Evaluation of various boosting ensemble algorithms for predicting flood hazard susceptibility areas

Quoc Bao Phama, Subodh Chandra Palb, Rabin Chakraborttyb, Akbar Norouzic,d, Mohammad Golsanh, Akinwale T. Ogunrinde, Saeid Janizadehg, Khaled Mohamed Khedherhi and Duong Tran Anhj

aInstitute of Applied Technology, Thu Dau Mot University, Thu Dau Mot City, Vietnam; bDepartment of Geography, The University of Burdwan, Bardhaman, West Bengal, India; cDepartment of Natural Engineering, Faculty of Natural Resources and Earth Science, Shahrekord University, Shahrekord, Iran; dExpert of water resource in Khoda Afarin county, East Azarbaijan Regional water company, Tabriz, Iran; eNatural Resources and Watershed Management Office, Astara, Guilan, Iran; fDepartment of Agricultural and Environmental Engineering, Federal University of Technology, Akure, Ondo State, Nigeria; gDepartment of Watershed Management Engineering and Sciences, Faculty in Natural Resources and Marine Science, Tarbiat Modares University, Tehran, Iran; hDepartment of Civil Engineering, College of Engineering, King Khalid University, Abha, Saudi Arabia; iDepartment of Civil Engineering, High Institute of Technological Studies, Mrezgua University Campus, Nabeul, Tunisia; jHo Chi Minh City University of Technology (HUTECH), Ho Chi Minh City, Vietnam

ABSTRACT

The purpose of the present study was to predict the areas affected by flood hazard in the Talar watershed, Mazandaran province, Iran, using Adaptive Boosting (AdaBoost), Boosted Generalized Linear Models (BGLM), Extreme Gradient Boosting (XGB) ensemble models, and the novel ensemble framework of deep decision trees include the Deep Boosting (DB) model. For this purpose, 14 flood conditioning variables were used as independent variables in flood hazard modeling. In addition, 130 flood points in the region were identified by field visits and available flood information, which were used as the dependent variable in modeling. The results showed that all used models have a good efficiency in predicting flood hazard. The area under curve (AUC) of BGLM, XGB, AdaBoost and DB models were 0.88, 0.87, 0.89 and 0.91, respectively, which indicated the highest efficiency of the DB model in flood hazard modeling in the study area. Relative importance of the variables showed that they have different effects in each model. Altitude and distance from the river are more important than other variables. However, these two variables have been selected as the most important variables based on machine learning models, but other variables may be influential in flood hazards.
1. Introduction

Floods are the most devastating natural disasters causing enormous threats to human lives and properties (Lee et al. 2012; Mosavi et al. 2018) and environmental socio-economic systems (Merkuryeva et al. 2015; Tehrany et al. 2015). Flood is a common disaster in the northern provinces of Iran due to high intensity rainfall, unsuitable watershed management and poor flood control structures (Tehrany et al. 2014). It is important to note that floods are part of nature. One of the main reasons for the increase of flood damage in recent years is changes in land use and residential construction in rivers and flood-prone areas. Therefore, flood hazard modeling is essential to identify flood susceptible areas of watersheds. Several factors affect the flood incidence including slope, altitude, land use, drainage network, distance from stream, and soil type (Rahmati and Pourghasemi 2017; Zhao et al. 2018). Climatic and basin factors play the most important roles in causing floods (Janizadeh et al. 2019; Pham et al. 2020). Flood control projects will benefit from understanding the climatic and basin factors that affect flood events (Merkuryeva et al. 2015). Hence, quantifying and evaluating the effects of these factors can be useful and be a high priority in flood mapping. At present time, with the advancement of the geo-information technologies such as Geographic Information Systems (GIS), Machine Learning (ML), remote sensing and statistical approaches, production of highly accurate flood prone area map is entirely achievable (Tehrany et al. 2014; Chapi et al. 2017; Zhao et al. 2018). However, these maps need to be detailed and suitable for understanding flood processes and how each factor influences the process of flood incidence, and appropriate model selection for developing the model.

Flood hazard modeling has been carried out for various areas at watersheds, regional and national scales (Shahabi et al. 2020) using different machine learning algorithms namely Generalized Linear Model (GLM; (Chapi et al. 2017), Support Vector Machine (SVM; (Tehrany et al. 2015); and Random Forest (RF Rashidi et al. 2016; Chapi et al. 2017; Chen et al. 2020); adaptive neuro-fuzzy inference systems (ANFIS) (Ahmadlou et al. 2019; Wang et al. 2019); ensemble models (Choubin et al. 2019; Chowdhuri et al. 2020). In this study, our focus was on using different boosting algorithms and comparing their performance in studying and preparing a flood susceptibility map. These algorithms have been used in many studies in various fields in recent years due to their high capabilities compared to other methods. Boosting comes with an easy-to-read and interpret algorithm, making its prediction interpretations easy to handle. The prediction capability is efficient through the use of its clone methods, such as bagging or random forest, and decision trees. Boosting is a resilient method that curbs over-fitting easily. Therefore, the present study focuses on other types of machine learning algorithms such as AdaBoost (Freund and Schapire 1997), Boosted Generalized Linear Models (BGLM) (Youssef et al. 2016), Extreme Gradient Boosting (XGB) (Chen and Guestrin 2016), Deep Boost (DP) (Cortes et al. 2014). Standard ensemble algorithms namely, AdaBoost incorporate functions selected from a base classifier hypothesis set \( H \). In successful cases, \( H \) is diminished and generated boosting stumps in which decision trees reach to depth one. For some difficult situations, simple boosting stumps are not adequate to reach a high level of accuracy due to a more complex hypothesis set, for instance, decision trees with depth limited by
certain comparatively grand number. For instance, Al-Abadi (2018) stated that AdaBoost classifier was the most suitable with respect to the statistical measures in mapping flood susceptibility in an arid region, followed by the random forest and rotation forest models. Also, Coltin et al. (2016) indicated that AdaBoost have high effectiveness, an extensively variety of conditions in flood mapping algorithms. On the other hand, some researches, namely Lee et al. (2017), Albers et al. (2016) and Wang et al. (2015), presented that machine learning models (e.g. random forest models) can be useful in flood susceptibility mapping and its analysis. As indicated above, each model may present different accuracies on the basis of its assumptions with respect to data distribution, interior factors, and main causes of development. Accordingly, there are no worldwide instructions for assessing these models, and each model has its superiority and drawbacks (Shahabi et al. 2020). However, incorporating individual models could engender more generalizable results, less sensitive and high accurate models (Araújo and New 2007; Buston and Elith 2011).

Nowadays, the use of ensemble models has been considered by researchers in various fields due to the use of different modeling algorithms that reduce the prediction generalization error (Hosseini et al. 2020; Mosavi et al. 2020). These studies show that different ensemble models have been considered by various researchers and they proved that it is appropriate for the study of flood hazard. However, the use of the deep boost ensemble model to model flood hazard has not been considered by researchers. Therefore, in this study, the main objectives are to (1) evaluate the efficiency of ensemble models Adaptive Boosting (AdaBoost), Boosted Generalized

Figure 1. Location of the case study: (a) Iran, (b) Mazandaran, Tehran and Semnan Province (c) Talar watershed.
Linear Models (BGLM), Extreme Gradient Boosting (XGB) and novel framework of deep decision trees (Deep Boosting) to model flood hazard in the Talar watershed, Mazandaran province, Iran, (2) identify the most significant independent variables affecting flood hazard in the study area, and (3) predict a flood hazard map for the Talar watershed.

2. Materials and methods

2.1. Case study

Talar watershed is one of the mountainous watersheds of Iran, which is located in the provinces of Mazandaran, Semnan and Tehran, so that most of this watershed is located in Mazandaran province. The study area is 1770 km² and is within range latitudes 35° 45′– 36° 20′ N, and longitude 52° 35′– 53° 25′ E (Figure 1). Elevation changes in this area are relatively high, so that the minimum height is 214 m and the maximum height of the area is 3541 m. Precipitation changes are inversely related to altitude, so that most of the precipitation occurs in low-lying areas of the Talar watershed and the average annual rainfall in this area is 609 mm. The main land uses in the Talar watershed include agriculture, rangeland, forest, orchard and residential areas, forest is a major part of the basin, however, due to human intervention and human misuse, forest cover has decreased in recent years (Figure 2). Destruction
and land use change have increased destructive phenomena such as floods in recent years in the region, which is one of the most important floods in the region related to the flood on 18 March 2019, which caused great loss of life and property (Shirzadi et al. 2018).

2.2 Methodology

In this study, 130 flood points within the flood-affected locations were selected by field survey and available data of the Regional Water Department of Mazandaran Province. The 130 non-flood points were prepared in ArcGIS 10.5 software by random selection. Data for modeling were divided into the training (70%) and validation (30%) randomly using caret package in R software (Choubin et al. 2019; Janizadeh et al. 2019).

The steps of the modeling framework related to AdaBoost, BGLM, XGB and novel deep decision tree (Deep Boost) is summarized in a flowchart which is presented in Figure 3.

2.2.1. Dataset preparation for spatial modeling

Flood susceptibility map is essential for development planning. A variety of flood conditioning parameters are required for constructing the flood susceptibility model (Tehrany et al. 2014). The computational GIS database is being used in a phase of implausibility that can influence the appropriate procedure (Mojaddadi et al. 2017).

In the geospatial environment, all flood-conditioning factors are considered to be completely unbiased. Here, 14 flood causing factors, namely slope, aspect, altitude, plan curvature, profile curvature, land use/land cover (LULC), drainage density, lithology, rainfall, topographic wetness index (TWI), stream power index (SPI), depth of soil, distance from river and distance from road, were used as input data to estimate...
flood prone zones (Figure 4, Table 1). The flood-related causal parameters are discussed below.

The digital elevation model (DEM) map was prepared with a pixel size of 12.5 m from the ALOSPALSAR sensor, the slope map, aspect, plan and profile curvature were prepared by this DEM.

Figure 4. Flood susceptibility affected variables: (a) altitude, (b) slope, (c) aspect, (d) plan curvature, (e) profile curvature, (f) drainage density, (g) distance from river, (h) distance from road, (i) land use, (j) lithology, (k) soil depth, (l) rainfall, (m) SPI, (n) TWI.
Environmental degradation is one of the main reasons for the increase in floods and its devastating effects. Improper construction and the creation of impervious surfaces such as buildings, streets and roads disrupt ecosystem function are increasing the risk of flooding. The distance from river and the distance from road were prepared using river and road layer in GIS software as raster layers based on Euclidean extension. Drainage density map was created using line density extension. SAGAGIS software was used to TWI and SPI layers. The TWI equation:

\[ TWI = \ln \left( \frac{as}{\tan \theta} \right) \]  \hspace{1cm} (1)

where \( as \) is the catchment area and \( \tan \theta \) represent slope in radians.

The SPI equation

\[ SPI = \frac{As}{\tan \sigma} \]  \hspace{1cm} (2)

In Equation (2), \( \sigma \) shows the slope in radian.

The soil depth map of the region was obtained from the Natural Resources Administration of the Province of Mazandaran, Iran. The lithological map was obtained from geological map of 1: 100,000 of the country mapping organization. Land use map was prepared by Landsat satellite images, OLI measurements and the Maximum Likelihood Classifier methods in the ENVI software. The interpolation method was used to prepare the precipitation map, based on observed data at seven climatological stations in Mazandaran province.
2.3. Multi-collinearity analysis

Multi-collinearity is seen as a state of interdependence that may occur apart from the origin or presence of two or more variables. Evaluation of a data set as to the nature of multi-collinearity must always serve as a unique function and step in any probability estimation (Haitovsky 1969). Scientists, statisticians and scholars from the various fields in which they conduct probabilities analysis would learn the reverse implications of multi-collinearity and create problems with symmetrical dependency detection. The main issue of multi-collinearity is the least square coefficient of optimization algorithm for variables associated with sequential linkages of major variances (Slinker and Glantz 1985). All other negative impacts are the result of these major variances: the projections are usually significant and may show signs of disagreement with known hypothetical variable characteristics; and the partial F-statistics were also highly dependent and incorrect for use in variable determination. The methods of estimation of the Tolerance (TOL) and the Variance Inflation Factor (VIF) shown as follows:

\[
TOL = 1 - R^2_J \tag{3}
\]

\[
VIF = \frac{1}{TOL} \tag{4}
\]

The VIF is a predictor that presents the application with a define how much bigger the Variance would be for multi-collinearity statistics as contrasted with orthonormal data (in which the VIF is a 1.0). If the VIF are not significantly bigger than 1.0, the multi-collinearity is no major issue (Slinker and Glantz 1985).

2.4. Machine learning algorithms

2.4.1. AdaBoost

Boosting is one of the reliable techniques that can solve the problem of classification and is mainly linked to the solution of two class classifications. Initially, this technique was introduced by Freund and Schapire (1995) in the form of AdaBoost algorithm. The directions AdaBoost is generalized from two-class to multiple class depends on the understanding or perception of AdaBoost’s performance in binary classification, which is still problematic. It has been observed that AdaBoost is similar to the forward stage wise multiplicative function algorithm, which reduces the speed of failure.

This type of technique is capable of increasing the accuracy of the predictive by integrating multiple weak classifier models (Zahid et al. 2020). The major point in AdaBoost is to optimize the distribution \((w)\) and train the set of data \((t)\) in individual iteration \((t)\) to properly assess the maximum level or unexpected outcomes. Originally, AdaBoost selects a small portion of the training sample \(D = (x_i, y_i)\) where the individual \(x_i\) instance is the vector of the character values belonging to the area \(X\) and each \(y_i\) category mark is connected to the \(x_i\) belonging to the \(Y\) and then sequentially trains the AdaBoost predictor by selecting a training set that depends on the
specific estimation of the final training. This assigns the maximum weight \((w)\) to incorrectly classified results, so that such findings would have a good classification probability for the next iteration. In addition, the weight of each iteration shall also be applied to the qualified classifier according to its accuracy. Its most accurate classification may result in higher weights. If all training data do not match any error until the maximum number of estimation methods specified is reached, this approach will be followed. Alternatively, for each classifier training, algorithms would iterate all feasible functions and, in each case, determine the loss of each function. Therefore, the optimal function is chosen as the very first poor classifier. Then role of the initial learner is to identify a fragile hypothesis \(h_t: X \rightarrow \{-1, +1\}\), which is acceptable for distribution \((D_t)\). The objective is to choose \(h_t\) in order to reduce the error \((\epsilon_t)\).

\[
\epsilon_t = P_{r \sim d_t}(h_t(x_t) \neq y_t)
\]  

\[
z_t = \frac{1}{2} \ln \left( \frac{1-\epsilon_t}{\epsilon_t} \right)
\]

Modifying the \(D\) distribution and highlighting the misclassified areas, the final equation for AdaBoost classifier is provided here.

\[
H(x) = \text{Sign} \left( \sum_{t=1}^{T} z_t h_t(x) \right)
\]

where \(h_t\) is a weak learner, the coefficient is \(z_t\) and the final hypothesis is outcome is \(H(x)\).

### 2.4.2. Boosted generalized linear models (BGLM)

The Generalized Linear Model is the extended part of the linear model which is associated with modified the regression structure to provide the ‘non-normal distribution’. The mathematical expressions for structuring the method are as follows (Youssef et al. 2016):

\[
Y = \Pr(Y = 1) = \frac{\prod e^{rX_1} \ldots \ldots e^{rX_n}}{1 + \prod e^{rX_1} \ldots \ldots e^{rX_n}}
\]

where \(Y = \pi\) was recognizable, i.e. the possibility of every other expected result to be unbalanced or stable due to some kind of combined effect of uncertainty variables. LOGIT is considered as a linear relationship for marginal reaction simulation to process the data at zero and 1. Severe principles thus, the GLM comprised of three components: (1) a conditional possibility for the response variable \(Y\), (2) a linear determinant, and (3) a link feature that offers the correlation among the linear predictor and the gaussian distribution mean (Nikita 2014; Youssef et al. 2016).

The mentioned model utilizes boosting segment-wise. Element-wise boosting indicates that one part of the prediction is fitted at a time, in this situation one conditional expression (Tutz and Groll 2010). The BGLM function enables communication
to fit linear (generalized) frameworks. It is established of a linear (nonspecific) framework of the covariates \( x = (x_1, ..., x_p)^T \) (Hofner et al. 2014):

\[
g(\mu) = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p
\]

With the (predicated) response assumption \( \mu = \mathbb{E}(y|x) \), the connection function and variables \( \beta \) (Hofner et al. 2014).

2.4.3. Extreme gradient boosting (XGB)

This method was initially proposed by Chen and Guestrin (2016) as an innovative algorithm which is considered to be an advanced gradient boosting associated with ‘K classification’ and ‘Regression tree’ With the additive procedure, it can be helpful for learners to develop a strong prediction based on a combination of weekly prediction sets. XGBoost is trying to avoid computational complexity but also to improve computing resources. This is achieved by modifying the optimization problem, which allows the combination of probabilistic and normalization terms, while at the same time maintaining optimum computing speeds, and simulations are also widely performed for the XGBoost features throughout the training process. The additive learning mechanism of XGBoost can be found elsewhere. Initially, the first learner matches the entire input storage, and these coefficients are then integrated into Model 2 to address the disadvantages of the weak learning process. This convenient process is repeated several times, until after a stop indicator has been found. The final estimation of the model is based on the total amount of the individual instructor’s prediction.

2.4.4. Deep boost (DP)

A simple boost stump sometimes cannot be considered to be more accurate to predict (Cortes et al. 2014). Then it is upsetting to provide a more complicated set of hypotheses, e.g. the pair of most decision trees, with a comparatively substantial percentage constraining deeper. However, the current training assurance for AdaBoost focuses not only on the range and number of training instances, but also on the difficulty of the calculated H interval of the VC component or the consistency of the Rademacher component (Cortes et al. 2014). From this perspective, the expansion of new hypotheses and procedures can play a key influence in the improvement of models.

2.5. Evaluation of efficiency

The performance of flood susceptibility models were validated by various statistical indices (Pourghasemi et al. 2017). These indices are positive predictive values (PPV), negative predictive values (NPV), Sensitivity (SST), specificity (SPF) and accuracy (ACC), which are measured to confirm the flood susceptibility model. All modeling steps in this study were performed using R software (R Core Team 2020). Variance inflation factors (VIFs) were computed using the usdm package version 1.1 (Naimi 2015), all models and the ensemble forecasts were performed using the caret package in R.
### Table 2. The results of multi-collinearity analysis in independent variables.

| Factors       | VIF  | TOL  |
|---------------|------|------|
| Altitude      | 3.82 | 0.26 |
| Slope         | 1.59 | 0.63 |
| Aspect        | 1.08 | 0.93 |
| Plan curvature| 1.93 | 0.52 |
| Profile curvature | 1.41 | 0.71 |
| Distance from river | 3.91 | 0.26 |
| Drainage density | 3.15 | 0.32 |
| Distance from road | 1.64 | 0.61 |
| Rainfall      | 2.98 | 0.34 |
| Land use      | 1.2  | 0.83 |
| Lithology     | 1.15 | 0.87 |
| Soil depth    | 1.45 | 0.69 |
| SPI           | 1.56 | 0.64 |
| TWI           | 2.74 | 0.36 |

### 3. Results

#### 3.1. Multi-collinearity analysis

Multi-collinearity is one of the precise methods for removing redundant variables of predictive models by considering the relationship between two or multiple independent variables (Roy et al. 2020). The TOL and VIF values are either ≤ 0.1 or ≥ 10, representing a high multi-collinearity (Amiri et al. 2019). From the multi-collinearity analysis of 14 flood causality factors, the range of VIF and TOL were 1.08 to 3.91 and 0.26 to 0.93, respectively. This indicates the maximum possibilities of efficient accuracy of the predicted models. The details of the VIF and TOL values of 14 flood causality factor are shown in Table 2. These variables are associated with permissible limit of VIF and TOL, so there is no such problem of multi-collinearity.

#### 3.2. Flood susceptibility modeling

The flood susceptibility assessment was carried out with the help of a variety of boosting ensemble algorithms, a novel deep learning framework and GIS to estimate vulnerable areas. However, there are some spatial differences between the susceptible areas of the predicted models. All raster outputs have been reclassified into different qualitative classes using a natural break method in the GIS environment (Figure 5 and Table 3).

In AdaBoost, the areal coverage of very high, high, moderate, low and very low are 152.04 (8.59%), 217.9 (12.31%), 222.66 (12.58%), 251.88 (14.23%) and 925.8 km² (52.30), respectively. A large part of the study area has very low to moderate sensitivity to floods. The very low to low susceptible areas (66.53%) are distributed in the southern parts of study area, whereas the rest of the portion (33.48%) of this region, mainly the major streams and its adjacent areas, are associated with moderate to very high susceptible areas.

In Boosted Generalized Linear Models (BGLM), the areas of very high, high, moderate, low and very low are 263.2 (14.87%), 366.6 (20.7%), 468.96 (26.49%), 422.57 (23.87%) and 248.95 km² (14.06), respectively. Most of the portion (62.06%) of this outcome is moderate to very high susceptible areas and which is concentrated in
Figure 5. Flood susceptibility map using the four models: AdaBoost, GLMboost, XGBoost and Deep boost.
Figure 5. Continued.
The remaining area of the watershed (37.93%) is associated with very low to low susceptible zones, which are located in the southern section of this study area.

### Table 3. Flood hazard classes’ area.

| Models       | Area   | Very low | Low  | Moderate | High | Very high |
|--------------|--------|----------|------|----------|------|-----------|
| AdaBoost     | Km²    | 925.8    | 251.88 | 222.66    | 217.9 | 152.04    |
|              | %      | 52.30    | 14.23 | 12.58     | 12.31 | 8.59      |
| GLM boost    | Km²    | 248.95   | 422.57 | 468.96    | 366.6 | 263.2     |
|              | %      | 14.06    | 23.87 | 26.49     | 20.71 | 14.87     |
| XG boost     | Km²    | 1124.75  | 176.79 | 113.75    | 97.73 | 257.26    |
|              | %      | 63.54    | 9.99  | 6.43      | 5.52  | 14.53     |
| Deep boost   | Km²    | 1027.01  | 312.22 | 154.19    | 156.19 | 120.67    |
|              | %      | 58.01    | 17.64 | 8.71      | 8.82  | 6.82      |

### Figure 6. The result of evaluation accuracy based on AUC for four models in testing stage.

### Table 4. Results of performance evaluation of four models used in flood hazard modeling.

| Algorithms   | Stage   | Parameters       | Sensitivity | Specificity | PPV  | NPV  | AUC  |
|--------------|---------|------------------|-------------|-------------|------|------|------|
| AdaBoost     | Training|                 | 0.89        | 0.92        | 0.91 | 0.92 | 0.94 |
|              | Testing |                 | 0.84        | 0.78        | 0.82 | 0.80 | 0.89 |
| GLM boost    | Training|                 | 0.88        | 0.77        | 0.78 | 0.88 | 0.90 |
|              | Testing |                 | 0.93        | 0.72        | 0.80 | 0.89 | 0.88 |
| XG boost     | Training|                 | 0.91        | 0.93        | 0.92 | 0.95 | 0.96 |
|              | Testing |                 | 0.77        | 0.80        | 0.82 | 0.74 | 0.87 |
| Deep boost   | Training|                 | 0.93        | 0.97        | 0.96 | 0.98 | 0.99 |
|              | Testing |                 | 0.88        | 0.86        | 0.88 | 0.86 | 0.91 |
In Extreme Gradient Boosting (XGB), the areal coverage of very high, high, moderate, low and very low are 257.26 (14.53%), 97.73 (5.52%), 113.75 (6.43%), 176.75 (9.99%) and 1124.75 km² (63.54), respectively. In this outcome, as the surface morphology and the river valley density expect the 73.53% of this region is associated with very low to low flood susceptible areas and which is located into the southern, eastern and western portions of the catchment area. The remaining 26.48% is associated with moderate to high flood susceptible areas and which is located only the major stream lines and surroundings.

In Deep Boost (DP), the areal coverage of very high, high, moderate, low and very low are 120.67 (6.82%), 156.19 (8.82%), 154.19 (8.71%), 312.22 (17.64%) and 1027.01 km² (58.01), respectively. In this model, 75.65% of this region are associated with very low to low flood susceptible areas and which is located into the southern, eastern and western portions of the catchment area. The moderate to high flood susceptible areas which are located only along the major stream lines and surroundings account for 24.35% of the area.

### 3.3. Evaluation of models accuracy

The evaluation of the predictive models was determined by using the following statistical parameters: Sensitivity, Specificity, PPV, NPV, and AUC. The AUC values of ROC of AdaBoost, GLM boost, XG boost and Deep boost are 0.89, 0.91, 0.88 and 0.87, respectively (Figure 6). From this analysis, it was observed that all the predictive methods are showed high accuracy and successfully predict the flood susceptibility (Table 4). The Deep boost model is the most efficient among used models according to its predictive ability. The sensitivity of AdaBoost, GLM boost, XG boost and Deep boost for training data sets are 0.89, 0.88, 0.91 and 0.93, respectively. The sensitivity of AdaBoost, GLM boost, XG boost and Deep boost for validation data sets are 0.84, 0.93, 0.77 and 0.88, respectively. The specificity of AdaBoost, GLM boost, XG boost and Deep boost for training data sets are 0.92, 0.77, 0.93 and 0.97, respectively. The PPV of AdaBoost, GLM boost, XG boost and Deep boost for training data sets are

![Figure 7. Chart of variables importance affecting flood susceptibility in Ada boost, GLM boost, XG boost, and Deep boost models.](image)
0.91, 0.78, 0.92 and 0.96, respectively. The PPV of AdaBoost, GLM boost, XG boost and Deep boost for validation data sets are 0.82, 0.80, 0.82 and 0.88, respectively. The NPV of AdaBoost, GLM boost, XG boost and Deep boost for training data sets are 0.92, 0.88, 0.95 and 0.98, respectively. The NPV of AdaBoost, GLM boost, XG boost and Deep boost for training data sets are 0.80, 0.89, 0.74 and 0.86, respectively. From the statistical analysis, we found that the Deep boost model is the most efficient model, followed by AdaBoost, XG boost and Deep boost algorithms (Table 4).

3.4. Variables importance analysis

The maximum importance of the variables in the AdaBoost model was found to be altitude (100), distance from river (75.02), drainage density (68.11) and distance from road (51.17), while the moderate to minor importance is associated with variables such as soil depth (26.22), slope (24.27), plan curvature (21.87), TWI (20.13), lithology (13.12), profile curvature (11.41), land use (2.37) and aspect (1.11). In the case of SPI, no such importance was found in the modeling of the susceptibility of floods.

In the GLM boost model, the maximum importance of the variables can be found in the drainage density (100), profile curvature (46.66) and TWI (36.45), while the moderate to less importance are associated in the variables, such as slope (27.21), lithology (9.88), soil depth (7.55), altitude (6.64) and distance from road (1.5). In the case of aspect, plan curvature, distance from river, rainfall, land use and SPI, there is no such importance was observed in flood susceptibility modeling.

In the XG boost model, the maximum importance of the variables was found only in the altitude (100) variable, while the moderate to less importance is associated in the other variables, such as distance from river (23.78), distance from road (19.97), plan curvature (10.78), slope (7.24), aspect (5.73), drainage density (5.53), lithology (5.07), TWI (4.95), rainfall (4.94), profile curvature (3.83), soil depth (3.04) and SPI (0.25). In the case of land use, there is no importance was observed in flood susceptibility modeling.

In the Deep Boost model, the maximum importance of the variables is found in the altitude (100), distance from river (66.54), drainage density (51.65), distance from road (42.26), soil depth (36.61) and rainfall (36.45), while the moderate to less importance is associated in the variables, such as slope (18.41), TWI (15.76), lithology (13.78), plan curvature (12.45), profile curvature (9.4), SPI (6.72) and aspect (2.5). Like in the XG boost model, land use has no importance for flood susceptibility modeling (Figure 7).

4. Discussion

In the present study, we evaluated the efficiency of various boosting algorithms including AdaBoost, BGLM, XGB, and Deep Boosting for flood susceptibility mapping in the Talar watershed, Mazandaran province, Iran. Based on evaluation criteria, the Deep boost model is the most efficient, followed by AdaBoost, XG boost, and Deep boost algorithms. The results of variable importance analysis of the most
accurate model (i.e. Deep Boost model) showed that the important variables are altitude, distance from the river, drainage density, and distance from the road.

The focus of this research is to introduce an efficient model for flood hazard evaluation based on different boosting ensemble algorithms in the Talar watershed, Mazandaran province, Iran. For increasing the accuracy of the predictive models, different ensemble hybrid approaches were considered. From this outcome, the Deep boost model indicated the maximum predictive accuracy followed by AdaBoost, GLM boost, and XG boost, respectively. The efficiency of the Deep Boost model has been proven in different studies in various fields (W. Chen et al. 2021; Y. Chen et al. 2021).

The reason for higher accuracy of Deep boost model is the positive importance of the conditioning factors for the flood hazard assessment, which is much higher than other models. Apart from this, the Deep boost algorithm have special features that easily increases the efficiency and accuracy of the model. Deep boost model is one of the ensemble models that in this model ensemble of deep decision trees or members of some other very rich or complicated families can make out in achieving a higher efficiency level (Cortes et al. 2014; Y. Chen et al. 2021; Y. Chen et al. 2021). The efficiency of the classifier is one of the determining elements of model accuracy. Independent sample estimation is one of the major issues of model performance regarding accuracy and predictive capability. The main advantage of a hybrid algorithm is that the predictive ability is always higher than the single model. From the previous studies, it can be said that the hybrid ensemble approaches are more efficient than any single model, so our main objective is to propose the most suitable model that can predict flood susceptibility.

In this study, the importance of variables affecting floods hazard was evaluated. Investigation of importance value shows that altitude, distance from river, and distance from road are very effective variables in modeling flood hazard in this study area. The results of the flood susceptibility maps of the Talar watershed show that low elevation areas of the watershed, which often have a low slope and are close to the watershed outlet, are highly sensitive to floods. The importance of altitude has been proven in other studies in the field of flood risk modeling in other studies (Khosravi et al. 2019; Sahana and Patel 2019; Das and Pal 2020; Yariyan et al. 2020; Kalantar et al. 2021). In accordance with the present results, Kalantar et al. (2021) demonstrated that altitude and distance from the river are the most important variables in flood hazard modeling in Brisbane, Australia. Physiographic characteristics such as altitude directly affect the hydrological regime and indirectly affect the climate of the region. In other words, physiographic and topographic conditions have a decisive effect on hydrological characteristics and water regime such as flood volume, runoff coefficient, intensity and weakness of flood discharge (Chen et al. 2020; Dodangeh et al. 2020). Another factor affecting flood risk in the study area is the distance from the river. The results of the flood susceptibility studies indicated that areas close to the river are highly sensitive to flood risk. Various researchers have shown in their studies that one of the important factors affecting the flood risk is the distance from the river and most areas near the river are highly sensitive to flood risk (Arabameri et al. 2020; Mosavi et al. 2020). In this study, the distance from the road
was also an important factor in modeling flood risk. Due to the fact that road construction creates areas with less water permeability than other areas, it is an important factor in causing floods (Pesaresi et al. 2016).

Flood is one of the major problems in every region of the world and it is associated with extreme loss of the property, economy and livelihood of the people (Few 2003). In this perspective, various methods were developed for estimating the flood susceptibility and the main objectives of those models are the same, i.e. predicting this scenario at more optimal level. Most researchers have tried to incorporate new algorithms to develop new models or create hybrid models from the existing models (Nourani et al. 2014). Floods cannot be completely avoided but we can minimize its effect by introducing the appropriate measures. For this purpose, the identification of vulnerable areas in accurate manners is necessary and first consideration for planning purpose. As a result, the use of boosting models (e.g. Deep Boost) has a high accuracy in spatial prediction and modeling, which is suggested as an ideal method in future studies.

5. Conclusion

In this study, four boosting ensemble algorithms were considered for flood hazard modeling. According to evaluation criteria, the deep boost ensemble approach is the most efficient among the other approaches in this study. All models had high accuracy, and 24.35% area of this catchment is connected with moderate to very high flood hazard zones. Based on result of Deep Boost model, the maximum importance of the variables is found in the altitude, distance from river, and drainage density. The main finding of present study is identification of flood susceptible areas. In recent years, due to human activities and environmental changes, the rate of natural disasters such as floods in Iran has increased. One of the main ways to reduce flood risk is to assess flood prone areas in different watersheds. Therefore, the outputs of this investigation can help planning of future development plans for flood mitigation in the Talar watershed. In terms of methodology, researchers are trying to assess the spatial modeling of flood hazard using different models in a number of studies and recommend an optimum one. The results of this study recommend Deep Boost algorithm, not only in the future study of flood hazard, but also in other fields of spatial modeling because of high accuracy of this algorithm.

Acknowledgements

The authors extend their thanks to the Deanship of Scientific Research at King Khalid University for funding this work through the large research groups under grant number RGP. 2/173/42.

Authors’ contributions

Quoc Bao Pham, Saeid Janizadeh, Subodh Chandra Pal conceptualized and performed the analysis; Saeid Janizadeh and Mohammad Golshan acquired the data; Akbar Norouzi, Mohammad Golshan, Subodh Chandra Pal, Rabin Chakrabortty, Duong Tran Anh and Akinwale T.
Ogunrinde wrote the manuscript and discussion, and analyzed the data; Subodh Chandra Pal, Quoc Bao Pham and Khaled Mohamed Khedher provided technical sights, as well as edited, restructured, and professionally optimized the manuscript. All authors discussed the results and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Consent to participate
All authors have participated the manuscript and agree with submission to Geomatics, Natural Hazards and Risk.

Consent to publish
All authors have approved the publication of this manuscript in the Geomatics, Natural Hazards and Risk.

Data availability statement
We have no permission to release data and codes.

Disclosure statement
The authors declare no competing interest.

Ethical approval
We confirm that this manuscript has not been published elsewhere and is not under consideration by another journal.

Funding
This research work was supported by the Deanship of Scientific Research at King Khalid University under Grant number RGP. 2/173/42.

ORCID
Quoc Bao Pham [6] http://orcid.org/0000-0002-0468-5962
Subodh Chandra Pal [6] http://orcid.org/0000-0003-0805-8007
Rabin Chakrabortty [6] http://orcid.org/0000-0002-6323-4838

References
Ahmadlou M, Karimi M, Alizadeh S, Shirzadi A, Parvinnejhad D, Shahabi H, Panahi M. 2019. Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). Geocarto Int. 34(11):1252–1272.
Al-Abadi AM. 2018. Mapping flood susceptibility in an arid region of southern Iraq using ensemble machine learning classifiers: a comparative study. Arab J Geosci. 11(9):218.
Albers SJ, Déry SJ, Petticrew EL. 2016. Flooding in the Nechako River Basin of Canada: a random forest modeling approach to flood analysis in a regulated reservoir system. Can Water Resour J/Revue Can Des Ressources Hydriques. 41(1–2):250–260.

Amiri M, Pourghasemi HR, Ghanbarian GA, Afzali SF. 2019. Assessment of the importance of gully erosion effective factors using Boruta algorithm and its spatial modeling and mapping using three machine learning algorithms. Geoderma. 340:55–69.

Arabameri A, Saha S, Chen W, Roy J, Pradhan B, Bui DT. 2020. Flash flood susceptibility modelling using functional tree and hybrid ensemble techniques. J Hydrol. 587:125007.

Araújo MB, New M. 2007. Ensemble forecasting of species distributions. Trends Ecol Evol. 22(1):42–47.

Buston PM, Elith J. 2011. Determinants of reproductive success in dominant pairs of clownfish: a boosted regression tree analysis. J Anim Ecol. 80(3):528–538.

Chapi K, Singh VP, Shirzadi A, Shahabi H, Bui DT, Pham BT, Khosravi K. 2017. A novel hybrid artificial intelligence approach for flood susceptibility assessment. Environ Model Softw. 95:229–245.

Chen Y, Chen W, Chandra Pal S, Saha A, Chowdhuri I, Adeli B, Janizadeh S, Dineva AA, Wang X, Mosavi A. 2021. Evaluation efficiency of hybrid deep learning algorithms with neural network decision tree and boosting methods for predicting groundwater potential. Geocarto Int. 1–21.

Chen Y, Chen W, Janizadeh S, Bhunia GS, Bera A, Pham QB, Linh NTT, Balogun A-L, Wang X. 2021. Deep learning and boosting framework for piping erosion susceptibility modeling: spatial evaluation of agricultural areas in the semi-arid region. Geocarto Int. 36: 1–26.doi: 10.1080/10106049.2021.1892212

Chen T, Guestrin C. 2016. Xgboost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. p. 785–794.

Chen W, Lei X, Chakrabortty R, Pal SC, Sahana M, Janizadeh S. 2021. Evaluation of different boosting ensemble machine learning models and novel deep learning and boosting framework for head-cut gully erosion susceptibility. J Environ Manag. 284(112015):112015.

Chen W, Li Y, Xue W, Shahabi H, Li S, Hong H, Wang X, Bian H, Zhang S, Pradhan B, et al. 2020. Modeling flood susceptibility using data-driven approaches of nave bayes tree, alternating decision tree, and random forest methods. Sci Total Environ. 701:134979.

Choubin B, Moradi E, Golshan M, Adamowski J, Sajedi-Hosseini F, Mosavi A. 2019. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. Sci Total Environ. 651(Pt 2): 2087–2096.

Chowdhuri I, Pal SC, Chakrabortty R. 2020. Flood susceptibility mapping by ensemble evidential belief function and binomial logistic regression model on river basin of eastern India. Adv Sp Res. 65(5):1466–1489.

Coltin B, McMichael S, Smith T, Fong T. 2016. Automatic boosted flood mapping from satellite data. Int J Remote Sens. 37(5):993–1015.

Cortes C, Mohri M, Syed U. 2014. Deep boosting.

Das B, Pal SC. 2020. Assessment of groundwater recharge and its potential zone identification in groundwater-stressed Goghat-I block of Hugli District, West Bengal, India. Environ Dev Sustain. 22(6):5905–5923.

Dodangeh E, Choubin B, Eigdir AN, Nabipour N, Panahi M, Shamshirband S, Mosavi A. 2020. Integrated machine learning methods with resampling algorithms for flood susceptibility prediction. Sci Total Environ. 705:135983.

Few R. 2003. Flooding, vulnerability and coping strategies: local responses to a global threat. Progress in Development Studies. 3(1):43–58.

Freund Y, Schapire RE. 1995. A decision-theoretic generalization of on-line learning and an application to boosting. European Conference on Computational Learning Theory. p. 23–37.

Freund Y, Schapire RE. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci. 55(1):119–139.
Haitovsky Y. 1969. Multicollinearity in regression analysis: comment. Rev Econ Stat. 51(4): 486–489.

Hofner B, Mayr A, Robinzonov N, Schmid M. 2014. Model-based boosting in R: a hands-on tutorial using the R package mboost. Comput Stat. 29(1–2):3–35.

Hosseini FS, Choubin B, Mosavi A, Nabipour N, Shamshirband S, Darabi H, Haghighi AT. 2020. Flash-flood hazard assessment using ensembles and Bayesian-based machine learning models: application of the simulated annealing feature selection method. Sci Total Environ. 711:135161.

Janizadeh S, Avand M, Jaafari A, Phong T, Van Bayat M, Ahmadisharaf E, Prakash I, Pham BT, Lee S. 2019. Prediction success of machine learning methods for flash flood susceptibility mapping in the Tafresh Watershed, Iran. Sustainability. 11(19):5426.

Kalantar B, Ueda N, Saeidi V, Janizadeh S, Shabani F, Ahmadi K, Shabani F. 2021. Deep Neural Network Utilizing Remote Sensing Datasets for Flood Hazard Susceptibility Mapping in Brisbane, Australia. Remote Sens. 13(13):2638.

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

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

Lee S, Kim J-C, Jung H-S, Lee MJ, Lee S. 2017. Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. Geomatics, Nat Hazards Risk. 8(2):1185–1203.

Merkuryeva G, Merkuryev Y, Sokolov BV, Potryasaev S, Zelentsov VA, Lektauers A. 2015. Advanced river flood monitoring, modelling and forecasting. J Comput Sci. 10:77–85.

Mojaddadi H, Pradhan B, Nampak H, Ahmad N, Ghazali AHB. 2017. Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS. Geomatics, Nat Hazards Risk. 8(2):1080–1102.

Mosavi A, Golshan M, Janizadeh S, Choubin B, Melesse AM, Dineva AA. 2020. Ensemble models of GLM, FDA, MARS, and RF for flood and erosion susceptibility mapping: a priority assessment of sub-basins. Geocarto Int. 35: 1–20.

Mosavi A, Ozturk P, Chau K. 2018. Flood prediction using machine learning models: literature review. Water. 10(11):1536.

Naimi B. 2015. USDM: Uncertainty analysis for species distribution models. R package version 1.1–15. R Documentation http://www.rdocumentation.org/packages/usdm.

Nikita E. 2014. The use of generalized linear models and generalized estimating equations in bioarchaeological studies. Am J Phys Anthropol. 153(3):473–483.

Nourani V, Baghanam AH, Adamowski J, Kisi O. 2014. Applications of hybrid wavelet–artificial intelligence models in hydrology: a review. J Hydrol. 514:358–377.

Pesaresi M, Ehrlich D, Florczyk AJ, Freire S, Kemper T, Julea A, Soille P, Syrris V. 2016. GHS built-up grid, derived from Landsat, Multitemporal (1975, 1990, 2000, 2014) (version R2015). European Commission, Joint Research Centre (JRC).

Pham BT, Avand M, Janizadeh S, Phong T, Van Al-Ansari N, Ho LS, Das S, Le H, Van Amini A, Bozychaloëi SK, others. 2020. GIS based hybrid computational approaches for flash flood susceptibility assessment. Water. 12(3):683.

Pourghasemi HR, Yousefi S, Kornejady A, Cerdà A. 2017. Performance assessment of individual and ensemble data-mining techniques for gully erosion modeling. Sci Total Environ. 609:764–775.

Rahmati O, Pourghasemi HR. 2017. Identification of critical flood prone areas in data-scarce and ungauged regions: a comparison of three data mining models. Water Resour Manag. 31(5):1473–1487.

Rashidi S, Vafakhah M, Lafdani EK, Javadi MR. 2016. Evaluating the support vector machine for suspended sediment load forecasting based on gamma test. Arab J Geosci. 9(11):583.
R Development Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Roy P, Chakrabortty R, Chowdhuri I, Malik S, Das B, Pal SC. 2020. Development of different machine learning ensemble classifier for gully erosion susceptibility in Gandheswari Watershed of West Bengal, India. In: Rout JK, Rout M, Das H, editors. Machine learning for intelligent decision science. Singapore: Springer; p. 1–26.

Sahana M, Patel PP. 2019. A comparison of frequency ratio and fuzzy logic models for flood susceptibility assessment of the lower Kosi River Basin in India. Environ Earth Sci. 78(10):1–27.

Shahabi H, Shirzadi A, Ghaderi K, Omidvar E, Al-Ansari N, Clague JJ, Geertsema M, Khosravi K, Amini A, Bahrami S, et al. 2020. Flood detection and susceptibility mapping using sentinel-1 remote sensing data and a machine learning approach: hybrid intelligence of bagging ensemble based on k-nearest neighbor classifier. Remote Sens. 12(2):266.

Shirzadi A, Soliamani K, Habibnejhad M, Kavian A, Chapi K, Shahabi H, Chen W, Khosravi K, Thai Pham B, Pradhan B, et al. 2018. Novel GIS based machine learning algorithms for shallow landslide susceptibility mapping. Sensors. 18(11):3777.

Slinker BK, Glantz SA. 1985. Multiple regression for physiological data analysis: the problem of multicollinearity. Am J Physiol Regul Integr Comp Physiol. 249(1):R1–R12.

Tehrany MS, Pradhan B, Jebur MN. 2014. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. J Hydrol. 512:332–343.

Tehrany MS, Pradhan B, Mansor S, Ahmad N. 2015. Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. Catena. 125:91–101.

Tutz G, Groll A. 2010. Generalized linear mixed models based on boosting. In: Kneib T, Tutz G, editors. Statistical modelling and regression structures. Berlin: Springer; p. 197–215.

Wang Y, Hong H, Chen W, Li S, Panahi M, Khosravi K, Shirzadi A, Shahabi H, Panahi S, Costache R. 2019. Flood susceptibility mapping in dingnan county (China) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm. J Environ Manage. 247:712–729.

Wang Z, Lai C, Chen X, Yang B, Zhao S, Bai X. 2015. Flood hazard risk assessment model based on random forest. J Hydrol. 527:1130–1141.

Yariyan P, Janizadeh S, Van Phong T, Nguyen HD, Costache R, Van Le H, Pham BT, Pradhan B, Tiefenbacher JP. 2020. Improvement of best first decision trees using bagging and dagging ensembles for flood probability mapping. Water Resour Manag. 34:1–17.

Youssef AM, Pourghasemi HR, Pourtaghi ZS, Al-Katheeri MM. 2016. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. Landslides. 13(5):839–856.

Zahid M, Chen Y, Khan S, Jamal A, Ijaz M, Ahmed T. 2020. Predicting risky and aggressive driving behavior among taxi drivers: do spatio-temporal attributes matter? Int J Environ Res Public Health. 17(11):3937.

Zhao G, Pang B, Xu Z, Yue J, Tu T. 2018. Mapping flood susceptibility in mountainous areas on a national scale in China. Sci Total Environ. 615:1133–1142.