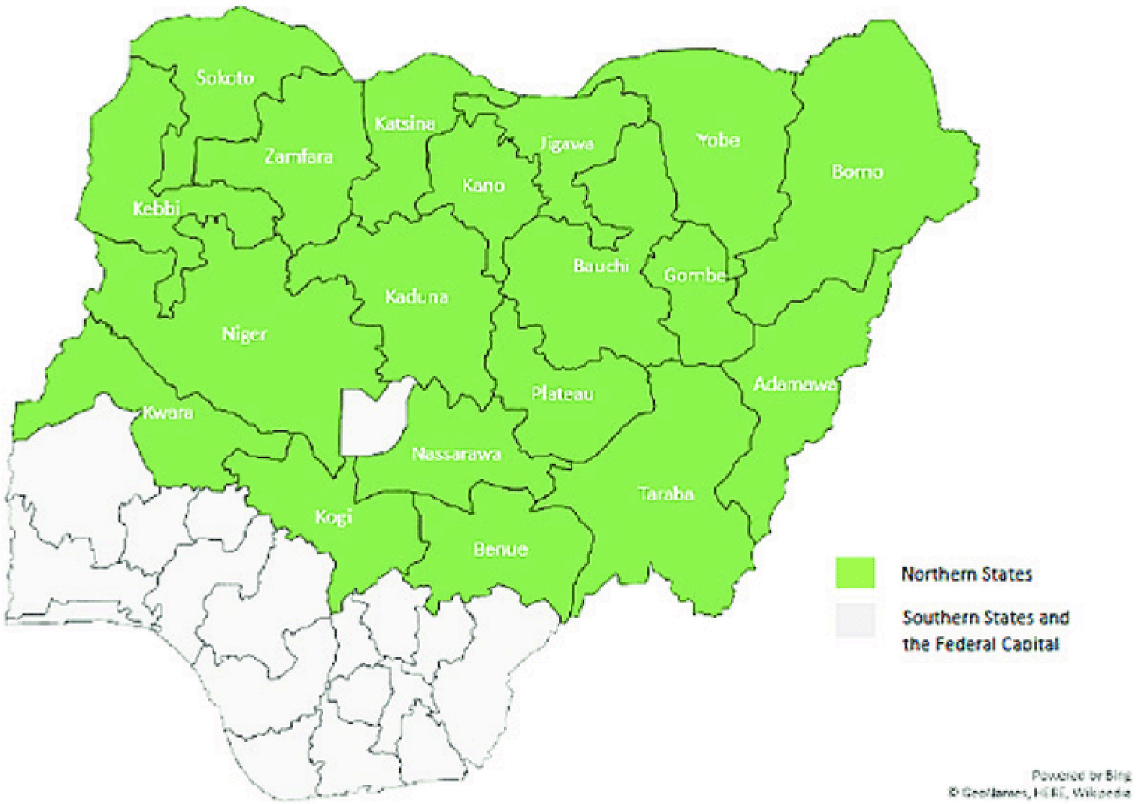


# Flood Area Extent and Degree Prediction in Northern Nigeria using GIS and Machine Learning Models

**By Team Adichie** AI Saturdays Lagos Flipped Cohort 9

13.09.2025

# Introduction

Natural disasters rank among the most devastating phenomena in human history. Events such as tsunamis, volcanic eruptions, hurricanes, sandstorms, hailstorms, and flooding inflict tremendous human and economic losses. Globally, damages from natural catastrophes now reach into the hundreds of billions of US dollars each year. According to recent United Nations Office for Disaster Risk Reduction (UNDRR) reporting, while earlier estimates placed direct losses from disasters at US$70–80 billion annually (1970–2000), this has surged to about US$180–200 billion per year over 2001–2020, and when indirect costs and cascading effects are included, the true cost may exceed US$2.3 trillion annually. [1]

Although Nigeria does not face some hazards such as tsunamis or volcanoes, it is far from immune to natural disasters. Flooding is an especially persistent challenge. Coastal regions—most notably Lagos and the Niger Delta—as well as the basins of the Niger and Benue Rivers are frequently affected. Contrary to the popular belief that Northern Nigeria is safe from flooding due to typically lower rainfall, recent events show that it is vulnerable too. For example, the May 2025 Mokwa flood in Niger State overwhelmed communities in the North-Central region: torrential rains, a collapsing dam, and inadequate drainage combined to destroy thousands of homes, submerge crops, displace over 3,000 people, and kill at least 150–200 persons. [2]

Given these recurrent events and their substantial socio-economic impacts, there is a pressing need to develop reliable methods to predict both the extent of flood-affected areas and the severity of flooding, especially in Northern Nigeria. The goal of this project, therefore, is to accurately model and forecast flood extents and flood degrees (e.g. depth or intensity) in Northern Nigeria, in order to aid preparedness, mitigation, and response.

# Problem Statement

Despite efforts by government agencies and humanitarian organizations, response strategies are often reactive rather than proactive due to limited predictive insights into where and to what extent floods may occur. Current flood risk assessments are often based on outdated hydrological models, manual mapping, or incomplete field surveys, making them inadequate for timely disaster preparedness and mitigation.

Thus, there is a pressing need for an integrated, data-driven approach that combines Geographic Information Systems (GIS) with Machine Learning models to accurately predict the spatial extent and severity (degree) of floods in Northern Nigeria. This would provide early warning systems, aid resource allocation, and support evidence-based decision-making for disaster risk reduction.

## Existing Solutions

Flood management in Nigeria has seen the use of hydrological and hydraulic models like HEC-RAS and SWAT [3], remote sensing and GIS mapping [4], machine learning methods such as Random Forest, SVM, and Neural Networks [5], and early warning systems from agencies like NEMA and NiMet [6]. While these approaches are useful, they often struggle with missing hydrological data, focus more on mapping after floods than actually predicting, rely on poorly calibrated models, and provide alerts that are too broad to help communities directly.

## Our Contribution

What we are doing differently is bringing satellite-derived flood data together with machine learning to not only map flood extent but also predict flood degree, while using exploratory data analysis to generate insights that truly fit the Northern Nigeria context.

## Objectives

The project aims to:

1. Develop a GIS-based Machine Learning framework for predicting the extent and degree of flooding in Northern Nigeria.

2. Integrate multi-source datasets (rainfall, elevation, river networks, soil types, land use/land cover, historical flood records, and remote sensing imagery) for accurate flood modeling.

3. Compare and evaluate multiple Machine Learning algorithms (e.g., Decision Trees, Random Forest, Gradient Boosting, SVM, Neural Networks) to identify the most effective model for local conditions.

4. Generate spatially explicit flood risk maps that classify areas based on predicted flood extent and severity.

5. Provide actionable insights to government agencies, NGOs, and local communities for proactive flood management, disaster preparedness, and sustainable urban and agricultural planning.

## Proposed Data Sources

I. **Historical Data, 1985 - 2024** from the EM-DAT, CRED and Dartmouth Flood Observatory, to identify and map historical flood locations to investigate the geographical link between flood likelihood and the factors that impact it

II. **Satellite images** - Sentinel-1, Landsat and Digital Elevation models from Digital Earth Africa and Copernicus Data Space Ecosystem and the NASA Earthdata websites. The digital elevation model (DEM) with a 30-m spatial resolution produced by the USGS Earthexplorer will be used to calculate the aspect ratio, slope, contour and curvature. Using QGIS and SAGA GIS software, we will produce the final maps of the slope, aspect, TWI, SPI, and curvature.

III. **Rainfall, Temperature, Landcover, water areas and soil type** will be collected from Nimet Nigeria, Global Climate data, Globeland30, OCHA, Nigeria and The Harmonised World Soil Database respectively.

## Proposed Methodology

**Data collection:** The first step is to gather relevant data for flood area extent (spatial) and degree modeling. The data will comprise Northern areas associated with flood occurrences. Data include satellite images, organizational data, and statistical records of flood events. These criteria are selected based on their relevance to flood events and their ability to provide valuable insights into flood-prone areas.

**Flood occurrence data extraction:** This includes flood points extraction These points were derived through the combination of Sentinel-1 and Landsat-8 satellite images, utilizing advanced image processing techniques and flood detection algorithms.

**Data preprocessing and preparation**: This step is important for quality and suitability of the data for modeling. This involved performing a multicollinearity test to assess the dependence between independent variables, and employing the step-wise weight assessment ratio analysis to determine the weights of the criteria.

**Model selection and optimization**: This step selects the non-parametric DT algorithm as the base model for flood-prone areas mapping. It also uses other algorithms like Random Forest, Gradient Boosting, SVM, Neural Networks.

**Model validation and performance evaluation:** The trained and optimized model is validated using independent validation data. Various evaluation metrics are employed to assess the performance of the model. These metrics include root mean square error (RMSE) and mean absolute error (MAE) to measure prediction accuracy, coefficient of determination (R2) to determine the goodness of fit.

## Expected Outcomes

1. A geospatial map identifying and visualizing areas at risk of flooding in Northern Nigeria.
2. A robust machine learning model capable of accurately predicting flood degrees under varying conditions.
3. Analytical insights into the environmental and hydrological drivers of flooding in the region.

## Community Impact

This project aims to empower communities across Northern Nigeria by providing actionable flood predictions that can save lives, protect livelihoods, and guide sustainable development. By delivering accurate, localized flood risk maps and early warnings, we hope to support farmers, local governments, and emergency responders in making informed decisions. Ultimately, our goal is to reduce vulnerability, enhance preparedness, and contribute to a safer, more resilient future for the people of Northern Nigeria

## Team Members

Omoregie Oluwabunmi

Lawal Fawaz

Olatona Deborah

Bello Ridwan

Ayodeji

## Acknowledgement

We extend our sincere gratitude to our mentors, colleagues, and the entire AI Saturdays Lagos Flipped Cohort 9 community for their invaluable guidance, encouragement, and support throughout this project. We also thank the various data providers—including NASA, USGS, Copernicus, NiMet, and other organizations—for making their resources accessible, without which this research would not have been possible. Finally, we acknowledge the resilience of the communities in Northern Nigeria, whose experiences and challenges inspired this work and remind us of its real-world importance.

# Reference

[1] UNDRR. *Global Assessment Report on Disaster Risk Reduction 2025*. UNDRR, 2025. (Estimates global economic losses from disasters, about US$2.3 trillion when both direct and indirect costs are considered.)

[2] "2025 Nigeria Floods (Mokwa)." Multiple media reports including AP, Premium Times, NEMA.

[3] *Flood Risk Assessment and Mapping in Hadejia River Basin, Nigeria, Using Hydro-Geomorphic Approach and Multi-Criterion Decision-Making Method* – states that certain sections of the river couldn't be simulated due to lack of hydrological data.

[4] *GIS-assisted Flood-risk Potential Mapping of Ilorin and its Environs, Kwara State* – uses LULC, elevation, slope etc to map vulnerability zones.

[5] *Application of GIS and Machine Learning to Predict Flood Areas in Nigeria* (Halima Ighile et al.) – uses ANN and logistic regression over historical flood records and environmental variables. Demonstrates some success, but mentions issues with data completeness

[6] *Flood Risk and Vulnerability Analysis of the Lower Usuma River in Gwagwalada Town Abuja* uses GIS & HEC-RAS and concludes that more discharge recording equipment and gauge stations are needed to reduce errors in flood risk assessment.