



Flood susceptibility mapping with ensemble machine learning: a case of Eastern Mediterranean basin, Türkiye

Hüseyin Özdemir¹ · Müsteyde Baduna Koçyiğit² · Diyar Akay³

Accepted: 20 June 2023 / Published online: 1 July 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

In recent years, with the advances in remote sensing and geospatial technology, various machine learning algorithms found applications in determining potentially flooded areas, which have an important place in basin planning and depend on various environmental parameters. This study uses ensemble models of decision trees (DT), gradient boosting trees (GBT), support vector machines (SVM) and artificial neural network (ANN) algorithms to generate flood susceptibility maps of the Eastern Mediterranean Basin located in the Eastern Türkiye where intense short-duration rainfall causes flash floods with devastating effects on the densely populated coastal region and agricultural areas. Results of test-set analyses showed that all algorithms were more successful with ensemble models included compared to the models alone. Among the ensemble models created, the ensemble ANN model substantially increased performed best when sued with training and test sets. It was observed from the flood susceptibility maps that the flood areas formed by ensemble models were more distributed than those created by a single machine learning algorithm, and with the help of ensemble models, the distribution of the parameters affecting the floods at the flood points more elucidating. Furthermore, the McNemar test was applied to assess the differences between the predictions of the generated models using test data. It is concluded, in general, that ensemble models created with the ANN and GBT algorithms can help decision-makers in identifying flood susceptibility areas.

Keywords Flood susceptibility mapping · Decision trees · Support vector machines · Artificial neural network · Gradient boosting trees

1 Introduction

Floods are natural events that interrupt regional economic and social frameworks, sometimes with devastating results. It is estimated that 31% of global economic losses due to natural events are caused by flood hazards and are the

costliest natural disaster (Dano et al. 2019). Floods occur at varying intervals and durations and with varying magnitudes. Therefore, floods are very complex events that are difficult to predict. However, reliable predictions of floods in a basin are critical for engineers to design effective and safe infrastructure and guide decision-makers concerning flood risk assessment and watershed management. Watershed management is a complex task and demands integrated thinking and sound science. It involves challenges such as appropriate approaches for flood control, effects of land-use changes, and floodplain regulations, which all require understanding the frequency, duration, and magnitude of flood events (Knuepfer and Montz 2008). Studies showed that factors such as meteorological, hydrological, geomorphological and human intervention affect the occurrence and magnitude of flash floods (Roy et al. 2020). Once the effects of those parameters on the occurrence of floods and the identification of flood areas are determined, various structural and non-structural measures can be taken

✉ Müsteyde Baduna Koçyiğit
musteydebaduna@gmail.com

Hüseyin Özdemir
huseyin661194@gmail.com

Diyar Akay
diyarakay@hacettepe.edu.tr

¹ Graduate School of Natural and Applied Sciences, Gazi University, Ankara, Turkey

² Civil Engineering Department, Faculty of Engineering, Gazi University, Ankara, Turkey

³ Department of Industrial Engineering, Faculty of Engineering, Hacettepe University, Beytepe, Ankara, Turkey

to reduce or prevent flash flood events sustainably. One of these measures is to map the areas of flash flood susceptibility.

A significant challenge in determining flood areas is to identify the relations between the factors above that affect flooding. With the development of geospatial technology, different statistical methods have been extensively used in flood susceptibility studies, with GIS tools and remote sensing data (Membele et al. 2021; Pradhan and Youssef 2011; Omran et al. 2011; Koçyiğit et al. 2021). However, these studies are insufficient to create a generalized model for small areas, as they generally determine the flood potential by using the parameters of sub-basins as a reference but without considering the effect of the relationships between parameters. Various studies showed that parameters that directly or indirectly affect flood events and impact a basin are interconnected rather than independent (Ouma and Tateishi 2014; Wang et al. 2018; Khosravi et al. 2019). Thus, new approaches and methods are required to consider the effects of parameter interdependency on flood events. In the last decade or so, Multi-Criteria Decision Making (MCDM) methods that take into account the effects of complex inputs have been widely used in determining flood-prone areas (Ouma and Tateishi 2014; Roy and Blaschke 2015; Jaiswal et al. 2015; Dandapat and Panda 2017; De Brito et al. 2018; Wang et al. 2018; Dano et al. 2019; Hoque et al. 2019; Khosravi et al. 2019; Ishtiaque et al. 2019; Meshram et al. 2020).

Since flash floods occur non-linearly over a wide range of spatial and temporal scales, accurate and advanced tools and models are required to identify floodplains (Ahmadlou et al. 2019). Hence, Machine Learning (ML) algorithms with complex input and fast learning capabilities have attracted much attention from researchers. ML algorithms are generally used in flood, landslide, gully erosion susceptibility and groundwater potential mapping (Su et al. 2015; Hong et al. 2016; Rahmati et al. 2017; Naghibi et al. 2017; Ali et al. 2020). Various ML algorithms have become widespread with the development of technology for flood susceptibility mapping, where many geo-environmental parameters are effective (Abdollahi et al. 2019; Nhu et al. 2020; Band et al. 2020; Chakraborty et al. 2021).

For flood prediction, some ML algorithms stand out due to their approaches and learning styles. One of these algorithms is the Support Vector Machines (SVM) algorithm, which has been used in studies to create linear and nonlinear models with the help of various kernel functions (Tehrany et al. 2015b; Su et al. 2015; Hong et al. 2016; Abdollahi et al. 2019). Another frequently used ML algorithm is the Artificial Neural Network (ANN) algorithm that has been shown to have the capacity to learn, memorize and reveal the relationship between data. With the help

of these features, the ANN algorithm has been used to determine flood susceptibility maps (Jahangir et al. 2019; Andaryani et al. 2021; Ahmed et al. 2021). In addition to the SVM and ANN algorithms, the ML algorithm frequently used in the previous studies is the Decision Tree (DT) algorithm, which generally adapts well to a given data set due to the simple decision rules it creates (Tehrany et al. 2013; Choubin et al. 2019). When creating a model with decision trees, complex models (larger decision trees) may be required to increase the model's success. However, more complex models tend to produce lower generalization performance (Kotsiantis 2013).

To enhance the generalization ability of ML algorithms, ensemble models have been used to prepare flash flood susceptibility maps instead of a single ML model (Choubin et al. 2019). Ensemble models are generally developed by integrating several statistical or ML models and appear to give improved results in obtaining the best performance with high prediction accuracy. Various new ensemble and statistical models have been used to prepare flood susceptibility maps (Tehrany et al. 2015a; Bui et al. 2016; Mojaddadi et al. 2017; Shafizadeh-Moghadam et al. 2018; Band et al. 2020; Prasad et al. 2021; El-Magd et al. 2021).

Studies showed ensemble models have a more complex structure and generally perform better than single ML models in determining flood and landslide susceptibility areas (Bui et al. 2016; Naghibi et al. 2017; Prasad et al. 2021). However, during the application of ensemble models, a number of significant problems may arise, such as the model not being suitable for the data, the requirement of extensive time to build the model that gives the best performance, the inadequacy of modelling data that have not undergone any preprocessing, and a prolonged meta-learning process. To overcome such problems, preprocessing procedures of ML models and optimization of hyperparameters have been enhanced by the use of Automatic Machine Learning (AutoML) algorithms (Gijssbers et al. 2019). These algorithms are frequently used in artificial intelligence studies, enhancing the learning phase for a given data set using different preprocessing and optimization techniques. The auto-sklearn algorithm, developed using the scikit-learn library formed for data analysis and ML algorithms in Python language, is one of the most frequently used AutoML algorithms (Feurer et al. 2015). By using this algorithm, preprocessing a given data set, creating ensemble models and optimizing hyperparameters of the ML algorithm can be performed very quickly and efficiently.

Of natural disasters in Türkiye, floods are second only to earthquakes of in terms of loss of life and property. Floods are weather-related hazards, so climate change will most likely have substantial effects on flood occurrences, frequency and magnitude. In a recent study conducted by the

Republic of Türkiye Ministry of Agriculture and Forestry General Directorate of Water Management (GDWM), future projections of the most important climate parameters, such as temperature and precipitation obtained using ensembles of regional climate models with a spatial resolution of 10 km within the framework of the Effect of Climate Change on Water Resources Project were analyzed (GDWM 2016). When those future projections of the annual total precipitation based on basins were analyzed, it was noted that the highest changes were in the Asi, Eastern Mediterranean and Ceyhan Basins. Despite a general decrease in precipitation in the Eastern Mediterranean Basin, increasing temperatures and evaporation can result in flash floods due to short, intense bursts of rainfall, especially in densely populated areas near the coast, as noted in recent years. So, in our study, the Eastern Mediterranean Basin was chosen as the case study considering the increasing effect of climate change on flash floods and the densely populated coastal regions prone to flash floods.

Our study aims to create ensemble models for flood susceptibility mapping of the Eastern Mediterranean Basin in Eastern Türkiye using DT, Gradient Boosting Trees (GBT), SVM and ANN algorithms. In the ArcGIS 10.8.1 environment, using a Digital Elevation Model (DEM) of the basin, the parameters thought to influence floods, such as elevation, slope, aspect, profile curvature, sediment transport index (STI), stream power index (SPI), topographic wetness index (TWI), terrain ruggedness index (TRI), distance from the river, surface runoff curve number (CN), drainage density and precipitation were obtained. Our main goal is to create ensemble models with better performance and high accuracy by preprocessing the data set and optimizing the hyperparameters using the Auto-sklearn algorithm. During the comparison and optimization of the models, the success of the training and test data was evaluated regarding the area under the Receiver Operating Characteristic (ROC) curve. In addition, the McNemar test, which can test the consistency between binary variables, was applied to observe whether the models used make a difference in estimating the test data due to their unique approaches in the learning process (Dietterich 1997).

2 Study area

In our study, the Eastern Mediterranean Basin was used as a case study area (Fig. 1). The Eastern Mediterranean Basin is located in the south of Türkiye between 36° and 37° north parallels and 32°–35° east meridians and has a drainage area of approximately 22 048 km². The basin houses 2.4% of Türkiye's population and constitutes approximately 3% of the country's surface area. The rivers

in the basin are short and have steep beds except for the Göksu and Berdan Rivers, the two largest rivers in the basin. There is no alluvial floor along the river beds in the basin, and the rivers pass through narrow valleys. The average altitude varies between 0–2000 m and exceeds 3000 m at the peaks (Koçyiğit et al. 2021). Mediterranean climate is observed in many parts of the Eastern Mediterranean Basin. Regarding soil structure, 14 large soil groups are determined depending on precipitation and climate in the Eastern Mediterranean region with red and red-brown Mediterranean soil, non-calcareous brown and brown forest soil, alluvial and colluvial soils being the most common. The Eastern Mediterranean Region is rich in forest area (42.1%) while country-wide approximately 27% of Türkiye is forested.

2.1 Flood inventory map

An accurate flood inventory map is one of the prerequisites for successful flood modeling and must be carefully prepared. Some important information can also be obtained from this inventory map, such as the flood's locations, frequencies, causes, and triggers (infiltration and heavy rainfall).

We obtained data for 188 flood events from the inventory of the General Directorate of State Hydraulic Works (DSİ) in Türkiye. A total of 352 points were used for analysis and within these 164 points had no flood occurrence. We considered the values of 12 parameters to be effective for flood prediction and these were determined at those points using. The flood inventory classes for each point within the borders of the Eastern Mediterranean Basin are shown in Fig. 1.

2.2 Flood conditioning factors

To prepare a flood susceptibility map, it is essential to obtain the principal factors affecting the flood event (Tehrany et al. 2015a). In our study, a DEM with 10 × 10 m pixels was created using the relevant data and topographic maps of the basin. Elevation, slope, aspect, profile curvature, STI, SPI, TWI, TRI, distance from the river, drainage density and CN maps were obtained with the DEM model using spatial analyst tools in ArcGIS version 10.8.1. An overview of the factors used for flood susceptibility mapping is shown in Table 1. The maps of the flood conditioning factors considered in the study are given in Fig. 2.

The elevation parameter, presented in Fig. 2a, is among the most influential parameters in flood susceptibility mapping. There is an inverse relationship between altitude and flood susceptibility, as surface runoff after precipitation flows from higher elevations to lower elevations. A

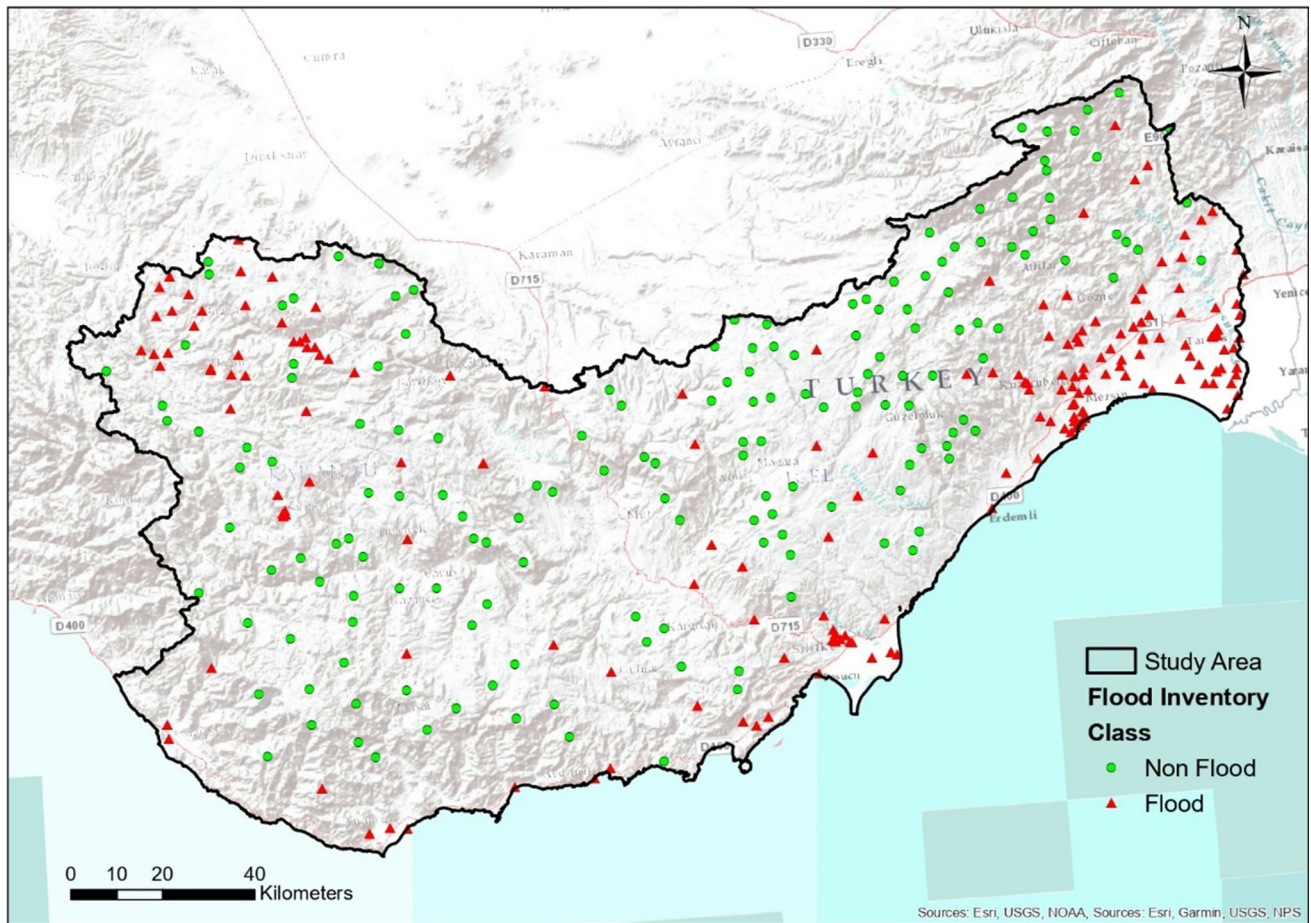


Fig. 1 Eastern Mediterranean Basin location and flood inventory

related factor that is strongly affects flood events is the slope of the land shown in Fig. 2b. As water flows from higher to lower elevations, the basin slope affects runoff and seepage rate. Hence, flat areas at low elevations can be inundated faster than areas at higher elevations with steeper slopes.

Hydrological processes, local climatic conditions, physiographic trends and soil moisture patterns are affected by aspect (Rahmati et al. 2016). We divided the aspect parameter into nine categories: Flat, North, Northeast, East, Southeast, South, Southwest, West and Northwest, as shown in Fig. 2c. Since curvature supports projections over water depths and model calibration, its inclusion is thought to be useful for accurately representing velocity (Horritt 2000). We delineated the profile curvature parameter as flat, convex and concave, as shown in Fig. 2d.

A geometric parameter that can be used to determine floodplains is the distance from the area to the river. The flood potential is predicted to decrease as the distance from the river increases since floods generally occur with the overflow of water around the rivers. The distance map from

the river was obtained using the drainage network generated in the ArcGIS environment, as shown in Fig. 2e.

SPI, shown in Fig. 2f, is an important factor in determining the flood potential of the region since it determines the level of the erosive power of the runoff and the degree of discharge relative to a particular area within the catchment area (Moore et al. 1991). STI is another critical parameter considered effective in flood formation as shown in Fig. 2g. During flooding, sediment transport occurs in the river bed due to high value of water velocity and water power. Indication of the sediment density to be transported on the map provides guidance on the sensitivity of the flood. TWI, shown in Fig. 2h combines local upslope contributing area and slope and is widely used to measure topographic control over hydrological processes (Sørensen 2006). TWI is a function of slope and upstream contributing flow area per unit width perpendicular to the flow direction. Areas with high wetness values are more exposed to flooding than areas with low TWI values (Samanta 2018).

TRI, a parameter developed by Riley et al. (1999), determines topographic heterogeneity. TRI refers to the

Table 1 An overview of the factors used for flood susceptibility mapping

Factors	Resolution	Obtaining
Elevation	10 × 10 m	ArcGIS 10.8.1 → DEM
Slope	10 × 10 m	DEM → Spatial Analyst Tools → Surface → Slope
Aspect	10 × 10 m	DEM → Spatial Analyst Tools → Surface → Aspect
Profile curvature	10 × 10 m	DEM → Spatial Analyst Tools → Surface → Curvature
SPI	10 × 10 m	DEM → Spatial Analyst Tools → Raster Calculator → f* (Flow Accumulation, Slope) → SPI
STI	10 × 10 m	DEM → Spatial Analyst Tools → Raster Calculator → f* (Flow Accumulation, Slope) → STI
TWI	10 × 10 m	DEM → Spatial Analyst Tools → Raster Calculator → f* (Flow Accumulation, Slope) → TWI
TRI	10 × 10 m	DEM → Spatial Analyst Tools → Raster Calculator → f* (DEM) → TRI
Distance from river	10 × 10 m	DEM → Terrain Preprocessing → Drainage Line → Spatial Analyst Tools → Distance → Euclidean Distance → Distance from River
Drainage density	50 × 50 m	DEM → Terrain Preprocessing → Drainage Line → Spatial Analyst Tools → Density → Line Density → Drainage Density
CN	25 × 25 m	DEM → Land-use and Soil Data → CN
Rainfall	10 × 10 m	Rainfall Data → Spatial Analyst Tools → Interpolation → IDW → Rainfall

*Where f represents the raster calculator, one of the spatial analyst tools available in ArcGIS

elevation difference between adjacent cells in the elevation model. A small TRI value represents a flat land surface, while a higher value represents an extremely uneven surface. Thus, a small TRI value in flat terrain is more susceptible to flooding than areas with a higher TRI value (Ali et al. 2020). Figure 2i shows the TRI map of the basin.

Drainage density is another important factor that has a direct impact on flooding. In general, areas with higher stream density are more prone to flooding. Drainage density can be expressed by dividing the total river length in a drainage basin by the total area of the drainage basin (Ali et al. 2020). The map of drainage density is given in Fig. 2j.

Another parameter affecting floods is the precipitation parameter. Precipitation is the most critical factor in creating a flood, and precipitation amount is the key factor for flooding. However, it is not certain how much an increase in precipitation will result in an increase in flooding. A basin map showing the annual precipitation data over the region was obtained using the recorded data at the stations operated by the Turkish State Meteorological Service. We chose annual precipitation as the influential factor in flood susceptibility mapping over different regions as shown in Fig. 2k.

Another important parameter in flood susceptibility is land use. Land use affects surface runoff and sediment transport and, consequently, the frequency of flooding. Instead of using soil and land use data as two different features in the models, the surface flow curve number (CN), as shown in Fig. 2l was obtained with based on these

two parameters and used in our models. Thus, the number of dimensions used in the model was reduced, and the effect of these two critical parameters was considered.

3 Methods

Ensemble models created with different ML models have a complex structure. Therefore, instead of creating ensemble models with different ML algorithms, various ML algorithms with strong, unique approaches and features can be used to create ensemble models. Frequently preferred ML algorithms such as ANN, SVM and DT have effective hyperparameters that increase performance and enable the formation of a generalized model. Generalized models with higher performance can be attained with ensemble models created by adjusting models with different preprocessing processes and hyperparameters. Hence, we chose the DT, GBT, SVM and ANN models of the ML algorithms for our study. Another reason for selecting those algorithms is that they are available in the Auto-sklearn algorithms (Feurer et al. 2015).

We used k-fold cross-validation to test the accuracy of the trained models and to determine whether or not any independent data set entered into the model generally represented the model (Kohavi 1995). Figure 3 shows the process we applied.

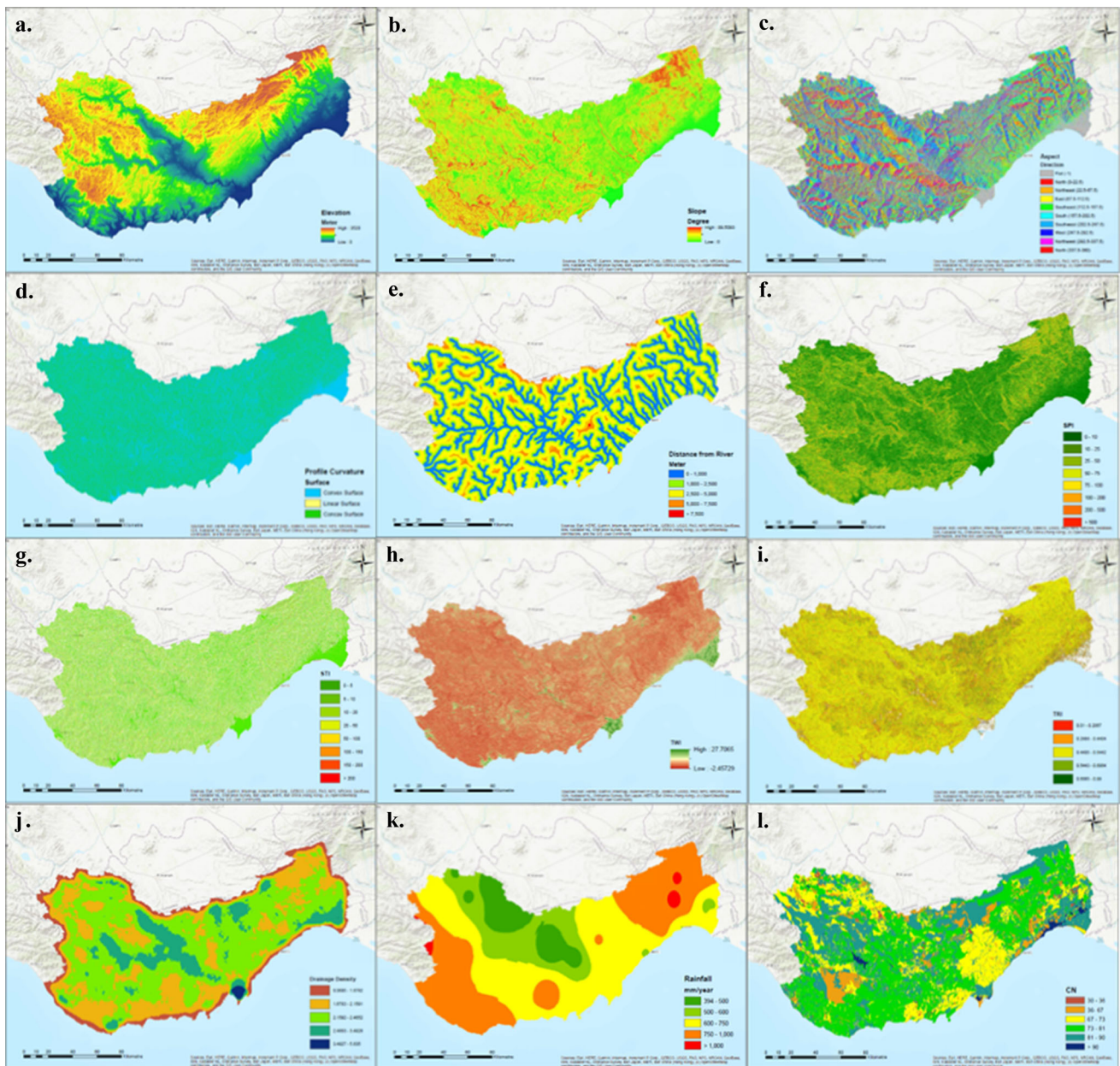


Fig. 2 Flood conditioning factors: **a** Elevation, **b** Slope, **c** Aspect, **d** Profile curvature, **e** Distance from the river, **f** SPI, **g** STI, **h** TWI, **i** TRI, **j** Drainage density, **k** Rainfall, **l** CN

3.1 Support vector machines (SVM)

SVM is a useful machine learning algorithm for two-group classification problems. SVM creates hyperplanes by separating the training set in the best way possible. It searches for the plane that makes the best distinction between these two classes according to the hyperplanes it creates (Cortes and Vapnik 1995).

For the given $\{(x_i, y_i)\}_{i=1,2,\dots,n}$ training set, SVM distinguishes two classes using Eq. 1 to identify the optimal hyperplane.

$$\min_{w,b,\xi} : \frac{\|w\|^2}{2} + c \sum_{i=1}^n \xi_i \quad (1)$$

where w is the parameter of the hyperplane equation and b is the offset of the hyperplane from the beginning. In the classification model created, some features may be misclassified. These misclassified features can be penalized with a specific penalty point, ξ_i in Eq. 1, representing the error of the misclassified feature, where c represents the penalty coefficient for situations that exceed the

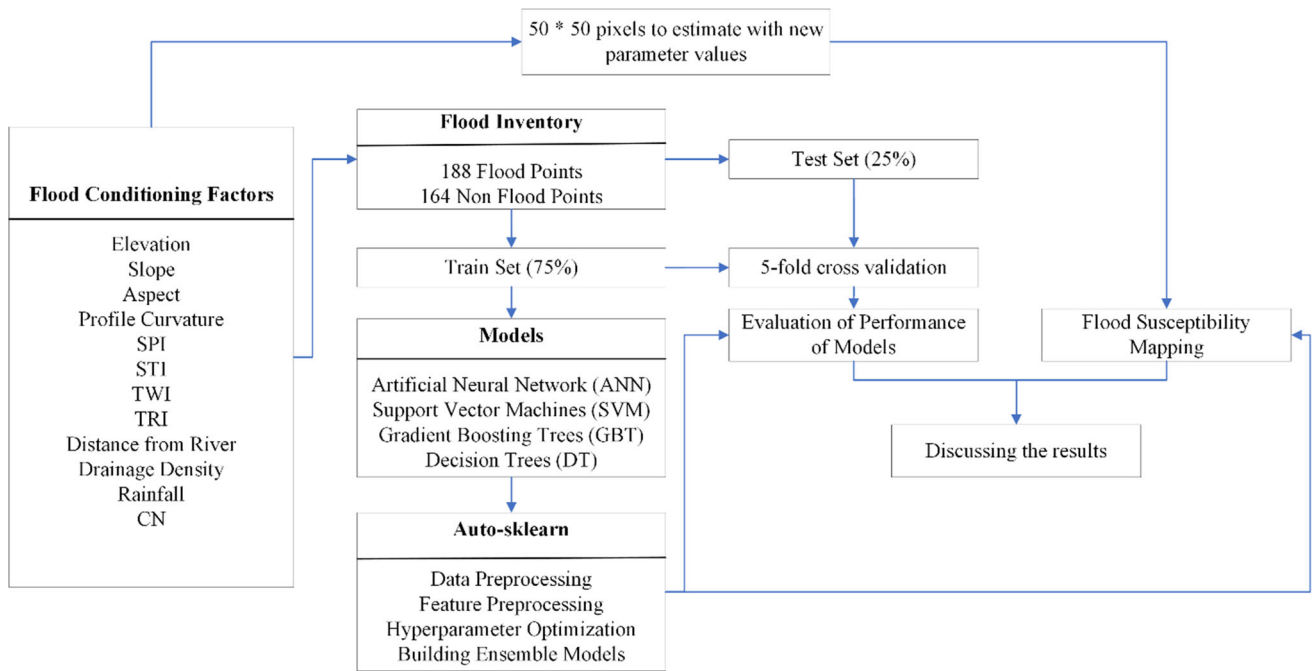


Fig. 3 A flow chart representing the general modeling strategy

constraints. The optimization problem specified by Eq. 1 must satisfy the constraint conditions given by Eq. 2.

$$y_i(w \cdot \varnothing(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0 \text{ for } i = 1, 2, \dots, n \quad (2)$$

where \varnothing is an unknown and nonlinear mapping function. The separation plane used in Eq. 1 is linear, however, due to effects uncertainty, it is an inadequate approach to accept only linear hyperplanes (Özdemir 2022). Therefore, different hyperplane approaches were tested using the nonlinear vector function K (kernel functions). The Kernel functions we are given in Eqs. 3–6 (Erdem et al. 2016; Pedregosa et al. 2011).

$$K(x_i, x_j) = x_i^T x_j \text{ LinearKernel} \quad (3)$$

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \text{ PolynomialKernel} \quad (4)$$

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \text{ RadialBasisKernel} \quad (5)$$

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r), \gamma > 0 \text{ SigmoidKernel} \quad (6)$$

where x_j is the parameter vector of the hyperplane equation, x_i^T is the transpose of the parameter matrix, r is a free parameter adjusting the influence of higher-order lower-order terms in the polynomial, γ (gamma) is a constant, and d is the degree of the polynomial kernel function.

3.2 Artificial neural networks (ANN)

ANN is an information-processing technology inspired by the information-processing technique of the human brain. Neurons form networks by connecting to each other in various ways. These networks have the capacity to learn, memorize and reveal the relationship between data. Given a set of features and a target, a neural network can construct a nonlinear model for classification or regression. Unlike the Logistic Regression algorithm, there may be one or more nonlinear layers, called hidden layers, between the input and output layers.

Learning in a neural network occurs by changing the connection weights after each data is processed, depending on the expected result according to the amount of error in the output. The model’s weights are developed by back-propagation, a generalization of the least mean squares algorithm.

For the given $\{(x_i, y_i)\}_{i=1,2,\dots,n}$ training set, the error of the output node can be expressed by Eq. 7 as the difference between the estimated and actual values.

$$e_j(n) = y_j(n) - \hat{y}_j(n) \quad (7)$$

Here \hat{y}_j is the estimated value and y_j is the actual value. Then, node weights can be adjusted according to the corrections that minimize the error value ξ in the entire output given by Eq. 8.

$$\xi(n) = \frac{1}{2} \sum_j e_j^2(n) \quad (8)$$

Using the gradient descent method, the change in each weight can be expressed by Eq. 9 (Rumelhart et al. 1986).

$$\Delta w_{ij}(n) = -\eta \frac{\partial \xi(n)}{\partial w_{ij}(n)} \quad (9)$$

Here, w_{ij} is the weight of the interconnected nodes, and η is the learning rate chosen to allow the weights to converge rapidly to a response without oscillations.

3.3 Decision trees (DT)

DT is a non-parametric supervised learning method used for classification and regression. The goal is to build a model that predicts the value of a target variable by learning simple decision rules derived from data features. DT learns from data to approximate a training set with a set of if-then-else decision rules. The deeper the tree, the more complex the decision rules and the more appropriate the model will be (Quinlan 1986). DT iteratively partitions the feature space so that samples with the same characteristics or similar target values are grouped. DT makes classification regarding any cross-entropy, Gini index or misclassification error.

3.4 Gradient boosting trees (GBT)

GBT is a machine learning technique for regression and classification problems. It generates a prediction model, which creates a group of weak prediction models as a decision tree using ensemble models. When the decision tree is a weak learner, gradient boosting, like other boosting methods, iteratively combines the weak learners into a single strong learner (Friedman 1999). It builds the model stepwise as other boosting strategies do and generalizes it by allowing optimization of an arbitrary differentiable loss function.

For the given $\{(x_i, y_i)\}_{i=1,2,\dots,N}$ training set, as in all other ML methods, the GBT algorithm aims to optimize the losses that will occur during the training of the models. The objective function to be optimized is expressed by Eq. 10.

$$obj = \sum_i^n L(y_i, \hat{y}_i^{(t)}) \quad (10)$$

where obj is the objective function to be optimized, $\hat{y}_i^{(t)}$ is a matrix of the predictions of the model created in t^{th} steps, y_i is its true value, and L is the loss function. Since it would be difficult to learn by optimizing all trees at once, a strategy that corrects what is learned at each step and adds

it to the next tree can be implemented. This additional strategy is illustrated by Eq. 11.

$$\hat{y}_i^{(t)} = \sum_{k=1}^t f_k(x_i) = \hat{y}_i^{(t-1)} + \eta f_t(x_i) \quad (11)$$

where f_t represents the function of the tree in t^{th} step, and η represents the learning rate that considers the contributions of the generated weak trees. Generally, the first-order Taylor series approach given by Eq. 12 can be used to optimize the loss function quickly (Pedregosa et al. 2011).

$$L(y_i, \hat{y}_i^{(t-1)} + \eta f_t(x_i)) = L(y_i, \hat{y}_i^{(t-1)}) + \eta f_t(x_i) g_i \quad (12)$$

where g_i is the first derivative of the loss function obtained in the previous learning and is given by Eq. (13) as

$$g_i = \frac{\partial}{\partial \hat{y}_i^{(t-1)}} [L(y_i, \hat{y}_i^{(t-1)})] \quad (13)$$

An important advantage of the definition given by Eq. 12 is that the value of the objective function depends only on the gradient of the loss function obtained in the previous step.

3.5 Building of ensemble models

The same type of ML model may require different constraints, weights or learning rates to generalize different data models. Such properties are called hyperparameters and need to be adjusted so that the model can best solve classification and regression problems. Therefore, it is necessary to continuously adjust the hyperparameters, train a set of models with different combinations of values, and then compare the model performance to select the best model (Wu et al. 2019). Using Auto-Sklearn 0.14.7, an AutoML algorithm that efficiently supports the creation of new ML applications, such optimization problems can be solved more easily (Feurer et al. 2015).

Raw data may not be in a format to which all algorithms can be applied. In ML applications, it is necessary to optimize the hyperparameters of the model to be used, as well as preprocessing data and its features. Preprocessing includes improving the performance of ML models by performing operations such as detecting and correcting corrupt or incorrect records, reviewing and adjusting data, and simplifying and correcting complex data.

Auto-sklearn has different data and feature preprocessing approaches. Data preprocessing consists of 4 main steps: scaling inputs, assigning missing values, categorical coding and balancing target classes. Feature preprocessing methods can be categorized as feature selection (select percentile, select rate), kernel approximation (Nystroem sampler, random kitchen sinks), matrix decomposition

(Principal Component Analysis (PCA), kernel PCA, fast Independent Component Analysis (fast ICA)), embeddings (Random tree embedding), feature clustering (Feature agglomeration), polynomial feature expansion (Polynomial feature) and methods that use a classifier for feature selection (Linear SVM, Extremely Random Trees) (Feurer et al. 2015).

3.6 Evaluation of model performance

In evaluating the performance of ML models, the metrics to be referenced should be carefully selected. One of the features of the auto-sklearn algorithm is that the metric to be taken as a reference can be selected while creating the models. The ROC curve is often used because it considers criteria such as accuracy, sensitivity and precision in the confusion matrices it creates. We evaluated the efficiency and precision of each model was the area under the curve (AUC).

During the evaluation of the models, the extent to which the models created do not differ much in their predictions for the given test set should be assessed. We used the McNemar test, frequently used on nominal data in statistics, to examine these differences. In comparing two binary classification algorithms, the test interprets whether the two models agree (or do not agree) in the same way (Dietterich 1997). The McNemar test does not indicate whether one model is more or less accurate or error-prone than the other. The McNemar test converts the estimation results of two binary classification models into a probability table and calculates the chi-square value given by Eq. 14.

$$\chi^2 = \frac{(n_{1/0} - n_{0/1})^2}{n_{1/0} + n_{0/1}} \quad (14)$$

where, $n_{1/0}$ represents the number of samples predicted by the second model as 0 while the first model predicts 1, and $n_{0/1}$ refers to the number of samples predicted by the second model as 1 while the first model predicts 0. The McNemar test can be interpreted for the corresponding probability value p at the test statistic χ^2 calculated by the test, given a choice of α significance level: If $p > \alpha$, the null hypothesis cannot be rejected, there is no significant difference between the models. If $p < \alpha$, the null hypothesis is rejected, there is a significant difference between the models. Comparing the models in our study, the value of the α significance level was taken as 0.05 for the 99% confidence interval.

4 Results and discussion

The 352 points obtained for the flood inventory were allocated as 264 training and 88 test sets according to the percentage distribution of flooded and non-flooded points. A fivefold cross-validation process was applied to verify whether the obtained training and test set adequately represent the data set.

Models created in Python 3.10 have a probability calculation feature for binary classification. Using this feature, for each model generated we were able to calculate the flooding potential for the parameter values to be defined. The Eastern Mediterranean Basin was divided into 50×50 pixels, and the parameters of each pixel were calculated with ArcGIS. Flooding potential was calculated with the order of these pixels whose parameters are known. Each pixel that had flood potential was then transferred back to the ArcGIS environment, and a flood susceptibility map was prepared. The models created for 4 different ML algorithms were sequentially mapped, and then differences between each other were examined.

The models are given in Fig. 4. According to the AUC, the ensemble model created by the ANN algorithm had the highest success with 0.936. The model that achieved the least success in AUC was the ANN single model, which was created without preprocessing and hyperparameter optimization.

Table 2 shows the performances of the models created by single ML and ensemble ML models for training and test sets, respectively. When the results are analyzed, it is seen that the success of ensemble models created with SVM and ANN algorithms in the training set has increased substantially. The success of the models with GBT and DT algorithms in the training set are similar between single and ensemble models. In all cases the success of the algorithms increased for the test set with ensemble models compared to the models created alone.

Among the ensemble models created in this study, the ensemble ANN model had the most increased success in the training and test set. In .

Table 3, preprocesses, hyperparameters and ensemble weights of the models created by the single and ensemble ANN models are given. Ensemble ANN model, which was created by combining 3 different models with different hyperparameters such as a correction term, hidden layer size, and learning rate, used a power transformer in the data preprocessing process, feature agglomeration in the feature preprocessing process, and gave the best results in approximating the training and test sets. These results show that creating ensemble models by preprocessing and hyperparameter optimization is a valuable exercise in improving performance.

Fig. 4 The ROC curves for the test data of the models

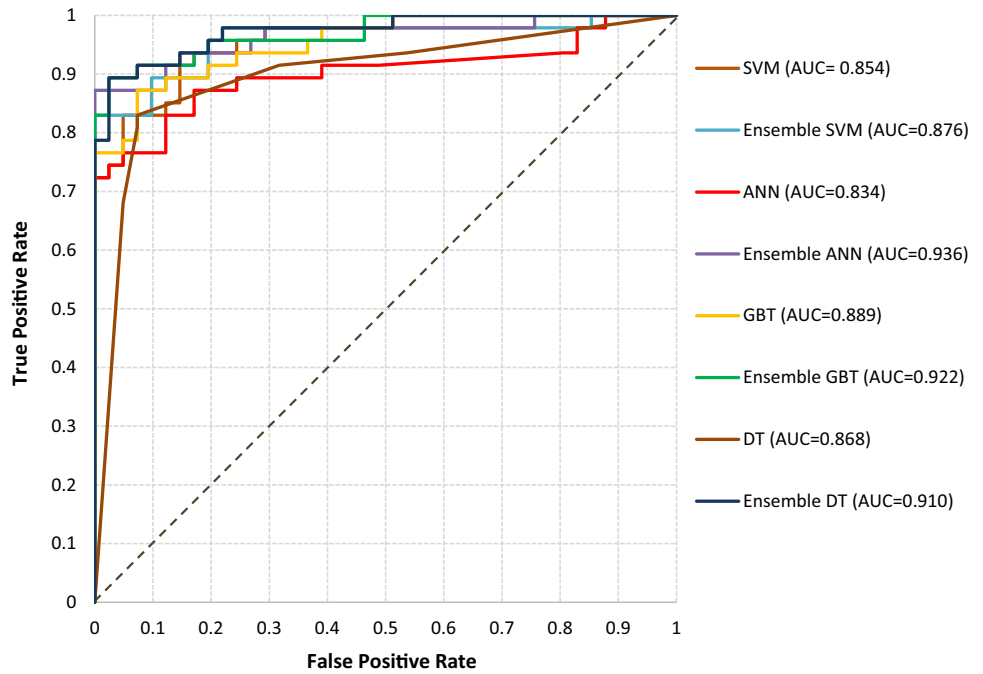


Table 2 Performance of models in training and test set

Model	Single model		Ensemble model	
	Train	Test	Train	Test
SVM	0,861	0,854	0,919	0,877
ANN	0,845	0,834	0,952	0,936
GBT	0,989	0,889	0,989	0,922
DT	0,921	0,868	0,923	0,910

The flood susceptibility maps created by ANN models shown in Fig. 5 are examined. The single ANN model generated areas with very sharp separations over the basin. Most flooding points in the inventory are close to the basin

drainage network. The single ANN model, which generally focuses on the success of the estimations, was found to be insufficient to represent the basin, in general, since it generally classifies by reference to points close to the drainage network. However, this situation differs in the flood susceptibility map obtained with the ensemble ANN model. While the ensemble ANN model predicted the flood areas, it discerned the distribution of the parameters affecting the flood in flood points more effectively with the help of combined models. This situation effectively creates areas that show a better distribution in the basin and consider the effects of all parameters on the flood. The ensemble ANN model generally determined the flood potential as high at low elevations, low slope, convex profile curvature, areas close to river drainage network,

Table 3 Preprocesses and hyperparameters of ANN models

		Single ANN	Ensemble ANN		
Ensemble weights		–	0,2	0,2	0,6
Data preprocessing	Rescaling type	–	–	Power transformer	Power transformer
	Categorical encoding	Yes	Yes	Yes	Yes
Feature preprocessing		–	Feature agglomeration	–	Feature agglomeration
Regularization term*		0,0001	0,04472	0,000455	$1,0848 \times 10^{-7}$
Hidden layer size		(38,2)	(46,1)	(30,1)	(167,1)
Learning rate		0,002248	0,02125	0,09518	0,0001767
Activation function		tanh	relu	tanh	relu
Maximum iteration		200	64	64	256

*Regularization (L2 regularization) term, which helps in avoiding overfitting by penalizing weights with large magnitudes

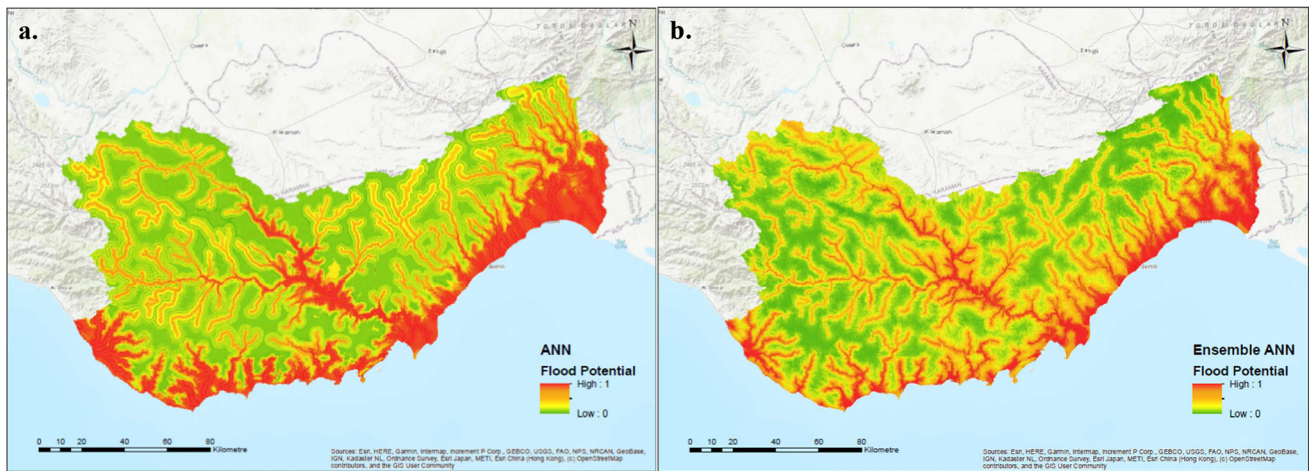


Fig. 5 Flood susceptibility maps of ANN models: **a** ANN, **b** Ensemble ANN

regions with high topographic wetness index and high drainage density.

In Table 4, preprocesses, hyperparameters and ensemble weights of the models created for single and ensemble SVM models are given. In the ensemble model, SVM has combined models with different types of kernel functions, which have undergone different data and feature preprocessing, into a single ensemble model. This explains the success of the model in the training and test set.

Flood susceptibility maps obtained from models created with the SVM algorithm are shown in Fig. 6. the SVM model created alone generally predicts the areas close to the basin drainage network with a potential close to 1, and the remaining areas with a potential close to 0. In the ensemble SVM model, it was observed that, unlike the single SVM model, areas with a distribution between 0.5 and 1 were predicted close to the drainage network. Flood susceptibility maps created by the SVM models show that

areas with high flood potential generally occurred at low elevations, with low slopes, convex profile curvatures, and close to the drainage network. However, in the ensemble SVM model created with Auto-sklearn, it was observed that the flood susceptibility areas form regions that include more parameters and can provide more clarity with a distribution between 0 and 1.

The preprocessing, hyperparameters and model weights of GBT models that show the most success in the training set are given in Table 5. The ensemble GBT model had increased performance in the test data by creating more complex tree models using a standard and robust scaler with different hyperparameters and preprocessing for feature selection.

Flood susceptibility maps of GBT models are shown in Fig. 7. Generally, the results of the two models are similar. However, the ensemble GBT model estimated the flood potential of the areas near major streams and the seaside

Table 4 Preprocesses and hyperparameters of SVM models

	Single SVM	Ensemble SVM			
Ensemble weights	–	0,2	0,4	0,2	0,2
Data preprocessing	Rescaling type	–	Standard scaler	Robust scaler	Standard scaler
	Categorical encoding	Yes	Yes	Yes	Yes
Feature preprocessing	–	Random tree embedding	Linear SVM	Nystroem sampler	Fast ICA
Kernel function	Polynomial	Polynomial	Polynomial	Linear	Linear
C (Penalty Coefficient)	250	249,344	0,4812	0,8116	1,43,258
Tolerance	0,0003	0,000296	0,000817	0,009447	0,08805
Gamma	–	0,093	0,017679	–	–
Degree	3	2	5	–	–
r	15	0,77,045	0,73,186	–	–
Shrinking	Yes	No	No	–	–

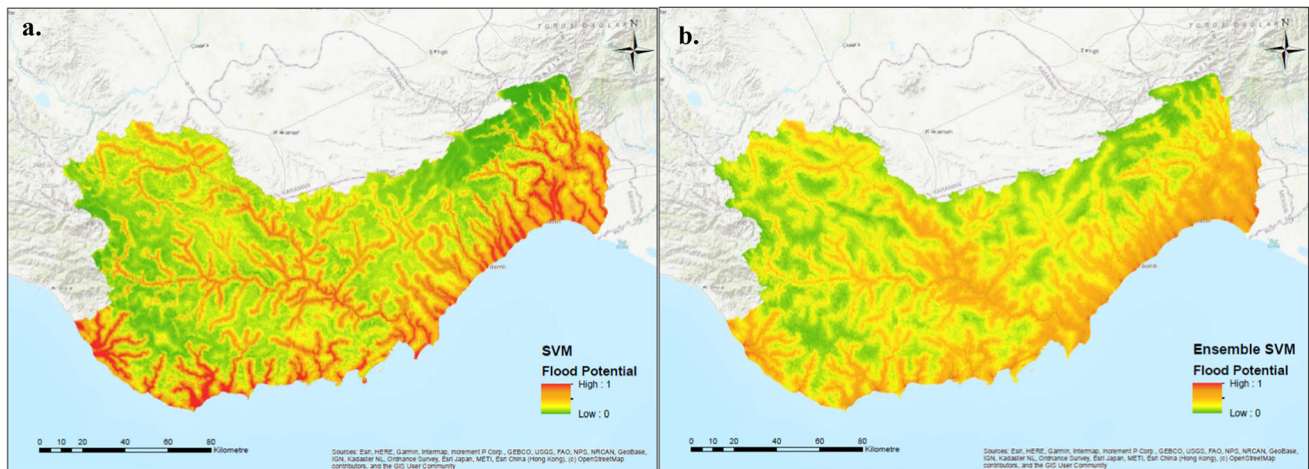


Fig. 6 Flood susceptibility maps of SVM models: **a** SVM, **b** Ensemble SVM

closer to 1 than did the single GBT model. Single and ensemble GBT models generally predicted high flood potential in low slope and convex profile curvature areas. These models also relied on the topographic wetness index parameter as well as the slope and profile curvature parameters. This is evident by the similarities of the single GBT model flood susceptibility map and the topographic wetness index map for large rivers.

In Table 6, preprocesses, hyperparameters and weights of the models created by single and ensemble DT models are given. That the ensemble DT model with different preprocessing processes and hyperparameters approaches the test set by creating more complex trees explains the success of the ensemble DT model in the test set. that the DT models generally predicted high flood potential in areas close to the sea, with low slopes and low elevations as shown in Fig. 8. Since susceptibility areas created by a single DT model cannot represent the basin generally, it can be concluded that it is insufficient in flood susceptibility mapping. This is because decision trees created during the learning phase insufficiently estimate the overall

watershed with the decision rules. However, it was observed that flood susceptibility areas created by the ensemble DT model exhibited a better distribution over the basin due to preprocessing and hyperparameter optimization.

To examine the difference in flood susceptibility areas predicted by the models more clearly, the flood susceptibility prediction of all models in the Mersin province and its vicinity are shown in Fig. 9. The ensemble SVM model, while determining the potential of floodplains, generally estimated areas with low slope and convex profile curvature as areas with flood potential close to 1 (Fig. 9b). The ensemble SVM model estimated areas with low slope and convex curvature to generally range from 0.85 to 0.95 in terms of flood potential. However, the areas with these features had indices very close to 1 in the SVM model created alone (Fig. 9a).

The single ANN model classifies areas similarly to the distance from the river map and ignores the effect of parameters other than elevation, aspect, distance from the river and profile curvature (Fig. 9c). On the other hand, the

Table 5 Preprocesses and hyperparameters of GBT models

	Single GBT	Ensemble GBT			
Ensemble weights	–	0,2	0,2	0,4	0,2
Data preprocessing	Rescaling type	–	Standard scaler	Robust scaler	–
	Categorical encoding	Yes	Yes	Yes	Yes
Feature preprocessing	–	–	Extremely random trees	–	Select Percentile
Maximum leaf nodes	–	31	36	1986	837
Regularization term	–	1.0×10^{-10}	$1,768 \times 10^{-8}$	0,41,663	$5,364 \times 10^{-5}$
Learning rate	0,08	0,1	0,05117	0,030356	0,019534
Minimum samples leaf	1	20	4	16	15
Maximum iteration	–	512	64	256	256

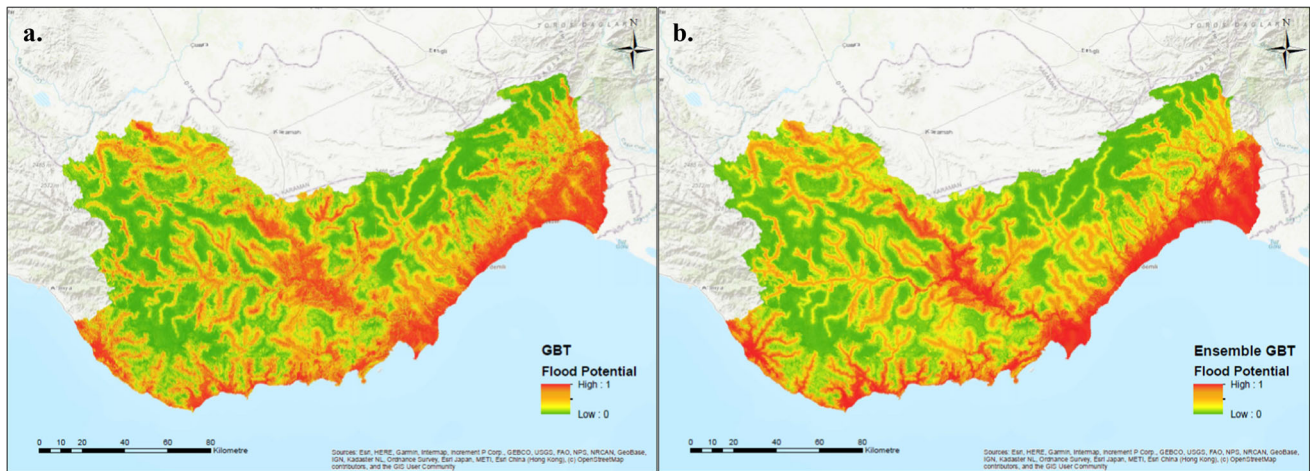


Fig. 7 Flood susceptibility maps of GBT models: **a** GBT, **b** Ensemble GBT

Table 6 Preprocesses and hyperparameters of DT models

		Single DT		Ensemble DT			
Ensemble weights		–	Select Rate	Linear SVM	Select percentile	Select percentile	Select percentile
Data preprocessing	Entropy	Gini	Gini	Gini	Entropy	Gini	–
	4	12	10	12	6	2	Var
Feature preprocessing		1	15	7	4	19	20
Criterion		5	20	10	16	2	18
Maximum depth		4	12	10	12	6	2
Minimum samples leaf		1	15	7	4	19	20
Minimum samples split		5	20	10	16	2	18

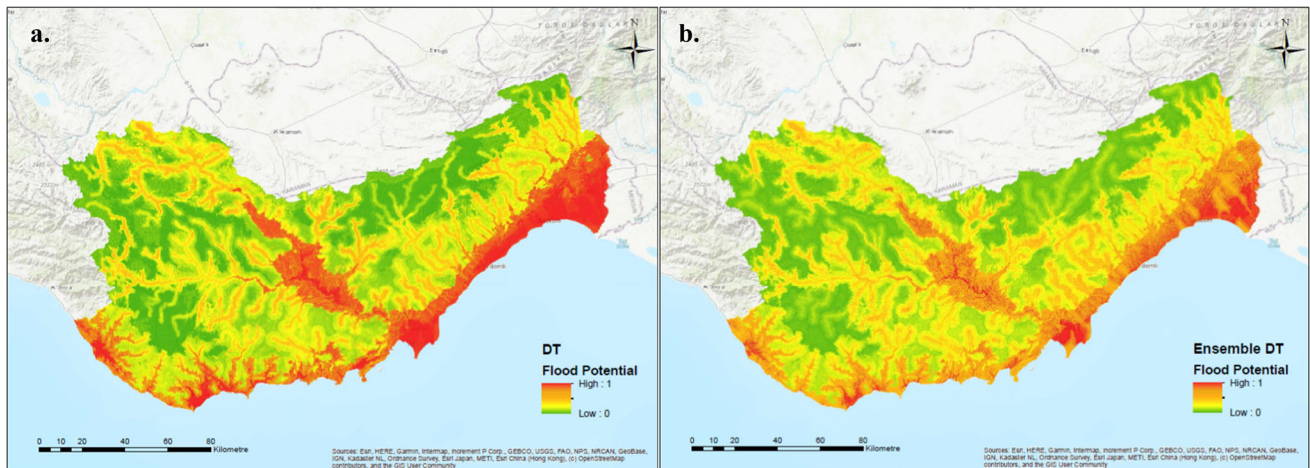
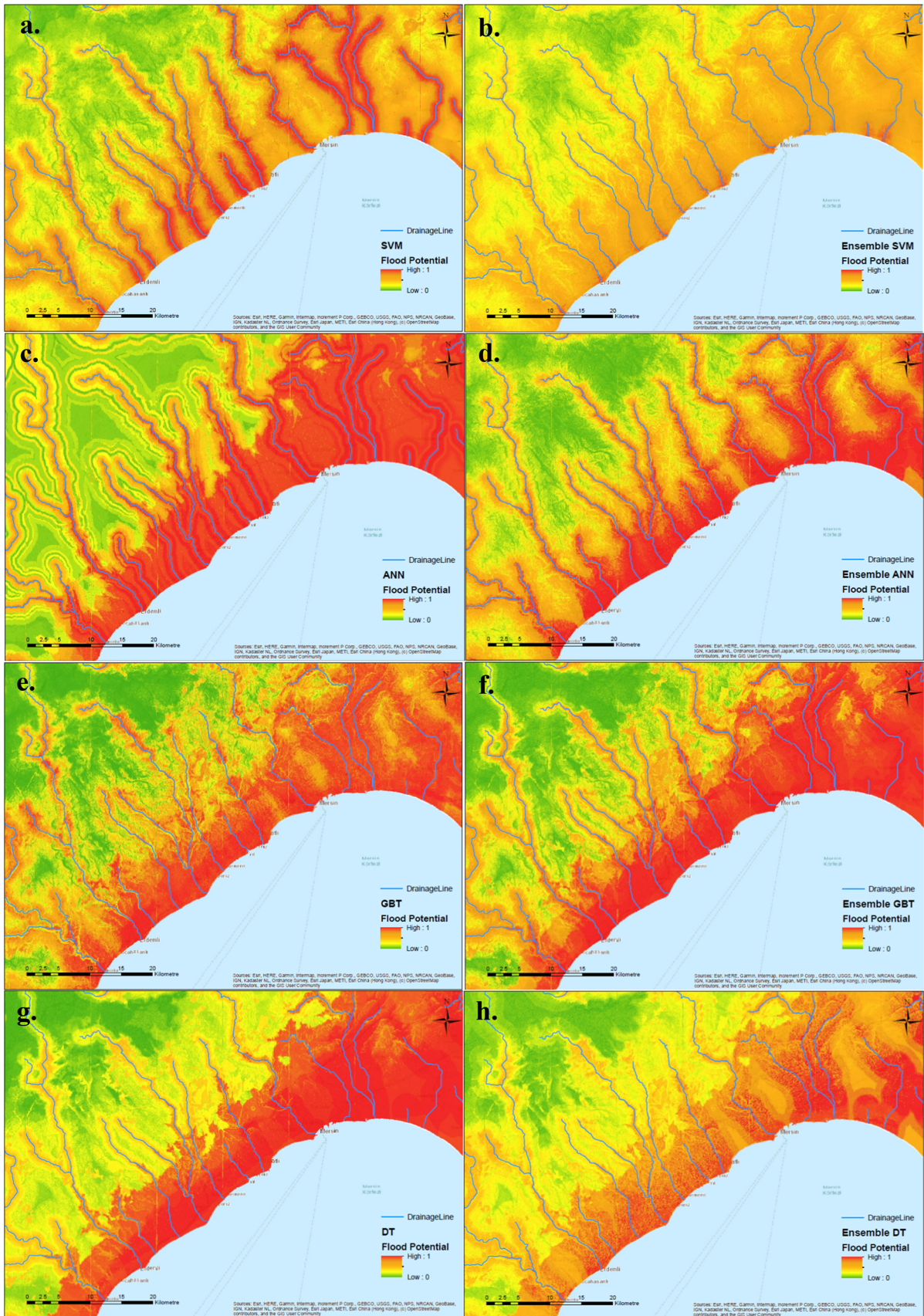


Fig. 8 Flood susceptibility maps of DT models: **a** DT, **b** Ensemble DT

ensemble ANN model considered more parameters and better reflected the overall watershed (Fig. 9d).

The single GBT and ensemble GBT models generally have similar flood hazard estimates, but the ensemble GBT model predicted areas with higher flood potential in near-

sea areas than the single GBT model. (Fig. 9e and f). In addition, in areas very close to the drainage network outside the areas close to the sea, the single GBT model resulted in areas with more flood potential than the ensemble GBT model.



◀ **Fig. 9** Flood susceptibility areas of Mersin province and its vicinity. **a** SVM, **b** Ensemble SVM, **c** ANN, **d** Ensemble ANN, **e** GBT, **f** Ensemble GBT, **g** DT, **h** Ensemble DT

Both DT models made similar estimates of areas with no flood potential. However, the models made different predictions in areas with high flood risk (Fig. 9g, and h).

The McNemar test results given in Table 7 indicate that the models created are consistent in estimating the test dataset with all values exceeding the critical alpha. The similarity of the GBT models visually evident in the flood susceptibility maps is also discernable in the McNemar test results of these two models. However, although predictions from the ANN and DT models created alone are consistent those of the other models, they are insufficient in flood susceptibility areas. This indicates that these algorithms successfully estimate test data but are inadequate in evaluating new situations with different parameters not included in the training set. This situation is not detected in the McNemar test results.

5 Conclusion

Floods that cause deadly events and socio-economic damage worldwide can be ameliorated by identifying potential flood hazard areas and taking necessary precautions. Determination of flood areas, which is a complex work dependent on many parameters, has an important place in watershed management, and can be achieved with appropriate data sets and correct approaches by using ML algorithms and ensemble models. The use of ensemble methods in different disciplines and study areas with the AutoML algorithms can guide decision-makers. In this context, findings of the study can assist decision makers at local response agencies such as councils, municipalities

and at national governmental bodies such as State Water Works, General Directorate of Water Management, Ministry of Environment, Urbanization and Climate Change by enabling the use of such powerful and advanced algorithms in the preliminary flood risk assessment phase, which is compulsory for effective flood and basin management plans. Planners and engineering consultants involved with flooding can also benefit from the results in terms of implementation of necessary actions by considering susceptible areas with higher accuracy.

One of the most important factors in determining the success of the models is that the models can predict the flooding points in the test data with high success rate. In the test set, the single ANN model correctly predicted the 36 points out of 47 flooded spots, while the Ensemble ANN model correctly predicted the 41 out of 47. Likewise, the single DT model correctly predicted the 38 of the flooded spots, while the Ensemble DT model predicted the 42 out of 47. Therefore, the number of points that are estimated to have no flood risk is also reduced by estimating the flooded points with higher accuracy. All ensemble models reduced the number of incorrectly predicted locations compared to their single models. The most outstanding of these models, the Ensemble ANN model, correctly predicted all the 41 non-flooding points in the test set.

We believe that results of the current study can be a valuable guide for local and central authorities in terms of strategic implications and they can be implemented to the Eastern Mediterranean Basin Flood Management Plan prepared by the Ministry of Agriculture and Forestry General Directorate of Water Management, Türkiye.

When the performances of the created models in the training and test sets are considered as a whole, results obtained by preprocessing and hyperparameter optimization with Auto-sklearn regarding the AUC results of the chosen algorithms showed good performance in adapting to the data. These results show how important it is to adapt a

Table 7 Pairwise comparison of each model with the McNemar test

Algorithms	GBT	ANN	SVM	DT	Ensemble GBT	Ensemble ANN	Ensemble SVM	Ensemble DT
GBT	–	0,581*	1	0,727	1	0,687	1	0,687
ANN		–	0,289	1	0,388	1	0,344	0,267
SVM			–	0,508	1	0,453	1	1
DT				–	0,453	1	0,581	0,219
Ensemble GBT					–	0,375	1	1
Ensemble ANN						–	0,453	0,219
Ensemble SVM							–	1
Ensemble DT								–

*The cell values show the probability value corresponding to χ^2 value calculated with the McNemar test in the Chi-square test statistic. For the test, the α significance level was 0.05 for the 99% confidence interval

given data set to machine learning and adjust the algorithms' hyperparameters in the models' performance. When the success rates of different algorithms in the validation set are examined, the ensemble models created by preprocessing and hyperparameter optimization with Auto-sklearn successfully estimate the validation set.

The flood susceptibility maps that were generated generally estimated regions close to Mersin, Mut, Erdemli, Mezitli, Silifke, Toroslar, Yenişehir and the Akdeniz provinces, Tarsus and Gazipaşa districts and their surrounding areas with high flood potential. The results obtained from this study and the flood management plan prepared for the Eastern Mediterranean Basin using hydrodynamic models can be evaluated together to develop a reliable watershed management plan so that governmental bodies and landowners can take necessary precautions for the areas having flooding risks in the basin. In this way, an accurate estimation of floods in a watershed can guide watershed management and planning.

The results generally show that the ensemble models are successful. However, these results are mainly focused on correct predictions. Therefore, the flood areas obtained in the study can be verified by comparing the flood areas to be obtained with the two- or three-dimensional hydraulic model used in the flood risk preliminary assessment studies.

In general, results showed that ensemble ML algorithms with preprocessing and hyperparameter optimization for the given data set more successfully modeled flood susceptibility areas in a watershed. However, the biggest problem limiting those who use ML algorithms in this context is the availability of a reliable data set that can accurately represent the basin. Attention should be given to the suitability of the data used to represent the studied study area more clarity in the prediction of flood susceptibility areas can be obtained throughout the basin by increasing the number of points in the flood inventory used in flood susceptibility mapping and diversity in the parameters affecting the flood. In this way, more generalized linear or non-linear models can be obtained using parameters that are thought to be effective on the flood instead of concentrating on a few parameters.

Author contribution MBK: Conceptualization, Methodology, Supervision, Writing—review & editing. DA: Conceptualization, Methodology, Software, Supervision Writing—review & editing. HÖ: Data analysis, Formal analysis, Software, Validation, Writing—original draft.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability The digital maps and relevant data used in the findings of this study were obtained from the governmental bodies in

Türkiye that are not publicly available. So, the data used in this study cannot be made available.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethical approval This research does not contain any studies with human participants or animals performed by any of the authors.

Consent to participate Not applicable.

Consent to publish Not applicable.

References

- Abdollahi S, Pourghasemi HR, Ghanbarian GA, Safaeian R (2019) Prioritization of effective factors in the occurrence of land subsidence and its susceptibility mapping using an SVM model and their different kernel functions. *Bullet Eng Geol Environ*. <https://doi.org/10.1007/s10064-018-1403-6>
- Ahmadi M, Karimi M, Alizadeh S, Shirzadic A, Parvinnehadd D, Shahabie 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*. <https://doi.org/10.1080/10106049.2018.1474276>
- Ahmed N, Hoque MAA, Arabameri A, Pal SC, Chakraborty R, Jui J (2021) Flood susceptibility mapping in Brahmaputra floodplain of Bangladesh using deep boost, deep learning neural network, and artificial neural network. *Geocarto Int*. <https://doi.org/10.1080/10106049.2021.2005698>
- Ali SA, Parvin F, Pham QB, Vojtek M, Vojteková J, Costache R, Linh NTT, Nguyen HQ, Ahmad A, Ghorbani MA (2020) GIS-based comparative assessment of flood susceptibility mapping using hybrid multi-criteria decision-making approach, naïve Bayes tree, bivariate statistics and logistic regression: a case of Topľa basin, Slovakia. *Ecol Indicators*. <https://doi.org/10.1016/j.ecoind.2020.106620>
- Andaryani S, Nourani V, Haghighi AT, Keesstra S (2021) Integration of hard and soft supervised machine learning for flood susceptibility mapping. *J Environ Manag* 291:112731. <https://doi.org/10.1016/j.jenvman.2021.112731>
- Band SS, Janizadeh S, Pal SC, Saha A, Chakraborty R, Melesse AM, Mosavi A (2020) Flash flood susceptibility modeling using new approaches of hybrid and ensemble tree-based machine learning algorithms. *Remote Sensing* 12:3568. <https://doi.org/10.3390/rs12213568>
- Bui DT, Pradhan B, Nampak H, Bui QT, Tran QA, Nguyen QP (2016) Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibility modeling in a high-frequency tropical cyclone area using GIS. *J Hydrol* 540:317–330. <https://doi.org/10.1016/j.jhydrol.2016.06.027>
- Chakraborty R, Pal SC, Rezaie F, Arabameri A, Lee S, Roy P, Saha A, Chowdhuri I, Moayedi H (2021) Flash-flood hazard susceptibility mapping in Kangsabati River Basin, India. *Geocarto Int*. <https://doi.org/10.1080/10106049.2021.1953618>
- Choubin B, Moradi E, Golshan M, Adamowski J, Hosseini FS, Mosavi A (2019) An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and

- regression trees, and support vector machines. *Sci Total Environ.* <https://doi.org/10.1016/j.scitotenv.2018.10.064>
- Cortes C, Vapnik V (1995) Support-Vector Networks. *Mach Learn* 20:273–297
- Dandapat K, Panda GK (2017) Flood vulnerability analysis and risk assessment using analytical hierarchy process. *Model Earth Syst Environ* 3(4):1627–1646. <https://doi.org/10.1007/s40808-017-0388-7>
- Dano UL, Balogun A, Matori A, Yusouf KW, Abubakar IR, Mohamed MAS, Aina YA, Pradhan B (2019) Flood susceptibility mapping using gis-based analytic network process: a case study of Perlis, Malaysia. *Water.* <https://doi.org/10.3390/w11030615>
- De Brito MM, Evers M, Almoradie ADS (2018) Participatory flood vulnerability assessment: a multi-criteria approach. *Hydrol Earth Syst Sci* 22(1):373–390. <https://doi.org/10.5194/hess-22-373-2018>
- Dietterich TG (1997) Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Comput* 10(7):1895–1923
- El-Magd SAA, Pradhan B, Alamri A (2021) Machine learning algorithm for flash flood prediction mapping in Wadi El-Laqaita and surroundings, Central Eastern Desert, Egypt. *Arab J Geosci.* <https://doi.org/10.1007/s12517-021-06466-z>
- Erdem M, Boran FE, Akay D (2016) Classification of risks of occupational low back disorders with support vector machines. *Hum Factors Ergon Manufact Serv Ind.* <https://doi.org/10.1002/hfm.20671>
- Feurer M, Aaron K, Eggensperger K, Jost S, Manuel B, Hutter F (2015) Efficient and robust automated machine learning. *Adv Neural Inf Process Syst* 28:2962–2970
- Friedman JH (1999) Greedy function approximation: a gradient boosting machine <http://www.salford-systems.com/doc/GreedyFuncApproxSS.pdf>.
- General Directorate of Water Management, Republic of Türkiye Ministry of Forestry And Water Management, (2016). The Effect of Climate Change on Water Resources Project Final Report, Ankara. https://www.tarimorman.gov.tr/SYGM/Belgeler/iklim%20de%C4%9Fi%C5%9Fikli%C4%9Finin%20su%20kay-naklar%C4%B1na%20etkisi/Iklim_NihaiRapor.pdf.
- Gijsbers P, LeDell E, Thomas J, Poirier S, Bischl B, Vanschoren J (2019) An open source AutoML benchmark. *arXiv preprint arXiv:1907.00909*.
- Hong H, Pradhan B, Jebur MN, Bui DT, Xu C, Akgun A (2016) Spatial prediction of landslide hazard at the Luxi area (China) using support vector machines. *Environ Earth Sci* 75(1):1–14. <https://doi.org/10.1007/s12665-015-4866-9>
- Hoque MAA, Tasfia S, Ahmed N, Pradhan B (2019) Assessing spatial flood vulnerability at Kalapara Upazila in Bangladesh using an analytic hierarchy process. *Sensors* 19(6):1302. <https://doi.org/10.3390/s19061302>
- Horritt MS (2000) Calibration of a two-dimensional finite element flood flow model using satellite radar imagery. *Water Resour Res.* <https://doi.org/10.1029/2000WR900206>
- Ishtiaque A, Eakin H, Chhetri N, Myint SW, Dewan A, Kamruzzaman M (2019) Examination of coastal vulnerability framings at multiple levels of governance using spatial MCDA approach. *Ocean Coast Manag* 171:66–79. <https://doi.org/10.1016/j.ocecoaman.2019.01.020>
- Jahangir MH, Reineh SMM, Abolghasemi M (2019) Spatial prediction of flood zonation mapping in Kan River Basin, Iran, using artificial neural network algorithm. *Weather Clim Extrem* 25:100215. <https://doi.org/10.1016/j.wace.2019.100215>
- Jaiswal RK, Ghosh NC, Lohani AK, Thomas T (2015) Fuzzy AHP based multi criteria decision support for watershed prioritization. *Water Resourc Manag.* <https://doi.org/10.1007/s11269-015-1054-3>
- Khosravi K, Shahabi H, Pham BT, Adamowski J, Shirzadi A, Pradhan B, Dou J, Ly H, Gróf G, Ho HL, Hong H, Chapi K, Prakash I (2019) A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods. *J Hydrol* 573:311–323. <https://doi.org/10.1016/j.jhydrol.2019.03.073>
- Knuepfer PL, Montz BE (2008) Flooding and watershed management. *J Contemp Water Res Educat* 138(1):45–51
- Koçyiğit MB, Akay H, Babaiban E (2021) Evaluation of morphometric analysis of flash flood potential of Eastern Mediterranean Basin using principle component analysis (Original in Turkish). *J Fac Eng Architec Gazi Univ* 36(3):1669–1686. <https://doi.org/10.17341/gazimmfd.829390>
- Kohavi R (1995) A Study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai* 14(2):1137–1145
- Kotsiantis SB (2013) Decision trees: a recent overview. *Artif Intell Rev* 39(4):261–283. <https://doi.org/10.1007/s10462-011-9272-4>
- Membele GM, Naidu M, Mutanga O (2021) Examining flood vulnerability mapping approaches in developing countries: a scoping review. *Int J Disaster Risk Reduct.* <https://doi.org/10.1016/j.ijdrr.2021.102766>
- Meshram SG, Alvandi E, Meshram C, Kahya E, Al-Quraishi AMF (2020) Application of SAW and TOPSIS in prioritizing watersheds. *Water Resour Manag.* <https://doi.org/10.1007/s11269-019-02470-x>
- 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. *Geomat Nat Haz Risk* 8(2):1080–1102. <https://doi.org/10.1080/19475705.2017.1294113>
- Moore IG, Grayson RB, Ladson AR (1991) Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrol Process* 5(1):3–30
- Naghibi SA, Ahmadi K, Daneshi A (2017) Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping. *Water Resour Manage* 31(9):2761–2775. <https://doi.org/10.1007/s11269-017-1660-3>
- Nhu V-H, Shirzadi A, Shahabi H, Singh SK, Al-Ansari N, Clague JJ, Jaafari A, Chen W, Miraki S, Dou J, Luu C, Górski K, Pham PT, Nguyen HD, Ahmad BB (2020) Shallow landslide susceptibility mapping: a comparison between logistic model tree, logistic regression, naïve Bayes tree, artificial neural network, and support vector machine algorithms. *Int J Environ Res Public Health.* <https://doi.org/10.3390/ijerph17082749>
- Omran A, Schröder D, El Rayes A, Geresh M (2011) Flood hazard assessment in Wadi Dahab, Egypt based on basin morphometry using GIS techniques. *GI Forum Program Committee.*
- Ouma YO, Tateishi R (2014) Urban flood vulnerability and risk mapping using integrated multi-parametric AHP and GIS: methodological overview and case study assessment. *Water.* <https://doi.org/10.3390/w6061515>
- Özdemir H (2022) Flood susceptibility mapping with ensemble based machine learning; Case of the Eastern Mediterranean (M.Sc.thesis original in Turkish). Gazi University, Ankara, Türkiye.
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in python. *J Mach Learn Res*, 2825–2830.
- Pradhan B, Youssef AM (2011) A100-year maximum flood susceptibility mapping using integrated hydrological and

- hydrodynamic models: Kelantan River Corridor, Malaysia. *J Flood Risk Manag.* <https://doi.org/10.1111/j.1753-318X.2011.01103.x>
- Prasad P, Loveson VJ, Das B, Kotha M (2021) Novel ensemble machine learning models in flood susceptibility mapping. *Geocarto Int.* <https://doi.org/10.1080/10106049.2021.1892209>
- Quinlan J (1986) Induction of decision trees. *Mach Learn* 1:81–106
- Rahmati O, Pourghasemi HR, Zeinivand H (2016) Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. *Geocarto Int.* <https://doi.org/10.1080/10106049.2015.1041559>
- Rahmati O, Tahmasebipour N, Haghizadeh A, Pourghasemi HR, Feizizadeh B (2017) Evaluation of different machine learning models for predicting and mapping the susceptibility of gully erosion. *Geomorphology* 298:118–137. <https://doi.org/10.1016/j.geomorph.2017.09.006>
- Riley SJ, DeGloria SD, Elliot R (1999) At ruggedness Index that quantifies topographic heterogeneity. *Int J Sci* 5(1–4):23–27
- Roy DC, Blaschke T (2015) Spatial vulnerability assessment of floods in the coastal regions of Bangladesh. *Nat Haz Risk* 6(1):21–44. <https://doi.org/10.1080/19475705.2013.816785>
- Roy P, Pal SC, Chakraborty R, Chowdhuri I, Malik S, Das B (2020) Threats of climate and land use change on future flood susceptibility. *J Clean Prod.* <https://doi.org/10.1016/j.jclepro.2020.122757>
- Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323(6088):533–536
- Samanta S, Pal DK, Palsamanta B (2018) Flood susceptibility analysis through remote sensing, GIS and frequency ratio model. *Appl Water Sci.* <https://doi.org/10.1007/s13201-018-0710-1>
- Shafizadeh-Moghadam H, Valavi R, Shahabi H, Chapi K, Shirzadi A (2018) Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. *J Environ Manag* 217:1–11. <https://doi.org/10.1016/j.jenvman.2018.03.089>
- Sørensen R, Zinko U, Seibert J (2006) On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrol Earth Syst Sci* 10(1):101–112
- Su C, Wang L, Wang X, Huang Z, Zhang X (2015) Mapping of rainfall-induced landslide susceptibility in Wencheng, China, using support vector machine. *Nat Haz* 76(3):1759–1779. <https://doi.org/10.1007/s11069-014-1562-0>
- Tehrany MS, Pradhan B, Jebur MN (2013) Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *J Hydrol* 504:69–79. <https://doi.org/10.1016/j.jhydrol.2013.09.034>
- Tehrany MS, Pradhan B, Jebur MN (2015) Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. *Stoch Environ Res Risk Assess.* <https://doi.org/10.1007/s00477-015-1021-9>
- Tehrany MS, Pradhan B, Mansor S, Ahmad N (2015b) Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. *Catena* 125:91–101. <https://doi.org/10.1016/j.catena.2014.10.017>
- Wang Y, Hong H, Chen W, Li S, Pamučar D, Gigović L, Drobnjak S, Tien Bui D, Duan H (2018) A hybrid GIS multi-criteria decision-making method for flood susceptibility mapping at Shangyou. *China Remote Sens* 11(1):62. <https://doi.org/10.3390/rs11010062>
- Wu J, Chen X, Zhang H, Xiong L, Lei H, Deng S (2019) Hyperparameter optimization for machine learning models based on bayesian optimization. *J Electron Sci Technol.* <https://doi.org/10.11989/JEST.1674-862X.80904120>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.