Flood Susceptibility Mapping a Novel Hybrid Approach Based On Swarm Intelligence Optimized

Subhashree Sukla[1], R Subba Rao[2] and Amitav Saran[3]
[1,3]Associate Professor, Department of Computer Science Engineering, Gandhi Institute For Technology (GIFT), Bhubaneswar
[2]Assistant Professor, Department of Computer Science Engineering, Gandhi Engineering College, Bhubaneswar

Publishing Date: August 27, 2015

A R T I C L E I N F O

Keywords:
Environmental modeling
Extreme learning
machines Particle swarm
optimization Flash flood
Vietnam

A B S T R A C T

Flash flood is a typical natural hazard that occurs within a short time with high flow velocities and is difficult to predict. In this study, we propose and validate a new soft computing approach that is an integration of an Extreme Learning Machine (ELM) and a Particle Swarm Optimization (PSO), named as PSO-ELM, for the spatial prediction of flash floods. The ELM is used to generate the initial flood model, whereas the PSO was employed to optimize the model. A high frequency tropical typhoon area at Northwest of Vietnam was selected as a case study. In this regard, a geospatial database for the study area was constructed with 654 flash flood locations and 12 influencing factors (elevation, slope, aspect, curvature, toposhade, topographic wetness index, stream power index, stream density, NDVI, soil type, lithology, and rainfall). The model performance was validated using several evaluators such as kappa statistics, root-mean-square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), and area under the ROC curve (AUC-ROC) and compared to three state-of-the-art machine learning techniques, including multilayer perceptron neural networks, support vector machine, and C4.5 decision tree. The results revealed that the PSO-ELM model has high prediction performance (kappa statistics = 0.801, RMSE = 0.281; MAE = 0.079, R^2 = 0.829, AUC-ROC = 0.954) and successfully outperformed the three machine learning models. We conclude that the proposed model is a new tool for the prediction of flash flood susceptibility at high frequency tropical typhoon areas.

1. Introduction

Flooding is one of the most hazardous natural disasters due to its widespread social and economic impacts, especially in the high frequency tropical cyclone regions of south and east Asia (Bubeck and Thieken, ; Hu et al.,; Peduzzi, ; Ward et al.,). Some estimations suggest that the flooding damage during the year 2013 was over 50 billion USD worldwide (Wasko and Sharma,). The damage may even be exuberated due to climate change induced by heavy rainfall and unplanned land-use activities (Bubeck and Thieken, ; Hartnett and Nash,). Among different types of floods, flash flooding is extremely dangerous due to its rapid onset characteristics and high flow velocities, which cause fatalities, economic losses, and severe damage to the environment (Destro et al., ; Jalayer et al., ; Santo et al.,). However, it remains a difficult task to accurately predict flash flood due to the complex nature of this phenomenon (Edouard et al.,). Thus, it is necessary to develop a high accuracy model for predicting the probability of flash flood occurrence and mapping its susceptibility. The result can assist the local authorities and the decision-makers in disaster risks management and for mitigating the effects of climate change.

Various studies have been conducted to predict the probability of flood occurrences that can be grouped into three main groups, traditional analysis, rainfall–runoff, and pattern classification (Tien Bui and Hoang,). In the first group, long-period time series data at hydrological stations are used to generate regression models. Therefore,

these models are capable to forecast discharge in both space and time scales. Although these models are typically used to predict flood problems for large-scale basins, the lack of long-period data often limits their applications. The second group of the models (e.g., HEC-RAS (Brunner, 1995) and MIKE (Zhou et al., 2012)) focuses on establishing relations between rainfall and runoff to predict both spatial and temporal floods. They are the most widely used models for flood forecasting and management. However, they still require long-period time series monitoring data at hydrological stations to achieve the highest level of predictive accuracy. In contrast to these models, the last group employs a new approach called "on-off" classification, which is independent of monitoring data at gauge stations. Using these models and based on the historical floods and geo-environmental data, a flood-prone area is delineated into the flood non-flood classes (Bui et al.,). Given the availability of multi-source remotely sensed data such optical and Synthetic Aperture Radar (SAR) images, various data-driven models have been proposed and suggested for flood prediction (de Musso et al., ; Khosravi et al., ; Lim and Lee, ; Tong et al.,). However, it is still difficult to accurately predict flash flooding in the mountainous areas.

In recent years, machine learning models have received attention for flood prediction due to their high prediction performance and capability to handle complicated relationships between input variables. These models include artificial neural networks (ANN) (Sahoo et al., 2006; Youssef et al., 2011), support vector machines (SVM) (Tehrany et al., 2015), logistic regression (Nandi et al), evidential belief function and decision trees (Rahmati and Pourghasemi), random forest and boosted-tree (Lee et al.,), GARP and QUEST (Darabi et al.,), WELLSVM (Zhao et al.,), and classification and re- gression trees (CART). However, some other studies indicate that a high level of predictive accuracy can be achieved via hybrid and ensemble machine learning models, e.g., ensemble of weights-of-evidence and support vector machines (Tehrany et al., 2014), hybrid Bayesian framework (Tien Bui and Hoang,), neuro-fuzzy system integrated with metaheuristic algorithms (Termeh et al., 2018; Tien Bui et al., b; Tien Bui et al.,), logistic model tree with bagging ensembles (Chapi et al.,), swarm optimized neural networks (Ngo et al., a), ensemble of multivariate discriminant analysis, CART, and SVM (Choubin et al.,), ensemble of multi-criteria decision making (Wang et al.,), and fuzzy rule based ensembles (Bui et al.,). A detailed review of the application of machine learning models for flood prediction can be found in Mosavi et al.

The recent advancements on machine learning have introduced the extreme learning machine (ELM) (Huang et al., 2004), which has proven its effectiveness in various real-word problems, though it has not yet been explored for flash flood modeling. The main advantage of the ELM is that it uses linear algebra to derive its output weights, instead of using the backpropagation algorithm or the Bayesian algorithm in traditional ANNs. In other words, the output weights are computed analytically, which results in a fast training process and high prediction performance in various environmental modeling problems. Deo and Şahin (2015) compared the ELM and ANNs for drought prediction and reported that the ELM produced better results. Abdullah et al. (2015) showed that the ELM is more efficient than feedforward back-propagation model in predicting evapotranspiration. Yaseen et al. reported that the ELM model has better accuracy for predicting compressive strength of concrete compared that that of multivariate adaptive regression spline (MARS), decision tree, and SVM models.

However, a disadvantage of the ELM model is that the weights of the input layer are randomly generated and the model do not update or optimize these weights in the training phase, causing an unstable performance during different training phases (Eshtay et al., ; Zhu et al., 2005). To overcome this limitation, metaheuristic optimization algorithms were suggested to search for the best parameters and optimize the weights of the input layer. Taormina and Chau (2015) integrated particle swarm optimization (PSO) and ELM for streamflow forecasting and reported that the integrated model outperforms the single ELM model. Niu et al. showed that PSO can significantly improve the prediction accuracy of the base ELM model in hydrologic analysis. Ghimire et al. combined self-adaptive differential evolutionary (SDE) with ELM for predicting solar radiation and showed an improved model performance than that of the SVM and Gaussian process models.

Nonetheless, to the best of our knowledge, the potential of the ELM model and its integration with optimization algorithms have not yet been explored for the prediction of flash flood susceptibility. Thus, this study was conducted to fill this gap by proposing and verifying a PSO optimized ELM model for the prediction of flash floods. PSO is selected because this is one of the most successful metaheuristic optimization algorithm used in literature (Cheng and Jin, 2015). In this regard, ELM is used to generate the model structure, while the PSO is employed to search and optimize the weights of the model. The application model is described via a case study from a high frequency tropical typhoon area in the northwest of Vietnam. Finally, the efficiency of the proposed PSO-EML model is evaluated using several performance metrics and compared to benchmarks, MLP-ANN, SVM, and C4.5 decision.

2. Mathematical background of the algorithms used

2.1. Extreme learning machines

Extreme learning machine (ELM) is a state-of-the-art learning algorithm for single-hidden layer feedforward neural networks (SLFNs), which has proven its remarkable performance in both classification and regression applications (Huang et al., 2006; Huang et al., 2012). This algorithm can produce good generalization performance at much faster learning speed than traditional least square support vector machine (LS-SVM) and proximal support vector machine (PSVM) (Huang et al., 2006).

Given pattern $\boldsymbol{x}$ of N input variables $\boldsymbol{x} \in R^N$, the output of an ELM with L hidden nodes and q out nodes is expressed as follows:

$$f_L(\boldsymbol{x}) = \sum_{i=1}^{L} \beta_i G(IW_i, BA_i, \boldsymbol{x}) \quad (1)$$

where IW, BA, OW, and G are the input weights, the bias, the output weights, and the activation function, respectively.

In ELM, the input weights are randomly generated, whereas the output weight, OW, is computed using the linear system (Eq. (2)) (Huang et al., 2011) below.

$$OW = \boldsymbol{H}\dagger\boldsymbol{T} \quad (2)$$

where $\boldsymbol{H}\dagger$ is the Moore–Penrose generalized inverse of $\boldsymbol{H}$.

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{h}(\boldsymbol{x}_1) \\ \vdots \\ \boldsymbol{h}(\boldsymbol{x}_N) \end{bmatrix} = \begin{bmatrix} G(\boldsymbol{a}_1, b_1, \boldsymbol{x}_1) & \dots & G(\boldsymbol{a}_L, b_L, \boldsymbol{x}_1) \\ \vdots & \dots & \vdots \\ G(\boldsymbol{a}_1, b_1, \boldsymbol{x}_N) & \dots & G(\boldsymbol{a}_L, b_L, \boldsymbol{x}_L) \end{bmatrix} N \times L \quad (3)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix} N \text{ x } q \text{ and } \boldsymbol{T} = \begin{bmatrix} \boldsymbol{t}_1^T \\ \vdots \\ \boldsymbol{t}_L^T \end{bmatrix} N \times q \quad (4)$$

where $\boldsymbol{H}$ is the output matrix of the EL; $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, …, $\boldsymbol{x}_N$ are flood conditioning factors; and T is the target values of the training dataset (i.e., flash flood and non-flash flood values).

2.2. Particle swarm optimization

This section presents the PSO algorithm that was used to optimize the ELM model for flash flood susceptibility mapping in this research. More specifically, PSO was used to optimize the input weights ($\boldsymbol{a}_i$) and bias of the input layer of the ELM model, whereas the weights of the output layer were computed by the ELM algorithm (the Moore–Penrose

generalized inverse mentioned in Section 2.1).

The PSO algorithm, proposed by Kennedy and Eberhart (1995), is a well-known stochastic optimization technique that works by initializing randomly a group of birds within a population (swarm) over the searching space called as a "particle". A hypothesis on the solution of the optimization problem can be given by the position of each particle in the swarm and is depicted by a different objective function value, i.e., *fitness value*. The value is defined when swarm flies across the error surface and particles are identified by their positions and velocities. These are vectors of the same number of elements as the dimension problems while the particles fly with a certain velocity and find the best position after some iterations. In each iteration, a particle adjusts its velocity vector following its momentum and the best position influence (P_b) on its neighbors (P_g), and then a new position that the particle is to fly is achieved.

Assuming the position of the i^{th} particle can be expressed as vector $\boldsymbol{X_i}$, $\boldsymbol{V_i}$ is its velocity, and $\boldsymbol{P}$ identifies the best positions of its neighbors, the employed velocity updates equation can be given by:

$$\boldsymbol{V_i} \leftarrow \chi\left(\boldsymbol{V_i} + \varphi\left(\left(\sum_{j=1}^{K_i} a_j \boldsymbol{P}_{nbrj}\right) - X_i\right)\right) \tag{5}$$

where the scalars χ and φ are constriction coefficients often set to 0.7298 and 4.1, respectively. K_i is the number of neighbors for particle i and nbr_j identifies i's j-th neighbor.

The contribution of each neighbor is weighted using a normalized random coefficient a_j which is the same for each particle dimension, and is defined as:

$$a_j = \frac{b_j}{\sum b_j},\ bj = U\ (0,1),\ \mathrm{j} = 1,\ \ldots, Ki \tag{6}$$

where U ranges from 0 to 1 identifying a uniform random number.

3. Study area and data used

3.1. General description of the study area

The study area is located in the northwest mountainous region of two districts namely, Bac Ha and Bao Yen (BHBY), in Lao Cai province, which is about 263 km from Hanoi, Vietnam. It lies between the longitudes 104°10′E and 105°37′E, and between the latitudes 22°5′N and 22°40′N, covering an area of about 1510.4 km^2 (Fig. 1).

The topography of the area is classified as hilly mountains, gorges, and highly dense rivers. The altitude ranges from 38.9 m to 1878.69 m above sea level with the mean is 538.1 m. Approximately 85.4% of the total study area has slopes ranging from 10 ^{o} to 40^o, while about 11.5% of the area has slopes less than 10^o and the areas with slopes greater than 40^o accounts for only 3.1% of the total study area (General Statistics Office, ; Tehrany et al., ; Tien Bui et al.,).

Administratively, the BHBY districts are divided into 37 communes and 2 towns and the total population is 136.06 thousand people. The average population densities were 83 people km^{-2} and 96 people km^{-2} for Bac Ha and Bao Yen, respectively (General Statistics Office,).

The study area is characterized by subtropical monsoon climate with two unique seasons, a rainy season lasts from November to April while a dry season starts from May to October. The average annual precipitation varies from 1440 mm to 2200 mm and the majority of precipitation generally occurs in the rainy reason. The study area located in a stormy center of the world and annually experiences heavy and extreme rainfalls. Although flash floods occur almost every year, there was a report of extreme flash floods during the last three days on October 10–12, with severe damage. The highest total precipitation at a measuring station was approximately 201 mm (General Statistics Office, 2017; Tehrany et al.,).

3.2. Data used

3.2.1. Flash-flood inventory map

Forecasting flash flooding using a machine learning approach requires a knowledge of the events that have occurred in the recent past (Borga et al., 2011; Khosravi et al., ; P. et al., 2011; Tien Bui and Hoang, 2017). Thus, preparation of a flash flooding inventory map is a compulsory task. However, flash flooding is generally characterized by both small temporal and spatial scales that are difficult to observe and map (Borga et al., 2011). Therefore, a collection of flash flooding inventories remains challenging in the current literature. Prior studies normally collected historical flash flood events during time-consuming and costly field surveys (e.g., Tehrany et al. (2015); Khosravi et al. . In this study, an inventory map with 654 flash flood locations, which was the outcomes from the stated-fund project of flash flooding (B2018-MDA-18DT) (Ngo et al.,), was used. Accordingly, flash flooding polygons occurred in the rainfall season of were identi- fied using the change detection techniques based on the remotely sensed data analysis of the Sentinel-1A synthetic aperture radar (SAR) imagery and the field surveys.

3.2.2. Influencing factors

Modeling of flash flooding requires characterizing geospatial information on key hydrological properties of the study area. Therefore, determination of influencing variables is an important task. In this study, a total number of twelve influencing factors were selected: elevation, slope, curvature, toposhade, aspect, topographic wetness index (TWI), stream power index (SPI), stream density, Normalized Difference Vegetation Index (NDVI), soil type, lithology, and rainfall.

Topography is a major factor of the hydrological process and strongly related to flash flooding events because its steep relief enhances the rapid concentration of waterflow (Destro et al.,). Therefore, topographic related factors i.e. elevation, slope, curvature, toposhade, aspect, TWI, and SPI should be used in such a modeling effort. In this study, a Digital Elevation Model (DEM) with a spatial resolution of 20 m for the study area was generated from the national topographic map with 1:50,000 scale provided by the Ministry of Natural Resources and Environment of Vietnam (MONRE). Seven geomorphometric factors shown in Fig. 2 were derived from the DEM.

Elevation and slope were selected in current work because the flow of water is driven by the gravity force, moving from higher to lower elevations (Ellabban et al., 2014), while slope controls the speeds of surface runoff. Additionally, flood-prone areas are usually located on the flat areas and low elevations (Tehrany et al., 2013). Curvature was employed since flash flooding areas are related to high topographic convergence (Manfreda et al., 2014). In this research, the maps of elevation and slope were prepared in ten classes (Fig. 2a, b), whereas the curvature map was classified into seven classes (Fig. 2c). These classes of the three maps were determined based on the natural break interval method available in ArcGIS. The aspect map (Fig. 2d) consisted of nine classes were used.

Toposhade and aspect were selected, as they are related to the convergence and directions of water flowing (Aryal and Mein, 2003). For this study, the toposhade and aspect maps (Fig. 2c, g) were categorized into nine classes. The TWI (Fig. 2e) and the SPI (Fig. 2f) are typical hydrological parameters that influence both the flow intensity and water accumulation (Martınez-Casasnovas et al., 2004). Therefore, they were selected for flash flood modeling in this research. The TWI (Beven et al., 1984) and the SPI (Moore et al., 1991) were calculated using the following equations:

$$TWI = \ln(a/\tan\beta) \tag{7}$$

$$SPI = A_s \tan\beta \tag{8}$$

where a is local upslope area, β is local slope, and $A_s$is the specific catchment area.

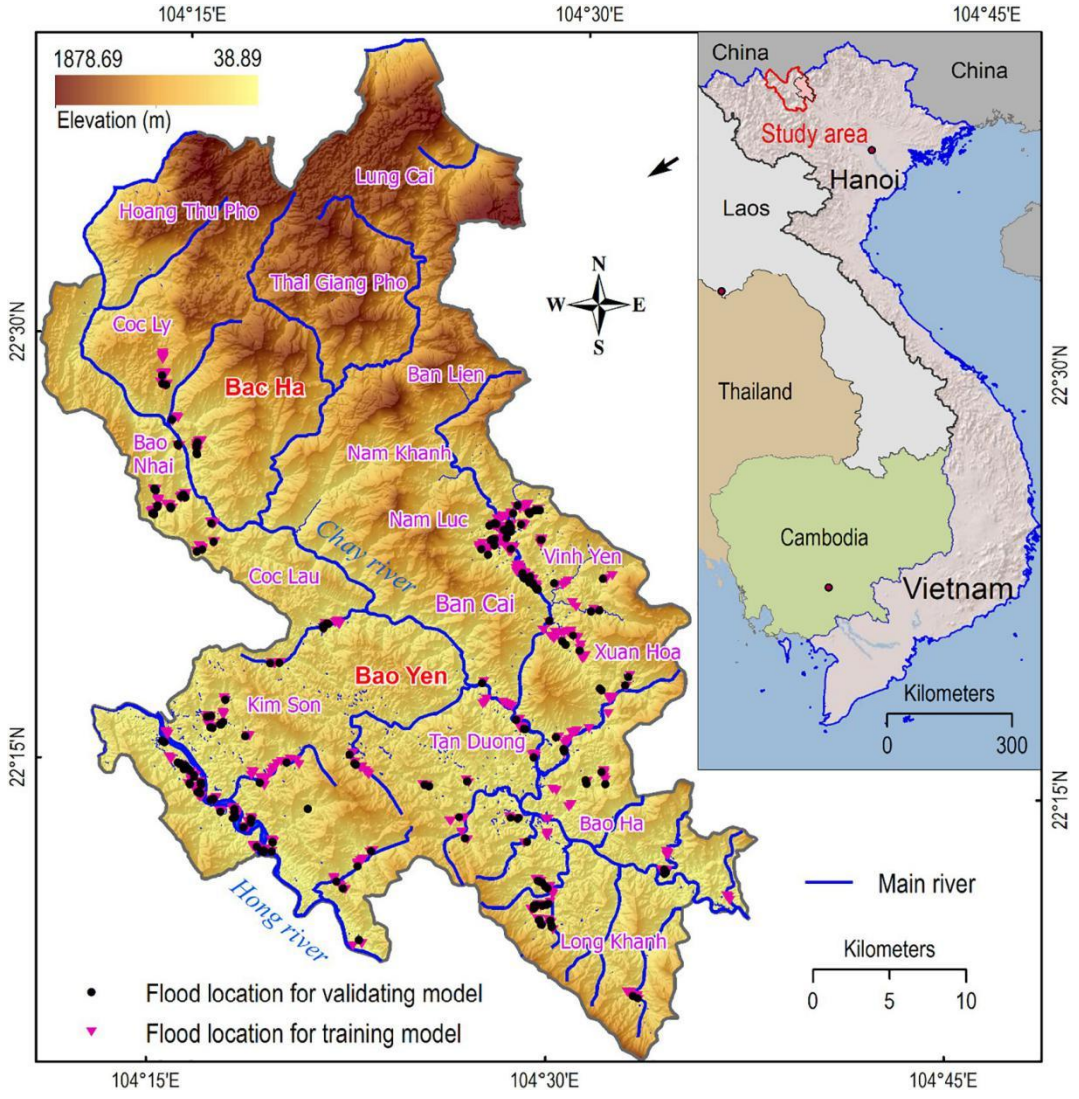

Fig. 1. Location of the study area and flood locations.

Stream density, which is computed by dividing the length of the river (km) over the basin area (km^2), is an important factor influencing flash flood occurrences. The regions with higher stream density are more likely to have a rapid response to rainstorm (Brody et al., 2007) and are more prone to flash floods. The stream density map with seven classes was considered for the current work. Furthermore, NDVI is an important indicator that reflects the vegetation coverage and flash floods are very likely to occur in the areas with low vegetation (Tehrany et al., 2013). In this analysis, the NDVI map (Fig. 2j) with eight classes was derived from the Landsat-8 Operational Land Imagery (30 m resolution and available from http://earthexplorer.usgs.gov) using the following equation (Reed et al., 1994):

$$NDVI = (NIR - RED)/(NIR + RED) \quad (9)$$

where NIR and RED are the surface reflectance of the near-infrared band and the red band, respectively.

Soil characteristics have long been recognized as an important factor that influences rainfall-runoff mechanisms (Geris et al., 2015), whereas lithology structure greatly affects the architecture of the drainage pattern (Pizzuto, 1995) that relates to the development of floodplain. In this research, the soil types map with 13 categories (Fig. 2k) was extracted from the national pedology map at a scale of 100,000, which was provided by the MONRE. The lithology map with 14 classes was generated from the national geological and mineral resources map of Vietnam at a scale of 200:000.

Because a flash flood is often associated with high-intensity and short rainstorms (Borga et al., 2011); thus, rainfall is a key influencing factor for flash flood modeling (Hapuarachchi et al., 2011). In our study area, high-intensity rainstorms occurred on 10, 11, and 12th October generated widespread severe flash flooding. In addition, rainfall lasted during the previous 9 days and the rainfall was ended after Oc- tober 12, 2017. Therefore, a total measured rainfall from 1 to 12th October 2017 at 16 rainfall stations in and around the study areas were used to generate the rainfall map (Fig. 2l).

4. Proposed hybrid artificial intelligence approach based on the PSO-ELM for predicting flash flood susceptibility

The overall workflow of the hybrid PSO-ELM model for predicting flash flood susceptibility is shown in Fig. 3 and consists of six main steps: i) construction of the geospatial database, ii) multicollinearity analysis and factor selection, iii) construction of the training and validation datasets; iv) implementation of the ELM model; v) optimization of the ELM model using PSO, and vi) validation of the model. The interested reader is referred to Huang (2013) for further information on ELM and its MATLAB scripts. In this study, the proposed PSO-ELM was programmed by the authors in MATLAB R2018.

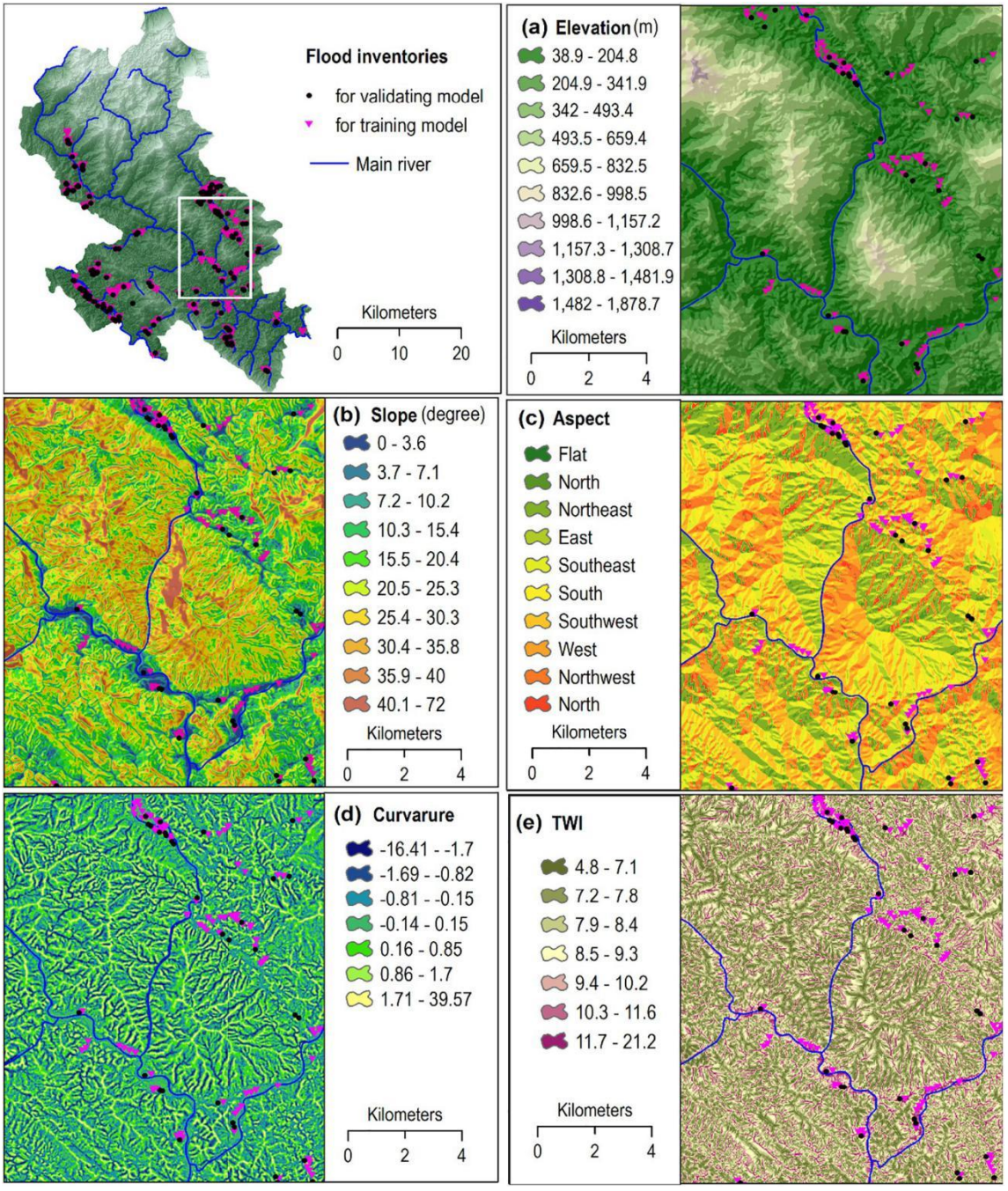

Fig. 2. Flash Flood influencing factors: **(a)** Elevation; **(b)** Slope; **(c)** Aspect; **(d)** Curvature; **(e)** Topographic wetness index (TWI); **(f)** Stream power index; **(g)** Toposhade; **(h) NDVI**; **(i)** Stream density; **(j)** Soil type; **(k)** Lithology; and **(l)** Rainfall.

4.1. Construction of the geospatial database

The geospatial database was constructed using the ESRI® file geo-database format (Zeiller and Murphy, 2010) available at the ArcCatalog 10.4. The database consists of the flash flood inventory map (654 flash flood locations) and 12 influencing factors affected flash flood of the study area. All factors were converted to a raster format with a grid size of 20 × 20 m and in the national geodetic system of VN2000 (UTM map projection, Zone 48 N, central meridian of 105°, and scale factor of 0.9996). Consequently, the size of the study area was formulated in 2367 rows and 3292 columns.

4.2. Multicollinearity analysis and factor selection

In this study, the possible multicollinearity among the flash-flood influencing factors was tested using the Variance Inflation Factors (VIF) and tolerances (Dormann et al., 2013). Previous studies reported by Bui et al. (2011) and Khosravi et al. (2018) suggesting that the VIF > 10 or tolerance < 0.1 shows a multicollinearity problem among the influencing factors.

The results from Table 1 indicate no collinearity exist among flash flood-influencing factors in the study area. Thus, all influencing factors were selected for flash-flood prediction modeling.

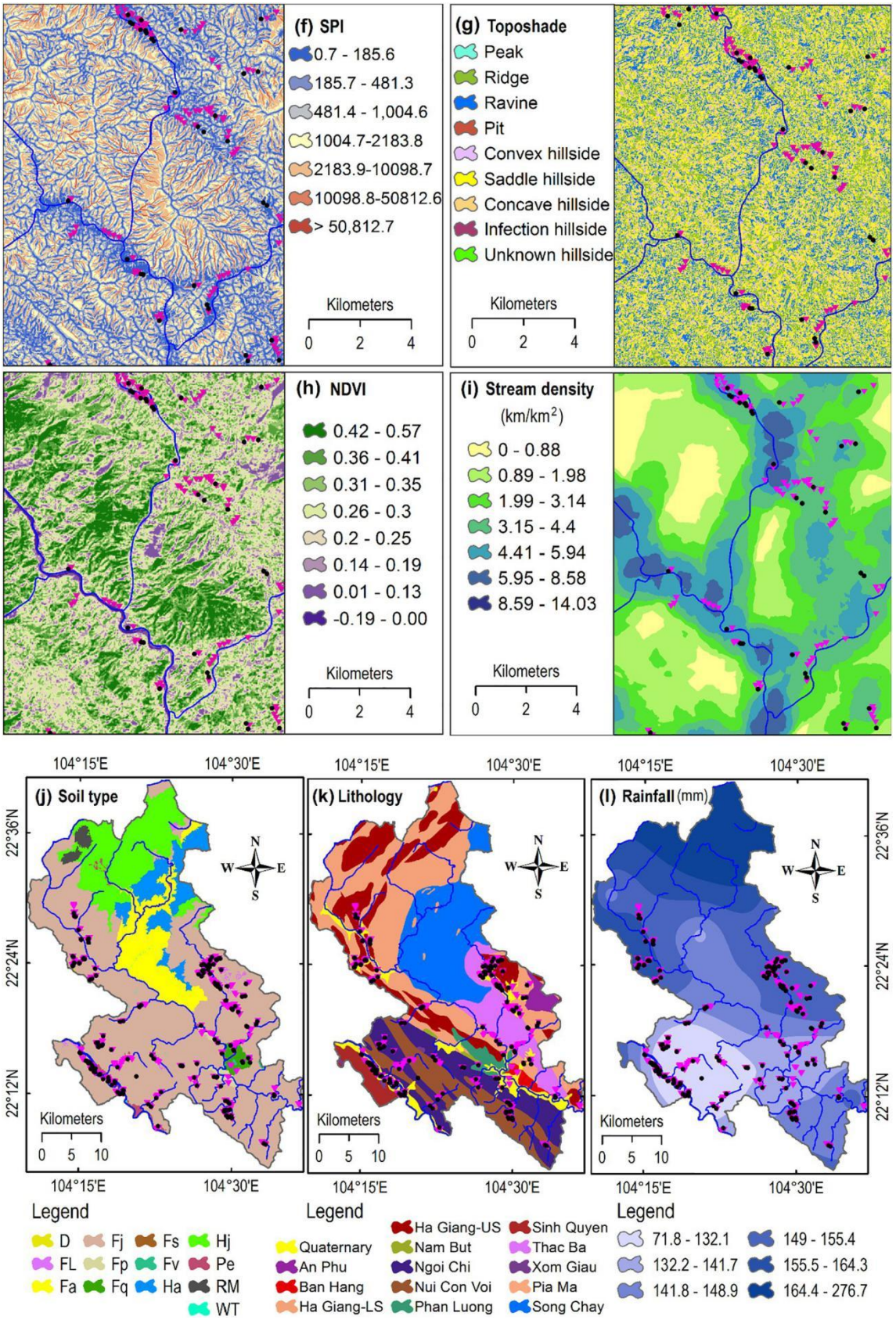

Fig. 2. (*continued*)

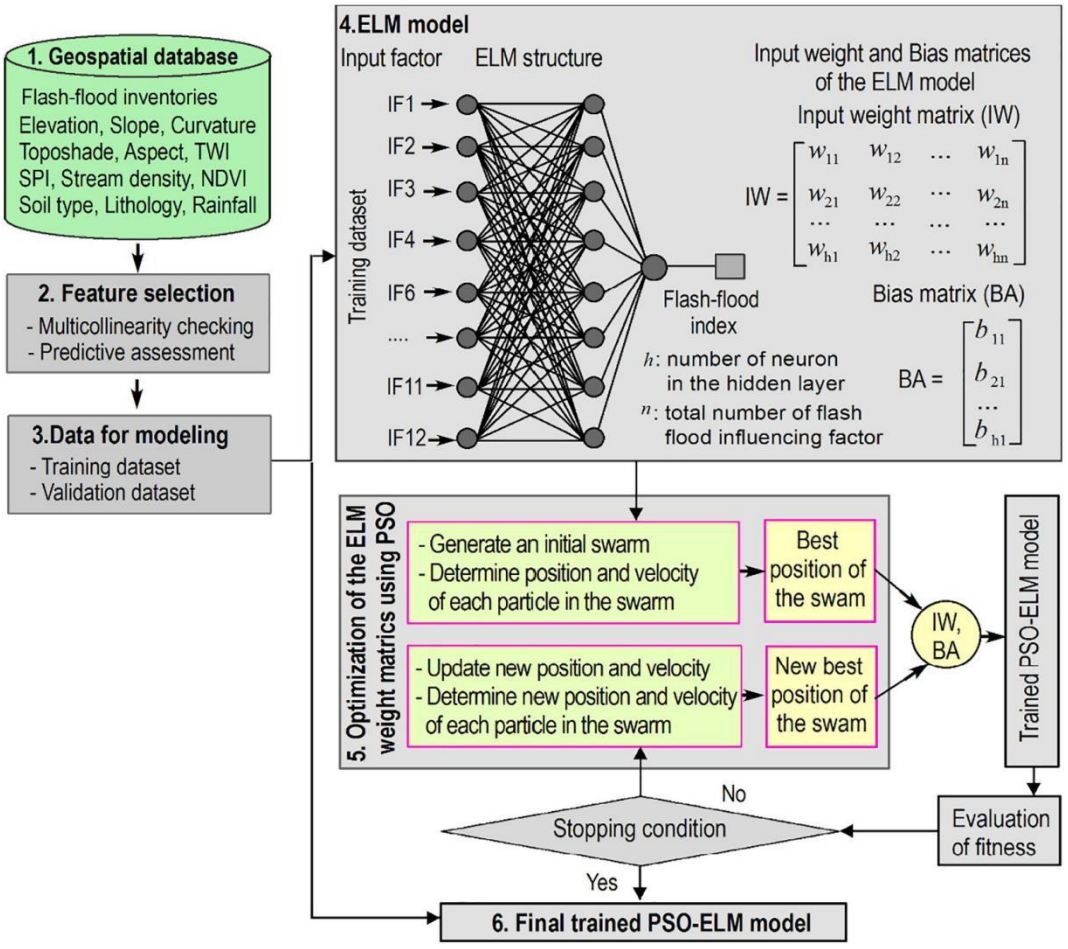

Fig. 3. Workflow of the proposed PSO-ELM model.

Table 1
Multicollinearity analysis for flash-flood influencing factors.

No	Flash flood influencing factors	Collinearity statistics	
		Tolerance	VIF
1	Elevation	0.429	2.334
2	Slope	0.146	6.842
3	Curvature	0.684	1.462
4	Toposhade	0.579	1.727
5	Aspect	0.840	1.190
6	Topographic wetness index (TWI)	0.169	5.902
7	Stream power index (SPI)	0.377	2.652
8	Stream density	0.545	1.836
9	NDVI	0.636	1.572
10	Soil type	0.795	1.258
11	Lithology	0.804	1.243
12	Rainfall	0.591	1.691

4.3. Construction of the training and validation datasets

The training and the validation datasets were computed using the influencing factors combined along with the flash flood inventory map. Seventy percent of the flood events were randomly selected and used for the training dataset, whereas the remaining (30%) was used for the testing set as suggested in (Razavi Termeh et al., 2018).

Because the flash-flood modeling here is considered as a machine learning based binary pattern recognition, the same number of non-flash flood was sampled to hinder the prediction model from potential bias (Ayalew and Yamagishi, 2005; Martinović et al., 2016). Then, attribute values of 12 flash flood influencing factors corresponding to each sample position were extracted. Values of "1" and "0" was assigned to the flood and non-flash samples, respectively. This resulted in training validation datasets with 916 and 392 samples, respectively.

4.4. Configuration of the ELM model

Given the training dataset with 12 flash-flood influencing factors as input neurons, the ELM model was configured by initiating values for the input weight matrix (IW) and bias (BA). The output weight matrix (OW) is computed analytically from the IW and BA by the original ELM algorithm. Since the performance of the ELM model is influenced by the amount of the hidden neuron used (L), a test was carried out between L versus RMSE (Eq. (12)) to determine the best L, and herein, 10 hidden neurons revealed the best result for the data at hand. The final structure of the ELM model was determined as 12 input neurons (flash flood influencing factors), 10 hidden neurons, and 2 output neurons (flash flood and non-flash flood).

4.5. Optimization of the ELM model using PSO

The aim of the optimization process using the PSO algorithm was to find the best values for $IW_{12\times10}$ and $BA_{10\times1}$ that minimize the differ- ence between the flood output value and the flood target value. This difference was measured by a proposed cost function (CF) as below:

$$CF = \sum_{i=1}^{n} \frac{(\bar{y}_i - t_i)^2}{n} \quad (10)$$

where y_i and t_i are the predicted value and the target value of the i^{th} sample; n is the total number of samples.

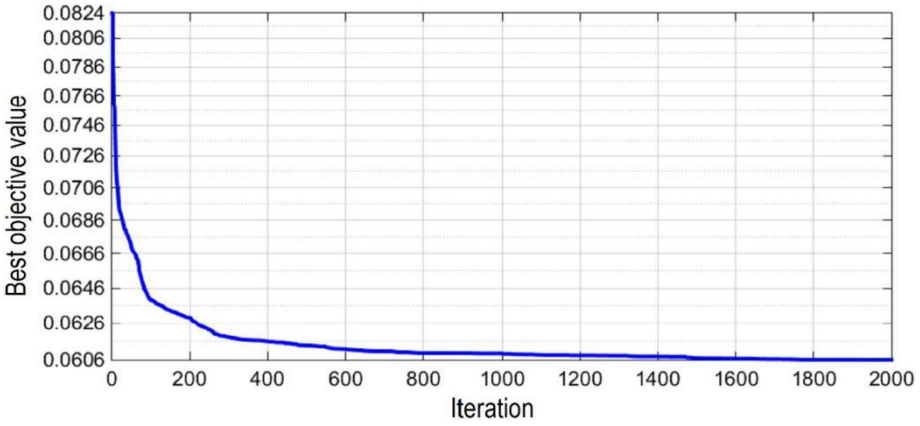

Fig. 4. Best objective value versus running iteration in the training phase of the PSO-ELM model.

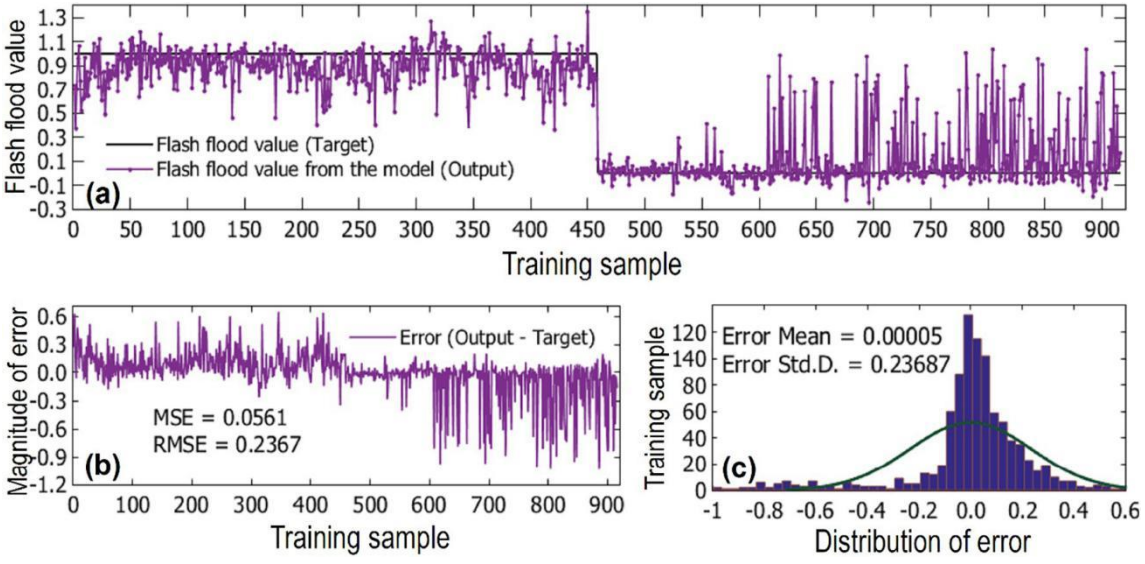

Fig. 5. Performance of the PSO-ELM model in the training dataset: **(a)** Target and computed values; **(b)** Magnitude of the error; **(c)** Distribution of the error.

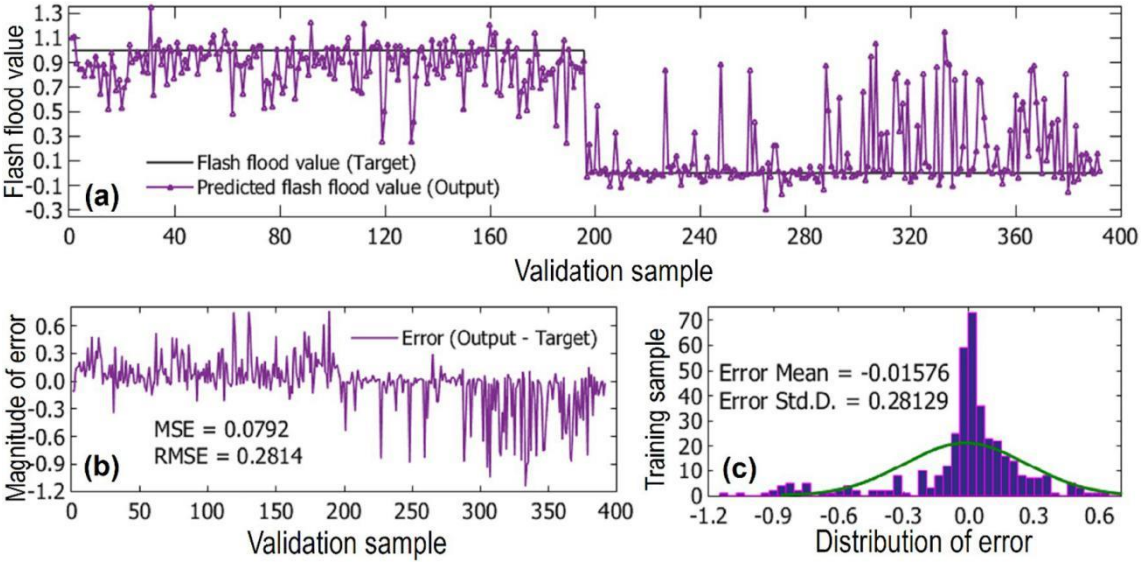

Fig. 6. Performance of the PSO-ELM model in the validation dataset: **(a)** Target and computed values; **(b)** Magnitude of the error; **(c)** Distribution of the error.

Table 2
Performance of the proposed PSO-EML model.

Statistical measure dataset	Training dataset	Validation
True positive	407	165
True negative	448	188
False positive	51	31
False negative	10	8
Sensitivity (%)	97.60	95.38
Specificity (%)	89.78	85.84
Accuracy (%)	93.34	90.05
RMSE	0.237	0.281
MAE	0.056	0.079
Kappa statistic	0.867	0.801
R^2	0.881	0.829

The best values are searched using PSO, resulting in optimizing the base ELM model. Thus, values of the $IW_{12\times10}$ and $BA_{10\times1}$ were trans- ferred to the coordinates of particles in the swarm (130 dimensions), in which, each particle with its coordinates was a solution for the two weight matrices, IW and BA, of the flash flood model. The basic para- meters of the PSO were adapted following the recommendation given in the literature (Tien Bui et al., 2018a), as follows: population size = 35, searching space = −1.0 to 1.0, maximum number of iteration =2000, inertia weight = 0.98, damping ratio = 0.75, $c_1 = 1$, and $c_2 = 2$.

4.6. Performance assessment and final trained PSO-ELM model

Performance of the obtained models was assessed using the root-mean-square error (*RMSE*), the mean absolute error (*MAE*), and the coefficient of determination (R^2) (Mohammadzadeh et al., 2014).

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - t_i)^2}{n}} \quad (11)$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - t_i| \quad (12)$$

$$R2 = 1 - \frac{\sum_{i=1}^{n}(y_i - t_i)^2}{\sum_{i=1}^{n}(t_i - \bar{t})^2} \quad (13)$$

where y_i and $\bar{y}$ are the predicted value of the i^{th} sample and the predicted mean value of the samples from the obtained models, respectively; t_i and $\bar{t}$ are the target value of the i^{th} sample and the target mean values, respectively; n is the total number of samples.

The Receiver Operating Characteristic (*ROC*) curve was also used to assess the performance of the predictive ability of the model. The *ROC* curve was generated by plotting the true positive (*TP*) rate against the false positive (*FP*) rate. Additionally, the area under the *ROC* curve (*AUC*) is the standard statistical measure was generated to validate and compare the selected machine learning algorithms used in this study (Khosravi et al., 2018). A higher AUC value depicts better goodness-of-fit of the model, and the prediction model with AUC values ranges from 0.8 to 0.9 indicates very good performance (Tien Bui et al., 2016).

Once the best performance PSO-ELM model was achieved, the model was used to compute the flash flood index for each pixel of the study area.

Table 3
Performance of the benchmark models in the validation dataset.

Performance	PSO-ELM	MLP-ANN	SVM	C4.5 Decision tree
Accuracy (%)	90.05	88.01	87.24	89.28
Kappa statistic	0.801	0.760	0.730	0.786
RMSE	0.281	0.314	0.307	0.305
MAE	0.079	0.129	0.181	0.164
AUC	0.954	0.938	0.930	0.912

Table 4
Pairwise comparison of the PSO-ELM model and the benchmarks using the t-test at the significant level of $\alpha = 5\%$.

No	Pairwise comparison	*t*	*p*-Value	Sig.
1	PSO-ELM vs. MLP-ANN	2.338	0.0199	Yes
2	PSO-ELM vs. SVM	−1.998	0.0464	Yes
3	PSO-ELM vs. C4.5 Decision tree	−2.331	0.0203	Yes

5. Results and discussion

5.1. Experimental results and model comparisons

Figs. 4, 5, and 6 show the performance of the hybrid PSO-ELM model in the training and the testing phases. It can be clearly seen from the graphs that the proposed model performs well in both the training and the validation datasets. The overall accuracies for the training and validation datasets were 93.34 and 90.05%, respectively. In addition, the PSO-ELM model achieved very high R^2 and low RMSE and MAE values in both training and testing phases.

The results of the training and validation phases of the PSO-ELM model are summarized in Table 2 that shows $R^2 = 0.881$, RMSE = 0.237, MAE = 0.056, and Kappa statistic = 0.867 for the training phase and $R^2 = 0.829$, RMSE = 0.281, MAE = 0.079, and Kappa statistic = 0.801 for the validation phase, indicating that a highly satisfactory flash-flood predictions with relatively small errors was achieved.

The global performance of the PSO-ELM model was measured using the AUC-ROC method (Fig. 7) and showed that the proposed model yielded very high values of AUC (> 0.95) for both the training and

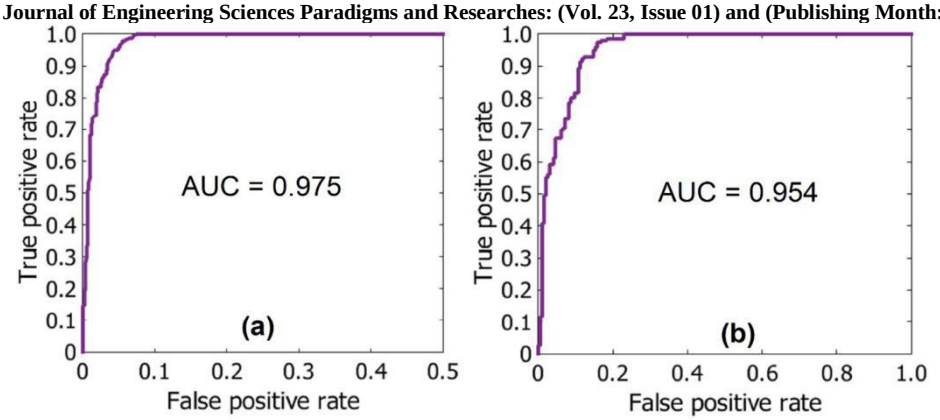

Fig. 7. AUC of the PSO-ELM model: **(a)** Training dataset; **(b)** Validation dataset.

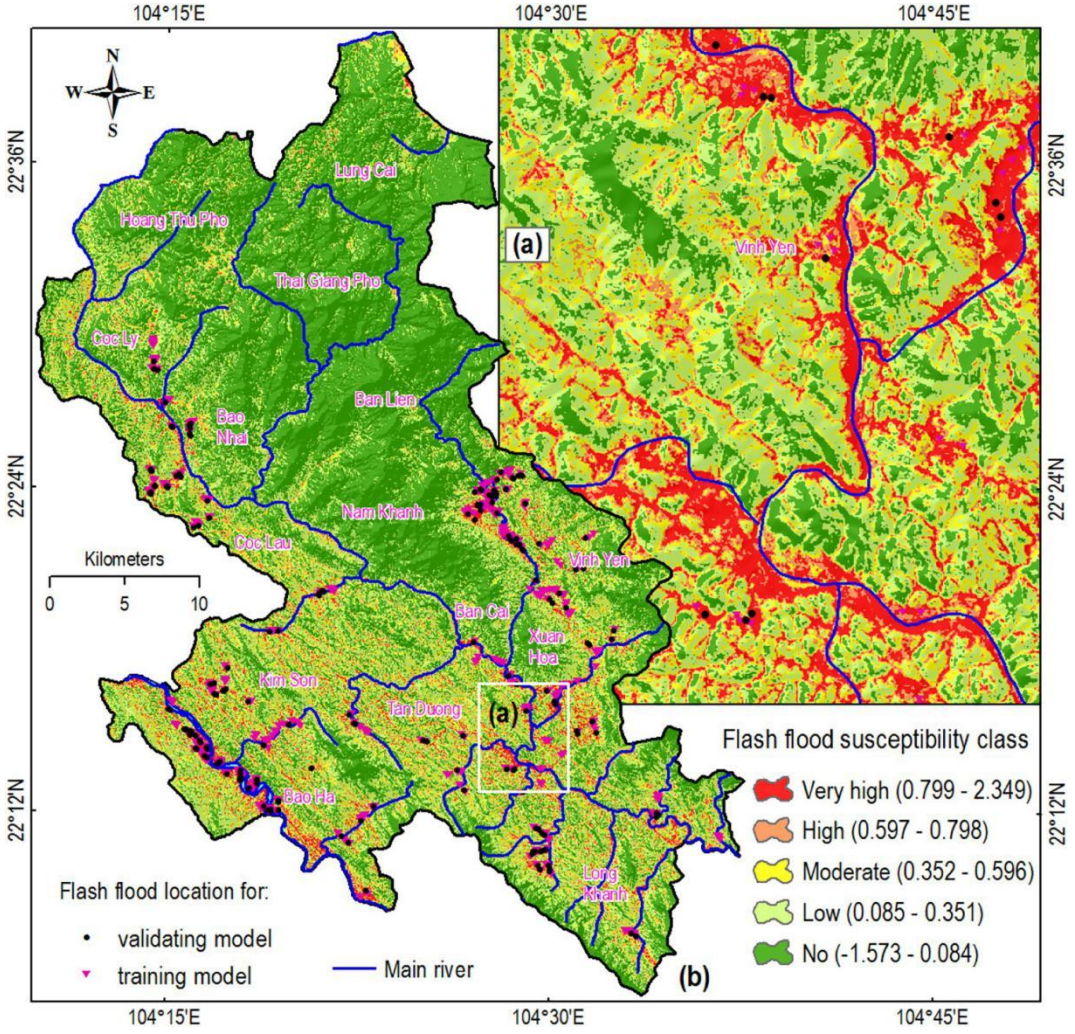

Fig. 8. Flash-flood susceptibility map for the study area using the proposed PSO-ELM model.

Table 5
Characteristics of flash-flood susceptibility classes for the study area derived from the PSO-ELM model.

No.	Flash-flood index	Flood-flood susceptibility map (%)	Verbal expressio	Flash-flood location	Areas (km^2)	Flash-flood
1	0.799–2.349	90–100	Very high	73.81	75.52	97.74
2	0.597–0.798	80–90	High	19.55	75.52	25.89
3	0.352–0.596	70–80	Moderate	5.59	151.04	3.70
4	0.085–0.351	55–70	Low	0.95	226.56	0.42
5	−1.573–0.084	0–55	No	0.01	981.76	0

validation datasets. Notably, there is a balance between the model performance in the training and validation phases, indicating that the PSO algorithm successfully optimized the ELM model. Overall, the results revealed that the PSO-ELM model achieved a high level of predictive accuracy and reliability in terms of AUC values.

The effectiveness of the proposed PSO-ELM model was compared to SVM, MLP-ANN, and C4.5 Decision Tree benchmark methods. These models were performed within the MATLAB software (Matwork, 2017). Regarding the implementation of the SVM, the model was constructed using with $C = 14$ and $\gamma = 5.126$ and using the Radial Basis Function (RBF) kernel function. For the MLP-ANN, an appropriate number of neurons in the hidden layer was found to be 10 through a test with detailed parameters suggested by Hoa et al. (2019). Regarding the C4.5 Decision tree, the hyper-parameters of the model were selected based on a suitable value of the minimal number of observations per tree leaf and was set at 2 with a confidence factor of 0.15 for the C4.5 Decision tree model.

Table 3 shows the performance of the hybrid PSO-ELM model and the benchmark machine learning models. As can be clearly seen that the PSO-ELM yielded the highest predictive accuracy in terms of the overall accuracy, the Kappa coefficient, the RMSE, the MAE, and the AUC for the validation dataset. To check if the performance of the PSO-ELM model and the benchmark models has statistical significance, the paired *t*-test (Altman, 1990) was performed at the significant level of $\alpha = 5\%$. The null hypothesis is that the performance difference of two flash-flood models is equal to 0, and then, *p*-value (two-tailed probability value) and the test statistic *t* value are computed. If *p*-value < 0.05 and the test statistic *t* exceeds the critical values (± 1.96), the null hypothesis is rejected and that indicate a significant difference between two models. Table 4 shows that the *p*-values are < 0.05 and the test

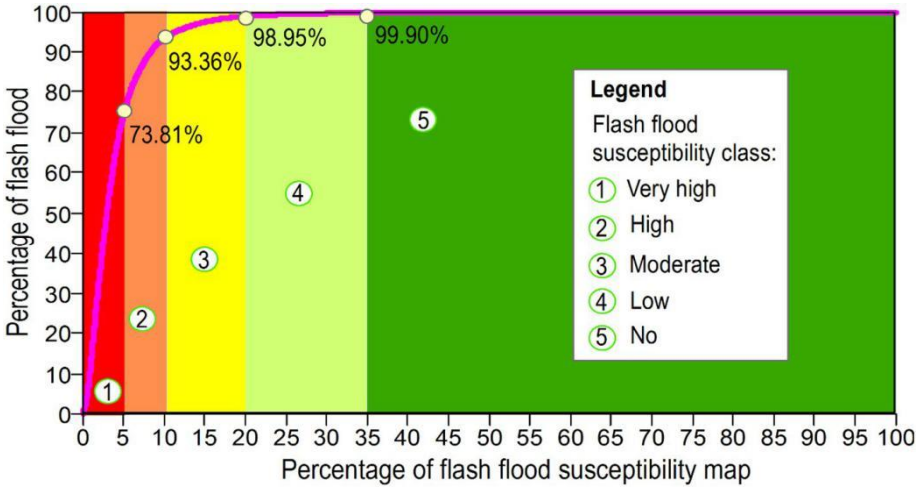

Fig. 9. The graphic curve for separating the flash flood susceptibility classes in the study area.

statistic *t* values are less than the critical values, demonstrating that the null hypothesis is rejected and the performance of the PSO-ELM model significantly differ from the benchmarks. Overall, these results demonstrated that the PSO-ELM model is the best suit for flash flood susceptibility in the study area, and support conclusions drawn about the application of such hybrid models for other types of natural hazards (Jaafari et al., a,b).

5.2. Construction of the flash flood susceptibility map

Since the PSO-ELM model performed well for both the training and validation datasets and outperformed the benchmark models, this model was further used to compute the flash flood susceptibility indices for all the pixels within the study area. To visualize the final flood susceptibility maps, these indices were transformed to ArcGIS 10.4 software package and converted to a raster format. The map was then reclassified into five classes that depict to very high (10%), high (10%), moderate (10%), low (15%), and no flood (55%) susceptibility across the study area (Fig. 8). The threshold values used for classifying these classes (Fig. 9) were determined by overlaying the historical flood locations and the flash-flood susceptibility indices map (Tien Bui and Hoang ; Tien Bui et al.,). The results showed that when 10% of the total area was classified into the very high and the high classes, 73.81 and 19.55% of the total historical flash flooding locations covered by these two classes, respectively. The moderate class occupied 10% of the region and accounted for only 5.59% of the flash-flood locations (Table 4).

The low susceptibility class covered 15% of the land area and consisted of only 0.95% of the flash-flood locations. Further analysis of flash-flood density revealed that there was a considerable increase in the flash-flood density from the moderate class to the very high class (Table 5). These indicate that the proposed PSO-ELM hybrid machine learning model successfully delineated the study area into different susceptible areas (Fig. 8).

6. Concluding remarks

This study proposed a new hybrid PSO-ELM model for predicting flash flood susceptibility in a high frequency tropical-cyclone area of Vietnam. We employed the ELM algorithm to generate the initial prediction model, while the PSO was used to search and optimize the hyper-parameters of the model. A geospatial database consisted of 12 conditioning factors was used to train and verify the newly proposed PSO-ELM model. The experimental results show that the proposed PSO-ELM machine learning model has high performance in terms of R^2, RMSE, MAE, and AUC performance metrics.

The main advantage of the hybrid PSO-ELM model is the PSO algorithm that helps to automatically search and find the input weights for the ELM, where the coordinates of each particle are values of the weights, and thus, each particle is a hyper-parameter for the ELM model. With a population of 35 particles and 2000 running iterations, a total of 70,000 solutions were searched and tried to find the optimized values for the input weights. Consequently, the high performance of the model indicates that the selection of influencing factors, the coding, and training process were carried out successfully.

The main disadvantage of the PSO-ELM model is the determination of hidden neurons that requires a time-consuming trial and error procedure. Further, the determination of the upper and lower bounds of the searching space that influence the convergence of the PSO-ELM model is a difficult task. Although the searching space of -1.0 to 1.0 used in this research resulted in high performance of the PSO-ELM model, we can certainly not conclude that this is the best searching space for the study area. Overall, by integrating PSO with ELM, the training process takes much time compared to that of the original EML algorithm. Thus, for large-scale landscapes, the training process of the PSO-ELM may be time-consuming and restricts the management plans that need a quick estimation of flood susceptibilities.

Despite these limitations, the PSO-ELM model can help the managers/authorities to achieve a high level of predictive accuracy of future floods that indeed greatly facilitate the development of management plans for flood-prone landscapes in the face of future climate change.

Acknowledgment

This research was supported by National Foundation of Science and Technology Development funded by the project 105.99-2014.25. The data analysis and write-up were carried out as a part of the second author's Ph.D. studies at Faculty of Geomatics and Land Administration, Hanoi University of Mining and Geology, Vietnam.

Conflicts of interests

The authors declare no conflict of interest.

References

Abdullah, S.S., Malek, M., Abdullah, N.S., Kisi, O., Yap, K.S., 2015. Extreme learning machines: a new approach for prediction of reference evapotranspiration. J. Hydrol. 527, 184–195.
Altman, D.G., 1990. Practical Statistics for Medical Research. CRC press.
Aryal, S.K., Mein, R.G., O'Loughlin, E.M., 2003. The concept of effective length in hill- slopes: assessing the influence of climate and topography on the contributing areas of

catchments. Hydrol. Process. 17, 131–151.
Ayalew, L., Yamagishi, H., 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology 65, 15–31.
Beven, K., Kirkby, M., Schofield, N., Tagg, A., 1984. Testing a physically-based flood forecasting model (TOPMODEL) for three UK catchments. J. Hydrol. 69, 119–143.
Borga, M., Anagnostou, E.N., Blöschl, G., Creutin, J.D., 2011. Flash flood forecasting, warning and risk management: the HYDRATE project. Environ. Sci. Pol. 14, 834–844.
Brody, S.D., Zahran, S., Maghelal, P., Grover, H., Highfield, W.E., 2007. The rising costs of floods: examining the impact of planning and development decisions on property damage in Florida. J. Am. Plan. Assoc. 73, 330–345.
Brunner, G.W., 1995. HEC-RAS River Analysis System. Hydraulic Reference Manual. Version 1.0. Hydrologic Engineering Center, Davis CA.
Bubeck, P., Thieken, A.H., . What helps people recover from floods? Insights from a survey among flood-affected residents in Germany. Reg. Environ. Chang. 18, 287–296.
Bui, D.T., Lofman, O., Revhaug, I., Dick, O., 2011. Landslide susceptibility analysis in the Hoa Binh province of Vietnam using statistical index and logistic regression. Nat. Hazards 59, 1413.
Bui, D.T., et al., 2016. Hybrid artificial intelligence approach based on neural fuzzy in- ference model and metaheuristic optimization for flood susceptibility modelling in a high-frequency tropical cyclone area using GIS. J. Hydrol. 540, 317–330.
Bui, D.T., Tsangaratos, P., Ngo, P.-T.T., Pham, T.D., Pham, B.T., 2019. Flash flood sus- ceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods. Sci. Total Environ. 668, 1038–1054.
Chapi, K., et al., . A novel hybrid artificial intelligence approach for flood suscept- ibility assessment. Environ. Model Softw. 95, 229–245.
Cheng, R., Jin, Y., 2015. A social learning particle swarm optimization algorithm for scalable optimization. Inf. Sci. 291, 43–60.
Choubin, B., et al., . An ensemble prediction of flood susceptibility using multi- variate discriminant analysis, classification and regression trees, and support vector machines. Sci. Total Environ. 651, 2087–2096.
Darabi, H., et al., 2019. Urban flood risk mapping using the GARP and QUEST models: a comparative study of machine learning techniques. J. Hydrol. 569, 142–154.
de Musso, N.M., et al., 2018. Spatial evolution of the December 2013 Metaponto plain (Basilicata, Italy) flood event using multi-source and high-resolution remotely sensed data. J. Maps 14, 219–229.
Deo, R.C., Şahin, M., 2015. Application of the extreme learning machine algorithm for the prediction of monthly effective drought index in eastern Australia. Atmos. Res. 153, 512–525.
Destro, E., et al., 2018. Coupled prediction of flash flood response and debris flow oc- currence: application on an alpine extreme flood event. J. Hydrol. 558, 225–237. Dormann, C.F., et al., 2013. Collinearity: a review of methods to deal with it and a si- mulation study evaluating their performance. Ecography 36, 27–46.
Edouard, S., Vincendon, B., Ducrocq, V., . Ensemble-based flash-flood modelling: taking into account hydrodynamic parameters and initial soil moisture uncertainties. J. Hydrol. 560, 480–494.
Ellabban, O., Abu-Rub, H., Blaabjerg, F., 2014. Renewable energy resources: current status, future prospects and their enabling technology. Renew. Sust. Energ. Rev. 39, 748–764.
Eshtay, M., Faris, H., Obeid, N.,. Improving extreme learning machine by compe- titive swarm optimization and its application for medical diagnosis problems. Expert Syst. Appl. 104, 134–152.
General Statistics Office, 2017. The Laocai Statistical Year Book 2016. vol. 2017 Statistical Publishing House.
Geris, J., Tetzlaff, D., McDonnell, J., Soulsby, C., 2015. The relative role of soil type and tree cover on water storage and transmission in northern headwater catchments. Hydrol. Process. 29, 1844–1860.
Ghimire, S., Deo, R.C., Downs, N.J., Raj, N., . Self-adaptive differential evolutionary extreme learning machines for long-term solar radiation prediction with remotely- sensed MODIS satellite and reanalysis atmospheric products in solar-rich cities. Remote Sens. Environ. 212, 176–198.
Hapuarachchi, H., Wang, Q., Pagano, T., 2011. A review of advances in flash flood forecasting. Hydrol. Process. 25, 2771–2784.
Hartnett, M., Nash, S., . High-resolution flood modeling of urban areas using MSN_Flood. Water Sci. Eng. 10, 175–183.
Hoa, P.V., et al., . Soil salinity mapping using SAR Sentinel-1 data and advanced machine learning algorithms: A case study at Ben Tre Province of the Mekong River Delta (Vietnam). Remote Sens. 11, 128.
Hu, P., Zhang, Q., Shi, P., Chen, B., Fang, J., . Flood-induced mortality across the globe: spatiotemporal pattern and influencing factors. Sci. Total Environ. 643, 171–182.
Huang, G.-B., 2013. Extreme machine learning. Nanyang Technological University, Singaporehttp://www.ntu.edu.sg/home/egbhuang/index.html.
Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2004. Extreme learning machine: A new learning scheme of feedforward neural networks, neural networks, 2004. In: Proceedings. 2004 IEEE International Joint Conference on. IEEE, pp. 985–990.
Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2006. Extreme learning machine: theory and ap- plications. Neurocomputing 70, 489–501.
Huang, G.-B., Wang, D.H., Lan, Y., 2011. Extreme learning machines: a survey. Int. J. Mach. Learn. Cybern. 2, 107–122.
Huang, G.B., Zhou, H., Ding, X., Zhang, R., 2012. Extreme learning machine for regression and multiclass classification. IEEE Trans. Syst. Man Cybern. B Cybern. 42, 513–529. Jaafari, A., Panahi, M., Pham, B.T., Shahabi, H., Bui, D.T., Rezaie, F., Lee, S., 2019a. Meta optimization of an adaptive neuro-fuzzy inference system with grey wolf optimizer and biogeography-based optimization algorithms for spatial prediction of landslide

susceptibility. Catena 175, 430–445.
Jaafari, A., Zenner, E.K., Panahi, M., Shahabi, H., b. Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability. Agricultural and Forest Meteorology 266, 198–207.
Jalayer, F., Aronica, G.T., Recupero, A., Carozza, S., Manfredi, G., . Debris flow damage incurred to buildings: an in situ back analysis. J. Flood Risk Manage. 11, S646–S662.
Kennedy, J., Eberhart, R., 1995. In: Particle Swarm Optimization, Neural Networks, 1995. Proceedings., IEEE International Conference on. vol. 4. pp. 1942–1948.
Khosravi, K., et al., . A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. Sci. Total Environ. 627, 744–755.
Lee, S., Kim, J.-C., Jung, H.-S., Lee, M.J., Lee, S., . Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. Geomat. Nat. Haz. Risk 8, 1185–1203.
Lim, J., Lee, K.-s., 2018. Flood mapping using multi-source remotely sensed data and logistic regression in the heterogeneous mountainous regions in North Korea. Remote Sens. 10, 1036.
Manfreda, S., et al., 2014. Investigation on the use of geomorphic approaches for the delineation of flood prone areas. J. Hydrol. 517, 863–876.
Martınez-Casasnovas, J., Ramos, M., Poesen, J., 2004. Assessment of sidewall erosion in large gullies using multi-temporal DEMs and logistic regression analysis. Geomorphology 58, 305–321.
Martinović, K., Gavin, K., Reale, C., . Development of a landslide susceptibility as- sessment for a rail network. Eng. Geol. 215, 1–9.
Matwork, . Statistics and Machine Learning Toolbox User's Guide. Matwork Inc.. https://www.mathworks.com/help/pdf_doc/stats/stats.pdf,
Mohammadzadeh, D., Bazaz, J.B., Alavi, A.H., 2014. An evolutionary computational approach for formulation of compression index of fine-grained soils. Eng. Appl. Artif. Intell. 33, 58–68.
Moore, I.D., Grayson, R., Ladson, A., 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrol. Process. 5, 3–30. Mosavi, A., Ozturk, P., Chau, K.-w., . Flood prediction using machine learning models: literature review. Water 10, 1536.
Nandi, A., Mandal, A., Wilson, M., Smith, D., 2016. Flood hazard mapping in Jamaica using principal component analysis and logistic regression. Environ. Earth Sci. 75, 465.
Ngo, P.-T., et al., 2015a. A novel hybrid swarm optimized multilayer neural network for spatial prediction of flash floods in tropical areas using Sentinel-1 SAR imagery and geospatial data. Sensors 18, 3704.
Ngo, P.-T.T., et al., 2015b. Research to Develop Ensemble Models for Flash Flood Prediction at Lao Cai Area Using Advanced Geomatics Tecnologies (B2015-MDA-18DT). Hanoi University of Mining and Geology, Hanoi, Vietnam.
Niu, W.-j., Feng, Z.-k., Cheng, C.-t., Zhou, J.-z., 2015. Forecasting daily runoff by extreme learning machine based on quantum-behaved particle swarm optimization. J. Hydrol. Eng. 23, 04018002.
P., H.H.A., W.Q, J., P.T, C., 2011. A review of advances in flash flood forecasting. Hydrol. Process. 25, 2771–2784.
Peduzzi, P., 2017. Flooding: prioritizing protection? Nat. Clim. Chang. 7, 625.
Pizzuto, J.E., 1995. Downstream fining in a network of gravel-bedded rivers. Water Resour. Res. 31, 753–759.
Rahmati, O., Pourghasemi, H.R., 2015. Identification of critical flood prone areas in data- scarce and ungauged regions: A comparison of three data mining models. Water Resour. Manag. 31, 1473–1487.
Razavi Termeh, S.V., Kornejady, A., Pourghasemi, H.R., Keesstra, S., 2015. Flood sus- ceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. Sci. Total Environ. 615, 438–451.
Reed, B.C., et al., 1994. Measuring phenological variability from satellite imagery. J. Veg. Sci. 5, 703–714.
Sahoo, G.B., Ray, C., De Carlo, E.H., 2006. Use of neural network to predict flash flood and attendant water qualities of a mountainous stream on Oahu, Hawaii. J. Hydrol. 327, 525–538.
Santo, A., Santangelo, N., Forte, G., De Falco, M., 2015. Post flash flood survey: the 14th and 2015 event in the Paupisi-Solopaca area (Southern Italy). J. Maps 13, 19–25.
Taormina, R., Chau, K.-W., 2015. Data-driven input variable selection for rainfall–runoff modeling using binary-coded particle swarm optimization and extreme learning machines. J. Hydrol. 529, 1617–1632.
Tehrany, M.S., Pradhan, B., Jebur, M.N., 2013. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and mul- tivariate statistical models in GIS. J. Hydrol. 504, 69–79.
Tehrany, M.S., Pradhan, B., Jebur, M.N., 2014. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. J. Hydrol. 512, 332–343.
Tehrany, M.S., 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.
Tehrany, M.S., Jones, S., Shabani, F., Martínez-Álvarez, F., Tien Bui, D., 2018. A novel ensemble modeling approach for the spatial prediction of tropical forest fire sus- ceptibility using LogitBoost machine learning classifier and multi-source geospatial data. Theor. Appl. Climatol. https://doi.org/10.1007/s00704-018-2628-9.
Termeh, S.V.R., Kornejady, A., Pourghasemi, H.R., Keesstra, S., 2015. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and me- taheuristic algorithms. Sci. Total Environ. 615, 438–451.

Tien Bui, D., Hoang, N.D., 2015. A Bayesian framework based on a Gaussian mixture model and radial-basis-function fisher discriminant analysis (BayGmmKda V1.1) for spatial prediction of floods. Geosci. Model Dev. 10, 3391–3409.

Tien Bui, D., et al., 2015. Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization. Landslides 14, 447–458.

Tien Bui, D., Nhu, V.-H., Hoang, N.-D., 2014a. Prediction of soil compression coefficient for urban housing project using novel integration machine learning approach of swarm intelligence and multi-layer perceptron neural network. Adv. Eng. Inform. 38, 593–604.

Tien Bui, D., et al., 2014b. Novel hybrid evolutionary algorithms for spatial prediction of floods. Sci. Rep. 8, 15364.

Tong, X., et al., 2014. An approach for flood monitoring by the combined use of Landsat 8 optical imagery and COSMO-SkyMed radar imagery. ISPRS J. Photogramm. Remote Sens. 136, 144–153.

Wang, Y., et al., 2014. A hybrid GIS multi-criteria decision-making method for flood susceptibility mapping at Shangyou, China. Remote Sens. 11, 62.

Ward, P.J., et al., 2015. A global framework for future costs and benefits of river-flood protection in urban areas. Nat. Clim. Chang. 7, 642.

Wasko, C., Sharma, A., 2017. Global assessment of flood and storm extremes with in- creased temperatures. Sci. Rep. 7, 7945.

Yaseen, Z.M., et al., 2014. Predicting compressive strength of lightweight foamed con- crete using extreme learning machine model. Adv. Eng. Softw. 115, 112–125.

Youssef, A.M., Pradhan, B., Hassan, A.M., 2011. Flash flood risk estimation along the St. Katherine road, southern Sinai, Egypt using GIS based morphometry and satellite imagery. Environ. Earth Sci. 62, 611–623.

Zeiller, M., Murphy, J., 2010. Modeling our World: The ESRI Guide to Geodatabase Concepts. ESRI Press, Redlands, California, USA, pp. 297.

Zhao, G., Pang, B., Xu, Z., Peng, D., Xu, L., 2014. Assessment of urban flood susceptibility using semi-supervised machine learning model. Sci. Total Environ. 659, 940–949.

Zhou, Q., Mikkelsen, P.S., Halsnæs, K., Arnbjerg-Nielsen, K., 2012. Framework for eco- nomic pluvial flood risk assessment considering climate change effects and adapta- tion benefits. J. Hydrol. 414, 539–549.

Zhu, Q.-Y., Qin, A.K., Suganthan, P.N., Huang, G.-B., 2005. Evolutionary extreme learning machine. Pattern Recogn. 38, 1759–1763.