Flood susceptibility mapping using meta-heuristic algorithms

Alireza Arabameri, Amir Seyed Danesh, M. Santosh, Artemi Cerda, Subodh Chandra Pal, Omid Ghorbanzadeh, Paramita Roy and Indrajit Chowdhuri

Department of Geomorphology, Tarbiat Modares University, Tehran, Iran; Faculty of Technology and Engineering, East of Guilan, University of Guilan, Rudbar-Vajargah, Iran; School of Earth Sciences and Resources, China University of Geosciences Beijing, Beijing, China; Department of Earth Sciences, University of Adelaide, Adelaide, South Australia, Australia; Soil Erosion and Degradation Research Group, Departament de Geografia, Universitat de Valencia, Valencia, Spain; Department of Geography, The University of Burdwan, Bardhaman, West Bengal, India; Department of Geoinformatics—Z_GIS, University of Salzburg, Salzburg, Austria

ABSTRACT

Flood is a common global natural hazard, and detailed flood susceptibility maps for specific watersheds are important for flood management measures. We compute the flood susceptibility map for the Kaiser watershed in Iran using machine learning models such as support vector machine (SVM), Particle swarm optimization (PSO), and genetic algorithm (GA) along with ensembles (PSO-GA and SVM-GA). The application of such machine learning models in flood susceptibility assessment and mapping is analyzed, and future research suggestions are presented. The model of flood susceptibility model was constructed based on fifteen causatives: slope, slope aspect, elevation, plan curvature, land use, and land cover, normalize differences vegetation index (NDVI), convergence index (CI), topographical wetness index (TWI), topographic positioning Index (TPI), drainage density (DD), distance to stream, terrain ruggedness index (TRI), terrain surface texture (TST), geology and stream power index (SPI) and flood inventory data which later is divided by 70% for training the model and 30% for validated the model. The model output was evaluated through sensitivity, specificity, accuracy, precision, Cohen Kappa, F-score, and receiver operating curve (ROC). The evaluation of flood susceptibility mapping through the receiver operating curve method along with flood density shows robust results from support vector machine (0.839), particle swarm optimization (0.851), genetic algorithm (0.874), SVM-GA (0.886), and PSO-GA (0.902). Compared have done with some methods commonly used in this susceptibility assessment. A high-quality, informative database is essential for the classification of flood types in flood susceptibility mapping that is very important and helpful to improve the model performances. The performance of the ensemble PSO-GA is better...
1. Introduction

Flood is one of the common and disastrous natural hazards in different regions of the globe. According to United Nations disaster and risk reduction data, floods from 1996 to 2015 have caused severe damage to property and human life (Hong et al. 2018; Chen et al. 2019). Preparation of flood susceptibility maps is important in reducing the damages arising from flood and exercise alert through suitable measurements in the susceptible areas. In this context, geo-environmental parameters are essential for the formulation of susceptibility maps considering the most significant causatives of the flood. The susceptibility map is helpful for regional planners and decision-makers to take suitable measures. Even though floods cannot be controlled, this map help in managing the pre-condition and post-operation of floods.

The assessment of floods is complex because of the diverse climatic and anthropogenic factors. With the ever-increasing global population, devastating floods can affect millions of lives and property loss. Due to global climate change, the increasing frequency and intensity of rainfall also lead to a flood. In the case of Iran, the climate is highly variable, with cold winters and heavy snowfall at subfreezing temperatures in the northwestern part. The weather is relatively mild during the spring season, whereas summer is dry and hot. In the southern part of the country, winters are mild, and summer is very hot, with an average daily temperature in July exceeding 38 °C. Though Iran has a continental climate, most precipitation occurs from October to April, and the annual average precipitation is 880 mm. When the intensity of rainfall goes above infiltration capacity, flood events occur. Anthropogenic activities like land use planning, unscientific construction, dam construction, cutting of trees for road construction, and rapid settlement plantation on the riverbank have all contributed to the increased impact on the flood. Where dam construction can change the normal hydrological process and impervious area, road construction and urbanization reduce infiltration rate and promotes flood with warm water that indirectly destroys the ecosystem process in the river (Hawley and Bledsoe 2011). Therefore, both climate and human activities contribute to flood occurrences.

The preparation of flood susceptibility maps using various machine learning models through statistical approaches has gained attention in recent years, rather than the earlier approach of using only statistical models. Several previous studies have adopted machine learning models to prepare flood susceptibility maps such as decision trees (Khosravi et al. 2018; Choubin et al. 2019), support vector machine (SVM) (Huang et al. 2010; Lin et al. 2013), random forest (Chapi et al. 2017; Liu et al. 2020), and artificial neural network (Campolo et al. 2003; Gebrehiwot et al. 2019). Several regional approaches were also applied in the catchment to estimate the flood probabilities, namely index flood (Dalrymple 1960; ) Region of Influence (Burn 1990; Zrinji and Burn 1994). In recent years many data driven methods considering bi-
variate statistical models, multi-criteria decision making (MCDM), Artificial Intelligence (AI), Computational Intelligence (CI), soft computing, Data mining (DI), Knowledge Discovery in Database (KDD), Intelligence Data Analysis (IDA) have shown promises to predict flood susceptibility. Recently, the hybrid computational method has been shown to perform more accurately preparing flood susceptibility maps. The sub-sampling and bootstrapping algorithm with machine learning is used for preparing such maps. Costache et al. (2020) point out the application of integrated frequency ratio and logistic regression model for FSM. The bagging and random subspace ensemble strategies instigated with Reduced-error pruning tree models have also been applied (Chen et al. 2019; Bui et al. 2019; Wang et al. 2019. Bui et al. (2019) and Wang et al. (2019) used meta-heuristic optimization algorithms to find the optimal parameters of intelligence models. Though these methods have different strategies, they are all helpful to predict the unknown areas defined by the probability for flooding.

A flood susceptibility map is a tool for risk reduction measures (Tehrany et al. 2014; Youssef et al. 2016; Dano et al. 2019; Mind’je et al. 2019; Roy, Chandra Pal, Chakraborty et al. 2020). In a nonlinear relationship between multiple variables and hazard level, the flood susceptibility assessment is considered as a quantitative evaluation of the degree of rainstorms and their adverse impact, including the root cause of the flood and final output as an environmental hazard. The methods of flood susceptibility maps were categorized into three parts: hydrological and hydrodynamic models, multi-criteria decision analysis, and machine learning models. In machine learning methods, feature engineering is an essential step that uses domain knowledge of the data to features to obtain more accurate results. This critical technique converts the raw data into specific data to the predictive models. Thus, the information from the study area is first converted from raster data to specific spatial resolution. The susceptibility prediction is regarded as a binary classification process to differentiate the grid cell responsible for flood disaster. Each grid cell comprises a feature value that is a conditioning factor.

Most of the previous studies provided a simple rate of contribution of two factors based on the time series, where the flood susceptibility map with causative characteristics represents the final output in the environment. The simple relative contribution rate helps to understand the effect of climate change and human activities to flood. Nevertheless, modern approaches have established the causatives as spatial training data and validated them through ensembles for accuracy.

A large amount of spatial data can be acquired through remote sensing (Zhang et al. 2019; Zhang et al. 2019; Zhao et al. 2020; Chao et al. 2021; Wang et al. 2021; Zhou et al. 2021; Zheng et al. 2021; Zhao et al. 2021) and geographical information system (GIS) (Li and Zhang 2008). The MCDM (Yang et al. 2022), statistical (Chen et al. 2022), deep learning (Zhong et al. 2021), and machine learning (Xiong et al. 2022) models and their hybrid combinations were used for better accuracy. Based on a careful evaluation of previous studies, the most popular and robust method was selected in this research work: SVM, particle swarm optimization, and genetic algorithm (GA) for preparing the flood susceptibility map to identify vulnerable zones.
This study aims to present an appropriate methodology for the delineation of susceptibility map assessment in the Kaiser watershed, Iran; therefore, for this purpose, we have used a machine learning model through remote sensing and a GIS platform. This approach will help to identify flood-prone or vulnerable zones by considering geo-environmental factors in the surrounding environment. Fifteen causatives were considered in the model for testing and training. The output of three machine learning models was validated through sensitivity, specificity, accuracy, Cohen Kappa, receiver operating curve (ROC), precision, and F score. The results from the present study can be helpful for regional planners, engineers, and future researchers by focusing on the vulnerable zone for effective management. The governmental agencies can also use the information for taking suitable preventive measures. This paper is organized as follows: section 2 has described the study area and dataset. Section 3 has represented the models and a brief discussion about them. Moreover, the result of the application of various techniques was presented in section 4 to address flood-prone areas in this region. Finally, a summary and conclusion have to end this research paper.

Based on the previous literature, the hybrid model performs better than the single model (). SVM advances a lot in dealing with the problem of nonlinear regression prediction. However, its generalization performance is sensitive to the selection of the parameters (Bui et al. 2016). The difficulties in capturing the critical modeling variables are noted as the major drawbacks to SVM (Choubin et al. 2019). Hence, it is urgent to apply optimization algorithms to search for the optimal parameters of SVM. The Genetic Algorithms (GA) is an optimization algorithm it simulates the mechanism of genetic variation and the theory of biological evolution. GA has the advantages of parallelism and global optimization. It has achieved excellent optimization results in plenty of studies (Zhang et al. 2014). In the current study, GA was utilized to select and optimize the parameters of SVM. Application of SVM-GA hybrid model in flash flood susceptibility mapping is novelty of this manuscript.

2. Materials and methods

2.1. Study area

The Kaiser watershed covering 3648.32 km² of the present study is located in the Mazandaran province, northern Iran. This area lies between 35°56’22″ to 36°49’20″N latitude and 53°00’34″ to 53°43’41″E longitude (Figure 1). The elevation ranges from 1 mt m.a.s.l to 3725 m.a.s.l and average elevation is 932.62 m.a.s.l. Half of the basin part is mountainous, whereas the northern part of the watershed has a smooth, gentle slope, forming part of the Alborz mountains. According to IRIMO 2012 (IRIMO 2012), the annual rainfall of this area is around 600 mm. the high intensity of rainfall over a short period and land-use changes in rangelands, e.g. deforestation, conversion to agricultural land and garden to residential areas, causes severe flood here. Maximum rainfall occurring month is January, February, March, and October, addition October is the wettest month within average monthly precipitation of 160 mm. The climatic features of the area s semi-humid, semi-arid, and the Mediterranean. Most of the study area is utilized for agriculture (around 45.35%), followed by
rangeland (38.07%), forest (15.95%), and residential area (0.64%). Geologically, the dominant part of the area consists of marl, calcareous sandstone, sandy limestone, and a minor conglomerate. Red conglomerate and sandstone are also found. In this study area, alfisols (58.30%), mollisols (25.76%), and rock outcrops/entisols (8.73%) are the most common soil types.

2.2. Database

Data preparation is an essential and forward step to prepare the flood susceptibility map. The slope, slope aspect, elevation, and plan curvature factors were extracted from the digital elevation model (DEM) using Arc GIS software. Using Euclidean tools, the drainage density and distance to stream were extracted. The land use land cover data were derived from Landsat 8 satellite images. The factors of NDVI were calculated using ENVI software. The topographic wetness index (TWI), Topographic position index (TPI), stream power index (SPI), terrain ruggedness index (TRI), terrain surface texture (TST), and convergence index (CI) were calculated from DEM using SAGA software. Various scale data were first vectorized on the Arc GIS platform. The flood conditioning factors were extracted and converted into raster data from the vector data. All the data were converted into raster data with the exact spatial resolution consistent with the DEM (Table 1). Based on previous data, expert knowledge, and characteristics of the flood distribution, the factors were reclassified. In this study, 70% flood and non-flood points were considered as training data set, and 30% flood and non-flood points were considered for validation of those datasets.
2.3. Methodology

We adopted the following workflow for this study.

I. Firstly, the flood location and causatives were identified through high-resolution satellite images, Google earth images, GPS, and based on an extensive field survey. The flooding and non-flood points were considered 70% and 30% in a random split.

II. The multi co-linearity analysis among the factors was done through the variance inflation system (VIF) and tolerance (TOL).

III. To get the susceptibility map for identifying the flooding zone, three models were used, namely, SVM, PSO, and GA.

IV. Two ensemble models were used for better performance with accuracy: the SVM-GA and PSO-GA (Figure 2).

V. Validation of results has a vital role in scientific works (Moayedi and Mosallanezhad 2017; Moayedi et al. 2021; Mi et al. 2022; Zhou et al. 2021). This model’s result was validated through sensitivity, specificity, accuracy, F score, kappa’s model, precision, and receiver operating curve.

2.3.1. Flood inventory map

For modeling and susceptibility mapping of floods in an area, an inventory map is essential and primary data for training and testing the applied models. This flood inventory map shows the spatial distribution of floods and their geometry. Through the susceptibility map, one can predict the vulnerable flood zone by studying historical flood events. In recent years, remote sensing and GIS have provided important platforms to create flood inventory maps. The geographical location of flood images was collected through Google earth images and local and regional organizations. Extensive field observation was done to compliment the global positioning earth system (GPS) and Google earth images. Based on this data, all-inclusive maps were created in the ArcGIS10.3 version. The final output inventory map was split into two parts as training and validation data in seven replicates with different training and validation balances. In this study, flood points were considered in the model 70% and 30% as training and validation data set. Moreover, the results from the three models

| Database            | Source                          | Purpose                                                                 | Spatial resolution/scale |
|---------------------|---------------------------------|-------------------------------------------------------------------------|--------------------------|
| ALOS PALSAR DEM     | Alaska Satellite Facility       | Slope, slope aspect, elevation, plan curvature, convergence index, topographical wetness index (TWI), stream power index (SPI), topographic position index (TPI), terrain ruggedness index (TRI), TST, | 12.5 mt.                |
| Sentinel 2A         | European Space Agency (ESA)    | Land use and land cover (LULC), normalized difference vegetation index (NDVI) | 10 mt.                  |
| ALOS PAL-SAR DEM    | Alaska Satellite Facility      | Distance to stream, drainage density                                    | 12.5 mt.                |
| Geological Society of Iran (GSI) | http://www.gsi.ir/ | Geology                                                                 | 1:100,000               |
were validated through sensitivity, specificity, accuracy, precision, F score, kappa model, and receiver operating curve.

### 2.3.2. Flood conditioning factors

A variety of environmental factors triggered the flooding condition, and these causatives ensure the reliability and accuracy of flood susceptibility mapping due to appropriate parameters (Chapi et al. 2017; Bui et al. 2020). We selected the flood conditioning factors in this research work based on previous studies and expert knowledge. Noted that a flat area is more probable for flood as water flows down from higher terrain. River flood disasters are more likely to occur in lower height and slope areas in Iran. The degree of slope surface is indicated by curvature and is an important factor to cause the flood. The minimum value is 0, and the maximum value is 72.45 in degree (Figure 3). The slope Aspect is another key factor, as the windward slope is prone to precipitation. Aspect is related to the intensity of solar radiation, which affects the surface vegetation and soil moisture. The elevation ranges from 1 mt to 3725 mt control the water flow rate and is responsible for the flooding. Furthermore, the terrain parameter, convergence index (CI), represents the agreement of aspect direction and control of the flooding process. The output value varies from -100 to 100 and is extracted from the DEM and satellite images with 12.5 m resolution.

Several authors (e.g. Tehrany et al. 2015; Mahmoud and Gan 2018) confirmed that the topographic wetness index (TWI) presents soil saturated situation associated with water accumulation in the basin. TWI is calculated as follows:
Figure 3. Flood conditioning factors: (a) slope, (b) aspect, (c) elevation, (d) plan curvature, (e) convergence index, (f) TWI, (g) SPI, (h) TPI, (i) TRI, (j) TST, (k) distance to stream, (l) drainage density, (m) LULC, (n) NDVI, and (o) geology.
Where $\propto$ presents upslope area per unit contour length and slope angle is represented by $\beta$.

In Iran, most of the floods occur due to heavy rainfall, and the annual average precipitation is chosen as a conditioning factor for the flood. The stability of the terrain is affected by the stream power index (SPI), which directly reflects the erosive force of the current. Studied the relationship between geomorphic factors and flood, he claimed that catastrophic channel variation is led by high stream power.

$$SPI = As \times \tan \sigma$$  \hspace{1cm} (2)  

Where $As$ and $\tan B$ indicate the area of the basin and slope gradient, respectively (Chakrabortty et al. 2020).
Water permeability, storage capacity, and drainage density are controlled by soil types (Zuo et al. 2020; Jiang et al. 2021). Lithology determines the shape of the channel and affects the development of floodplains (Heitmuller et al. 2015). The distance of the river as a factor controls the flood discharging and its expansion. The surface vegetation coverage and density were shown by normalizing differences in vegetation index (NDVI). The relationship between NDVI was established in previous studies (Gao et al. 2012). Caprario and Finotti (2019) noted that sparse vegetation cover has a high potential for flooding since its poor water storage capacity. The NDVI is calculated as follows:

\[
NDVI = \frac{RNIR - RR}{RNIR + RR}
\]  

\( RNIR \) is the spectral reflectance of the near-infrared band, and \( RR \) is the red band in the electromagnetic spectrum (Pal et al. 2018; Malik et al. 2019).

The drainage density on the earth’s surface is directly related to the hydrological process (Yin et al. 2022), and the range varies from 0 to 1.82. High drainage density defines a more vulnerable area or zoning to flood. Water infiltration and evapotranspiration are affected by various types of land use and land cover of any area. Nine types of land use zone have been identified, namely agricultural area, fellow land, built-up area, water bodies, dense forest, scattered forest, mixed forest rangeland, and plantation. Moreover, the geology of any region helps construct river channels and develops the flood plains. The geology of our study area can be classified into nine zones. The topographic position index can be used to determine the ruggedness of the terrain. It can be calculated by comparing each cell slope to the mean slope. The topographic ruggedness index expresses the elevation difference between adjacent cells of a DEM. The value of TRI varies from 0 to 85.28 and another factor, topographic surface texture, refers to both the profile shape and surface roughness which control flood intensity and speed of flowing. Here the value of TST 0 to 68.63 plays a significant role to occur the flood in this region frequently.

### 2.3.3. Multi-collinearity analysis

It is a standard criterion for excluding the co-linearity and selecting influential factors for the probability estimation model. If the value of variance inflation function (VIF) is more than 5 and the TOL value is 0.1 then it can be concluded that factors create the multi co-linearity problem. Fifteen flood conditioning parameters were selected to prepare the flood susceptibility map: slope, slope aspect, elevation, plan curvature, SPI, TWI, CI, TST, TRI, TPI, NDVI, LULC, and distance to stream, drainage density, and geology. Two indices were included to estimate the multi co-linearity assessment of the factors, such as variance inflation function (VIF) tolerance (TOL). The calculation is as follows:

\[
TOL = 1 - R_j^2
\]  

\[
VIF = \frac{1}{TOL}
\]
where, $R^2_j$ indicate the regression value of $j$ on supplementary parameters in a given dataset (Roy, Chandra Pal, Arabameri et al. 2020; Chowdhuri et al. 2021).

### 2.3.4. Importance of flood susceptibility causal parameters by random forest

In this research, random forest (RF), a robust machine learning method, was partially applied to determine the relative importance of flood causal parameters. The RF is the multivariate non-parametric statistical analysis. First of all, the algorithm was developed by Breiman (Breiman 2001). The RF algorithm generated the decision tree (DT) with the help of training datasets (Chowdhuri, Pal, Arabameri et al. 2020). The target variables indicated (1) and absence (0) of the flood and the individual flood causal parameters input as independent variables. The RF algorithm and its generalization error are expressed here. In this analysis, the out-of-bag error is 5.14%, where the prediction result is 94.86% is corrected. The average values of mean decrease Gini (MDG) and mean decrease accuracy (MDA) were used to calculate the importance of LS causative factors for the construction of the flood susceptibility model.

\[
GE = P_{x,y}(mg(x,y) < 0) \tag{6}
\]

\[
mg(x,y) = av_k I_j(x) - \max_{j\neq y} av_k I_k(x) = j \tag{7}
\]

Where, $x$ and $y$ represent flood causal parameters specify the probability over $x$ and $y$ space, $mg$ indicates the marginal function, and $I$ (*) represents the indicator function (Breiman 2001).

### 2.3.5. Flood susceptibility modeling

#### 2.3.5.1. Support vector machine

Among machine learning methods (Moayedi et al. 2020), a support vector machine (SVM) is a data mining machine learning method including a set of linear indicators functions, and for the issues of solution estimation, it can be used (Vapnik et al. 1995). The SVM is also known as the maximum margin method (Mohammadi et al. 2019), and it yields higher performance and better results with limited data points. Based on the statistical learning theory SVM maps the data-set into a high dimensional feature space through nonlinear transformers to create the best hyperplane (Ghorbanzadeh et al. 2019). When margins between the defined classes of the problem are maximal, the best hyperplane can be obtained. SVMs are unidirectional in character and consist of two layers that implement various activation functions: linear, polynomial, radial, sigmoid, and radial basis function (RBF). Among them, the RBF kernel function is more prevalent in SVM for its nonlinear classifier or regression line characteristics. Though SVM includes four kernel categories approach to understand the classification efficiency and susceptibility mapping along with the type as a string. SVM performs and generalized well on the out of sample data. It is slightly different from SVR, a regression algorithm for working with continuous values and SVM is for classification and can handle high dimensional data, which proves a great help taking into its account and application in the field of machine learning.
2.3.5.2. Particle swarm optimization. Particle swarm optimization is an evolutionary computing algorithm derived from bird predation behavior study. The birds foraging process, heuristic characteristics, and random search of evolutionary algorithms inspired the construction of this model (Abido 2002). In this algorithm, the bird is considered an independent particle, and a group of birds is taken as a particle swarm. The ‘group’ and ‘individuals’ are common like another evolutionary algorithm. Each particle can be regarded as a searching individual in the N-dimensional search space. The optimal historical position of the particles and optimal historical position of the population helps to adjust the flight speed of the particle dynamically. Particle swarm optimization includes two parts, namely speed and position. The speed of movement is represented by speeds and positions representing the direction of movement. The mathematical calculation is as follows:

\[
V^* = wV + c_1 r_1 (pbest - X) + c_2 r_2 (Gbest - X)
\]

\[
X^* = X + V^*
\]

Where \( pbest \) represents the historical best position of the single-particle and \( Gbest \) represents the historical best position of particle swarm, \( c_1 \) and \( c_2 \) are considered as learning factors. \( r_1 \) and \( r_2 \) are the random probabilities values that vary from 0 to 1. \( w \) indicates the inertial weight. The current position and velocity are represented by \( X \) and \( V \), respectively. \( X^* \) and \( V^* \) represents the updated position and velocity of the particle.

Each particle helps to search for the optimal solution which is called an individual extreme and in the particle swarm, the optimal individual extreme is taken for the current global optimal solution. The speed and position are continuously updated by iterating which helps to get the optimal solution and finally meets the terminate conditions. To adapt to the environment, every particle helps one another in this model and achieves the optimal search of complex solutions in complex spaces (Armaghani et al. 2014).

2.3.5.3. Genetic algorithm. From an evolitional perspective, the genetic algorithm developed a computational model (Mühlenbein 1997). This type of algorithm encodes the potential solution or candidate solution to a specific problem on the chromosome-like data structure. Whitley (1994) noted that the genetic algorithm preserves critical information by applying recombination operators to those structures.

The GA searches for a candidate solution space by simulating the evolution process in nature to get the best solution. A set of initial solutions is produced and distributed in solution space for the optimal solutions. Here each population member is considered as a chromosome including several genes. Each gene shows the behavior and traits of the chromosome. A fitness function is defined for chromosomes based on objective functions. The calculation of the population member depends on the fitness value. For the next generation, the chromosome with a better fitness value is considered as the parent chromosome. The three main operators in orders will affect the new generation of the population. New solutions characterize the new population.
or chromosome. The selection, mutation, and cross-over are the three leading operators of the genetic algorithm. The initial solutions change depends on the convergence of the optimal solutions instill the fitness is acquired (Marchetti and Wanke 2020). The genetic algorithm model optimizes the solutions by repeating and applying defined operators. This algorithm has a global exploitation ability without evaluating many points in each space. It is possible to find a solution to a complex combinatorial optimization problem in this situation. The mutation operator prevents this algorithm from being trapped in the local optimum by increasing candidate solutions.

2.3.5.3.1. GA-SVM model. In order to build an effective SVM model, the parameters (C and $\gamma$) of the model need to be chosen properly in advance (Lin 2001). The parameter C determines the tradeoff cost between minimizing the training error and complexity of the SVM model. With a bigger C value, the predictive accuracy of the training sample is higher. However, this may cause an over-training problem. The parameter of the RBF kernel function defines a nonlinear mapping from input space to high-dimensional feature space. The value of effects the shape of RBF function. Hence, the parameters (C and $\gamma$) have a powerful influence on the efficiency and generalization performance of the SVM model. At present, the choice of the parameters lacks the guide of mature theory, mainly depending on experiences. A grid-search technique was presented by Lin (2001). However, the grid algorithm is time-consuming and does not perform very well (Gu et al. 2011).

According to some related research in different fields, GA is proved to be a better choice to determine the parameters (Lessmann et al. 2005). It can reduce the blindness of human-made choices and improve the predictive performance of the SVM model. Therefore, we choose GA to search for the optimal parameters of the SVM model for flash flood prediction in this study. The algorithm can be realized by a parameter optimization procedure designed by Li of Beijing Normal University based on the libsvm-mat toolbox developed by Lin of National Taiwan University (Chang and Lin 2001).

We built a single-factor SVM model for the flash flood prediction and determined the parameters (C and $\gamma$) of the model by GA. The GA had a generation number of 100, the population size of 20. The search range of C and $\gamma$ parameters is [0, 100]. We obtained a best C parameter of 6.8231, and a best $\gamma$ parameter of 0.12945. The model with the best parameters has the smallest mean square error (MSE) (0.00034).

We also built a multi-factor SVM model for the flash flood prediction and determined the parameters (C and $\gamma$) of the model by GA. The parameters of GA and the search range of C and $\gamma$ parameters are identical to those of the single-factor GA–SVM. We obtained a best C parameter of 63.6, and a best $\gamma$ parameter of 0.0045 and MSE 0.00021 for multi factors SVM by GA.

2.3.5.3.2. PSO-SVM model. Considering the sensitivity of SVM to the model parameters, the PSO method was proposed to establish the PSO–SVM prediction model in this paper. Then, the PSO was applied to search the SVM parameters: the penalty factor C, the kernel function $\gamma$ of the radial basis RBF and the parameters of the loss function p. Where the evolutionary generation quantity of the PSO algorithm was
100 and the population number was 20. The search result was: \( C_{\text{best}} = 23.3241, \gamma_{\text{best}} = 0.0265, p_{\text{best}} = 1743. \)

2.3.6. Validation and accuracy assessment. The validation and accuracy of the results have a crucial role in scientific works (Zhou et al. 2018; Xu et al. 2021; Liu et al. 2021; Cao et al. 2021). The validation and accuracy of predictive results and evaluation of MLA and optimization models applied in the study were conducted by AUROC, Accuracy, Sensitivity, Specificity, F-Score, Kappa, MAE (%), NSC (%). The ROC curve is the graphical presentation of model validation. The AUC of ROC is the quantitative measure of the graphical presentation. The following equations are incorporated for calculating the other statistical indices of model evaluation (Costache et al. 2020; Hoque et al. 2020; Lei et al. 2020; Pham et al. 2020; Chowdhuri, Pal, and Chakrabortty et al. 2020).

\[
ACC = \frac{TP + TN}{P + N} \tag{10}
\]

\[
SST = \frac{TP}{TP + FN} \tag{11}
\]

\[
SCF = \frac{TN}{FP + TN} \tag{12}
\]

\[
F_1 = \frac{2TP}{2TP + FP + FN} \tag{13}
\]

\[
Kappa = \frac{P_a - P_{\text{exp}}}{1 - P_{\text{exp}}} \tag{14}
\]

\[
P_a = \frac{TP + TN}{TP + TN + FN + FP} \tag{15}
\]

\[
P_{\text{exp}} = \frac{(TP + FN)(TP + FP) + (FP + TN)(FN + TN)}{\sqrt{(TP + TN + FN + FP)^2}} \tag{16}
\]

\[
AUC = \frac{\sum TP + \sum TN}{P + N} \tag{17}
\]

Where, TP is the true positive, TN is true negative, FP is false positive, FN is false negative and P and N signify the presence and absence of flood, accordingly (Arabameri et al. 2020; Arabameri et al. 2021).
3. Results

3.1. Multi-collinearity analysis

As individual variables are associated in a predictive model, multi-collinearity occurs. Since factors must be independent, this association is a concern. Based on which control variables would be in the framework, the estimated coefficients will fluctuate wildly. The coefficient gets markedly susceptible to the model’s sudden variations. It decreases the accuracy of the predicted coefficients, thus reducing the regression model’s predictive strength. The multi-collinearity of the multiple parameters was performed by considering VIF and TOL. The ranges of TOL and VIF in this flood susceptibility assessment are 0.205 to 0.897 and 1.114 to 4.772, respectively (Table 2). With the basis of TOL limit and VIF there is no problem of multicollinearity in this flood susceptibility assessment.

3.2. Spatial mapping of gully erosion susceptibility

The flood susceptibility mapping in this region was done by considering SVM, PSO, GA, SVM-GA, and PSO-GA algorithms. In all the predictive models, the northern portion is mainly associated with very high to high flood susceptible zone, and the rest of the region is associated with moderate to very low flood susceptible zones (Figure 4). The spatial coverage of very low, low, moderate, high and very high flood susceptible zone in SVM model are 1127.89 km² (30.92%), 1151.20 km² (31.55%), 728.00 km² (19.95%), 260.46 km² (7.41%) and 380.78 km² (10.17%) (Table 3). The spatial coverage of very low, low, moderate, high and very high flood susceptible zone in PSO model are 1034.03 km² (28.34%), 1270.35 km² (34.82%), 781.25 km² (21.41%), 177.26 km² (4.86%) and 385.43 km² (10.57%). The spatial coverage of very low, low, moderate, high and very high flood susceptible zone in GA model are 1081.49 km² (29.64%), 1100.61 km² (30.17%), 662.68 km² (18.61%), 355.63 km² (9.75%) and 447.91 km² (12.28%). The spatial coverage of very low, low, moderate, high and very high flood susceptible zone in ensemble of SVM-GA model are 888.70 km² (24.36%), 1163.67 km² (31.90%), 962.03 km² (26.37%), 306.81 km² (8.41%)

| FCFs          | Tolerance | VIF  |
|---------------|-----------|------|
| TWI           | 0.205     | 4.878|
| TPI           | 0.414     | 2.418|
| TST           | 0.250     | 3.993|
| TRI           | 0.213     | 4.700|
| SPI           | 0.230     | 4.772|
| Slope         | 0.221     | 4.520|
| Plan curvature| 0.481     | 2.077|
| NDVI          | 0.482     | 2.075|
| Elevation     | 0.277     | 3.604|
| Drainage density| 0.646    | 1.548|
| Distance to stream| 0.652   | 1.534|
| Convergence index | 0.753   | 1.327|
| Aspect        | 0.897     | 1.114|
| LULC          | 0.400     | 2.503|
| Geology       | 0.363     | 2.752|
and $327.11 \text{ km}^2$ (8.96%) (Figure 5). The spatial coverage of very low, low, moderate, high, and very high flood susceptible zone in the ensemble of PSO-GA model is $1045.38 \text{ km}^2$ (28.65%), $11.64 \text{ km}^2$ (31.93%), $823.44 \text{ km}^2$ (22.57%), $309.49 \text{ km}^2$ (8.48%) and $305.26 \text{ km}^2$ (8.37%).

### 3.3. Model evaluation performance result

In the method of statistical modeling, accuracy assessment plays a significant role. By selecting the correct indicators, the efficiency of a predictive outcome is measured and compared. Thus, it is important to use the correct parameters for a specific predictive approach to obtain an accurate performance. Evaluating appropriate predictive models is necessary since different types of information points are likely to be considered by the particular predictive model. In this analysis, AUC and different statistical indices, such as sensitivity (SST), specificity (SPF), accuracy (ACC), precession (PRE), F-score (F), and kappa index (KI) have been considered to find the most optimal model. Testing for these indices indicates that all four models worked well. The value
of different statistical indices as validation methods for the training and validation datasets is shown in Figure 6. In the training phase, the PSO-GA model has the highest SST (0.923), SPF (0.913), ACC (0.918), PRE (0.913), F (0.918), and KI (0.836) indicating the best performance model. This model correctly classified 92.3 per cent of flood cells and 91.3 per cent of non-flood cells showing an outstanding agreement between the excepted and observed flood events. The SVM-GA is the second performer model followed by GA, PSO, and SVM model considering all the statistical indices.

On the other hand, the model performance consistent with the training stage during the validation process. As a result, the PSO-GA model also the highest SST (0.923), SPF (0.902), ACC (0.912), PRE (0.901), F (0.911), and KI (0.825), indicating the best optimal flood susceptibility model. As the training stage, SVM-GA is the
second-ranked flood predicting model followed by GA, PSO, and SVM model considering all the statistical indices. The AUC values from ROC of SVM, PSO, GA, SVM-GA, and PSO-GA for the training datasets are 0.868, 0.884, 0.895, 0.901, and 0.916, respectively. The AUC values of SVM, PSO, GA, SVM-GA, and PSO-GA for the validation datasets are 0.839, 0.851, 0.874, 0.886, and 0.902, respectively (Figure 7). The PSO-GA model has 91.6 percent success for running the flood susceptibility model and 90.2 percent accurate in predicting the flood occurrences. From the analysis, it is pretty clear that the novel ensemble PSO-GA model is the best optimal for flood susceptibility model than the other ensemble and standalone models. Another ensemble model, SVM-GA, had the second-highest goodness-of-fit and predictive performance with the AUC values of 90.1 (training) and 88.6 (validation), respectively. As a result, combining evolutionary optimization algorithms with a machine learning model outperforms standalone SVM and PSO models. Even though SVM and PSO are highly structured and capable of solving multidimensional problems, their performance was unsatisfactory. However, when combined with metaheuristic algorithms, both goodness-of-fit and predictive performance improve to acceptable levels by resolving the parameter tuning and optimization issues of SVM and PSO.

3.4. Relative importance of the variables

The relative importance of the selected variables was estimated by considering the RF algorithm. The importance of some variables i.e. drainage density, LULC, TWI, TRI, distance to stream, plan curvature, and NDVI in flood susceptibility assessment, is maximum and the important values of these parameters are 0.890, 0.760, 0.612, 0.564, 0.502, 0.453 and 0.421 respectively. Rests of the parameters, i.e. convergence index, TST, elevation, slope, geology, TPI, SPI, and aspect, are associated with moderate to lower importance in flood susceptibility assessment; the importance values of these parameters are 0.394, 0.358, 0.312, 0.284, 0.222, 0.204, 0.151 and 0.124 respectively (Figure 8).

4. Discussion

Flood is one of the most dangerous natural hazards that frequently occur in regions with high slopes like those in Iran (Pourghasemi et al. 2019). Flood susceptibility
mapping is an important tool for management after flood occurrence and for effective planning before floods (Saha et al. 2021). Sometimes the complex conditions make it very difficult to make reliable predictions of flood occurrences and causative factors (Ouma and Tateishi 2014).

The flood generating factors are complex, and this mechanism and is different in different regions (Merz et al. 2010). Generally, they are considered as internal and external factors. The factors that are common for flood in our study area are: slope, slope aspect, elevation, plan curvature, land use and land cover, differences vegetation index, convergence index, topographical wetness index, topographic positioning Index, drainage density, distance to stream, terrain ruggedness index, terrain surface texture, geology, and stream power index. Among these, the most important causatives by the model performances are drainage density, land use, and land cover. The changes in land-use patterns and hydrological processes also play a major role in flash floods. External factor such as human activities also has a major control on flooding. The destruction of grassland or forest, road construction, and rapid settlement on the riverbank side are also major threats to flood. If the input data are more reliable, then the result will be more accurate. However, there is no standard rule for selecting non-flood locations and a random process selected these. Many previous studies noted that non-flood points are that locations where no flood occurred. Future studies require improved input data to select reliable non-flood locations.

In this study, to prepare the flood susceptibility map, we used three machine learning models and two ensembles. The machine learning model performances and success rate were compared through the validation such as particle swarm optimization (success rate 0.85% and prediction rate 0.88%), support vector machine (success rate 0.839% prediction rate 0.86%), genetic algorithm (success rate 0.87% and prediction rate 0.89%). The novel ensembles (SVM-GA and PSO-GA) for mapping flood

Figure 8. Relative importance of FCFs.
susceptibility were also considered in the Kaiser watershed. In this work, the ensemble model PSO-GA (prediction rate 0.91% and success rate 0.90%) and SVM-GA (prediction rate 0.90% and 0.88%) yielded better results than the single model. The result also demonstrated that the ensembles outperformed other models with higher accuracy. The percentage of susceptibility classes varies from very high to very low classes. The spatial distribution of these three models and ensemble performance are as follows: vulnerability zoning by SVM (very low 30.92%, very high 10.17% area) vulnerability zone by PSO (very low 28.34% and very high 10.57%), vulnerable are by GA (very low 29.64% and very high 12.28%), vulnerable zone prepared by ensembles SVM-GA (very low 24.36% and very high 8.96%) and vulnerable zone by PSO-GA (very low 28.65% and very high 8.37%). These data suggest that more practical and reliable results are yielded by the main models. The ML method can reduce cost and time with an effective mitigation plan in flood-targeted areas. Some factors are most influential in flood susceptibility assessment; drainage density, LULC, TWI, TRI, distance from the stream, etc. The different resecures established similar kind of findings in different parts of the world (Shafapour Tehrany et al. 2019; Tehrany et al. 2019). There is a direct impact of LULC on flood occurrences has been already established by different researchers (Roy et al. 2021). Flood susceptibility prediction research is currently being undertaken in America, Europe, Africa, and Asia (Alfieri et al. 2017). Though the effects of climate change coupled with the influence of predicted LULC on flooding are still debatable, the effects of climatic instability and seasonality necessitate further investigation (Roy, Chandra Pal, Chakrabortty et al. 2020). From the storm rainfall event and its associated flash flood, the large scale erosion can be continued and which is a negative effect on the environment and its associated ecosystem (Pal and Chakrabortty 2019a; Pal and Chakrabortty 2019b; Chakrabortty et al. 2020; Chakrabortty, Pal, Chowdhuri, Pal, and Chakrabortty et al. 2020). The changing pattern of LULC and conversion of forest cover into settlement and agricultural areas can drastically increase the amount and intensity of flood. On the other hand, the increase in impervious areas can changes the nature of runoff and, the increase in impervious areas can changes the nature of runoff, which ultimately influences the magnitude of the flood.

Our study may be beneficial for land-use planners and governmental policymakers to adopt suitable measures for flood management. The decision on the construction or settlement zones as well as dams and other structures in the susceptible or vulnerable zones can also be evaluated based on our results. These time-saving and modern approaches in machine learning can help reduce the loss of life and property and financial liabilities.

5. Conclusion

The flood susceptibility map formulation is a non-engineering approach for preventing floods and reducing socio-economic losses. This method combined both synthetic and analytical methods by considering numerous factors that lead to flooding. The most prominent model can be considered for solving uncertainty between quantitative and qualitative factors. This study focused on the Kaiser watershed located in the
Mazandaran province of northern Iran, where the flood susceptibility map was constructed by applying fifteen parameters. The main goal of this study is to evaluate three models (SVM, PSO, and GA) and ensemble models (PSO-GA and SVM-GA). The results show that the ensemble models represent better accuracy with a value of PSO-GA (0.92) and SVM-GA (0.886). Except for ensemble SVM-GA and PSO-GA, all the probability maps appeared to be visually similar. This could be due to the inefficiency of the standalone model to detect the probable flood areas precisely.

The ensemble model is good at predicting flood locations, as evidenced by the seven types of validation methods. The best combination for a sound flood susceptibility model is to consider more precise flood causative factors, locations of historical flood events, and new ensemble machine learning models. So, shortly, flood susceptibility analysis, landslide susceptibility analysis, soil erosion, gully erosion, and potentiality analysis of agriculture, mineral, and groundwater could all be done using the above mentioned ensemble methods. Because of the lack of spatial flood location data in this study, the flood susceptibility maps do not exceed 90% accuracy in predicting future flood occurrences. We obtained the specific indicators for causing the flood from the model performances. The drainage density (0.89), land use and land cover (0.76), and topographical wetness index are revealed as the more responsible factors for causing flood in this region. The resulting map has been classified to different conditions of flood in this region. Moreover, for the evaluation of the functional approaches compared to the traditional approach, the flood susceptibility map using various machine learning models and ensembles and result was validated. A reliable flood susceptibility map provides a valuable tool for land-use managers and governmental agencies to adopt effective management measures for reducing the risk. Our results could be helpful in the land use prior to the planning of these areas, which are likely to be affected by the flood. In conclusion, the result is attained from the study, demonstrating the superiority of the machine learning model and flood susceptibility mapping through integrated ensemble models yields with better accuracy will help prevent and mitigate future floods.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This research was partly funded by the Austrian Science Fund (FWF) through the Doctoral College GIScience (DK W 1237-N23) at the University of Salzburg.

Data availability
The data that support the findings of this study are available on request from the authors.

References
Abido MA. 2002. Optimal power flow using particle swarm optimization. Int J Electric Power Energy Syst. 24(7):563–571.
Alfieri L, Bisselink B, Dottori F, Naumann G, de Roo A, Salamon P, Wyser K, Feyen L. 2017. Global projections of river flood risk in a warmer world. Earths Future. 5(2):171–182.

Arabameri A, Asadi Nalivan O, Chandra Pal S, Chakrabortty R, Saha A, Lee S, Pradhan B, Tien Bui D. 2020. Novel machine learning approaches for modelling the gully erosion susceptibility. Remote Sens. 12(17):2833.

Arabameri A, Pal SC, Rezaie F, Chakrabortty R, Chowdhuri I, Blaschke T, Ngo PTT. 2021. Comparison of multi-criteria and artificial intelligence models for land-subsidence susceptibility zonation. J Environ Manage. 284:112067.

Armaghani DJ, Hajihassani M, Bejbarbaneh BY, Marto A, Mohamad ET. 2014. Indirect measure of shale shear strength parameters by means of rock index tests through an optimized artificial neural network. Measurement. 55:487–498.

Breiman L. 2001. Random forests. Machine Learn. 45(1):5–32.

Bui DT, Hoang N-D, Martínez-Álvez F, Ngo P-TT, Hoa PV, Pham TD, Samui P, Costache R. 2020. A novel deep learning neural network approach for predicting flash flood susceptibility: a case study at a high frequency tropical storm area. Sci Total Environ. 701:134413.

Bui DT, Ngo P-TT, Pham TD, Jaafari A, Minh NQ, Hoa PV, Samui P. 2019. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. CATENA. 179:184–196.

Bui DT, Pham BT, Nguye QP, Hoang ND. 2016. Spatial prediction of rainfall-induced shallow landslides using hybrid integration approach of least-squares support vector machines and differential evolution optimization: a case study in central Vietnam. Int J Digit Earth. 9:1077–1097.

Burn DH. 1990. Evaluation of regional flood frequency analysis with a region of influence approach. Water Resour Res. 26(10):2257–2265.

Cao B, Zhang W, Wang X, Zhao J, Gu Y, Zhang Y. 2021. A memetic algorithm based on two_Arch2 for multi-depot heterogeneous-vehicle capacitated arc routing problem. Swarm Evolution Comput. 63:100864.

Campolo M, Soldati A, Andreussi P. 2003. Artificial neural network approach to flood forecasting in the River Arno. Hydrol Sci J. 48(3):381–398.

Caprario J, Finotti AR. 2019. Socio-technological tool for mapping susceptibility to urban flooding. J Hydrol. 574:1152–1163.

Chakrabortty R, Pal SC, Chowdhuri I, Malik S, Das B. 2020. Assessing the importance of static and dynamic causative factors on erosion potentiality using SWAT, EBF with uncertainty and plausibility, logistic regression and novel ensemble model in a sub-tropical environment. J Indian Soc Remote Sens. 48(5):765–789.

Chakrabortty R, Pal SC, Sahana M, Mondal A, Dou J, Pham BT, Yunus AP. 2020. Soil erosion potential hotspot zone identification using machine learning and statistical approaches in eastern India. Nat Hazards. 104(2):1259–1294.

Chakrabortty R, Pradhan B, Mondal P, Pal SC. 2020. The use of RUSLE and GCMs to predict potential soil erosion associated with climate change in a monsoon-dominated region of eastern India. Arab J Geosci. 13(20):1–20.

Choubin B, Mosavi A, Alamdarloo EH, Hosseini FS, Shamshirband S, Dashtekian K, Ghamisi P. 2019. Earth fissure hazard prediction using machine learning models. Environ Res. 179(Pt A):108770.

Chao L, Zhang K, Wang J, Feng J, Zhang M. 2021. A Comprehensive evaluation of five evapotranspiration datasets based on ground and GRACE satellite observations: implications for improvement of evapotranspiration retrieval algorithm. Remote Sens. 13(12):2414.

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 Software. 95:229–245.

Chen W, Hong H, Li S, Shahabi H, Wang Y, Wang X, Ahmad BB. 2019. Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles. J Hydrol. 575:864–873.
Chen Z, Liu Z, Yin L, Zheng W. 2022. Statistical analysis of regional air temperature characteristics before and after dam construction. Urban Clim. 41:101085.

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, Arabameri A, Saha A, Chakrabortty R, Blaschke T, Pradhan B, Band S. 2020. Implementation of artificial intelligence based ensemble models for gully erosion susceptibility assessment. Remote Sens. 12(21):3620.

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 Space Res. 65(5):1466–1489.

Chowdhuri I, Pal SC, Chakrabortty R, Malik S, Das B, Roy P. 2021. Torrential rainfall-induced landslide susceptibility assessment using machine learning and statistical methods of eastern Himalaya. Nat Hazards. 107(1):697–722.

Chang C. C, Lin C. J. 2001. LIBSVM: a library for support vector machines. ACM Trans Int Syst Technol. 2:1–39.

Costache R, Pham QB, Avand M, Linh NTT, Vojtek M, Vojteková J, Lee S, Khoi DN, Nhi PTT, Dung TD. 2020. Novel hybrid models between bivariate statistics, artificial neural networks and boosting algorithms for flood susceptibility assessment. J Environ Manage. 265: 110485.

Dano UL, Balogun A-L, Matori A-N, Wan Yusouf K, Abubakar IR, Said Mohamed MA, Aina YA, Pradhan B. 2019. Flood susceptibility mapping using GIS-based analytic network process: a case study of Perlis, Malaysia. Water. 11(3):615.

Gao Y, Huang J, Li S, Li S. 2012. Spatial pattern of non-stationarity and scale-dependent relationships between NDVI and climatic factors—A case study in Qinghai-Tibet Plateau, China. Ecol Indic. 20:170–176.

Gebrehiwot A, Hashemi-Beni L, Thompson G, Kordjamshidi P, Langan TE. 2019. Deep convolutional neural network for flood extent mapping using unmanned aerial vehicles data. Sensors. 19(7):1486.

Ghorbanzadeh, Omid, Thomas Blaschke, Khalil Gholamnia, Sansar Raj Meena, Dirk Tiede, and Jagannath Aryal. 2019. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. Remote Sensing. 11(2): 196.

Gu J, Zhu M, Jiang L. 2011. Housing price forecasting based on genetic algorithm and support vector machine. Expert Syst Appl. 38(4):3383–3386.

Hawley RJ, Bledsoe BP. 2011. How do flow peaks and durations change in suburbanizing semi-arid watersheds? A southern California case study. J Hydrol. 405(1-2):69–82.

Heitmuller FT, Hudson PF, Asquith WH. 2015. Lithologic and hydrologic controls of mixed alluvial–bedrock channels in flood-prone fluvial systems: bankfull and macrochannels in the Llano River watershed, central Texas, USA. Geomorphology. 232:1–19.

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

Hoque MA-A, Pradhan B, Ahmed N. 2020. Assessing drought vulnerability using geospatial techniques in northwestern part of Bangladesh. Sci Total Environ. 705:135957.

Huang Z, Zhou J, Song L, Lu Y, Zhang Y. 2010. Flood disaster loss comprehensive evaluation model based on optimization support vector machine. Expert Syst Appl. 37(5):3810–3814.

IRIMO. 2012. I.R. of Iran Meteorological Organization [Internet]. http://www.semnanmet.ir.

Jiang S, Zuo Y, Yang M, Feng R. 2021. Reconstruction of the Cenozoic tectono-thermal history of the Dongpu Depression, Bohai Bay Basin, China: constraints from apatite fission track and vitrinite reflectance data. J Petrol Sci Eng. 205:108809.

Khosravi K, Pham BT, Chapi K, Shirzadi A, Shahabi H, Revhaug I, Prakash I, Bui DT. 2018. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. Sci Total Environ. 627:744–755.
Lei X, Chen W, Avand M, Janizadeh S, Kariminejad N, Shahabi H, Costache R, Shahabi H, Shirzadi A, Mosavi A. 2020. GIS-Based Machine Learning Algorithms for Gully Erosion Susceptibility Mapping in a Semi-Arid Region of Iran. Remote Sens. 12(15):2478.

Lessmann S, Stahlbock R, Crone SF. 2005. Optimizing hyperparameters of support vector machines by genetic algorithms. In IC-AI, p. 74–82.

Li Z, Zhang K. 2008. Comparison of three GIS-based hydrological models. J Hydrol Eng. 13(5):364–370.

Lin G-F, Chou Y-C, Wu M-C. 2013. Typhoon flood forecasting using integrated two-stage Support Vector Machine approach. J Hydrol. 486:334–342.

Lin PT. 2001. Support vector regression: systematic design and performance analysis. [Unpublished Doctoral Dissertation]. Department of Electronic Engineering, National Taiwan University.

Liu D, Fan Z, Fu Q, Li M, Faiz MA, Ali S, Li T, Zhang L, Khan MI. 2020. Random forest regression evaluation model of regional flood disaster resilience based on the whale optimization algorithm. J Cleaner Prod. 250:119468.

Liu L, Moayedi H, Rashid ASA, Rahman SSA, Nguyen H. 2020. Optimizing an ANN model with genetic algorithm (GA) predicting load-settlement behaviours of eco-friendly raft-pile foundation (ERP) system. Eng Comput. 36(1):421–433.

Liu L, Xiang H, Li X. 2021. A novel perturbation method to reduce the dynamical degradation of digital chaotic maps. Nonlinear Dyn. 103(1):1099–1115.

Mahmoud SH, Gan TY. 2018. Urbanization and climate change implications in flood risk management: developing an efficient decision support system for flood susceptibility mapping. Sci Total Environ. 636:152–167.

Malik S, Pal SC, Das B, Chakrabortty R. 2019. Intra-annual variations of vegetation status in a sub-tropical deciduous forest-dominated area using geospatial approach: a case study of Sali watershed, Bankura, West Bengal, India. Geol Ecol Landscape. 2019:1–12.

Marchetti D, Wanke P. 2020. Efficiency of the rail sections in Brazilian railway system, using TOPSIS and a genetic algorithm to analyse optimized scenarios. Transp Res E: Logist Transp Rev. 135:101858.

Merz B, Hall J, Disse M, Schumann A. 2010. Fluvial flood risk management in a changing world. Nat Hazards Earth Syst Sci. 10(3):509–527.

Mi C, Chen J, Zhang Z, Huang S, Postolache O. 2022. Visual sensor network task scheduling algorithm at automated container terminal. IEEE Sensors J. 22(6):6042–6051.

Mind’je R, Li L, Amanambu AC, Nahayo L, Nsengiyumva JB, Gasirabo A, Mindje M. 2019. Flood susceptibility modeling and hazard perception in Rwanda. Int J Disaster Risk Reduct. 38:101211.

Mohammadi A, Shahabi H, Bin Ahmad B. 2019. Land-cover change detection in a part of Cameron Highlands, Malaysia using ETM + satellite imagery and support vector machine (SVM) algorithm. EnvironmentAsia. 12(2):145–154.

Moayedi H, Abdullahi MM, Nguyen H, Rashid ASA. 2021. Comparison of dragonfly algorithm and Harris hawks optimization evolutionary data mining techniques for the assessment of bearing capacity of footings over two-layer foundation. Eng Comput. 37(1):437–447.

Moayedi H, Gör M, Lyu Z, Bui DT. 2020. Herding behaviors of grasshopper and Harris hawk for hybridizing the neural network in predicting the soil compression coefficient. Measurement. 152:107389.

Moayedi H, Mosallanezhad M. 2017. Uplift resistance of belled and multi-belled piles in loose sand. Measurement. 109:346–353.

Mühlenbein H. 1997. Genetic algorithms.

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. 6(6):1515–1545.

Pal SC, Chakrabortty R. 2019a. Simulating the impact of climate change on soil erosion in sub-tropical monsoon dominated watershed based on RUSLE, SCS runoff and MIROC5 climatic model. Adv Space Res. 64(2):352–377.
Pal SC, Chakrabortty R. 2019b. Modeling of water induced surface soil erosion and the potential risk zone prediction in a sub-tropical watershed of Eastern India. Model Earth Syst Environ. 5(2):369–393.

Pal SC, Chakrabortty R, Malik S, Das B. 2018. Application of forest canopy density model for forest cover mapping using LISS-IV satellite data: a case study of Sali watershed, West Bengal. Model Earth Syst Environ. 4(2):853–865.

Pham BT, Avand M, Janizadeh S, Phong TV, Al-Ansari N, Ho LS, Das S, Le HV, Amini A, Bozchaloei SK, et al. 2020. GIS based hybrid computational approaches for flash flood susceptibility assessment. Water. 12(3):683.

Pourghasemi HR, Gayen A, Panahi M, Rezaie F, Blaschke T. 2019. Multi-hazard probability assessment and mapping in Iran. Sci Total Environ. 692:556–571.

Roy P, Chandra Pal S, Arabameri A, Chakrabortty R, Pradhan B, Chowdhuri I, Lee S, Tien Bui D. 2020. Novel ensemble of multivariate adaptive regression spline with spatial logistic regression and boosted regression tree for gully erosion susceptibility. Remote Sens. 12(20):3284.

Roy P, Chandra Pal S, Chakrabortty R, Chowdhuri I, Malik S, Das B. 2020. Threats of climate and land use change on future flood susceptibility. J Cleaner Prod. 272:122757.

Roy P, Pal SC, Arabameri A, Rezaie F, Chakrabortty R, Chowdhuri I, Saha A, Malik S, Das B. 2021. Climate and land use change induced future flood susceptibility assessment in a sub-tropical region of India. Soft Comput. 25(8):5925–5925.

Saha A, Pal SC, Arabameri A, Blaschke T, Panahi S, Chowdhuri I, Chakrabortty R, Costache R, Arora A. 2021. Flood susceptibility assessment using novel ensemble of hyperpipes and support vector regression algorithms. Water. 13(2):241.

Shafapour Tehrany M, Kumar L, Neamah Jebur M, Shabani F. 2019. Evaluating the application of the statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods. Geomatics Nat Hazards Risk. 10(1):79–101.

Tehrany MS, Jones S, Shabani F. 2019. Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. Catena. 175:174–192.

Tehrany MS, Lee M-J, Pradhan B, Jebur MN, Lee S. 2014. Flood susceptibility mapping using integrated bivariate and multivariate statistical models. Environ Earth Sci. 72(10):4001–4015.

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.

Vapnik V, Guyon I, Hastie T. 1995. Support vector machines. Mach Learn. 20(3):273–297.

Wang Y, Hong H, Chen W, Li S, Panahi M, Khosravi K, Shrizadi 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 P, Wang L, Leung H, Zhang G. 2021. Super-resolution mapping based on spatial-spectral correlation for spectral imagery. IEEE Trans Geosci Remote Sens. 59(3):2256–2268.

Whitley D. 1994. A genetic algorithm tutorial. Stat Comput. 4(2):65–85.

Xiong Z, Mo F, Zhao X, Xu F, Zhang X, Wu Y. 2022. Dynamic texture classification based on 3D ICA-learned filters and fisher vector encoding in big data environment. J Sign Process Syst.

Xu J, Wu Z, Chen H, Shao L, Zhou X, Wang S. 2021. Study on strength behavior of basalt fiber-reinforced loess by digital image technology (DIT) and scanning electron microscope (SEM). Arab J Sci Eng. 46(11):11319–11338.

Yang L, Xiong Z, Liu G, Hu Y, Zhang X, Qiu M. 2022. An analytical model of page dissemination for efficient big data transmission of C-ITS. IEEE Trans Intell Transport Syst.

Yin L, Wang L, Keim BD, Konsoer K, Zheng W. 2022. Wavelet analysis of dam injection and discharge in three gorges dam and reservoir with precipitation and river discharge. Water. 14(4):567.

Youssef AM, Pradhan B, Sefry SA. 2016. Flash flood susceptibility assessment in Jeddah city (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models. Environ Earth Sci. 75(1):12.
Zrinji Z, Burn DH. 1994. Flood frequency analysis for ungauged sites using a region of influence approach. J Hydrol. 153(1-4):1–21.
Zhao T, Shi J, Lv L, Xu H, Chen D, Cui Q, Jackson T, Yan G, Jia L, Chen L, et al. 2020. Soil moisture experiment in the Luan River supporting new satellite mission opportunities. Remote Sens Environ. 240:111680.
Zhao X, Xia H, Pan L, Song H, Niu W, Wang R, Li R, Bian X, Guo Y, Qin Y. 2021. Drought monitoring over yellow river basin from 2003–2019 using reconstructed MODIS land surface temperature in Google Earth engine. Remote Sens. 13(18):3748.
Zhao X, Xia H, Pan L, Song H, Niu W, Wang R, Li R, Bian X, Guo Y, Qin Y. 2021. Drought monitoring over yellow river basin from 2003–2019 using reconstructed MODIS land surface temperature in Google Earth engine. Remote Sens. 13(18):3748.
Zhao X, Xia H, Pan L, Song H, Niu W, Wang R, Li R, Bian X, Guo Y, Qin Y. 2021. Drought monitoring over yellow river basin from 2003–2019 using reconstructed MODIS land surface temperature in Google Earth engine. Remote Sens. 13(18):3748.
Zhao X, Xia H, Pan L, Song H, Niu W, Wang R, Li R, Bian X, Guo Y, Qin Y. 2021. Drought monitoring over yellow river basin from 2003–2019 using reconstructed MODIS land surface temperature in Google Earth engine. Remote Sens. 13(18):3748.
Zhao X, Xia H, Pan L, Song H, Niu W, Wang R, Li R, Bian X, Guo Y, Qin Y. 2021. Drought monitoring over yellow river basin from 2003–2019 using reconstructed MODIS land surface temperature in Google Earth engine. Remote Sens. 13(18):3748.
Zhao X, Xia H, Pan L, Song H, Niu W, Wang R, Li R, Bian X, Guo Y, Qin Y. 2021. Drought monitoring over yellow river basin from 2003–2019 using reconstructed MODIS land surface temperature in Google Earth engine. Remote Sens. 13(18):3748.
Zhao X, Xia H, Pan L, Song H, Niu W, Wang R, Li R, Bian X, Guo Y, Qin Y. 2021. Drought monitoring over yellow river basin from 2003–2019 using reconstructed MODIS land surface temperature in Google Earth engine. Remote Sens. 13(18):3748.
Zhao X, Xia H, Pan L, Song H, Niu W, Wang R, Li R, Bian X, Guo Y, Qin Y. 2021. Drought monitoring over yellow river basin from 2003–2019 using reconstructed MODIS land surface temperature in Google Earth engine. Remote Sens. 13(18):3748.