



UDC 528.92

## SIMULATION OF FLOOD-PRONE AREAS USING MACHINE LEARNING AND GIS TECHNIQUES IN SAMANGAN PROVINCE, AFGHANISTAN

Vahid ISAZADE<sup>1</sup>, Abdul Baser QASIMI<sup>2</sup>✉, Abdulla AI KAFY<sup>3,4</sup>, Pinliang DONG<sup>5</sup>,  
Mustafa MOHAMMADI<sup>6</sup>

<sup>1</sup>Department of Geographical Science, Kharazmi University, Tehran, Iran

<sup>2,6</sup>Department of Geography, Education Faculty, Samangan University, Afghanistan

<sup>3</sup>Department of Urban and Regional Planning, Rajshahi University of Engineering and Technology (RUET), Rajshahi, 6204, Bangladesh

<sup>4</sup>ICLEI South Asia, Rajshahi City Corporation, Rajshahi, 6203, Bangladesh

<sup>5</sup>Department of Geography, University of North Texas, North Texas, USA

### Article History:

- received 31 January 2023
- accepted 04 March 2024

**Abstract.** Flood events are the most sophisticated and damaging natural hazard compared to other natural catastrophes. Every year, this hazard causes human-financial losses and damage to croplands in different locations worldwide. This research employs a combination of artificial neural networks and geographic information systems (GIS) to simulate flood-vulnerable locations in the Samangan Province of Afghanistan. First, flood-influencing factors, such as soil, slope layer, elevation, flow direction, and land use/cover, were evaluated as influential factors in simulating flood-prone areas. These factors were imported into GIS software. The Fishnet command was used to partition the information layers. Furthermore, each layer was converted into points, and this data was fed into the perceptron neural network along with the educational data obtained from Google Earth. In the perceptron neural network, the input layers have five neurons and 16 nodes, and the outputs showed that elevation had the lowest possible weight ( $R_2 = 0.713$ ) and flow direction had the highest weight ( $R_2 = 0.913$ ). This study demonstrated that combining GIS and artificial neural networks results in acceptable performance for simulating and modeling flood susceptible areas in different geographical locations and significantly helps prevent or reduce flood hazards.

**Keywords:** flood, perceptron artificial neural network, digital elevation model, Samangan, Afghanistan.

✉Corresponding author. E-mail: [qasimi.abdul.a@gmail.com](mailto:qasimi.abdul.a@gmail.com)

## 1. Introduction

Flooding is a natural disaster when the water level in a particular location exceeds the usual and expected level, as indicated by the usual frequency index. According to investigators and planners, flooding is a devastating disaster that can occur from any source, including rainfall or intentional acts (Fernandes et al., 2018; Desai et al., 2015; Santos & Reis, 2018). Some influencing natural factors like vegetation, elevation, drainage density, soil texture, and distance operate as flood triggers in various places (Hosseini et al., 2020). Floods are recognized to be the most catastrophic natural hazard among the various hydrological hazards (Mishra & Sinha, 2020). Severe loss of human lives and destruction of infrastructures are the main consequences of floods (Hirabayashi et al., 2013; Costache, 2019). According to statistics from the United Nations

Annual Tome for Catastrophe Risk Devaluation (UNISDR), over 150,061 flood incidents occurred worldwide between 1995 and 2018, resulting in the deaths of around 157,000 individuals, accounting for 11.1 percent of global disaster victims (Hong et al., 2018). As climate change on a global scale has significant impacts on precipitation intensity (Chang et al., 2012), the increased frequency of urban flood catastrophes could be exacerbated by rising sea levels and the ineffectiveness of antiquated infrastructures (Hallegatte et al., 2011; Schubert & Sanders, 2012).

Moreover, urban areas are being expanded within the flood-prone regions due to population growth, which amplifies the range of damage influenced by urban flooding (Neumann et al., 2015). Disruption of supplying power and water, traffic jams, flood contamination with sewage waters, destruction of transportation infrastructures, and loss of life are the expected consequences of urban flooding

(Jonkman & Vrijling, 2008). According to the statistics reported by various research, floods impact around 200 million people around the world every year (Tien Bui et al., 2019). Direct effects of floods include deaths and financial loss, the devastation of crops, the spread of infectious illnesses close to the water's route, and the deterioration of environmental ecosystems (Shafapour Tehrany, 2019). Furthermore, it is expected that by 2050, the rates and intensity of floods will have deteriorated due to climate change predictions, land use template changes, and population increase, resulting in enormous losses (US\$ 1 trillion) (Jahandideh-Tehrani et al., 2019; Xu et al., 2019; Huang et al., 2019).

Many studies on flood risk evaluation use hydrodynamic and hydrological models and artificial neural networks. According to a literature review, predictive models and reliable artificial neural network algorithms are frequently required to improve new technologies to generate flood maps of flood-prone areas that facilitate better decisions to mitigate and monitor these events. For example, Shahabad et al. (2020) presented hybrid bagging approaches based on four Kernels of the Nearest Neighbor (KNN) algorithm (weighted, cosine, coarse, and cubic). They found that the Bagging-Cubic ensemble method had the most significant capability in predicting flooded areas in the Hares area.

Dano et al. (2019) investigated the analytic network process of artificial neural networks (ANP) and GIS integrations for flood prediction and visualization. Expert knowledge and questionnaires based on the ANP scientific model were used to assign relative weights to flood-influencing factors. Despite the model's apparent simplicity, its reliance on expert opinion makes it incompatible with quantitative approaches and unsuitable for large-scale modeling (Sachdeva et al., 2017). Tehrany et al. (2015) used Support Vector Machine (SVM) frequency ratio (FR) [PD3]. This methodology was used to evaluate the suitability and effectiveness of the proposed ensembled models on a dataset including topographic wetness index (TWI), rainfall, land use/cover, stream power index (SPI), river, curvature, DEM, slope, soil type, and geology to predict flood susceptible areas. Moreover, Tehrany et al. (2019) used a GIS-based ensembled approach (RBF, sigmoid kernel, SVM with linear, evidential belief function, and polynomial) to map flood-prone areas in the Brisbane watershed, Australia. Aspect, slope, SPI, curvature, TWI, land use/cover, soil type, geology, rainfall, distance from the road, and distance from the river were among the 12 digital elevation model factors they employed. Khosravi et al. (2019) evaluated the performance of three knowledge-based approaches (TOPSIS, VIKOR, and SAW) with two machine learning models, naive byes (NB) and naive Bayes tree (NBT), to map flood-prone areas in the Ningdu watershed, China. They discovered that machine learning methods outperformed the knowledge-based strategy based on three experts in speed and accuracy.

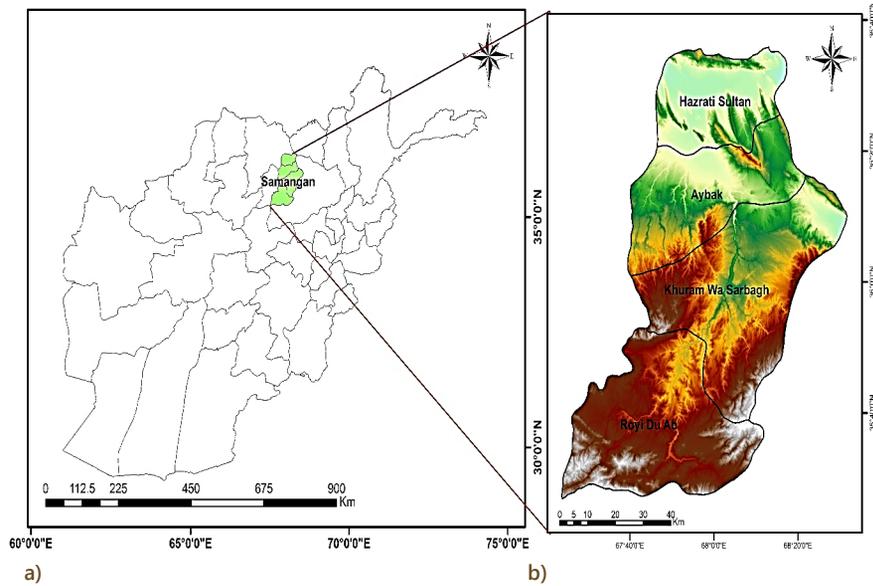
Metin et al. (2020) demonstrated in the Elbe basin, Germany, using a pragmatic spatial reliance flood modeling

approach, that flood consequences are exceeded for return periods less than 1 in 200 years and underestimated by static flood maps for return periods greater than 1 in 50 years. They discovered that the tail of the damage dispersal (>200 – years) was nevertheless binary when they combined the T-year depletion and the T-year loss. Furthermore, Bui et al. (2020) used four swarm intelligence and optimized DLNN algorithms to predict flash flood areas in tropical climates (Grasshopper optimization algorithm (GOA)).

Particle swarm optimization (PSO), social spider optimization (SSO), and grey wolf optimization (GWO) are the instances). In terms of accuracy, their targeted approaches outperformed independent benchmarks such as SVM, PSO, and RF. In Shanghai, China, Seejata et al. (2018) and Wang et al. (2018) collaborated to develop a hybrid approach combining network analytical process, multi-criteria decision analysis, and weighted linear combination (WLC). Few studies have been conducted to predict flood-prone areas in Afghanistan where developed systems to monitor flood occurrence are unavailable, and the accessibility to digital terrain data is limited. In the last two decades, the death toll from floods in Afghanistan has been horrific, with hundreds of villages destroyed in the last 20 years. Samangan Province in northern Afghanistan suffers heavy casualties from yearly devastating seasonal floods. The most vulnerable area runs parallel to the Samangan River and sustains the most damage each year. Maximum entropy, artificial neural networks (ANN), regression, frequency ratio, logistic decision trees (DT), support vector machine (SVM), and MLPNN have been utilized for flood slope mapping (Shafapour Tehrany, 2019). MLPNN and GIS are widely used, and the results are more accurate in flood simulation and workable (Cabrera & Lee, 2020). This study evaluates flood-prone areas using a multilayer perceptron neural network (MLPNN) model and GIS, and high-risk areas are simulated.

## 2. Study area

Samangan Province is located between the longitude of 67°24' and 46" and 68°34' and 8" of the eastern hemisphere and the latitude of 35°26'46" and 36°35'10" in the northern hemisphere. Samangan is one of the northern provinces of Afghanistan, which shares borders with Balkh Province to the north, Baghlan Province to the east, Sar-e-Pul and Balkh provinces in the west, Bamiyan Province to the south, and its center is Aibak city. Samangan Province covers an area of (13437.8) km<sup>2</sup> and is divided into six districts, three of which include the cities of Roy e do aab, Khorram-Sarbagh, and the city of Hazrat Sultan, with Abik serving as the research's focal point. Samangan Province covers an area of (13437.8) km<sup>2</sup> and is divided into six districts, three of which include the cities of Roy e do aab, Khorram-Sarbagh, and the city of Hazrat Sultan, with Abik serving as the research's focal point (<https://www.mindat.org/loc-123204.html>) (Figure 1). According to the National Statistics Office's yearbook, the study area's population is



**Figure 1.** a – presents the geographical location of the Samangan Province; b – indicates study area

415343. (National Statistics and Information Authority of Afghanistan, 2019). and its population according to the yearbook of the National Statistics Office (415343) (National Statistics and Information Authority of Afghanistan, 2019).

### 3. Materials and methods

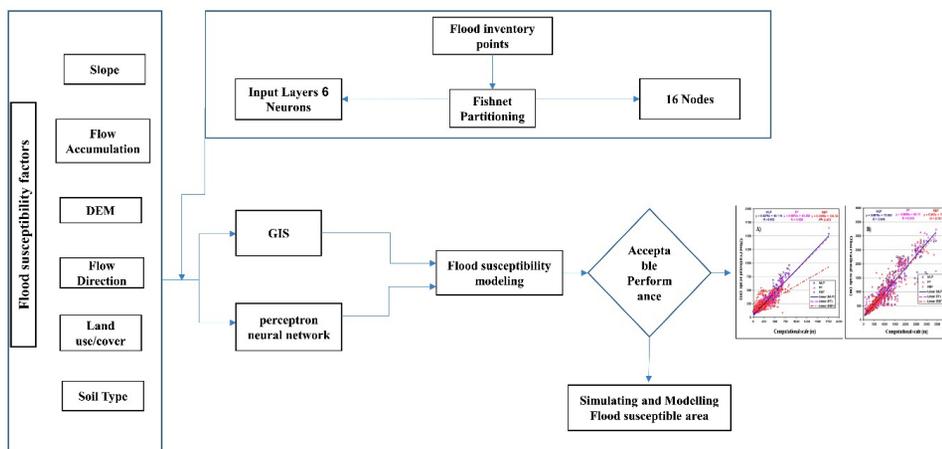
#### 3.1. General flowchart

The research methodology is based on the premise that the present and past are the keys to the future. Specifying and determining the factors' effectiveness is considered critical in developing a flood model. These criteria are defined according to the knowledge acquired from the literature review. Field studies for flood simulation and modeling are considered essential but impossible in all environmental conditions. Decision makers need to know where the next flood could occur to reduce or prevent casualties. Using GIS software, 30-meter resolution raster datasets were created, including elevation, soil type, land

use/cover, flow direction, flow accumulation, and slope of the research area. The rasters were partitioned using a 250×250 fishnet in GIS so that each pixel had a value in the attribute table. Then, each partitioned data layer is input to the perceptron neural network in a matrix to extract flood-prone areas. This process is illustrated in the general flowchart in Figure 2. Some significant elements of the flowchart are described in the following sections.

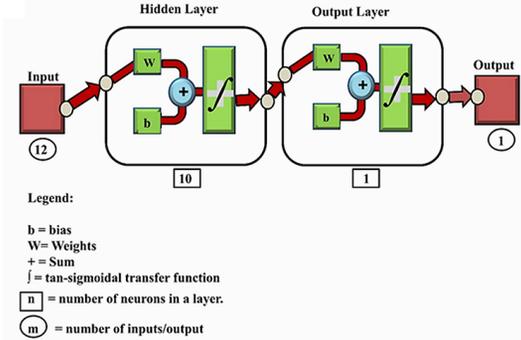
#### 3.2. Artificial neural network (ANN) models

A group of population-based meta-heuristic optimization techniques is known as neural network models. In most situations, neural network models are widely used as practical computational tools in various scientific areas for predicting intent (Zhang et al., 2020). The ANN approaches are inspired by biological nervous systems and exhibit self-compatibility, self-learning, fitness, error tolerance, and compatibility extension capabilities. In multivariate fitting analysis, improved learning efficacy reveals a high level of performance (Zhao et al., 2019). ANN model



**Figure 2.** General flowchart

optimization aims to discover the best possible solutions (Shirwaikar et al., 2019). In order to address a variety of real-world problems, ANN algorithms are widely used in data mining. In this research, it was utilized to simulate flooding. The architecture of artificial neural networks is presented in Figure 3.



**Figure 3.** The structure of an artificial neural network neuron

### 3.3. Multi-layer Perception Neural Network (MLPNN)

disasters and flood modeling (Shirzadi et al., 2017). The MLP approach uses neurons to represent input, hidden, and output layers. This method is associated with non-linear function activities. This method is used mainly in solving complicated problems based on its capability to demonstrate problems using a stochastic procedure (Costache et al., 2019). For crucial tribulations such as toughness estimation, this generally results in approximations. Cybenko (1989) proved that the MLPNN has a general purpose for calculations and can be used to establish statistical methods by regression study. When the response variables are specified, the categorization is considered an exact case of regression. The MLPNN establishes improved classifier algorithms. An extensively used ANN is the MLPNN comprising perceptron neurons (Mukerji et al., 2009). The mathematical structure of the multilayer perceptron (MLP) is presented as follows:

$$Y = f \left( \sum_{z=1}^n m_z x_z + b \right), \quad (1)$$

where  $Y$  represents the output,  $x_z$  represents the input vector ( $z = 1 \dots n$ ),  $f$  represents the transfer function,  $m_z$  is the vector weight, and  $b$  is the bias.

The training algorithm minimizes the global error  $E$  as follows:

$$E = \frac{1}{N} \sum_{n=1}^N E_n, \quad (2)$$

where  $N$  denotes the number of training samples;  $E_n$  denotes the error associated with the training pattern  $N$ ; and  $E_n$  is expressed as:

$$E_n = \frac{1}{2} \sum_{g=1}^n (o_g - t_g), \quad (3)$$

where  $n$  denotes the total number of output nodes,  $g$  indicates the  $g^{\text{th}}$  output node,  $o_g$  is the network output at the  $g^{\text{th}}$  output node, and  $t_g$  is the desired outputs at the  $g^{\text{th}}$  output node (Rezaeianzadeh et al., 2014).

### 3.4. Effective parameters in floods

In this study, six input parameters were used, including flow accumulation, slope, soil, land use, flow direction, and digital elevation model. All these factors effectively determine flooded areas (Figure 3). Spatial variables and datasets for flood simulation are stated in (Table 1).

**Table 1.** Spatial variables and flood simulation datasets

No	Parameters	Data Type	Spatial Resolution
1	Slope	Raster	30 m
2	Flow Accumulation	Raster	30 m
3	DEM	Raster	30 m
4	Flow Direction	Raster	30 m
5	Land use	Raster	30 m
6	Soil	Raster	Textural class

#### 3.4.1. Slope

The slope is an essential geomorphological attribute contributing to the runoff velocity and intensifying flood probability (Choubin et al., 2019; Rahmati et al., 2016). Flood velocity exhibits an inverse relationship with slope angle, whereby the likelihood of flooding rises as the slope angle decreases and declines as the slope angle increases (Costache, 2019). Due to its low elevation and flat topography, the Samangan river basin is considered to be more vulnerable to floods than the surrounding areas (Figure 4a).

#### 3.4.2. Soil

Soil classes are the most significant affecting elements that regulate runoff mechanisms (Xie et al., 2019). The water saturations are directly controlled by soil type features, which influence the precipitation runoff rate through other factors such as weather, erosion, rainfall intensity, and morphology (Phillips et al., 2019). There is a lower possibility of flash flooding where the saturation rate is high. This research uses the soil map of Afghanistan developed by the Geological Survey of the United States (USGS), and four soil classes are distributed in the study area (Figure 4b).

#### 3.4.3. Land use/cover

In flood modeling and susceptibility mapping, land use/cover is a significant influencing factor (Panahi et al., 2021). Specific types of land use have an impact on flooding by influencing runoff. The land use map in this study was created using the Google Earth engine and a two-year average of Landsat 8 images. It comprises four classes: vegetation, bare soils, snow-covered, and impervious surface (Figure 4c).

### 3.4.4. Digital Elevation Model (DEM)

Topography plays a significant role in flood intensity. Topographic influencing factors directly impact runoff velocity and speed. Flooded areas are located mainly in low elevated and low slope regions. Topographic maps and digital elevation models (DEM) are valuable tools for determining topographical flood-affecting elements responsible for flood events in a particular location. Since the flood model requires DEM data to determine flood-susceptible zones on the DEM (Pradhan, 2009), a DEM was created from 1:25000 topographic maps. (Figure 4d).

### 3.4.5. Flow direction

to a particular pixel, which indicates not only the drainage system but also the path of a future flood event (Kourgias & Karatzas, 2011). Floodplains, out-of-bank flow, and within the channel in valley floors are significant influencers of the flood (Van Appledorn et al., 2019). Expressly, in piedmont environments, either through channelized overland flows or water movement (similar to the current research area) where the flow directions differ extensively.

ArcGIS 10.8 is used to create the flow direction raster using the Shuttle Radar Topographic Mission (SRTM) digital elevation model (DEM) (Figure 4e).

### 3.4.6. Flow accumulation

Flow accumulation plays a vital role in specifying the speed and accumulation of a river network (Costache et al., 2020). A flow accumulation raster was created from the DEM data using ArcGIS 10.8 software by estimating accumulated flow for all pixels. The obtained raster was classified into two classes based on the natural breaks (Figure 4f).

## 4. Results and discussion

### 4.1. Flood simulation

in the input layer (land use, slope, soil, flow accumulation, flow direction, and digital elevation model) contributes to flooding in the basin. The output layer comprises a single stream that denotes the result as a flow area. The latent layer and the number of neurons are used to define the

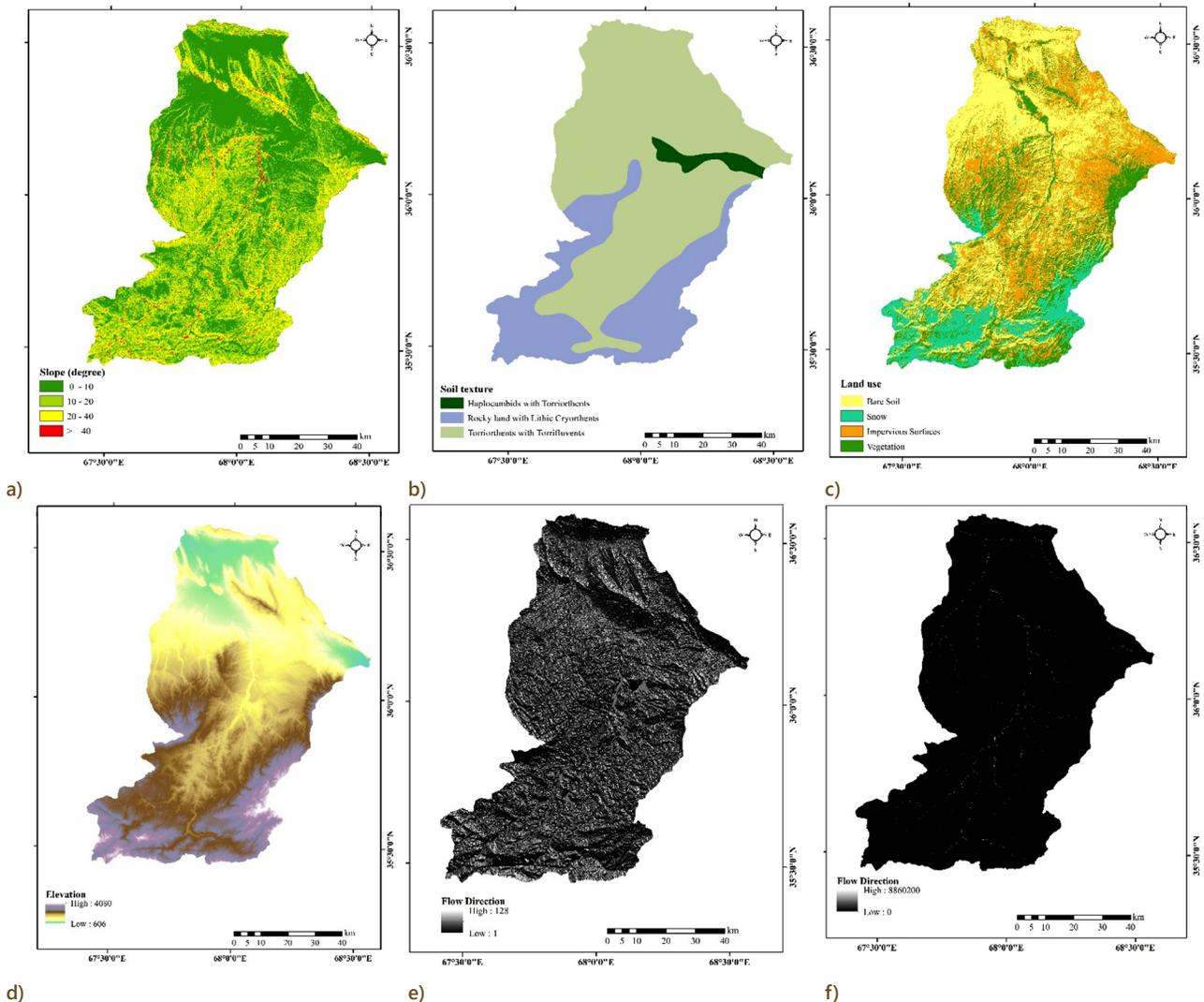
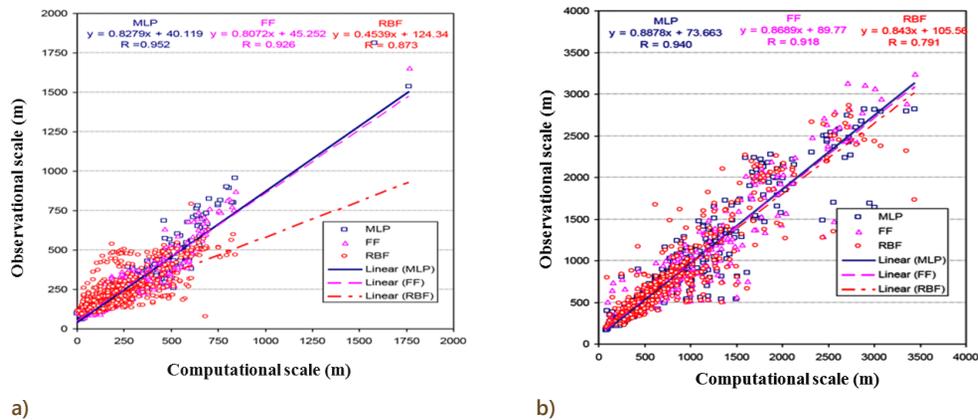


Figure 4. Flood conditioning factors: a – slope; b – soil; c – land use; d – elevation; e – flow direction; f – flow accumulation



**Figure 5.** Perceptron neural network output for observational and computational scaling

sophisticated association between variables and the outcome  $d$ . As a result, the number of neurons in the inputs is determined by the quantity of input layer resources. In the hidden and output layers, data is actively processed. Trial and error are frequently used to identify the number of hidden layers and neurons.

#### 4.2. Perceptron neural network training

After achieving the minimum error and completing the training, the artificial neural network uses the forward source storage structure to classify the entire dataset. In this study, an ANN was trained using a 6-N-N-1 structure, where  $N$  is the total number of nodes in the hidden layer, and 6 indicates the number of input layers. The algorithm was evaluated multiple times with varied numbers of neurons in both hidden layers to determine the optimum architecture of the neurons. This was evaluated based on accurate trading and testing processes. The training process was evaluated 16 times, and neurons were changed in each stage, and then the least square mean error was produced. The number of neurons in the two initial layers was evaluated using values from 5 to 21 and 4 to 12, respectively. In order to configure ANN, the training process evaluated initial independent conditions repeatedly until the best network performance was chosen. The procedure of decreasing the minimum square error in validation and training sets was applied for optimum decision. This training was terminated when the square mean error reached a minimum. This was accomplished by adding an initial stop procedure to the MATLAB software. This indicates that the network has been overtrained. In the training phase, it performs admirably as an ANN, but it is unable to retain that level of performance when using other datasets. In order to prevent errors, the data is divided into three distinct sections: training, validation, and testing. In order to prevent errors, the data is divided into three distinct sections: training, validation, and testing. The training validation error is reduced in the first stage. After a predetermined number of iterations, the training phase is terminated when a validation error reaches the minimum. MATLAB software was used to implement the Perceptron MLP. The inputs of the model are prepared

in GIS to facilitate easier calculation. The Marquardt algorithm is used to scale (normalize) the data, train the model, and perform the processing required to achieve the model output. During artificial neural network modeling, each of the six input nodes represents flood-causing factors such as slope, digital elevation model, soil, flow accumulation, flow direction, and land use. These parameters are I1, I2, I3, I4, and I5, in that order. The letters I6 are highlighted. The weight connection changes between the hidden and input layers are presented in Table 2. Table 3 provides that, with the exception of the slope parameter, there are only minor changes in the maximum and minimum weight of the connection between the nodes of the input and hidden layers (I1). The difference between the maximum and minimum connection weights for the slope parameter (I1) is greater than the other influencing factors. The output of the model training is shown in Figure 5. In flood prediction research, ANFIS has been widely used all over the world and has performed reasonably well in comparison to other AI modeling techniques (Nayak et al., 2005; Noymanee et al., 2017).

#### 4.3. Perceptron neural network testing

After completing the ANN training steps, different datasets (experimental data) were utilized to develop and define the accuracy of the model. The performance of the network was assessed using a new dataset. These datasets had similar characteristics to the previously used training data but were not employed during the model training. ANN demonstrated the ability to detect identical values at the training stage, which was a significant finding in the testing of this data. This outcome gives an  $R^2$  of 1, which is the most feasible output and represents an acceptable accuracy of prediction. In Figure 6, river flows are simulated, ANN predictions are made, and regression maps are displayed.

#### 4.4. Evaluation of simulation performance

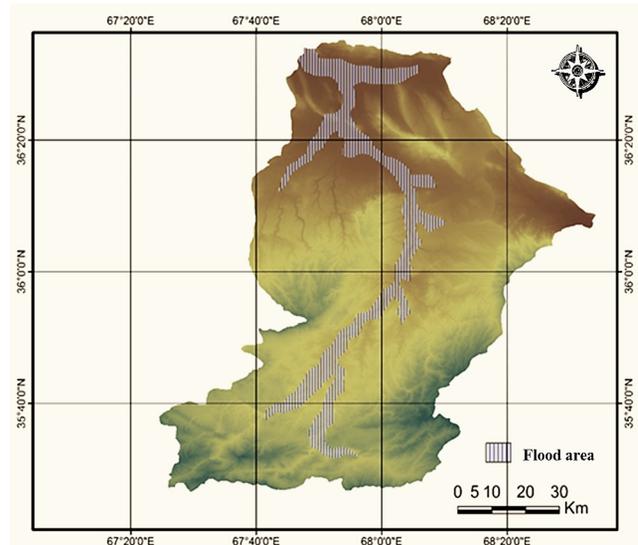
To evaluate the model accuracy regarding prediction error or the variance between actual and predicted values. In order to achieve feasible outputs, the most prevalent

**Table 2.** Input layers (I), hidden layers (AH) Weights communication

Node	I1 (Slope)	I2 (DEM)	I3 (Soil)	I4 (Flow a)	I5 (Land use)	I6 (Flow d)
HA1	22.5432	1.3E-07	1.5E-08	-2.8E-02	-3.2E-09	-2.2E-09
HA2	0.4563	-3.9E-03	-2.4E-02	1.6E-04	5.1E-03	4.1E-03
HA3	-1.3421	1.0E-02	7.3E-05	2.1E-07	7.9E-02	4.9E-03
HA4	-9.6581	-1.8E-03	-2.4E-02	1.3E-03	5.3E-05	5.3E-05
HA5	2.5642	-9.1E-08	-1.7E-04	4.4E-03	1.8E-02	1.8E-04
HA6	3.8731	-3.1E-05	-5.9E-07	3.8E-05	1.7E-02	1.7E-01
HA7	4.8741	-3.1E-09	-2.9E-04	3.8E-08	6.7E-03	3.6E-03
HA8	0.1331	1.2E-05	2.8E-08	5.9E-02	1.1E-05	1.2E-05
HA9	-2.2341	-1E-06	-3.5E-02	-9.7E-02	6.5E-08	3.5E-05
HA10	-23.7621	0.653	-9.4E-03	3.3E-05	-1.7E-01	-1.6E-01
HA11	-0.6541	1.1E-01	5.6E-03	-2.2E-05	2.2E-04	2.1E-03
HA12	5.9832	3.3E-07	-1.1E-06	3.2E-03	1.653	1.6 E 5
HA13	4.4147	2.1E-01	1.5E-01	1.2E-06	-1.1E-01	1.2E-01
HA14	-0.4352	2.5E-05	-5.3E-04	-8.6E-03	1.4E-04	1.7E-02
HA15	-0.1276	3.8E-01	9.3E-02	1.1E-06	-4.8E-02	4.8E-01
HA16	9.6532	6.1E-06	5.2E-01	2.8E-02	-4.9E-06	-4.2E-03

approaches were employed, including determination coefficient (R<sup>2</sup>), squared sum error (SSE), root-mean-squared error (RMSE), and mean squared error (MSE). In order to evaluate the efficacy of the model, a multilayer perceptron neural network (MLP) was computed using training and validation data to predict flood events (Table 3). The coefficient of determination (R<sup>2</sup>) shows that the prediction results of the MLP model match the training data very well. The SSE, MSE, and RMSE also suggest that the neural network models generally had minor errors (Table 4).

There is a very complex relationship between these watershed intrusions, geomorphology, etc., and they significantly impact each other and runoff. Interaction and adaptation of these parameters are essential for the hydrological modeling of the basin. The results of this study show the simulation of floods in Samangan, Afghanistan, with a neural network. The results suggest that the elevation factor is the most essential in floods. Elevation has the lowest weight (R<sup>2</sup> = 0.713) compared with flow direction (R<sup>2</sup> = 0.845) and slope (R<sup>2</sup> = 0.871). As a result, sensitivity evaluation is required to determine the weighting of input variables. These results can be utilized to facilitate gradient management and land planning. In addition, the methods used in this study can be applied to general purposes and planning. If this artificial neural network model can be combined with an alarm system or alarm sensor, the risk of floods in the study area can be somewhat reduced. Although the rest of the modeling approaches accurately simulated the flood region, Hong et al. (2018) found that NNbr, MLR, and Fmamdani models demonstrated more significant deviation. The analytical metrics MSE, RMSE, R<sup>2</sup>, NSE, MAE, and CA, were used to assess the flood prediction accuracy of all the models used.

**Figure 6.** Simulated flood prone area in Samangan, Afghanistan**Table 3.** Assessment and evaluation of model performance for perceptron neural network during training and testing

Parameter	Train	Test
R <sup>2</sup>	1	1
SSE	5.1E-07	5.2E-03
MSE	3.5E-08	3.5E-08
RMSE	2.5E-17	2.6E-17

**Table 4.** Analysis results and analysis for input programs

Parameter	SSE	MSE	RMSE	R <sup>2</sup>
Slope	2.54	0.00022	0.01352	0.871
DEM	3.43	0.00031	0.01521	0.713
Soil	1.87	0.00017	0.01861	0.913
Flow accumulation	5.32	0.00035	0.03321	0.845
Land Use	3.12	0.00017	0.02321	0.891

## 5. Conclusions

Floods are a significant hazard to communities and their assets, particularly in highly populated urban areas with increased impenetrability. Surfaces increase flood occurrence by intensifying surface runoff. In this research, the performance of the perceptron artificial neural network and GIS for predicting flood-prone areas in the Samangan Province of Afghanistan was evaluated using slope, elevation, flow direction, soil, and land use parameters. The results showed that integrating artificial neural network techniques with GIS in flood simulation and modeling has helped prevent floods in different spatial environments. The results from this research show that the most significant factor for causing floods in spatial environments is the elevation factor, which in this study has the lowest weight ( $R^2 = 0.713$ ), and the highest weight parameter is related to the flow direction ( $R^2 = 0.913$ ). It can be concluded that the artificial neural network MLP can simulate and model flood-prone areas with insufficient spatial data and little information about the study area, such as the physiographic characteristics of the basin. By combining the MLP and GIS neural network models, it is possible to identify areas where floods are likely to occur. By applying proper management regulations, it is possible to provide suitable conditions for establishing the ecological balance of the basin and protecting it from flood hazards.

## Acknowledgements

We would like to thank anonymous reviewers for their comments that improved this paper.

## Data availability

Data will be available upon request.

## Compliance with ethical standards

The authors have no conflict of interest.

## References

- Bui, Q. T., Nguyen, Q. H., Nguyen, X. L., Pham, V. D., Nguyen, H. D., & Pham, V. M. (2020). Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. *Journal of Hydrology*, *581*, Article 124379. <https://doi.org/10.1016/j.jhydrol.2019.124379>
- Cabrera, J. S., & Lee, H. S. (2020). Flood risk assessment for Davao Oriental in the Philippines using geographic information system-based multi-criteria analysis and the maximum entropy model. *Journal of Flood Risk Management*, *13*(2), Article e12607. <https://doi.org/10.1111/jfr3.12607>
- Chang, H., Lafrenz, M., Jung, I. W., Figliozzi, M., Platman, D., & Pederson, C. (2012). Potential impacts of climate change on flood-induced travel disruptions: a case study of Portland, Oregon, USA. In *Geography of climate change* (pp. 231–245). Routledge. <https://doi.org/10.4324/9780203723364>
- 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. *Science of the Total Environment*, *651*, 2087–2096. <https://doi.org/10.1016/j.scitotenv.2018.10.064>
- Costache, R. (2019). Flood susceptibility assessment by using bivariate statistics and machine learning models – a useful tool for flood risk management. *Water Resources Management*, *33*(9), 3239–3256. <https://doi.org/10.1007/s11269-019-02301-z>
- Costache, R., Pham, Q. B., Avand, M., Linh, N. T. T., Vojtek, M., Vojteková, J., Lee, S., Khoi, D. N., Thao Nhi, P. T., & Dung, T. D. (2020). Novel hybrid models between bivariate statistics, artificial neural networks and boosting algorithms for flood susceptibility assessment. *Journal of Environmental Management*, *265*, Article 110485. <https://doi.org/10.1016/j.jenvman.2020.110485>
- Costache, R., Pham, Q. B., Sharifi, E., Linh, N. T. T., Abba, S., Vojtek, M., Vojteková, J., Nhi, P. T. T., & Khoi, D. N. (2019). Flash-flood susceptibility assessment using multi-criteria decision making and machine learning supported by remote sensing and GIS techniques. *Remote Sensing*, *12*(1), Article 106. <https://doi.org/10.3390/rs12010106>
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, *2*(4), 303–314. <https://doi.org/10.1007/BF02551274>
- Dano, U. L., Balogun, A. L., Matori, A. N., Wan Yusouf, K., Abubakar, I. R., Said Mohamed, M. A., Aina, Y. A., & Pradhan, B. (2019). Flood susceptibility mapping using GIS-based analytic network process: A case study of Perlis, Malaysia. *Water*, *11*(3), Article 615. <https://doi.org/10.3390/w11030615>
- Desai, B., Maskrey, A., Peduzzi, P., De Bono, A., & Herold, C. (2015). *Making development sustainable: The future of disaster risk management, global assessment report on disaster risk reduction*. United Nations Office for Disaster Risk Reduction (UNISDR), Genève, Suisse. <https://archive-ouverte.unige.ch/unige:78299>
- Fernandes, O., Murphy, R., Adams, J., & Merrick, D. (2018, August). Quantitative data analysis: CRASAR small unmanned aerial systems at hurricane Harvey. In *2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)* (pp. 1–6). IEEE. <https://doi.org/10.1109/SSRR.2018.8468647>
- Hallegatte, S., Ranger, N., Mestre, O., Dumas, P., Corfee-Morlot, J., Herweijer, C., & Wood, R. M. (2011). Assessing climate change impacts, sea level rise and storm surge risk in port cities: A case study on Copenhagen. *Climatic Change*, *104*(1), 113–137. <https://doi.org/10.1007/s10584-010-9978-3>
- Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., & Kanae, S. (2013). Global flood risk under climate change. *Nature Climate Change*, *3*(9), 816–821. <https://doi.org/10.1038/nclimate1911>
- Hong, H., Panahi, M., Shirzadi, A., Ma, T., Liu, J., Zhu, A. X., Chen, W., Kougiyas, I., & Kazakis, N. (2018). Flood susceptibility assessment in Hengfeng area coupling adaptive neuro-fuzzy inference system with genetic algorithm and differential evolution. *Science of the Total Environment*, *621*, 1124–1141. <https://doi.org/10.1016/j.scitotenv.2017.10.114>
- Hosseini, F. S., Choubin, B., Mosavi, A., Nabipour, N., Shamshirband, S., Darabi, H., & Haghghi, A. T. (2020). Flash-flood hazard assessment using ensembles and Bayesian-based machine learning models: Application of the simulated annealing feature selection method. *Science of the Total Environment*, *711*, Article 135161. <https://doi.org/10.1016/j.scitotenv.2019.135161>
- Huang, K., Li, X., Liu, X., & Seto, K. C. (2019). Projecting global urban land expansion and heat island intensification through

2050. *Environmental Research Letters*, 14(11), Article 114037. <https://doi.org/10.1088/1748-9326/ab4b71>
- Jahandideh-Tehrani, M., Zhang, H., Helfer, F., & Yu, Y. (2019). Review of climate change impacts on predicted river streamflow in tropical rivers. *Environmental Monitoring and Assessment*, 191(12), 1–23. <https://doi.org/10.1007/s10661-019-7841-1>
- Jonkman, S. N., & Vrijling, J. K. (2008). Loss of life due to floods. *Journal of Flood Risk Management*, 1(1), 43–56. <https://doi.org/10.1111/j.1753-318X.2008.00006.x>
- Khosravi, K., Shahabi, H., Pham, B. T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.-B., Gróf, G., Ho, H. L., Hong, H., Chapi, K., & Prakash, I. (2019). A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods. *Journal of Hydrology*, 573, 311–323. <https://doi.org/10.1016/j.jhydrol.2019.03.073>
- Kourgialas, N. N., & Karatzas, G. P. (2011). Flood management and a GIS modelling method to assess flood-hazard areas – a case study. *Hydrological Sciences Journal–Journal des Sciences Hydrologiques*, 56(2), 212–225. <https://doi.org/10.1080/02626667.2011.555836>
- Metin, A. D., Dung, N. V., Schröter, K., Vorogushyn, S., Guse, B., Kreibich, H., & Merz, B. (2020). The role of spatial dependence for large-scale flood risk estimation. *Natural Hazards and Earth System Sciences*, 20(4), 967–979. <https://doi.org/10.5194/nhess-20-967-2020>
- Mishra, K., & Sinha, R. (2020). Flood risk assessment in the Kosi alluvial plains using Analytical Hierarchy Process (AHP) framework, North Bihar, India. *Geomorphology*, 350, Article 106861. <https://doi.org/10.1016/j.geomorph.2019.106861>
- Mukerji, A., Chatterjee, C., & Singh Raghuvanshi, N. (2009). Flood forecasting using ANN, neuro-fuzzy, and neuro-GA models. *Journal of Hydrologic Engineering*, 14(6), 647–652. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000040](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000040)
- National Statistics and Information Authority of Afghanistan. (2019). *Central Yearbook of the Central Statistics Country*. <http://nsia.gov.af/library>
- Nayak, P. C., Sudheer, K. P., Rangan, D. M., & Ramasastri, K. S. (2005). Short-term flood forecasting with a neurofuzzy model. *Water Resources Research*, 41(4). <https://doi.org/10.1029/2004WR003562>
- Neumann, B., Vafeidis, A. T., Zimmermann, J., & Nicholls, R. J. (2015). Future coastal population growth and exposure to sea-level rise and coastal flooding—a global assessment. *PLoS ONE*, 10(3), Article e0118571. <https://doi.org/10.1371/journal.pone.0118571>
- Noymanee, J., Nikitin, N. O., & Kalyuzhnaya, A. V. (2017). Urban pluvial flood forecasting using open data with machine learning techniques in pattani basin. *Procedia Computer Science*, 119, 288–297. <https://doi.org/10.1016/j.procs.2017.11.187>
- Panahi, M., Jaafari, A., Shirzadi, A., Shahabi, H., Rahmati, O., Omidvar, E., Lee, S., & Bui, D. T. (2021). Deep learning neural networks for spatially explicit prediction of flash flood probability. *Geoscience Frontiers*, 12(3), Article 101076. <https://doi.org/10.1016/j.gsf.2020.09.007>
- Phillips, T. H., Baker, M. E., Lautar, K., Yesilonis, I., & Pavao-Zuckerman, M. A. (2019). The capacity of urban forest patches to infiltrate stormwater is influenced by soil physical properties and soil moisture. *Journal of Environmental Management*, 246, 11–18. <https://doi.org/10.1016/j.jenvman.2019.05.127>
- Pradhan, B. (2009). Groundwater potential zonation for basaltic watersheds using satellite remote sensing data and GIS techniques. *Central European Journal of Geosciences*, 1(1), 120–129. <https://doi.org/10.2478/v10085-009-0008-5>
- Rahmati, O., Pourghasemi, H. R., & Zeinivand, H. (2016). Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. *Geocarto International*, 31(1), 42–70. <https://doi.org/10.1080/10106049.2015.1041559>
- Rezaeianzadeh, M., Tabari, H., Arabi Yazdi, A., Isik, S., & Kalin, L. (2014). Flood flow forecasting using ANN, ANFIS and regression models. *Neural Computing and Applications*, 25(1), 25–37. <https://doi.org/10.1007/s00521-013-1443-6>
- Sachdeva, S., Bhatia, T., & Verma, A. K. (2017). Flood susceptibility mapping using GIS-based support vector machine and particle swarm optimization: A case study in Uttarakhand (India). In *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1–7). IEEE. <https://doi.org/10.1109/ICCCNT.2017.8204182>
- Santos, P. P., & Reis, E. (2018). Assessment of stream flood susceptibility: a cross-analysis between model results and flood losses. *Journal of Flood Risk Management*, 11, S1038–S1050. <https://doi.org/10.1111/jfr3.12290>
- Schubert, J. E., & Sanders, B. F. (2012). Building treatments for urban flood inundation models and implications for predictive skill and modeling efficiency. *Advances in Water Resources*, 41, 49–64. <https://doi.org/10.1016/j.advwatres.2012.02.012>
- Seejata, K., Yodying, A., Wongthadam, T., Mahavik, N., & Tantaanee, S. (2018). Assessment of flood hazard areas using analytical hierarchy process over the Lower Yom Basin, Sukhothai Province. *Procedia Engineering*, 212, 340–347. <https://doi.org/10.1016/j.proeng.2018.01.044>
- Shahabad, S. I., Zhang, Z., Keshavarzkermani, A., Ali, U., Mahmoodkhani, Y., Esmaeilzadeh, R., Bonakdar, A., & Toyserkani, E. (2020). Heat source model calibration for thermal analysis of laser powder-bed fusion. *The International Journal of Advanced Manufacturing Technology*, 106(7–8), 3367–3379. <https://doi.org/10.1007/s00170-019-04908-3>
- Shafapour Tehrany, M., Kumar, L., & Shabani, F. (2019). A novel GIS-based ensemble technique for flood susceptibility mapping using evidential belief function and support vector machine: Brisbane, Australia. *PeerJ*, 7, Article e7653. <https://doi.org/10.7717/peerj.7653>
- Shirwaikar, R. D., Acharya, D., Makkithaya, K., Surulivelrajan, M., & Srivastava, S. (2019). Optimizing neural networks for medical data sets: A case study on neonatal apnea prediction. *Artificial Intelligence in Medicine*, 98, 59–76. <https://doi.org/10.1016/j.artmed.2019.07.008>
- Shirzadi, A., Shahabi, H., Chapi, K., Bui, D. T., Pham, B. T., Shaheidi, K., & Ahmad, B. B. (2017). A comparative study between popular statistical and machine learning methods for simulating volume of landslides. *Catena*, 157, 213–226. <https://doi.org/10.1016/j.catena.2017.05.016>
- Tehrany, M. S., Kumar, L., & Shabani, F. (2019). A novel GIS-based ensemble technique for flood susceptibility mapping using evidential belief function and support vector machine: Brisbane, Australia. *PeerJ*, 7, Article e7653. <https://doi.org/10.7717/peerj.7653>
- Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2015). Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. *Stochastic Environmental Research and Risk Assessment*, 29(4), 1149–1165. <https://doi.org/10.1007/s00477-015-1021-9>
- Tien Bui, D., Khosravi, K., Shahabi, H., Daggupati, P., Adamowski, J. F., Melesse, A. M., Thai Pham, B., Pourghasemi, H. R., Mahmoudi, M., Bahrami, S., Pradhan, B., Shirzadi, A., Chapi K., & Lee, S. (2019). Flood spatial modeling in northern Iran using remote sensing and gis: A comparison between evidential

- belief functions and its ensemble with a multivariate logistic regression model. *Remote Sensing*, 11(13), Article 1589. <https://doi.org/10.3390/rs11131589>
- Van Appledorn, M., Baker, M. E., & Miller, A. J. (2019). River-valley morphology, basin size, and flow-event magnitude interact to produce wide variation in flooding dynamics. *Ecosphere*, 10(1), Article e02546. <https://doi.org/10.1002/ecs2.2546>
- Wang, Y., Hong, H., Chen, W., Li, S., Pamučar, D., Gigović, L., Drobnjak, S., Tien Bui, D., & Duan, H. (2018). A hybrid GIS multi-criteria decision-making method for flood susceptibility mapping at Shangyou, China. *Remote Sensing*, 11(1), Article 62. <https://doi.org/10.3390/rs11010062>
- Xie, X., He, Z., Chen, N., Tang, Z., Wang, Q., & Cai, Y. (2019). The roles of environmental factors in regulation of oxidative stress in plant. *BioMed Research International*, 2019, Article 9732325. <https://doi.org/10.1155/2019/9732325>
- Xu, L., Wang, X., Liu, J., He, Y., Tang, J., Nguyen, M., & Cui, S. (2019). Identifying the trade-offs between climate change mitigation and adaptation in urban land use planning: An empirical study in a coastal city. *Environment International*, 133, Article 105162. <https://doi.org/10.1016/j.envint.2019.105162>
- Zhang, Y., Jin, Z., & Chen, Y. (2020). Hybrid teaching-learning-based optimization and neural network algorithm for engineering design optimization problems. *Knowledge-Based Systems*, 187, Article 104836. <https://doi.org/10.1016/j.knosys.2019.07.007>
- Zhao, B., Ren, Y., Gao, D., Xu, L., & Zhang, Y. (2019). Energy utilization efficiency evaluation model of refining unit Based on Contourlet neural network optimized by improved grey optimization algorithm. *Energy*, 185, 1032–1044. <https://doi.org/10.1016/j.energy.2019.07.111>