ISERME 2024 Proceedings
2nd September 2024 – Sapporo, Japan
https://doi.org/10.31705/ISERME.2024.32

Session 4 – A

Forecasting the Impact of Land Utilization on Flood Vulnerability through Machine Learning and Remote Sensing in Athuraliya and Akuressa Divisional Secretariat

Sujahath[1] MSM, Senarathna[1] HDK, Saranga[1] KHGR and *Dissanayaka[1] DMDOK
[1]*Department of Earth Resources Engineering, University of Moratuwa, Sri Lanka.*
*Corresponding author – Email: dmdok@uom.lk

Abstract

Akuressa and Athuraliya Divisional Secretariats in Sri Lanka frequently experience severe damage to human lives, infrastructure, and economic growth due to floods. These floods are often caused by elements like land use patterns, urbanization, and environmental degradation. This study aims to establish the connection between flood vulnerability and land use as importantly necessary for effective disaster management and mitigation strategies. Therefore, this research provides useful knowledge on flood vulnerability prediction based on land use patterns that can be used by policymakers, urban planners, and disaster management authorities for decision-making on proactive measures that will reduce the negative impacts caused by flooding while building resilience in the region.

The primary purpose of this research is its innovative and essential because no earlier study has applied these cutting-edge techniques to assess flood risks in this area. Consequently, there is a significant gap in the current knowledge base and practice. Therefore, this research is intended to understand the land utilization situation in the area and how it affects flood vulnerability, identify environmental key variables that contribute to flood susceptibility, and use machine learning models including XGBoost, Random Forest, and CatBoost for predicting flood susceptibility. The latter also uses DEM derived factors with geological, soil, land use/land cover data, distance from roads and rivers to provide a closer understanding of flood conditioning factors within the study area.

The XGBoost algorithm gave an accuracy score of 0.91 throughout the other utilized Machine Learning models, confirming how well machine learning performs when it comes to predictions. The results from the machine learning model were then used to determine the feature importance according to each conditioning factor that influences floods. Based on these feature importance values; a future risk map was generated using ArcGIS software. Therefore, this research indicates that prediction-based planning is more effective than post event-based recovery measures in building resilient and sustainable communities prone to flooding like Akuressa and Athuraliya Divisional Secretariat. In addition, these findings show that Machine Learning (ML) and Remote Sensing (RS) have potential for improving on-flood forecasting techniques as well as mitigating measures.

Keywords: Flood Vulnerability; Land Utilization; Remote Sensing; Machine Learning.

1. Introduction

Floods are among the most severe natural disasters, having grave effects on social, economic, and environmental aspects worldwide. They cause huge financial losses to businesses, transportation, and the overall economy, as well as human deaths, body injuries, and disease outbreaks[1]. Sri Lanka has shown an increase in extreme hydrological events in the river Basin due to seasonal rainfall variations driven by the southwest and northeast monsoons. Moreover, flood susceptibility is further increased by human activities such as deforestation, urbanization, and poor land use practices[2].

Regions like Matara, Sri Lanka, with diverse topography and complex hydrology only serve to enhance their vulnerability to floods. The wet zone of Matara, characterized by hills, rivers, and valleys, especially exposes low-lying areas and riverine plains to flooding[3]. For instance, effective flood risk management requires proactive measures integrating land use planning with hydrology coupled with vulnerability assessment for future mitigation efforts. Modern technologies based on big data are paramount for sustainable planning, while flood vulnerability assessments benefit from GIS-based mapping, which assists crisis responders and decision-makers.

Machine learning (ML) techniques and remote sensing (RS) methodologies have become pivotal in flood impact mitigation. ML algorithms like Random Forest and XGBoost assess flood susceptibility by analyzing environmental factors and creating accurate predictive models[4]. Combining ML with RS data, including satellite imagery and digital elevation models (DEMs), is crucial for precise flood vulnerability assessment. These models generate detailed flood vulnerability maps, providing spatial information on land cover, topography, and hydrological features[5]. Such models support disaster preparedness and resilience-building efforts, enhancing overall disaster management strategies.

This study focuses on the Akuressa and Athuraliya Divisional Secretariats within the Matara district. It aims to forecast flood vulnerability based on land utilization patterns using machine learning techniques and remote sensing methodologies.

2. Methods and Materials

2.1 Study Area

Akuressa and Athuraliya Divisional Secretariats are in the Matara District of Southern Province, Sri Lanka. These areas are particularly susceptible to frequent flooding due to their geographical and climatic conditions. The geolocation of Akuressa is between Northern latitudes 6°04’00” - 6°08’00” and Eastern longitudes 80°28’00” - 80°32’00”, while Athuraliya is situated between Northern latitudes 6°00’00” - 6°04’00” and Eastern longitudes 80°28’00” - 80°32’00” [6].

These regions fall within the catchment area of the Nilwala River, one of the major rivers in Sri Lanka, which significantly influences their flood dynamics. The Nilwala River and its tributaries traverse through Akuressa and Athuraliya, contributing to the high flood susceptibility of these areas, especially during the south-west monsoon season (May to September), which brings intense and prolonged rainfall [7].

The natural topography and river network play a critical role in the hydrology of the region. The presence of numerous small rivers and streams that feed into the Nilwala River exacerbates flood risks during heavy rainfall periods. Effective management of the Nilwala River catchment is crucial to mitigating the frequent flooding experienced by the residents of Akuressa and Athuraliya.

2.2 Flood Inventory

Maps of the flood inventory are essential for creating models of flood vulnerability. They provide historical records of flood incidents and, by analyzing the connections between those events and conditioning factors can be applied to forecast the likelihood that floods will occur in the future [8]. In this research study we used flood points recorded between 2003 and 2017 to create a map of the flood inventory. The flood data points were collected from a publicly available database.

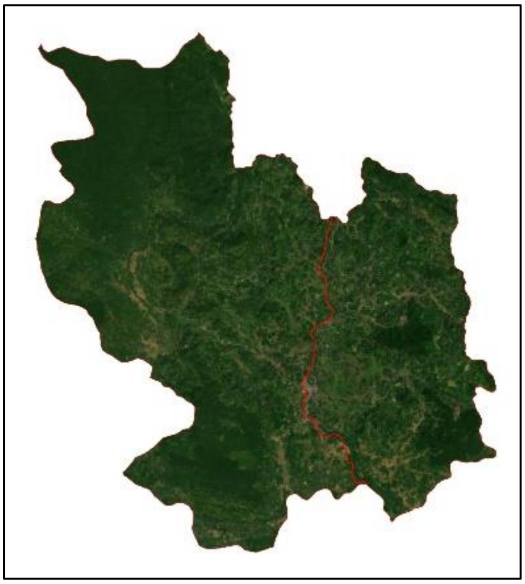

Figure 1 Study Area Map

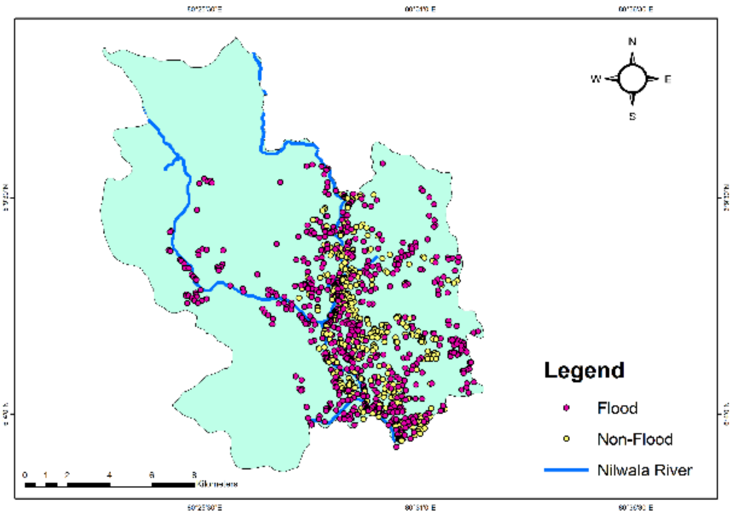

Figure 2 Flood Event Locations

To construct the model, 986 flood points in total were gathered. Since the suggested model is binary, it is also necessary to gather non-flooding locations, such as those in slope and high elevation areas that are never impacted by flooding. Finally, the data points were split into two sets: 30% were used as validation data to assess the models, and 70% were used as training data to train the models.

2.3 Flood Conditioning Factors

2.3.1 Slope

Slope is an important factor in floods because it affects the speed and direction of water flow. Slope refers to the degree of steepness of the land surface. Slope becomes crucial when we contemplate flooding scenarios[9]. Steeper slopes generally lead to faster water flow, increasing the velocity of flow. This rapid flow can trigger heightened erosion, with the water crashing down, carrying away soil and rocks. It facilitates quicker water movement to lower areas, potentially inducing flooding. However, the effects of slope on flooding can vary depending on other factors[10]. Various methods related to GIS can be used to integrate slope data into flood risk models. In this study, a 30 m-resolution DEM was used to calculate the slope in the study area (Fig. 3a).

2.3.2 Altitude

Altitude refers to the height above sea level of a particular location. In mountainous regions, altitude plays a crucial role in influencing floods. Studies show that higher altitudes can lead to increased severity of floods due to various factors. in high-elevation mountain streams, altitude impacts sediment retention and floodplain disturbance patterns, with valley confinement being a key predictor of floodplain dynamics[11]. Understanding how altitude affects floods is crucial for effective flood forecasting and mitigation strategies in mountainous areas. GIS utilizes Digital Elevation Model (DEM) data to create altitude maps. By analyzing DEM, GIS can generate detailed topographic images reflecting elevation variations in the study area (Fig. 3b).

2.3.3 Curvature

Curvature is the degree of convexity or concavity of the land surface, impacts various phenomena. Curvature plays a vital role in modeling flood risk as it affects the flow of water and the amount of erosion during flood events[12]. In areas with a high positive curvature, the land surface is convex, sloping upward in all directions. And areas with high negative curvatures are concave, where the land surface slopes downward[13]. A negative value indicates that the surface is upwardly convex at that cell. A positive profile indicates that the surface is upwardly concave at that cell. A value of zero indicates that the surface is linear. Understanding curvature through DEM data is crucial for modeling flood risk, as positive curvature accelerates water flow, increasing erosion risk, while negative curvature slows flow, increasing the risk of water accumulation (Fig. 3c).

2.3.4 Aspect

Aspect map classification into nine categories (flat, north, northeast, east, southeast, south, southwest, west, northwest) using ArcGIS toolbox based on DEM map is discussed, influencing local climate and hydrology[14]. For instance, north-facing slopes generally exhibit higher biomass, coverage, and species diversity compared to south-facing slopes[15]. Additionally, the hydrological behavior and pore water pressure dynamics differ between south-facing and north-facing slopes, influencing landslide initiation conditions (Fig. 3d).

2.3.5 Distance to River

Distance to the river is a critical factor in flood risk modeling. GIS integrates drainage network data and elevations to identify high-risk flood areas near rivers, streams, and canals[16]. Distance to rivers impacts flood vulnerability by influencing land-use decisions. Proximity intensifies flood risks, while increased distance can reduce exposure, affecting community resilience and infrastructure vulnerability in flood-prone areas[17]. Distance to rivers plays a significant role in land use planning decisions, affecting infrastructure vulnerability and community resilience in flood-prone regions (Fig. 3e).

2.3.6 Topographic Wetness Index

Topographic Wetness Index (TWI) plays a crucial role in assessing flood vulnerability in catchments. TWI evaluates the influence of topography on water flow accumulation and runoff, indicating areas with low moisture-holding capacities due to hypsometrical characteristics of the landforms. In the context of flood assessment, TWI helps in understanding the physical behavior of catchments during heavy rainfall events, aiding in the identification of areas prone to floods[18]. Therefore, TWI serves as a significant morphometric factor in evaluating flood vulnerability and implementing suitable mitigation measures in catchment areas prone to flash floods and other water-related disasters (Fig. 3f).

2.3.7 Slope length and Slope steepness (LS) Factor

The slope length and slope steepness factors play crucial roles in influencing flood characteristics. The slope length factor, often denoted as the L-factor, and the slope steepness factor, known as the S-factor, are fundamental parameters in soil erosion and flood modeling. The L-factor defines the impact of slope length on flood response processes, while the S-factor measures the effect of slope steepness on flood-prone areas. These factors are essential in determining flow concentration times, flood depths, and the spatial distribution of flood-prone points in urban areas vulnerable to pluvial flooding[19]. Therefore, understanding and incorporating these factors are crucial for effective flood risk assessment and mitigation strategies (Fig. 3g).

2.3.8 Rainfall

Precipitation plays a crucial role in flood risk modeling by influencing river overflow and the severity of flooding. Incorporating rainfall data into flood risk models involves utilizing historical records, real-time monitoring, and climate projections to assess the impact of precipitation on flooding events. Changes in precipitation patterns, especially extreme rainfall events, can lead to increased flood risks[20]. Increases in precipitation extremes contribute significantly to river floods, impacting flood magnitudes. Interpretable machine learning helps understand these effects and predict future flood risks in a warming climate[21]. In this study, rainfall data from 2003 to 2023 were collected from two rainfall data collection stations in the study area of the Meteorological Department of Sri Lanka.

2.3.9 Land Use & Land Cover (LULC)

Land use and land cover (LULC) significantly influence flood vulnerability, as changes in these parameters affect the hydrological cycle, runoff patterns, and soil infiltration rates. Urbanization, deforestation, and agricultural expansion can increase the impervious surface area, reducing the land's capacity to absorb rainfall and thereby escalating flood risks [17]. Conversely, natural vegetation and wetlands play a crucial role in mitigating floods by enhancing water retention and groundwater recharge [18]. Understanding LULC dynamics is essential for predicting and managing flood risks, particularly in flood-prone regions like Athuraliya and Akuressa.

In this study, we conducted a comprehensive LULC analysis in Athuraliya and Akuressa Divisional Secretariats from 2017 to 2023. Utilizing data from the "Esri | Sentinel-2 Land Cover Explorer" with a 10-meter accuracy, we classified the land cover into four main categories: Trees, Crops, Built, and Rangelands. This categorization helps in understanding the spatial and temporal changes in land cover and their potential impact on flood vulnerability. The LULC data were analyzed for each year within the specified period, resulting in the creation of seven detailed maps that visually represent the changes over time (Fig. 4).

2.3 Methods

The four steps of assessing how land cover and climate change affect flood susceptibility are as follows: (i) gathering and pre-processing the data; (ii) creating the Random Forest (RF), CatBoost and XGBoost models; (iii) evaluating the fit of the suggested models; and (iv) analyzing the results.

(i) Flood conditioning factors for this study were derived from SRTM (30m) DEM, Meteorological Department of Sri Lanka, and Global flood inventory database. A global flood database is used to get the Flood point data.

(ii) The Machine Learning algorithms used for this study are CatBoost, RF and XGBoost, all those ML models were coded using Python Programming Language and ML libraries (TensorFlow, NumPy, Sci-kit Learn, Pandas and Seaborn)

(iii) Typical model evaluation indices for RF, CatBoost and XGBoost usually involve AUC-ROC, accuracy, precision, F1-score, and recall. These measurement parameters determine the performance of the models in correctly classifying data points while ensuring that a balance is struck between precision and recall as well as measuring the trade-off between true positive and false positive rates [22] [23].

(iv) Analysis: All three ML algorithms were used to create the flood susceptibility maps for 2017 and 2021 following validation. Using the natural break approach, the model's output values, which range from 0 to 1, have been classified into five categories: very low, low, moderate, high, and very high.

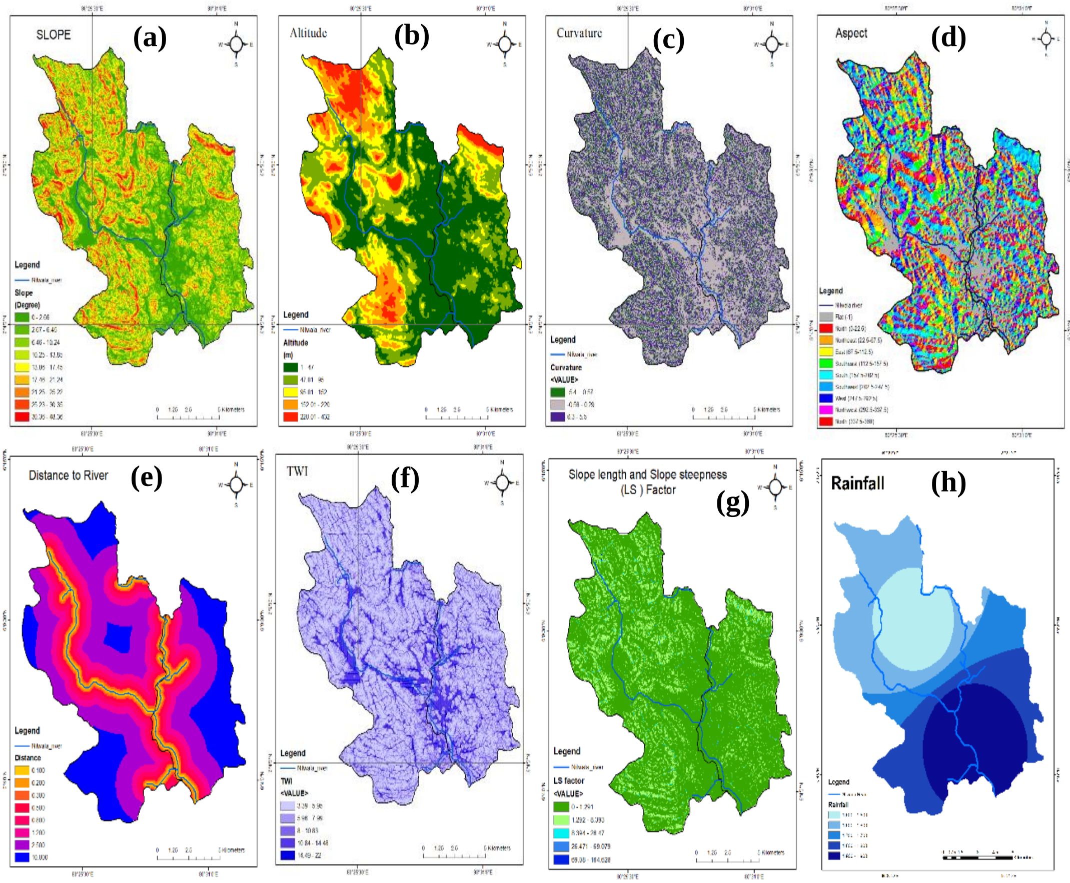

Figure 3 Flood Conditioning Factors

2.4.1 CatBoost

CatBoost is a powerful version of the gradient boosting algorithm, which can handle categorical characteristics in machine learning systems. Developed by Yandex, it has an innovative approach that uses ordered boosting and processing of category features, thus making it a tool of choice for dealing with high-dimensional data with categorical variables

[24]. Unlike other traditional gradient boosting algorithms, CatBoost automatically handles categorical data without one-hot encoding or pre-processing which simplifies work processes and reduces overfitting risk. With such techniques as ordering boosting and computing of feature importance among others incorporated into CatBoost, it delivers accuracy and speed above other alternatives making it applicable in various fields such as e-commerce, finance, advertising for tasks like classification, regression ranking etc.

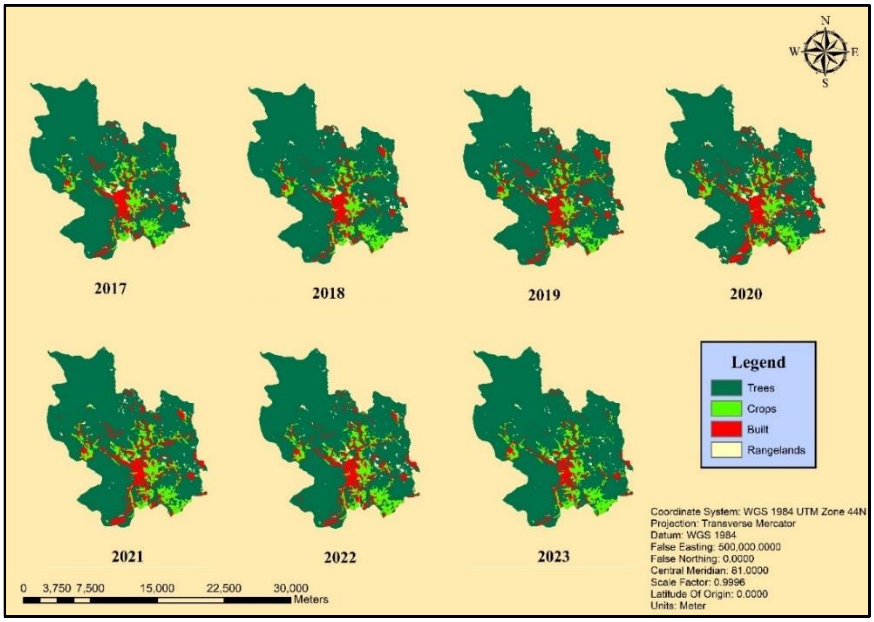

Figure 4 Analysis of LULC over years

2.4.2 Random Forest

Leo Breiman developed the random forest algorithm in 2001, and it is a popular and effective machine learning technique for prediction [25]. It is an ensemble method that combines multiple decision trees to improve predictive accuracy. It is a supervised learning algorithm that generates multiple decision trees and uses majority voting to combine their outcomes [26]. The algorithm has been shown to be effective in improving classification accuracy and reducing learning and classification time. Random Forest can be used to tackle regression and classification problems. It combines a random subspace with a bagging technique.

2.4.3 XGBoost (XGB)

The XGBoost algorithm stands out as a powerful and efficient machine learning algorithm widely applied to classification and regression problems. It amalgamates the strengths of boosting and gradient boosting algorithms to construct a highly accurate predictive model. Chen[22] introduced XGBoost as a distributed system optimized for fast parallel tree construction, fault tolerance, and scalability. Arif Ali[27] highlighted XGBoost's exceptional capability in modeling complex systems, its superior prediction accuracy, and its adaptability in a Python environment. These studies collectively emphasize the significance of XGBoost in machine learning.

3. Results and Discussion

3.1 Factor Selection

It is crucial to choose the right flood occurrence variables since the model's accuracy depends on the quality of the data. In this study, the importance of the conditioning component was assessed using XGBoost. The results shows that rainfall, distance from river, LS factor, curvature, aspect and LULC are the most crucial factors that affects the flood risk in our study area.

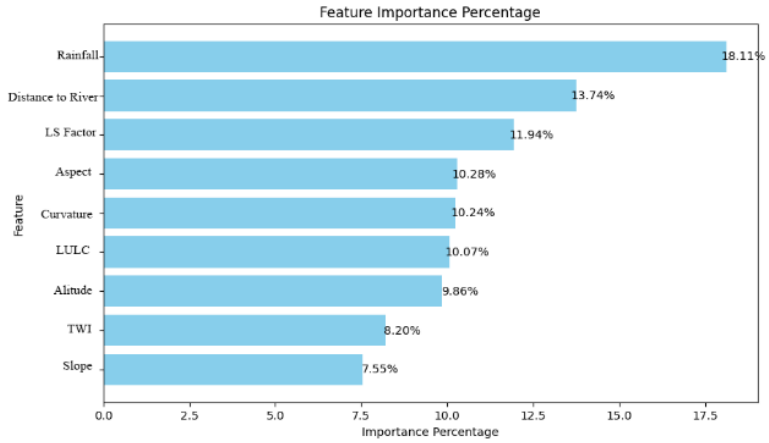

Figure 5 Feature Importance of Conditioning factors

3.2 Risk Maps

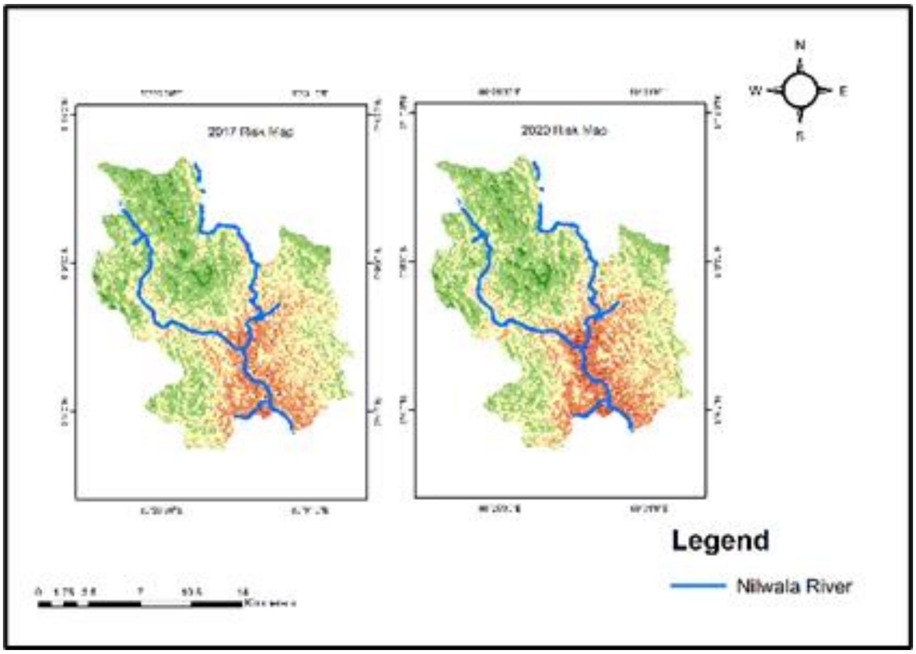

Figure 6 Flood risk maps for the years 2017 and 2020

These risk maps were generated from the results we get from XGBoost machine learning model which gave a high accuracy among applied ML algorithms.

3.3 Accuracy Assessment

The area under the curve (AUC) and the receiver operating characteristics (ROC) curve have been used to evaluate the validity of the model. The AUC value has been utilized to represent the model's predictive capacity and is regarded as a means of validating and contrasting models. AUC-ROC and Root Mean Square Error methods are used to assess and compare the flood susceptibility prediction capabilities of the models. The models' AUC-ROC and (RMSE) findings are displayed below.

Table 1 Accuracy assessment for Training data

	AUC	**RMSE**
SVM	0.78	0.32
CatBoost	0.86	0.21
XGBoost	0.92	0.17

Table 2 Accuracy assessment for Validating data

	AUC	**RMSE**
SVM	0.65	0.29
CatBoost	0.84	0.21
XGBoost	0.91	0.18

Above tables suggest that in both training and validating phases of the data, XGBoost algorithm outperformed the other ML algorithms in terms of AUC score and RMSE.

Table 3 Class wise areas of used ML models – 2017

Classes	**RF**	**CatBoost**	**XGB**
Vey Low	5123	5364	5298
Low	9847	11240	10269
Moderate	11687	9968	10554
High	8746	9002	8954
Very High	4692	4421	4620

Table 4 Class wise areas of used ML models – 2020

Classes	**RF**	**CatBoost**	**XGB**
Vey Low	5971	4986	5476
Low	11246	10486	9856
Moderate	9645	10742	12475
High	9994	9856	8945
Very High	6201	5412	5883

Table 5 Percentage Difference of Area Classes of flood risk from 2017 to 2020

Classes	**RF (%)**	**CatBoost (%)**	**XGB (%)**
Vey Low	+16.58	-7.15	+3.50
Low	+14.16	-6.67	-3.02
Moderate	-7.47	+7.73	+18.2
High	+14.26	+9.5	-0.1
Very High	+32.18	+22.42	+27.4

Table 3 and Table 4 indicate that the area of flood risk has increased from 2017 to 2020. While table 5 clearly shows how the chances of very high-risk floods have been increased from 2017 to 2020.

3.5 LULC change and Flood risk.

Table 6 LULC Area (sq km) Changes in 2017 & 2023

	2017	**2023**
Trees	176.128	172.164
Crops	17.569	18.609
Built	17.713	20.561
Rangelands	3.183	3.261

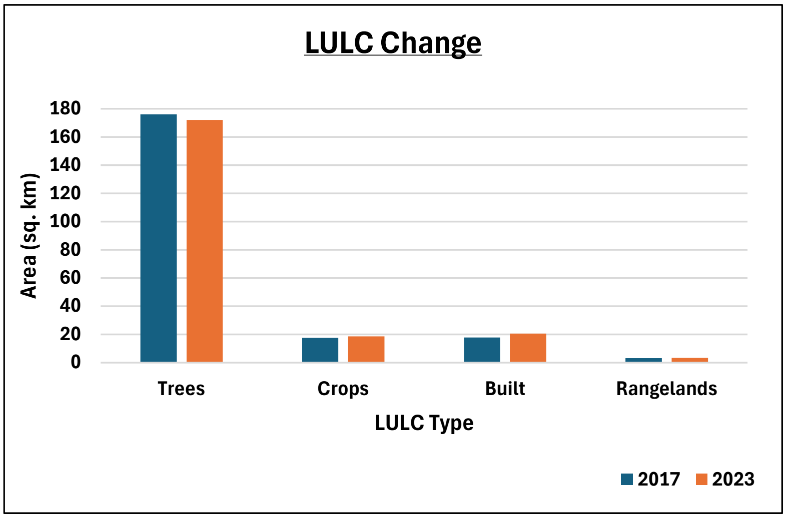

Figure 7 LULC Area changes in 2017 & 2023

3.6 Discussion

Identifying flood-related issues and creating a risk map is one of the fundamental ways for dealing with this natural disaster. Several strategies for estimating the severity of flood danger have been provided by various researchers. In this study, by implementing machine learning model XGBoost, the flood risk maps in Nilwala river catchment of Athuraliya and Akuressa DS have been produced. The study's archived results show that nine useful factors—rainfall, Topographic Wetness Index (TWI), land use, slope, LS factor, altitude, distance from rivers, curvature, and aspect—were obtained using open-source GIS data portals, the USA National Weather Services, and the Google Earth Engine cloud system. The likelihood of a flood occurring in each pixel of the domain was assessed after RF machine learning model are developed using the Python programming language after the required pre-processing and training dataset were created in ArcGIS software. The training dataset that is used has a significant impact on the machine learning models' output. Because of this, the caliber and volume of the training data are crucial. There should not be any duplicate data and the training data should be large enough to finish the learning process. Consequently, the quantity and quality of the training data available in the target area must be taken into consideration when creating a machine learning model to generate a flood risk map. Flood frequency is more influenced by Soil type, aspect TWI, and land use than by other inputs, according to an analysis of the parameters' relative relevance in the RF model.

Out of the 9 components, these four effective factors account for fifty percent of the risk of flooding. It should be mentioned that the significance of each attribute varies depending on the region, and in many studies conducted in various locations. Further research suggests that a more accurate estimate of the flood occurrence may be obtained by removing the less significant factors from the flood risk mapping and adding some additional influencing layers of information, such as soil type, river water discharge rate, land subsidence map, and population distribution map.

4. Conclusion

This study used machine learning and remote sensing models to assess flood sensitivity to the effects of climate change and land-use change, focusing on the southern part of the country, specifically the Athuraliya and Akuressa DS areas. Nine condition parameters and 989 flood points from 2017 to 2021 are used as input data in the suggested models. The findings indicate that the most significant variables influencing the likelihood of flooding in the research region were Rainfall, distance to a river, LS factor, Aspect, Curvature and LULC.

AUC values over 0.90 indicate that the influence of LULC and climate change on floods were successfully evaluated by combining machine learning techniques (RF, CatBoost, and XGB) with remote sensing data. The technique may be applied elsewhere to assess flood risk, especially in places with little data. In Akuressa and Athuraliya, regions with extremely high flood susceptibility rose by almost between 27-33% between 2017 and 2020. The precipitation and the expanded construction are the causes of the changes.

References

[1] M. Huda, N. Rather, and S. Eslamian, “Social Aspects of Flooding,” 2022, pp. 145–170. doi: 10.1201/9781003262640-8.

[2] R. Panditharathne, M. Gunathilake, I. Chathuranika, U. Rathnayake, M. Babel, and M. Jha, “Trends and Variabilities in Rainfall and Streamflow: A Case Study of the Nilwala River Basin in Sri Lanka,” Hydrology, vol. 10, pp. 1–17, Jan. 2023, doi: 10.3390/hydrology10010008.

[3] M. D. K. Gunathilaka and W. T. S. Harshana, “Climate Change Effect on the Urbanization: Intensified Rainfall and Flood Susceptibility in Sri Lanka,” 2023, pp. 49–151.

[4] S. Saravanan et al., “Flood susceptibility mapping using machine learning boosting algorithms techniques in Idukki district of Kerala India,” Urban Climate, vol. 49, p. 101503, May 2023, doi: 10.1016/j.uclim.2023.101503.

[5] M. I. Hariyono, D. Ramdani, F. E. S. Silalahi, A. A. Kurniawan, N. Indriasari, and M. Buswari, “Land use and land cover change analysis of flood prone area using remote sensing data and machine learning in Malang Raya, East Java, Indonesia,” IOP Conf. Ser.: Earth Environ. Sci., vol. 1173, no. 1, p. 012051, May 2023, doi: 10.1088/1755-1315/1173/1/012051.

[6] “Survey Department of SriLanka.” Accessed: May 29, 2024.

[7] “Climate of Sri Lanka.” Accessed: May 29, 2024.

[8] S. Ghosh, S. Saha, and B. Bera, “Flood susceptibility zonation using advanced ensemble machine learning models within Himalayan foreland basin,” Natural Hazards Research, vol. 2, no. 4, pp. 363–374, Dec. 2022, doi: 10.1016/j.nhres.2022.06.003.

[9] A. Arabameri, K. Rezaei, A. Cerdà, C. Conoscenti, and Z. Kalantari, “A comparison of statistical methods and multi-criteria decision making to map flood hazard susceptibility in Northern Iran,” Sci Total Environ, vol. 660, pp. 443–458, Apr. 2019, doi: 10.1016/j.scitotenv.2019.01.021.

[10] M. Ajmal, M. Waseem, D. Kim, and T.-W. Kim, “A Pragmatic Slope-Adjusted Curve Number Model to Reduce Uncertainty in Predicting Flood Runoff from Steep Watersheds,” Water, vol. 12, no. 5, Art. no. 5, May 2020, doi: 10.3390/w12051469.

[11] J. Lee et al., “Estimation of Real-Time Rainfall Fields Reflecting the Mountain Effect of Rainfall Explained by the WRF Rainfall Fields,” Water, vol. 15, no. 9, Art. no. 9, Jan. 2023, doi: 10.3390/w15091794.
[12] W. Yang, W. Giarè, S. Pan, E. Di Valentino, A. Melchiorri, and J. Silk, “Revealing the effects of curvature on the cosmological models,” Physical Review D, vol. 107, Mar. 2023, doi: 10.1103/PhysRevD.107.063509.
[13] A. Clemente, O. Penacchio, M. Vila-Vidal, R. Pepperell, and N. Ruta, “Explaining the curvature effect : perceptual and hedonic evaluations of visual contour,” Apr. 2023, doi: 10.31234/osf.io/pdhkf.
[14] Z. Wan-cun, “Influences of Slope and Aspect on Distribution and Change of Land Use and Cover in Daninghe River Watershed,” Journal of Soil and Water Conservation, 2004, Accessed: Apr. 22, 2024.
[15] J. Yang, Y. A. El-Kassaby, and W. Guan, “The effect of slope aspect on vegetation attributes in a mountainous dry valley, Southwest China,” Sci Rep, vol. 10, no. 1, p. 16465, Oct. 2020, doi: 10.1038/s41598-020-73496-0.
[16] S. Lee and F. Rezaie, “Data used for GIS-based Flood Susceptibility Mapping,” GEO DATA, vol. 4, pp. 1–15, Mar. 2022, doi: 10.22761/DJ2022.4.1.001.
[17] Y. Atencia, L. Rojas, J. Ramos, I. Rojas-Rodríguez, and A. Álvarez, “Use of Soil Infiltration Capacity and Stream Flow Velocity to Estimate Physical Flood Vulnerability under Land-Use Change Scenarios,” Water, vol. 15, Mar. 2023, doi: 10.3390/w15061214.
[18] M. Sisay, Catchment Morphometric Characterization of the Akaki River in the Upper Awash Sub-Basin, Ethiopia. 2022. doi: 10.21203/rs.3.rs-2007118/v1.
[19] P. Panagos, P. Borrelli, and K. Meusburger, “A New European Slope Length and Steepness Factor (LS-Factor) for Modeling Soil Erosion by Water,” Geosciences, vol. 5, no. 2, Art. no. 2, Jun. 2015, doi: 10.3390/geosciences5020117.
[20] “Global Compound Floods from Precipitation and Storm Surge: Hazards and the Roles of Cyclones in: Journal of Climate Volume 34 Issue 20 (2021).” Accessed: May 29, 2024. [Online].
[21] S. Jiang and J. Zscheischler, “Revealing how precipitation extremes impact river floods in a warming climate with interpretable machine learning,” p. EGU-2343, May 2023, doi: 10.5194/egusphere-egu23-2343.
[22] T. Chen and C. Guestrin, “XGBoost : Reliable Large-scale Tree Boosting System,” 2015. Accessed: Jan. 10, 2024.
[23] G. Biau and E. Scornet, “A Random Forest Guided Tour,” TEST, vol. 25, Nov. 2015, doi: 10.1007/s11749-016-0481-7.
[24] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: unbiased boosting with categorical features,” in Advances in Neural Information Processing Systems, Curran Associates, Inc., 2018. Accessed: May 29, 2024.
[25] M. Schonlau and R. Y. Zou, “The random forest algorithm for statistical learning,” The Stata Journal, vol. 20, no. 1, pp. 3–29, Mar. 2020, doi: 10.1177/1536867X20909688.
[26] V. Kulkarni and P. K. Sinha, “Effective Learning and Classification using Random Forest Algorithm,” Jun. 2014. Accessed: Jan. 10, 2024.
[27] Z. Arif Ali, Z. H. Abduljabbar, H. A. Tahir, A. Bibo Sallow, and S. M. Almufti, “eXtreme Gradient Boosting Algorithm with Machine Learning: a Review,” ACAD J NAWROZ UNIV, vol. 12, no. 2, pp. 320–334, May 2023, doi: 10.25007/ajnu.v12n2a1612.