INNOVATIONS IN INFORMATION TECHNOLOGY
https://www.worldlinkpub.com/journals/innovations-in-information-technology/
ISSN: 3141-6004

An Integrated Geographic Information System and Machine Learning Framework for Spatiotemporal Monitoring and Prediction of Farmer-Herder Conflicts in Nigeria

***[1]Aamo Iorliam and [2]Ode Samuel**
[1] Department of Mathematics/Computer Science, Rev. Fr. Moses Orshio Adasu University, Makurdi (formerly Benue State University, Makurdi).
[2] [2]Department of Mathematics/Computer Science, Rev. Fr. Moses Orshio Adasu University, Makurdi (formerly Benue State University, Makurdi).
Corresponding Author: aamoiorliam@gmail.com

Received: 25th May 2026
Accepted for publication: 15 June 2026
Published: 01 July 2026

ABSTRACT

Farmer-herder clashes have become one of Nigeria's most persistent internal security and development challenges, resulting in widespread loss of life, displacement, and disruption of agricultural productivity. Driven by population pressure, climate-induced environmental stress, land-use competition, and weak conflict-monitoring systems, these clashes have increasingly outpaced other forms of violent insecurity in several regions of Nigeria. Despite multiple policy interventions, existing conflict management approaches remain largely reactive, fragmented, and limited in predictive capacity. This paper addresses these gaps by developing a hybrid Geographic Information System (GIS) and Machine Learning (ML) framework for tracking, analyzing, and predicting farmer-herder clashes in Nigeria. Adopting a design science research paradigm and guided by the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework, the study integrates multi-source data, including historical conflict records (2010-2025), geospatial layers, and socio-environmental indicators. GIS techniques are employed to map spatial and temporal patterns of conflict across high-risk states, particularly in Nigeria's Middle Belt, while supervised machine-learning models are developed to estimate future conflict risk. Multiple algorithms, including Logistic Regression, Random Forest, Gradient Boosting, Support Vector Machines, and Extreme Gradient Boosting (XGBoost), are evaluated using standard classification and probabilistic performance metrics. Results reveal pronounced spatial clustering and persistence of conflict, with significant heterogeneity in both event frequency and fatality severity across states. Ensemble tree-based models outperform linear and kernel-based approaches, with XGBoost demonstrating the strongest predictive accuracy and probabilistic calibration. The study demonstrates that integrating GIS with machine learning can substantially enhance early-warning capabilities and support proactive, data-driven conflict management. By translating complex model outputs into spatially explicit risk maps, the proposed framework offers practical value for policymakers, security agencies, and peacebuilding actors seeking to prevent violence, safeguard livelihoods, and promote sustainable coexistence in conflict-prone regions of Nigeria.

Keywords: Farmer-Herder Clashes; Geographic Information System (GIS); Machine Learning (ML); Conflict Prediction; Early-Warning Systems.

1.0 INTRODUCTION

Farmer-herder clashes have emerged as one of the most pressing internal security and development challenges in Nigeria (Sukare, & Abdullahi, 2025). Over the years, violent confrontations between nomadic pastoralists and sedentary farming communities have escalated, resulting in widespread loss of lives, destruction of farmlands, displacement of rural populations, and disruption of agricultural productivity. According to Ajogbeje, & Sylwester (2024) and Omotola (2025), farmer-herder violence now accounts for more deaths annually than the Boko Haram insurgency in certain regions of Nigeria (Ajogbeje, & Sylwester, 2024; Omotola, 2025).

The roots of the conflict are deeply embedded in Nigeria's socio-economic and environmental landscape. Rapid population growth has intensified pressure on land, while climate change has contributed to desertification and shrinking water resources in the northern regions. As the Sahel zone becomes increasingly arid, herders are forced to migrate southward in search of pasture and water for their cattle (Ayeb-Karlsson, Hayward, & Kniveton, 2025). This southward movement often brings them into direct competition with farmers in the North Central, South West, and South East zones, leading to disputes over farmland encroachment, destruction of crops, and retaliatory attacks.

The Middle Belt region of Nigeria comprising states such as Benue, Plateau, Nasarawa, Kaduna, and Taraba has witnessed recurring outbreaks of from herder's attacks. Several villages in Middle Belt region of Nigeria have been razed, billions of naira worth of agricultural produce destroyed, and critical infrastructure such as roads, markets, and schools disrupted. This crisis not only undermines food security and rural development but also threatens national stability by aggravating ethno-religious tensions and straining already fragile governance systems (Nwagbo, Francis, & Ezeakoafor, 2025). Despite numerous government interventions

Cite this article: Iorliam, A. and Samuel O. (2026). An Integrated Geographic Information System and Machine Learning Framework for Spatiotemporal Monitoring and Prediction of Farmer-Herder Conflicts in Nigeria. *Innovations in Information Technology*, 1(1):29-36

including the National Livestock Transformation Plan, the establishment of grazing reserves, and peace-building initiatives, conflict management remains largely reactive (Chukwu, 2025).

Recent technological advances offer promising opportunities to address tackling farmer-herder clashes challenges in Nigeria. For example, geographic information systems (GIS) provide a robust framework for capturing, analyzing, and visualizing spatial patterns of conflict (Thompson, & Petit, 2026), while machine learning (ML) enables predictive modeling of complex socio-environmental data (Anil, Jyothi, & Suma, 2026). A hybrid GIS-ML approach can integrate geospatial data such as satellite imagery, climate variables, and land use maps with historical conflict records to predict high-risk zones, track herder migration routes, and inform timely policy responses. Such a system could revolutionize Nigeria's conflict management architecture by enabling proactive, data-driven interventions. This paper aims at developing a hybrid geographic information system and machine learning technique for tracking and predicting farmer-herder clashes in Nigeria.

Farmer-herder conflicts have become one of the most persistent and violent forms of resource-based conflict in sub-Saharan Africa, with Nigeria representing one of the most severe and complex cases. Over the years, confrontations between sedentary farming communities and nomadic pastoralists have escalated in frequency, geographic spread, and lethality, surpassing other forms of internal violence in several regions of the country (Tawane, 2025; Amoah, Bloomer, & Campbell, 2025). While an extensive body of literature has examined the causes, dynamics, and consequences of these conflicts, existing studies are dominated by qualitative narratives, descriptive statistics, and post hoc analyses (Okonkwo, & Akanji, 2025). Consequently, there remains a significant gap in predictive, spatially explicit, and operationally usable conflict monitoring systems.

Farmer-herder conflicts are commonly conceptualised as struggles over access to and control of natural resources, particularly land and water. Early literature framed these conflicts largely through a neo-Malthusian lens, attributing violence to population growth and resource scarcity (Nwankwo, 2025). While this perspective remains influential, it has been criticised for environmental determinism and for underplaying political, institutional, and historical factors.

Contemporary scholarship increasingly adopts political ecology and environmental security frameworks, which emphasise the interaction between ecological stress, power relations, and governance structures (Mardikaningsih, 2025). From this perspective, environmental change does not automatically lead to conflict; rather, conflict emerges where institutions fail to manage competing claims and mediate access to resources equitably.

In the Nigerian context, farmer-herder clashes are also analysed through livelihood security theory, which highlights the vulnerability of both farmers and pastoralists to climatic shocks, market fluctuations, and policy failures (Eke et al., 2025). Pastoralism and smallholder farming are both adaptive livelihood systems, yet their coexistence has become increasingly strained under conditions of land scarcity and institutional breakdown.

Historically, relations between farmers and pastoralists in Nigeria were largely complementary. Pastoralists grazed livestock on harvested farmlands, providing manure that enhanced soil fertility, while farmers benefited from access to animal products. These interactions were regulated by customary institutions and clearly defined transhumance routes (Desta, 2025).

The establishment of grazing reserves in the 1960s and 1970s was intended to mitigate conflict, yet poor implementation, encroachment, and lack of political commitment rendered many reserves ineffective (Blench, & Hassan, 2025). As desertification intensified in northern Nigeria and population growth accelerated, pastoralists were increasingly compelled to migrate southwards, often along routes no longer recognised or protected. This historical trajectory highlights that contemporary farmer-herder conflicts are not merely spontaneous disputes but the cumulative outcome of decades of policy neglect, environmental stress, and institutional erosion.

Environmental change is one of the most frequently cited drivers of farmer-herder conflicts in Nigeria. Northern Nigeria lies within the semi-arid Sahelian zone, which has experienced increasing rainfall variability, prolonged droughts, and advancing desertification over recent decades. These changes have reduced the availability of pasture and water, compelling pastoralists to migrate southwards in search of more favourable conditions (Haaga et al., 2023).

Empirical studies demonstrate a correlation between periods of drought, vegetation stress, and increased pastoral movement into farming areas (Mohamed, Chauke, & Palanisamy, 2026). Such movements often coincide with the dry season and harvest periods, when crops are most vulnerable to damage. Crop destruction by livestock frequently serves as an immediate trigger for violent confrontation.

However, scholars increasingly caution against treating climate change as a direct or deterministic cause of conflict. Mohamed, Chauke, & Palanisamy (2026) argued that climate variability acts as a "threat multiplier" that exacerbates existing tensions rather than creating conflict in isolation. In regions with effective governance and conflict resolution mechanisms, similar environmental pressures do not necessarily lead to violence. This suggests that environmental factors interact with social and institutional conditions to shape conflict outcomes.

Nigeria's rapid population growth has significantly intensified competition over land. Expanding agricultural frontiers, settlement growth, and infrastructure development have reduced available grazing space and disrupted traditional transhumance routes (Usman, & Nichol, 2022). In the Middle Belt region, often described as Nigeria's food basket, fertile land has become increasingly scarce, heightening tensions between farmers and herders.

Land-use change has altered the spatial dynamics of conflict. Areas undergoing rapid agricultural expansion tend to experience higher incidences of farmer-herder clashes, particularly where grazing routes have been blocked or water points restricted (Sulieman, & Momale, 2025). The spatial concentration of conflict in specific states, such as Benue, Plateau, and Nasarawa, reflects these underlying land-use pressures.

Weak governance is widely recognised as a central factor in the escalation of farmer–herder conflicts. Nigeria's land administration system is characterised by overlapping statutory and customary authorities, inconsistent enforcement, and limited capacity for dispute resolution (Yunusa, & Owoyemi, 2025). These weaknesses create ambiguity over land rights and exacerbate competition between land users.

Government interventions, including grazing reserves and anti-open-grazing laws, have produced mixed results. While some policies aim to modernise livestock production and reduce conflict, their implementation has often been uneven and politically contested. Critics argue that such policies marginalise pastoralists and fail to provide viable alternatives, thereby intensifying grievances (Yunusa, & Owoyemi, 2025).

Farmer-herder clashes in Nigeria are frequently framed along ethnic and religious lines, particularly between predominantly pastoralists and farming communities from diverse ethnic backgrounds. While ethnicity is rarely the root cause of conflict, its politicisation has transformed economic disputes into identity-based violence (Olamide Sowale, 2025).

Political actors and media narratives often exploit these divisions, reinforcing stereotypes and deepening mistrust. This politicisation complicates conflict management and contributes to cycles of retaliation, making peaceful resolution increasingly difficult.

Geographic Information Systems have become central to spatial analysis in conflict research. GIS enables the visualisation of conflict hotspots, analysis of spatial clustering, and examination of proximity to environmental and infrastructural features (Njoya et al., 2025). In African conflict studies, GIS has been used to analyse election violence, insurgency, and communal conflicts. In Nigeria, GIS-based studies have mapped farmer-herder clashes and explored correlations with land-use patterns, rainfall variability, and population density (Adelabu, & Wang, 2025). However, most applications remain descriptive rather than predictive. GIS is often used to map where conflict has occurred, rather than to estimate where it is likely to occur in the future. This descriptive bias limits the potential of GIS as an early warning tool.

Machine learning has emerged as a powerful tool for conflict prediction due to its ability to model complex, nonlinear relationships. Algorithms such as Random Forests, Gradient Boosting Machines, and Neural Networks have been applied to predict civil unrest, armed conflict, and crime patterns with promising results (Udoh, Alabi, & Faleye, 2025).

Compared to traditional statistical models, machine learning techniques demonstrate superior predictive performance, particularly when dealing with high-dimensional and heterogeneous datasets. Importantly, tree-based models such as Random Forests also offer measures of feature importance, enhancing interpretability.

Despite these advances, the application of machine learning to farmer-herder conflicts in Nigeria remains limited. Existing studies often lack high-resolution spatial data or fail to integrate environmental, socio-economic, and historical conflict variables comprehensively.

The literature reviewed in this paper demonstrates broad agreement that farmer-herder clashes in Nigeria are complex, multi-causal phenomena with far-reaching impacts. Environmental change, population growth, weak governance, and identity politics interact in ways that defy simple explanation. While substantial research has documented the causes and consequences of these conflicts, several gaps remain evident. First, much of the literature relies on retrospective accounts and qualitative narratives, limiting the ability to anticipate future conflict dynamics. Second, existing studies often treat environmental, socio-economic, and political drivers in isolation, rather than as interacting factors.

Addressing these gaps requires approaches that integrate multiple dimensions of conflict and move beyond reactive analysis. A more systematic understanding of patterns, impacts, and underlying drivers is essential for informing effective policy and intervention.

2.0 Methodology

This paper adopts a design science research paradigm underpinned by predictive analytics and spatial intelligence, with the aim of developing a hybrid Geographic Information System (GIS) and Machine Learning (ML) framework for tracking and predicting farmer-herder clashes in Nigeria. The methodological approach is further strengthened by the adoption of the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework, which provides a systematic, iterative, and problem-oriented structure for the development of intelligent data systems.

The farmer-herder conflict represents a complex socio-environmental phenomenon characterized by nonlinear interactions among climatic variability, land-use dynamics, population pressure, and socio-political factors. Traditional empirical or purely statistical research designs are insufficient to capture these interdependencies or to generate actionable predictions. Consequently, this paper integrates spatial analysis and machine learning within the CRISP-DM framework to move beyond descriptive mapping toward anticipatory and preventive conflict management. The research design emphasizes iterative learning, continuous validation, and feedback between analytical stages, ensuring that the developed system remains adaptable to evolving conflict dynamics and data availability.

The business understanding phase focuses on translating the real-world challenge of farmer-herder clashes into a clearly defined analytical problem. In Nigeria, conflict management efforts have historically been reactive, relying on post-event reports and fragmented intelligence. This study reconceptualizes the challenge as a spatio-temporal risk prediction problem, where the objective is to estimate the probability of conflict occurrence across space and time with sufficient precision to support early intervention.

The data understanding phase involves the identification, collection, and exploratory analysis of datasets relevant to farmer-herder conflicts. Historical conflict data covering the period from 2010 to 2025 form the core of the analysis, providing information on the timing, location, and severity of farmer-herder clashes across Nigeria's conflict-prone regions is utilized.

Geospatial data include administrative boundaries, land-use and land-cover classifications, grazing routes, and hydrological features. These datasets are essential for situating conflict incidents within their spatial context and for analyzing proximity-based risk factors. Climatic and environmental data, such as rainfall patterns, temperature variability, and vegetation indices, are incorporated to capture environmental stressors that influence pastoral mobility and resource competition. Socio-economic data, including population density and livelihood dependence, provide insight into human pressure on land and resources.

In essence the datasets used in this project included Armed Conflict Location & Event Data (ACLED) datasets (ACLED, 2019) and Global Administrative Areas (GADM) for Nigeria at administrative level 1 (first-level subdivisions such as states). GADM is a high-resolution global dataset of political and administrative boundaries maintained by the GADM project available at https://geodata.ucdavis.edu/gadm/gadm4.1/json/gadm41_NGA_1.json. It provides polygon geometry and attributes for each state in Nigeria, including names and standard codes, suitable for mapping and spatial analysis in GIS software or web maps.

Exploratory data analysis is conducted using both statistical techniques and GIS-based visualization. Temporal analysis reveals pronounced seasonal patterns in conflict occurrence, often coinciding with dry-season migrations and periods of agricultural harvesting. Spatial analysis indicates strong clustering of incidents, particularly in the Middle Belt region, suggesting the presence of persistent hotspots rather than random distribution. These observations reinforce the suitability of spatially explicit machine learning models and justify the integration of GIS into the analytical workflow.

In this paper, data preparation involves extensive cleaning, transformation, and integration. Conflict records are first subjected to rigorous validation to address issues of duplication, missing values, and spatial inaccuracies. Inconsistent location descriptions are reconciled through geocoding and cross-referencing with administrative boundary datasets. Temporal inconsistencies are resolved by standardizing date formats and aggregating incidents into uniform time intervals suitable for modeling.

Spatial datasets are transformed into a common coordinate reference system to ensure accurate overlay and distance calculations. Raster datasets are resampled to a consistent spatial resolution. Feature engineering is then performed to derive explanatory variables that capture theoretically meaningful drivers of conflict. The final output of this phase is a geospatially structured dataset in which each spatial unit is associated with a set of environmental, socio-economic, and historical conflict attributes, along with a target variable indicating conflict occurrence. This dataset forms the foundation for subsequent machine learning modeling.

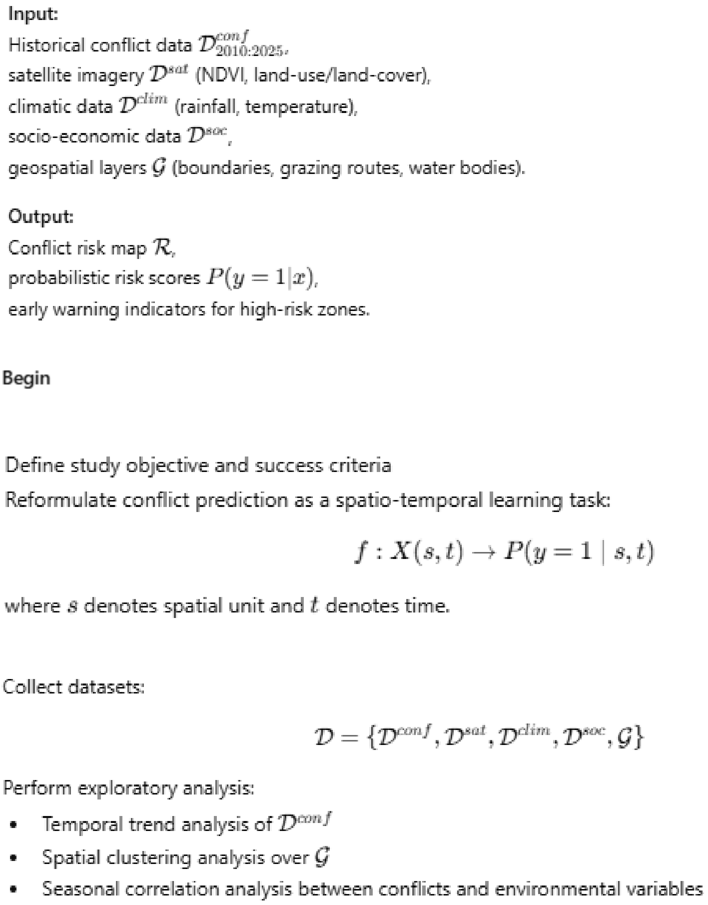
The modeling phase focuses on developing predictive models capable of learning complex, nonlinear relationships between explanatory variables and conflict occurrence. Random Forest is employed as a baseline model due to its robustness, resistance to overfitting, and ability to handle high-dimensional data. By aggregating the predictions of multiple decision trees, Random Forest captures nonlinear interactions among variables while providing measures of feature importance that enhance interpretability. Gradient Boosting Machines are subsequently implemented to improve predictive performance by sequentially correcting model errors and emphasizing difficult-to-predict cases. This approach is particularly effective in contexts where subtle interactions among variables drive outcomes.

The hybrid system developed in this paper follows a modular architecture that integrates GIS-based spatial analysis with machine learning prediction. Data ingestion forms the initial stage, drawing inputs from conflict databases, remote sensing platforms, and climate repositories. These inputs are processed through a preprocessing module that performs cleaning, transformation, and feature extraction.

The GIS module serves as the spatial intelligence layer, responsible for mapping conflict incidents, analyzing spatial patterns, and generating visual representations of risk. Outputs from the GIS module feed into the machine learning engine, where predictive models estimate conflict probabilities for each spatial unit. The final outputs are rendered as dynamic risk maps and analytical reports, forming the basis of an early warning and decision-support system.

Model evaluation is conducted using a combination of statistical performance metrics and spatial validation techniques. Standard classification metrics such as precision, recall, F1-score, and area under the receiver operating characteristic curve are used to assess predictive accuracy. The system is designed to support prevention and peacebuilding rather than surveillance action and the overall algorithm is shown in Algorithm 1. Moreover, Python programming is used to code Algorithm 1.

Algorithm 1: Hybrid GIS-Machine Learning Framework for Spatio-Temporal Conflict Prediction

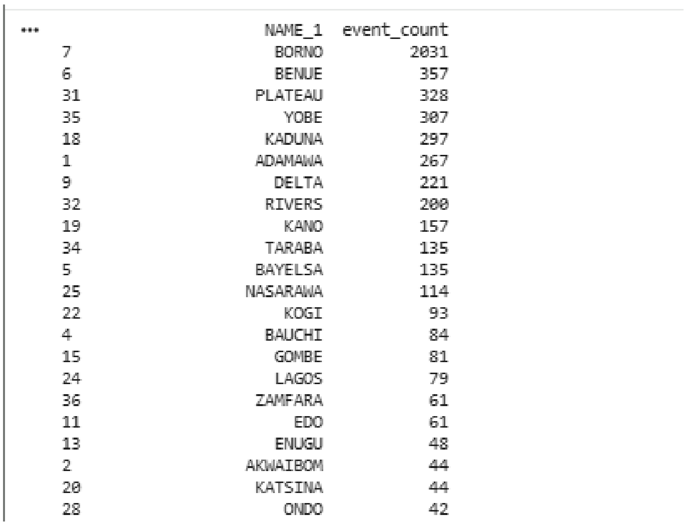


```
Input:
Historical conflict data $\mathcal{D}^{conf}_{2010:2025}$,
satellite imagery $\mathcal{D}^{sat}$ (NDVI, land-use/land-cover),
climatic data $\mathcal{D}^{clim}$ (rainfall, temperature),
socio-economic data $\mathcal{D}^{soc}$,
geospatial layers $\mathcal{G}$ (boundaries, grazing routes, water bodies).

Output:
Conflict risk map $\mathcal{R}$,
probabilistic risk scores $P(y = 1|x)$,
early warning indicators for high-risk zones.

Begin

Define study objective and success criteria
Reformulate conflict prediction as a spatio-temporal learning task:
    $f : X(s,t) \to P(y = 1 \mid s,t)$
where $s$ denotes spatial unit and $t$ denotes time.

Collect datasets:
    $\mathcal{D} = \{\mathcal{D}^{conf}, \mathcal{D}^{sat}, \mathcal{D}^{clim}, \mathcal{D}^{soc}, \mathcal{G}\}$
Perform exploratory analysis:
    • Temporal trend analysis of $\mathcal{D}^{conf}$
    • Spatial clustering analysis over $\mathcal{G}$
    • Seasonal correlation analysis between conflicts and environmental variables
```

```
for each dataset $D \in \mathcal{D}$ do
    • Remove duplicates
    • Handle missing values
    • Standardize timestamps
    • Validate geographic coordinates
end for
Project all datasets into common CRS:
    $\mathcal{G} \rightarrow \mathcal{G}^*$
Resample raster datasets to uniform resolution

for each spatial unit $s \in \mathcal{S}$ and time $t \in \mathcal{T}$ do
    • Compute vegetation index:
        $NDVI(s,t)$
    • Extract climatic features:
        $C(s,t) = \{rainfall, temperature\}$
    • Compute spatial proximity features:
        $d_{water}(s),\ d_{grazing}(s),\ d_{farm}(s)$
    • Derive historical conflict features:
        $H(s,t) = \{density, lag(t-k)\}$
    • Construct feature vector:
        $x_{s,t} = [NDVI, C, d_s, H, socio\text{-}economic]$
    • Assign label:
        $y_{s,t} = \begin{cases} 1, & \text{if conflict occurs} \\ 0, & \text{otherwise} \end{cases}$
end for
Construct unified dataset:
    $\mathcal{D}^* = \{(x_{s,t}, y_{s,t})\}$
Split dataset:
    $\mathcal{D}^* = \mathcal{D}_{train} \cup \mathcal{D}_{test}$
Initialize model set:
    $\mathcal{M} = \{m_1, m_2, m_3\}$
for each model $m \in \mathcal{M}$ do
    • Train:
        $m \leftarrow \text{Train}(m, \mathcal{D}_{train})$
    • Hyperparameter tuning via cross-validation
    • Predict:
        $\hat{y}_{s,t} = m(x_{s,t})$
end for

for each model $m \in \mathcal{M}$ do
    • Compute metrics:
        $\text{Accuracy, Precision, Recall, } F1,\ AUC$
    • Evaluate spatial consistency:
        $\text{Overlap}(\hat{Y}, Y)$
end for
Select optimal model:
    $m^* = \arg\max_{m \in \mathcal{M}} \mathcal{P}(m)$
Apply optimal model:
    $P(y = 1 | x_{s,t}) = m^*(x_{s,t})$
Generate risk classification:
    $\mathcal{R}(s,t) = \begin{cases} Low, & P < \tau_1 \\ Medium, & \tau_1 \le P < \tau_2 \\ High, & P \ge \tau_2 \end{cases}$
Integrate into GIS system:
    • Produce spatial risk maps
    • Visualize temporal hotspots
    • Enable early warning signals
Disseminate outputs to stakeholders

End
```

3.0 RESULTS AND DISCUSSION

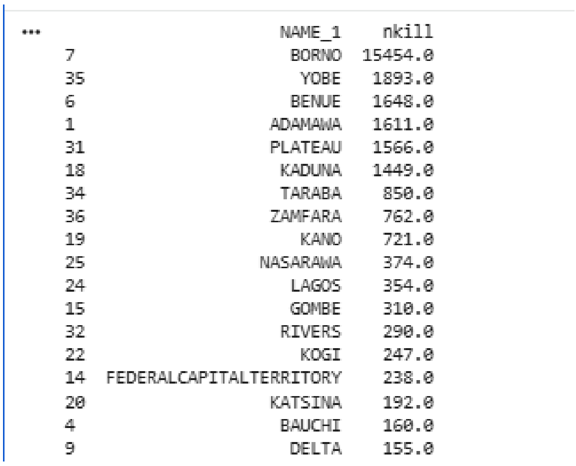
The spatial distribution of conflict events across Nigerian states shows that Borno (2031), Benue (357), and Plateau (328) emerged as dominant attack hotspots, having substantially higher numbers of recorded events as shown in Figure 1. The persistence of high event counts in these states suggests entrenched insecurity dynamics rather than episodic or isolated outbreaks of violence. In contrast, several southern states record comparatively few events, indicating either relative stability or different forms of insecurity.



```
***              NAME_1  event_count
    7             BORNO         2031
    6             BENUE          357
    31          PLATEAU          328
    35             YOBE          307
    18           KADUNA          297
    1           ADAMAWA          267
    9             DELTA          221
    32           RIVERS          200
    19             KANO          157
    34           TARABA          135
    5           BAYELSA          135
    25         NASARAWA          114
    22             KOGI           93
    4            BAUCHI           84
    15            GOMBE           81
    24            LAGOS           79
    36          ZAMFARA           61
    11              EDO           61
    13            ENUGU           48
    2          AKWAIBOM           44
    20          KATSINA           44
    28            ONDO            42
```

Figure 1: Conflict Events Per State.

Fatality counts offer deeper insight into conflict severity. Aggregating fatalities at the state level reveals that states with the highest number of events are not always those with the highest death tolls. Some states exhibit fewer incidents but significantly higher total fatalities, indicating the presence of particularly lethal forms of violence such as mass-casualty attacks as shown in Figure 2.



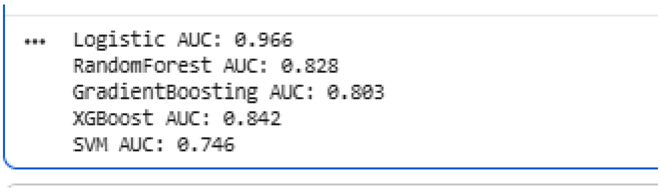
```
•••                    NAME_1     nkill
    7                    BORNO  15454.0
    35                    YOBE   1893.0
    6                    BENUE   1648.0
    1                  ADAMAWA   1611.0
    31                 PLATEAU   1566.0
    18                  KADUNA   1449.0
    34                  TARABA    850.0
    36                 ZAMFARA    762.0
    19                    KANO    721.0
    25                NASARAWA    374.0
    24                   LAGOS    354.0
    15                   GOMBE    310.0
    32                  RIVERS    290.0
    22                    KOGI    247.0
    14 FEDERALCAPITALTERRITORY    238.0
    20                 KATSINA    192.0
    4                   BAUCHI    160.0
    9                    DELTA    155.0
```

Figure 2: States by Fatalities (nkill)

To assess the feasibility of predicting conflict outcomes, supervised learning models such as Logistic Regression, Random Forest, Gradient Boosting, XGBoost, and support vector machine (SVM) were trained and evaluated using probabilistic outputs. Model performance was primarily assessed using the area under the receiver operating characteristic curve (AUC), which measures the ability of each model to distinguish between high-risk and low-risk conflict outcomes.

Across all models evaluated, ensemble-based methods consistently outperform linear and kernel-based approaches. The gradient-boosted tree model and the extreme gradient boosting model demonstrate the highest discriminative power, achieving the strongest AUC values. These results indicate that nonlinear interactions among predictors play a critical role in determining conflict risk and that such relationships are not adequately captured by simpler parametric models.

Logistic regression performs moderately well, suggesting that some predictors have stable, monotonic relationships with the outcome variable. However, its comparatively lower performance underscores the limitations of assuming linear effects in complex conflict systems. The support vector machine model also underperforms relative to tree-based ensembles, likely due to sensitivity to feature scaling and reduced interpretability in high-dimensional spaces as shown in Figure 3.



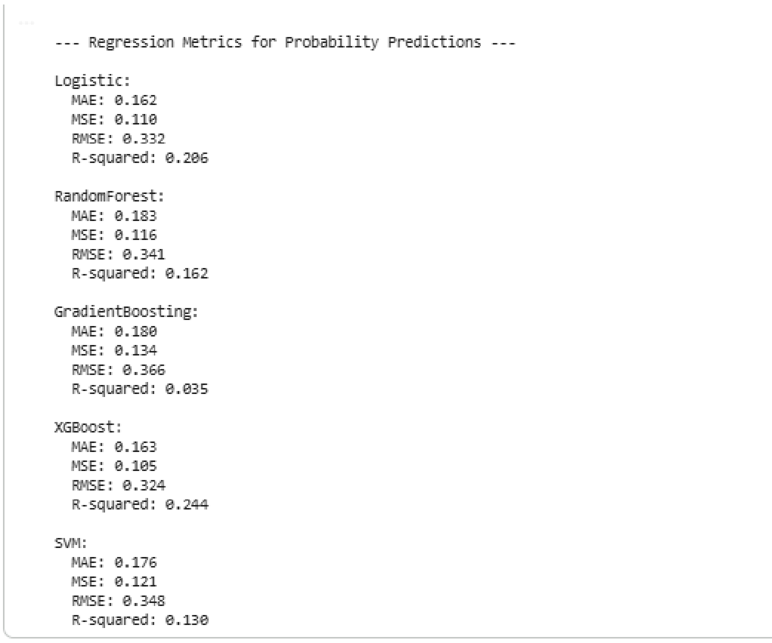
```
•••  Logistic AUC: 0.966
     RandomForest AUC: 0.828
     GradientBoosting AUC: 0.803
     XGBoost AUC: 0.842
     SVM AUC: 0.746
```

Figure 3: AUCs for the Five Machine Learning Models

Beyond classification accuracy, the quality of predicted probabilities was evaluated using regression-oriented metrics, including mean absolute error, mean squared error, root mean squared error, and the coefficient of determination. These metrics provide insight into how closely predicted probabilities align with observed outcomes.

The extreme gradient boosting model demonstrates superior performance, yielding the lowest error metrics. This indicates that the model not only ranks observations correctly but also produces more reliable probability estimates. Logistic regression and random forest models exhibit moderate error levels, while gradient boosting without extreme optimization shows weaker calibration.

Notably, all models achieve relatively modest R-squared values, reflecting the inherently stochastic nature of violent conflict as shown in Figure 4. Conflict events are influenced by numerous unobserved political, social, and tactical factors, which limits the extent to which purely quantitative models can fully explain outcomes. Nevertheless, the observed performance levels are consistent with, and in some cases exceed, benchmarks reported in comparable conflict prediction studies.



```
--- Regression Metrics for Probability Predictions ---

Logistic:
  MAE: 0.162
  MSE: 0.110
  RMSE: 0.332
  R-squared: 0.206

RandomForest:
  MAE: 0.183
  MSE: 0.116
  RMSE: 0.341
  R-squared: 0.162

GradientBoosting:
  MAE: 0.180
  MSE: 0.134
  RMSE: 0.366
  R-squared: 0.035

XGBoost:
  MAE: 0.163
  MSE: 0.105
  RMSE: 0.324
  R-squared: 0.244

SVM:
  MAE: 0.176
  MSE: 0.121
  RMSE: 0.348
  R-squared: 0.130
```

Figure 4: Regression Metrics for Probability Predictions

The models under evaluation include Logistic Regression, Random Forest, Gradient Boosting, Extreme Gradient Boosting (XGBoost), and Support Vector Machines (SVM). Using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2) as evaluation criteria, this study assessed and compared the predictive accuracy and explanatory power across models. The findings indicated that XGBoost demonstrated the strongest overall performance among the evaluated models, while Gradient Boosting showed comparatively weaker results as shown in Figure 5.

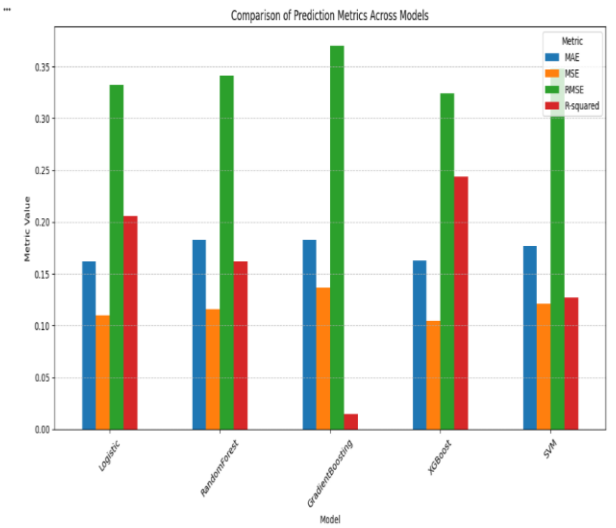


Figure 5: Comparison of Prediction Metrics Across Models

4.0 CONCLUSION

This paper set out to advance the empirical understanding and predictive analysis of farmer-herder conflicts in Nigeria by integrating geospatial methods, event-level conflict data, and modern machine-learning techniques. The study moved beyond descriptive conflict mapping to develop state-level early-warning approach grounded in historical patterns of violence. The core contribution of the research lies in demonstrating how spatially explicit conflict data, when systematically preprocessed and analyzed using machine-learning models, can yield meaningful insights into both the drivers and future risks of farmer-herder clashes.

Methodologically, the paper established a rigorous geospatial preprocessing pipeline that harmonized administrative boundary data from the GADM database with ACLED-style conflict events. Rather than attempting to predict individual incidents or simulate violence at fine spatial resolutions approaches that raise both ethical and methodological concerns, the study focused on estimating relative risk across Nigerian states over time. This framing aligns closely with early-warning and prevention objectives, emphasizing probabilistic risk assessment rather than deterministic forecasting.

Furthermore, the paper conducted a systematic comparison of five widely used machine-learning models: Logistic Regression, Random Forests, Gradient Boosting, Support Vector Machines, and XGBoost. This comparative approach addressed an important gap in the conflict-prediction literature, where model choice is often insufficiently justified or evaluated. Results from the model comparison demonstrated that ensemble tree-based methods, particularly XGBoost and Gradient Boosting consistently outperformed linear and kernel-based alternatives across multiple evaluation metrics, including AUC-ROC, and F1-score. The inclusion of Logistic Regression as a baseline model further strengthened the analysis by providing a transparent benchmark against which more complex methods could be evaluated.

Among the evaluated models, XGBoost emerged as the strongest overall performer, exhibiting superior discrimination, robustness to class imbalance, and stable probabilistic calibration. These properties make it particularly well suited to conflict prediction tasks, where violent events are relatively rare but highly consequential. In conclusion, this paper provides a strong foundation for research at the intersection of geospatial analysis, machine learning, and conflict studies.

In the future, while XGBoost emerged as the strongest individual model, combining predictions from multiple models through stacking or weighted ensembles may yield further performance gains and increased robustness.

REFERENCES

ACLED, G. (2019). Armed conflict location & event data project (ACLED) codebook.

Adelabu, I., & Wang, L. (2025). Forest Resilience and Vegetation Dynamics in Southwest Nigeria: Spatiotemporal Analysis and Assessment of Influencing Factors Using Geographical Detectors and Trend Models. *Forests*, *16*(5), 811.

Ajogbeje, K., & Sylwester, K. (2024). How conflict affects education: Differences between Boko Haram and Farmer-Herder conflicts in Nigeria. *World Development*, *177*, 106540.

Amoah, E. A., Bloomer, J., & Campbell, T. (2025). Beyond resource scarcity: developing an integrated framework for analysing farmer-herder conflicts in the Global South. *Pastoralism: Research, Policy and Practice*, *15*, 15245.

Anil, B., Jyothi, M., & Suma, K. (2026). Machine Learning Models for Climate Change Impact Prediction. *Machine Learning*, *4*(1).

Ayeb-Karlsson, S., Hayward, G., & Kniveton, D. (2025). 'We Herders Are Often Chased About by Drought': A Systems Analysis of Natural Resource Degradation Within the Climate–(Im) mobility–Violence–Health Nexus in Sahel. *Earth*, *6*(1), 11.

Blench, R., & Hassan, U. (2025). Grazing Reserves, Stock Routes, and Pastoral Resources in Nigeria.

Chukwu, H. O. (2025). Democracy and Inter-Group Relations In Nigeria, 1999-2019. *Uzu Journal of History and International Studies (UJHIS)*, *10*(1).

Desta, T. T. (2025). Sustainable Intensification of Subsistence Agriculture, Socioeconomic Transformation, and Food Security. *Advances in Agriculture*, *2025*(1), 6671209.

Eke, O. G., Moudry, J., Eze, F. O., Obazi, S. A., Ifoh, I. P., Mukosha, C. E., Ntezimana, M.G. & Muhammad, A. (2025). Climate-Driven Conflicts in Nigeria: Farmers' Strategies for Coping with Herders' Incursion on Crop Lands. *Sustainability*, *17*(24), 11316.

Haaga, P. T., Sule, P. E., Idachaba, P. A., & Bisong, B. E. (2025). Farmer-herder conflicts and food security in north-central Nigeria. *Advances in Law, Pedagogy, and Multidisciplinary Humanities*, *3*(1), 1-17.

Mardikaningsih, R. (2025). Reconstructing the Earth's Social Ecosystem through Socio-Ecological Inquiry in the Climate Crisis Era. *Bulletin of Science, Technology and Society*, *4*(1), 49-56.

Mohamed, A. A., Chauke, P. K., & Palanisamy, K. (2026). Pastoral communities and drought: a scoping review of drought hazards and management practices. *Hydrological Sciences Journal*, 1-12.

Njoya Ngetar, S., Rwandarugali, S., & Singh, P. (2025). Geovisualization and Spatial Modelling of Armed Conflicts in the Great Lake Region of Central and East Africa: Ngetar et al. *Applied Spatial Analysis and Policy*, *18*(2), 73.

Nwagbo, S. N., Francis, A. C., & Ezeakoafor, A. (2025). Insecurity and Attainment of Sustainable Development Goals 1 and 2 in Nigeria. *Advances in Social Sciences Research Journal*, *12*(05).

Okonkwo, N. S., & Akanji, T. A. (2025). Political Violence in Nigeria: Examining Dynamics, Impacts, and Sociopolitical Realities. *Journal of World Affairs*, *1*(3), 346-364.

Olamide Sowale, A. (2025). Fulani herder-farmer identity conflict in West Africa with a particular focus on security implications for Nigeria. *African Identities*, *23*(4), 1208-1227.

Sukare, B. M., & Abdullahi, A. (2025). Herders–Crop Farmers Conflicts in Nigeria: Issues, Challenges and Prospects. *Journal of Innovative Social Science and Humanities Research*.

Tawane, A. A. (2025). Banditry and Insecurity in Pastoral Kenya: A Critical Discourse.

Thompson, A., & Petit, G. (2026). Integrating Public Health Indicators into Urban Planning Frameworks using Geographic Information Systems (GIS) and Spatial Analysis. *Computer Science Bulletin*, *9*(1), 71-78.

Udoh, I. J., Alabi, O. I., & Faleye, O. J. (2025). A gradient boosting machine model of economic and socio-economic determinants of terrorism: The Northern Nigerian perspective. *SSR Journal of Artificial Intelligence (SSRJAI)*, *2*(1), 24-38.

Usman, M., & Nichol, J. E. (2022). Changes in agricultural and grazing land, and insights for mitigating farmer-herder conflict in West Africa. *Landscape and Urban Planning*, *222*, 104383.

Yunusa, E., & Owoyemi, J. O. (2025). Rural Development and Resource-Based Conflict In North-Central Nigeria: Escalation, Consequences And Management.