Flash-flood hazard assessment using Ensembles and Bayesian-based machine learning models: application of the simulated annealing feature selection method

Farzaneh Sajedi Hosseini [a], Bahram Choubin [b], Amir Mosavi [c,d], Narjes Nabipour [e,*], Shahaboddin Shamshirband [f,g,*], Hamid Darabi [h], Ali Torabi Haghighi [h]

[a] Department of Reclamation of Arid and Mountainous Regions, Faculty of Natural Resources, University of Tehran, Karaj, Iran

[b] Soil Conservation and Watershed Management Research Department, West Azarbaijan Agricultural and Natural Resources Research and Education Center, AREEO, Urmia, Iran

[c] School of the Built Environment, Oxford Brookes University, Oxford OX3 0BP, UK

[d] Kalman Kando Faculty of Electrical Engineering, Obuda University, Budapest, Hungary

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

[f] Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam

[g] Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam

[h] Water, Energy and Environmental Engineering Research Unit, University of Oulu, P.O. Box 4300, FIN-90014 Oulu, Finland

* Corresponding authors, Email: narjesnabipour@duytan.edu.vn; shahaboddin.shamshirband@tdtu.edu.vn

Flash-flood hazard assessment using Ensembles and Bayesian-based machine learning models: application of the simulated annealing feature selection method

Abstract

Flash-floods are increasingly recognized as a frequent natural hazard worldwide. Iran has been among the most devastated regions affected by the major floods. While the temporal flash-flood forecasting models are mainly developed for warning systems, the models for assessing hazardous areas can greatly contribute to adaptation and mitigation policy-making and disaster risk reduction. Former researches in the flash-flood hazard mapping have heightened the urge for the advancement of more accurate models. Thus, the current research proposes the state-of-the-art ensemble models of boosted generalized linear model (GLMBoost) and random forest (RF), and Bayesian generalized linear model (BayesGLM) methods for higher performance modeling. Furthermore, a pre-processing method, namely simulated annealing (SA), is used to eliminate redundant variables from the modeling process. Results of the modeling based on the hit and miss analysis indicates high performance for both models (accuracy= 90–92%, Kappa= 79–84%, Success ratio= 94–96%, Threat score= 80–84%, and Heidke skill score= 79–84%). The variables of distance from the stream, vegetation, drainage density, land use, and elevation have shown more contribution among others for modeling the flash-flood. The results of this study can significantly facilitate mapping the hazardous areas and further assist watershed managers to control and remediate induced damages of flood in the data-scarce regions.

Keywords: Flash-flood; hazard; ensemble machine learning; Bayesian; simulated annealing

1. Introduction

Abnormality in precipitation is rapidly increasing worldwide (Hao et al., 2019; Li et al., 2019; Lyubchich et al., 2019). Besides, climate change is alternating the hydrometeorological patterns in terms of frequency, irregularity, and severity of precipitation which has led to the rise of the life-threatening hydrological disasters (Hennequin et al., 2018; Serago and Vogel, 2018; Shkolnik et al., 2018). On the other hand, the vulnerability to hydrological disasters, e.g., the flood has recently been magnified due to the rapid urbanization and population growth, particularly in the developing countries (Ahmadalipour et al., 2019; Casagrande et al., 2017; Kubwarugira et al., 2019). Thus, for the purpose of mitigation and planning to extreme events, more than ever, there is an urge for the advancement of reliable modeling techniques to accurately identify the hazardous areas (Chiang and Ling, 2017; Frigerio et al., 2018; Henriksen et al., 2018).

Spatial assessment of flood hazard is of utmost importance for the urban and the built environment planning and land use management, infrastructures engineering and design, and the advancement of the mitigation structures to optimally reduce the devastation (Al-Juaidi et al., 2018; Muhamad et al., 2019; Sozer et al., 2018). Advancement of the novel methods and continued progress in improving the methods for hazard susceptibility mapping are especially vital for flash-floods hazard mitigation due to their higher destructive power in a brief period of time compared to the river and coastal floods, for instance (Abuzied et al., 2016; Youssef et al., 2016). The accordance of flash-flood follows a complex interaction of the meteorology with hydrology (Doswell III et al., 1996). Multi-criteria decision-making analysis methods, (e.g., Alves et al., 2018; Kanani-Sadat et al., 2019; Roslee and Norhisham, 2018; Tang et al., 2018; Tiryaki and Karaca, 2018), the statistical methods, e.g., frequency ratio, regression logistics, Shannon's entropy, generalized linear model, statistical index, weights-of-evidence, weighting factor, multivariate discriminant analysis, flexible discriminant analysis, multivariate logistic regression, generalized additive model, and further bivariate and multivariate statistical approaches (Giovannettone et al., 2018; Shafapour Tehrany et al., 2019; Youssef et al., 2016), the fuzzy rule-based systems (Bui et al., 2019b, Sahana and Patel, 2019), time series (Kuenzer et al., 2013, Kwak et al., 2014, Sghaier et al., 2018, Sciance and Nooner, 2018), physical models for rainfall-runoff modeling, (e.g., Hofierka and Knutová, 2015; Zhou et al., 2012; Motevalli and Vafakhah, 2016), and the soft computing and machine learning methods, e.g., artificial neural networks (ANNs), backpropagation ANNs, support vector machines (SVM), least squares SVM (LSSVM), classification and regression trees (CART), random forest (RF), decision trees (DT), Naïve Bayes (NB), adaptive neuro-fuzzy inference system (ANFIS), quick unbiased efficient statistical tree (QUEST), and genetic algorithm rule-set production (GARP) (Hong et al., 2018; Darabi et al., 2019; Chen et al., 2017; Lee et al., 2017; Yan et al., 2018) are among the most popular methods used for flood susceptibility mapping to identify flood-prone areas. A number of recent comparative studies reported promising results using machine learning methods (Khosravi et al., 2018, Khosravi et al., 2019; Shafapour Tehrany et al., 2019, Siahkamari et al., 2018, Tehrany and Kumar, 2018; Chen et al., 2017). Consequently, machine learning has become the key instrument in susceptibility mapping (Chapi et al., 2017; Alfieri et al., 2015; Lindenschmidt et al., 2016). Machine learning methods have shown promising results in dealing with the complexity raised in modeling the flash-flood hazard maps (Mahmood and Rahman 2019; Mahmood et al. 2019) which encompasses the multiple spheres of total environment, e.g., Anthroposphere, Hydrosphere, Atmosphere, and Lithosphere (Bui et al., 2018a; Kanani-Sadat et al., 2019; Ngo et al., 2018). Recently the accuracy of perdition models for flood susceptibility mapping has been dramatically increased using the emerging novel hybrid machine learning models (Bui et al., 2019a, Bui et al., 2018b, Chen et al., 2019,), as well as ensemble models (Al-Abadi, 2018; Bui et al., 2019b, Choubin et al., 2019b; Razavi Termeh et al., 2018). Hybrid models are generally created through combination of the regular machine learning models with the soft computing techniques, multi-criteria decision-making analysis methods, optimization algorithms, and/or other machine learning methods or an integration of multiple of them (Chen et al., 2019; Costache, 2019; Ngo et al., 2018). Ensemble models are often developed using either of the bagging, boosting, or random subspace methodologies to employ more than one learning system to achieve higher performance and accuracy for predictors (Buchen and Wohlrabe, 2011, Bui et al., 2019b).

The future trend in the data-driven models for the flood susceptibility mapping has heightened the need for advancing sophisticated machine learning models (Shafizadeh-Moghadam, et al., 2018; Valavi, et al., 2019; Khosravi et al., 2018; Bui et al., 2019b). Surveys such as that conducted by Mosavi et al. (2018), and various comparative studies, e.g., (Khosravi et al., 2019, Chen et al., 2017) suggest that, the ensemble and Bayesian variations of the machine learning models generally provide higher accuracy and performance compared to their conventional forms. However, the existing literature provides minor knowledge on the performance of various techniques of ensemble models (bagging and boosting) and Bayesian, considering the flash-flood hazard assessment. Although a number of researches have used bagging and boosting for increasing the quality of the prediction, very limited studies exist on the comparative study of bagging, boosting and Bayesian methods. Thus, there is a general lack of research in this regard, and there has been little discussion about potential, differences, and individual characteristics of boosting, bagging and Bayesian in modeling the susceptibility mapping. Consequently, the main objective of this study is to provide a comparative study between the boosting, bagging and Bayesian-based models. The comparative study is conducted between the novel method of Bayesian generalized linear model (BayesGLM), and the ensemble methods of boosted generalized linear model (GLMBoost), and random forest (RF) (bagging-based method). The proposed comparative study employs a promising feature selection (FS) method (Kira and Rendell, 1992, Liu and Motoda, 2007), namely simulated annealing (SA) (Van Laarhoven and Aarts, 1987) to eliminate redundant variables from the modeling process.

Section two presents the material and methods used in this study, which begins by describing the study area and then proceeds with explaining the methodology, modeling process, and validating the results. Section three analyzes the results and discusses them. Finally, in section four, the conclusion gives a summary and areas for further researches.

2. Material and methods

2.1. Study area

Gorganroud River Basin, located in the north of Iran within Golestan Province, extends between latitudes of 36º 25' to 38º 15' N and longitudes of 56º 26' to 54º 10' E. It has an area about 11,290 km^2 which drainages the Eastern Alborz Mountains into the Caspian Sea. The elevation changes between -96 m a.s.l. and 3669 m a.s.l. for western regions (Caspian Sea) towards southern areas (Fig. 1). According to the long-term (1988-2018) weather stations' data in the study area (presented in Fig. 1), the mean annual rainfall is about 500 mm with a mean temperature of approximately 17.8°C. The main climates of the study area are including semi-arid, Mediterranean, semi-humid, and humid.

This watershed is known as one of the most affected regions by floods in Iran (Safaripour et al., 2012), which experienced many floods. From 1991 to 2019, more than 120 large and small floods occurred in this watershed (Jannati, 2019). For example, during the flood that occurred on 11

August 2001, the Gorganroud discharge reached about 3020 m^3, and the width of the river increased from 10 meters (the normal width) to 400 meters. This flood killed more than 500 people, which is considered as the most casualty flood in Iran (Jannati, 2019). Another example is on 17 March 2019, which most of the cities such as Aqqala and Gomishan, at least 70 villages, about 12000 houses, infrastructures, agricultural areas, and gardens were damaged along the Gorganroud river (Donya-e-eqtesad, 2019). However, identifying the hazardous areas in this most extremely flooded area is most important for reducing the damages.

Fig. 1 SOMEWHERE HERE

2.2. Methodology

The procedural approach taken in the present research can be summarized as (i) collection and preparation of the required data for the flash-flood modeling in the study area, (ii) extraction of the flooded locations using Sentinel 2 images, (iii) consideration of the factors affected flash-flood, (iv) feature selection (FS) using simulated annealing (SA) method, (v) machine learning modeling of flash-flood, (vi) validation of the results using hit and miss analysis, (vii) extraction of the hazard areas induced by flash-flood in the study area. These steps are explained in details as follow:

2.2.1. Preparation of flash-flood inventory map

Due to the lack of recorded location of flood occurrences, we extracted the inundation area using the Modified Normalized Difference Water Index (MDNWI) of Sentinel-2 satellite through Google Earth Engine (GEE, 2019a) environment. Many litterateurs have demonstrated that MNDWI is more appropriate to extract water bodies (e.g., Du et al., 2014; Du et al., 2016; Xu, 2006; Li et al., 2013; Singh et al., 2015). Radiometrically calibrated and terrain corrected Sentinel-2 Level-1C dataset is stored within GEE. GEE provides free cloud computing facilities for research

(Clement et al., 2018), however, to remove the effects of cloud, pixels with less than 2% cloud were filtered and used. Also, we used the quality band (QA60) to mask the clouds and cirrus (respectively bits 10 and 11 in the QA60). MDNWI is defined as (Xu, 2006):

$$\text{MNDWI}=\frac{B_{Green}-B_{SWIR}}{B_{Green}+B_{SWIR}} \quad (1)$$

where B_{Green} and B_{SWIR} are respectively the Top-Of-Atmosphere (TOA) reflectance of the green (i.e., Band 3) and Shortwave-Infrared (SWIR; Band 11) bands. The bandwidth for green (SWIR) band, central wavelength, and spatial resolution are respectively equal to 35 (90) nm, 560 (1610) nm, and 10 (20) m (Du et al., 2016). MNDWI varies between -1 to 1, which values greater than zero is considered as water (Du et al., 2016; Clement et al., 2018).

Fig. 2a indicates the inundated area extracted by MNDWI during a period from 11 March 2019 to 10 April 2019, which flash-flood affected large parts of Iran. In this period, the Gorganrood River Basin was most severely faced with flood disasters, and flooding has surrounded about 70 villages. Also, some of the cities, such as Aqqala in this watershed was submerged during this period (for around more than one month) (Fig. 3).

After identifying the inundated area, the number of 368 flash-flood locations were randomly considered from the inundated pixels (Fig. 2b), and their locations were confirmed through many field surveys and reports from Iran's Minister of Energy (IMOE). It should be stated that the existing water bodies were masked from MNDWI, so the considered flash-flood locations were not from the existing water bodies. Also, the equal number of the flood occurrence locations, 368 non-flood occurrence locations were randomly considered from the non-inundated pixels (Fig. 2b). These flood/non-flood points were considered as the dependent variable and used for flash-flood modeling using machine learning (ML) models. Section 2.2.4 (i.e., Flash-flood modeling) provides more details about the flash-flood modeling.

Fig. 2 SOMEWHERE HERE

Fig. 3 SOMEWHERE HERE

2.2.2. Flash-flood influencing factor (FFIF)

According to the data availability in the study area and due to the literature review, number of 15 factors including elevation, slope, aspect, topographic roughness index (TPI), topographic position index (TPI), flow accumulation (FA), topographic wetness index (TWI), drainage density (Dd), distance from stream (Dfs), precipitation, normalized difference vegetation index (NDVI), soil depth, soil type, land use, and lithology (Fig. 4) were considered as predictors to model flash-flood in the Gorganroud River Basin.

The topographical factors, including elevation, slope, aspect, TPI, and TRI, are important factors that affect flood occurrences. An ASTER Digital Elevation Model (DEM) with a cell size of 30 m × 30 m was used to extract the topographic factors. The elevation is among the most essential factors in flood modeling (Fig. 4a), and the probability of flood events in areas with high elevation is almost impossible (Botzen et al., 2012). Water flow moves from high elevations towards low elevations, and therefore, the possibility of flood occurrence is naturally higher in flat regions. The slope layer is another factor which plays a major role in flood event through effects on movement and velocity of runoff, speed of the water (Torabi Haghighi et al., 2018). The slope layer changes from 0 to 433 percent (0 to 77 degrees) in the study area (Fig. 4b). The different aspects (Fig. 4c) have different effects on the flood due to the difference in receiving solar energy and rainfall in each aspect (Mojaddadi et al., 2017). The TRI indicates the roughness of the ground which affects flood movement (Kalantar et al., 2017). The value of the TRI layer (Fig. 4d) in the study area varies from 0 to 73. The TPI (Fig. 4e) shows regions that have high (the positive values) and low (the negative values) elevation than average of their surroundings. The TPI layer in the study area changes from -52 to 53. The TPI and TRI layers were generated by SAGA-GIS software.

The FA layer (Fig. 4f) indicates several accumulated pixels in upstream of a given pixel. So, this can be a good index to show the areas with highly accumulated water (Kia et al., 2012). The TWI layer (Fig. 4g) represents the wetness conditions due to the topography, which is important in surface runoff generation (Nampak et al. 2014; Sajedi-Hosseini et al., 2018; Alilou et al., 2019). It was generated in the SAGA-GIS environment. The drainage density (Fig. 4h) is related to the slope, elevation, and structures of lithology. A lot of floods occur in a high drainage density area due to the large accumulation of water. When the drainage density of an area is high, it demonstrates a high runoff and low infiltration rate and vice versa (Prasad et al, 2008). Naturally, areas close to rivers and streams have more probability of flooding. The Dfs layer was created using the Euclidian distance tool in ArcGIS (Fig. 4i). The precipitation is the motive of a flood, which in this study its map for flooding periods (Fig. 4i) was created using precipitation data of weather stations (Fig. 1) obtained from the Iran Meteorological Organization through Inverse Distance Weighting (IDW) method.

The event of a flood is oppositely related to the density of vegetation (Kia et al., 2012). Hence, NDVI (Fig. 4K), as an index indicating vegetation conditions, was extracted for flooding period (March 2019) using Landsat 8 satellite images through Google Earth Engine (GEE, 2019b). The soil is an effective factor in the generation or infiltration of runoff (Csáfordi et al., 2012). The most dominant soil types of the study area are Mollisol, Alfisol, Inceptisol, Entisol, Aridisol, Rock Outcrop, and Salt flat (Fig. 4m). Soil type (Fig. 4m) and soil depth (Fig. 4l) affects the drainage process because of different characteristics such as penetrability degree, texture, and structure. Land use is another important factor for a flood event. Conditions of runoff vary with different land uses (Wang et al, 2015). Precipitation on the barren land run over the surface quickly compared to the forest land (Kia et al., 2012). In this study, the land use map includes nine dominant classes, namely forest, agriculture, rangeland, dry farming, orchard, bareland, water bodies, woodland, and residential areas (Fig. 4n). The lithology is another factor which can affect flood sensitive areas. The lithology units based on the rock permeability is also required in flood hazard assessment. According to the lithology map, most of the study area located in west is Q_m unit (Quaternary- swamp and marsh). In this study, the Iranian Water Resources Management Company (IWRMC) provided the essential data on the soil, land use, and the lithology maps.

Fig. 4 SOMEWHERE HERE

2.2.3. Feature selection using simulated annealing (SA) method

In this study, at first, we considered FFIFs based on the literature reviews. Due to the lack of a universal guideline for selecting factors in flood hazard assessment studies, the FS methodology was conducted on data to select key variables and avoid the effects of redundant factors on flash-flood modeling. Therefore, the SA method was carried out to select key features. SA is an efficient global optimization method based on a random search technique, widely used to identify the optimum in relatively large design space (Kirkpatrick et al., 1983, Aarts and Korst, 1988). SA particularly suitable in avoiding the local solutions traps through displacing toward the uphill following the probability $p = exp(-\Delta E/T)$, where T is representing the annealing parameter, and the ΔE would be the value of the uphill movement. Uphill movements are regulated for organized progress towards the optimum while avoiding the big movement to maintain the accuracy of the solution. To do so a primary solution is randomly selected. The value of the cost function is accordingly calculated aiming at the minimum value. In every step, the value of the cost function is compared with the value of the neighboring points. The new values must be less than the former values, or the Boltzman's probability (Aarts and Korst, 1988) must be satisfied to be accepted. The workflow will continue to find the global minimum value of the function. Literature includes several studies where SA has been used in hydrological forecasting for optimizing the input feature subset selection and used to reduce the redundant variables from the modeling process (Zhu and Wu, 2013, Huang et al., 2018, Choubin et al., 2020). In this study, the SA method was conducted using a k-fold (10-fold) cross-validation methodology by training data set (70% of the data) within the R environment through the Caret package (Kuhn, 2015).

2.2.4. Flash-flood modeling

After preparing the predictand (flood/non-flood locations) and predictors (FFIF selected by SA) variables respectively as output and inputs data, flash-flood modeling was conducted using machine learning (ML) models. In this study, the flooded and non-flooded points (Fig. 2b) are converted into a binary scale (or presence-absence). So, the values of 0 and 1 were assigned into the non-flood and flood occurrence locations respectively. Then, the corresponding values of the predictors in the location of the non-flood and flood points were extracted. From the whole datasets, 70% of the data (including 258 flood and 258 non-flood occurrence locations) was considered for training objectives and rest 30% of the data (including 110 flood and 110 non-flood occurrence locations) was used to evaluate data. A k-fold cross-validation methodology (k=10) was used to train the ML models.

2.2.4.1. Boosted Generalized Linear Model (GLMBoost)

For fitting generalized linear model (GLM) (Lee and Nelder, 2006), GLMBoost generally uses a functional gradient descent method for optimizing the overall loss functions through implementing component-wise least squares while the variable selection can be carried out simultaneously (Bühlmann, 2006). GLMBoost (Hothorn and Bühlmann, 2006) represents an ensemble form of the GLM (McCullagh, 1984; McCullagh 2019), which transforms this ordinary linear regression into a model, suitable for high-dimensional data sets (Bühlmann and Yu, 2003; Bühlmann, 2006). Ensemble modeling is the strategy of simultaneously using a number of classifiers to improve the accuracy and quality of prediction models (Dietterich, 2000; Zhang and Ma, 2012; Buchen and Wohlrabe, 2011). One of the efficient ways of developing ensembles is the boosting technique which employs a gradient descent algorithm in function space (Breiman, 2004; Bühlmann and Hothorn, 2007; Freund and Schapire, 1995). GLMBoost is shown to be a very fast algorithm with exceptional computation characteristics of high efficiency (Dettling and Bühlmann, 2003; Hao et al., 2014; Cengiz Colak et al., 2017). Furthermore, it is easy to build and does not need to run the algorithm multiple times for cross-validation. Due to the various advantages of GLMBoost for modeling the high-dimensional phenomenon, it is expected to be a suitable candidate for flash-flood susceptibility modeling. In this study, the GLMBoost model is run by the mboost R package (Hothorn et al., 2010), and its mstop parameter (i.e., number of Boosting iterations) was tuned by the tuning function of Caret R package (Kuhn, 2015).

2.2.4.2. Bayesian Generalized Linear Model (BayesGLM)

The GLM as a generalized and flexible linear regression (Lee and Nelder, 2006) has shown to highly benefit from the Bayesian techniques for efficient predicting the unknown parameters of the model (Antonio et al., 2005, Merl et al., 2008, Scollnik, 2005, Verrall, 2004). Modeling in a

Bayesian framework generally provides the opportunity of powerful yet low-cost computation which makes it suitable for high-dimensional data sets, e.g., hydrological data sets (Barbetta et al., 2018, Bolle et al., 2018, Liu and Merwade, 2018, Sikorska and Seibert, 2018). The Bayesian statistical analysis of *GLM, known as BayesGLM*, have recently become popular in a range of applications and have been applied to complex prediction modeling problems, e.g. health informatics and applied statistics, with promising results (Suleiman et al., 2019, Gelman et al., 2008, Ryu et al., 2018). In this study, the Bayesglm model was performed by the 'arm' package (Gelman and Hill, 2006) in the R software environment.

2.2.4.3. Random Forest (RF)

RF (Ho, 1995) is a popular ensemble machine learning method for mapping flood susceptibility (Lee et al., 2017; Rahmati and Pourghasemi, 2017). As a non-parametric and accurate classification and regression method, it has gained recognition in outperforming various machine learning methods in hydrological modeling and flood prediction systems (De Silva and Hornberger, 2019; Tyralis et al., 2019). For efficient regression and classification modeling, RF constructs a group of decision trees (DTs) in the framework of the random subspace method (Ho, 1995). The DTs in RF benefit from the controlled variance, which improves the prediction quality and troubleshoots the overfitting issues (Ho, 1998). RF represents bagging, a set of random samples, and features selection approach to ensemble learning (Breiman, 2001). For the RF model building, often the two-thirds of the data set goes for building DTs, and one-thirds goes for performance evaluation. In the next step, the sum of the DTs performed, and the best performing model is identified according to the most votes of all trees. The model has two parameters, i.e., ntree which specifies the number of trees and mtry which indicates the number of predictors randomly sampled for splitting at each tree node (Breiman, 2001). Furthermore, the out-of-bag (OOB) error rate is used to optimize parameters (Canion et al., 2019, Yang et al., 2019). In this study, the RF model was run using the randomForest package (Liaw and Wiener, 2002) in the R software environment, and their parameters were tuned by the tuning function of Caret R package (Kuhn, 2015).

2.2.5. Validation of the results

Evaluation of the modeling results was done using holdout data sets (which are not used in the calibration phase). Hit and miss analysis were used to assess the result of ML models. Metrics used in this study are including Accuracy (Acc, Eq. 1), Kappa (K, Eq. 2), success ratio (SR, Eq. 4), threat score (TS, Eq. 5) and, Heidke skill score (HSS, Eq. 6). Accuracy indicates what fraction of the modeled values are correct (Efron et al., 1986). Kappa is the degree of agreement between modeled and observed flood occurrences (Viera and Garrett, 2005). The success ratio indicates information about the likelihood of the modeled floods. Threat score indicates how were well modeled the observed flood occurrences (Stanski, 1989). HSS indicates the fraction of correct perditions after eliminating the random predictions (Heidke, 1926).

$$\text{Acc} = \frac{H+CN}{H+CN+M+CN} \qquad (1)$$

$$\text{K} = \frac{Acc - P_e}{1 - P_e} \qquad (2)$$

$$P_e = \frac{(H + FA)(H + M) + (M + CN)(FA + CN)}{(H + FA + M + CN)^2} \qquad (3)$$

$$\text{SR} = \frac{H}{H+FA} \qquad (4)$$

$$\text{TS}=\frac{H}{H+M+FA} \qquad (5)$$

$$\text{HSS}=\frac{2[(H\times CN)\text{-}(FA\times M)]}{[(H+M)(M+CN)+(H+FA)(FA+CN)]} \qquad (6)$$

where *H*, *FA*, *M*, and *CN* are respectively number of hits, false alarms, misses, and correct negatives that are computed using a contingency table (Johnson and Olsen, 1998). P_e indicates expected agreement between modeled and observed values (Viera and Garrett, 2005). Acc, K, SR, and TS range between zero (no skill) and 1 (perfect), while HSS varies between -1 and 1 (perfect) (Choubin et al., 2019a).

3. Results and discussion

3.1. Simulated annealing (SA) results

Using the SA method, the selected features were identified in each fold through the fitness values of Accuracy and Kappa. Table 1 indicates the optimum number of the features in each fold based on the Accuracy and Kappa metrics using the SA method. The minimum and the maximum number of the features identified by the SA method were equal to 6 (in Fold02) and 11 (Fold07) features, respectively. For example, Fold01 with a number of 7 features (including elevation, Dd, Dfr, land use, NDVI, precipitation, and TRI) had a higher modeling performance (respectively the Accuracy and Kappa values were equal to 0.90 and 0.80, Table 1) than using a less or greater number of the features. Therefore, the best number of selected features for flash-flood modeling can be between 6 and 11 features. Since the average number of the features for all folds was greater than 8 (equal to 8.4) (Table 1), so nine first features were selected as key features based on their occurrence frequencies in 10 folds (with 100 iterations, totally 1000 runs) (Fig. 5). Hence, nine important selected features (which had at least 50% frequency of occurrence in the folds) in this study were

NDVI, distance from stream (Dfs), elevation, precipitation, drainage density, soil type, flow accumulation (FA), topographic wetness index (TWI), and land use respectively with occurrence frequency of 100%, 100%, 100%, 100%, 80%, 60%, 60%, 50%, and 50% in the all folds (Fig. 5). Furthermore, the physical mechanisms of various features involved in the formation of floods can be individually discussed. The NDVI, soil type, and land use are the most effective factors in the generation or infiltration of runoff (Csáfordi et al., 2012). On the other hand, the Dfs and elevation are among the essential factors in modeling, in which the probability of flood events in the areas with low elevations and also close to rivers or streams are naturally high (Botzen et al., 2012; Choubin et al., 2019b). Yet, the precipitation and TWI features are important in surface runoff generation (Nampak et al. 2014). The drainage density is associated with lithology and topographic factors (such as slope and elevation), which its higher values may demonstrate a high runoff and low infiltration rate (Prasad et al., 2008). And finally, the FA factor shows the areas with highly accumulated water (Kia et al., 2012).

Table 1 SOMEWHERE HERE

Fig. 5 SOMEWHERE HERE

3.2. Model evaluation results

Model evaluation was conducted using hit and miss analysis after calibrating using key features identified by the SA method. Table 2 shows the performance of the predictive models for the testing data set. The accuracy (Acc) was equal to 90 % for both GLMBoost and BayesGLM models, while Random Forest (RF) indicated a higher accuracy (Acc= 92 %). Kappa (K) indicated a good performance ($0.55 < K < 0.85$) for the models according to the Monserud and Leemans (1992) (Table 2).

Success ratio (SR) indicated that 94%, 95%, and 96% of the modeled flood occurrences were actually observed respectively for the BayesGLM, RF, and GLMBoost models. The threat score (TS) showed that the correct forecasted floods by the GLMBoost, BayesGLM, and RF models are quale to 81%, 80%, and 84%, respectively. Also, the Heidke skill score (HSS) results highlighted that after eliminating the random predictions, the correct predicated floods are equal to 80%, 79%, and 0.84% respectively for the GLMBoost, BayesGLM, and RF models (Table 2).

Application of the GLMBoost and BayesGLM in flash-flood modeling was novel, and direct compassion with scholars was not possible. The BayesGLM model has never been used for flood susceptibility mapping or any other hydrological modeling up until now. But the popularity of boosting in the advancement of ensemble machine learning methods for hydrological modeling including the flood prediction has been fast-growing due to their accuracy (Antonetti et al., 2019; Berkhahn et al., 2019; Gomez et al., 2019; Lee et al., 2017; Peng et al., 2019; Tian et al., 2019). However, the promising results of GLMBoost in bioinformatics and biomedical applications have been confirmed to the artificial intelligence community (Hao et al., 2014; Dettling and Bühlmann, 2003). Among the used models, the application of RF is well established for flood susceptibility hazard assessment, and the literature includes adequate RF models with high accuracy and promising results in this realm (Al-Abadi, 2018, Chapi et al., 2017; Zhao et al., 2018). The popularity of RF in modeling the flash-flood susceptibility has also been increased during the past few years due to its simplicity, robustness and capacity to deal with complex data structures (Laudan et al., 2017, Muñoz et al., 2018, Terti et al., 2019). Thus, choosing the RF as a benchmark method was highly beneficial in this study to better explore the potential of the new methods, i.e., GLMBoost and BayesGLM. Generally, the comparison of the applied models in this study revealed that a good and close performance of them to model flash-flood locations (Table 2). The successful results of the RF model are in agreement with Wang et al. (2015), Lee et al. (2017), and Feng et al. (2015).

Table 2 SOMEWHERE HERE

3.3. Spatial assessment of flash-flood hazard

After validating the predictive models, spatial probability maps (from 0 to 1) of flood occurrence were predicted using the calibrated models and predictive variables for the whole study area. Then the probability maps were converted into five classes (i.e., very low, low, medium, high, and very high) with equal interval scheme within the ArcGIS software, so flash-flood hazard maps with 30 × 30 m cell size were produced (Fig. 6).

According to the flash-flood hazard maps, the area with very low and low hazard classes covers about 82.7% (9332.6 km^2), 83.4% (9414.3 km^2), and 84.4% (9530.2 km^2) of the total area by the BayesGLM, GLMBoost, and RF models. Moderate class predicted by the BayesGLM model is higher than the GLMBoost and RF models (respectively 8.9%, 5.7%, and 5.2% of the study area). Summation of the high and very high classes have respectively 10.9% (1230.0 km^2), 10.3% (1167.5 km^2), and 8.4% (950.8 km^2) of the study area for the GLMBoost, RF, and BayesGLM models, which mostly are placed around the rivers in the down streams of the study area (Fig. 6).

Fig. 6 SOMEWHERE HERE

3.4. Contribution analysis of predictive variables

The importance of the predictive variables based on the percent decrease in area under the curve (DAUC) of the receiver operating characteristics (ROC) was analyzed. Results highlighted that the most importance variables in the modeling process was Dfs (DAUC = 81.48%) and NDVI

(DAUC = 81.16%), while variables of Dd (DAUC = 75.84%), land use (DAUC = 73.62%), and elevation (DAUC = 73.33%) had a moderate importance (Fig. 7). Also, DAUC indicated that TWI, precipitation, FA, and soil type were in the next orders in view of DAUC (respectively equal to 57.18%, 54.29%, 50.65%, and 50.16%) (Fig. 7). In this regard, Siahkamari et al. (2018) and Choubin et al. (2019b) indicated that Dfs was one of the most important variables in flood modeling.

Probability curves based on the GLMBoost model for each variable are shown in Fig. 8. P (1) and P (0) respectively indicate the probability of flood occurrence and non-occurrence. Following the present results, with increasing elevation, the probability of flood occurrence increased (Fig. 8a). While flood decreased with increasing flow accumulation, TWI, drainage density, and precipitation (Fig. 8b to Fig. 8d, and Fig. 8f). On the contrary, when the distance from stream increased, the flood occurrence decreased (Fig. 8e). This is in accordance with the results of Hong et al. (2018) and Darabi et al. (2019). Furthermore, variations of NDVI indicated that with increasing it, the flood occurrence decreased (Fig. 8g). This is well-indicated with the role of vegetation in the control and infiltration of surface water (Islam and Sado, 2000).

For categorical variables (i.e., soil type and land use), the probability values indicate mean probability in each class (Fig. 8h and Fig. 8i). Soil types of Salt Flats and Aridisols are shown a high probability of flood occurrence (Fig. 8h). As can be seen from the land use map, water/wetland, and residential areas had the most probability of flood occurrence, while forest indicated less probability (Fig. 8i).

Fig. 6 SOMEWHERE HERE

Fig. 8 SOMEWHERE HERE

4. Conclusion

This study used three state-of-the-art machine learning models (i.e., GLMBoost, RF, and BayesGLM) for modeling flash-flood in an area that is strongly affected by the flood. The application of the GLMBoost and BayesGLM in this study was novel. Moreover, the simulated annealing (SA) method as a novel and successful method was used to eliminate redundant variables from the flood modeling process for the first time. Results of the flash-flood modeling revealed that the applied models had a good and close performance (e.g., Accuracy = 90% for both models). Variables of Dfs, NDVI, Dd, land use and elevation had more contribution, among others. Although results indicated a good performance of the modeling, lack of soil data such as infiltration and soil hydrological groups, which have effects on surface runoff, was one of the limitations of the study. Furthermore, due to the lack of recorded flood locations, this study tried to extract flooded areas from remotely sensed data for a short period (from March to April). Still, flooding is affected by various variables such as return period (RP), in future investigations, it might be possible to use a different RPs of recorded precipitation and flood occurrence locations for calibrating the models and extracting different hazardous area based on the RPs. Nevertheless, our findings can significantly facilitate understanding the hazardous area and help watershed managers to control and remediate induced damages of flood in a data-scarce region.

References

Aarts, E. & Korst, J. 1988. Simulated annealing and Boltzmann machines.

Abuzied, S., Yuan, M., Ibrahim, S., Kaiser, M. & Saleem, T. 2016. Geospatial risk assessment of flash floods in Nuweiba area, Egypt. Journal of Arid Environments, 133, 54-72.

Ahmadalipour, A., Moradkhani, H., Castelletti, A. & Magliocca, N. 2019. Future drought risk in Africa: Integrating vulnerability, climate change, and population growth. Science of the Total Environment, 662, 672-686.

Al-Abadi, A. M. 2018. Mapping flood susceptibility in an arid region of southern Iraq using ensemble machine learning classifiers: a comparative study. Arabian Journal of Geosciences, 11.

Alfieri, L., Feyen, L., Dottori, F. & Bianchi, A. 2015. Ensemble flood risk assessment in Europe under high end climate scenarios. Global Environmental Change, 35, 199-212.

Alilou, H., Rahmati, O., Singh, V.P., Choubin, B., Pradhan, B., Keesstra, S., Ghiasi, S.S. and Sadeghi, S.H., 2019. Evaluation of watershed health using Fuzzy-ANP approach considering geo-environmental and topo-hydrological criteria. *Journal of environmental management, 232*, pp.22-36.

Al-Juaidi, A. E. M., Nassar, A. M. & Al-Juaidi, O. E. M. 2018. Evaluation of flood susceptibility mapping using logistic regression and GIS conditioning factors. Arabian Journal of Geosciences, 11.

Alves, P. B. R., De Melo Filho, H., Tsuyuguchi, B. B., Rufino, I. A. A. & Feitosa, P. H. C. 2018. Mapping of flood susceptibility in Campina Grande County – PB: A spatial multicriteria approach. Boletim de Ciencias Geodesicas, 24, 28-43.

Antonetti, M., Horat, C., Sideris, I. V. & Zappa, M. 2019. Ensemble flood forecasting considering dominant runoff processes - Part 1: Set-up and application to nested basins (Emme, Switzerland). Natural Hazards and Earth System Sciences, 19, 19-40.

Antonio, K., Beirlant, J. & Hoedemakers, T. 2005. "A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving," R. J. Verrall, July 2004. North American Actuarial Journal, 9, 130-142.

Barbetta, S., Coccia, G., Moramarco, T. & Todini, E. 2018. Real-time flood forecasting downstream river confluences using a Bayesian approach. Journal of Hydrology, 565, 516-523.

Berkhahn, S., Fuchs, L. & Neuweiler, I. 2019. An ensemble neural network model for real-time prediction of urban floods. Journal of Hydrology, 575, 743-754.

Bolle, A., Das Neves, L., Smets, S., Mollaert, J. & Buitrago, S. 2018. An impact-oriented Early Warning and Bayesian-based Decision Support System for flood risks in Zeebrugge harbour. Coastal Engineering, 134, 191-202.

Botzen WJW, Aerts JCJH, Van Den Bergh JCJM. 2012. Individual preferences for reducing flood risk to near zero through elevation. Mitig Adapt Strateg Glob Change. 2:229–244. doi:10.1007/s11027-012-9359-5.

Botzen WJW, Aerts JCJH, van den Bergh JCJM. 2012. Individual preferences for reducing flood risk to near zero through elevation. Mitig Adapt Strateg Glob Change. 2:229–244. doi:10.1007/s11027-012-9359-5.

Breiman, L. 2001. Random forests. Machine Learning, 45, 5-32.

Breiman, L. 2004. The 2002 wald memorial lectures population theory for boosting ensembles. Annals of Statistics, 32, 1-11.

Buchen, T. & Wohlrabe, K. 2011. Forecasting with many predictors: Is boosting a viable alternative? Economics Letters, 113, 16-18.

Bühlmann, P. & Hothorn, T. 2007. Boosting algorithms: Regularization, prediction and model fitting. Statistical Science, 22, 477-505.

Bühlmann, P. & Yu, B. 2003. Boosting with the L2 loss: Regression and classification. Journal of the American Statistical Association, 98, 324-339.

Bühlmann, P. 2006. Boosting for high-dimensional linear models. Annals of Statistics, 34, 559-583.

Bui, D. T., Khosravi, K., Li, S., Shahabi, H., Panahi, M., Singh, V. P., Chapi, K., Shirzadi, A., Panahi, S., Chen, W. & Bin Ahmad, B. 2018a. New hybrids of ANFIS with several optimization algorithms for flood susceptibility modeling. Water (Switzerland), 10.

Bui, D. T., Ngo, P. T. T., Pham, T. D., Jaafari, A., Minh, N. Q., Hoa, P. V. & Samui, P. 2019a. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. Catena, 179, 184-196.
Bui, D. T., Panahi, M., Shahabi, H., Singh, V. P., Shirzadi, A., Chapi, K., Khosravi, K., Chen, W., Panahi, S., Li, S. & Ahmad, B. B. 2018b. Novel Hybrid Evolutionary Algorithms for Spatial Prediction of Floods. Scientific Reports, 8.
Bui, D. T., Tsangaratos, P., Ngo, P. T. T., Pham, T. D. & Pham, B. T. 2019b. Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods. Science of the Total Environment, 668, 1038-1054.
Canion, A., Mccloud, L. & Dobberfuhl, D. 2019. Predictive modeling of elevated groundwater nitrate in a karstic spring-contributing area using random forests and regression-kriging. Environmental Earth Sciences, 78.
Casagrande, L., Tomasella, J., Dos Santos Alvalá, R. C., Bottino, M. J. & De Oliveira Caram, R. 2017. Early flood warning in the Itajaí-Açu River basin using numerical weather forecasting and hydrological modeling. Natural Hazards, 88, 741-757.
Cengiz Colak, M., Karaaslan, E., Colak, C., Arslan, A. K. & Erdil, N. 2017. Handling imbalanced class problem for the prediction of atrial fibrillation in obese patient. Biomedical Research (India), 28, 3293-3299.
Chapi, K., Singh, V. P., Shirzadi, A., Shahabi, H., Bui, D. T., Pham, B. T. & Khosravi, K. 2017. A novel hybrid artificial intelligence approach for flood susceptibility assessment. Environmental Modelling and Software, 95, 229-245.
Chen, W., Hong, H., Li, S., Shahabi, H., Wang, Y., Wang, X. & Ahmad, B. B. 2019. Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles. Journal of Hydrology, 575, 864-873.
Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D. T., Duan, Z. & Ma, J. 2017. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena, 151, 147-160.
Chiang, Y. C. & Ling, T. Y. 2017. Exploring flood resilience thinking in the retail sector under climate change: A case study of an estuarine region of Taipei City. Sustainability (Switzerland), 9.
Choubin, B., Abdolshahnejad, M., Moradi, E., Querol, X., Mosavi, A., Shamshirband, S., Ghamisi, P. 2020. Spatial hazard assessment of the PM10 using machine learning models in Barcelona, Spain. Science of the Total Environment. DOI: 10.1016/j.scitotenv.2019.134474
Choubin, B., Borji, M., Mosavi, A., Sajedi-Hosseini, F., Singh, V.P. and Shamshirband, S., 2019a. Snow avalanche hazard prediction using machine learning methods. Journal of Hydrology, 577, p.123929.
Choubin, B., Moradi, E., Golshan, M., Adamowski, J., Sajedi-Hosseini, F. and Mosavi, A., 2019b. An Ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. Science of the Total Environment, 651, pp.2087-2096.
Clement, M.A., Kilsby, C.G. and Moore, P., 2018. Multi-temporal synthetic aperture radar flood mapping using change detection. Journal of Flood Risk Management, 11(2), pp.152-168.
Costache, R. 2019. Flash-Flood Potential assessment in the upper and middle sector of Prahova river catchment (Romania). A comparative approach between four hybrid models. Science of The Total Environment, 659, 1115-1134.

Csáfordi, P., Pődör, A., Bug, J., & Gribovsyki, Z. (2012). Soil erosion analysis in a small forested catchment supported by ArcGIS Model Builder. Acta Silvatica et Lignaria Hungarica, 8(1), 39-56.
Darabi, H., Choubin, B., Rahmati, O., Haghighi, A. T., Pradhan, B. & Kløve, B. 2019. Urban flood risk mapping using the GARP and QUEST models: A comparative study of machine learning techniques. Journal of hydrology, 569, 142-154.
De Silva, T. M. & Hornberger, G. M. 2019. Identifying El Niño-Southern Oscillation influences on rainfall with classification models: Implications for water resource management of Sri Lanka. Hydrology and Earth System Sciences, 23, 1905-1929.
Dettling, M. & Bühlmann, P. 2003. Boosting for tumor classification with gene expression data. Bioinformatics, 19, 1061-1069.
Dietterich, T. G. 2000. Ensemble methods in machine learning C3 - Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).
Donya-e-eqtesad (2019, Apr 6). Retrieved from https://www.donya-e-eqtesad.com/fa/tiny/news-3511460.
Doswell Iii, C. A., Brooks, H. E. & Maddox, R. A. 1996. Flash flood forecasting: An ingredients-based methodology. Weather and Forecasting, 11, 560-581.
Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W. and Li, X., 2016. Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. Remote Sensing, 8(4), p.354.
Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W. and Li, X., 2016. Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. Remote Sensing, 8(4), p.354.
Du, Z., Li, W., Zhou, D., Tian, L., Ling, F., Wang, H., Gui, Y. and Sun, B., 2014. Analysis of Landsat-8 OLI imagery for land surface water mapping. Remote sensing letters, 5(7), pp.672-681.
Efron, B. and R. Tibshirani, 1986: Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Statistical Science, 1, 54-77.
Feng, Q., Liu, J. and Gong, J., 2015. Urban flood mapping based on unmanned aerial vehicle remote sensing and random forest classifier—A case of Yuyao, China. *Water*, *7*(4), pp.1437-1455.
Freund, Y. & Schapire, R. E. 1995. A decision-theoretic generalization of on-line learning and an application to boosting C3 - Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). In: VITANYI, P. (ed.). Springer Verlag.
Frigerio, S., Schenato, L., Bossi, G., Mantovani, M., Marcato, G. & Pasuto, A. 2018. Hands-on experience of crowdsourcing for flood risks. An android mobile application tested in Frederikssund, Denmark. International Journal of Environmental Research and Public Health, 15.
Gelman, A. and Hill, J., 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge university press.
Gelman, A., Jakulin, A., Pittau, M. G. & Su, Y. S. 2008. A weakly informative default prior distribution for logistic and other regression models. Annals of Applied Statistics, 2, 1360-1383.
Giovannettone, J., Copenhaver, T., Burns, M. & Choquette, S. 2018. A Statistical Approach to Mapping Flood Susceptibility in the Lower Connecticut River Valley Region. Water Resources Research, 54, 7603-7618.
Gomez, M., Sharma, S., Reed, S. & Mejia, A. 2019. Skill of ensemble flood inundation forecasts at short- to medium-range timescales. Journal of Hydrology, 568, 207-220.

Google Earth Engine (GEE). Sentinel-2 MSI: Multi Spectral Instrument, Level-1C. 2019a. Available at: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2 [accessed 20 May 2019].

Google Earth Engine (GEE). USGS Landsat 8 Surface Reflectance Tier 1. 2019b. Available at: https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C01_T1_SR [accessed 20 May 2019].

Hao, M., Wang, Y. & Bryant, S. H. 2014. An efficient algorithm coupled with synthetic minority over-sampling technique to classify imbalanced PubChem BioAssay data. Analytica Chimica Acta, 806, 117-127.

Hao, W., Hao, Z., Yuan, F., Ju, Q. & Hao, J. 2019. Regional frequency analysis of precipitation extremes and its spatio-temporal patterns in the Hanjiang river basin, China. Atmosphere, 10.

Heidke, P., 1926. Berechnung des Erfolges und der Güte der Windstärkevorhersagen im Sturmwarnungsdienst. Geografiska Annaler, 8(4), pp.301-349.

Hennequin, T., Sørup, H. J. D., Dong, Y. & Arnbjerg-Nielsen, K. 2018. A framework for performing comparative LCA between repairing flooded houses and construction of dikes in non-stationary climate with changing risk of flooding. Science of the Total Environment, 642, 473-484.

Henriksen, H. J., Roberts, M. J., Van Der Keur, P., Harjanne, A., Egilson, D. & Alfonso, L. 2018. Participatory early warning and monitoring systems: A Nordic framework for web-based flood risk management. International Journal of Disaster Risk Reduction, 31, 1295-1306.

Ho, T. K. 1995. Random decision forests C3 - Proceedings of the International Conference on Document Analysis and Recognition, ICDAR. IEEE Computer Society, 278-282.

Ho, T. K. 1998. The random subspace method for constructing decision forests. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20, 832-844.

Hofierka, J. & Knutová, M. 2015. Simulating spatial aspects of a flash flood using the Monte Carlo method and GRASS GIS: A case study of the Malá Svinka Basin (Slovakia). Open Geosciences, 7, 118-125.

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

Hothorn, T. & Bühlmann, P. 2006. Model-based boosting in high dimensions. Bioinformatics, 22, 2828-2829.

Hothorn, T., Bühlmann, P., Kneib, T., Schmid, M. and Hofner, B., 2010. Model-based boosting 2.0. Journal of Machine Learning Research, 11(Aug), pp.2109-2113.

Huang, C. L., Hsu, N. S., Liu, H. J. & Huang, Y. H. 2018. Optimization of low impact development layout designs for megacity flood mitigation. Journal of Hydrology, 564, 542-558.

Islam, M.M. and Sado, K., 2000. Flood hazard assessment in Bangladesh using NOAA AVHRR data with geographical information system. Hydrological Processes, 14(3), pp.605-620.

Jannati, H. (2019, Apr 12). History of the devastating floods in Iran. "Political Studies and Research Institute of Iran". Retrieved from http://ir-psri.com/?Page=ViewNews&NewsID=6283.

Johnson, L.E. and Olsen, B.G., 1998. Assessment of quantitative precipitation forecasts. Weather and forecasting, 13(1), pp.75-83.

Kalantar, B., Pradhan, B., Naghibi, S.A., Motevalli, A., Mansor, S. 2017. Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support

vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). Geomat. Nat. Haz. Risk. 5705, 1–21.
Kamyar Kalantar-Zadeh and Denis Fouque, 2017, Nutritional Management of Chronic Kidney Disease, The new england journal of medicine
Kanani-Sadat, Y., Arabsheibani, R., Karimipour, F. & Nasseri, M. 2019. A new approach to flood susceptibility assessment in data-scarce and ungauged regions based on GIS-based hybrid multi criteria decision-making method. Journal of Hydrology, 572, 17-31.
Khosravi, K., Pham, B. T., Chapi, K., Shirzadi, A., Shahabi, H., Revhaug, I., Prakash, I. & Tien Bui, D. 2018. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. Science of the Total Environment, 627, 744-755.
Khosravi, K., Shahabi, H., Pham, B. T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H. B., Gróf, G., Ho, H. L., Hong, H., Chapi, K. & Prakash, I. 2019. A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. Journal of Hydrology, 573, 311-323.
Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A. (2012). An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. Environmental Earth Sciences, 67(1), 251-264.
Kira, K. & Rendell, L. A. 1992.The feature selection problem: Traditional methods and a new algorithm. Aaai, 129-134.
Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. 1983. Optimization by simulated annealing. science, 220, 671-680.
Kubwarugira, G., Mayoussi, M. & El Khalki, Y. 2019. Assessing flood exposure in informal districts: a case study of Bujumbura, Burundi. Journal of Applied Water Engineering and Research.
Kuenzer, C., Guo, H., Huth, J., Leinenkugel, P., Li, X. & Dech, S. 2013. Flood mapping and flood dynamics of the mekong delta: ENVISAT-ASAR-WSM based time series analyses. Remote Sensing, 5, 687-715.
Kuhn, M., 2015. Caret: classification and regression training. Astrophysics Source Code Library.
Kundzewicz, Z. W., Su, B., Wang, Y., Xia, J., Huang, J. & Jiang, T. 2019. Flood risk and its reduction in China. Advances in Water Resources, 130, 37-45.
Kwak, Y., Park, J. & Fukami, K. 2014. Near real-time flood volume estimation from MODIS time-series imagery in the indus river basin. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 578-586.
Laudan, J., Rözer, V., Sieg, T., Vogel, K. & Thieken, A. H. 2017. Damage assessment in Braunsbach 2016: Data collection and analysis for an improved understanding of damaging processes during flash floods. Natural Hazards and Earth System Sciences, 17, 2163-2179.
Lee, S., Kim, J. C., Jung, H. S., Lee, M. J. & Lee, S. 2017. Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. Geomatics, Natural Hazards and Risk, 8, 1185-1203.
Lee, Y. & Nelder, J. A. 2006. Double hierarchical generalized linear models. Journal of the Royal Statistical Society. Series C: Applied Statistics, 55, 139-185.
Li, W., Du, Z., Ling, F., Zhou, D., Wang, H., Gui, Y., Sun, B. and Zhang, X., 2013. A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. Remote Sensing, 5(11), pp.5530-5549.

Li, Y., Huang, C., Ngo, H. H., Pang, J., Zha, X., Liu, T. & Guo, W. 2019. In situ reconstruction of long-term extreme flooding magnitudes and frequencies based on geological archives. Science of the Total Environment, 670, 8-17.

Liaw, A., Wiener, M., 2002b. The randomforest package. R news 2 (3), 18–22.

Lindenschmidt, K. E., Das, A., Rokaya, P. & Chu, T. 2016. Ice-jam flood risk assessment and mapping. Hydrological Processes, 30, 3754-3769.

Liu, H. & Motoda, H. 2007. Computational methods of feature selection, CRC Press.

Liu, Z. & Merwade, V. 2018. Accounting for model structure, parameter and input forcing uncertainty in flood inundation modeling using Bayesian model averaging. Journal of Hydrology, 565, 138-149.

Luu, C., Von Meding, J. & Mojtahedi, M. 2019. Analyzing Vietnam's national disaster loss database for flood risk assessment using multiple linear regression-TOPSIS. International Journal of Disaster Risk Reduction.

Lyubchich, V., Newlands, N. K., Ghahari, A., Mahdi, T. & Gel, Y. R. 2019. Insurance risk assessment in the face of climate change: Integrating data science and statistics. Wiley Interdisciplinary Reviews: Computational Statistics, 11.

Mahmood, S., & Rahman, A. 2019. Flash flood susceptibility modeling using geo-morphometric and hydrological approaches in Panjkora Basin, Eastern Hindu Kush, Pakistan. Environmental earth sciences, 78(1), 43.

Mahmood, S., Rahman, A. U., & Shaw, R. 2019. Spatial appraisal of flood risk assessment and evaluation using integrated hydro-probabilistic approach in Panjkora River Basin, Pakistan. Environmental Monitoring and Assessment, 191(9), 573.

Mccullagh, P. 1984. Generalized linear models. European Journal of Operational Research, 16, 285-292.

McCullagh, Peter. Generalized linear models. Routledge, 2019.

Merl, D., Prado, R. & Escalante, A. A. 2008. Assessing the effect of selection at the amino acid level in malaria antigen sequences through Bayesian generalized linear models. Journal of the American Statistical Association, 103, 1496-1507.

Mojaddadi, H., Pradhan, B., Nampak, H., Ahmad, N., & Ghazali, A. H. B. (2017). Ensemble machine-learning-based geospatial approach for flood risk assessment using multisensor remote-sensing data and GIS. Geomatics, Natural Hazards and Risk, 1–23.

Monserud, R.A. and Leemans, R., 1992. Comparing global vegetation maps with the Kappa statistic. Ecological modelling, 62(4), pp.275-293.

Mosavi, A., Ozturk, P. & Chau, K. W. 2018. Flood prediction using machine learning models: Literature review. Water (Switzerland), 10.

Motevalli, A. & Vafakhah, M. 2016. Flood hazard mapping using synthesis hydraulic and geomorphic properties at watershed scale. Stochastic Environmental Research and Risk Assessment, 30, 1889-1900.

Muhamad, N., Lim, C. S., Reza, M. I. H. & Pereira, J. J. 2019. (The needs of disaster susceptibility map as an input in land use management: A case study of Universiti kebangsaan Malaysia). Sains Malaysiana, 48, 33-43.

Muñoz, P., Orellana-Alvear, J., Willems, P. & Célleri, R. 2018. Flash-flood forecasting in an andean mountain catchment-development of a step-wise methodology based on the random forest algorithm. Water (Switzerland), 10.

Nampak, H., Pradhan, B., & Manap, M. A. (2014). Application of GIS based data driven evidential belief function model to predict groundwater potential zonation. Journal of Hydrology, 513, 283-300.

Ngo, P. T. T., Hoang, N. D., Pradhan, B., Nguyen, Q. K., Tran, X. T., Nguyen, Q. M., Nguyen, V. N., Samui, P. & Bui, D. T. 2018. A novel hybrid swarm optimized multilayer neural network for spatial prediction of flash floods in tropical areas using sentinel-1 SAR imagery and geospatial data. Sensors (Switzerland), 18.

Peng, A., Zhang, X., Peng, Y., Xu, W. & You, F. 2019. The application of ensemble precipitation forecasts to reservoir operation. Water Science and Technology: Water Supply, 19, 588-595.

Prasad, B., & Sangita, K. (2008). Heavy metal pollution index of ground water of an abandoned open cast mine filled with fly ash: a case study. Mine water and the Environment, 27(4), 265-267.

Qomnews (2019, Murch 24). Retrieved from http://www.qumpress.ir/302087

Rahmati, O. & Pourghasemi, H. R. 2017. Identification of Critical Flood Prone Areas in Data-Scarce and Ungauged Regions: A Comparison of Three Data Mining Models. Water Resources Management, 31, 1473-1487.

Razavi Termeh, S. V., Kornejady, A., Pourghasemi, H. R. & Keesstra, S. 2018. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. Science of the Total Environment, 615, 438-451.

Roslee, R. & Norhisham, M. N. 2018. Flood susceptibility analysis using multi-criteria evaluation model: A case study in Kota Kinabalu, Sabah. ASM Science Journal, 11, 123-133.

Ryu, D., Bilgili, D., Ergönül, Ö., Liang, F. & Ebrahimi, N. 2018. A Bayesian Generalized Linear Model for Crimean–Congo Hemorrhagic Fever Incidents. Journal of Agricultural, Biological, and Environmental Statistics, 23, 153-170.

Safaripour, M., Monavari, M., Zare, M., Abedi, Z. and Gharagozlou, A., 2012. Flood Risk Assessment Using GIS (Case Study: Golestan Province, Iran). Polish Journal of Environmental Studies, 21(6).

Sahana, M. & Patel, P. P. 2019. A comparison of frequency ratio and fuzzy logic models for flood susceptibility assessment of the lower Kosi River Basin in India. Environmental Earth Sciences, 78.

Sajedi-Hosseini, F., Choubin, B., Solaimani, K., Cerdà, A. and Kavian, A., 2018. Spatial prediction of soil erosion susceptibility using a fuzzy analytical network process: Application of the fuzzy decision making trial and evaluation laboratory approach. *Land degradation & development*, *29*(9), pp.3092-3103.

Sciance, M. B. & Nooner, S. L. 2018. Decadal flood trends in Bangladesh from extensive hydrographic data. Natural Hazards, 90, 115-135.

Scollnik, D. P. M. 2005. "A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving," R. J. Verrall, July 2004. North American Actuarial Journal, 9, 143-145.

Serago, J. M. & Vogel, R. M. 2018. Parsimonious nonstationary flood frequency analysis. Advances in Water Resources, 112, 1-16.

Sghaier, M. O., Hammami, I., Foucher, S. & Lepage, R. 2018. Flood extent mapping from time-series SAR images based on texture analysis and data fusion. Remote Sensing, 10.

Shafapour Tehrany, M., Shabani, F., Neamah Jebur, M., Hong, H., Chen, W. & Xie, X. 2017. GIS-based spatial prediction of flood prone areas using standalone frequency ratio, logistic regression, weight of evidence and their ensemble techniques. Geomatics, Natural Hazards and Risk, 8, 1538-1561.

Shafizadeh-Moghadam, H., Valavi, R., Shahabi, H., Chapi, K. and Shirzadi, A., 2018. Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. Journal of environmental management, 217, pp.1-11.

Shkolnik, I., Pavlova, T., Efimov, S. & Zhuravlev, S. 2018. Future changes in peak river flows across northern Eurasia as inferred from an ensemble of regional climate projections under the IPCC RCP8.5 scenario. Climate Dynamics, 50, 215-230.

Siahkamari, S., Haghizadeh, A., Zeinivand, H., Tahmasebipour, N. and Rahmati, O., 2018. Spatial prediction of flood-susceptible areas using frequency ratio and maximum entropy models. Geocarto international, 33(9), pp.927-941.

Sikorska, A. E. & Seibert, J. 2018. Value of different precipitation data for flood prediction in an alpine catchment: A Bayesian approach. Journal of Hydrology, 556, 961-971.

Singh, K.V., Setia, R., Sahoo, S., Prasad, A. and Pateriya, B., 2015. Evaluation of NDWI and MNDWI for assessment of waterlogging by integrating digital elevation model and groundwater level. Geocarto International, 30(6), pp.650-661.

Sozer, B., Kocaman, S., Nefeslioglu, H. A., Firat, O. & Gokceoglu, C. Preliminary investigations on flood susceptibility mapping in Ankara (Turkey) using modified analytical hierarchy process (M-AHP). In: SARAN, S., PADALIA, H. & KUMAR, A. S., eds., 2018. International Society for Photogrammetry and Remote Sensing, 361-365.

Stanski, H.R., L.J. Wilson, and W.R. Burrows, 1989: Survey of common verification methods in meteorology. World Weather Watch Tech. Rept. No.8, WMO/TD No.358, WMO, Geneva, 114 pp.

Suleiman, M., Demirhan, H., Boyd, L., Girosi, F. & Aksakalli, V. 2019. Bayesian logistic regression approaches to predict incorrect DRG assignment. Health Care Management Science, 22, 364-375.

Tang, Z., Yi, S., Wang, C. & Xiao, Y. 2018. Incorporating probabilistic approach into local multi-criteria decision analysis for flood susceptibility assessment. Stochastic Environmental Research and Risk Assessment, 32, 701-714.

Tehrany, M. S. & Kumar, L. 2018. The application of a Dempster–Shafer-based evidential belief function in flood susceptibility mapping and comparison with frequency ratio and logistic regression methods. Environmental Earth Sciences, 77.

Terti, G., Ruin, I., Gourley, J. J., Kirstetter, P., Flamig, Z., Blanchet, J., Arthur, A. & Anquetin, S. 2019. Toward Probabilistic Prediction of Flash Flood Human Impacts. Risk Analysis, 39, 140-161.

Tian, J., Liu, J., Yan, D., Ding, L. & Li, C. 2019. Ensemble flood forecasting based on a coupled atmospheric-hydrological modeling system with data assimilation. Atmospheric Research, 224, 127-137.

Tiryaki, M. & Karaca, O. 2018. Flood susceptibility mapping using GIS and multicriteria decision analysis: Saricay-Çanakkale (Turkey). Arabian Journal of Geosciences, 11.

Torabi Haghighi, A., Menberu, M. W., Darabi, H., Akanegbu, J., & Kløve, B. (2018). Use of remote sensing to analyse peatland changes after drainage for peat extraction. Land degradation & development, 29(10), 3479-3488.

Torabi Haghighi, A., Menberu, M. W., Darabi, H., Akanegbu, J., & Kløve, B. (2018). Use of remote sensing to analyse peatland changes after drainage for peat extraction. Land degradation & development, 29(10), 3479-3488.

Tyralis, H., Papacharalampous, G. & Tantanee, S. 2019. How to explain and predict the shape parameter of the generalized extreme value distribution of streamflow extremes using a big dataset. Journal of Hydrology, 574, 628-645.
Uzielli, M., Rianna, G., Ciervo, F., Mercogliano, P. & Eidsvig, U. K. 2018. Temporal evolution of flow-like landslide hazard for a road infrastructure in the municipality of Nocera Inferiore (southern Italy) under the effect of climate change. Natural Hazards and Earth System Sciences, 18, 3019-3035.
Valavi, R., Shafizadeh-Moghadam, H., Matkan, A., Shakiba, A., Mirbagheri, B. and Kia, S.H., 2019. Modelling climate change effects on Zagros forests in Iran using individual and ensemble forecasting approaches. *Theoretical and Applied Climatology*, *137*(1-2), pp.1015-1025.
Van Laarhoven, P. J. & Aarts, E. H. 1987. Simulated annealing. Simulated annealing: Theory and applications. Springer.
Verrall, R. J. 2004. A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving. North American Actuarial Journal, 8, 67-89.
Viera, A.J. and Garrett, J.M., 2005. Understanding interobserver agreement: the kappa statistic. Fam med, 37(5), pp.360-363.
Wang, Z., Lai, C., Chen, X., Yang, B., Zhao, S., & Bai, X. (2015). Flood hazard risk assessment model based on random forest. Journal of Hydrology, 527, 1130-1141.
Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International journal of remote sensing, 27(14), pp.3025-3033.
Yan, J., Jin, J., Chen, F., Yu, G., Yin, H. & Wang, W. 2018. Urban flash flood forecast using support vector machine and numerical simulation. Journal of Hydroinformatics, 20, 232-245.
Yang, J., Griffiths, J. & Zammit, C. 2019. National classification of surface–groundwater interaction using random forest machine learning technique. River Research and Applications.
Youssef, A. M., Pradhan, B. & Sefry, S. A. 2016. Flash flood susceptibility assessment in Jeddah city (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models. Environmental Earth Sciences, 75, 1-16.
Zhang, C. & Ma, Y. 2012. Ensemble machine learning: Methods and applications, Springer US.
Zhao, G., Pang, B., Xu, Z., Yue, J. & Tu, T. 2018. Mapping flood susceptibility in mountainous areas on a national scale in China. Science of the Total Environment, 615, 1133-1142.
Zhou, Q., Mikkelsen, P. S., Halsnæs, K. & Arnbjerg-Nielsen, K. 2012. Framework for economic pluvial flood risk assessment considering climate change effects and adaptation benefits. Journal of Hydrology, 414, 539-549.
Zhu, C. & Wu, J. 2013. Hybrid of genetic algorithm and simulated annealing for support vector regression optimization in rainfall forecasting. International Journal of Computational Intelligence and Applications, 12.

Tables:

Table 1 Selected features in each fold using the SA method

Fold	Number of the selected features	Selected features	Accuracy	Kappa

Fold01	7	Elevation, Dd, Dfr, Landuse, NDVI, Precipitation, TRI	0.90	0.80
Fold02	6	Aspect, Elevation, Dfr, NDVI, Precipitation, Soil type	0.89	0.79
Fold03	8	FA, Elevation, Dd, Dfr, NDVI, Precipitation, TRI, Soil type	0.81	0.61
Fold04	9	FA, Elevation, Dd, Soil depth, Dfr, Landuse, NDVI, Precipitation, TPI	0.90	0.80
Fold05	10	FA, Elevation, Dd, Soil depth, Dfr, Lithology, NDVI, Precipitation, TRI, TWI	0.84	0.68
Fold06	7	Elevation, Dd, Dfr, Landuse, NDVI, Precipitation, TWI	0.93	0.86
Fold07	11	FA, Elevation, Dd, Soil depth, Dfr, Landuse, Lithology, NDVI, Precipitation, TRI, Soil type	0.92	0.84
Fold08	9	Aspect, Elevation, Dd, Soil depth, Dfr, Landuse, NDVI, Precipitation, Soil type	0.85	0.69
Fold09	7	FA, Elevation, Dfr, NDVI, Precipitation, Slope, Soil type	0.95	0.91
Fold10	10	FA, Aspect, Elevation, Dd, Dfr, NDVI, Precipitation, TPI, TRI, Soil type	0.87	0.74
Average	8.4	–	0.89	0.77

Table 2 Performance of the predictive models for the test data set

Statistic	GLMBoost	BayesGLM	Random Forest
Accuracy (Acc)	0.90	0.90	0.92
Kappa (K)	0.80	0.79	0.84
Success ratio (SR)	0.96	0.94	0.95
Threat score (TS)	0.81	0.80	0.84
Heidke skill score (HSS)	0.80	0.79	0.84

Figures:

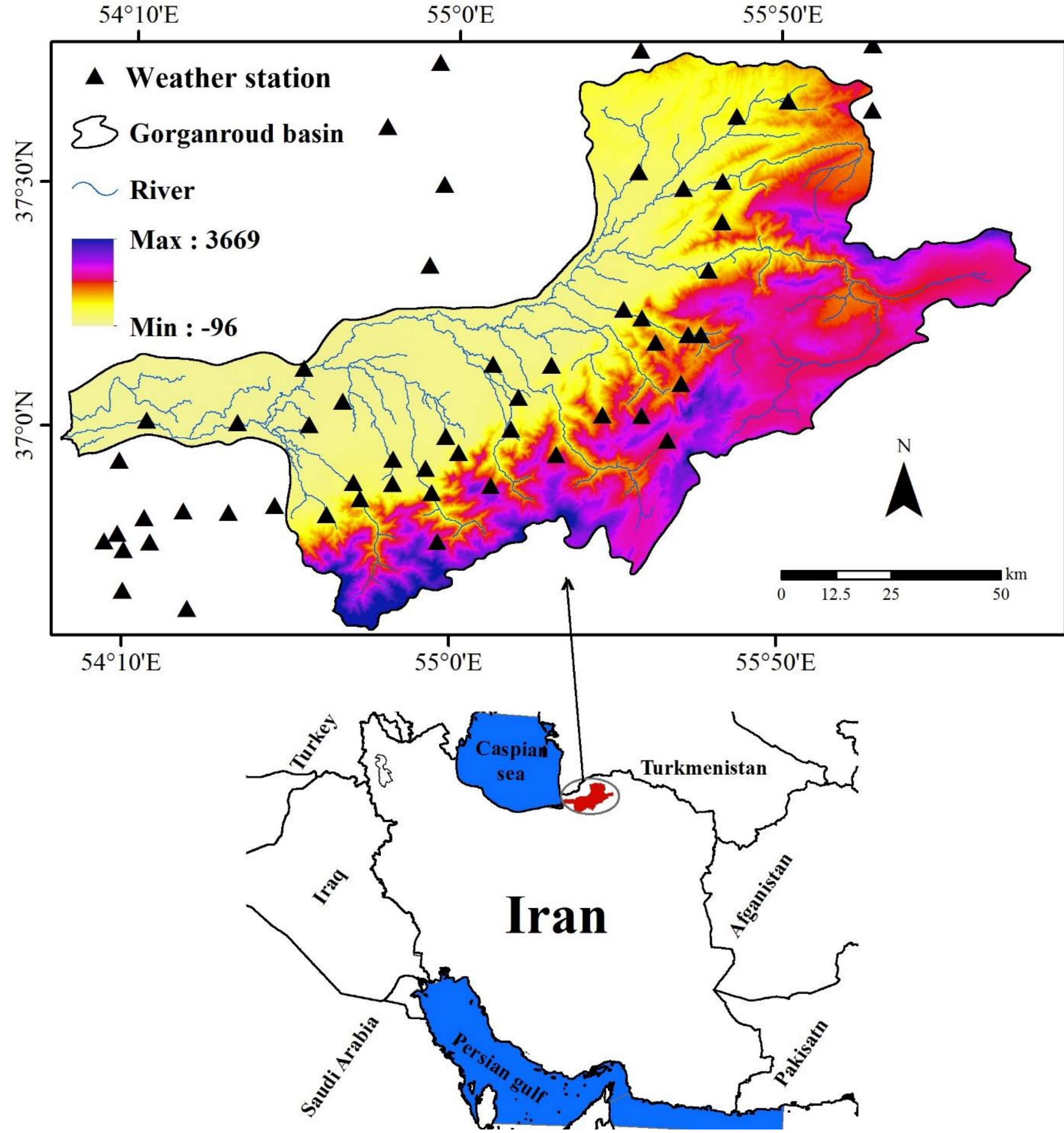


Fig. 1 Location of the Gorganroud River Basin

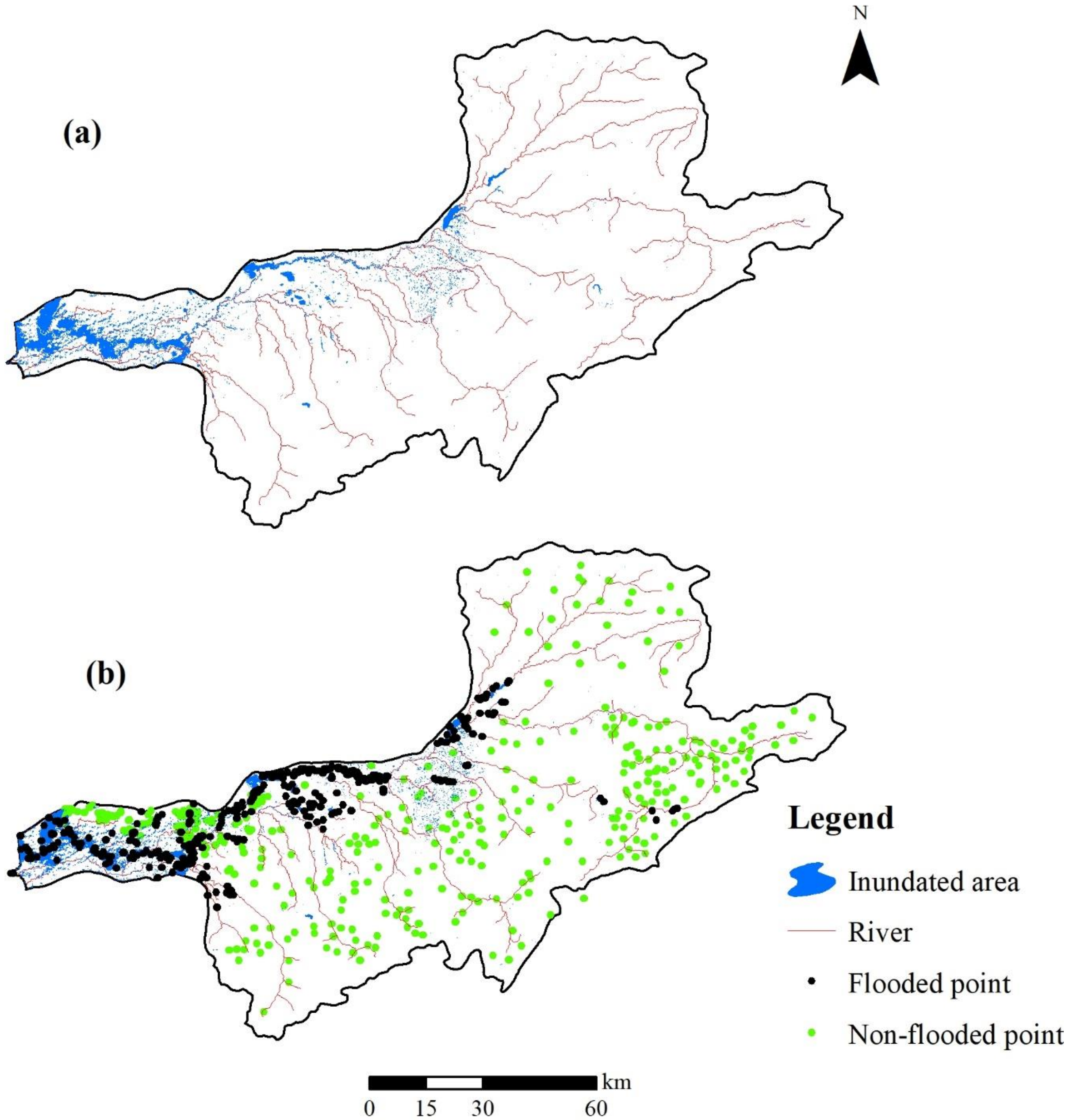
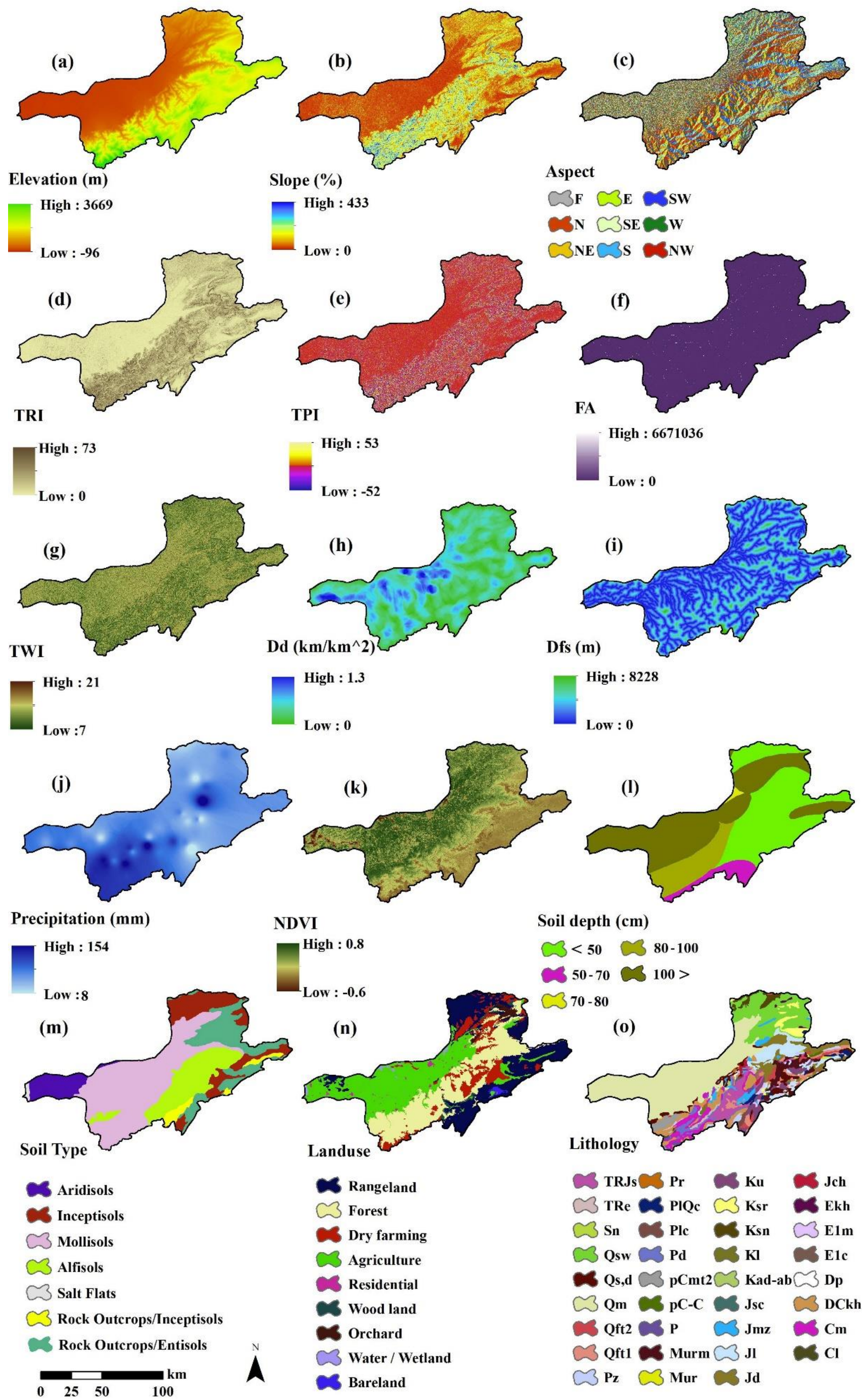


Fig. 2 Extracted inundated area by Sentinel-2 images during a period from 11 March 2019 to 10 April 2019 (a) and location of the flooded and non-flooded points (b)

Fig. 3 Flooded area in the Aqqala city (Qomnews, 2019).



(a)
Elevation (m)
High : 3669
Low : -96
(b)
Slope (%)
High : 433
Low : 0
(c)
Aspect
F
E
SW
N
SE
W
NE
S
NW
(d)
TRI
High : 73
Low : 0
(e)
TPI
High : 53
Low : -52
(f)
FA
High : 6671036
Low : 0
(g)
TWI
High : 21
Low :7
(h)
Dd (km/km^2)
High : 1.3
Low : 0
(i)
Dfs (m)
High : 8228
Low : 0
(j)
Precipitation (mm)
High : 154
Low :8
(k)
NDVI
High : 0.8
Low : -0.6
(l)
Soil depth (cm)
< 50
80 - 100
50 - 70
100 >
70 - 80
(m)
Soil Type
Aridisols
Inceptisols
Mollisols
Alfisols
Salt Flats
Rock Outcrops/Inceptisols
Rock Outcrops/Entisols
(n)
Landuse
Rangeland
Forest
Dry farming
Agriculture
Residential
Wood land
Orchard
Water / Wetland
Bareland
(o)
Lithology
TRJs
Pr
Ku
Jch
TRe
PlQc
Ksr
Ekh
Sn
Plc
Ksn
E1m
Qsw
Pd
Kl
E1c
Qs,d
pCmt2
Kad-ab
Dp
Qm
pC-C
Jsc
DCkh
Qft2
P
Jmz
Cm
Qft1
Murm
Jl
Cl
Pz
Mur
Jd
N
km
0 25 50 100

Fig. 4 Flash flood influencing factor: a) elevation, b) slope, c) aspect, d) topographic roughness index (TPI), e) topographic position index (TPI), f) flow accumulation (FA), g) topographic wetness index (TWI), h) drainage density (Dd), i) distance from stream (Dfs), j) precipitation, k) normalized difference vegetation index (NDVI), l) soil depth, m) soil type, n) land use, and o) lithology.

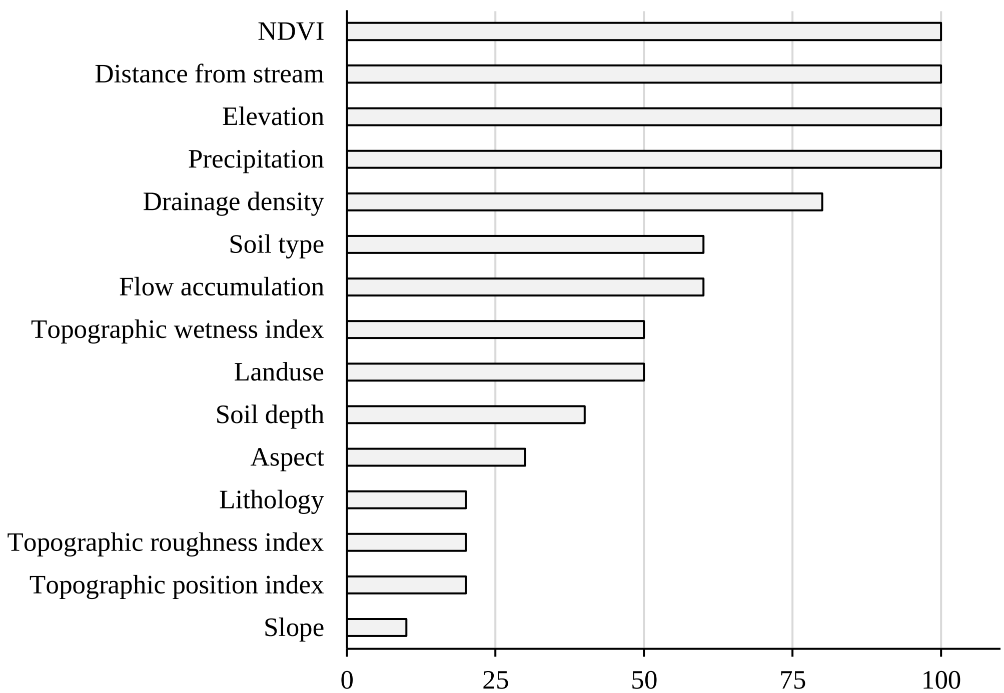


Fig. 5 The frequency (%) of selected features in all folds.

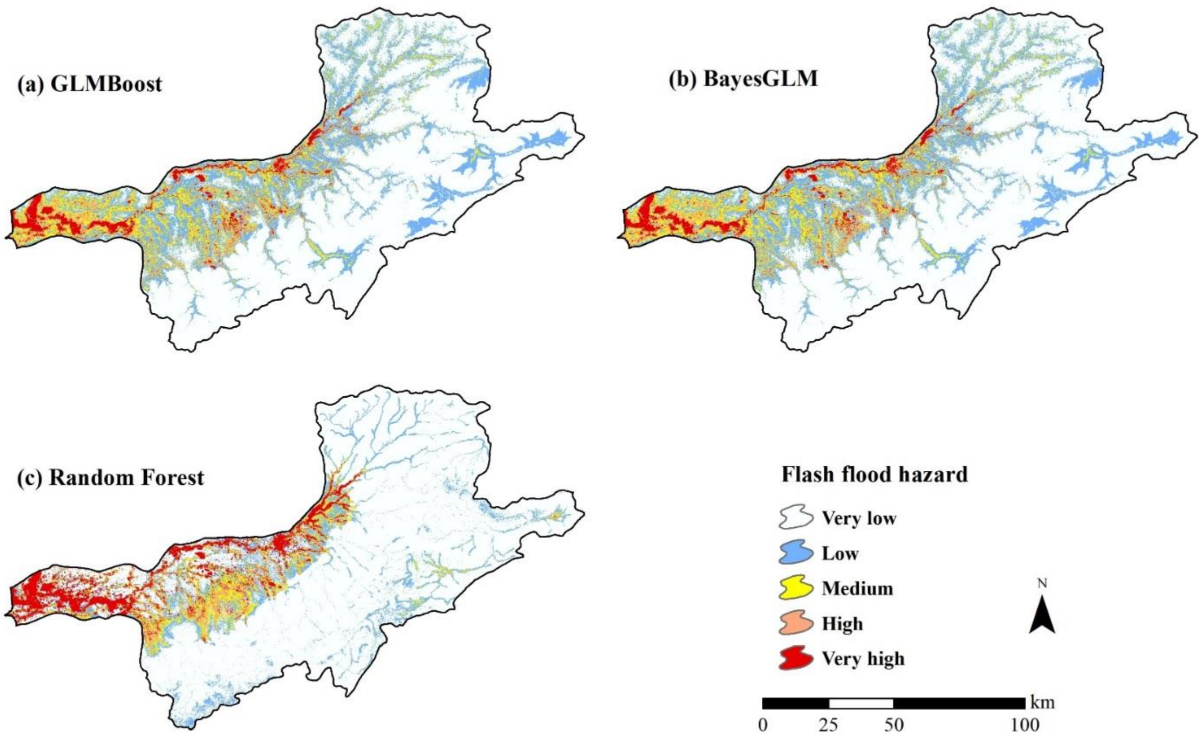


Fig. 6 Spatial prediction of flash flood hazard: (a) GLMBoost, (b) BayesGLM, and (c) Random forest

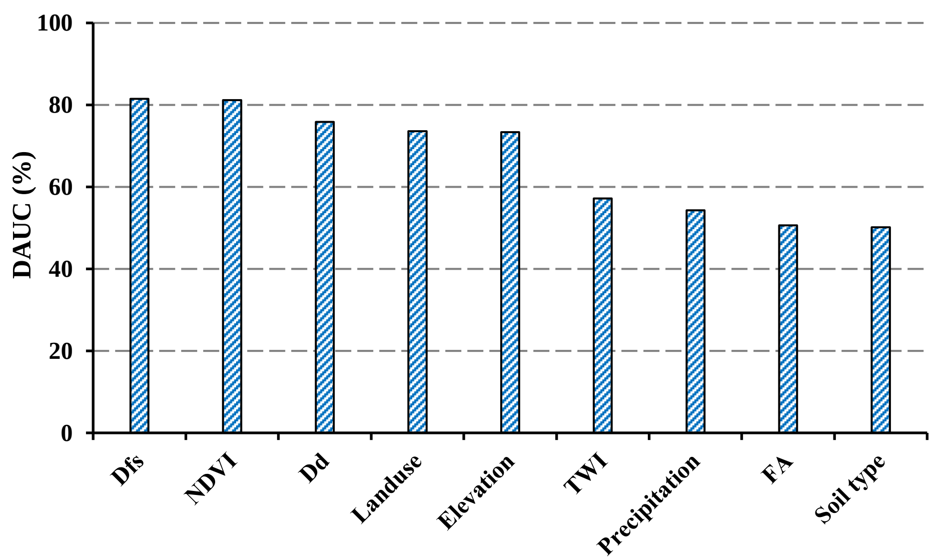


Fig. 7 Variable importance based on the BayesGLM model.

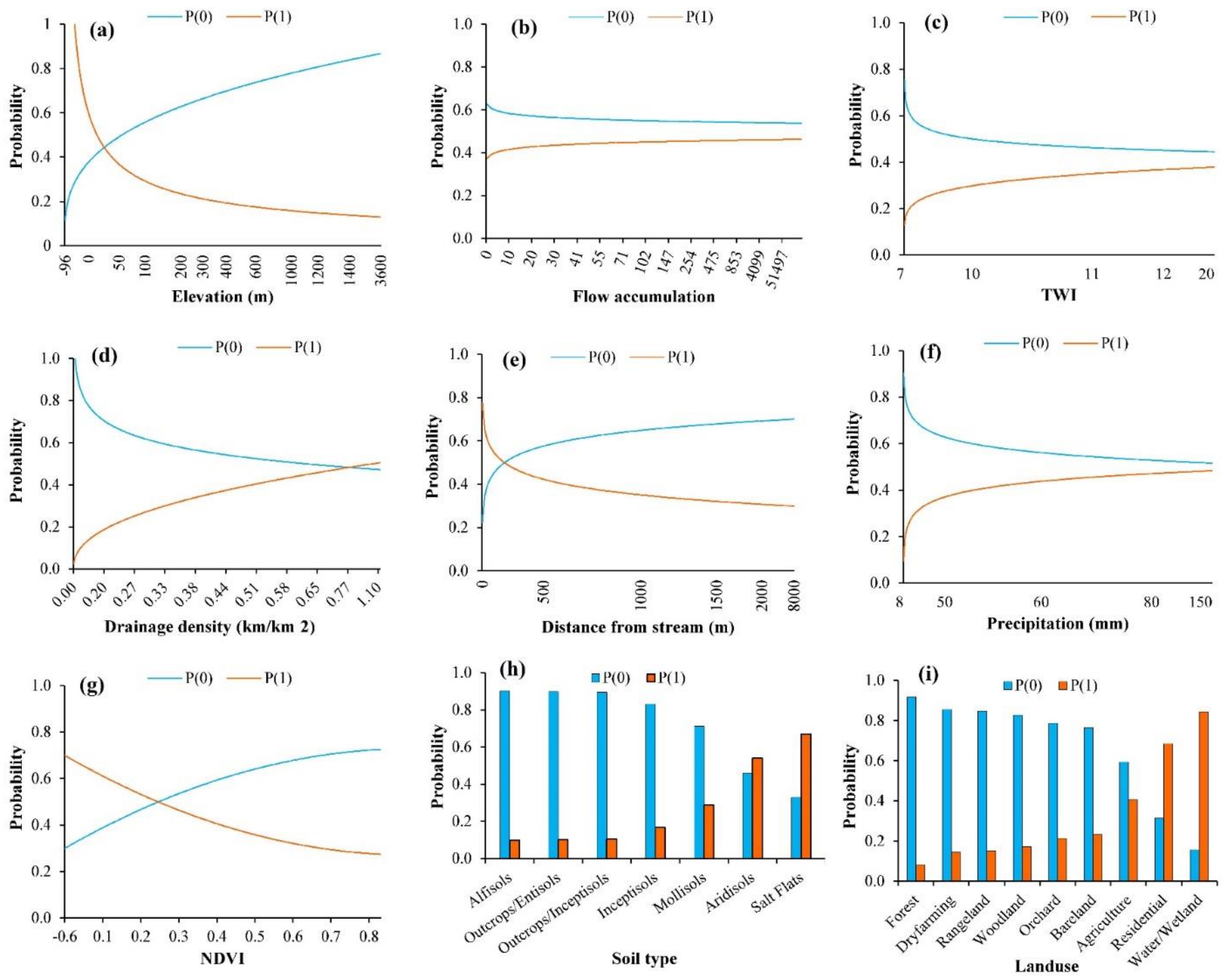


Fig. 8 GLMBoost probability curves for each variable: (a) elevation, (b) flow accumulation, (c) topographic wetness index (TWI), (d) drainage density, (e) distance from stream, (f) precipitation, (g) NDVI, (h) soil type, and (i) landuse.