African Journal of Education Science and Technology (AJEST) Vol. 8 No. 4 (May, 2026)
University of Eldoret, Kenya, Mount Kenya University, Kenya, Chukwuemeka Odumegwu Ojukwu University, Nigeria, Kyambogo University, Uganda and University of Makeni, Sierra Leone

Flood Susceptibility Assessment Using Geospatial Techniques and Predictive Modeling in Bunyala Sub-County, Busia County, Kenya

Magdaline Perpetua Agoya[1], Benjamin Mwasi[1], Nyaberi Daniel Mogaka[2]

[1] Department of Environmental Planning, Sustainability & Geoinformatics, University of Eldoret, Eldoret, Kenya

[1] Department of Environmental Planning, Sustainability & Geoinformatics, University of Eldoret, Eldoret, Kenya

[2] Department of Environmental Science, University of Eldoret, Eldoret, Kenya

*Corresponding Author**:** perpetua.agoya@gmail.com

Abstract

Flooding remains one of the most destructive natural hazards globally and is a major contributor to disaster-related losses in Kenya. Bunyala Sub-County in Busia County is particularly vulnerable due to recurrent flooding associated with River Nzoia, River Yala, and Lake Victoria backflow. Despite previous geospatial studies in the area, limited attention has been given to temporal changes in flood zones, the influence of flood drivers, and the application of predictive models in flood susceptibility mapping. This study assessed flood susceptibility in Bunyala Sub-County using Geographic Information Systems (GIS), Remote Sensing, and predictive modeling. Multi-temporal Landsat imagery (2000, 2010, and 2020), Digital Elevation Model (DEM), rainfall, GPS, and OpenStreetMap datasets were processed using ArcGIS Pro, Google Earth Engine, and SPSS to map flood zones and evaluate the influence of slope, elevation, distance to water bodies, rainfall, and land use/land cover (LULC). Logistic regression was used to determine the relationship between flood occurrence and the selected flood drivers, while a weighted overlay approach was applied to generate a Flood Susceptibility Index (FSI) raster map, which was further reclassified into a final flood susceptibility map. Results revealed a net increase of +3.70 km² in flood zones between 2000 and 2020, with a decline of -8.72 km² between 2000 and 2010 followed by an increase of +12.42 km² between 2010 and 2020. LULC emerged as the strongest-associated flood driver in the model, followed by slope and distance to water bodies. The LR model achieved an overall classification accuracy of 78.7% and an AUC value of 0.841, indicating good predictive performance. The final flood susceptibility map classified 62% of the study area as high susceptibility, 34% as moderate susceptibility, and 4% as low susceptibility. The study demonstrates the effectiveness of integrating GIS, remote sensing, and predictive modelling in flood susceptibility assessment and provides valuable information for land use planning, flood risk management and climate adaptation in flood-prone areas.

Keywords: *Flood Susceptibility; Logistic Regression; Remote Sensing; Land Use/Land Cover (LULC); Geographic Information Systems (GIS).*

Introduction

Flooding is a major environmental challenge affecting populations, livelihoods and infrastructure, particularly in low-lying areas. Globally, over two billion people were affected by floods between 1998 and 2017, with 32.3 million people affected in 2023 alone (IRDR, 2023). In East Africa, frequent flooding has caused major displacement and damage to infrastructure, including around Lake Victoria. In Kenya, the western region, particularly the River Nzoia catchment, experiences frequent flooding, while projected changes in rainfall patterns may increase flood risks in downstream areas (Onywere et al., 2013).

Bunyala Sub-County (Budalangi Constituency) in Busia County is notably vulnerable because of its location within a low-lying floodplain and proximity to the Rivers Nzoia and Yala and Lake Victoria. Flooding occurs frequently, with river flooding and Lake Victoria backflow contributing to inundation. Severe flooding in 2020 caused widespread displacement and damage within the floodplain (Wanzala, 2020), highlighting the need for improved assessment of flood patterns and areas susceptible to flooding.

Geospatial techniques provide effective methods for flood extent mapping and analyzing its spatial and temporal variation. Predictive modelling complements these techniques by evaluating relationships between observed flood occurrence and environmental factors such as land use/land cover (LULC), slope, elevation, distance from water bodies and rainfall. Understanding these factors and their influence is important in determining flood occurrence and developing effective flood management and early warning mechanisms.

Several studies have employed geospatial technologies to map flooding in Bunyala Sub-County (Mutiso, 2011; Gaya, 2019; Nabwire, 2019; Likuyi, 2018, 2025). However, critical gaps remain in understanding how flood zones have changed over time, their relationship with selected flood drivers and predictive modelling of future flood-zone patterns. This study therefore analyzed flood zone changes between 2000, 2010 and 2020, established the relationship between observed flood zones and selected flood drivers, and predicted future flood zone patterns based on the prevailing conditions in selected factors using geospatial and predictive modelling techniques in Bunyala Sub-County, Kenya. The study contributes to flood risk management by providing an updated flood susceptibility map to support disaster preparedness, land use planning and climate adaptation in flood-prone areas.

Materials and Methods

Study Area

The study was carried out in Bunyala Sub-County (Budalangi Constituency), Busia County, Kenya, located between latitudes 0°30′S and 0°11′30″N and longitudes 33°56′30″E and 34°10′30″E. The area forms part of the lower River Nzoia basin and is characterized by low-lying terrain, extensive floodplains, and proximity to Rivers Nzoia and Yala as well as Lake Victoria. The region experiences recurrent flooding associated with river overflows and Lake Victoria backflow, making it one of the most flood-prone areas in Kenya.

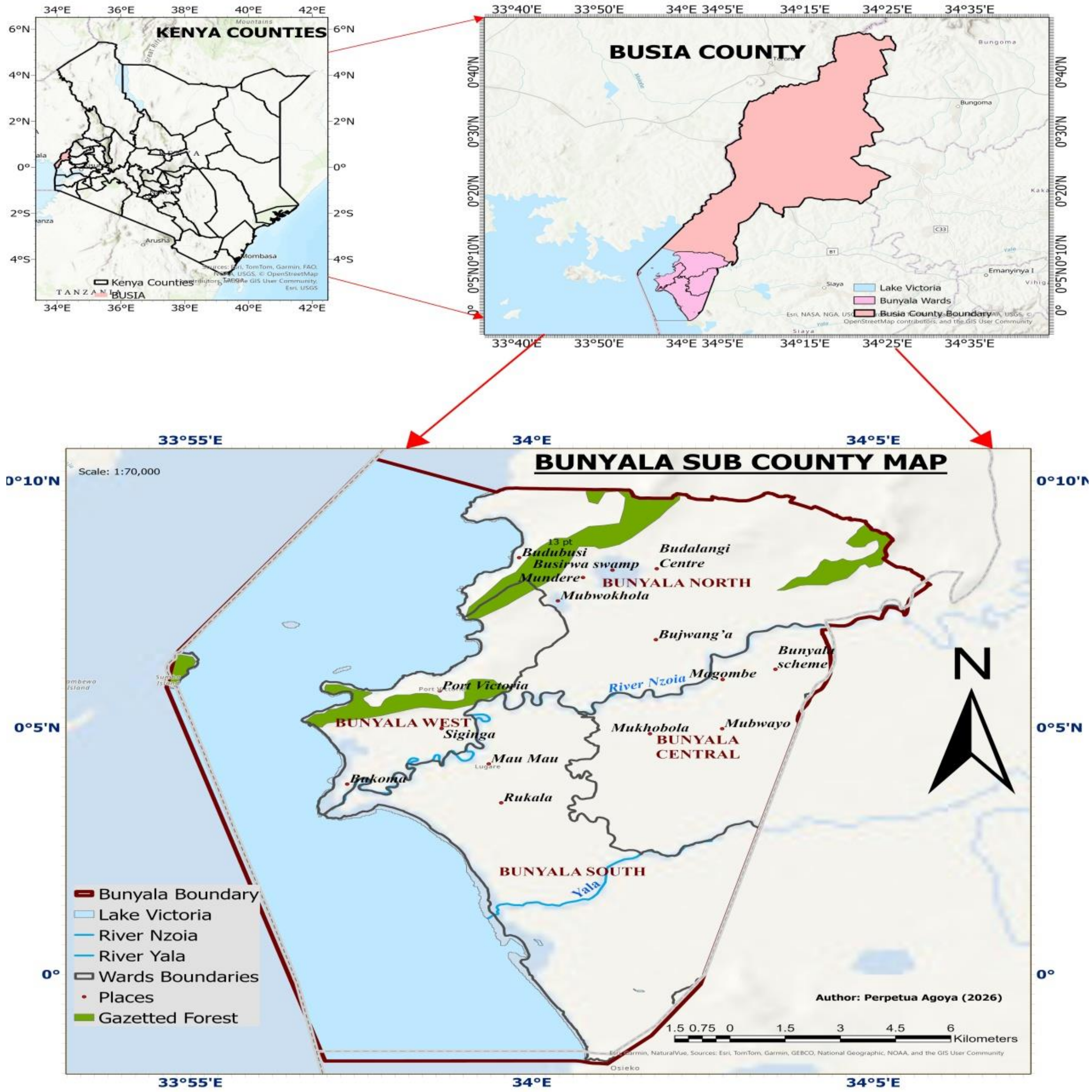


Figure 4: Location of Bunyala Sub County (Budalangi), Busia County, Kenya

Data Sources

Multi-source geospatial datasets were used, including Landsat satellite imagery, Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM), Climate Hazards Group InfraRed Precipitation with Station Data (CHIRPS), OpenStreetMap (OSM), and GPS ground-truth data. The characteristics of the Landsat imagery are summarized in Table 1.

Table 5: **Summary of satellite imagery used in the study**

Year	Satellite & Sensor	Source	Bands Used	Cloud Cover (%)	Spatial Resolution	Processing Level	Date Acquired
2000	Landsat 7 (ETM+)	USGS Earth Explorer	B4 (NIR), B3 (Red), B2 (Green)	<10%	30 m	Level-2 Surface Reflectance	06/03/2000
2010	Landsat 5 (TM)	USGS Earth Explorer	B4 (NIR), B3 (Red), B2 (Green)	<10%	30 m	Level-2 Surface Reflectance	21/01/2010
2020	Landsat 8 (OLI)	USGS Earth Explorer	B5 (NIR), B4 (Red), B3 (Green)	<10%	30 m	Level-2 Surface Reflectance	18/02/2020

Image Processing and Land Use/Land Cover Classification

Landsat images for 2000, 2010, and 2020 preprocessing involved cloud masking, band selection, clipping to the study area boundary, and projection to WGS 84 UTM Zone 36N. False Colour Composite (FCC) images were generated to enhance visual interpretation. LULC classification was performed using supervised classification with the Random Forest algorithm in Google Earth Engine. Five classes were identified: water bodies, bareland, natural vegetation, agriculture, and built-up areas (Table 2).

Table 6: Original LULC classes

Code	Class
0	Water Bodies
1	Bareland
2	Natural Vegetation
3	Agriculture
4	Built-up Areas

Classification accuracy was assessed using GPS-derived ground truth points and confusion matrices. Overall accuracy, user's accuracy, producer's accuracy, and Kappa statistics were computed to evaluate classification performance.

Flood Zone Extraction and Change Detection

Flood zones were extracted from the classified LULC maps by isolating the water class and removing permanent water bodies including Lake Victoria and Rivers Nzoia and Yala using OSM hydrographic data. The resulting binary flood maps were used to quantify flooded areas for 2000, 2010, and 2020.

Flood zone changes were determined by comparing flood extents between 2000-2010 and 2010-2020. Net and percentage changes were calculated to evaluate temporal flood dynamics.

Flood Drivers Analysis

Five flood drivers were selected based on their relevance to flood occurrence: elevation, slope, distance to water bodies, rainfall, and LULC. The environmental datasets used in the analysis are summarized in Table 3.

Table 7: Flood driver datasets used in the study

Data Type	Dataset / Source	Year(s) Used	Resolution	Purpose
Elevation (DEM) (m)	SRTM (USGS Earth Explorer)	Static	30 m	Elevation and slope extraction
Slope(degrees)	SRTM DEM	Static	30 m	Gradient of the terrain
Distance to Water(m)	OpenStreetMap (OSM)	Static	30 m	Hydrological proximity driver
Rainfall (mm)	CHIRPS via Google Earth Engine (March/May) season	2020	~5 km	Climatic driver
LULC	Classified LULC Image (Objective 1)	2010	30 m	Categorical driver for pre-flood conditions

The 2010 LULC layer was used as the pre-flood land-cover condition for the 2020 flood-zone layer. For Logistic Regression, the five LULC classes were collapsed into three categories based on their hydrological response and surface characteristics, consistent with approaches employed in flood-susceptibility studies that group LULC classes according to their influence on runoff and infiltration (Mulu et al., 2025). The resulting classes are shown in Table 4.

Table 8: Collapsed LULC classes used in Logistic Regression

Original Classes	Collapsed LULC Class	Description
Water (0)	0	Permanent water bodies
Bareland (1), Built-up (4)	1	Impervious surfaces
Natural Vegetation (2), Agriculture (3)	2	Semi-permeable to permeable surfaces

A balanced stratified random sampling procedure was applied to the 2010-2020 binary flood raster to generate 600 sample points, comprising 300 flooded (1) and 300 non-flooded (0) points. Values of the selected flood drivers were extracted at the sample locations and compiled into a master dataset for Logistic Regression analysis. After removing missing or invalid observations, 545 points (270 non-flooded and 275 flooded) were retained for analysis. The spatial distribution of the sampled points is shown in Figure 2.

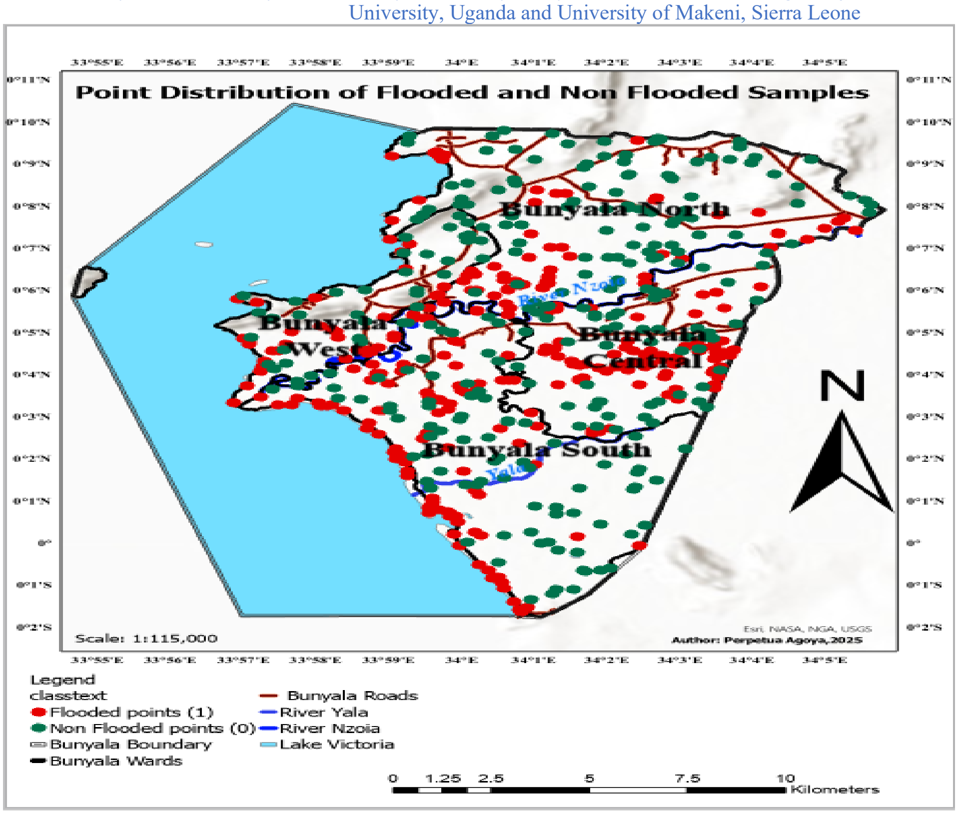


Figure 5: Distribution of balanced sample points for flooded (1) and non-flooded (0) classes in Bunyala Sub-County

Logistic Regression Modelling

Binary Logistic Regression was applied in SPSS to examine the relationship between flood occurrence in 2020 and the selected flood drivers for the 2010-2020 period**.** Flood occurrence was treated as the dependent variable (flooded = 1; non-flooded = 0), while elevation, slope, distance to water bodies, rainfall, and LULC were treated as independent variables.

The Logistic Regression Model was expressed as:

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n)}}$$

where:

- P= probability of flood occurrence
- β_0= intercept
- β_n= regression coefficients
- X_n= flood driver variables

Model performance was evaluated using classification accuracy, Receiver Operating Characteristic (ROC) analysis, Area Under the Curve (AUC), and the Hosmer-Lemeshow goodness-of-fit test.

Flood Susceptibility Mapping

The Logistic Regression coefficients were used to inform the relative importance of the selected flood drivers. Higher weights were assigned to factors with greater model coefficients, reflecting their relative contribution to flood occurrence. The assigned weights were normalized to sum to 100% and integrated using a weighted overlay approach in a GIS environment. Rainfall was excluded from the final weighted overlay because it was not statistically significant in the Logistic Regression model, although it is recognized as an important flood driver. This approach ensured that the spatial modelling was informed by the drivers supported by the model.

Table 9: Flood-driver weights used in the overlay

Flood Driver	Raw B Value	Assigned Weight (%)
LULC (combined)	6.452	40

Slope	0.241	25
Distance to Water	0.072	20
Elevation	0.020	15

The Flood Susceptibility Index (FSI) was calculated as a weighted sum of the selected flood drivers using the following equation:

$$FSI = (0.40 \times LULC) + (0.25 \times Slope) + (0.20 \times Distance\ to\ Water) + (0.15 \times Elevation)$$

The continuous FSI raster layer was first classified into five susceptibility levels ranging from very low to very high based on the FSI values (1.451-5.000), representing increasing potential for flooding. The FSI was further reclassified into three final flood susceptibility classes by grouping the five initial susceptibility levels into broader high, moderate and low categories, using the thresholds shown in table 6 to support practical decision-making and interpretation based on commonly used flood modeling thresholds.

Table 10: Flood susceptibility classification

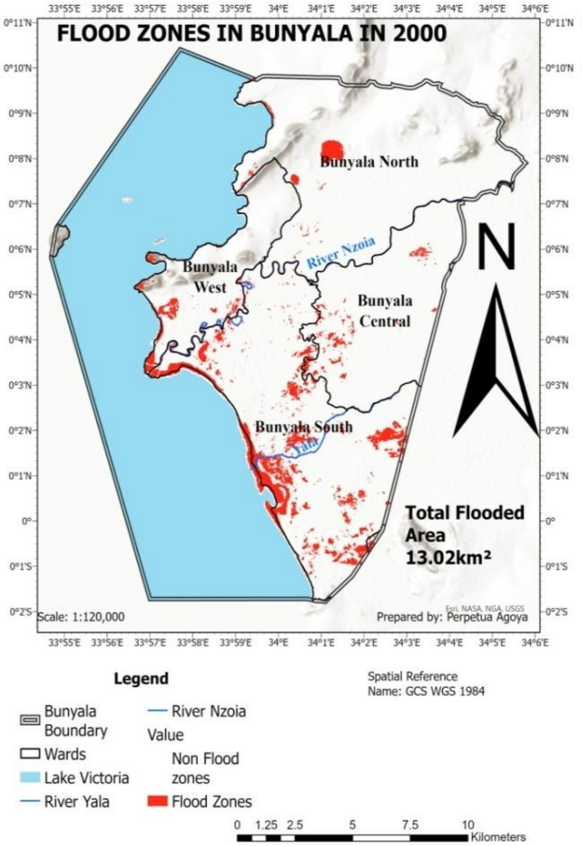
Range	**Class**	**Susceptibility Category**	**Color**
≤ 2.5	1	Low	Green
2.6 – 3.5	2	Moderate	Yellow
> 3.5	3	High	Red

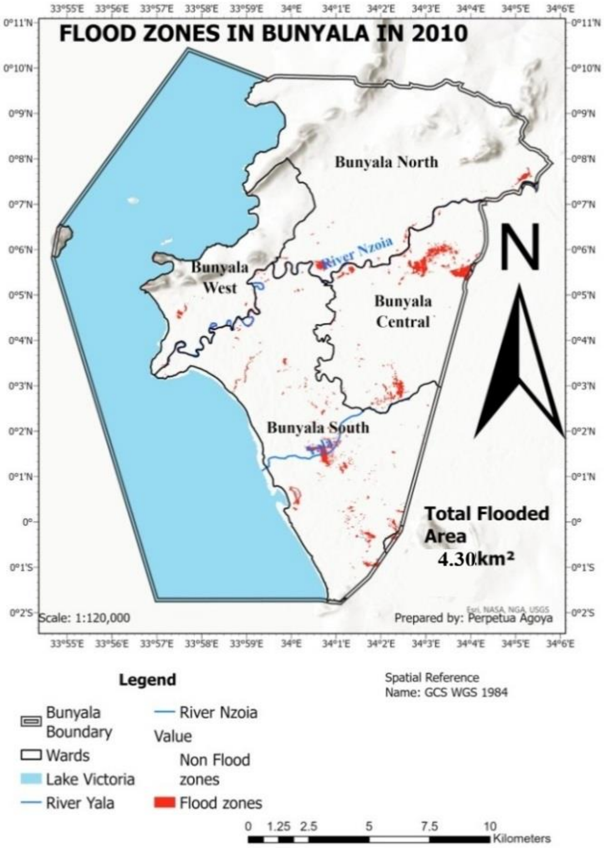
The resulting FSI Map was classified into low, moderate, and high flood susceptibility categories to generate the final flood susceptibility map of Bunyala Sub-County.

Results and Discussion

Changes in Flood Zones between 2000, 2010 and 2020

Flood-zone analysis showed substantial temporal variation in the spatial extent of flooding across Bunyala Sub-County. The spatial distribution of flood zones for 2000, 2010 and 2020 is shown in Figure 2.


FLOOD ZONES IN BUNYALA IN 2000
Bunyala North
Bunyala West
River Nzoia
Bunyala Central
Bunyala South
Yala
N
Total Flooded Area 13.02km²
Scale: 1:120,000
Prepared by: Perpetua Agoya
Legend
Bunyala Boundary
Wards
Lake Victoria
River Yala
River Nzoia
Value
Non Flood zones
Flood Zones
Spatial Reference
Name: GCS WGS 1984
0 1.25 2.5 5 7.5 10 Kilometers


FLOOD ZONES IN BUNYALA IN 2010
Bunyala North
Bunyala West
River Nzoia
Bunyala Central
Bunyala South
N
Total Flooded Area 4.30km²
Scale: 1:120,000
Prepared by: Perpetua Agoya
Legend
Bunyala Boundary
Wards
Lake Victoria
River Yala
River Nzoia
Value
Non Flood zones
Flood zones
Spatial Reference
Name: GCS WGS 1984
0 1.25 2.5 5 7.5 10 Kilometers

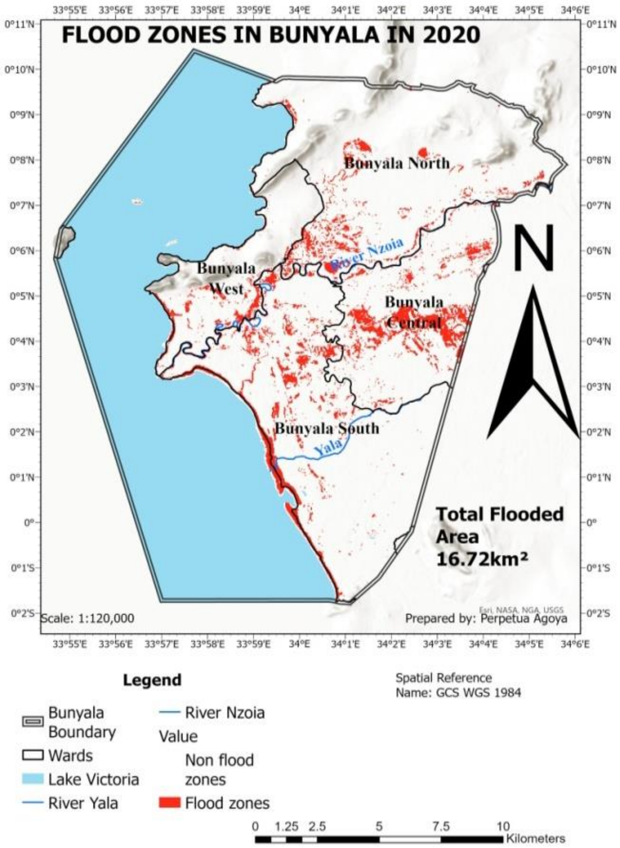


Figure 6: Spatial distribution of flood zones in Bunyala in 2000, 2010 and 2020

Table 7: Flood zone changes between 2000-2010-2020

Period	Start Area (km²)	End Area (km²)	Net Change (km²)
2000-2010	13.02	4.30	-8.72
2010-2020	4.30	16.72	+12.42
2000-2020	13.02	16.72	+3.70

Overall, flood zones increased by 3.70 km² between 2000 and 2020. Although flood extent declined between 2000 and 2010 by 8.72 km², a significant resurgence occurred between 2010 and 2020, resulting in a gain of 12.42 km². Spatially, flood zones were more persistent in Bunyala South and Central, while relatively few flooded zones were observed in North and West. The concentration of flooding in Bunyala South and Central was mainly associated with low-lying areas, although river overflow and Lake Victoria backflow also contributed to inundation.

The decline in flood zones between 2000-2010 may be associated with lower rainfall, river discharge and Lake Victoria water levels in that decade, while the subsequent increase in the following decade (2010-2020) may reflect increased rainfall and runoff, rising lake levels and backflow into the low-lying floodplain. These findings support observations by Okumu (2017) and Mutiso (2011), who argued that structural flood control measures such as dykes provide only temporary relief when broader hydrological and land use drivers remain unaddressed.

Relationship between observed Flood Zones and Flood Drivers

Binary Logistic Regression was applied to determine the relationship between flood occurrence and selected flood drivers. The results are presented in Table 8.

Table 8: Logistic Regression results

Flood Driver	B	Exp(B)	p-value
Elevation	0.020	1.020	0.544
Slope	0.241	1.272	0.001
Rainfall	0.000	1.000	0.343
Distance to Water	0.072	1.075	<0.001
Built-up/Bareland	4.836	125.940	<0.001
Agriculture/Vegetation	1.616	5.030	0.001

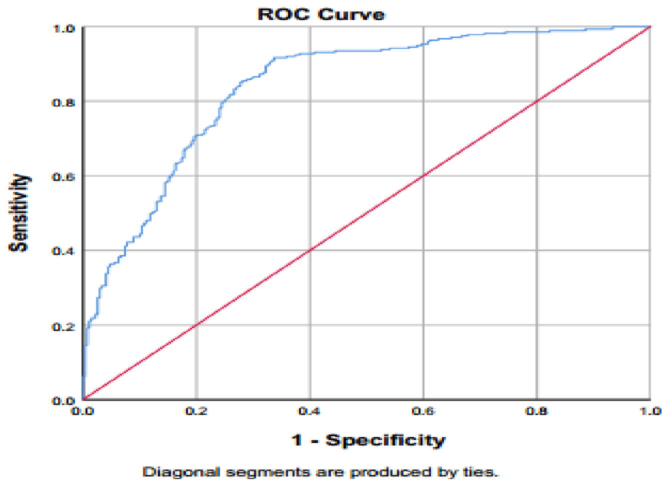
The results identified LULC as strongest-associated flood driver in the model in Bunyala Sub-County. Built-up and bareland areas had approximately 126 times higher odds of experiencing flooding than the reference category (B = 4.836, Exp(B) = 125.940, $p < 0.001$), while agriculture and vegetation areas had approximately 5 times higher odds of experiencing flooding than the reference category (B = 1.616, Exp(B) = 5.030, p = 0.001). Slope (B = 0.241, Exp(B) = 1.272, p = 0.001) and distance to water bodies (B = 0.072, Exp(B) = 1.075, $p < 0.001$) were statistically significant, whereas elevation showed a minimal statistical contribution (B = 0.020, Exp(B) = 1.020, p = 0.544) while rainfall (B = 0.000, Exp(B) = 1.000, p = 0.343) was not significant.

The strong influence of LULC as a flood driver was consistent with observations by Nabwire (2019) and Likuyi and Juma (2025), who associated land-cover changes and reduced surface permeability with increased flood vulnerability. Distance to water bodies and slope were also significantly associated with flood occurrence in the study area. The significance of slope reflects the influence of terrain on water accumulation and movement, while the significance of distance to water bodies indicates that the spatial relationship with major water bodies may influence the occurrence of floods. Although elevation showed a minimal statistical contribution, the observed flood patterns indicate that flooding affects areas across both low and moderate elevations. In contrast, although rainfall is an important flood driver, its influence was not statistically significant in the model, suggesting that local rainfall alone may not explain flood occurrence in Bunyala, where upstream hydrological processes and Lake Victoria backflow may also contribute (Nabwire, 2019; Wanzala, 2020**).** The performance of the Logistic Regression model is presented in table 9.

Table 9: Logistic Regression model performance

Indicator	Value
Nagelkerke R^2	0.455
Classification Accuracy	78.7%
Flooded Areas Correctly Classified	85.5%
Non-Flooded Areas Correctly Classified	71.9%
AUC	0.841
Hosmer-Lemeshow p-value	0.583

The model achieved an overall classification accuracy of 78.7%, correctly classifying 85.5% of flooded areas and 71.9% of non-flooded areas. The ROC analysis produced an Area Under the Curve (AUC) of 0.841, indicating good model discrimination, while the non-significant Hosmer-Lemeshow test (p = 0.583) confirmed good model calibration. These results demonstrate that the selected flood drivers adequately explain flood occurrence patterns in Bunyala Sub-County.



Area Under the Curve

Test Result Variable(s): Predicted probability

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.841	.017	.000	.808	.874

Figure 7: ROC curve of the logistic regression model for flood susceptibility assessment in Bunyala.

3.3 Flood Susceptibility Mapping and Predicted Future Flood Patterns

Statistically significant Logistic Regression coefficients were rescaled and incorporated into a weighted overlay model to generate the Flood Susceptibility Index (FSI) map and the predicted flood susceptibility map for Bunyala Sub-County. The resulting spatial distribution is presented in Figure 5 and 6 respectively.

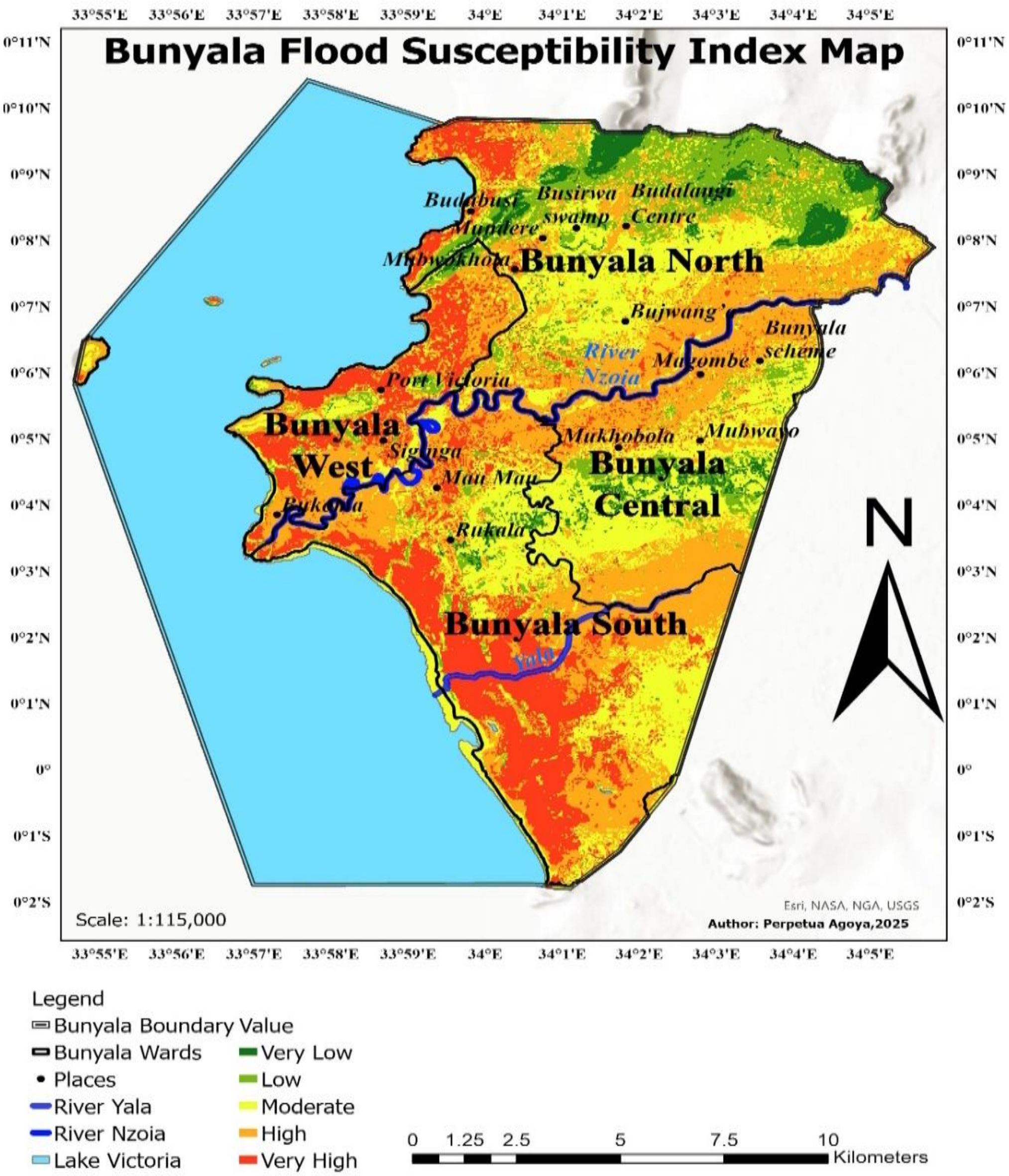


Figure 8: Flood Susceptibility Index (FSI) map for Bunyala Sub County

The resulting FSI Map revealed substantial spatial variation in flood susceptibility across the study area. The flood susceptibility assessment indicated that approximately 62% of the study area was classified as high flood susceptibility, 33% as moderate and 4% as low flood susceptibility as summarized in Table 10.

Table 10: Flood susceptibility classification

Susceptibility Class	Area (km²)	Percentage (%)
Low	7.55	4.09
Moderate	61.70	33.45
High	115.25	62.46

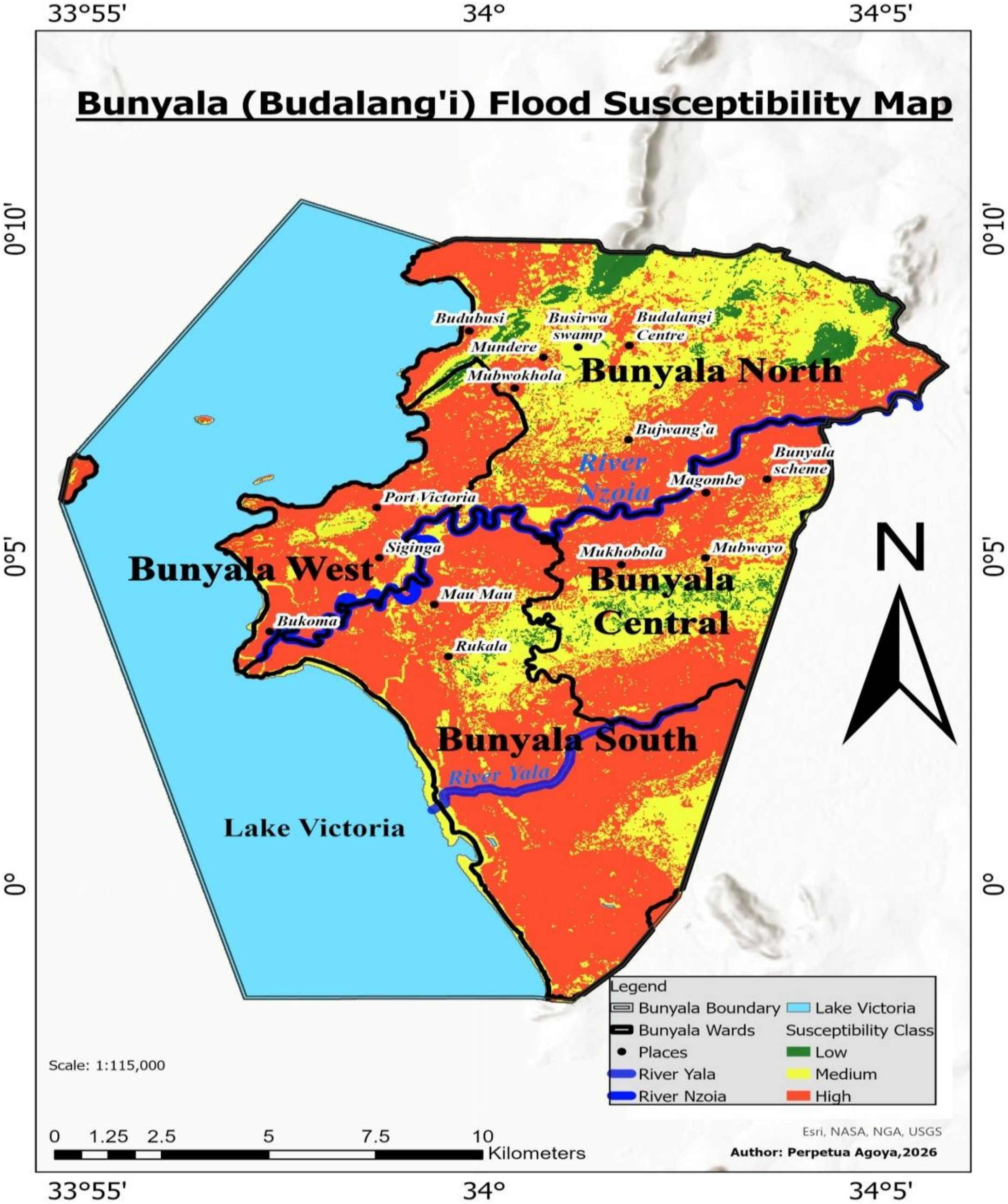


Figure 9: Flood susceptibility map of Bunyala Sub-County

High flood susceptibility zones were concentrated along Rivers Nzoia and Yala, the Lake Victoria shoreline, low-lying terrain, and areas dominated by built-up and bareland surfaces. Moderate susceptible areas were mainly associated with agricultural/ vegetated areas, while low susceptibility areas were mostly located in relatively higher elevations areas and away from permanent water bodies. These findings are consistent with previous studies in Bunyala (Mutiso, 2011; Nabwire, 2019; Likuyi and Juma, 2025), which identified the lower floodplain as highly vulnerable to flooding. The spatial correspondence between the predicted high susceptibility areas and previously identified flood prone areas suggests consistency in the spatial pattern of flood susceptibility across the study area.

Overall, the results demonstrate that flood susceptibility in Bunyala is shaped by the interaction of land-use characteristics, topography, proximity to water bodies, and broader hydrological processes. The identified high-susceptibility areas, particularly along Rivers Nzoia and Yala, the Lake Victoria shoreline, and low-lying areas like in Bunyala South and Central wards, represent locations with greater potential for future flood occurrence under the prevailing conditions of the selected flood drivers**.** The resulting flood susceptibility map therefore provides spatial information for anticipating future flood-prone areas and supporting land use planning, flood-risk management, disaster risk management, and climate adaptation in Bunyala Sub-County.

Conclusion and Recommendations

Conclusion

Flood zones in Bunyala Sub-County showed substantial temporal and spatial variation between 2000 and 2020. The extent of floods declined from 13.02 km² in 2000 to 4.30 km² in 2010, before increasing to 16.72 km² in 2020, resulting in a net increase of 3.70 km² over the study period. The results demonstrate that structural interventions alone may be insufficient to contain flood susceptibility over the long term. This demonstrates that flooding remains a persistent challenge in the sub-county, particularly in areas influenced by major water bodies, including rivers Nzoia &Yala and Lake Victoria and broader hydrological processes.

The logistic regression model established that LULC, mainly built-up and bareland areas, was the strongest determinant of flood occurrence, subsequently followed by agriculture and vegetation. Distance to water bodies and slope were also statistically significant, while elevation showed a minimal statistical contribution and rainfall was not statistically significant in the model. Although rainfall is an important flood driver, its non-significant contribution suggests that local rainfall alone may not explain flood occurrence in Bunyala, where wider upstream hydrological processes and Lake Victoria backflow may also contribute. These findings confirm that human-induced land use changes, particularly the expansion of impermeable and exposed surfaces, are progressively intensifying flood vulnerability in the study area.

Predictive modelling through the integration of Logistic Regression results with GIS-based weighted overlay analysis produced a Flood Susceptibility Index (FSI) and flood susceptibility map showing substantial spatial variation under the prevailing conditions of the selected flood drivers. High susceptibility areas covered 115.25 km² (62.46%), moderate-susceptibility areas 61.70 km² (33.45%), and low-susceptibility areas 7.55 km² (4.09%). The predicted pattern showed greater potential for future flood occurrence in areas associated with high susceptibility, while moderate- and low-susceptibility areas represented comparatively lower potential under the prevailing conditions. The findings demonstrate the value of integrating geospatial techniques and predictive modelling to identify areas with greater potential for future flooding and to support land-use planning, flood mitigation, and climate adaptation in Bunyala Sub-County.

Recommendations

The study recommends the use of the generated flood susceptibility map as a decision-support tool for land use planning, flood mitigation, and climate adaptation in Bunyala Sub-County. High-susceptibility areas that were identified in the flood susceptibility map should be designated as areas with restricted development, particularly for built-up expansion and vulnerable infrastructure and. Data from flood susceptibility and geospatial models should be integrated into zoning, land use planning, infrastructure design, and urban expansion plans to support data-driven decision-making.

Community-sensitized awareness programmes and Early Warning Systems should be enhanced to strengthen preparedness and reduce flood-related losses. Riparian zones should be protected through enforcement of land use regulations and restoration initiatives to reduce flood vulnerability associated with exposed and impermeable land surfaces.

At the policy level, flood susceptibility maps should be integrated into national and county disaster risk reduction and climate adaptation frameworks. Policies promoting upstream catchment protection should also be enhanced to complement downstream flood mitigation efforts in Bunyala and the wider Lake Victoria Basin.

Further research should explore rainfall and hydrological dynamics within the upstream River Nzoia catchment, including the Cherangany Hills, and integrate socio-economic vulnerability into flood-susceptibility assessment.

References

Gaya, C. O. (2019). *Application of GIS and remote sensing in flood management in the Lake Victoria Basin* (Doctoral dissertation). Jomo Kenyatta University of Agriculture and Technology.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). Wiley.

Integrated Research on Disaster Risk. (2023). *2023 global natural disaster assessment report*.

Likuyi, I. A. (2018). *Flood risk mapping and factors that motivate the affected community to continue living in the flood risk zones in Budalangi Sub County, Busia County, Kenya* (Master's thesis). Moi University.

Likuyi, I. A., & Juma, S. (2025). Mapping land use land cover within flood risks and safe zones in Budalangi Sub County, Kenya. *East African Journal of Environment and Natural Resources, 8*(1), 360-373.

Mulu, A., Kassa, S. B., Wossene, M. L., Adefris, S., & Meshesha, T. M. (2025). Identification of flood vulnerability areas using analytical hierarchy process techniques in the Wuseta watershed, Upper Blue Nile Basin, Ethiopia. *Scientific Reports, 15*(1), 28680. https://doi.org/10.1038/s41598-025-13822-6.

Mutiso, F. M. (2011). *Integration of remote sensing (RS) and geographic information system (GIS) in monitoring and management of floods: Case study Budalangi area* (Master's thesis). University of Nairobi.

Nabwire, C. A. (2019). *Assessment of River Nzoia Basin morphometric characteristics and flooding risks in Budalang'i area, Busia County* (Master's thesis). University of Nairobi.

Okumu, M. (2017). Assessing the root cause of persistent floods and strategic community-based interventions in Bunyala, Busia County, Kenya. *International Journal of Innovative Research and Development*.

Onywere, S. M., Getenga, Z. M., Mwakalila, S., Twesigye, C. K., Nakiranda, J. K., & Opere, A. (2013). Erosive coping after the 2011 floods in Kenya. *International Journal of Global Warming, 5*(4), 452-474. https://doi.org/10.1504/IJGW.2013.057285.

Wanzala, J. (2020, August 19). Floods on record high Lake Victoria expose need for water cooperation. *Thomson Reuters Foundation*.