Flood Susceptibility Mapping Using Explainable Machine Learning Models

***Kurugama[1] KAKM, Kazama[1] S, and Chaminda[2] SP**

[1]Department of Civil and Environmental Engineering, Tohoku University, Japan
[2]Department of Earth Resources Engineering, University of Moratuwa, Sri Lanka

*Corresponding author – kurugama.arachchige.kumudu.madhawa.r4@dc.tohoku.ac.jp

Abstract
Flooding is one of the most frequently encountered natural disasters globally. Frequent severe flood occurrences in Rathnapura city, Sri Lanka caused damages to both human lives and infrastructures. Data-driven models have been showing their ability of flood susceptibility mapping (FSM) in data-scare regions as an alternative to traditional hydrological models, but they are not widely used by stakeholders due to their black-box nature. This research suggests utilising the shapley additive explanation (SHAP) method to interpret the results generated by the CatBoost machine learning model and to assess the influence of different variables on flood susceptibility mapping. A flood inventory (445 flooded locations) and thirteen flood conditioning factors were used to implement the model and results were validated using the area under curve (AUC) method, which showed a success rate and prediction rate of 93.1% and 92.5%, respectively. SHAP plots indicated that the regions with lower elevations and topographic roughness values, gentler slopes, closer proximity to rivers, and moderate rainfall are more susceptible to flooding. According to the results obtained, we suggest incorporating SHAP-based data-driven models in forthcoming studies on FSM to enhance the interpretations of model outcomes.

Keywords: AUC, Flood susceptibility mapping, GIS, Gradient boosting, Machine learning

1 Introduction

Floods rank among the most devastating and catastrophic calamities on a global scale due to their immense damage to the infrastructure and the natural environment [1]. Although flood risk is global, most flood-exposed people live in south and east Asian regions, where approximately 90% of human losses are caused by natural hazards, principally floods. The likelihood and severity of flooding are anticipated to increase because of the effects of climate change, deforestation, poor management of land use, and rapid urbanisation [2]. Therefore, prior identification of flood-prone regions and factors that drive flood occurrences, is one of the critical steps in developing flood mitigation strategies to reduce the impact of floods and effectively allocate resources to future flood events.

In recent years, new approaches for flood susceptibility mapping have been developed including knowledge-based qualitative methods, physically based hydrological models [3] and historical data-based statistical and machine learning (ML) methods [4]. Despite the superior performance of hydrological models, their usage is hindered by the time-consuming nature of their implementation and the need for a substantial amount of data for calibration and validation, which restricts their applicability in regions with a lack of data. Qualitative techniques like the analytic hierarchy process (AHP) and analytical network process (ANP) are straightforward, cost-effective, less time-consuming, and easier to implement for flood susceptibility studies and more suited for regional studies.

However, qualitative models demand expert knowledge when selecting influencing factors and their attributes which could generate a significant level of bias in the prediction results [5]. Statistical models like frequency ratio (FR) and weight of evidence (WoE) are widely used and their accuracy has been demonstrated in several studies [6]. Nevertheless, most statistical methods are based on the assumption of linearity, which is inadequate for understanding floods since they are complex and multidimensional events. On the other hand, ML-based models like random forest (RF), support vector machine (SVM) and artificial neural networks (ANN) are more accurate than other FSM models [7], particularly in data-scare regions. While ML models often yield more accurate outcomes, they are commonly referred to as "black boxes" owing to their limited model interpretability or explainability and, therefore rarely selected by stakeholders [8]. The Shapley additive explanation (SHAP) model is widely utilised as a transparent and interpretable model. It is increasingly employed in diverse areas of geo-hazard research, including earthquake damage estimations, drought mapping and vegetation classification [9] etc. In this study, we used an ensemble tree-based CatBoost model to develop a flood susceptibility map for Rathnapura and then applied SHAP to interpret the model outcomes.

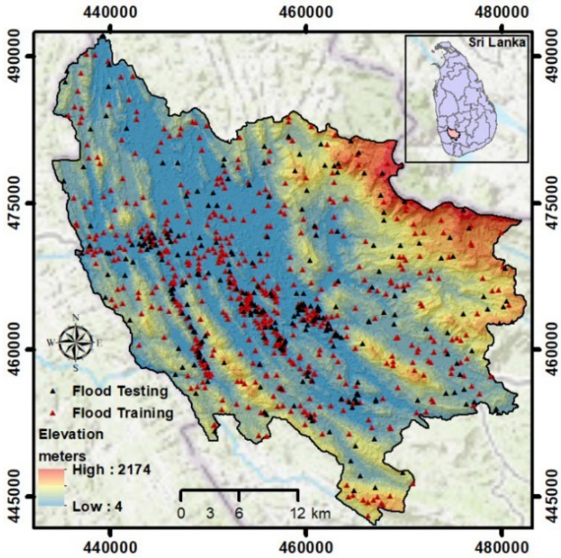

Figure 1: Study area

2 Methodology

2.1 Study area

The study area is Rathnapura city, which is located in the upper catchment of Kalu river, Western Province, Sri Lanka (Fig. 1). The weather of Rathnapura city is hot, humid, and cloudy. Rathnapura District receives an annual average rainfall of 3100 mm and experiences distinct wet and dry seasons. The temperature normally ranges from 71°F to 93°F throughout the year. The research area is centred between the latitudes 6° 55' 51"N and 6° 28' 36"N, longitudes 80° 10' 37"E and 80° 36' 57"E which encompasses nine Divisional Secretariat Divisions in Rathnapura District. Seasonal flooding is one of the major threats in this region and Rathnapura city experienced devastating floods following significant rainfall in the years 2003, 2010, 2014, and 2017, and was identified as a good application site for flood susceptibility mapping.

2.2 Flood inventory data

The initial stage of FSM involves recognizing flood spots or areas by referring to the historical documentation of previous flooding incidents. The prediction outcomes of machine learning and statistical modelling are significantly impacted by the accuracy of the historical flood location data [10]. A total of 445 flood location points were chosen as the flood inventory based on documentary sources and field survey data collected from the Disaster Management Centre (DMC), Sri Lanka. Similarly, the same number of points (445 points) were selected as non-flooded points across the region to improve the model training process. The flood layer was made up of 0 and 1 values, with 1 indicating the presence of flooding and 0 indicating the absence of flooding across the area. The flood inventory map was then divided into 70% and 30% sections for training and validation purposes, respectively.

2.3 Flood conditioning factors

Choosing the flood conditioning factors (FCFs) for FSM is crucial and has a direct effect on the accuracy of the model. Flood occurrence is governed by several factors including watershed characteristics, catchment area, topographic factors, meteorological factors, land use land cover (LULC) types, and human-induced factors. In the FSM domain, there are no universally established guidelines or benchmarks for selecting FCFs. As a result, a total of 13 FCFs were chosen for analysis, encompassing hydrological, topographical, landform, and anthropogenic factors. The hydrological FCFs include rainfall, topographic wetness index (TWI), sediment transportation index (STI), and distance from rivers. The topographical FCFs are altitude, aspect, curvature, slope, and topographic roughness index (TRI). LULC, soil texture, and normalised difference vegetation index (NDVI) are the landform FCFs. Distance from the river is an anthropogenic factor. Each of these FCFs was transformed into raster format with 30m x 30m spatial resolution using Arc GISPro v.3.0. All topographic factors were derived from the digital elevation model (DEM), which was obtained from the SRTM (Shuttle Radar Topography Mission) DEM (https://earthexplorer.usgs.gov/) at a spatial resolution of 30 m. Arc GIS Pro v.3.0 software was used to derive hydrological factors such as STI, TWI, and TRI using the following equations.

$$TWI = ln\left(\frac{A_s}{\tan\beta}\right) \tag{1}$$

$$STI = \left(\frac{A_s}{22.13}\right)^{0.6} * \left(\frac{\sin\beta}{0.0896}\right)^{1.3} \tag{2}$$

where, As is the specific catchment area and β (radian) is the slope gradient (in degrees) at each pixel.

$$TRI = \sqrt{Abs(max^2 - min^2)} \tag{3}$$

where max and min are the maximum and minimum values of the cells in the 3x3 rectangle altitude neighbourhoods, respectively. NDVI, which is a popular vegetation index is calculated as the following equation.

$$NDVI = \frac{NIR-RED}{NIR+RED} \tag{4}$$

Where RED and NIR are red and near-infrared bands of Landsat 08 (TIRS & OLI) Satellite image. The soil and land use maps were obtained from the Survey Department.

2.4 Feature selection

Multicollinearity analysis was conducted to assess the relationships among the FCFs based on variance inflation factor (VIF) and tolerance (TOL). In general, when there is strong multicollinearity, it becomes challenging for the model to make precise estimations as it may incorrectly depict the significant variables in the statistical models. If the VIF > 5 or the TOL < 0.1, it indicates that the factor is suffering from multicollinearity issues and should be removed [10].

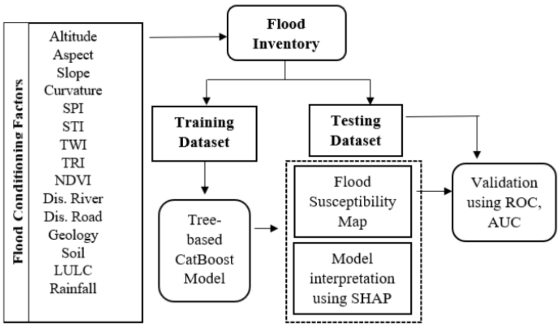

Figure 2: Methodology

2.5 CatBoost model

CatBoost algorithm is an enhanced approach of gradient boosting decision trees (GBDT) [11]. This model is created incrementally through a stagewise approach, with each stage refining the approximation more precisely. Suppose that we are looking at a dataset that contains the following samples:

$$Q = \{(X_j, y_j)\}_{j=1,2,\ldots,m} \tag{5}$$

where $X_j = (x_j^1, x_j^2, \ldots, x_j^n)$ is a vector with n features and $y_j \epsilon \mathcal{R}$ indicates the labelled vector set which can be binary [flooded-1, non-flooded-0]. The primary goal of the learning method is to develop a function $H : \mathcal{R}^n \rightarrow \mathcal{R}$ that reduces the expected loss which can be depicted in the following equation.

$$L(H) := \mathbb{E}L(y, H(X)) \tag{6}$$

where $L(.,.)$ represents a smooth loss function, X and y are some testing data samples from the training dataset Q. The process generates a series of successive estimations $H^t : \mathcal{R}^m \rightarrow \mathcal{R}\ (t = 0, 1, ..)$ through iterative steps.

$$H^t = H^{t-1} + \alpha g^t \tag{7}$$

where α is the step size and the function $g^t : \mathcal{R}^n \rightarrow \mathcal{R}$ is a base predictor, which is chosen from a group of functions G to decrease or limit the expected loss can be explained as in the following equation.

$$g^t = arg\ min_{g \in G} \mathbb{E}L(y, H^{t-1}(X) + g(X)) \tag{8}$$

2.6 SHapley Additive exPlanations

The explainability of ML models comes under Explainable Artificial Intelligence (XAI) which demonstrates the importance of conditioning factors to overall predicted results, allowing the analyst to determine what the ML model considers when assessing flood susceptibility [7]. In this study, the SHAP technique is employed to gain insights into the predicted class outputs generated by the proposed ML model. By calculating SHAP values, we can determine the individual influence (positive or negative) of each distinct sample on the likelihood of susceptibility. The Shapley value is calculated based on the average marginal contribution across all possible permutations of the features, as following equation.

$$\phi_i(f, x) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} [\upsilon(S \cup \{i\} - \upsilon(S))] \tag{9}$$

where ϕ_i is the contribution of feature i, N is the set of all features, n is the number of features in N, S is the subset of N containing feature i, and $\upsilon(N)$ is the base value.

2.7 Model validation

In this study, the model outcomes were validated using the success and prediction rate curves under the area under curve method (AUC) using the training and testing data sets, respectively. This demonstrates a perfect classification where AUC = 1, and a classification by chance where AUC = 0.5. The overall research methodology is illustrated in Fig 2.

3 Results

3.1 Multicollinearity test

Table 1 displays the outcomes of the multicollinearity analysis of the thirteen FCFs. According to the results, it is evident that altitude had the greatest VIF (3.520) and the smallest TOL (0.284) values. Nevertheless, both values were within the threshold limits (5 and 0.1, respectively), which suggests that there is no significant multicollinearity among the thirteen FCFs. Therefore, based on the above results, all thirteen FCFs were considered in the modelling process.

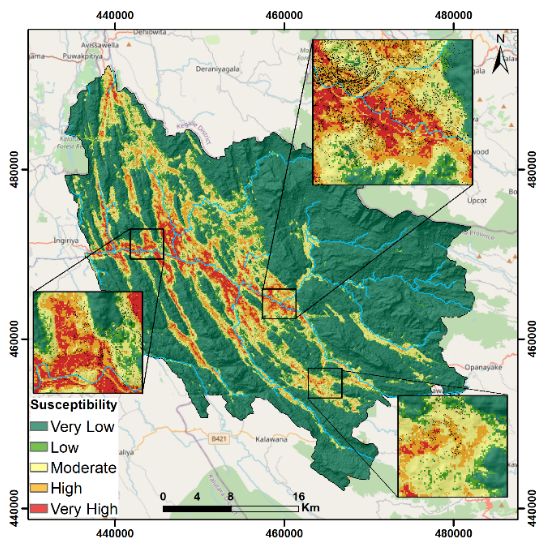

Figure 3: Flood susceptibility map

Table 1: Multicollinearity test results

FCF	TOL	VIF
Rainfall	0.760	1.315
Dist. River	0.563	1.775
Slope	0.591	1.689
LULC	0.803	1.244
Altitude	0.284	3.520
TRI	0.501	1.993
TWI	0.511	1.956
Aspect	0.793	1.260
Soil	0.832	1.201
STI	0.548	1.822
Dist. Road	0.399	2.500
Curvature	0.900	1.111
NDVI	0.832	1.201

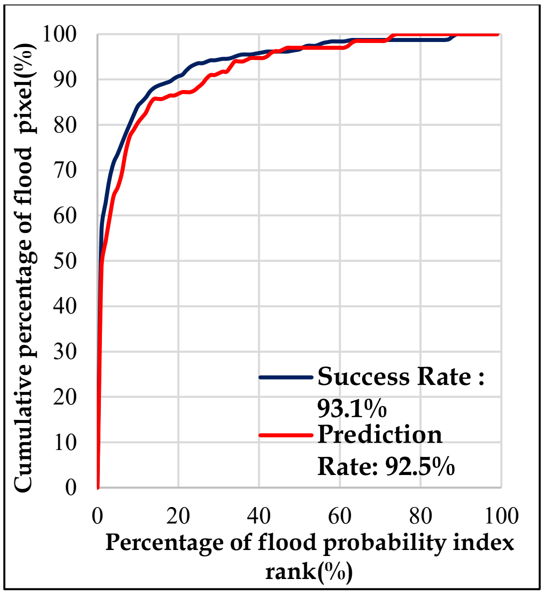

Figure 4: Success and prediction rate curves under the AUC method

3.2 Flood susceptibility map

The flood probability map, which was obtained as the model outcome was then classified into five flood-susceptible zones using the natural breaks classification technique (Fig. 3). According to Fig. 3, 8.9% of the study area was covered with high and very high susceptibility classes where 15% of the total area was classified as moderate risk area and most of the region (75.6%) was classified as low and very low susceptible to flooding.

3.3 Model validation results

In the model validation stage, the AUC calculations presented in Fig. 4, provide evidence of the impressive performance exhibited by the CatBoost model. In this study, the model outcomes were validated using the success and prediction rate curves under the area under the curve method. According to the results, our approach obtained 93.1% and 92.5% success and prediction rate values respectively.

3.4 Feature importance using SHAP

We calculated the SHAP values to evaluate the feature importance of the testing dataset using the CatBoost model. As previously stated, the SHAP values not only indicate the significance of features but also reveal whether a feature has a positive or negative impact on the predicted values. Fig. 5 displays the SHAP values assigned to input factors, organized based on their contribution level. The x-axis displays the SHAP value, whereas the y-axis represents the conditioning factors.

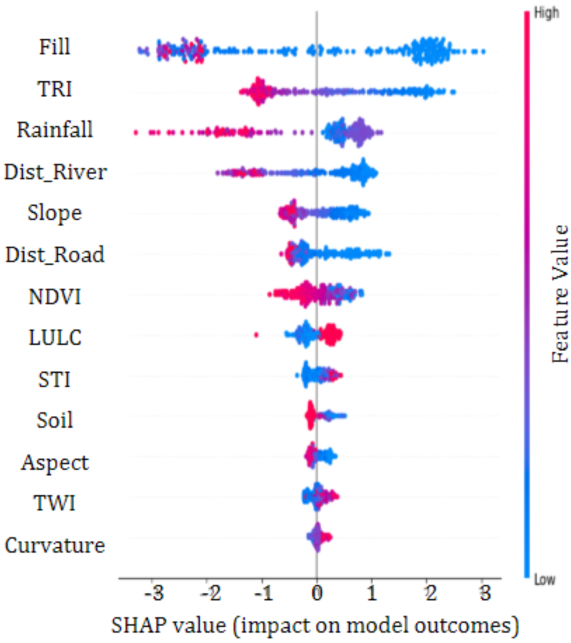

Figure 5: Shapley additive explanation summary plot

Every dot on the graph symbolizes a data point, and the colour of the dot indicates the value of a variable, where blue shades indicate lower values and red shades indicate higher values. According to the results, elevation, TRI, rainfall, distance from the river, and slope variables show a greater impact on flood occurrence, distance from the road, NDVI, LULC and soil variables show moderate effect while STI, TWI and curvature show the least importance. Such plots can be used to examine the relationships between a target and variables of interest.

4 Discussion

Prior identification of flood hazard-prone regions is one of the critical steps in developing flood mitigation strategies to reduce the impact of floods and effectively allocate resources to future flood events. To our knowledge, this is the first study to apply the explainable artificial intelligence (XAI) model and ML models for FSM in Sri Lanka. we used an ensemble tree-based CatBoost classifier for FSM of Rathnapura, Sri Lanka which has a long history of flood events causing immense damage to infrastructure and loss of life. The model achieved a prediction accuracy greater than 92% in both the success and prediction stages demonstrating its reliability.

So far, data-driven models in FSM were considered black boxes because they were difficult to explain locally. In this study, we investigated the importance of individual factors in predicting flood susceptibility and explained the model's predictions based on the input feature values using the SHAP method. According to the results, elevation, TRI, and rainfall were the top three contributing factors. The study area is located in an undulating terrain which is characterized by a series of hills, valleys, ridges, and depressions that create a dynamic and uneven landscape. Thus, elevation is a significant factor when assessing flood susceptibility. Elevation determines the natural flow of water. Areas at higher elevations generally act as water sources, while low-lying areas with more residential density tend to accumulate water causing floods. Thus, elevation helps to determine the drainage patterns as well as the slope and gradient of the study region. TRI quantifies the degree of surface roughness within the study area.

The TRI helps to identify areas that can impede or redirect the flow of water during floods. Features such as ridges, hills, and vegetation cover impact the flow pattern. Furthermore, TRI indirectly reflects the infiltration capacity of the land. Here, areas with high roughness values often have more vegetation, which can absorb and retain water, reducing runoff and flood risk.

Rainfall can be considered as one of the triggering factors of flooding. The city area experiences moderate rainfall, and the surrounding mountains experience high levels of rainfall due to orographic lifting. Thus, the river water level increases, as well as cities located in the valley area (downstream of these mountains) inhibit the rapid runoff of water and tend to have an increased risk of flooding. This effect is getting accelerated due to poor drainage systems and the impervious surfaces in the city area.

In addition to that, lower proximity to rivers, roads, and areas with lower slope gradient values showed a higher relationship towards flood occurrence. According to the landform factors, regions with residential areas and paddy fields and lower NDVI values tend to have a high risk of flooding reflects the fact that metropolitan areas, where there are more impermeable surfaces that increase the s runoff, and vegetated areas provide varying degrees of protection against flooding. Furthermore, alluvial soil type showed a higher effect towards flood occurrence.

According to the results, although lower STI values, higher TWI values, flat aspect and flat curvature regions show higher effects on flood occurrence, their contribution level on the overall prediction is less compared to the other factors. However, some studies showed that using fewer predicting features could improve the model performance [19]. Thus, our suggestion is to conduct additional research to perform an analysis of feature selection, focusing solely on features that have a significant impact on the model’s prediction.

5 Conclusion

Floods rank high as one of the most devastating recurring natural calamities on a global scale. Sri Lanka experiences significant impacts from flood occurrences. Therefore, the development of accurate and interpretable flood susceptibility maps would facilitate the design of effective flood management and mitigation plans. In this study, we used an ensemble CatBoost ML model to develop a flood susceptibility map for the Rathnapura area, Sri Lanka using a total of 13 flood-triggering factors and 445 historical flood events. The primary achievement of this study involves utilising the SHAP explainable algorithm to ascertain the process through which the model attained its outcomes, as well as identifying the variables that had the greatest impact.

The primary outcome of this research revealed that the CatBoost model attained a 92.5% AUC value, indicating a high level of accuracy in its predictions. Additionally, the SHAP summary plots demonstrated that, overall, elevation, TRI and rainfall emerged as the top three dominant factors with flood susceptibility. The generated flood susceptibility map revealed that approximately 9% of the area is classified as high and very high flood susceptibilities. These zones were mainly distributed in the majority of the central area and some parts of the western and northern lower-lying residential and agricultural land areas. Using the SHAP model alongside machine learning models in flood susceptibility modelling can enhance our comprehension of the fundamental mechanisms and factors influencing flood risk, particularly in regions where data is limited. The SHAP model's capacity to incorporate feature interactions enables practitioners to develop more efficient and focused strategies for flood management.

Acknowledgements

The authors wish to thank the Disaster Management Center and Department of Meteorology, Sri Lanka for providing the necessary data for the analysis.

References

[1] A. M. M Youssef, B. Pradhan, and S. A. Sefry, "Flash flood susceptibility assessment in Jeddah City (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models", *J. Environmental Earth Sciences*, vol. 75, no. 1, pp. 1-16, 2016.

[2] D. Rincón, J. F. Velandia, I. Tsanis, and U. T. Khan, "Stochastic flood risk assessment under climate change scenarios for Toronto, Canada using capra", *Water*, vol. 14, no.2, pp. 227, 2022.

[3] R. Jayakrishnan, R. Srinivasan, C. Santhi, and J. G. Arnold, "Advances in the application of the SWAT model for water resources management", *J. Hydrol Process*. vol. 19, pp.749–762, 2005.

[4] H. E. Aydin and M. C. Iban, "Predicting and analyzing flood susceptibility using boosting-based ensemble machine learning algorithms with Shapley additive explanations". *Natural Hazards*, 2022.

[5] V.M. Chowdary, D. Chakraborthy, A. Jeyaram, Y.V.N.K. Murthy, J.R. Sharma, and V.K. Dadhwal, "Multi-Criteria Decision-Making Approach for Watershed Prioritization Using Analytic Hierarchy Process Technique and GIS", *Water Resour. Manag*, vol. 27, pp. 3555–3571, 2013.

[6] M.S. Tehrany, B. Pradhan, and M.N. Jebur, "Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in gis", *J. Hydrol*,vol. 512, pp. 332–343, 2014.

[7] M.S. Tehrany, B. Pradhan, and M.N. Jebur, "Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method", *Stoch. Environ. Res. Risk Assess*, vol. 29, pp.1149–1165, 2015.

[8] B. Pradhan, S. Lee, A. Dikshit and H. Kim, "Spatial flood susceptibility mapping using an explainable artificial intelligence (XAI) model." *Geoscience Frontiers*, vol. 14, no.6, 2023.

[9] R. Abedi, R. Costache, H. Shafizadeh-Moghadam and Q. B. Pham, "Flash-flood susceptibility mapping based on XGBoost, random forest and boosted regression trees" *Geocarto International*, vol.37, no.19, pp. 5479–5496, 2021.

[10] Y.H. Shao and N.Y. Deng, "A coordinate descent margin based-twin support vector machine for classification", *Neural Network*s, vol. 25, pp. 114-121, 2012.

[11] M. Saber, T. Boulmaiz, M. Guermoui, K. I. Abdrabo, S. A. Kantoush, T. Sumi, H. Boutaghane, D. Nohara, and E. Mabrouk, "Examining lightgbm and CatBoost models for Wadi Flash flood susceptibility prediction" *Geocarto International*, vol.*37*, no.25, pp. 7462–7487, 2021.