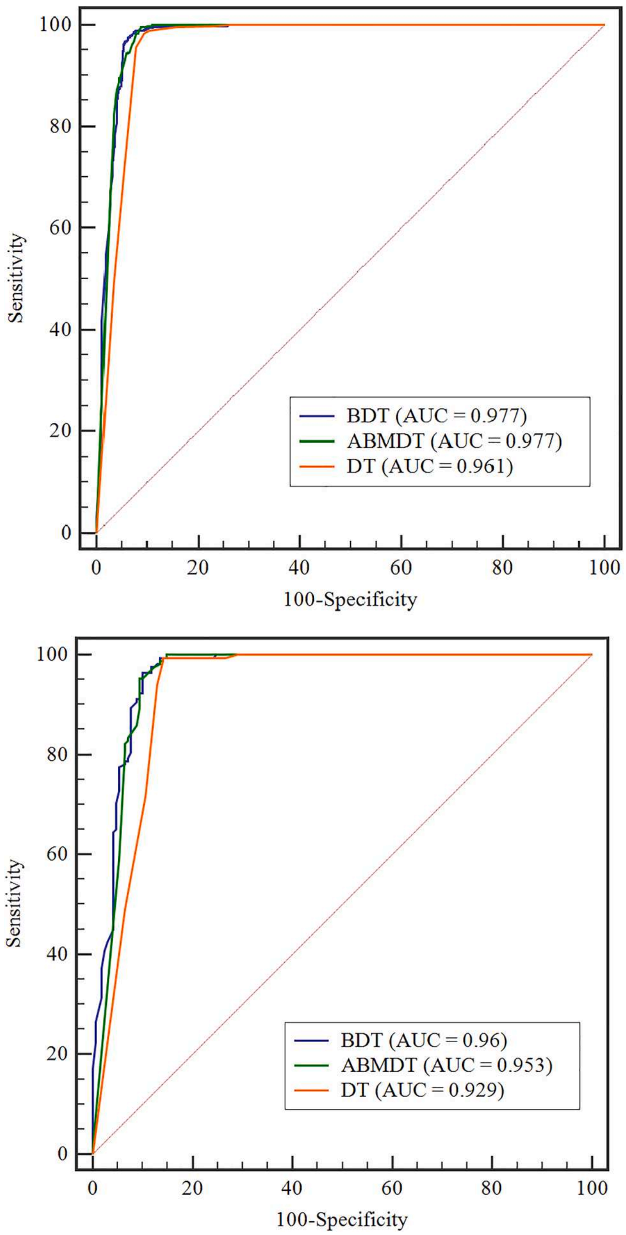

Fig. 9. Analysis of ROC of the models: (a) training dataset and (b) validating dataset.

Table 4
Z-values, *p*-values and significant levels of the different models using Wilcoxon signed-rank test and training dataset.

No	Comparative pairs	*Z*-value	*p*-value	Significance level
1	BDT vs ABMDT	14.337	<0.0001	Yes
2	BDT vs DT	21.434	<0.0001	Yes
3	ABMDT vs DT	6.049	<0.0001	Yes

Table 5
Z and p values and significant levels of the different models using Wilcoxon signed-rank test and testing dataset.

No	Comparative pairs	Z-value	p-value	Significance level
1	BDT vs ABMDT	5.926	<0.0001	Yes
2	BDT vs DT	10.246	<0.0001	Yes
3	ABMDT vs DT	4.028	0.0001	Yes

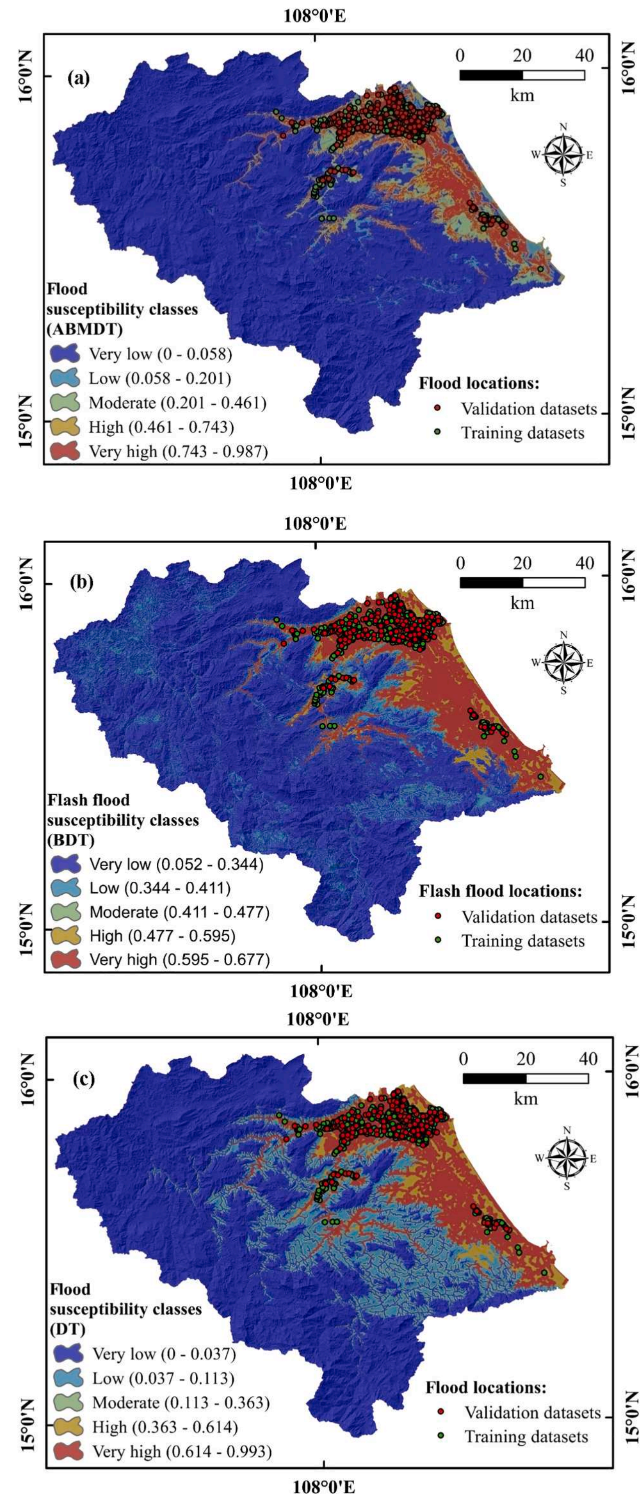

Fig. 10. Flood susceptibility maps using different models: (a) ABMDT, (b) BDT, and DT.

15.9% of the study area is located in the very low susceptibility area, 45.94% in the low susceptibility area, 19.8% in the moderate susceptibility area, and 4.682% and 14.49% in the high susceptibility area and very high susceptibility area. For the ABMDT model, 80.13% of the study area is in the very low susceptibility area, 2.794%, 2.83%, 4.617%, and9.626% is in the low, moderate, high and very high susceptibility areas, respectively. For the DT model, 64.5% of the study area is located in the very low susceptibility area, 11.09%, 5.421%, 4.045%,

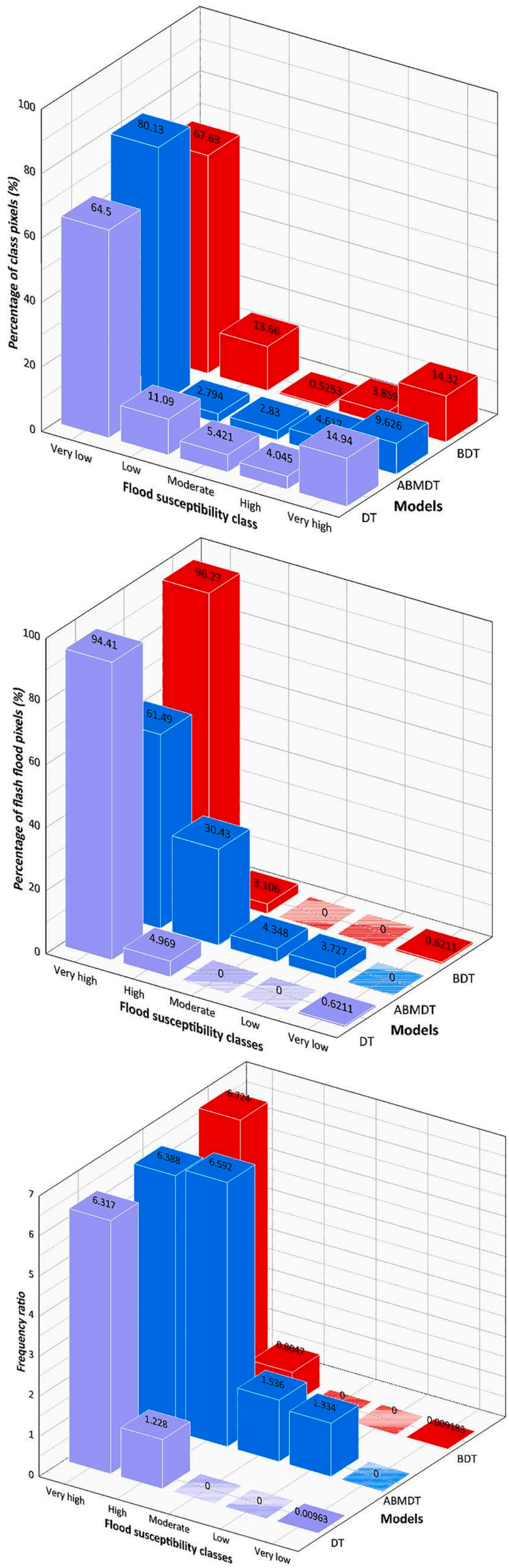

Fig. 11. Analysis of the performance of flash flood susceptibility maps using different models.

and 14.94% are located in the low, moderate, high and very high susceptibility areas, respectively (Fig. 11). It is observed that high and very high susceptibility zones are located along rivers in floodplain areas. It can be observed that the maps produced by the BDT model show similar results. Therefore, it can be reasonably stated that the flood susceptibility map produced by the BDT model is the most reliable compared with other models; thus, it was used for developing flood risk map in this study.

4.2. Flood consequences assessment map

In this study, we used Super Decisions software to calculate AHP algorithms (Whitaker and Adams, 2005) and the Weighted Sum tool in ArcGIS software to integrate the weighted indicator layers into a GIS framework. The flood consequences on human health and financial impact were analyzed as described in the conceptual framework AHP works by setting priorities for multiple criteria evaluated by experts to derive the best decision (Saaty, 1990). The weights of the criteria in the AHP method are based on subjective assessment of some experts (Godfrey et al., 2015; Kienberger et al., 2009; Kokangül et al., 2017; Moghadas et al., 2019) or based on the author experience assessment (Dewan, 2013; Kandilioti and Makropoulos, 2012; Li et al., 2013). In the present study, we used the author experience-based evaluation. We also refer to the related studies using AHP to evaluate flood risk criteria (Dewan, 2013; Godfrey et al., 2015; Kienberger et al., 2009; Kokangül et al., 2017; Moghadas et al., 2019; Velasquez and Hester, 2013).

The consequences on human health are assessed by the combination of population density, poverty rate, road density and number of medical staff (Fig. 12(a)). The consequences on financial impact are evaluated by considering three indicators of industrial facility density, service and commercial density and land-use categories (Fig. 12(b)). The weights of criteria and sub-criteria are presented in Table 6. We used the integrated AHP-WLC to combine indicators in the GIS environment. The *Weighted Sum tool* in ArcGIS software was used to overlay the weighted indicators weighted in Table 6. The final flood consequence map is displayed in Fig. 12(c).

4.3. Flood risk assessment map

Flood risk was assessed in this study by combining the hybrid machine learning BDTmodel-based flood susceptibility map and the flood consequences map. The flood risk assessment map was created by overlaying the flood susceptibility map and flood consequences map using *Weighted Sum* technique in GIS environment (Fig. 13). The risk scores were normalized to a range of 0–1. Out of the 1,056,609 ha total area of the province, 915,264 ha (86.62%) is in very low risk areas, 54,132 ha (5.12%) in low risk areas, 48,989 ha (4.64%) in medium risk areas, 29,308 ha (2.77%) in high risk areas and 8,916 ha (0.84%) in extremely high risk areas.

5. Discussion

Flood is one of the most destructive natural hazards in Vietnam, requiring systematic study to prevent loss of life and destruction of property for sustainable development. Therefore, developing flood risk maps is a necessary step to help decision-makers in flood risk management (Kourgialas and Karatzas, 2011). Even though various methods and approaches have been proposed and applied to assess flood risk; however, the complex geo-morphological conditions and dynamic human activities which cause changes in the morphology and destruction of forests pose difficulties in predicting flood risks accurately (Zahar et al., 2008). For this reason, various algorithms were developed and applied for the flood risk assessment and development of accurate predictive models. In this study, two advanced hybrid models (BDT and ABMDT) and single DT model were validated and compared to select the best model and map which was integrated with the MCDA (AHP) to

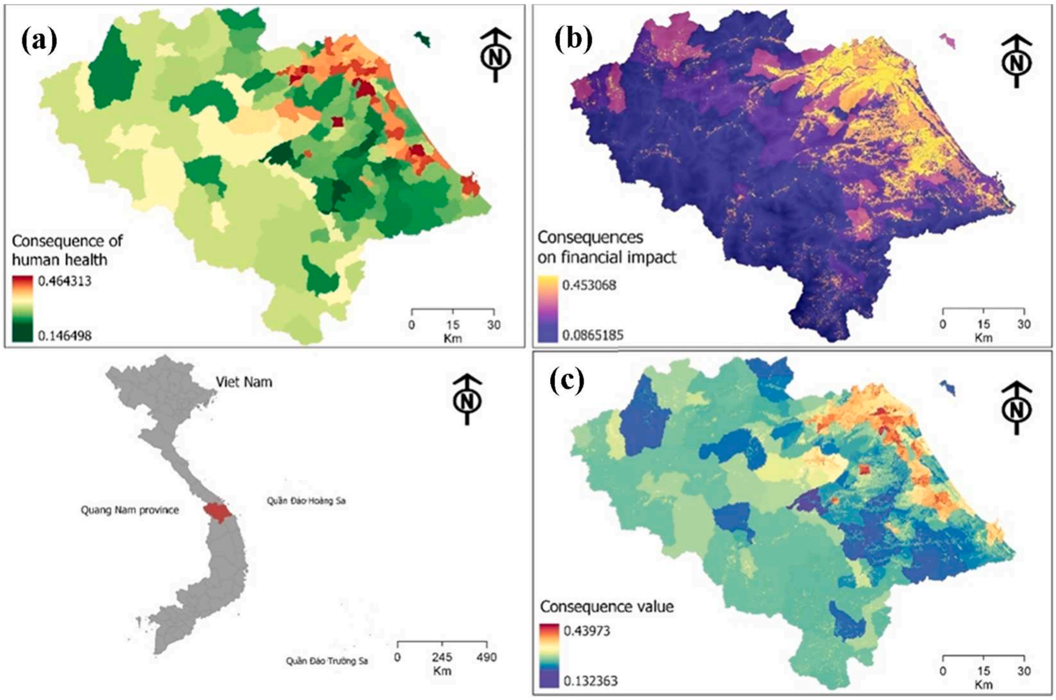

Fig. 12. Flood consequence maps used in this study: (a) consequences on human health, (b) consequences on financial impact, and (c) final flood consequence map.

Table 6
Weights of the indicators using AHP model for flood consequences assessment.

Component	Criteria	Weight	Sub-criteria	Weight
Facilities and Health	Density Population	0.50803	<50	0.05586
			50–200	0.12324
			200–500	0.21898
			>500	0.60192
	Provety Rate	0.24489	<10%	0.07545
			10–20%	0.11236
			20–40%	0.24447
			>40%	0.56772
	Healthcare facilities	0.09259	<1	0.56932
			1–3	0.22404
			3–5	0.13298
			>5	0.07366
	Road density	0.15449	<500	0.50677
			500–1000	0.26412
			1000–2000	0.14279
			>2000	0.08632
Financial impacts	Landuse	0.5	Agricultural land	0.28708
			Build-up	0.49997
			Forest and Vegetation	0.09098
			Water bodies	0.12197
	Industrial facility density	0.25	<5	0.07780
			5–10	0.12479
			10–20	0.30557
			>20	0.49184
	Service facility density	0.25	<50	0.08632
			50–100	0.14279
			100–200	0.26412
			>200	0.50677

develop flood risk map of the Quang Nam province, Vietnam.

Validation results show that all developed models (BDT, ABMDT, and DT) performed well for flood susceptibility assessing and modelling. It is reasonable as the DT model is an effective popular machine learning algorithm to solve the classification problem using transforming the data on the table. Its advantages are: (i) the DT model requires less effort for the data preprocessing process and (ii) it does not require data standardization which is very easy to understand and intuitive. The DT algorithm is sometimes more complicated compared to the other algorithms, and takes time to form the prediction model. In this study, the results also show that the BDT model is the best compared to other models for flood susceptibility mapping in the study area. This is appropriate because the BDT model used Bagging ensemble which is an excellent ensemble tool to improve the accuracy of individual classification prediction and the Radial Basis Function kernel used in Bagging ensemble can enhance the stability of the DT model. Simultaneously, the Bootstrap sampling technique used in the Bagging model was used to reduce the sensitivity of single classification to noise in training data sets (Pham and Prakash, 2019). The model variance is reduced, which can improve the accuracy of the individual model like DT (Breiman, 1996). The advantage of the Bagging model is presented in the change from the basic classification generalization error to the generalization error calculated on the smaller training data. Therefore, it can be used for the classification, which has the learning curves of reduced (Thai Pham, 2019). Thus, the BDT is observed as the best model for flood susceptibility modelling in this study compared with other models (ABMDT and DT). However, it is not suitable for linear classifications because of its classifications. In addition, it cannot be adapted for small or very large training data (Pham and Prakash, 2019; Skurichina and Duin, 2002). This finding of this study is also in comparison with other published works (Arabameri et al., 2020; Chen, 2019; Yariyan, 2020).

In this study, flood susceptibility map was combined with the flood consequences assessment map using AHP and WLC techniques to develop a flood risk map. Whereas, the flood susceptibility map was generated by the best hybrid ML model (BDT). Flood consequences reflect the impacts of human health and financial impacts (Winsemius et al., 2013). We devised criteria for human health impact, including population density, road density, poverty rate and healthcare facility; and criteria for financial impact including industrial facility density, service and commercial facility density, and land-use categories. In literature, some studies have additionally focused on flood risk assessment in terms of combining flood vulnerability and flood hazards such as Scheuer et al. (2011), Masuya (2014) and Lee et al. (2013). Several studies analyzed flood risk in the combination of hazard, exposure and vulnerability (Foudi et al., 2015; Winsemius et al., 2013). Some recent studies focused on the consequences of flood risk since the consequences

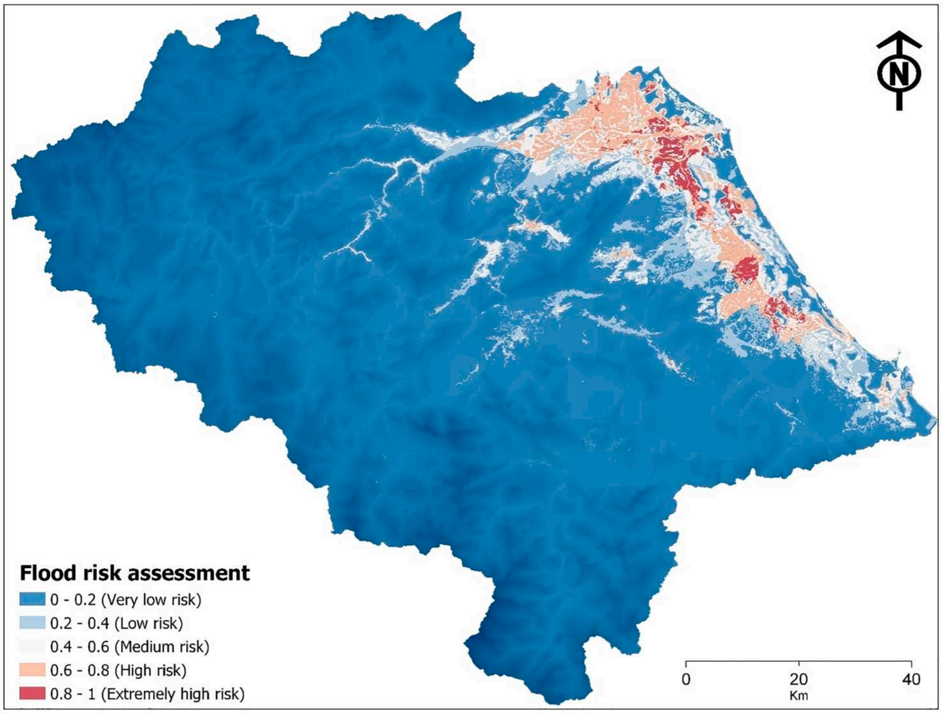

Fig. 13. Final spatial flood risk assessment map of the study area.

reflect the potential impacts of flood hazard (Albano et al., 2014; Jato-Espino et al., 2019). In this study, we assessed flood risk by combining flood susceptibility and flood consequences, which is a suitable approach for flood risk assessment in Quang Nam province, Vietnam. Finally, MCDA is used to incorporate the flood susceptibility and consequences components. In flood risk modelling, MCDA methods are often used in the risk analysis due to several advantages, such as criteria and sub-criteria systematization (Ishizaka and Labib, 2009), suitability for GIS integration (Malczewski, 2006), and consistency in judgment (Koczkodaj, 2017).

The flood risk assessment map shows high flood risk zones/locations which are required to be managed on a priority basis. In general, the integration of a hybrid intelligence model for flood susceptibility map and MCDA for flood consequences map is a good approach for developing a reliable flood risk map in data-scarce areas also. In addition, when we compared the flood risk assessment map of this study with the Quang Nam inundation map of Chau et al. (2013) and Luu et al. (2018), it is observed that the high flood risk areas are located close to and along rivers that are in and around floodplain areas. The result is compatible with previous studies such as Ho and Umitsu (2011), Chau et al. (2013) and Luu et al. (2018). The flooding occurs in the low lying areas of Vu Gia-Thu Bon river basin and affects districts of Dai Loc, Hoi An, Dien Ban, Duy Xuyen, Nong Son, Que Son, and Tam Ky.

6. Concluding remarks

In this study, we developed a soft computing based flook risk framework in which the flood risk map was generated by combing a flood susceptibility map and a flood consequences map. The flood susceptibility map was created using a hybrid AI model (BDT), and the flood consequences map was created by using AHP technique. The case study was Quang Nam province, Vietnam.

Validation results show that all the proposed models performed well for constructing flood susceptibility map. However, the performance of the BDT model (AUC 0.96) is the best in comparison to other models, namely ABMDT (AUC:0.953) and DT (AUC:0.929). Thus, the flood susceptibility map produced by BDT is the best map which was combined with a consequences map to develop a reliable flood risk map of the study area.

In general, the AI or ML integrated with MCDA approach is a suitable framework and approach for developing a reliable flood risk map that can be used and applied in other flood-prone areas considering local geo-environmental conditions. Advantage of this approach is that it does not require time series meteorological and streamflow data, or updated river cross-section data, which are not readily available in many data-scarce areas. However, this study has two main limitations. First, we can not analyze the frequency of flood events in the flood susceptibility mapping since AI or ML models use flood and non-flood locations as a dependent variable and flood influencing factors as predictors. Second, the proposed integrated approach needs to be validated in other flood-prone areas of Vietnam.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 105.08-2019.03.

References

Albano, R., Sole, A., Adamowski, J., Mancusi, L., 2014. A GIS-based model to estimate flood consequences and the degree of accessibility and operability of strategic emergency response structures in urban areas. Nat. Hazards Earth Syst. Sci. 14 (11).
Arabameri, A. et al., 2020. Flash flood susceptibility modelling using functional tree and hybrid ensemble techniques. 125007.
Bafitlhile, T.M., Li, Z., 2019. Applicability of ε-support vector machine and artificial neural network for flood forecasting in humid, semi-humid and semi-arid basins in China. Water 11 (1), 85.
Bennett, N.D., et al., 2013. Characterising performance of environmental models. Environ. Modell. Software 40, 1–20.

Bouwer, L.M., Bubeck, P., Aerts, J.C.J.H., 2010. Changes in future flood risk due to climate and development in a Dutch polder area. Global Environ. Change 20 (3), 463–471. https://doi.org/10.1016/j.gloenvcha.2010.04.002.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140.

Brown, J.D., Damery, S.L., 2002. Managing flood risk in the UK: Towards an integration of social and technical perspectives. Trans. Inst. British Geograph. 27 (4), 412–426.

Bubeck, P., Botzen, W.J.W., Suu, L.T.T., Aerts, J.C.J.H., 2012. Do flood risk perceptions provide useful insights for flood risk management? Findings from central Vietnam. J. Flood Risk Manage. 5 (4), 295–302. https://doi.org/10.1111/j.1753-318X.2012.01151.x.

Bui, D.T., et al., 2019. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. Catena 179, 184–196.

Bui, Q.-T. et al., 2019b. Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. Journal of Hydrology: 124379.

Bui, Q.-T., Nguyen, Q.-H., Pham, V.-M., 2019c. Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2019.124379.

Bui, Q.-T., Pham, V.-M., 2018. Hybrid model to optimize object-based land cover classification by meta-heuristic algorithm: An example for supporting urban management in Ha Noi, Viet Nam. Int. J. Digital Earth. https://doi.org/10.1080/17538947.2018.1542039.

Chau, V.N., Holland, J., Cassells, S., Tuohy, M., 2013. Using GIS to map impacts upon agriculture from extreme floods in Vietnam. Appl. Geogr. 41, 65–74. https://doi.org/10.1016/j.apgeog.2013.03.014.

Chen, W. et al., 2019. Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles. 575: 864-873.

Choubin, B., et al., 2019. An Ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. Sci. Total Environ. 651, 2087–2096.

Costache, R., 2019. Flood susceptibility assessment by using bivariate statistics and machine learning models-a useful tool for flood risk management. Water Resour. Manage. 1–18.

de Moel, H., et al., 2015. Flood risk assessments at different spatial scales. Mitig. Adapt. Strat. Glob. Change 20 (6), 865–890.

Degiorgis, M. et al., 2012a. Classifiers for the detection of flood-prone areas using remote sensed elevation data. Journal of hydrology, 470: 302-315.

Degiorgis, M. et al., 2013. Flood hazard assessment via threshold binary classifiers: case study of the Tanaro river basin. 62(S2): 1-10.

Degiorgis, M. et al., 2012b. Classifiers for the detection of flood-prone areas using remote sensed elevation data. 470: 302-315.

Dewan, A., 2013. Floods in a Megacity: Geospatial Techniques in Assessing Hazards, Risk and Vulnerability. Springer, Netherlands. DOI:https://doi.org/10.1007/978-94-007-5875-9_6.

Dietterich, T.G., 1998. Approximate statistical tests for comparing supervised classification learning algorithms. Neural Comput. 10 (7), 1895–1923.

Falah, F., et al., 2019. Artificial neural networks for flood susceptibility mapping in data-scarce urban areas, spatial modeling in GIS and R for earth and environmental sciences. Elsevier 323–336.

Fan, H., Wang, H., 2011. Preditcing protein subcellular location by AdaBoost.M1 algorithm. DOI:10.1109/AIMSEC.2011.6011378.

Foudi, S., Osés-Eraso, N., Tamayo, I., 2015. Integrated spatial flood risk assessment: The case of Zaragoza. Land Use Policy 42, 278–292.

Gain, A.K., Hoque, M., 2013. Flood risk assessment and its application in the eastern part of D haka C ity, B angladesh. J. Flood Risk Manage. 6 (3), 219–228.

Galar, M., Fernandez, A., Barrenechea, E., Bustince, H., Herrera, F., 2011. A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches. IEEE Trans. Syst. Man Cybernet., Part C (Applications and Reviews) 42 (4), 463–484.

Garrote, L., Bras, R.L., 1995. A distributed model for real-time flood forecasting using digital elevation models. J. Hydrol. 167 (1–4), 279–306.

Godfrey, A., Ciurean, R.L., van Westen, C.J., Kingma, N.C., Glade, T., 2015. Assessing vulnerability of buildings to hydro-meteorological hazards using an expert based approach – An application in Nehoiu Valley, Romania. Int. J. Disaster Risk Reduct. 13, 229–241. https://doi.org/10.1016/j.ijdrr.2015.06.001.

Gorsevski, P.V., Gessler, P.E., Foltz, R.B., Elliot, W.J., 2006. Spatial prediction of landslide hazard using logistic regression and ROC analysis. Trans. GIS 10 (3), 395–415.

Grothmann, T., Reusswig, F., 2006. People at risk of flooding: Why some residents take precautionary action while others do not. Nat. Hazards 38 (1–2), 101–120.

Hajat, S., et al., 2005. The human health consequences of flooding in Europe: A review, Extreme weather events and public health responses. Springer 185–196.

Ho, L.T.K., Umitsu, M., 2011. Micro-landform classification and flood hazard assessment of the Thu Bon alluvial plain, central Vietnam via an integrated method utilizing remotely sensed data. Appl. Geogr. 31 (3), 1082–1093. https://doi.org/10.1016/j.apgeog.2011.01.005.

Hoa, P., et al., 2019. Soil salinity mapping using SAR Sentinel-1 Data and Advanced Machine Learning Algorithms: A case study at ben tre province of the mekong river Delta (Vietnam). Remote Sens. https://doi.org/10.3390/rs11020128.

Hong, H., et al., 2018. Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China). Catena 163, 399–413.

Hosseini, F.S., et al., 2019. Flash-flood hazard assessment using ensembles and Bayesian-based machine learning models: application of the simulated annealing feature selection method. In: Science of the total environment, p. 135161.

Hosseini, F.S., et al., 2020. Flash-flood hazard assessment using ensembles and Bayesian-based machine learning models: application of the simulated annealing feature selection method. Sci. Total Environ. 711.

Huang, I.B., Keisler, J., Linkov, I., 2011. Multi-criteria decision analysis in environmental sciences: Ten years of applications and trends. Sci. Total Environ. 409 (19), 3578–3594.

Ishizaka, A., Labib, A., 2009. Analytic hierarchy process and expert choice: Benefits and limitations. ORI 22 (4), 201–220. https://doi.org/10.1057/ori.2009.10.

Jaafari, A., Zenner, E., Panahi, M., Shahabi, H., 2018. Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability. Agric. For. Meteorol. 198–207. https://doi.org/10.1016/j.agrformet.2018.12.015.

Janizadeh, S., et al., 2019. Prediction success of machine learning methods for flash flood susceptibility mapping in the tafresh watershed. Iran. Sustainability 11 (19), 5426.

Jato-Espino, D., Lobo, A., Ascorbe-Salcedo, A., 2019. Urban flood risk mapping using an optimised additive weighting methodology based on open data. J. Flood Risk Manage. 12.

Jongman, B., Ward, P.J., Aerts, J.C., 2012. Global exposure to river and coastal flooding: Long term trends and changes. Global Environ. Change 22 (4), 823–835.

Kandilioti, G., Makropoulos, C., 2012. Preliminary flood risk assessment: The case of Athens. Nat. Hazards 61 (2), 441–468. https://doi.org/10.1007/s11069-011-9930-5.

Kappes, M.S., Papathoma-Koehle, M., Keiler, M., 2012. Assessing physical vulnerability for multi-hazards using an indicator-based methodology. Appl. Geogr. 32 (2), 577–590.

Khosravi, K., et al., 2018. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. Sci. Total Environ. 627, 744–755.

Khosravi, K., Pourghasemi, H.R., Chapi, K., Bahri, M., 2016. Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: A comparison between Shannon's entropy, statistical index, and weighting factor models. Environ. Monit. Assess. 188 (12), 656.

Kienberger, S., Lang, S., Zeil, P., 2009. Spatial vulnerability units – Expert-based spatial modelling of socio-economic vulnerability in the Salzach catchment. Austria. Nat. Hazards Earth Syst. Sci. 9 (3), 767–778. https://doi.org/10.5194/nhess-9-767-2009.

Koczkodaj, W.W., et al., 2017. On normalization of inconsistency indicators in pairwise comparisons. Int. J. Approximate Reason. 86, 73–79.

Kohavi, R., 1997. The Power of Decision Tables. Proceedings of European Conference on Machine Learning. DOI:10.1007/3-540-59286-5_57.

Kokangül, A., Polat, U., Dağsuyu, C., 2017. A new approximation for risk assessment using the AHP and Fine Kinney methodologies. Saf. Sci. 91, 24–32. https://doi.org/10.1016/j.ssci.2016.07.015.

Kourgialas, N., Karatzas, G., 2011. Flood management and a GIS modelling method to assess flood-hazard areas-a case study. Hydrol. Sci. J.-J. Des Sci. Hydrol. 56, 212–225. https://doi.org/10.1080/02626667.2011.555836.

Lee, G., Jun, K.-S., Chung, E.-S., 2013. Integrated multi-criteria flood vulnerability approach using fuzzy TOPSIS and Delphi technique. Nat. Hazards Earth Syst. Sci. 13 (5), 1293.

Lekkas, D., Onof, C., Lee, M., Baltas, E., 2004. Application of artificial neural networks for flood forecasting. Global Nest J. 6 (3), 205–211.

Li, G.-F., Xiang, X.-Y., Tong, Y.-Y., Wang, H.-M., 2013. Impact assessment of urbanization on flood risk in the Yangtze River Delta. Stoch. Env. Res. Risk Assess. 27 (7), 1683–1693. https://doi.org/10.1007/s00477-013-0706-1.

Luu, C., von Meding, J., 2018. A flood risk assessment of quang nam, vietnam using spatial multicriteria decision analysis. Water 10 (4), 461.

Luu, C., Von Meding, J., Kanjanabootra, 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. Natural Hazards 90 (3), 1031–1050. https://doi.org/10.1007/s11069-017-3083-0.

Makropoulos, C., Butler, D., 2006. Spatial ordered weighted averaging: Incorporating spatially variable attitude towards risk in spatial multi-criteria decision-making. Environ. Modell. Software 21 (1), 69–84.

Malczewski, J., 2000. On the use of weighted linear combination method in GIS: Common and best practice approaches. Trans. GIS 4 (1), 5–22.

Malczewski, J., 2006. GIS-based multicriteria decision analysis: A survey of the literature. Int. J. Geograph. Inf. Sci. 20 (7), 703–726.

Manfreda, S., Samela, C.J.J.o.F.R.M., 2019. A digital elevation model based method for a rapid estimation of flood inundation depth. 12: e12541.

Masuya, A., 2014. Flood vulnerability and risk assessment with spatial multi-criteria evaluation. Dhaka Megacity. Springer 177–202.

Masuya, A., Dewan, A., Corner, R.J., 2015. Population evacuation: evaluating spatial distribution of flood shelters and vulnerable residential units in Dhaka with geographic information systems. Nat. Hazards 78 (3), 1859–1882.

Merz, B. et al., 2014. Floods and climate: emerging perspectives for flood risk assessment and management.

Merz, B., Hall, J., Disse, M., Schumann, A., 2010. Fluvial flood risk management in a changing world. Natural Hazards and Earth System Sciences.

Meyfroidt, P., Vu, T.P., Hoang, V.A., 2013. Trajectories of deforestation, coffee expansion and displacement of shifting cultivation in the Central Highlands of Vietnam. Global Environ. Change 23 (5), 1187–1198. https://doi.org/10.1016/j.gloenvcha.2013.04.005.

Moghadas, M., Asadzadeh, A., Vafeidis, A., Fekete, A., Kötter, T., 2019. A multi-criteria approach for assessing urban flood resilience in Tehran. Iran. Int. J. Disaster Risk Reduct. 35 https://doi.org/10.1016/j.ijdrr.2019.101069.
Mohammady, M., Pourghasemi, H.R., Pradhan, B., 2012. Landslide susceptibility mapping at Golestan Province, Iran: A comparison between frequency ratio, Dempster-Shafer, and weights-of-evidence models. J. Asian Earth Sci. 61, 221–236.
Naghibi, S.A., Vafakhah, M., Hashemi, H., Pradhan, B., Alavi, S.J., 2019. Water resources management through flood spreading project suitability mapping using frequency ratio, k-nearest Neighbours, and Random Forest Algorithms. Nat. Resour. Res. 1–19.
Nandi, A., Mandal, A., Wilson, M., Smith, D., 2016. Flood hazard mapping in Jamaica using principal component analysis and logistic regression. Environmental Earth Sciences, 75. DOI:10.1007/s12665-016-5323-0.
Nardi, F., Vivoni, E.R., Grimaldi, S.J.W.R.R., 2006. Investigating a floodplain scaling relation using a hydrogeomorphic delineation method. 42(9).
Nguyen, H., Ardillier-Carras, F., Touchart, L., 2018. Les paysages de rizières et leur évolution récente dans le delta du fleuve Gianh. Cybergeo. https://doi.org/10.4000/cybergeo.29826.
Nicholls, R.J., Hoozemans, F.M., Marchand, M., 1999. Increasing flood risk and wetland losses due to global sea-level rise: Regional and global analyses. Global Environ. Change 9, S69–S87.
Ozturk, D., Batuk, F., 2011. Implementation of GIS-based multicriteria decision analysis with VB in ArcGIS. Int. J. Inform. Technol. Decision Making 10 (06), 1023–1042.
Papathoma-Köhle, M., 2016. Vulnerability curves vs. vulnerability indicators: Application of an indicator-based methodology for debris-flow hazards. Nat. Hazards Earth Syst. Sci. 16 (8).
Pham, B.T., Bui, D.T., Dholakia, M., Prakash, I., Pham, H.V., 2016. 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. 34 (6), 1807–1824.
Pham, B.T., Bui, D.T., Prakash, I., 2018. Bagging based Support Vector Machines for spatial prediction of landslides. Environ. Earth Sci. 77 (4), 146.
Pham, B.T., Jaafari, A., Prakash, I., Bui, D.T., 2019a. A novel hybrid intelligent model of support vector machines and the MultiBoost ensemble for landslide susceptibility modeling. Bull. Eng. Geol. Environ. 78 (4), 2865–2886.
Pham, B.T., Prakash, I., 2019. A novel hybrid model of bagging-based naïve bayes trees for landslide susceptibility assessment. Bull. Eng. Geol. Environ. 78 (3), 1911–1925.
Pham, B.T., et al., 2019b. Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learning approaches. CATENA 175, 203–218.
Pham, Q.B., et al., 2019c. Potential of hybrid data-intelligence algorithms for multi-station modelling of rainfall. Water Resour. Manage. 33 (15), 5067–5087. https://doi.org/10.1007/s11269-019-02408-3.
Pradhan, B., 2009. Flood susceptible mapping and risk area delineation using logistic regression, GIS and remote sensing. J. Spatial Hydrol. 9, 1–18.
Rahmati, O., Pourghasemi, H.R., Zeinivand, H.J.G.I., 2016. Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province. Iran. 31 (1), 42–70.
Rang, M.S., Kang, M.G., Park, S.W., Lee, J.J., Yoo, R.H., 2006. Application of grey model and artificial neural networks to flood forecasting 1. JAWRA J. Am. Water Resourc. Assoc. 42 (2), 473–486.
Ronco, P., et al., 2014. The KULTURisk Regional Risk Assessment methodology for water-related natural hazards-Part 1: Physical-environmental assessment. Hydrol. Earth Syst. Sci 18 (12), 5399–5414.
Saaty, T.L., 1990. How to make a decision: The analytic hierarchy process. Eur. J. Oper. Res. 48 (1), 9–26. https://doi.org/10.1016/0377-2217(90)90057-I.
Saaty, T.L., 2003. Decision-making with the AHP: Why is the principal eigenvector necessary. Eur. J. Oper. Res. 145 (1), 85–91.
Saaty, T.L., Vargas, L.G., 2006. Decision making with the analytic network process. Springer.
Saha, S., Saha, A., Hembram, T.K., Pradhan, B., Alamri, A.M., 2020. Evaluating the performance of individual and novel ensemble of machine learning and statistical models for landslide susceptibility assessment at rudraprayag district of garhwal himalaya. Appl. Sci. 10 (11), 3772.
Sayers, P., Hall, J., Meadowcroft, I., 2002. Towards risk-based flood hazard management in the UK, Proceedings of the institution of civil engineers-civil engineering. Thomas Telford Ltd, pp. 36-42.
Schanze, J., 2006. Flood risk management–a basic framework, Flood risk management: hazards, vulnerability and mitigation measures. Springer 1–20.
Scheuer, S., Haase, D., Meyer, V., 2011. Exploring multicriteria flood vulnerability by integrating economic, social and ecological dimensions of flood risk and coping capacity: From a starting point view towards an end point view of vulnerability. Nat. Hazards 58 (2), 731–751.
Shrestha, B.B., et al., 2014. Fundamental analysis for flood risk management in the selected river basins of Southeast Asia. J. Disaster Res. 9 (5), 858–869.
Shrifi Garmdareh, E., Vafakhah, M., Eslamian, S., 2019. Assessment the performance of support vector machine and artificial neural network systems for regional flood frequency analysis (A Case Study: Namak Lake Watershed). JWSS-Isfahan Univ. Technol. 23 (1), 351–366.
Siahkamari, S., Haghizadeh, A., Zeinivand, H., Tahmasebipour, N., Rahmati, O.J.G.i., 2018. Spatial prediction of flood-susceptible areas using frequency ratio and maximum entropy models. 33(9): 927-941.
Skurichina, M., Duin, R.P., 2002. Bagging, boosting and the random subspace method for linear classifiers. Pattern Anal. Appl. 5 (2), 121–135.
Thai Pham, B., et al., 2019. Landslide susceptibility assessment by novel hybrid machine learning algorithms. Sustainability 11 (16), 4386.
Thieken, A.H., Mariani, S., Longfield, S., Vanneuville, W., 2014. Preface: Flood resilient communities–managing the consequences of flooding. Nat. Hazards Earth Syst. Sci 14, 33–39.
Tien Bui, D., et al., 2018. Novel hybrid evolutionary algorithms for spatial prediction of floods. Sci. Rep. 8 https://doi.org/10.1038/s41598-018-33755-7.
Tien Bui, D., et al., 2016. Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibility modelling in a high-frequency tropical cyclone area using GIS. J. Hydrol. 540 https://doi.org/10.1016/j.jhydrol.2016.06.027 (Accepted, in press).
Toosi, A.S., Calbimonte, G.H., Nouri, H., Alaghmand, S., 2019. River basin-scale flood hazard assessment using a modified multi-criteria decision analysis approach: A case study. J. Hydrol. 574, 660–671.
Tran, P., Shaw, R., Chantry, G., Norton, J., 2009. GIS and local knowledge in disaster management: A case study of flood risk mapping in Viet Nam. Disasters 33 (1), 152–169.
Urbanowicz, R.J., Meeker, M., La Cava, W., Olson, R.S., Moore, J.H., 2018. Relief-based feature selection: Introduction and review. J. Biomed. Inform. 85, 189–203.
Van Aalst, M., Burton, I., 2002. The Last Straw: Integrating Natural Disaster Mitigation with Environmental Management:(with Examples from Dominica, the Dominican Republic and St. Disaster Management Facility, World Bank, Lucia).
Van Dao, D., et al., 2020. A spatially explicit deep learning neural network model for the prediction of landslide susceptibility. Catena 188.
Velasquez, M., Hester, P.T., 2013. An analysis of multi-criteria decision making methods. Int. J. Oper. Res. 10 (2), 56–66.
Victoriano, J.M., Santos, M.L.C.D., Vinluan, A.A., Carpio, J.T., 2019. Predicting pollution level using random forest: A case study of marilao river in Bulacan Province, Philippines. Int. J. Comput. Sci. Res. 3 (1), 151–162.
Wang, Q., Li, W., Wu, Y., Pei, Y., Xie, P., 2016. Application of statistical index and index of entropy methods to landslide susceptibility assessment in Gongliu (Xinjiang, China). Environ. Earth Sci. 75 (7), 599.
Wang, Y., et al., 2019. A hybrid GIS multi-criteria decision-making method for flood susceptibility mapping at Shangyou. China. Remote Sens. 11 (1), 62.
Whitaker, R., Adams, W., 2005. Developers of Superdecisions software. Decisions Foundation, Pittsburgh.
Winsemius, H., Van Beek, L., Jongman, B., Ward, P., Bouwman, A., 2013. A framework for global river flood risk assessments. Hydrol. Earth Syst. Sci. 17 (5), 1871–1892.
Woodruff, J.M., 2005. Consequence and likelihood in risk estimation: A matter of balance in UK health and safety risk assessment practice. Saf. Sci. 43 (5), 345–353. https://doi.org/10.1016/j.ssci.2005.07.003.
Xiong, J., Li, J., Cheng, W., Wang, N., Guo, L., 2019. A gis-based support vector machine model for flash flood vulnerability assessment and mapping in China. ISPRS Int. J. Geo-Inf. 8 (7), 297.
Yang, Q., Shao, J., Scholz, M., Plant, C., 2011. Feature selection methods for characterizing and classifying adaptive Sustainable Flood Retention Basins. Water Res. 45 (3), 993–1004.
Yariyan, P. et al., 2020a. Improvement of best first decision trees using bagging and dagging ensembles for flood probability mapping. Water Resources Management, 34 (9): 3037-3053.
Yariyan, P. et al., 2020b. Improvement of best first decision trees using bagging and dagging ensembles for flood probability mapping. 34(9): 3037-3053.
Zahar, Y., Ghorbel, A., Albergel, J.J.J.o.H., 2008. Impacts of large dams on downstream flow conditions of rivers: aggradation and reduction of the Medjerda channel capacity downstream of the Sidi Salem dam (Tunisia). 351(3-4): 318-330.