Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Flood Susceptibility Mapping in Arid Region of Pakistan through ensemble Machine Learning Model.

Andaleeb Yaseen
Wuhan University
Jianzhong LU (lujzhong@whu.edu.cn)
Wuhan University
Chen Xiaoling
Wuhan University

Research Article

Keywords: Flash Flood, Logistic Regression, Support Vector Machine, Multi-Layer Perceptron, Ensemble Classifier.

Posted Date: January 25th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-928677/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at Stochastic Environmental Research and Risk Assessment on February 11th, 2022. See the published version at https://doi.org/10.1007/s00477-022-02179-1.

Flood Susceptibility Mapping in Arid Region of Pakistan through ensemble Machine Learning Model.

Andaleeb Yaseen[a], Jianzhong Lu[a]*, Xiaoling Chen[a]

a The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan 430079, China.

*corresponding author, Email: lujzhong@whu.edu.cn, Tel: +86-27-68778755

Abstract

Floods are among the most destructive natural hazards. Therefore, their prediction is pivotal for flood management and public safety. Factors contributing to flooding are different for every region as they depend upon the characteristics of each region. Therefore, this study evaluated the factors contributing to flood and the precise location of high and very high flood susceptibility regions in Karachi. A new ensemble model (LR-SVM-MLP) is introduced to develop the susceptibility map and evaluate influencing factors. This ensemble model was formed by employing a stacking ensemble on Logistic Regression (LR), Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP). A spatial database was generated for the Karachi watershed, which included; twelve conditioning factors as independent variables, 652 flood points and the same number of non-flood points as dependent variables. This data was then randomly divided into 70% and 30% to train and validate models, respectively. To analyse the collinearity among factors and to scrutinize each variable's predictive power, multicollinearity test and Information Gain Ratio were applied, respectively. After training, the models were evaluated on various statistical measures and compared with benchmark models. Results revealed that the proposed ensemble model outperformed Logistic Regression (LR), Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP) and produced a precise and accurate map. Results of the ensemble model showed 99% accuracy in training and 98% accuracy in validation datasets. This ensemble model can be used by flood management authorities and the government to contribute to future research studies.

Keywords: Flash Flood, Logistic Regression, Support Vector Machine, Multi-Layer Perceptron, Ensemble Classifier.

1 Introduction

Floods are considered destructive natural hazards that causes millions of human deaths and result in billions of economic losses worldwide. The flood threat cannot be put aside because of future climate change Odoh and Chilaka (2012). It is predicted that the majority of the world will be threatened by the flood frequency and intensity in future**.** Information about inundation can be obtained through different techniques of remote sensing, either airborne or space-borne Schumann and Moller (2015).

The generation of susceptibility maps for flash floods is challenging, especially in large regions, as flash floods are quite complicated because they are area-dependent and arise nonlinearly to the variety of Spatio-temporal scales (Ahmadlou et al. 2019). Recently, promising results of machine learning models have been reported in literature worldwide for solving problems related to natural hazards.

For flood modelling, some of the qualitative Multicriteria decision models include Analytical Hierarchy Process (AHP) (Kazakis et al. 2015; Rahmati et al. 2016), Fuzzy AHP (Ekmekcio˘glu et al. 2020), SAW (Meshram et al. 2020), Interval Rough AHP (Sepehri et al. 2020), Frequency Ratio (FR) (Lee et al. 2012; Tehrany et al. 2015a), Weights of Evidence (WOE) (Rahmati et al. 2016; Tehrany et al. 2014a), and quantitative AI models for example; Logistic Regression (LR) (Fekete 2009; Tehrany et al. 2014a), Neuro-Fuzzy Logic (Tien Bui et al. 2016a; Mukerji et al. 2009), Decision Trees (DT) (Tehrany et al. 2013), Support Vector Machine (SVM) (Tehrany et al. 2015b; Tehrany et al. 2015a) were used in the past. Among other ML models, MLP is one of the most used models for flood modelling as it has the capability of processing non-linear and multivariate data and has the potential of universal modelling (Youssef et al. 2011). Due to its prediction accuracy, SVM is becoming an emerging choice for hydrologists (Zhao et al. 2018; Tehrany et al. 2015; Choubin et al. 2019). However, as different models act differently in given scenarios and carry some drawbacks, no consensus has yet been reached on a single model for flood susceptibility modelling. Therefore, scientists now address these problems by forming hybrids of different machine learning models together. Ensemble or hybrid machine learning models showed high accuracy and better performance than conventional methods in many previous studies (Pham et al. 2017; Saha et al. 2021).

Hybrid methods are extensively used in literature for flood modelling. Some ensemble methods, for example, adaptive neuro-fuzzy interference systems and their optimization algorithms have

become famous for their effective prediction (Bui et al. 2015). Chapi et al. (2017) applied bagging ensemble on Logistic Model Tree, which performed best compared with other models. Similarly, Ngo et al. (2018) developed a new hybrid approach (FA-LM-ANN) by integrating Firefly Algorithm (FA), Levenberg-Marquardt (LM), and Artificial Neural Network (ANN) to study flood susceptibility, which proved to be the best model as compared to its benchmark models.

For this study, Logistic Regression (LR), Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) were chosen. Logistic Regression is simple but extremely effective in evaluating the relationship between dependent and independent variables for the classification problem (Chapi et al. 2017). For flood susceptibility, SVM uses the flood predictors' non-linear transformations in higher dimensional feature space (Yilmaz 2010; Ghorbanzadeh et al. 2019; Nguyen et al. 2019). SVM reduces the test error by finding optimal hyper-plane that could separate flood and non-flood (support vectors) (Kalantar et al. 2018). MLP has been widely used for natural hazard prediction because it is highly capable of modelling the non-linear relationship between an explanatory variable and target variable (Kia et al. 2012).

The stacking classifier technique is used to make an ensemble of these models. The stacking classifier is reliable because it generates predictions based on two-level information. First, the base classifier predicts and then the meta classifier refines any biases occur in the prediction of base classifier (Hu et al. 2020). To make the prediction more efficient and reliable, choosing suitable models for base and meta classifier is essential.

The Novelty of this research is 1) Ensemble of these three models using stacking classifier was never done before. 2) The effectiveness of this hybrid (LR_SVM-MLP) for flood susceptibility was never accessed. 3) To that of the authors' best knowledge, no single study has been done for flood susceptibility analysis in Pakistan using machine learning models and contributing factors. The main objectives of this study were to; 1) To compare the performance of the new ensemble model on flood susceptibility with commonly used models 2) To evaluate the importance of flood conditioning factors and their contribution to flood susceptibility in an arid region of Pakistan.

For the training process, 12 conditioning factors and 652 flood events were used. All models were trained individually, and their performances were compared with the ensemble model using several statistical measures. The fundamental purpose of this technique is not only to highlight flood-prone

regions precisely and accurately but also to analyze the contributing factors so that authorities and government can plan flood management accordingly.

2 Methods

2.1 Study Area

Located in the 24.86 N and 67.010 E in world coordinates, Karachi lies in the southern region of Pakistan. An increment from 466.5 Km^2 to 666.18 Km^2 was observed in the built-up area of Karachi. It has a very humid to hot and worst climate in the summer season with very low annual precipitation (Raza et al. 2019).

Two main rivers which pass from Karachi are Malir and Liyari. Malir flows from the east towards the south, while Liyari flows North to the Southwest (Tariq et al.2016). Karachi's population jumped from 450,000 in 1947 (Hassan. 2017) to more than 16 million in 2017 (Shahbaz et al. 2017) and is expected to reach more than 20 million by the year 2025 (Mangi et al. 2020). Karachi has the major seaport of Pakistan and faces floods every 2-3 years. Since 2000, Karachi faced floods in 2000, 2006, 2007, 2011, 2013, 2019, and 2020. Total 42 people were killed in the flood of 2019 and 41 in the flood of 2020 (DAWN 2020). (Figure 1)

2.2 Flood conditioning factors and database generation

Factors that affect flood probability vary as they mainly depend on the characteristics of particular region (Tien Bui et al.2016a). Therefore, it is vital to determine the influencing factors for each watershed for the accurate flood susceptibility mapping of that area (Chapi et al.2017).

This study is primarily based on geospatial data, extracted from Digital Elevation Model (DEM) and some other factors. A total of 12 factors were analyzed, which are elevation, Slope, Curvature, Stream Power Index (SPI), Topographic Wetness Index (TWI), Rainfall, Lithology, soil Type, Distance from Stream, Stream Density, Normalized Difference Vegetation Index (NDVI), and LandUse.

Table 1: Sources of different data used in the study.

Factors	Source	Links
Elevation Slope Curvature Stream Power Index (SPI) Topographic Wetness Index (TWI)	Digital Elevation Model (DEM 30m)	Advance Spaceborne Thermal Emission and Reflection (ASTER) as Global Elevation Model (GDEM) https://asterweb.jpl.nasa.gov/gdem.asp

Distance to River Stream Density		
Lithology	Geological Map	https://pubs.er.usgs.gov/publication/ofr97470C
Soil Type	Soil Map	Food and agricultural Organization (FAO) website http://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/en/
Land Use NDVI	OLI of Landsat 8 image	U.S. Geological Survey (USGS) https://glovis.usgs.gov/app?fullscreen=0
Rainfall	GIOVANNI	https://giovanni.gsfc.nasa.gov/giovanni/

This data was processed in ArcGIS 10.8, and the database was constructed as the matrix of 3459 columns and 3260 rows and spatial database of cell size (30m*30m). It is essential to assemble data based on past flood events for future flood prediction Manandhar Bikram (2010). For this study, the flood of 2019 and 2020 were used for data collection. As a large number of dependent variables gives more accurate results, Total 652 flood points were collected from the areas that regularly faced floods in the past, and a similar number of non-flood points were collected to avoid biases of data. This data was then randomly classified as training data (70%) and validation data (30%); this method of random partition division repeatedly split the data without replacing the training and test dataset. The choice of contributing factors for this study was made using the information of previous literature (Tehrany et al., 2014; Rahmati et al., 2016; Khosravi et al., 2016a).

The elevation is considered one of the most essential factors in flood analysis (Dodangeh et al. 2020). Previous research has shown that elevation and flood occurrence has an inverse relation with elevation; floods usually occur at relatively low elevated areas compared to the high elevation regions of the same area (Chen et al. 2020). The elevation of Karachi lies from -10m to 501 meters. The elevation was constructed with five intervals from -10 – 59m, 59.001 – 122m, 122.01m – 198m, 198.01m – 304m, and 304.01m – 501m (Fig. 2).

The surface of the slope area is highly pertinent to flood as floodwater flow with greater speed from the steep surface, and water tends to get absorbed in plain areas, led more damage (Stevaux et al. 2020). The slope of the study area was classified into five intervals: 0 - 2.018, 2.018 - 5.55, 5.55 - 12.11, 12.11 - 23.46, 23.46 - 64.33. Curvature effect the flooding water budge; therefore, its study is necessary for flood modelling (Ahmadlou et al. 2019). The curvature value of study area lied in five intervals: 0.5485 - 20.159, 0.014878 - 0.54849, -0.25192 - 0.014877, -0.65213 - -0.25193, -13.859 - -0.65214. The topographic wetness index ascertained the spatial wetness status

of the basin area, which affects the occurrence of floods in the region (Meles et al. 2020). TWI was calculated as follows:

$$\text{TWI} = Ln\left(\frac{\partial}{tan\beta}\right) \quad (1)$$

Where, ∂ is the cumulative upslope area drainage through the point per unit contour length and tanβ is the slope angle at the point. TWI was classified into the five classes: 2.8124 - 6.9503, 6.9504 - 8.541, 8.5418 - 10.451, 10.452 - 13.157, and 13.158 - 23.104. (Fig. 2). SPI determined the erosive power and discharge relative to a particular area (Poudyal et al. 2010).

$$\text{SPI} = \text{As} \tan \beta \quad (2)$$

Where As is the area of specific basin, and β is the local slope gradient (in degree).

Distance from the river is essential for identifying flood-prone areas in the watershed Tehrany et al. (2015a). Distance from the river was extracted in ArcGIS environment using Euclidean distance buffer and classified into eight classes: 0-0.5km, 0.5-1km, 1-1.5km, 1.5-2km, 2.5-3km, 3-3.5km, 3.5-4km, >4km. Stream density is another critical influencing factor while studying the flood susceptibility of an area. It is calculated by Elmore et al. (2013).

$$\text{Stream density} = \frac{stream\ length\ (m)}{Basin\ Area\ (km^2)} \quad (3)$$

Stream density was extracted from DEM in ArcGIS environment using line density. The five classes of stream density are: 0 - 20.603, 20.604 - 48.338, 48.339 - 77.657, 77.658 - 114.11, 114.12 - 202.07. The probability of flooding significantly increases with increasing rainfall events and the duration of rain (Lu et al.2020). Rainfall with the monthly average of the last ten years was used in this study. Rainfall data was downloaded from GIOVANNI, and interpolation was calculated in the ArcGIS environment using Inverse Distance Weighted (IDW). This method has been proved to be effective for rainfall interpolation by Chen and Liu et al. (2012). The interpolation result was converted into five intervals which are, 15.083 - 15.509mm, 15.51 - 15.962mm, 15.963 - 16.424mm, 16.425 - 16.796mm, and 16.797 - 17.394mm.

Lithology and soil type affects the hydrology of the basin. Areas with highly permeable subsoils and higher resistant rocks allow minor drainage (Çelik et al. 2012; Srivastava et al. 2014). The lithology and soil regulate flooding by regulating the erodibility and permeability in a watershed Stefanidis and Stathis (2013). Lithological data was obtained from the USGS site, while soil data was obtained from Food and Agriculture Organization (FAO). Both of these data were processed to extract data from the study area. The study area comprises four types of rocks and two

types of soil (Fig. 2). Land Use types are significant indicators while assessing the probability of a flood in any area (Rahmati et al.2015). Land use types and NDVI greatly influence and control the infiltration; for example, areas with high forest land allow more infiltration and thus less runoff compared to the areas with the concrete surface (Tehrany et al. 2014a; Tehrany et al. 2014b). For land use analysis, the data was obtained from Raza et al.(2019); they classified Karachi into vegetation class, water class, settlement, and barren land. Similarly, NDVI is used to assess the relationship between vegetation and flooding. (Tehrany et al.2013). The index of NDVI ranges between -1 to 1. NDVI was calculated using Landsat 8 OLI imagery. After preprocessing of imagery, NDVI was calculated in ArcGIS using the formula:

$$NDVI = \frac{(\text{near infrared} \_ \text{red})}{(\text{near infrared} + \text{red})} \quad (4)$$

2.2 Multicollinearity Test

If the study uses multiple independent factors, the presence of collinearity can significantly influence the final results. Therefore, it is necessary to ascertain the collinearity among those factors (Arabameri et al. 2019; Wang et al. 2021). Variance Inflation Factor (VIF) helps in detection of multicollinearity among independent variables that can be further used in the model (Arabameri et al.2020a). Factor having VIF values greater than 4 pose severe concerns with multicollinearity. If the test reports any collinear variable, then this specific variable should be removed and not used for the prediction purpose in the model (Arabameri et al. 2019; Bui et al. 2019). In this study, twelve conditioning factors were considered for analysis; therefore, it was essential to check their collinearity. The VIF among variables can determined by using the following equation. Where Tolerance is referred to as variability shown by an independent variable and not explained by other independent variables.

$$\text{VIF} = \frac{1}{\text{Tolerance}} \quad (5)$$

2.3 Information Gain Ratio

As the present study consists of numerous factors, there is the possibility that some factors can reduce the model's performance; therefore, to reduce uncertainty and noise in the result, it is necessary to determine the predictive power of each factor. If the value of the factor is 0, it will be removed from the study. Many previous researchers select this value to determine the factors that have zero influence on the study (Chapi et al. 2017; Khosravi et al. 2016b). The predictive power of influencing factors can be evaluated using Information Gain Ratio (IGR); the higher the value of IGR, the higher the predictive power for the flood (Chapi et al.2017).

Let D be the training dataset composed of n number of samples, and n (Yi, D) is the number of samples present in the training data D, belongs to the label of class Yi (flood or non-flood). The IGR for each conditioning factor, for example, the slope, will obtain as follows:

$$IGR\ (D, Slope) = \frac{Entropy\ (D) - Entropy\ (D, Slope)}{SplitEntropy(D, Slope)} \quad (6)$$

$$Entropy\ (D) = -\sum_{i=1}^{2} \frac{n(Yi, Slope)}{|D|} log2 \frac{n(Yi, Slope)}{|D|} \quad (7)$$

$$Entropy\ (D, Slope) = \sum_{j=1}^{m} \frac{Dj}{|D|} Entropy(D) \quad (8)$$

$$SplitEntropy(D, Slope) = -\sum_{j=1}^{m} \frac{|Dj|}{|D|} log2 \frac{|Dj|}{|D|} \quad (9)$$

2.4 Frequency Ratio

To analyse the relative importance of each factor and the contribution of each class of factors towards flood, flood frequency ratios were analysed (Gayen et al. 2019). The value of the frequency ratio can be expressed as the percentage of flood pixels shared in the domain of pixels of the whole area. FR can be calculated as follows:

$$FR = \frac{Li/Ci}{L/C} \quad (10)$$

Where Li represents the flood cells in the i^{th} category, whereas Ci represents total cells in the i^{th} category, L is the total flood cells, and C is the total cells. The FR values greater than 1 represent the large concentration of that class in the flood area, while if the value of the class is less than 1 it suggests that this class has a very little contribution of flood cells in the data layer.

3 Theoretical background of the models

3.1 Logistic Regression (LR) Model

Logistic regression is extensively used to analyze binary variables (Das et al. 2010; Dai et al. 2001; Chen et al. 2017), and LR is used to evaluate the relationship between dependent and independent variables. In this study, a standard LR model is used to evaluate the relationship between the conditioning factors, an independent variable (predictor values) of this study, and flood occurrence, a dependent variable.

Based on these conditioning factors (predictor values), the absence or presence of characteristics will be predicted by the maximum likelihood method (Xu et al. 2013). Results of LR will be evaluated based on the probability of the dependent variable as they are constrained to fall more in the category of 0 or 1 (Shahabi et al. 2015). The higher the coefficient, the more impact it has on the flood occurrence. The following equations is used to derive the probability for flood from the LR coefficient.

$$P = e^{z}/1+e^{z} \quad (11)$$

$$Z=b_0+b_1X_1+b_2X_2+\ldots..+b_nX_n \quad (12)$$

Where P represents the probability of flood, Z is the linear combination or linear logistic model, b_0 represents the intercept of the model, n shows the number of flood conditioning factors, b tells the weights of each condition factor, and X represents the flood influencing factors such as elevation.

3.2 Support Vector Machine (SVM)

Support Vector Machine is a widely popular model of machine learning. A hyper-plane has been generated using the training datasets when converting from the actual support vector machine datasets to high dimensional feature space has occurred (Choi et al. 2020). The functional performance of the model depends mainly on the usage of a suitable kernel. It is a common observation that, like other neural networks, SVM also faces the problem of over-fitting and under-fitting. The theory of SVM has been documented by several researchers Kecman (2001) in the

high-dimensional feature space, the hypothetical space of support vector machine is limited to the linear function. These hypotheses are then trained with a learning algorithm based on optimisation theory. These algorithms apply statistically extract learning theory. In this way, optimizing the machine to generalize is achieved by the fine-tuning of the learning machine. An appropriate choice of kernel allows the non-separable data in the original input space to become separable in feature space. The kernel can be defined as the function that directly calculates the inner product from the input points.

3.3 Multi-Layer Perceptron (MLP)

Artificial Neural Networks act as the black box to impersonate the human brain's structure and function (Kia et al. 2012). MLP has high stability while having a smaller structure than most other neural networks (Wang et al.2021), thus selected for the present study. The structure of an MLP consists of an input layer, hidden layer, and output layer. Flood conditioning factors are associated with the hidden layer, the output layer is the flood or non-flood, and the main aim of the hidden layer is to convert input into output layer (Fig. 3). The weight adjustments among neurons respond to errors between target output values and actual output values. The training of the neural nets in MLP consists of basically two major steps; 1. Use the forward propagation for the input data (conditioning factors) through the hidden layer to get the output, and then to estimate the difference, this output is compared with pre-set values 2. Adjust the weights of the connection so that the best results can be obtained with the minimum difference. The classification function of MLP for the present study can be written as (Pham et al. 2017).

$$t_i = f(x_i) \quad (13)$$

xi is the ith the vector of flood conditioning factor, whereas ti, i=0 is for non-flood pixels and 1 for flood pixels, and f (xi) is a hidden function optimized by adjustable network weights for the given architecture during the training process.

The error which occurs during the training input pattern is equal to the difference between the network output *ok* and target out dk and can be expressed by the following equation.

$$e_k = (d_k - o_k) \quad (14)$$

Weights adjustments among layers can reduce the error propagating from the output back through a neural network (Lee et al. 2003). Adjustments of weights will be done by the following equation

$$\Delta wij(n+1) = \eta(\delta j * oi) + \alpha \Delta wij(n) \quad (15)$$

where $\Delta wij\ (n+1)$ and $\Delta wij\ (n)$ are the weight changes of epochs $(n+1)$ and (n), respectively, η is the learning rate, $\delta$is the rate of change in the error, and α is the momentum coefficient.

3.4 Stacking classifier for flood susceptibility mapping

Stacking classifier is the global classifier that combines the output of different classifiers. In the stacking classifier, the final prediction is made using two steps. In the first step, base learning is used to predict the values from the dataset to feed as the input for the second-stage learner. The stacking classifier's overall performance depends on the selection of models Dou (2020). The meta-model is often simple, so it can provide the smooth interpretation of predictions made by the base classifiers. Therefore, it is recommended that linear models should be used as meta classifier for instant linear regression for regression task, as it predicts numeric value and logistic regression for classification tasks that predicts class labels. The powerful and complex models are usually used as base classifiers pourghasemi (2017); therefore, the SVM and MLP are used as base classifiers in this study. SVM and MLP proved to be an excellent combination as ensemble models in previous research as well (chen, 2017; pourghasemi 2017; Hu, 2020). In this study, Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) are used as base classifiers while Logistic Regression (LR) is used as a meta-classifier Fig. 4

3.5 Comparison and evaluation of the models

All models are compared based on their accuracy, confusion matrix (cross table), sensitivity which is true positive (TP) or recall, specificity (True Negative), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Area Under Receiver Operating Curve (AUROC) for both training and validation dataset. The confusion matrix shows the four types of results; True Positive, False Positive, True Negative, and True Positive. Based on these possible four results, the precision and recall are formulated as (Onan 2015).

$$Precision = \frac{TP}{TP+FP} \quad (16)$$

$$sensitivity = \frac{TP}{TP + FN} \quad (17)$$

$$specificity = \frac{TN}{TN + FP} \quad (18)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

Moreover, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are used for flood mapping (Tien Bui et al. 2016b; Kia et al. 2012). MAE are used for validation purposes, while RMSE is sensitive to the large dataset Chai and Draxler (2014)

$$RMSE = \sqrt{1/n}\sum_{i=1}^{n}\left(X_{predicted} - X_{actual}\right)^2 \quad (20)$$

$$MAE = \sqrt{1/n}\ \sum_{i=1}^{n} |\mathrm{X_{predicted}} - \mathrm{X_{actual}}| \quad (21)$$

While n represents the total number of samples in training or validation datasets, $X_{predicted}$ is the predicted values of training or validation datasets, and Xa_{ctual} is the actual values we got as output values from the model.

Receiver Operating Characteristic (ROC) curve is another way of finding the quality and predictive power of probabilistic models (Pham et al. 2017; Shahabi et al. 2015). Graphically, the sensitivity is plotted on the x-axis while the specificity rate is plotted on the y-axis (Gorsevski et al.2006). The quantitative index, AUROC, measures the general performance of the model (Pham et al. 2017). The more the AUROC, the better performance of the model is it ranges from 0.5, which is an inaccurate model, to 1, which is a perfect model (Tien Bui et al. 2016a) AUROC can be calculated as:

$$AUROC = \sum TP + \sum TN / P + N \quad (22)$$

Fig. 5 represents basic steps used for this study

3.6 Sensitivity analysis

In this study, the sensitivity of flood conditioning factors was analysed by the jackknife test Park No-Wook (2015). It is believed to have a high capability to deal with a wide range of practical problems (Bandos et al. 2017). This test is done using the percentage of relative decrease of the AUC (write full name) to determine the contribution of factors following equation Park No-Wook 2015):

$$PRDi = 100\ \times = \frac{[\mathrm{AUC_{all}} - \mathrm{AUC_i}]}{\mathrm{AUC_{all}}} \quad (23)$$

In this equation AUC_{all} is the value of AUC calculated from the prediction by all factors. AUC_i and PRD_i are the AUC values and percentage of relative decrease of AUC respectively, where the ith factor removed from the process of prediction

4 Results

4.1 Data Preparation

In this study, a total of 12 influencing factors and 652 flood points were selected (based on the flood of 2020) to generate the flood susceptibility map of Karachi, Pakistan. As evaluating flood-prone areas is based on the application of binary models, 652 points of non-flood points were used for the analysis. These non-flood locations were selected randomly at relatively high elevation areas which does not faced floods in past. The data was divided into 70% for training and 30% for testing models. This division is essential for the construction of a proper database.

4.2 Selection of conditioning factor

Based on VIF values, it is found that no factor had a VIF value greater than 4. Therefore no factor had the problem of multicollinearity. The highest VIF is shown by rainfall (3.9), while the lowest is shown by NDVI (1.14), but overall all the values showed negligible collinearity (Table 1). This result shows that all factors are independent of each other.

Table 2: prediction power of 12 conditioning factors

Variables	**VIF**
Curvature	1.360
Elevation	3.265
Rainfall	3.905
Slope	3.265
SPI	3.206
Stream Density	1.922
Twi	3.046
LULC	2.195
Soil	3.458
NDVI	1.149
TPI	1.70
Distance from river	1.562

4.3 Multicollinearity of effective Factors

To access the effect of conditioning factors on the flood. The predictive power of all factors was analyzed using Information Gain Ratio (IGR) method on the training dataset. Multicollinearity test helps in eliminating the unwanted factors or the factors whose presence create noise or are unproductive for the data (Pham et al.2016). Results showed that some factors had a strong influence on the study, while some showed zero contribution.

LULC showed the highest value of IGR (0.55), which means LULC has the highest impact on flood occurrence, followed by elevation and rainfall (0.5). While 'distance from the river' is removed because it showed zero contribution in flood, as shown in Fig. 6, similar results were obtained by Bui et al.(2015).

4.4 Spatial Relationship between Flood Probability and FCFs

FR of each class shows its contribution in the probability of flood probability. The highest FR values among slopes (100%) lie in the class 0-2. Similarly, all the flood pixels lie in the class range -10m to 56m in case of elevation, which showed that low lying regions of the study area is more prone to flood, and in the case of lithology, all flood points fell in the class of Paleogene sedimentary rocks; all other classes of these factors showed negligible contribution towards flood (Table 3). Other factors have the distribution of flood pixels in more than two classes. This assessment gave us an absolute understanding of the contribution of each class of all the influencing factors in the flood occurrence in the study area Table 3

Table 3: Spatial relationship between conditioning factors and floods by frequency ratio model.

Factor	class	Pixel in Domain		Flood Pixel		Frequency Ratio
		No	%	No	%	
Slope	0 - 2.0183	4013495	97.44	288900	100	1.0
	2.0184 - 5.5504	88322	2.144	0	0	0
	5.5505 - 12.11	14741	0.35	0	0	0
	12.111 - 23.463	2014	0.04	0	0	0
	23.464 - 64.334	210	0.00	0	0	0
curvature	0.5485 - 20.159	41	0.00	33300	11.52	0.99
	0.014878 - 0.54849	541051	13.13	255600	88.47	0
	-0.25192 - 0.014877	3577625	86.86	0	0	0
	-0.65213 - -0.25193	63	0.00	0	0	0
	-13.859 - -0.65214	2	4.85	0	0	0
Soil	Calcaric regosol	1332732	32.35	280000	89.14	0.98
	Lithosols	2786049	67.64	8900	10.85	0.01
Lithology	Paleogene sedimentary rocks	3688904	89.56	243000	84.11	0.84
	Neogene Sedimentary rock	94038	2.28	45900	15.88	0.15
	Quatemary Sediments	93538	2.27	0	0	0
	water	93804	2.27	0	0	0
		148498	3.60	0	0	0
Land Use	Water	69758	1.72	900	0.31	0.03
	Settlement	722266	17.86	258300	89.40	0.87
	Barren Land	3001990	74.27	22500	7.78	0.01
	Vegetation	247940	6.13	7200	2.49	0.07
Stream Density	0 - 20.603	1410838	34.92	0	0	0
	20.604 - 48.338	1497870	37.08	184500	64.26	0.64
	48.339 - 77.657	894851	22.15	81000	28.21	0.28

	77.658 - 114.11 114.12 - 202.07	205764 29863	5.09 0.73	17100 4500	5.95 1.56	0.05 0.01
TWI	2.8124 - 6.9503 6.9504 - 8.5417 8.5418 - 10.451 10.452 - 13.157 13.158 - 23.104	1317735 2402148 360033 36486 2380	31.99 58.32 8.74 0.88 0.05	76500 192600 18900 900 0	26.47 66.66 6.54 0.31	0.26 0.37 0.24 0.11 0
SPI	9.2792 - 19.246 6.843 - 9.2791 5.2188 - 6.8429 3.8899 - 5.2187 0.41993 - 3.8898	1552721 2338970 208299 18072 720	37.69 56.78 5.05 0.43 0.01	147600 133200 7200 900 0	51.09 46.10 2.49 0.31	0.40 0.24 0.14 0.21 0
Distance from River (Km)	0-0.5 0.5-1 1-1.5 1.5-2 2.5-3 3-3.5 3.5-4 >4	803109 671489 535889 400699 288449 205989 132179 1080979	19.49 16.30 13 9.7 7 5 3.20 26.24	44000 57500 40300 41500 30500 75100 0 0	15.23 19.90 13.94 14.36 10.55 25.99 0 0	0.06 0.10 0.09 0.13 0.13 0.46 0 0
Rainfall (mm/)	15.083 - 15.509 15.51 - 15.962 15.963 - 16.424 16.425 - 16.796 16.797 - 17.394	363464 320593 690938 1975558 768228	8.82 7.78 16.77 47.96 18.65	63225 73225 65225 87225 0	21.88 25.34 22.57 30.19 0	0.32 0.42 0.17 0.08 0
elevation	-10 – 59 59.001 – 122 122.01 – 198 198.01 – 304 304.01 - 501	2210702 1257046 389341 196127 65566	53.67 30.51 9.45 4.76 1.59	288900 0 0 0 0	100 0 0 0 0	1 0 0 0 0
NDVI	-0.5777 - -0.2742 -0.2741 - -0.197 -0.1969 - -0.1651 -0.165 - -0.1305 -0.1304 - 0.1012	2463 31549 1680915 2313539 7618	0.06 0.78 41.64 57.32 0.18	1800 163800 118800 900 0	0.63 57.41 41.64 0.31 0	0.12 0.87 0.01 0.01 0

4.5 Stacking ensemble of the models

Although single machine learning models gives decisive predictions, integrating these machine learning algorithms or with some other statistical techniques gives better performance (Hong et al. 2018; Arabameri et al. 2019). Ensembles methods gave much better and precise outcomes for the flood susceptibility analysis in the past Pham et al. (2017). Therefore, an ensemble model was constructed for flood susceptibility of the study area using the stacking ensemble technique.

For the stacking of models, Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) were trained as a base classifier on the training dataset, whereas Logistic Regression was trained on the outcome of the base classifier. Different optimizations of models were applied to get the best result and avoid overfitting.

To obtain the best performance from SVM, SVM was trained by using four types of kernels, namely polynomial kernel (PL), Radial based function (RBF), Linear Kernel (LN), and sigmoid kernel (SIG). Except for RBF, all other kernels showed the problem of overfitting or under-fitting, 'RBF' kernel showed perfect performance with a root mean square value of 0.18. Some previous research also tried different kernels (Tien Bui et al. 2012; Tehrany et al. 2015), and the results showed that the radial base kernel gave the best performance. Numerous studies confirm that RBF overweighs other kernel functions in the case of flood susceptibility (Chen et al. 2020; Yang and Cervone 2019). Similarly, MLP was executed several times with the different number of neurons and hidden layers, and the final choice was based on the highest accuracy and lowest RMSE. After many trials and errors, we found that MLP in our case gave the best performance with 20 hidden layers and 30 output neurons. (Table 4)

Table 4: Results of hyper parameter tuning of models

Algorithm	Parameters
LR	C=1.0, class_weight=None, dual=False, fit_intercept=True, intercept_scaling=1, l1_ratio=None, max_iter=100, multi_class='auto', n_jobs=None, penalty='l2', random_state=None, solver='lbfgs', tol=0.0001, verbose=0, warm_start=False
SVM	C=1.0, break_ties=False, cache_size=200, class_weight=None, coef0=0.0, decision_function_shape='ovr', degree=3, gamma='scale', kernel='rbf', max_iter=-1, probability=False, random_state=None, shrinking=True, tol=0.001, verbose=False
MLP	activation='relu', alpha=0.0001, batch_size='auto', beta_1=0.9, beta_2=0.999, early_stopping=False, epsilon=1e-08, hidden_layer_sizes=(11, 20, 30), learning_rate='constant', learning_rate_init=0.001, max_fun=15000, max_iter=100, momentum=0.9, n_iter_no_change=10, nesterovs_momentum=True, power_t=0.5, random_state=None, shuffle=True, solver='adam', tol=0.0001, validation_fraction=0.1, verbose=False, warm_start=False

4.6 Model validation and comparison

After training the models on their best optimization and hyper-parameter tuning, their performances were compared by the difference in their accuracy, precision, recall, Etc. The evaluation was performed both on the training and validation datasets because the training dataset represents the model's fitting skill and the validation dataset indicates the model's generalizations ability. In the case of the validation dataset highest accuracy was 98 obtained by ensemble model, SVM and MLP showed an accuracy of 96%, while LR showed 93.9% accuracy. For the training

samples, the ensemble model ranked first with an accuracy of 99%, followed by MLP (98%), SVM (97%), and LR (94.8%). In the case of precision, the ensemble model showed 98% precision for the validation dataset and 99% for the Training dataset, which showed that this model gave exceptionally precise locations of flood susceptibility.

Table 5: Model performance for training dataset for new ensemble model and other benchmark models.

Training data	**LR**	**SVM**	**MLP**	**LR-SVM-MLP**
True Negative(TN)	446	469	468	457
True positive(TP)	444	446	445	470
False Positive(FP)	25	2	3	1
False Negative(FN)	32	30	31	19
sensitivity	96.01	96	93.4	96.1
Specificity	93.68	99	99.3	99.7
precision	0.95	0.94	0.98	0.98
recall	0.96	0.96	0.97	0.07
Accuracy	94.8	97	98	99
Root Mean Square Error (RMSE)	0.24	0.183	0.189	0.14
Mean Absolute Error (MAE)	0.060	0.033	0.035	0.021

Table 6: Model performance for training dataset for new ensemble model and other benchmark models.

Validation data	**LR**	**SVM**	**MLP**	**LR-SVM-MLP**
True Negative(TN)	193	205	204	219
True Positive (TP)	193	193	198	171
False Positive(FP)	13	1	2	1
False Negative(FN)	8	8	3	16
sensitivity	93.2	96	98.5	91.4
Specificity	94.6	99.5	99	99.8
precision	0.94	0.92	0.96	0.99
recall	0.93	0.93	0.98	0.95
Accuracy	93.8	96.6	96	98
Root Mean Square Error (RMSE)	0.22	0.148	0.110	0.198
Mean Absolute Error (MAE)	0.051	0.022	0.012	0.039

Receiver operating characteristics curve was calculated on both training and validation datasets. In both training and validation datasets, it is observed that the ensemble model obtained the highest performance. The results of sensitivity and specificity of the ensemble model showed that the number of correctly classified flood pixels was 99% for the training dataset and 99.8% for validation datasets. Similarly, 97.6 % and 99% were correctly classified as non-flood pixels for training and validation datasets, respectively, by ensemble model (Table 5, Table 6, and Figure 7). It is necessary to categorize the final prediction map of flood susceptibility into different classes to easily visualize flood probability in the study area (Pham et al. 2017; Tehrany et al. 2014b).

Final flood susceptibility maps were constructed in five classes based on quantile classification; for example, classes of ensemble model were categorized in very low or non-susceptibility category (0.0- 0.001), low (0.001-0.221), moderate (0.221-0.862), high (0.862-0.992), and very high (0.992-0.999). Table 7 represents the percentage of the area that lies in each class. The ensemble model showed that almost 16% of the region comes under the category of very high and 23% under the class of high susceptibility; both of these regions collectively made 39% of the high susceptibility zone.

Table 7: percentage of area in each zone

Class	LR (Km^2)	%	SVM (Km^2)	%	MLP (Km^2)	%	Ensemble (Km^2)	%
Very low	630258.68	17	742548.2	20.03	674299.39	18.19	722960.96	19.50
Low	852427.64	22.9	744375.7	20.08	799383.23	21.56	767380.096	20.70
Medium	802612.35	21.65	747083.5	20.15	760295.55	20.51	745017.92	20.09
High	716998.33	19.3	788454.3	21.27	742257.6	20.02	850915.008	22.95
Very High	704484.73	19	684320.1	18.46	730546.04	19.70	620507.84	16.73

4.7 Sensitivity analysis result

For flood susceptibility, choosing a suitable conditioning factor is very critical (Kourgialas and Karatzas, 2012). Therefore, jackknife test was used in this study to calculate the sensitivity of eleven conditioning factors. It is a simple method and reduces the bias in the estimator, which could cause if we apply a complex method. Fig 9 indicate the relative importance of each factor. Land Use was the most critical factor in this study; it had the highest contribution in all model's

445 predictions, followed by elevation and rainfall. Overall, LULC, Rainfall, Elevation and NDVI are
446 the most important contributor in all the models.

447 # 5 Discussion
448
449 Identifying flood-prone areas and their zonation is crucial for reducing damage caused by flooding.
450 There are several methodologies suggested by different researchers for the development of flood
451 susceptibility maps around the world (Tehrany et al. 2015b; Chen et al. 2020). Remote Sensing and
452 GIS applications provide powerful tools for predicting and analyzing many multidimensional
453 incidents like flooding, which are influenced and controlled by multiple factors (Arabameri et al.
454 2020; Arabameri et al. 2019). But machine learning models, primarily ensemble of single
455 algorithms with each other or with some other statistical techniques, gives a more remarkable
456 performance, especially in the case of flood susceptibility analysis (Hong et al. 2018; Arabameri
457 et al. 2019; Pham et al. 2017; Tien Bui et al. 2012; Islam et al. 2020; Wang et al. 2021)
458 For better visualisation of susceptibility areas, the final map was classified into different classes
459 (Pham et al. 2017; Tehrany et al. 2014b). There are different techniques of reclassification, such as
460 standard deviation, geometric intervals, equal intervals, and quantiles (Francis et al. 2015). All
461 these methods show different results, and therefore it is crucial to analyse which methods best suit
462 the particular study (Tehrany et al. 2015a). The equal intervals are suitable for the data with
463 standard data distribution; natural breaks better suit certain jumps (Tehrany et al. 2015b). The
464 geometric intervals are suitable for not normally distributed data, such as continuous data, because
465 it reduces variance (Russell et al. 2012). However, the literature shows that the quantile method
466 performs best in the case of susceptibility maps (Tehrany et al. 2015b; Chapi et al. 2017). Therefore
467 final maps were generated using the quantile method of classification.
468 In all the maps, it was observed that the high susceptibility areas were located on the southern side,
469 which is the low elevation zone. Also, the results of sensitivity analysis showed that Landuse,
470 elevation, rainfall and NDVI are major contributing factors. Similar results were obtained by some
471 previous research (Arabameri et al., 2020; Santos et al., 2019; Zhao et al., 2018). Most of the
472 settlement area of Karachi is in the southern part Raza et al. (2019), due to which there is very few
473 barren and green areas which result in less seepage of water. This area also receives the highest
474 rainfall compared to the other areas. Along with intense settlements, which reduces seepage, this
475 area faces the problem of the poor drainage system, clogged main holes, and nullah's heaped with

garbage, due to which low laying areas of Karachi is always at the risk of flooding Arif Hasan (2020)

In comparison, the Northern sides, which are mountainous regions of Karachi, fell in the lowest susceptibility zone. The comparison of susceptibility maps showed that the ensemble model calculated the susceptibility areas more precisely and accurately as compared to single machine learning models (Chapi et al. 2017; Arabameri et al. 2020; Prasad et al. 2021)

The precise and accurate framework for flood modelling is a fundamental task because unpredicted floods are responsible for damage in terms of human and infrastructure damage and huge economic loss (Tehrany et al. 2013; Youssef et al. 2016). Therefore, the precision of the proposed model was calculated on both training and validation datasets. Many previous research has shown that statistical metrics, for example, False Positive Rate (FP), False Negative (FN), True Positive (TP) and True Negative (TN), Receiver Operating Curve (ROC), is the best parameters for performance analysis (Althuwaynee et al. 2014; Costache et al. 2020; Wang et al. 2021; Chapi et al. 2017)

The comparison showed that the accuracy of the ensemble model was 99%, which is almost 5%, 2%, and 1% more than LR, SVM, and MLP, respectively. In the case of the ROC curve ensemble model, it gave way better results than single models, which shows that the total number of true positive and true negative cases is more in the ensemble model compared to individual models. This increased and better performance of ensemble and hybrid models are confirmed by using these statistical measures in many previous studies (Chapi et al. 2017; Arabameri et al. 2020; Prasad et al. 2021).

Also, this ensemble model gave a much better performance than previously applied ensemble techniques; in the case of chapi et al 2017 bagging ensemble was applied with the logistic model tree, which gave an accuracy of 95%. Similarly, the ensemble approach using a weighted average by Choubin et al 2019 showed an AUC of 0.91, Islam et al 2020., used dagging ensemble, which showed RMSE 0.189.

6 Conclusion

For flood susceptibility mapping, an ensemble model (LR-SVM-MLP) was used, which was formed by a stacking ensemble of LR, SVM, and MLP. This model was then compared with its benchmarked individual models, which showed that the ensemble model performed best and has higher reliability than all other models.

This research contributes several ways; for example, there is no previous record of any research done on this region for flood forecasting using machine learning models. No previous research has been conducted to evaluate the conditioning or influencing factor contributing to this region's flood, although this region suffers floods almost every year or every second year. This research gives some of the best models which can be used individually to predict flood in this region, but the proposed ensemble model outweigh all other models, and it highlights high and very high flood-prone regions in the watershed.

The main limitation of this study is that it did not consider the sanitary condition of Karachi, as blockage of nullah's, is one of the significant reasons for Karachi's flood. As future work, this research can be enhanced if these models are combined with 2D and 3D modelling systems for better and real-world visualization. The model's effectiveness can be strengthened if it combines with other models that can tell lag time as lag time is one of the most critical factors in the case of flash flood monitoring and prediction.

Field-based surveys are costly and time-consuming work, while the produced maps have fine details of areas; therefore, they can help management, policymakers, government, and other relevant authorities to provide better flood prevention measures and to mitigate any damage that can be caused by the flood.

7 Declarations

Conflicts of interest/Competing interests: No conflict of interest

Availability of data and material: Not Applicable

Code availability: Not Applicable

Ethics approval Not Applicable

Consent to participate Not Applicable

Consent for publication Not Applicable

Funding.

This work was funded by the National Key Research and Development Program (2018YFC1506506), the Frontier Project of the Applied Foundation of Wuhan (2019020701011502), the Key Research and Development Program of Jiangxi Province (20201BBG71002), and the LIESMARS Special Research Funding.

References

Ahmadlou M, Karimi M, Alizadeh S, Shirzadi A, Parvinnejhad D, Shahabi H, Panahi M (2019) Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). Geocarto Int. 34:1252–1272. https://doi.org/10.1080/10106049.2018.1474276

Althuwaynee OF, Pradhan B, Park HJ, Lee JH (2014) A novel ensemble decision tree-based CHi-squared Automatic Interaction Detection (CHAID) and multivariate logistic regression models in landslide susceptibility mapping. Landslides 11:1063-1078.

Arabameri A, Chen W, Loche M, Zhao X, Li Y, Lombardo L, Cerda A, Pradhan B, Bui D T (2020a) Comparison of machine learning models for gully erosion susceptibility mapping. Geosci Front 11:1609–1620. https://doi.org/10.1016/j.gsf.2019.11.009

Arabameri A, Rezaei K, Cerdà A, Conoscenti C, Kalantari Z (2019) A comparison of statistical methods and multi-criteria decision making to map flood hazard susceptibility in Northern Iran. Sci Total Environ 660:443–458. https://doi.org/10.1016/j.scitotenv.2019.01.021

Arabameri A, Saha S, Mukherjee K, Blaschke T, Chen W, Ngo P T.T, Band S S (2020b) Modeling spatial flood using novel ensemble artificial intelligence approaches in northern Iran. Remote Sens 12:1–30. https://doi.org/10.3390/rs12203423

Arif Hasan (2020) Urban flooding: the case of Karachi. International Institute for Environment and Development. https://www.iied.org/urban-flooding-case-karachi. Accessed 26 December 2021

ARM Towfiqul Islam, S Talukdar, S Mahato, S Kundu, KU Eibek, QB Pham, A Kuriqi, NTT Linh (2020) Flood susceptibility modelling using advanced ensemble machine learning models. https://doi.org/10.1016/j.gsf.2020.09.006

Bandos AI. Guo B. Gur D. (2017). Jackknife variance of the partial area under the empirical receiver operating characteristic curve. Stat. Methods Med. Res. 26 (2), 528–541

Beven K J, Kirkby M J (1979) A physically based, variable contributing area model of basin hydrology. Hydrol. Sci. Bull 24:43–69. https://doi.org/10.1080/02626667909491834

Bui D T, Pradhan B, Revhaug I, Nguyen D B, Pham H V, Bui Q N (2015) A novel hybrid evidentiabelief function-based fuzzy logic model in spatial prediction of rainfall-induced shallow landslides in the Lang Son city area (Vietnam). Geomatics Nat Hazards Risk 6:243–

271. https://doi.org/10.1080/19475705.2013.843206

Bui DT, Ngo P-TT, Pham TD, Jaafari A, Minh NQ, Hoa PV, Samui P (2019) A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. Catena 179:184–196. https://doi.org/10.1016/j.catena.2019.04.009

Çelik H E, Coskun G, Cigizoglu H K, Ağiralioğlu N, Aydin A, Esin A I (2012) The analysis of 2004 flood on Kozdere Stream in Istanbul. Nat Hazards 63:461–477. https://doi.org/10.1007/s11069-012-0165-x

Chai T, Draxler R R (2014) Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. Geosci Model Dev 7:1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Chapi K, Singh V P, Shirzadi A, Shahabi H, Bui D.T, Pham B.T, Khosravi K (2017) A novel hybrid artificial intelligence approach for flood susceptibility assessment. Environ. Model Softw 95: 229–245. https://doi.org/10.1016/j.envsoft.2017.06.012.

Chen FW, and Liu CW (2012) Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. Paddy and Water Environ. 10(3):209-222. https://doi.org/10.1007/s10333-012-0319-1

Chen W, Pourghasemi HR, Zhao Z (2017) A GIS-based comparative study of Dempster-Shafer, logistic regression and artificial neural network models for landslide susceptibility mapping. Geocarto Int 32:367–385. https://doi.org/10.1080/10106049.2016.1140824

Chen W, Zhao X, Tsangaratos P, Shahabi H, Ilia I, Xue W, Wang X, Ahmad B Bin (2020) Evaluating the usage of tree-based ensemble methods in groundwater spring potential mapping. J Hydrol 583:124602. https://doi.org/10.1016/j.jhydrol.2020.124602

Choi IH, Koo J, Bin, Woo JW, Son, JA Bae, DY Yoon, YG. Oh TK (2020) Damage evaluation of porcelain insulators with 154 kV transmission lines by various support vector machine (SVM) and ensemble methods using frequency response data. Appl Sci 10. https://doi.org/10.3390/app10010084

Choubin B, Moradi E, Golshan M, Adamowski J, Sajedi-Hosseini F, Mosavi A (2019) An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. Sci. Total Environ 651:2087–2096. https://doi.org/10.1016/j.scitotenv.2018.10.064

Costache R, Pham QB, Avand M, Linh NTT, Vojtek M, Vojtekov´a J. Lee S, Khoi DN, Nhi PTT,

597 Dung TD (2020) Novel hybrid models between bivariate statistics, artificial neural networks
598 and boosting algorithms for flood susceptibility assessment. J. Environ. Manag. 265: 110485.
599 Dai FC, Lee CF, Li J, Xu Z W (2001) Assessment of landslide susceptibility on the natural
600 terrain of Lantau Island, Hong Kong. Environ Geol 40: 381–391.
601 https://doi.org/10.1007/s002540000163
602 Das I, Sahoo S, van Westen C, Stein A, Hack R (2010) Landslide susceptibility assessment using
603 logistic regression and its comparison with a rock mass classification system, along a road
604 section in the northern Himalayas (India). Geomorphology 114: 627–637.
605 https://doi.org/10.1016/j.geomorph.2009.09.023
606 DeLeo JM (1993) Receiver operating characteristic laboratory (ROCLAB): Software for
607 developing decision strategies that account for uncertainty. Proc. - 2nd Int. Symp. Uncertain.
608 Model. Anal. ISUMA 1993 318–325. https://doi.org/10.1109/ISUMA.1993.366750
609 Dodangeh E, Choubin B, Eigdir AN, Nabipour N, Panahi M, Shamshirband S, Mosavi A (2020)
610 Integrated machine learning methods with resampling algorithms for flood susceptibility
611 prediction. Sci Total Environ. 705:135983. https://doi.org/10.1016/j.scitotenv.2019.135983
612 Dou J, Yunus AP, Bui, DT et al (2020) Improved landslide assessment using support vector
613 machine with bagging, boosting, and stacking ensemble machine learning framework in a
614 mountainous watershed, Japan. Landslides. 17:641–658. https://doi.org/10.1007/s10346-019-
615 01286-5
616 Ekmekcioğlu Ö, Koc K, & Özger M (2021) Stakeholder perceptions in flood risk assessment: A
617 hybrid fuzzy AHP-TOPSIS approach for Istanbul, Turkey. Int J of Disaster Risk Reduction,
618 60, 102327. https://doi.org/10.1016/j.ijdrr.2021.102327
619 Elmore AJ, Julian JP, Guinn SM, Fitzpatrick MC, (2013) Potential Stream Density in Mid-Atlantic
620 U.S. Watersheds. PLoS One 8:1–15. https://doi.org/10.1371/journal.pone.0074819
621 Fekete A (2009) Validation of a social vulnerability index in context to river-floods in Germany.
622 Nat Hazards Earth Syst Sci 9: 393–403. https://doi.org/10.5194/nhess-9-393-2009
623 Francis J, Tontisirin N, Anantsuksomsri S, Vink J, Zhong V (2015) Alternative Strategies for
624 Mapping ACS Estimates and Error of Estimation 247–273. https://doi.org/10.1007/978-94-
625 017-8990-5_16
626 Gayen A, Pourghasemi HR, Saha S, Keesstra S, Bai S (2019) Gully erosion susceptibility
627 assessment and management of hazard-prone areas in India using different machine learning

algorithms. Sci Total Environ 668:124–138. https://doi.org/10.1016/j.scitotenv.2019.02.436

Ghorbanzadeh O, Blaschke T, Gholamnia K, Meena SR, Tiede D, Aryal J. (2019). Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. Remote Sens-Basel. 11(2):196.

Gorsevski PV, Gessler PE, Foltz RB, Elliot WJ (2006) Spatial prediction of landslide hazard using logistic regression and ROC analysis. Trans GIS 10:395–415. https://doi.org/10.1111/j.1467-9671.2006.01004.x

Hassan MM (2017) Monitoring land use/land cover change, urban growth dynamics and landscape pattern analysis in five fastest urbanized cities in Bangladesh. Remote Sens. Appl Soc Environ 7: 69–83. https://doi.org/10.1016/j.rsase.2017.07.001

Hong H, Liu J, Bui DT, Pradhan B, Acharya TD, Pham BT, Zhu AX, Chen W, Ahmad B, Bin (2018) Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China). Catena 163: 399–413. https://doi.org/10.1016/j.catena.2018.01.005

Hu Q, Zhou Y, Wang, S, & Wang F (2020) Machine learning and fractal theory models for landslide susceptibility mapping: Case study from the Jinsha River Basin. Geomorphology. 351:106975. https://doi.org/10.1016/j.geomorph.2019.106975

Hussan Q, Ali I, Mughen I (2020) 10 more die as mjaor parts of Karachi remain powerless a day after rain record.

Jonkman SN (2005) Global perspectives on loss of human life caused by floods. Nat. Hazards 34:151–175. https://doi.org/10.1007/s11069-004-8891-3

Joy S, Xx L (2004) Application of Remote Sensing in Flood Management with Special Reference to Monsoon Asia: A Review. Nat Hazards 283–301. https://doi.org/10.1023/B:NHAZ.0000037035.65105.95

Kalantar B, Pradhan B, Naghibi SA, Motevalli A, Mansor S (2018) Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). Geomat Nat Hazards Risk. 9:49–69. https://doi.org/10.1080/19475705.2017.1407368

Kazakis N, Kougias I, Patsialis T (2015) Assessment of flood hazard areas at a regional scale using an index-based approach and Analytical Hierarchy Process: Application in Rhodope-Evros region, Greece. Sci Total Environ 538: 555–563.

https://doi.org/10.1016/j.scitotenv.2015.08.055
Kecman V (2001) Learning and Soft Computing: Support Vector Machines, Neural Networks and Fuzzy Logic Models. The MIT Press, Cambridge, MA.
Khosravi K, Pourghasemi HR, Chapi K, Bahri M (2016) Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: a comparison between Shannon's entropy, statistical index, and weighting factor models. Environ. Monit Assess 188: 1–21.
Kia MB, Pirasteh S, Pradhan B, Mahmud AR, Sulaiman WNA, Moradi A (2012) An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. Environ. Earth Sci 67: 251–264. https://doi.org/10.1007/s12665-011-1504-z
Kourgialas NN. Karatzas GP (2012) Flood management and a GIS modeling method to assess flood hazard areas – a case study. Hydrol. Sci. J. 56 (2), 212–225.
Lauría E J, Presutti E, Kapogiannis M, and Kamath A (2018) Stacking Classifiers for Early Detection of Students at Risk. In CSEDU (1). 390-397. DOI: 10.5220/0006781203900397
Lee MJ, Kang JE, Jeon S (2012) Application of frequency ratio model and validation for predictive flooded area susceptibility mapping using GIS. Int Geosci Remote Sens Symp 895–898. https://doi.org/10.1109/IGARSS.2012.6351414
Lee S, Ryu JH, Lee MJ, Won JS (2003) Use of an artificial neural network for analysis of the susceptibility to landslides at Boun, Korea. Environ. Geol 44: 820–833. https://doi.org/10.1007/s00254-003-0825-y
Lu Y, Bookman R, Waldmann N, and Marco S (2020) A 45 kyr laminae record from the Dead Sea: Implications for basin erosion and floods recurrence. Quaternary Science Reviews 229: 106143. https://doi.org/10.1016/j.quascirev.2019.106143
Manandhar Bikram (2010) Flood plain analysis and risk assessment of Lothar Khola 77. https://doi.org/10.13140/2.1.1664.1289
Mangi MY, Yue Z, Kalwar S, Lashari ZA (2020) Comparative analysis of urban development trends of Beijing and Karachi metropolitan areas. Sustain. 12. https://doi.org/10.3390/su12020451
Meles MB, Younger SE, Jackson CR, Du E, Drover D (2020) Wetness index based on landscape position and topography (WILT): Modifying TWI to reflect landscape position. J Environ. Manage 255: 109863. https://doi.org/10.1016/j.jenvman.2019.109863
Meshram SG, Alvandi E, Meshram C et al (2020) Application of SAW and TOPSIS in Prioritizing

690 Watersheds. Water Resour Manage 34:715–732. https://doi.org/10.1007/s11269-019-02470-
691 x
692 Mukerji A, Chatterjee C, Raghuwanshi NS (2009) Flood Forecasting Using ANN, Neuro-Fuzzy,
693 and Neuro-GA Models. J Hydrol Eng 14: 647–652. https://doi.org/10.1061/(asce)he.1943-
694 5584.0000040
695 Ngo P TT, Hoang ND, Pradhan B, Nguyen QK, Tran XT, Nguyen QM, Nguyen VN, Samui P,
696 Tien Bui D (2018) A novel hybrid swarm optimized multilayer neural network for spatial
697 prediction of flash floods in tropical areas using sentinel-1 SAR imagery and geospatial data.
698 Sensors 18: 3704. https://doi.org/10.3390/s18113704
699 Nguyen H, Bui XN, Tran QH, Moayedi H. (2019). Predicting blast-induced peak particle velocity
700 using BGAMs, ANN and SVM: a case study at the Nui Beo open-pit coal mine in Vietnam.
701 Environ Earth Sci. 78 (15):1–14.
702 Odoh SI, Chilaka FC (2012) Climate Change and Conflict in Nigeria : A Theoretical and Empirical
703 Examination of the Worsening Incidence of Conflict between Fulani Herdsmen and Farmers
704 in. Oman Chapter Arab. J Bus Manag Rev 2: 110–124. https://doi.org/10.12816/0002246
705 Onan A (2015) A fuzzy-rough nearest neighbor classifier combined with consistency-based subset
706 evaluation and instance selection for automated diagnosis of breast cancer. Expert Syst Appl
707 42: 6844–6852. https://doi.org/10.1016/j.eswa.2015.05.006
708 Park NW. (2015). Using maximum entropymodeling for landslide susceptibility mapping with
709 multiple geoenvironmental data sets. Environ Earth Sci. 73 (3), 937–949.
710 Pham BT, Pradhan BT, Bui D, Prakash I, Dholakia MB (2016) A comparative study of different
711 machine learning methods for landslide susceptibility assessment: a case study of Uttarakhand
712 area (India). Environ Modell Softw 84:240–250.
713 https://doi.org/10.1016/j.envsoft.2016.07.005
714 Pham BT, Tien Bui D, Prakash I, Dholakia MB (2017) Hybrid integration of Multilayer Perceptron
715 Neural Networks and machine learning ensembles for landslide susceptibility assessment at
716 Himalayan area (India) using GIS. Catena 149: 52–63.
717 https://doi.org/10.1016/j.catena.2016.09.007
718 Poudyal CP, Chang C, Oh HJ, Lee S (2010) Landslide susceptibility maps comparing frequency
719 ratio and artificial neural networks: A case study from the Nepal Himalaya. Environ Earth Sci
720 61: 1049–1064. https://doi.org/10.1007/s12665-009-0426-5

721 Pourghasemi H R, Yousefi S, Kornejady A, & Cerdà A (2017) Performance assessment of
722 individual and ensemble data-mining techniques for gully erosion modeling. Sci of Total
723 Environ. 609:764-775. 10.1016/j.scitotenv.2017.07.198
724 Prasad P, Loveson VJ, Das B, Kotha M (2021) Novel ensemble machine learning models in flood
725 susceptibility mapping, Geocarto International. Taylor & Francis.
726 https://doi.org/10.1080/10106049.2021.1892209
727 Rahmati O, Nazari Samani A, Mahdavi M, Pourghasemi HR, Zeinivand H (2015) Groundwater
728 potential mapping at Kurdistan region of Iran using analytic hierarchy process and GIS. Arab.
729 J Geosci 8: 7059–7071. https://doi.org/10.1007/s12517-014-1668-4
730 Rahmati O, Zeinivand H, Besharat M (2016) Flood hazard zoning in Yasooj region, Iran, using
731 GIS and multi-criteria decision analysis. Geomatics, Nat Hazards Risk 7: 1000–1017.
732 https://doi.org/10.1080/19475705.2015.1045043
733 Raza D, Karim RB, Nasir A, Khan SU, Zubair MH, Amir R (2019) Satellite Based Surveillance
734 of LULC with Deliberation on Urban Land Surface Temperature and Precipitation Pattern
735 Changes of Karachi. Geography & Natural Disasters 9: 1–8. https://doi.org/10.4172/2167-
736 0587.1000237
737 Razavi Termeh SV, Kornejady A, Pourghasemi HR, Keesstra S (2018) Flood susceptibility
738 mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic
739 algorithms. Sci Total Environ 615: 438–451. https://doi.org/10.1016/j.scitotenv.2017.09.262
740 Russell EL, Chiang LH, Braatz RD (2012) Data-driven Methods for Fault Detection and Diagnosis
741 in Chemical Processes. Springer Science & Business Media 1:1430-9491.10.1007/978-1-
742 4471-0409-4
743 Saha S, Paul GC, Pradhan B, Abdul Maulud KN, and Alamri AM (2021) Integrating multilayer
744 perceptron neural nets with hybrid ensemble classifiers for deforestation probability
745 assessment in Eastern India. Geomatics, Natural Hazards and Risk12: 29-62.
746 https://doi.org/10.1080/19475705.2020.1860139
747 Santos PP, Reis E, Pereira S, Santos M (2019) A flood susceptibility model at the national scale
748 based on multicriteria analysis. Sci Total Environ 667:325–337.
749 Schumann GJP, Moller DK (2015) Microwave remote sensing of flood inundation. Phys Chem
750 Earth 83–84, 84–95. https://doi.org/10.1016/j.pce.2015.05.002
751 Sepehri M, Malekinezhad H, Jahanbakhshi F (2020) Integration of interval rough AHP and fuzzy

752 logic for assessment of flood prone areas at the regional scale. Acta Geophys. 68, 477–493.
753 https://doi.org/10.1007/s11600-019-00398-9
754 Shahabi H, Hashim M, Ahmad B, Bin (2015) Remote sensing and GIS-based landslide
755 susceptibility mapping using frequency ratio, logistic regression, and fuzzy logic methods at
756 the central Zab basin, Iran. Environ Earth Sci 73: 8647–8668. https://doi.org/10.1007/s12665-
757 015-4028-0
758 Shahbaz M, Chaudhary AR, Ozturk I, (2017) Does urbanization cause increasing energy demand
759 in Pakistan? Empirical evidence from STIRPAT model. Energy 122: 83–93.
760 https://doi.org/10.1016/j.energy.2017.01.080
761 Srivastava PK, Han D, Rico-Ramirez MA, Islam T (2014) Sensitivity and uncertainty analysis of
762 mesoscale model downscaled hydro-meteorological variables for discharge prediction.
763 Hydrol Process 28:4419–4432. https://doi.org/10.1002/hyp.9946
764 Stefanidis S, Stathis D (2013) Assessment of flood hazard based on natural and anthropogenic
765 factors using analytic hierarchy process (AHP). Nat Hazards 68: 569–585.
766 https://doi.org/10.1007/s11069-013-0639-5
767 Stevaux JC, Macedo H, de A Assine ML, Silva A (2020) Changing fluvial styles and backwater
768 flooding along the Upper Paraguay River plains in the Brazilian Pantanal wetland.
769 Geomorphology 350. https://doi.org/10.1016/j.geomorph.2019.106906
770 Tariq S Khan M A and Alamgir A (2016) Physico-chemical profile of Malir river and chinna creek.
771 Procedia Environmental Sciences 34:514-524. https://doi.org/10.1016/j.proenv.2016.04.045
772 Tehrany MS, Lee MJ, Pradhan B, Jebur MN, Lee S (2014a) Flood susceptibility mapping using
773 integrated bivariate and multivariate statistical models. Environ. Earth Sci 72: 4001–4015.
774 https://doi.org/10.1007/s12665-014-3289-3
775 Tehrany MS, Pradhan B, Jebur MN (2013) Spatial prediction of flood susceptible areas using rule
776 based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in
777 GIS. J Hydrol 504:69–79. https://doi.org/10.1016/j.jhydrol.2013.09.034
778 Tehrany MS, Pradhan B, Jebur MN (2014b) Flood susceptibility mapping using a novel ensemble
779 weights-of-evidence and support vector machine models in GIS. J Hydrol 512:332–343.
780 https://doi.org/10.1016/j.jhydrol.2014.03.008
781 Tehrany MS, Pradhan B, Jebur MN (2015a) Flood susceptibility analysis and its verification using
782 a novel ensemble support vector machine and frequency ratio method. Stoch. Environ. Res

783 Risk Assess 29: 1149–1165. https://doi.org/10.1007/s00477-015-1021-9
784 Tehrany MS, Pradhan B, Mansor S, Ahmad N (2015b) Flood susceptibility assessment using GIS-
785 based support vector machine model with different kernel types. Catena 125: 91–101.
786 https://doi.org/10.1016/j.catena.2014.10.017
787 Tien Bui D, Pradhan B, Lofman O, Revhaug I (2012) Landslide susceptibility assessment in
788 Vietnam using support vector machines, decision tree, and nave bayes models. Math Probl
789 Eng 2012. https://doi.org/10.1155/2012/974638
790 Tien Bui D, Pradhan B, Nampak H, Bui QT, Tran QA, Nguyen QP (2016a) Hybrid artificial
791 intelligence approach based on neural fuzzy inference model and metaheuristic optimization
792 for flood susceptibilitgy modeling in a high-frequency tropical cyclone area using GIS. J
793 Hydrol 540: 317–330. https://doi.org/10.1016/j.jhydrol.2016.06.027
794 Tien Bui D, Tuan TA, Klempe H, Pradhan B, Revhaug I (2016b) Spatial prediction models for
795 shallow landslide hazards: a comparative assessment of the efficacy of support vector
796 machines, artificial neural networks, kernel logistic regression, and logistic model tree.
797 Landslides 13: 361–378. https://doi.org/10.1007/s10346-015-0557-6
798 Tucker CJ, Sellers PJ (1986) Satellite remote sensing of primary production. Int J Remote Sens 7:
799 1395–1416. https://doi.org/10.1080/01431168608948944
800 Wang Y, Fang Z, Hong H, Costache R, Tang X (2021) Flood susceptibility mapping by integrating
801 frequency ratio and index of entropy with multilayer perceptron and classification and
802 regression tree. J Environ Manage 289: 112449.
803 https://doi.org/10.1016/j.jenvman.2021.112449
804 Wang Y, Hong H, Chen W, Li S, Panahi M, Khosravi K, Shirzadi A, Shahabi H, Panahi S,
805 Costache R (2019) Flood susceptibility mapping in Dingnan County (China) using adaptive
806 neuro-fuzzy inference system with biogeography based optimization and imperialistic
807 competitive algorithm. Environ Manag 247: 712–729.
808 https://doi.org/10.1016/j.jhydrol.2016.06.027
809 Xu C, Xu X, Dai F, Wu Z, He H, Shi F, Wu X, Xu S (2013) Application of an incomplete
810 landslide inventory, logistic regression model and its validation for landslide susceptibility
811 mapping related to the May 12, 2008 Wenchuan earthquake of China. Nat Hazards 68: 883–
812 900. https://doi.org/10.1007/s11069-013-0661-7
813 Yang L, Cervone G (2019) Analysis of remote sensing imagery for disaster assessment using deep

learning: a case study of flooding event. Soft Comput 13393–13408. https://doi.org/10.1007/s00500-019-03878-8

Yilmaz I (2010) Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. Environ Earth Sci. 61(4), 821–836. https://doi.org/10.1016/j.scitotenv.2018.10.064

Youssef AM Pradhan B, Sefry SA (2016) Flash flood susceptibility assessment in Jeddah city (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models. Environ Earth Sci 75: 1–16. https://doi.org/10.1007/s12665-015-4830-8

Youssef AM, Pourghasemi HR, Pourtaghi ZS, Al-Katheeri MM (2016) Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. Landslides 13:839–856. https://doi.org/10.1007/s10346-015-0614-1

Youssef AM, Pradhan B, Hassan AM (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. https://doi.org/10.1007/s12665-010-0551-1

Zhao G, Pang B, Xu Z, Yue J, Tu T (2018) Mapping flood susceptibility in mountainous areas on a national scale in China. Sci Total Environ 615: 1133–1142. https://doi.org/10.1016/j.scitotenv.2017.10.037

Figures

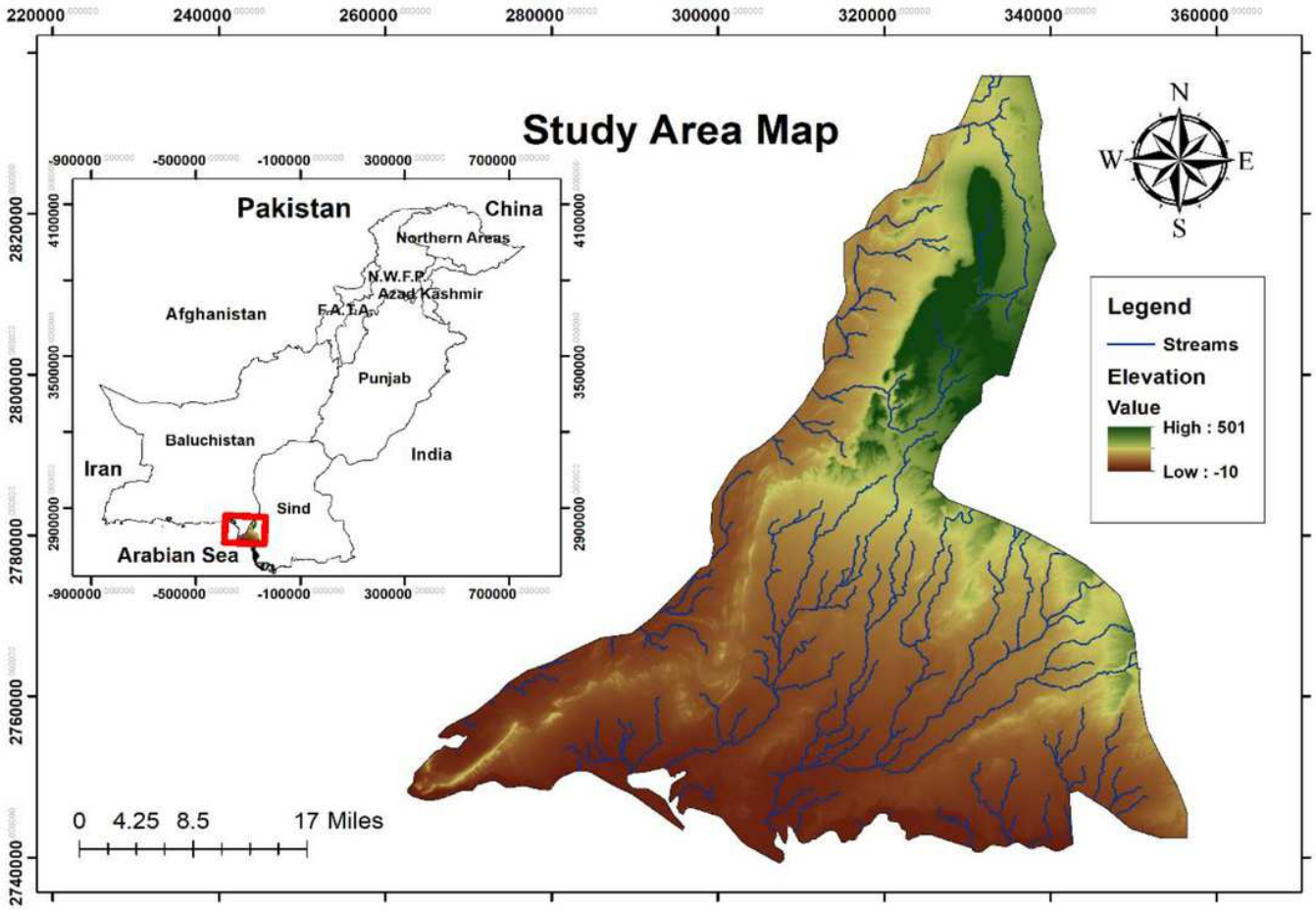

Figure 1

Map of Study Area

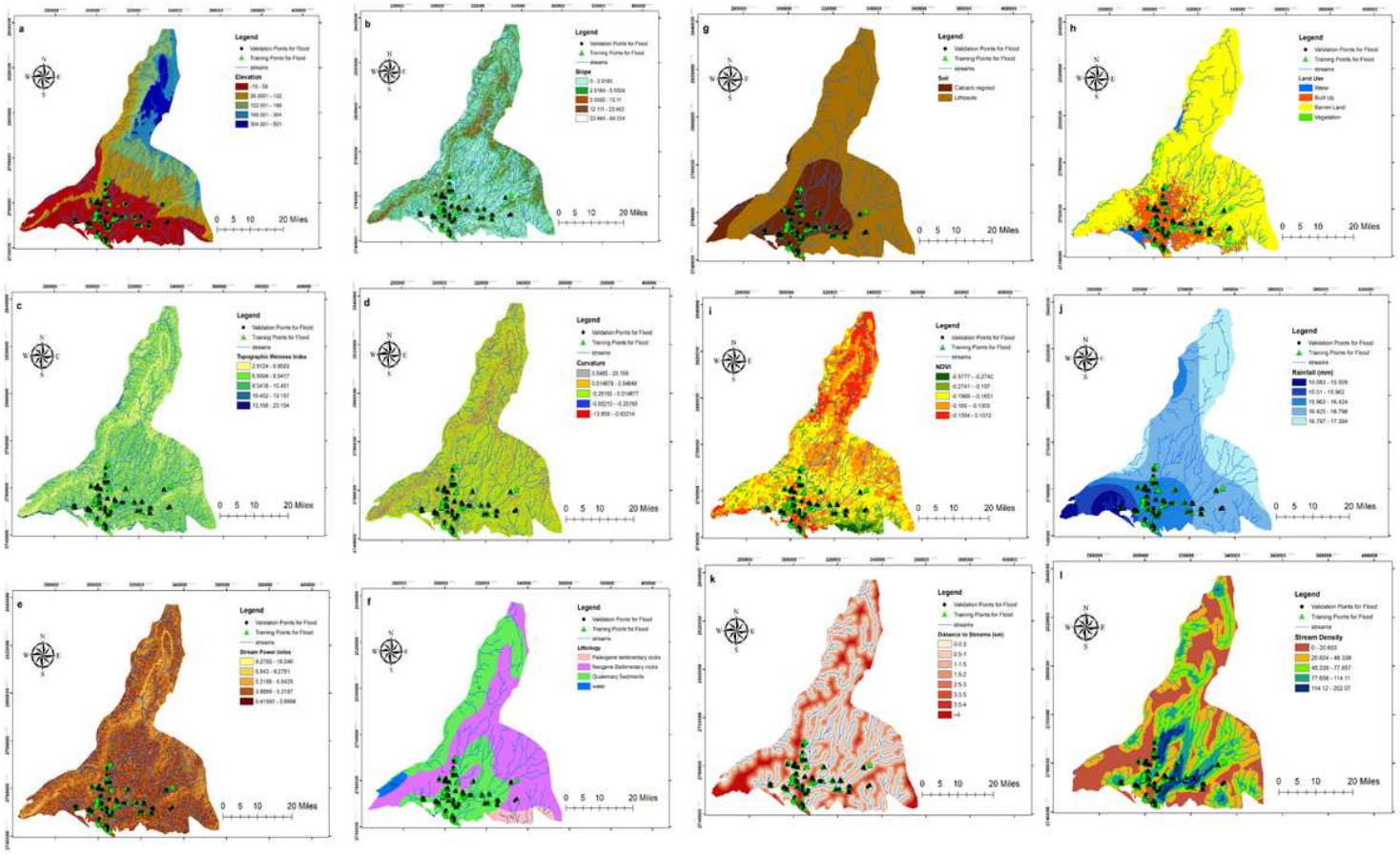

Figure 2

Influencing factors for flood susceptibility mapping: (a) Elevation, (b) Slope, (c) Topographic Wetness Index, (d) curvature, (e) Stream Power Index, (f) Lithology, (g) Soil, (h) Land Use, (i) NDVI, (j) Rainfall, (k) Distance from River, (l) Stream Density

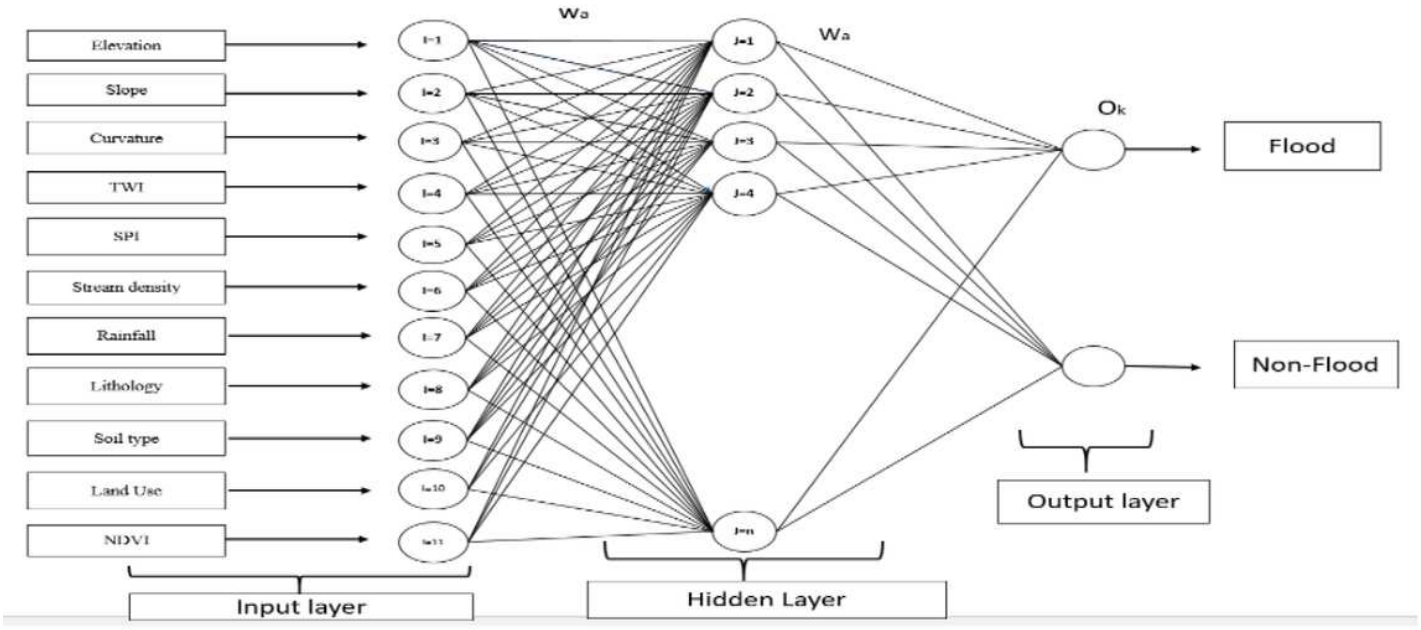

Figure 3

The general structure of an MLP model used for flash flood modelling

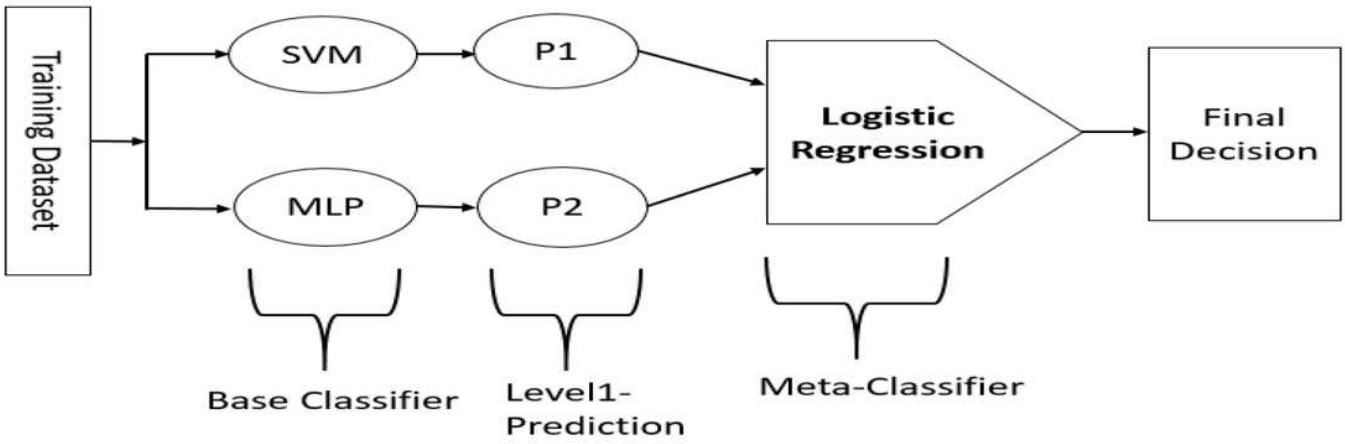

Figure 4

The basic steps in training of stacking classifier

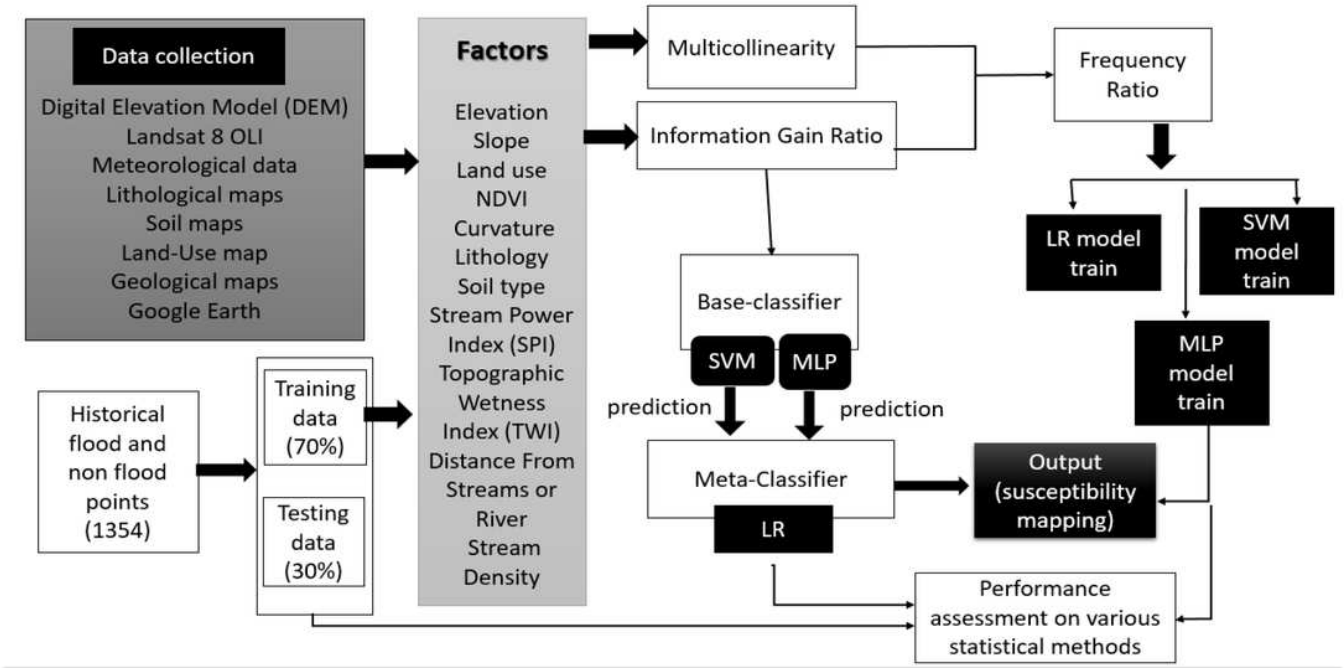

Figure 5

Framework used in this study

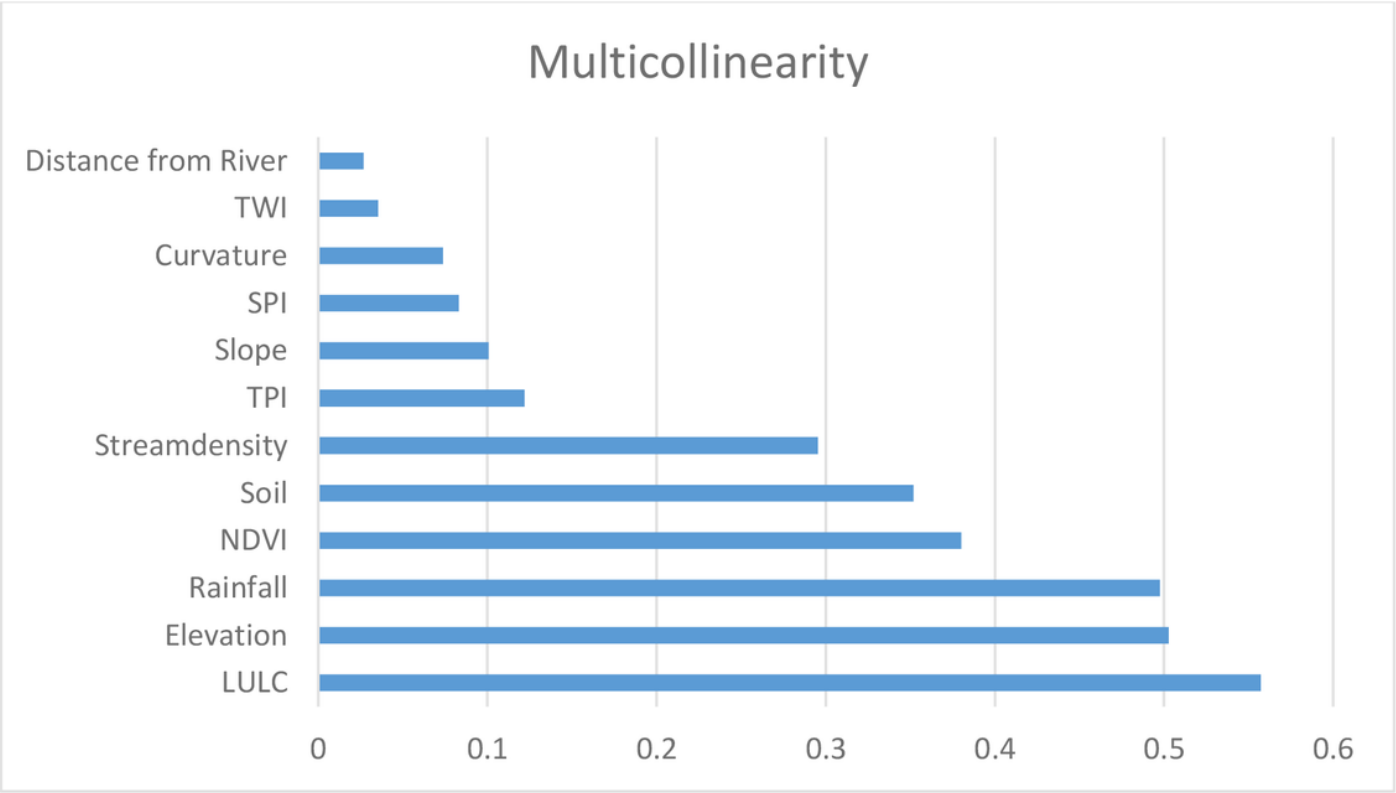

Figure 6

Multicollinearity test among conditioning factors

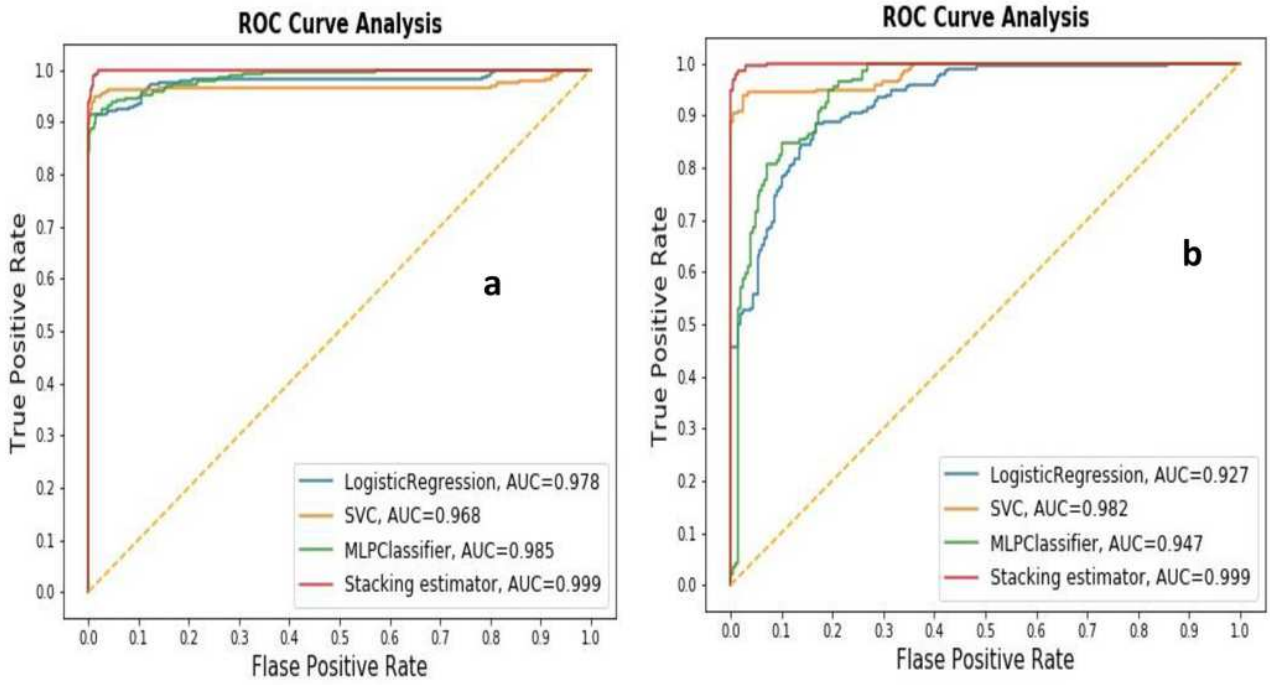

Figure 7

Receiver operating characteristic (ROC) curve a: Validating dataset, b: training dataset

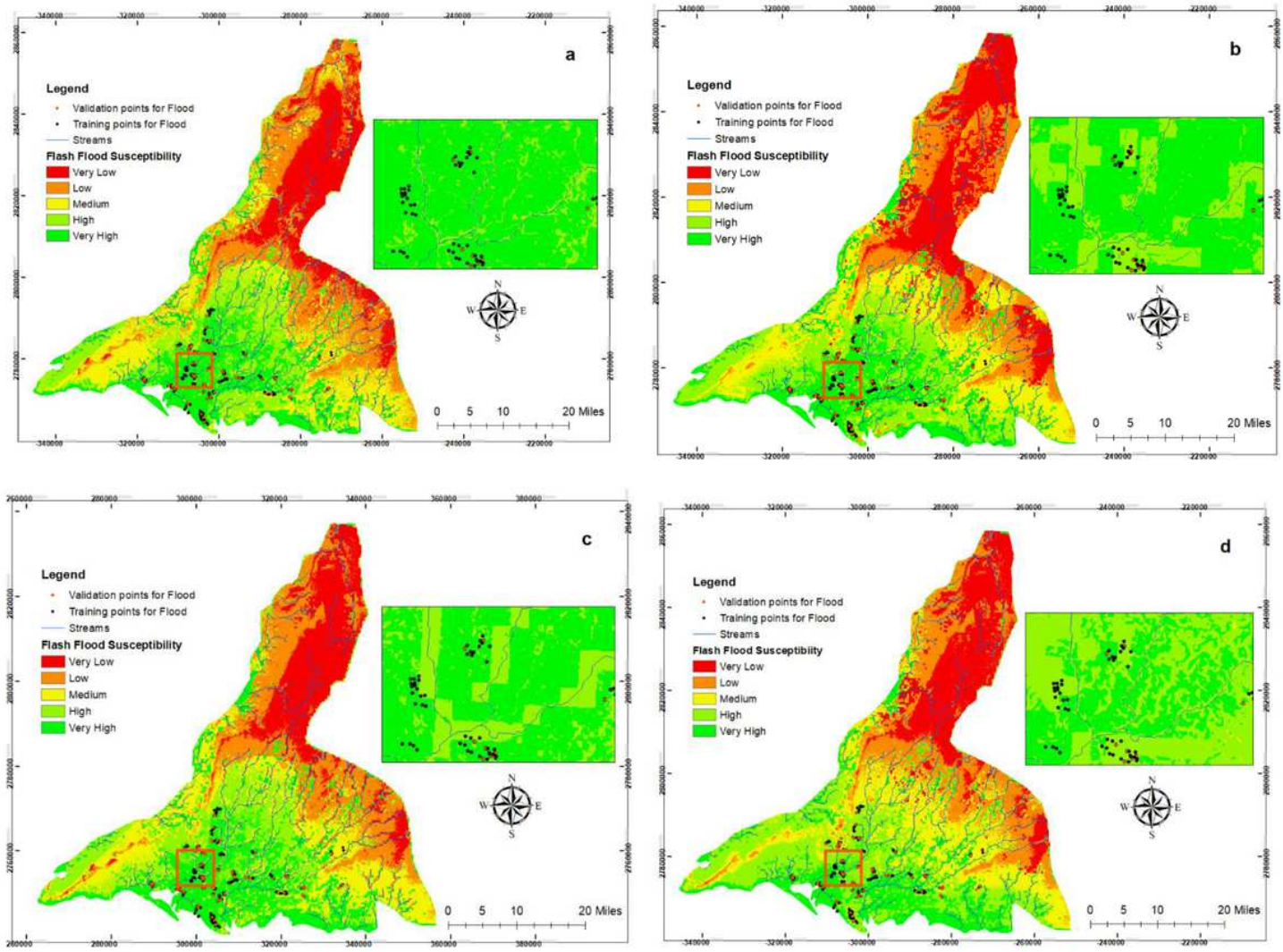

Figure 8

Flash Flood Susceptibility Maps a: Logistic Regression Map, b: Support Vector Machine Map, c: Multi-Layer Perceptron map d: ensemble map

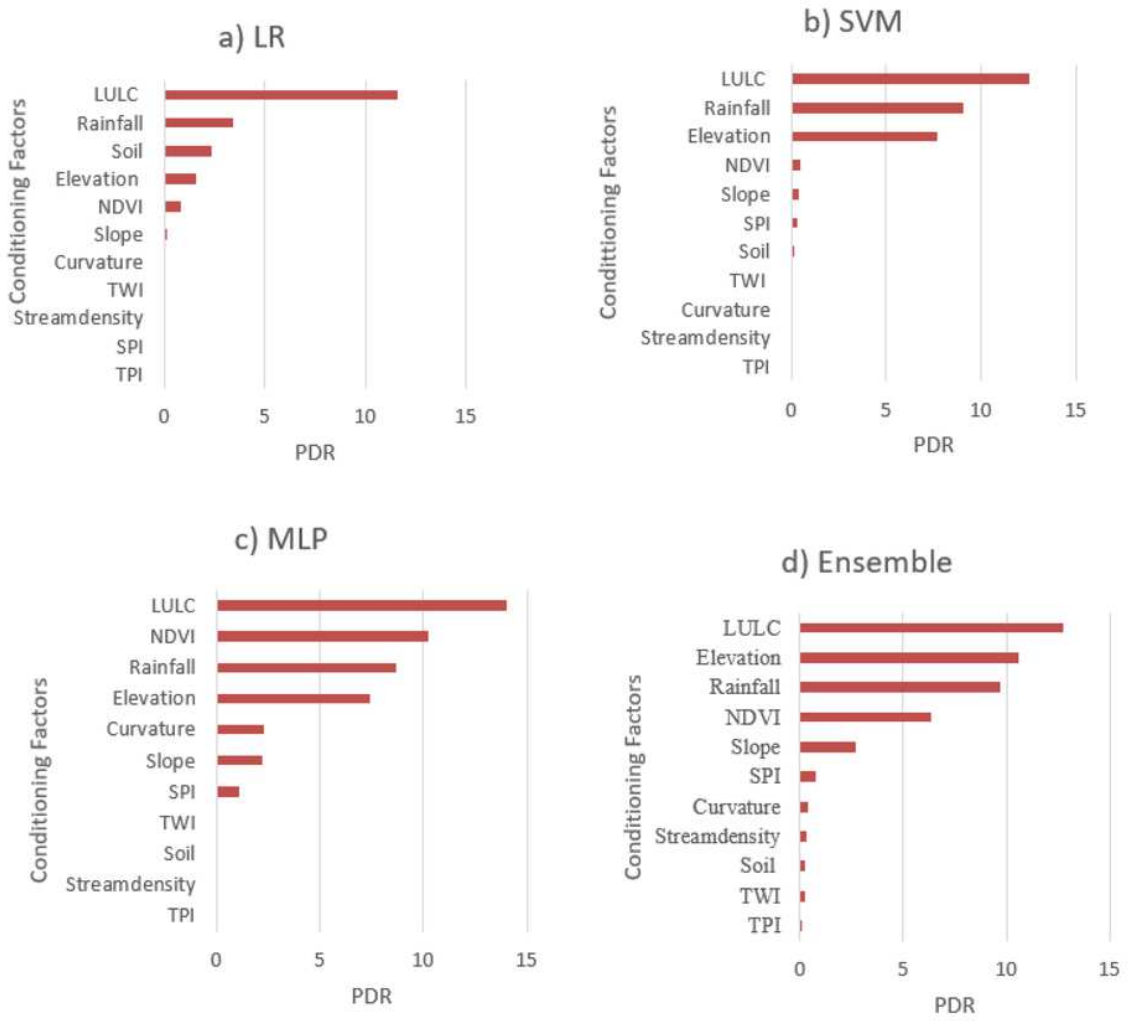

Figure 9

Sensitivity analysis results from the Jackknife test: a) Logistic Regression (LR), b) support vector machine (SVM), c) Multi-Layer Perceptron (MLP), and d) Ensemble