ENGINEER - Vol. LV, No. 02, pp. [01-09], 2022

DOI: http://doi.org/10.4038/engineer.v55i2.7503

Application of GIS and Logistic Regression for Flood Susceptibility Mapping in Nilwala River Basin, Sri Lanka

H.D. Abeysiriwardana and N.T.S. Wijesekera

Abstract: Flood susceptibility mapping is a crucial element in flood management. This study aimed to investigate the applicability of logistic regression (LR) and GIS tools to produce a flood susceptibility map for Nilwala river basin in Sri Lanka. Four conditioning factors: elevation, slope angle, distance from river and land-use, were selected as flood conditioning factors. Flood inventory database consisted of 205 flood and 199 non-flood locations; out of these, 70% was selected as the training dataset, and the remaining 30% was taken as the testing dataset. The flood susceptibility map was developed in ArcGIS 10.5, based on the LR coefficients derived in R statistical language. The four flood conditioning factors were statistically significant at $P < 0.05$. The Receiver Operating Characteristics curve method was used to validate the model. Area under the Success Rate Curve and Prediction Rate Curve were 86% and 87%, respectively. The Area Under the Curves values reveal the model's excellent compatibility and predictability. Therefore, a high degree of confidence can be placed on the model to identify flood vulnerable areas in Nilwala basin. Hence, the developed susceptibility map might be a vital decision-making tool for water managers.

Keywords: Flood susceptibility, Slope stability, GIS, Logistic regression, Area under the curve

1. Introduction

Flood is a major destructive natural disaster which causes substantial economic losses and casualties globally every year [1]. Sri Lanka experiences flood events annually, mainly during the two monsoon seasons—the Southwest (May–September) and the Northeast (November–February) monsoons [2]. Intense extreme rainfall events striking due to pressure variations in the Bay of Bengal is also a flood triggering phenomenon in Sri Lanka [3]. Every year, most natural hazards related property damages and fatalities have been reported due to floods. Over the five years from 2016 – 2020, Sri Lanka was affected by catastrophic floods in 2019 (16 deaths), 2018 (resulted in 24 deaths) and 2017 (resulted in more than 210 deaths) [4]. Turner and Annamalai [5] have predicted a rise in seasonal precipitation and extreme rainfall events over South India and Sri Lanka. Thus, disastrous floods are likely to be increased with climatic changes and anthropogenic activities. These factors emphasise the utmost importance of flood forecasting and preparedness, in which flood susceptibility mapping plays a significant role.

Flood susceptibility analysis involves identifying flood causative factors and their contribution to flood occurrence [6], [7]. Different methods, such as Frequency Ratio [8], Logistic Regression [9], [10], Random Forest [11], and Analytical Hierarchy Process [12], [7] have been widely used in flood susceptibility analysis to derive the relationship between flood occurrences and selected causative factors. Recently, with the development of Geographical Information Systems (GIS) technologies, many researchers [8], [9], [10], [12], [13] have employed GIS-based techniques in flood susceptibility mapping.

Eng. (Ms.) H.D. Abeysiriwardana, B.Sc. Eng. (Hons) (Moratuwa), PG.Dip. (Moratuwa), MSc. (Manchester), Lecturer, Sri Lanka Institute of Information Technology (SLIIT)
Email:hdabeysiriwardana@gmail.com
https://orcid.org/0000-0002-0858-0156
Eng. (Prof.) N.T.S. Wijesekera, C.Eng., FIE(SL), MICE(UK), B.Sc. Eng.Hons. (Sri Lanka), P.G. Dip. (Moratuwa), M.Eng.(Tokyo), Ph.D.(Tokyo), Chairman Construction Industry Development Authority (CIDA)
Email:ohanw2@gmail.com
https://orcid.org/0000-0003-0964-4331

In the tropics, extreme rainfall events are the primary trigger of both pluvial and fluvial floods. A fluvial flood is the inundation of nearby lands due to overflowing of waterbodies. Pluvial flood occurs when rainfall intensity exceeds infiltration capacity of the ground irrespective of presence of a water body [14]. The factors that could make an area fundamentally predisposed to floods, such as topography (e.g., elevation, slope), geology, land-use, river network density, distance from water bodies and soil type [7], are referred to as conditioning factors hereafter. Since the rainfall changes both temporally and spatially, in this study, flood susceptibility was evaluated based only on conditioning factors. This basin-scale study was done in Nilwala river basin, which is frequently subjected to both pluvial and fluvial floods. The selection of conditioning parameters was based on the resolution of the available data. For example, resolutions of soil and geology maps of Sri Lanka are coarser for a basin-scale study. In this study, logistic regression - a multivariate analysis that predicts the probability of a discrete dependent variable (e.g., in this case, presence or absence of floods, usually assigned 1 and 0 respectively, for numerical representation) based on a set of interdependent variables, was employed together with GIS technology [15].

2. Study Area

Nilwala river basin, with an area of about 1032 km^2, is in Southern Sri Lanka and falls onto the wet zone of the country. The basin topography is characterised by flat lands, mountains, hills, and valleys. The altitude ranges from 0 to 925 m above mean sea level. The slope angle varies from 0 to 62 degrees. The upper reach of the catchment receives a mean annual rainfall about 3000 mm, while in the lower reaches the mean annual rainfall is about 1900 mm. Most rain comes in March - June and August - December, with an average monthly rainfall exceeding 200 mm [16]. Soil types in the watershed include Red-Yellow podzolic soils, Alluvial soils, Bog/half-dog soils, beach, and dune sands. The land-use consists of paddy, tea, rubber, coconut, marshy lands, forests, grasslands, water bodies and residential areas, whereas residential areas cover about 30% of the total watershed.

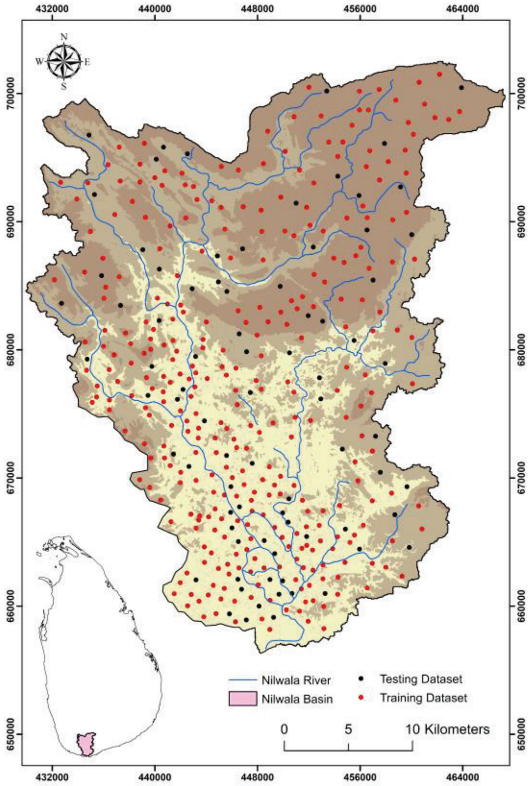

Figure 1 - Study Area - Nilwala River Basin, Sri Lanka

3. Materials and Methods

3.1 Flood Conditioning Factors

Considering the relevance to flood susceptibility, four factors: elevation, slope degree, land-use and distance from river (Figure 2), were selected based on the literature review. Many previous studies [10], [17], [18] have considered elevation as one of the most influencing factors of flood. As water flows from higher to lower elevations, lower elevations are more vulnerable to flooding [18]. Slope angle is another factor that was considered by many researchers [6], [10], [18]. Slope affects infiltration and velocity of rainfall-runoff. Steeper the slope, lesser the time for infiltration and higher the velocity of water flow, making gentle slopes and flat areas vulnerable to flooding [19]. Land-use is a determining factor in both infiltration and runoff generation. For example, build up lands decrease the infiltration capacity, increasing the runoff generation [15]. Distance from river is another vital flood conditioning factor, as the lands in closer proximity to water bodies have higher flood potential even during a moderately high rainfall event [17].

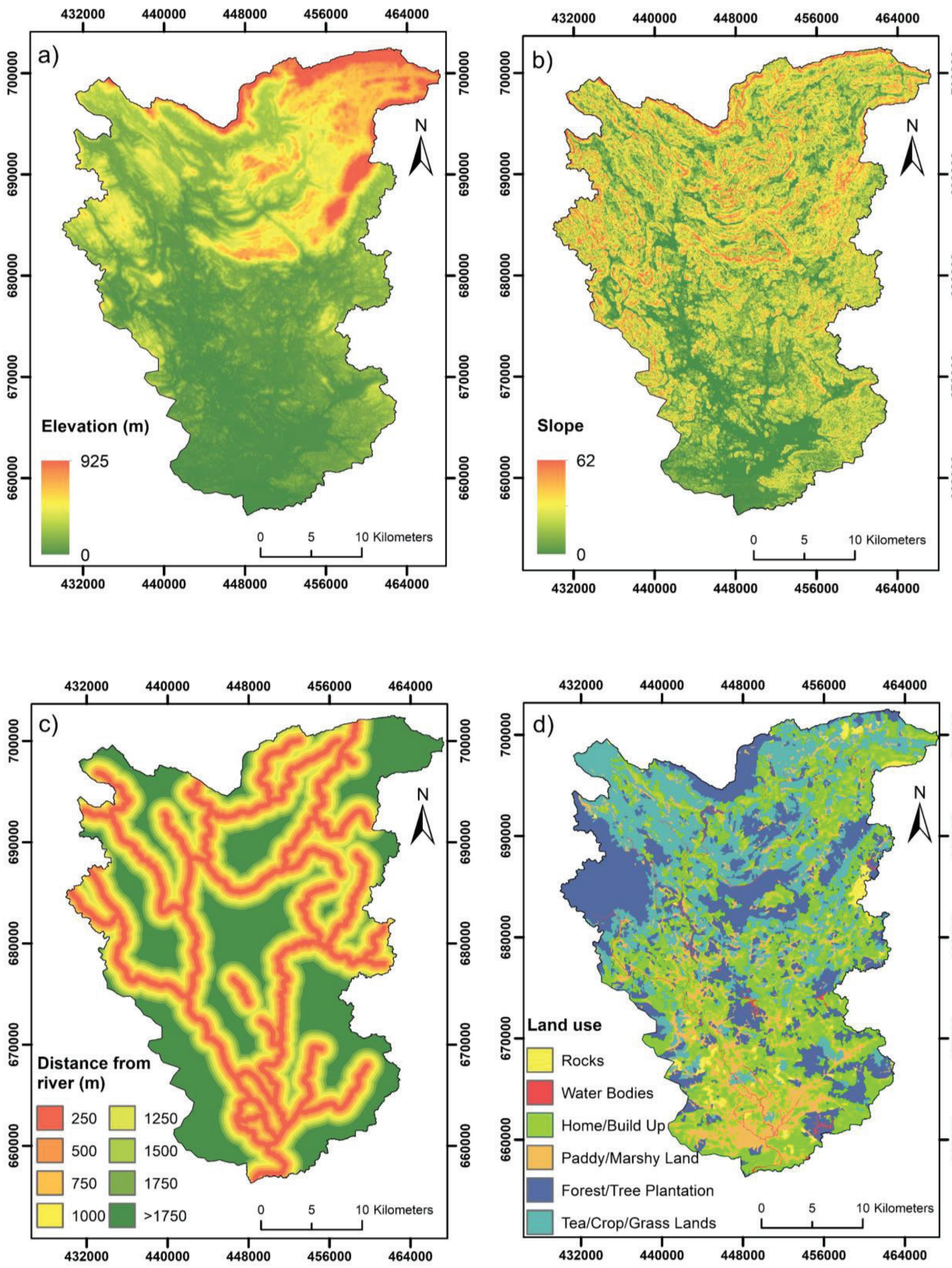

Figure 2 - Thematic Layers of the Flood Conditioning Factors a) Elevation; b) Slope degree; c) Distance from river; d) Land-use

A Digital Elevation Model (DEM) with a resolution of 30 m, the land use map of Sri Lanka and the stream network map of Sri Lanka were obtained from the Survey Department of Sri Lanka. A database of flood condition factors was constructed as a spatial database containing 30 m x 30 m grid cells. Elevation was extracted from the DEM, and slope angle was derived from the DEM, using the "slope" tool in ArcGIS. The land use map and the stream network maps were in the vector format. The land-use map was converted to a raster file and reclassified into six classes, as shown in Figure 2d. Using the stream network vector file, eight buffer zones around the stream lines were generated in ArcGIS with the buffer tool. Then the generated stream buffer vector file was converted to a raster layer to create a conditioning factor thematic layer, namely the distance from river (Figure 2c).

3.2 Flood Inventory Database

A precise flood inventory database is essential to identify a fairly accurate relationship between flood conditioning factors and flood occurrences, thus to achieve more accurate predictions [17]. In this study, a flood inventory database was constructed based on a single flood event that occurred in 2007, from 24 - 30 May, using a flood inundation map developed by the Dartmouth Flood Observatory (DFO) [20]. The model grid was converted to a point vector layer in ArcGIS. Afterwards, points for the inventory database were selected randomly, so that points represent historic flood locations, no flood locations, different elevations, different slopes, different land uses and finally covering the entire basin at an appropriate resolution. The flood inventory database consisted of 404 data points, out of which 205 were flood locations. All the spatial analyses were carried out in ArcGIS in the WGS 84 UTM Zone 44N coordinate system.

Prepared conditioning factor thematic layers were loaded into ArcGIS. Then, values of each conditioning factor were extracted to the points in the flood inventory database, then saved as an excel database to perform logistic regression analysis. The dataset was randomly split into two: 70% as the training data set which was used to calibrate the model, and 30% as the testing data set which was used to validate the model [17], [19].

3.3 Logistic Regression

Logistic regression predicts the probability of a binary dependent variable based on a set of independent variables, which can be either binary or scalar [21]. In this study, the dependent variable represents the presence or absence of flood, and four conditioning factors were taken as the independent variables. The logistic regression model can be written as in Equation (1).

$$p = \frac{1}{1 + e^{-z}} \quad \text{... (1)}$$

where p is the probability of flood occurrence, which varies from 0 to 1; z is the linear combination of independent variables, which varies from -∞ to +∞, and denotes as in Equation (2).

$$z = b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n \quad \text{... (2)}$$

where, b_0 is the intercept of the model, b_i (i = 1, 2, ……, n) are the corresponding coefficients to the respective independent variables, n is the number of independent variables, and x_i (i = 1, 2, ……, n) are the independent variables.

The statistical programming language R (currently developed by the R Development Core team) was used for logistic regression analysis. R comes with many packages that provide a coherent and integrated collection of functions and tools for data analysis. This study used the "glm()" function available in the "ISLR package" in R to develop the logistic regression model. The general use of the "glm()" function is to fit a generalised linear model; however, when utilising the specific argument - "family = binomial" - in the function, the function instructs R to fit a logistic model [22].

3.4 Model Validation

The logistic regression model was validated using Receiver Operating Characteristics (ROC) curve method [6], [9], [13], [17], [18]. A ROC curve plots True Positive Rate (TPR) (also known as sensitivity) as a function of False Positive Rate (FPR) (equals to 1- specificity) [23]. TPR and FPR can be calculated as in Equations (3) and (4).

$$TPR = \frac{TP}{TP + FN} \quad \text{... (3)}$$

$$FPR = \frac{FP}{TN + FP} \quad \text{... (4)}$$

Where, TP is true positive, TN is true negative, FP is false positive, and FN is false negative. In ROC, the Area Under the Curve (AUC) quantifies the performance of the model. AUC of the Prediction Rate Curve (PRC), which is based on the testing dataset, validates the model, while the AUC of the Success Rate Curve (SRC), which is based on the training dataset, evaluates goodness of fit, assuming that the model is correct. AUC generally ranges from 0.5 to 1 [24]. Based on AUC values, the model performances can be categorised as: average if 0.6 < AUC < 0.7, good if 0.7 < AUC < 0.8, very good if 0.8 < AUC < 0.9 and excellent if 0.9 < AUC < 1 [9]. Serval R packages facilitate generating of ROC curves, and the "pROC" package [22] was utilised for this study.

3.5 Mapping in GIS

Finally, the conditioning factor thematic layers were overlaid using the raster calculator in ArcGIS based on the predictive model equation (Equation (7)). Then, the generated flood susceptibility raster layer was divided into five susceptibility classes based on the probabilities and using the natural breaks method.

The flood susceptibility map was cross-checked with the land-use map to investigate the correspondence of flood susceptibility classes with land-use categories. Land-use wise composition of each flood susceptibility class was calculated using Equation (5). Further, the share of each land-use class corresponding to each flood susceptibility class out of the total area of the respective land-use class in the watershed was calculated using Equation (6).

$$\%FC_i = \frac{ALu_j}{AFC_i} \times 100 \qquad \ldots (5)$$

where, $\%FC_i$ is the percentage area composition (%) of the flood susceptibility class i, i is the index denoting the flood susceptibility classes ($i = 1 - 5$), ALu_j is the area of the land-use class j in the flood susceptibility, class i, j is the index denoting the land-use classes, and AFC_i is the total area of the flood susceptibility class i.

$$\%Lu_jFC_i = \frac{ALu_jFC_i}{TALu_j} \times 100 \qquad \ldots (6)$$

where, $\%Lu_jFC_i$ is the percentage of the land-use class j in the flood susceptibility class i, ALu_jFC_i is the area of the land-use class j in the flood susceptibility class i, $TALu_j$ is the total area of the land-use class j in the watershed.

4. Results and Discussion

The coefficients generated by the logistic regression analysis are shown in Table 1. All the independent variables had p-values lesser than 0.05 ($p < 0.05$), indicating all the variables are statistically significant in determining flood susceptibility. The generated success and prediction curves are shown in Figure 3. AUC values for the SRC and PRC are 0.860 (86%) and 0.870 (87%), respectively. According to the ROC classification mentioned in Al-Juaidi et al. [9], the AUC value of the SRC curve suggests a very good fit. Likewise, the AUC value of the PRC curve suggests an excellent predictive power. The predictive equation can be written as:

$$FS = \frac{1}{1 + e^{-(0.724 - 0.08S - 0.076E - 0.119D_r + 0.443Lu)}} \qquad \ldots (7)$$

Where, FS is flood susceptibility, S is slope, E is elevation, D_r is distance from river, and Lu is land-use type.

A negative coefficient indicates an inverse relationship between the corresponding conditioning factor and flood occurrence. Agreeing with previous similar studies [7], [9], [18], [25], elevation takes a negative coefficient, indicating that flooding is unlikely at higher elevations. The negative coefficient of slope implies that flatlands and gentle slopes are highly vulnerable to flooding and agrees with the findings of Vojtek and Vojteková [7] and Cao et al. [18]. The effects of fluvial floods, the most common in the study basin, decrease as the distance from the river increases. This relationship is vibrantly indicated by the negative coefficient assigned to D_r.

Furthermore, when comparing Figures 2 and 4, it is evident that very high to moderate susceptibility classes are mainly located at lower elevations (less than 30 m) and flatlands. Low susceptibility classes can be primarily found at elevations from 30 m to 100 m and lower elevations in the coastal region which are more than 2 km away from the river. High grounds (more than 100 m) with higher slope angles are the least susceptible to floods. Although the very high elevations categorised as very low susceptibility could be interpreted as no flood risk zones, classifying susceptibility maps from no flood to high flood has not been practised in the literature. The study was adhered to the general practice in susceptibility mapping which is to classify susceptibility map from (very) low to (very) high [7], [9], [18].

Table 1 - Coefficients of Logistic Regression Model

	Coefficient	Pr (>|z|)
Intercept	0.724	1.10×10^{-1}
Slope	-0.008	8.46×10^{-7}
Elevation	-0.076	2.18×10^{-4}
Distance from river	-0.119	3.89×10^{-2}
Land use	0.443	2.07×10^{-7}

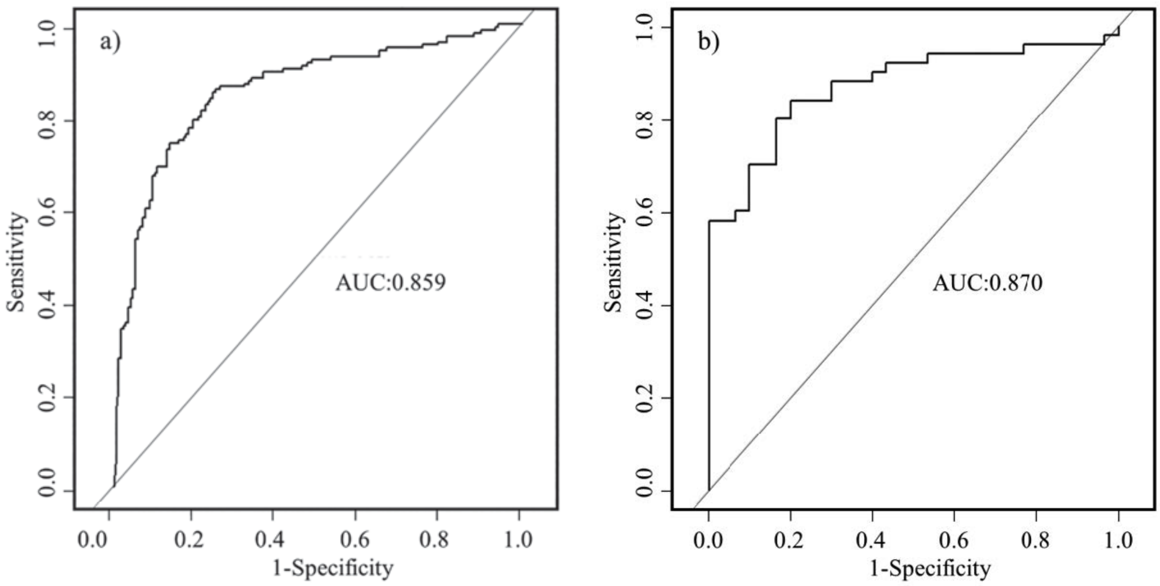

Figure 3 - Results of the ROC Analysis: a) Success Rate Curve using the Training Dataset; b) Predicted Rate Curve using the Validating Dataset

Table 2 - The Classification System used to Categorise Flood Susceptibility Map and the Area Coverage of each Susceptibility Class

Susceptibility Class	Area (km^2)	Area coverage (%)
Very Low	560.8	54.3
Low	300.2	29.0
Moderate	77.2	7.5
High	66.8	6.5
Very High	27.4	2.7

Table 3 - Land use wise composition of moderate, high, and very high flood susceptibility classes

Flood Class	Area Composition (%)				
	Residential	Cultivated Lands			
		Paddy	Coconut	Rubber	Tea
Moderate	21.00	38.00	5.00	4.00	4.00
High	5.00	76.00	0.40	0.40	0.60
Very high	3.00	83.00	0.20	0.12	0.07

Approximately 80% of the total study area has been marked as very low to low susceptible classes. The very high susceptibility class covers only about 3% of the total watershed. High and moderate susceptibility areas accounted for nearly 6.5% and 7.5% of the catchment (Table 2). Therefore, only about 14% of the total Nilwala basin has been characterised by moderate to very high flood susceptibility. Nevertheless, it is paramount to investigate the land-use composition of areas identified as moderate to very high susceptible for efficient flood management. The land-use wise composition of moderate, high, and very high flood susceptibility classes is shown in Table 3. About 83%, 76% and 38% of the very high, high, and moderate susceptibility classes comprise paddy fields. Approximately 21% of moderate susceptible lands is residential areas, which emphasises the need for proper preparedness for future flood events. As per Table 4, about 7% of residential areas in the watershed is moderate to very highly vulnerable to flooding. As for the cultivations, percentages of tea, rubber, and coconut in moderate to very high flood susceptibility classes are about 2%, 12% and 7%, respectively. Paddy is the most vulnerable cultivation to flood, with 15% of total paddy fields in the watershed has been fallen into the very high flood susceptibility class, followed by 33.4% and 20% of total paddy lands being in the high and moderate susceptibility classes. This necessitates proper flood mitigation approaches in the Nilwala basin to minimise the damages to paddy cultivation.

Table 4 - Area Coverage of Land-use Classes Corresponding to Moderate, High, and very High Flood Susceptibility Classes

Land-use class	Total Area (km²)	Area (%) of each flood class		
		Moderate	High	Very High
Residential	307.61	5.40	1.00	0.30
Paddy	151.21	20.00	33.40	15.00
Coconut	59.98	6.30	0.50	0.09
Rubber	28.94	10.60	0.85	0.10
Tea	153.23	2.00	0.30	0.01

The model outputs and the model validation results proved that the method employed by this study could be applied with a high degree of confidence for future flood mitigation as a preliminary planning tool – for example, to identify the priority areas that require evacuation or to identify suitable relief centre locations for future flood events. In addition, the model can be slightly modified utilising the road network layer and GIS tool to identify possible evacuation routes during flood disasters. When constructing the flood inventory, the study considered only one flood event that occurred in 2017. Hence, it should be noted that the generated flood susceptibility map is mainly valid for similar flood events. This limitation could be prevented with further improvements to the model. As the next step of this study, the model could be calibrated incorporating different flood events so that the model possesses the ability to predict flood susceptibility at different rainfall thresholds. As mentioned by Bernhofen et al. [26], a DEM with a resolution of 30 m is sufficient for geomorphological floodplain mapping for a catchment area of 10 km². Thus, the use of 30 m DEM for susceptibility mapping in the Nilwala basin (approximately 1030 km²) can be considered satisfactory. Hence, further improvements are also possible at the current resolution, which is also cost effective.

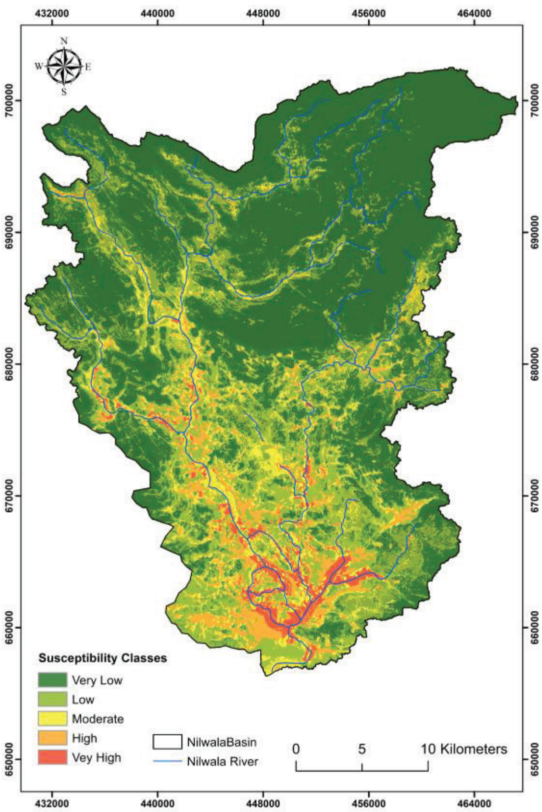

Figure 4 – Predicted Flood Susceptibility Map of Nilwala Basin

5. Conclusions

- The logistic regression analysis showed that the four conditioning factors, i.e., elevation, slope angle, distance from river and land-use, were statistically significant (at $p < 0.05$) for flood susceptibility mapping in Nilwala basin.
- The model was validated using the ROC method, which gave AUC of SRC and PRC as 0.86 and 0.87, respectively. This displayed a very good compatibility and predictability, indicating the potential to successfully classify flood susceptible areas in Nilwala basin.
- The present work showed that the classified susceptibility map coupled with the land-use map could be utilised with a high degree of confidence in flood disaster management in Nilwala basin to identify highly vulnerable residential areas which require evacuation and very low risk areas in which flood relief centres could be established.

References

1. Ceola, S., Laio, F. and Montanari, A., "Satellite Nighttime Lights Reveal Increasing Human Exposure to Floods Worldwide", *Geophys. Res. Lett*., Vol 41, No. 20, 2014, pp. 7184–7190. doi: 10.1002/2014GL061859.

2. Alahacoon, N., Matheswaran, K., Pani, P. and Amarnath, G., "A Decadal Historical Satellite Data and Rainfall Trend Analysis (2001–2016) for Flood Hazard Mapping in Sri Lanka", *Remote Sens*., Vol. 10, No. 3, 2018, pp. 448. doi: 10.3390/rs10030448.

3. Samarasinghe, S. M. J. S., Nandalal, H. K., Weliwitiya, D. P., Fowze, J. S. M., Hazarika, M. K. and Samarakoon, L., "Application of Remote Sensing and GIS for Flood Risk Analysis: A Case Study at Kalu- Ganga River, Sri Lanka" *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, ISPRS Technical Commission VIII Symposium Networking the World with Remote Sensing*, Kyoto, Japan, Vol. XXXVIII, No. 8, August 9–12, 2010, pp. 110 – 114.

4. Center for Excellence in Disaster Management & Humanitarian Assistance (CFE-DM), Sri Lanka Disaster Management Reference Handbook, March 2021.

5. Turner, A. G. and Annamalai, H., "Climate Change and the South Asian Summer Monsoon", Nat Clim Chang., Vol. 2, No. 8, 2012, pp. 587–595. doi: 10.1038/nclimate1495.

6. Shafizadeh-Moghadam, H., Valavi, R., Shahabi, H., Chapi, K. and Shirzadi, A., "Novel Forecasting Approaches using Combination of Machine Learning and Statistical Models for Flood Susceptibility Mapping", *J. Environ. Manage*., Vol. 217, 2018, pp. 1–11. doi: 10.1016/j.jenvman.2018.03.089.

7. Vojtek, M. and Vojteková, J., "Flood Susceptibility Mapping on a National Scale in Slovakia using the Analytical Hierarchy Process", *Wate*r, Vol. 11, No. 2, 2019, pp. 364. doi: 10.3390/w11020364.

8. Rahmati, O., Pourghasemi, H. R. and Zeinivand, H., "Flood Susceptibility Mapping using Frequency Ratio and Weights-of-Evidence Models in the Golastan Province, Iran", *Geocarto Int*., Vol. 31, No 1, 2016, pp. 42–70. doi: 10.1080/10106049.2015.1041559.

9. Al-Juaidi, A. E. M., Nassar, A. M. and Al-Juaidi, O. E. M., "Evaluation of Flood Susceptibility Mapping using Logistic Regression and GIS Conditioning Factors", *Arab. J. Geosci*., Vol. 11, No. 24, 2018, pp. 765. doi: 10.1007/s12517-018-4095-0.

10. Tehrany, M. S., Kumar, L., Jebur, M. N. and Shabani, F., "Evaluating the Application of the Statistical Index Method in Flood Susceptibility Mapping and Its Comparison with Frequency Ratio and Logistic Regression Methods", *Geomatics, Nat. Hazards Risk*., Vol. 10, No. 1, 2019, pp. 79–101. doi: 10.1080/19475705.2018.1506509.

11. Chen, W., Li, Y., Xue, W., Shahabi, H., Li, S., Hong, H., Wang, X., Bian, H., Zhang, S., Pradhan, B., and Ahmad, B. B., "Modeling Flood Susceptibility using Data-Driven Approaches of naïve Bayes tree, Alternating Decision Tree, and Random Forest Methods", *Sci. Total Environ*., Vol. 701, 2020, pp. 134979. doi: 10.1016/j.scitotenv.2019.134979.

12. Hammami, S., Zouhri, L., Souissi, D., Souei, A., Zghibi, A., Marzougui, A., and Dlala, M., "Application of the GIS Based Multi-Criteria Decision Analysis and Analytical Hierarchy Process (AHP) in the Flood Susceptibility Mapping (Tunisia)", *Arab. J. Geosci*., Vol. 12, No.21, 2019, pp. 653. doi: 10.1007/s12517-019-4754-9.

13. Tehrany, M. S., Pradhan, B. and Jebur, M. N., "Flood Susceptibility Mapping using a Novel Ensemble Weights-of-Evidence and Support Vector Machine Models in GIS", *J. Hydrol*., Vol. 512, 2014, pp. 332–343. doi: 10.1016/j.jhydrol.2014.03.008.

14. Tanaka, T., Kiyohara, K. and Tachikawa, Y., "Comparison of Fluvial and Pluvial Flood Risk Curves in Urban Cities Derived from a Large Ensemble Climate Simulation Dataset: A Case Study in Nagoya, Japan", *J. Hydrol*., Vol. 584, 2020, pp. 124706. doi: 10.1016/j.jhydrol.2020.124706.

15. Chapi, K., Singh, V. P., Shirzadi, A., Shahabi, H., Bui, D. T., Pham B. T. and Khosravi, K., "A Novel Hybrid Artificial Intelligence Approach for Flood Susceptibility Assessment", *Environ Model Softw*, Vol. 95, 2017, pp. 229–245. doi: 10.1016/j.envsoft.2017.06.012.

16. Ratnayake, U., Sachindra, D. A. and Nandalal, K. D. W., "Rainfall Forecasting for Flood Prediction in the Nilwala Basin", *International Conference on Sustainable Built Environment (ICSBE-2010)*, Kandy, 2010, pp. 13-14, December 2010, pp. 355 – 362.

17. Tehrany, M. S., Pradhan, B., Mansor, S., and Ahmad, N., "Flood Susceptibility Assessment using GIS-Based Support Vector Machine Model with different Kernel Types", *CATENA*, Vol. 125, 2015, pp. 91–101. doi: 10.1016/j.catena.2014.10.017.

18. Cao, C., Xu, P., Wang, P., Chen, J., Zheng, L. and Niu, C., "Flash Flood Hazard Susceptibility Mapping using Frequency Ratio and Statistical Index Methods in Coalmine Subsidence Areas", *Sustainability*, Vol. 8, No. 9, 2016, pp. 948. doi: 10.3390/su8090948.

19. Khosravi, K., Nohani, E., Maroufinia, E and Pourghasemi H. R., "A GIS-Based Flood Susceptibility Assessment and its Mapping in Iran: A Comparison Between Frequency Ratio and Weights-of-Evidence Bivariate Statistical Models with Multi-Criteria Decision-making Technique", *Nat Hazards*, Vol. 83, No. 2, 2016, pp. 947–987. doi: 10.1007/s11069-016-2357-2.

20. Dartmouth Flood Observatory (DFO), "DFO Flood Event", Dartmouth Flood Observatory, University of Colorado, Boulder, Colorado, USA, 2017, https://floodobservatory.colorado.edu/Events/2017SriLanka4474/2017SriLanka.html, accessed on 10. 04. 2021.

21. Menard, S., "Applied Logistic Regression Analysis. Sage University Paper Series on Quantitative Applications in Social Sciences", Thousand Oaks, California, Vol. 106, 1995, pp: 99.

22. STHDA, "Articles - Classification Methods Essentials", Statistical tools for high-throughput data analysis, 2021, http://www.sthda.com/english/articles/36-classification-methods-essentials/151-logistic-regression-essentials-in-r/. Accessed 25.04.2021.

23. Oh, H.-J. and Lee, S., "Assessment of Ground Subsidence using GIS and the Weights-of-Evidence Model", *Eng Geol*., Vol. 115, No. 1–2, 2010, pp. 36–48. doi: 10.1016/j.enggeo.2010.06.015.

24. Arabameri, A., Saha, S., Roy, J., Chen, W., Blaschke, T., and Bui, D. T., "Landslide Susceptibility Evaluation and Management using Different Machine Learning Methods in the Gallicash River Watershed, Iran", *Remote Sens*., Vol. 12, No. 3, 2020, pp. 475. doi: 10.3390/rs12030475.

25. Botzen, W. J. W., Aerts, J. C. J. H., and Van den Bergh, J. C. J. M. "Individual Preferences for Reducing Flood Risk to Near Zero through Elevation", *Mitig. Adapt. Strateg. Glob. Chang*., Vol. 18, 2013, pp. 229–244. https://doi.org/10.1007/s11027-012-9359-5.

26. Bernhofen, M. V., Trigg, M. A., Sleigh, P. A., Sampson, C. C., and Smith, A. M. "Global Flood Exposure from Different Sized Rivers". Nat. Hazards Earth Syst. Sci. 2021. https://doi.org/10.5194/nhess-2021-102.