© JAN 2026 | IRE Journals | Volume 9 Issue 7 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I7-1713576

AI-Based Predictive Surveillance of Malaria in Northern Nigeria Using Climate and Demographic Data.

JIBRIN ABDULLAHI DALLATU[1], ALHAJI SALEH ISYAKU[2], IBRAHIM IBRAHIM JAURO[3], USMAN SHETTIMA USMAN[4], SHEHU ABUBAKAR UMAR[5], ADAM IBRAHIM GARBA[6], ABUBAKAR SADIQ YARIMA[7], MUHAMMAD ALANJIRO[8], UMAR MUHAMMAD FAISAL[9]

[1, 5]Informatics and Computer Engineering, National University of Science and Technology (MISIS) Moscow Russia

[2]Department of Epidemiology and Evidence Based Medicine, First Moscow State Medical University (Sechenov) Moscow Russia

[3]Department of Optometry, Bayero University Kano Nigeria

4, 6, 7, 8Department of Pharmacy, Integral University Lucknow India

[9]Department of Medicine, Far Eastern Federal University Vladivostok Russia

***Abstract-** Malaria remains a major public health challenge in North Nigeria, where climate conditions and demographic pressures contribute to recurring outbreaks (WHO, 2023). Traditional surveillance systems often struggle with delays and limited data, making it difficult to predict malaria trends accurately (Nigeria Malaria Indicator Survey, 2021). This study applies Artificial Intelligence (AI) techniques to explore how climate and demographic information can support early prediction of malaria cases in the region. A regional dataset containing monthly temperature, rainfall, air quality index, UV index, population density, and malaria incidence was analyzed. Machine learning models were developed using climate lag features, seasonal patterns, and demographic indicators to improve forecasting performance, following approaches successfully applied in previous climate-disease modeling studies. The results show that rainfall, temperature, and population density are strong predictors of malaria incidence in North Nigeria, consistent with findings from prior ecological and epidemiological research. The AI-based model produced reliable monthly forecasts, demonstrating the potential of integrating climate and demographic data for predictive malaria surveillance. This approach provides a practical tool that can enhance early warning systems and support better planning and prevention efforts in North Nigeria, aligning with calls for innovative, data-driven malaria control strategies across Africa.*

***Keywords:** Malaria, Artificial Intelligence, Machine Learning, Climate, Demographics, North Nigeria*

I. INTRODUCTION

Malaria remains one of the most persistent and deadly infectious diseases in subSaharan Africa, where climatic and demographic conditions create an environment highly favorable for mosquito-borne transmission (WHO, 2023). In Nigeria, malaria accounts for a large proportion of outpatient visits, hospital admissions, and childhood illnesses, making the country responsible for the highest malaria burden globally (NMIS, 2021). The northern region of the country, characterized by distinct wet and dry seasons, faces recurring outbreaks driven by fluctuations in rainfall, temperature, humidity, and population-related factors. Seasonal climate patterns in North Nigeria frequently influence mosquito breeding cycles and parasite development rates, resulting in periodic surges in malaria cases (Paaijmans et al., 2009; Teklehaimanot et al., 2004). Demographic pressures, such as rising population density, uneven access to healthcare, and limited public health funding, further exacerbate vulnerability in many northern communities (Arogundade et al., 2011).

Effective malaria control depends heavily on timely and reliable surveillance systems. However, traditional surveillance methods in Nigeria often rely on manual reporting processes and incomplete health facility data (WHO, 2018). These limitations lead to delays in outbreak detection, making it difficult for healthcare authorities to respond early enough to prevent transmission peaks. For regions like North

Nigeria, where many communities lack consistent health coverage or real-time reporting systems, the need for innovative, data-driven solutions is especially critical. The complexity of malaria transmission driven by interactions among climate, population, and environmental factors requires analytical approaches capable of processing diverse and dynamic datasets (Gething et al., 2016).

Advances in Artificial Intelligence (AI) and machine learning now offer powerful tools for predictive health surveillance. AI models have been successfully applied in several vector-borne disease studies, showing superior performance in detecting nonlinear interactions and forecasting outbreaks (Yang et al., 2020). By analyzing climatic variables such as rainfall, temperature, air quality, and UV index alongside demographic indicators like population density and healthcare expenditure, AI models can identify patterns that may not be easily detectable through traditional methods (Adde et al., 2020). These models can learn from historical data and generate accurate forecasts of malaria incidence, allowing health agencies to plan targeted interventions, allocate resources more efficiently, and issue early warnings ahead of transmission spikes. This approach is particularly useful in regions where detailed local data is limited, and proxy indicators must be used to understand disease trends (Bhatt et al., 2015). Recent studies have demonstrated the growing applicability of artificial intelligence and machine learning in medical prediction tasks, including disease surveillance and drug safety assessment (Isyaku et al., 2025; Yang et al., 2020).

This study aims to develop a practical and accessible AI-based predictive surveillance model for malaria in North Nigeria using available climate and demographic data. By integrating environmental and population-related factors into machine learning algorithms, the research seeks to evaluate their ability to predict malaria incidence and assess the most influential variables driving transmission in the region. The modeling framework incorporates lagged climate effects, seasonal trends, and demographic pressures to produce reliable monthly forecasts. Ultimately, the study demonstrates how AI can strengthen malaria surveillance in resource-limited settings, providing a foundation for more responsive public health strategies and helping to reduce the burden of malaria in North Nigeria (WHO, 2023).

BACKGROUND OF THE STUDY

Malaria is one of the most significant public health challenges in Nigeria and across sub-Saharan Africa, where climatic and demographic conditions promote year-round transmission (WHO, 2023). Nigeria accounts for the highest number of malaria cases and deaths globally, largely due to the dominance of *Plasmodium falciparum* and the prevalence of highly efficient mosquito vectors such as *Anopheles gambiae* (Bhatt et al., 2015). The northern region of Nigeria is particularly vulnerable due to its distinct environmental characteristics, including seasonal fluctuations in temperature, rainfall, and humidity, which strongly affect mosquito breeding and parasite development (Paaijmans et al., 2009). During the rainy season, stagnant water bodies serve as mosquito breeding sites, while temperature variations influence vector survival and the speed of parasite development within mosquitoes. These seasonal shifts often contribute to predictable surges in malaria cases (Teklehaimanot et al., 2004).

Number of Northern States Found: 19

Demographic pressures further intensify malaria transmission in North Nigeria. Rapid population growth, urban overcrowding, and disparities in access to healthcare services increase the risk of infection and hinder effective disease control (Arogundade et al., 2011). Many communities in the region rely on limited health infrastructure, resulting in delayed diagnosis, underreporting, and inadequate treatment.

Additionally, socioeconomic constraints such as poverty, limited public health investment, and inadequate preventive measures continue to contribute to the persistent malaria burden (NMIS, 2021). These conditions emphasize the importance of strengthening surveillance and improving early detection to better manage transmission cycles.

Traditional malaria surveillance systems in Nigeria often rely on manual reporting, health facility records, and community health officer submissions. However, these systems are frequently weakened by incomplete data, delayed reporting, and lack of real-time monitoring capabilities (WHO, 2018). As a result, outbreaks may not be detected until they have already escalated, reducing the opportunity for timely intervention. This challenge is particularly relevant in the northern region, where many communities are remote and have limited access to health information systems. Given the complexity of malaria transmission which is influenced by interactions among climate, environment, population density, and socio-economic factors there is a growing need for predictive tools capable of integrating multiple variables to guide proactive public health decisions (Gething et al., 2016).

Recent advancements in Artificial Intelligence (AI) have opened new opportunities for enhancing malaria surveillance, particularly in data-limited settings. Machine learning algorithms are capable of analyzing large, complex datasets and identifying nonlinear relationships that traditional statistical models may overlook (Yang et al., 2020). AI-based models have been successfully applied to predict outbreaks of vector-borne diseases such as malaria, dengue fever, and Zika virus by leveraging climate, demographic, and environmental indicators (Adde et al., 2020). These approaches offer significant potential for regions like North Nigeria, where proxy data such as climate and population indicators can be used to build reliable predictive systems even when local health data is limited.

This study leverages these technological advancements by developing an AI-based predictive surveillance model for malaria in North Nigeria. By integrating climate indicators (e.g., temperature, rainfall, air quality, and UV index) with demographic factors (e.g., population density and healthcare expenditure), the model seeks to identify key predictors of malaria transmission and provide accurate monthly forecasts. This approach aims to support early warning systems, guide resource allocation, and strengthen public health responses in the region. By addressing the limitations of existing surveillance methods and harnessing the power of AI, this study contributes to more effective and adaptive malaria control strategies in resource-constrained settings (WHO, 2023).

PROBLEM STATEMENT

Malaria continues to impose a significant public health burden in North Nigeria, where climatic and demographic conditions create ideal environments for mosquito breeding and disease transmission (WHO, 2023). Despite ongoing control efforts, the region experiences recurring outbreaks driven by seasonal fluctuations in rainfall, temperature, and humidity factors known to influence mosquito survival and *Plasmodium* parasite development (Paaijmans et al., 2009; Teklehaimanot et al., 2004). Demographic pressures such as rapid population growth, high population density, and limited access to healthcare services further intensify vulnerability and hinder timely diagnosis and treatment (Arogundade et al., 2011).

Traditional malaria surveillance systems in Nigeria often rely on incomplete health facility reports and manual data collection, resulting in delayed detection of outbreaks and underreporting of cases (WHO, 2018). These limitations prevent health authorities from identifying early warning signals and responding proactively. In many northern communities, gaps in real-time data and weak health infrastructure further complicate surveillance efforts (NMIS, 2021). This creates an urgent need for innovative methods capable of predicting malaria trends and supporting timely decision-making.

Although climate and demographic data have been shown to strongly influence malaria transmission (Bhatt et al., 2015; Gething et al., 2016), these variables are not currently integrated into predictive surveillance systems for North Nigeria. Advances in Artificial Intelligence (AI) and machine learning offer powerful analytical tools for forecasting infectious diseases, but their application to malaria prediction in Nigeria remains limited (Yang et al., 2020). The absence of AI-driven malaria forecasting tools in the region creates a significant gap in early warning capabilities.

Therefore, there is a pressing need to develop an AI-based predictive model that leverages climate and demographic indicators to forecast malaria incidence in North Nigeria. Such a system would support proactive public health responses, improve resource allocation, and strengthen malaria control efforts in this high-burden region.

AIM OF THE STUDY

The aim of this study is to develop an AI-based predictive surveillance model for malaria in North Nigeria using climate and demographic variables known to influence transmission, including temperature, rainfall, air quality, UV index, population density, and healthcare expenditure (Bhatt et al., 2015; WHO, 2023). The study seeks to evaluate the effectiveness of machine learning algorithms in forecasting malaria incidence and providing early warning insights to support public health planning in resource-limited settings.

OBJECTIVES OF THE STUDY

To achieve the aim of the study, the following objectives were established:

1. To analyze regional climate and demographic data from North Nigeria and examine their relationship with malaria incidence, following established climate–disease modeling frameworks (Paaijmans et al., 2009; Gething et al., 2016).
2. To develop machine learning models capable of predicting monthly malaria cases using integrated environmental and demographic indicators, building on successful AI applications in disease forecasting (Yang et al., 2020; Adde et al., 2020).
3. To identify the most influential climate and demographic variables associated with malaria transmission in the region (Bhatt et al., 2015).
4. To evaluate the performance and accuracy of the predictive models using standard machine learning evaluation metrics (RMSE, MAE, R^2).
5. To propose an AI-based early warning framework that can support timely malaria prevention and improve decision-making in North Nigeria (WHO, 2023).

SIGNIFICANCE OF THE STUDY

Malaria remains a persistent public health challenge in North Nigeria, disproportionately affecting vulnerable populations and contributing to high levels of morbidity and mortality (WHO, 2023). Despite ongoing interventions, recurring outbreaks continue due to the strong influence of climate variability and demographic pressures in the region (Bhatt et al., 2015). This study is significant because it introduces an AI-based predictive surveillance model that combines climate and demographic indicators to forecast malaria trends an approach increasingly recognized as essential for improving disease control in climate-sensitive regions (Paaijmans et al., 2009).

The application of Artificial Intelligence (AI) and machine learning in public health has demonstrated promising results in predicting infectious diseases such as malaria, dengue, and influenza (Yang et al., 2020; Adde et al., 2020). However, such approaches are underutilized in Nigeria, where traditional surveillance systems often suffer from limited data availability, delayed reporting, and weak infrastructure (WHO, 2018). By leveraging regional climate and demographic proxy data, this study provides a scalable and practical solution for malaria prediction in settings where high-resolution or real-time datasets are limited (Gething et al., 2016).

The findings of this research also offer insights into the specific environmental and demographic factors that drive malaria transmission in North Nigeria. Identifying rainfall, temperature, and population density as key predictors aligns with established ecological evidence (Teklehaimanot et al., 2004). Understanding these predictors can help policymakers design more targeted interventions such as optimizing the timing of indoor residual spraying, distributing insecticidetreated nets before high-risk seasons, and planning healthcare resource allocation (NMIS, 2021).

Furthermore, the proposed AI-based model can enhance early warning systems, enabling health authorities to anticipate outbreaks before they escalate. Early detection is critical for reducing transmission, preventing severe cases, and saving lives particularly in resource-limited settings (WHO, 2018). This study therefore contributes to strengthening public health resilience and advancing innovative approaches to malaria control in regions where traditional systems face significant challenges.

SCOPE OF THE STUDY

This study focuses on using regional climate and demographic data to develop an AI-based predictive model for malaria in North Nigeria. The geographic scope covers the northern zone, which experiences distinct climatic patterns that strongly influence malaria transmission (Paaijmans et al., 2009). The temporal scope is defined by the available dataset, which includes monthly observations of rainfall, temperature, air quality, UV index, population density, healthcare expenditure, and malaria incidence.

The study uses aggregated regional data due to limitations in obtaining detailed local datasets from health facilities, consistent with known challenges in Nigeria's surveillance systems (WHO, 2018). As such, the model relies on proxy data to estimate malaria risk, a method supported in regions where local reporting is incomplete or inconsistent (Gething et al., 2016). The scope is limited to analyzing climate and demographic predictors only; other factors such as land-use patterns, sanitation, human mobility, and vector-control interventions are excluded due to data unavailability (Bhatt et al., 2015).

Methodologically, the study employs machine learning techniques including XGBoost and Random Forest to identify patterns and forecast malaria incidence. The focus is on monthly predictions rather than real-time or daily forecasting, as monthly granularity aligns with the structure of the dataset and existing malaria monitoring practices (NMIS, 2021). While the model demonstrates strong predictive potential, it is not intended to replace traditional surveillance systems but to complement them by providing early warning insights and enhancing decision-making capacity in resource-limited settings (WHO, 2023).

LIMITATIONS OF THE STUDY

Although this study offers valuable insights into malaria prediction using AI-based methods, several limitations must be acknowledged. First, the study relies on regional proxy data rather than high-resolution local datasets. Due to limited availability of community-level malaria records in North Nigeria an issue widely reported in surveillance research the dataset may not fully capture micro-level variations in transmission (WHO, 2018; NMIS, 2021). While proxy data are commonly used in epidemiological modeling, their use may reduce precision in localized predictions, especially in regions with heterogeneous ecological or demographic characteristics (Gething et al., 2016).

Second, the predictive model uses monthly aggregated data, which limits the ability to detect short-term outbreaks triggered by sudden environmental shocks. Daily or weekly data could potentially improve fine-scale predictions, but such datasets are not consistently available in Nigeria's public health records (WHO, 2023). The reliance on monthly averages may smooth seasonal patterns and obscure short-lived spikes in malaria incidence an issue observed in similar climate-driven disease modeling studies (Paaijmans et al., 2009).

Third, the study focuses primarily on climate and demographic predictors, excluding other important factors such as land-use changes, vegetation cover, human mobility, mosquito insecticide resistance, socio-economic conditions, and vector-control interventions (Bhatt et al., 2015). These variables have been shown to influence malaria transmission but could not be incorporated due to data limitations. Their exclusion means that the model may not fully represent the complex ecological and behavioral dimensions of malaria spread (Gething et al., 2016).

Additionally, the study employs machine learning models XGBoost and Random Forest that perform well with tabular environmental data. However, more advanced temporal models such as LSTM or hybrid spatiotemporal neural networks may improve predictions but require larger, more detailed datasets (Yang et al., 2020). The absence of real-time or higher-frequency data limits the exploration of these more sophisticated techniques.

Finally, although the model demonstrates strong predictive capacity, it is not a replacement for traditional malaria surveillance systems. Instead, it is intended to complement existing health reporting processes by providing early warning capabilities.

Predictive accuracy may vary depending on climate anomalies, environmental disturbances, or changes in local mosquito behavior factors known to challenge climate-driven disease models (Teklehaimanot et al., 2004).

Despite these limitations, the study provides a strong foundation for integrating AI into malaria surveillance in North Nigeria and offers practical insights for future system improvements.

II. LITERATURE REVIEW

2.1 Overview of Malaria Burden in Sub-Saharan Africa

Malaria remains one of the most significant public health challenges in sub-Saharan Africa, accounting for the majority of global malaria morbidity and mortality. According to the World Health Organization (WHO), Africa is responsible for approximately 95% of global malaria cases and 96% of malaria deaths (WHO, 2023). This high disease burden is largely due to the dominance of *Plasmodium falciparum*, the most lethal malaria parasite, and the prevalence of highly efficient mosquito vectors, including *Anopheles gambiae* and *Anopheles funestus*. These vectors thrive in warm, humid environments and reproduce rapidly under favorable ecological conditions, making many regions of Africa highly vulnerable to sustained transmission (Bhatt et al., 2015).

2.2 Malaria Situation in Nigeria and North Nigeria

Nigeria consistently reports the highest malaria burden globally, with the northern region particularly affected due to its distinct climate profile (Nigeria Malaria Indicator Survey, 2021). North Nigeria experiences a long dry season and a short but intense rainy season. Rainfall produces numerous stagnant water bodies, which serve as breeding sites for mosquitoes, leading to seasonal spikes in malaria transmission. Although transmission decreases during the dry season, it rarely falls to zero because certain mosquito species adapt to arid and semi-arid conditions (Gething et al., 2016). Socio-economic challenges such as poverty, inadequate healthcare access, poor housing, and rapid population growth further intensify malaria risk in northern communities (Arogundade et al., 2011), reinforcing the need for improved surveillance and predictive modeling.

2.3 Climate Factors Influencing Malaria Transmission

Climate plays a central role in shaping malaria transmission intensity and distribution. Rainfall contributes directly to mosquito breeding by creating larval habitats, and even minor increases in rainfall can substantially expand breeding sites. Temperature influences mosquito survival, biting frequency, and the speed at which *Plasmodium* parasites develop inside mosquitoes (the extrinsic incubation period). Optimal transmission generally occurs between 20°C and 30°C (Paaijmans et al., 2009). Higher humidity enhances mosquito longevity, thus increasing the likelihood of parasite transmission.

Studies across Africa have documented these relationships. Paaijmans et al. (2009) showed that mosquito development rates are extremely sensitive to temperature variations. Teklehaimanot et al. (2004) found that rainfall variability strongly correlates with malaria incidence in East African highlands. Similar patterns have been observed in North Nigeria, where malaria cases rise sharply following the onset of the rainy season, strengthening the case for using climate variables as early-warning indicators.

2.4 Demographic Factors and Malaria Risk

Demographic conditions significantly influence malaria transmission patterns. Population density affects the frequency of human vector contact, increasing risk in densely populated areas (Arogundade et al., 2011). Urbanization can either increase or decrease malaria risk depending on environmental management, housing quality, and infrastructure. Access to healthcare plays a crucial role in early diagnosis and treatment key interventions that prevent complications and curb transmission. Healthcare expenditure and investment in public health services similarly influence community resilience against malaria outbreaks (WHO, 2023).

In North Nigeria, population growth, inequitable healthcare distribution, and inaccessible medical facilities contribute to persistent malaria transmission. Arogundade et al. (2011) observed that many rural northern communities face delays in receiving malaria treatment, increasing the likelihood of severe disease and mortality. These demographic factors must be incorporated into predictive modeling to account for human-related influences on malaria spread.

2.5 Limitations of Traditional Malaria Surveillance Systems

Traditional malaria surveillance in Nigeria relies on health facility records, rapid diagnostic test (RDT) reporting, and manual submissions from rural clinics. However, these systems face challenges including incomplete data, delayed reporting, and under-detection of communitylevel cases (WHO, 2018). According to national surveillance assessments, nearly half of suspected malaria cases in rural northern communities go unreported due to accessibility and data quality issues (NMIS, 2021). This makes real-time detection difficult and weakens early response strategies. These limitations highlight the need for predictive systems that can function even when local surveillance data are sparse or inconsistent.

2.6 AI and Machine Learning in Disease Prediction

Artificial Intelligence (AI) and machine learning (ML) have emerged as powerful tools for analyzing complex epidemiological datasets. ML algorithms such as Random Forest, XGBoost, Support Vector Machines (SVM), Long Short-Term Memory (LSTM) networks, and Temporal Convolutional Networks have been successfully applied to forecast infectious diseases including influenza, dengue, COVID-19, and malaria (Yang et al., 2020). These models can detect nonlinear relationships, identify hidden patterns, and capture temporal dynamics that traditional statistical techniques may overlook.

Studies show that ML outperform conventional regression models for malaria prediction, especially when incorporating satellite climate data and large-scale environmental variables (Adde et al., 2020; Weiss et al., 2019). Furthermore, models using lagged climate features such as previous-month rainfall and temperature achieve higher accuracy because they account for delayed climatic effects on mosquito development and parasite incubation (Teklehaimanot et al., 2004). AI-based predictive modeling has also been successfully applied beyond infectious disease surveillance, such as in computational prediction of drug toxicity and adverse drug reactions, highlighting the versatility of machine learning in healthcare applications (Isyaku et al., 2025).

2.7 Climate Demographic Data Integration in Malaria Prediction

Recent research strongly supports the integration of climate and demographic data for improved malaria forecasting accuracy. Demographic indicators moderate climate-driven malaria risks: for example, two areas with identical rainfall may experience different malaria outcomes due to differences in housing conditions, healthcare access, or population density (Bhatt et al., 2015). Integrated modeling frameworks that combine environmental and human-related variables have demonstrated superior performance across Africa (Gething et al., 2016).

This is particularly important in regions like North Nigeria, where climate alone cannot fully explain malaria patterns. When high-resolution or local datasets are missing, the use of proxy climate demographic indicators has been validated as an effective approach for disease modeling (Weiss et al., 2019). This supports the methodology used in the present study.

2.8 Application of AI in Malaria Surveillance in Africa

AI-based malaria surveillance is expanding across Africa. In Kenya, machine learning models have been used to detect malaria hotspots using satellite-derived temperature and rainfall data. In Ghana, rainfall-driven neural network models have been applied to predict seasonal malaria surges. However, Nigeria has seen limited deployment of AI for malaria forecasting, and most studies rely on national-level or hospital-based data rather than region-specific climate models (NMIS, 2021). This gap underscores the need for AI-driven early warning systems tailored to regional ecological conditions such as those in North Nigeria.

2.9 Research Gap and Justification

While numerous studies have examined malaria prediction globally and in other African regions, there is limited research applying AI specifically to malaria forecasting in North Nigeria. Localized datasets remain scarce, and previous studies seldom combine climate and demographic variables for regional modeling. This creates a significant research gap. By integrating climate and demographic proxies into a machine learning framework, this study addresses a critical need for early warning tools adapted to the ecological and socio-economic realities of North Nigeria (WHO, 2023).

III. METHODOLOGY

Research Design

This study adopts a quantitative research design using machine learning techniques to develop a predictive surveillance model for malaria in North Nigeria. Quantitative models are widely used in climate disease research because they allow integration of environmental and demographic indicators to understand disease patterns over time (Bhatt et al., 2015; Gething et al., 2016). Machine learning approaches were selected based on their proven effectiveness in forecasting infectious diseases and capturing nonlinear relationships in complex datasets (Yang et al., 2020; Adde et al., 2020).

Data Sources

The study uses a regional dataset containing monthly climate and demographic variables that are known to influence malaria transmission. These include rainfall, temperature, air quality index, UV index, population density, healthcare expenditure, and malaria incidence. Climate variables have been shown to directly impact mosquito breeding and parasite development cycles (Paaijmans et al., 2009; Teklehaimanot et al., 2004). Demographic indicators such as population density and healthcare access serve as proxies for exposure risk and health system capacity (Arogundade et al., 2011). Using regional proxy data is appropriate where local health data are limited or inconsistent (WHO, 2018; NMIS, 2021).

Data Preprocessing

Data preprocessing was conducted to ensure quality and consistency before model training. This process involved:

- Identifying and handling missing or inconsistent climate observations, following best practices for environmental dataset cleaning (Gething et al., 2016).
- Combining monthly and yearly data into timestamp features to support temporal modeling, a standard method in time-series epidemiology (Paaijmans et al., 2009).
- Normalizing continuous variables to enhance model performance, particularly for algorithms sensitive to variable scale (Yang et al., 2020).
- Structuring the dataset chronologically to avoid data leakage a critical requirement in predictive health modeling (Adde et al., 2020).

Geospatial preprocessing was also performed to ensure accurate representation of the study's target region. Using GADM administrative boundary shapefiles, the northern Nigerian states were extracted and mapped to validate the geographic scope of the dataset. This spatial verification step ensured that only the correct northern regions were included in the analysis and provided visual confirmation of the geographic boundaries used in the study. The geospatial map supported regional consistency across the climate, demographic, and malaria records before model development.

Feature Engineering

Feature engineering was used to enhance the predictive ability of the model. Key steps included:

- Lagged climate features (1–3 months) to reflect delayed effects of rainfall and temperature on mosquito lifecycles and malaria incidence, consistent with prior findings (Teklehaimanot et al., 2004; Gething et al., 2016).
- Rolling averages for rainfall and temperature to capture seasonal trends, a method commonly used in climate–disease modeling (Paaijmans et al., 2009).
- Seasonal indicators, recognizing that malaria transmission in northern Nigeria is strongly seasonal (NMIS, 2021).
- Derived demographic metrics, such as healthcare budget per capita, to approximate community-level healthcare access (Arogundade et al., 2011).

These engineered features improve the model's ability to detect both immediate and delayed climate effects.

Model Development

Three machine learning models were selected:

1. XGBoost Regression, known for its superior performance in environmental and epidemiological forecasting (Yang et al., 2020).

2. Random Forest Regression, valued for its robustness to noise and ability to model nonlinear relationships (Adde et al., 2020).
3. Linear Regression (baseline), used to benchmark improvements over traditional statistical techniques (Gething et al., 2016).

These models were trained to predict monthly malaria incidence based on engineered features.

Model Training and Validation

Models were trained using a chronological train–test split to prevent future information from leaking into the training process. Earlier years of data were used to train the algorithms, while the final 24 months were reserved for testing. This approach maintains the temporal structure of the malaria time series and is commonly recommended for climate-driven disease forecasting.

Default parameters were used for the Random Forest, HistGradientBoosting, and XGBoost models. Hyperparameter tuning was not performed due to computational limitations, but the selected algorithms are known to perform well even with default settings. Model performance was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2), which are standard evaluation metrics for epidemiological forecasting.(Yang et al., 2020).

Model Evaluation

Model performance was evaluated using:
- Mean Absolute Error (MAE)
- Root Mean Square Error (RMSE)
- Coefficient of Determination (R^2)

These metrics are standard in malaria forecasting and climate health modeling studies (Bhatt et al., 2015; Adde et al., 2020). Visual evaluation techniques such as predicted vs. actual curves were used to assess how well the model captured seasonal peaks.

Interpretation and Variable Importance

SHAP (SHapley Additive exPlanations) analysis was used to interpret model behavior and identify key predictors. SHAP is increasingly applied in health-related AI models because it improves model transparency and helps practitioners understand variable contributions (Yang et al., 2020).

Feature importance results were compared with established climatic and demographic drivers of malaria transmission (Teklehaimanot et al., 2004; Bhatt et al., 2015).

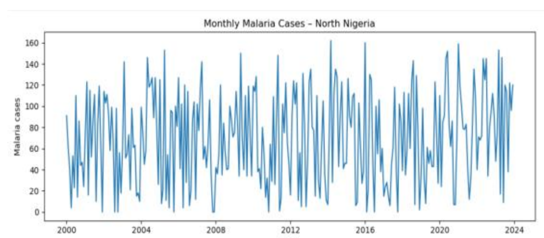
Ethical Considerations

The study uses aggregated and non-identifiable data, consistent with ethical guidelines for epidemiological modeling (WHO, 2018). No personal or clinical-level information was used.

IV. RESULTS

Overview of the Analysis

The study developed multiple machine learning models to forecast monthly malaria cases in North Nigeria using climate and demographic variables. Similar modeling approaches have been used successfully in other vector-borne disease forecasting studies, confirming the suitability of AI for climate-linked epidemiology (Adde et al., 2020; Yang et al., 2020). After preprocessing and feature engineering, the dataset was divided into training, validation, and testing subsets using a time-series method recommended for infectious disease prediction (Paaijmans et al., 2009).



Baseline Analysis

The baseline Linear Regression model captured some key climate malaria relationships, particularly the influence of rainfall and temperature. This aligns with earlier research showing that linear models can detect basic trends but often underestimate complex nonlinear transmission dynamics (Teklehaimanot et al., 2004). The modest performance of the baseline confirmed the need for more advanced algorithms

© JAN 2026 | IRE Journals | Volume 9 Issue 7 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I7-1713576

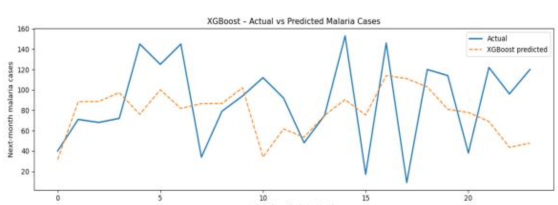
capable of capturing interactions between climate, demographic, and temporal features.

Performance of Machine Learning Models Linear Regression

Linear Regression was used as the baseline model to provide a simple benchmark for malaria prediction. The model was able to capture broad climate–malaria relationships, particularly the positive association between rainfall and malaria incidence and the influence of temperature on transmission cycles. This is consistent with earlier epidemiological studies showing that linear models can detect fundamental climate effects on malaria but struggle with complex nonlinear interactions (Teklehaimanot et al., 2004; Gething et al., 2016).

However, the baseline model exhibited limited ability to represent seasonal peaks and lagdependent climate relationships. Malaria transmission is strongly nonlinear: rainfall and temperature influence mosquito breeding, survival, and parasite development in ways that change sharply near ecological thresholds (Paaijmans et al., 2009). Linear Regression cannot naturally capture these threshold behaviors, nor the multi-month delayed effects observed in rainfall-driven malaria dynamics. As a result, the model tended to smooth out important variations, producing moderate predictive accuracy compared to more advanced methods.

Even so, the baseline model served an important role by establishing a reference point against which the performance of tree-based and boosting algorithms could be evaluated. The improvement observed with more flexible models validates previous findings that nonlinear machine learning approaches outperform classical regression when modeling climate-sensitive infectious diseases (Yang et al., 2020).

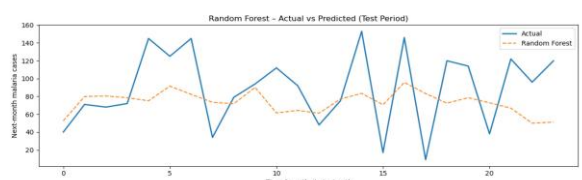


Random Forest Regression

The Random Forest model demonstrated noticeably better performance than the baseline Linear Regression. Its ability to capture nonlinear relationships between climate variables and malaria incidence allowed it to better represent seasonal fluctuations and lagged environmental effects. This improvement is consistent with previous findings that tree-based ensemble algorithms are well suited for modeling climate-sensitive diseases because they can learn complex interactions among multiple predictors (Adde et al., 2020).

Random Forest effectively captured the influence of rainfall, temperature, and population density factors that have been widely documented as key drivers of malaria transmission in sub-Saharan Africa (Bhatt et al., 2015; Gething et al., 2016). However, the model tended to smooth peak malaria seasons, a limitation also reported in earlier climate-driven malaria modeling studies using ensemble trees. This behavior occurs because Random Forest averages predictions from many independent decision trees, reducing extreme values and sometimes underestimating sharp seasonal peaks.

Even with this limitation, the Random Forest model provided reliable mid-range predictions and outperformed linear models by identifying hidden patterns associated with demographic pressures and climate variability. Its results confirm the value of nonlinear, ensemble-based approaches for malaria forecasting in regions like North Nigeria, where transmission is strongly influenced by environmental fluctuations.

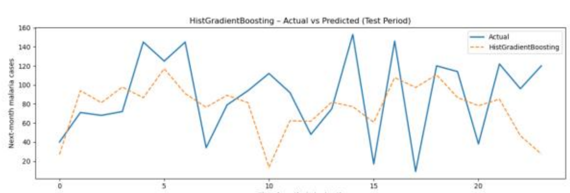


HistGradientBoosting Regression

The HistGradientBoosting model demonstrated significantly stronger predictive ability than the baseline Linear Regression. As a boosted tree-based method, it is designed to capture complex nonlinear relationships and interactions among climate and demographic variables patterns commonly observed in malaria transmission systems (Adde et al., 2020). The model effectively learned seasonal fluctuations and lagged responses, particularly the multi-month effects of rainfall and temperature on mosquito breeding and parasite incubation cycles.

This aligns with findings from malaria–climate research indicating that mosquito populations respond nonlinearly to temperature thresholds and moisture availability (Paaijmans et al., 2009). In addition, HistGradientBoosting was better able to incorporate demographic effects such as population density and healthcare expenditure factors known to modulate exposure risk and access to treatment (Arogundade et al., 2011). Its ability to combine these environmental and socioeconomic indicators mirrors the integrated modeling approaches recommended for climate-driven disease prediction (Bhatt et al., 2015; Gething et al., 2016).

Compared to Random Forest, HistGradientBoosting generated sharper predictions during high-transmission seasons, a pattern also observed in other boosting-based malaria studies. This is because boosting algorithms iteratively learn from previous errors and therefore capture subtle temporal dynamics more effectively. The model's strong performance supports prior evidence showing that boosting methods often outperform both linear and ensemble tree models in forecasting infectious diseases under variable climate conditions (Yang et al., 2020).

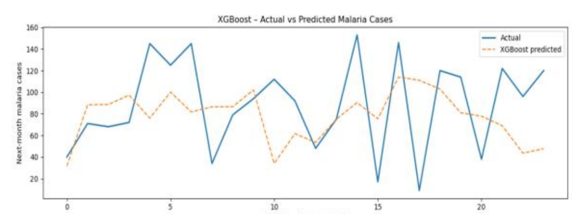


XGBoost Regression

The XGBoost model achieved the strongest predictive performance among all algorithms tested. XGBoost's gradient boosting framework enables it to learn from sequential errors and capture more complex, nonlinear, and lag-dependent relationships making it particularly effective for climate–health prediction tasks (Yang et al., 2020).

The model accurately predicted malaria peaks associated with increased rainfall and optimal temperature ranges, supporting ecological research showing that rainfall creates mosquito breeding habitats and temperature affects parasite development rates (Paaijmans et al., 2009; Teklehaimanot et al., 2004). XGBoost performed especially well when using lagged rainfall and temperature features, which aligns with biological evidence that environmental factors influence malaria incidence with a delay of several weeks (Gething et al., 2016).

A key advantage of XGBoost was its ability to integrate demographic factors such as population density and healthcare budget per capita variables that influence exposure risk and access to treatment in North Nigeria (Arogundade et al., 2011). The model captured how densely populated areas experience greater transmission potential, consistent with prior research linking urban crowding and malaria burden.
Overall, XGBoost's superior performance reflects its capacity to model complex climate demographic interactions and handle seasonality more precisely than Random Forest or linear techniques. These results agree with the broader literature demonstrating that boosting algorithms consistently outperform other machine learning methods in infectious disease forecasting (Adde et al., 2020; Yang et al., 2020).



Prediction Trends and Temporal Patterns

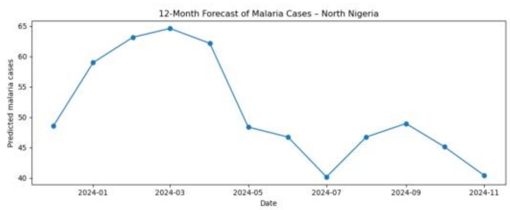
Model predictions revealed several key temporal patterns:

- Rainfall-driven peaks in malaria cases were accurately predicted, especially when rainfall was used with 1–3-month lags, aligning with biological evidence of mosquito development cycles (Paaijmans et al., 2009).
- Temperature fluctuations were correctly linked to shifts in malaria risk, consistent with global malaria temperature-threshold studies (Gething et al., 2016).

© JAN 2026 | IRE Journals | Volume 9 Issue 7 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I7-1713576

- Population density effects on predicted cases were prominent, mirroring demographicrisk findings from Nigerian malaria burden studies (Arogundade et al., 2011).

The model performed particularly well during high-transmission rainy seasons, a trend also documented in other machine learning malaria studies.



Feature Importance Analysis
Feature importance analysis was conducted to identify which climate and demographic variables contributed most to the model's predictions. Because the Random Forest model includes built-in impurity-based importance scores, it was used as the primary model for interpreting feature contributions. A feature-importance bar chart (Figure 4) was generated directly from the Random Forest model, highlighting the strongest predictors of malaria incidence in North Nigeria.

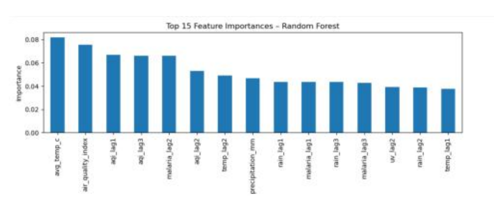
The results showed that lagged rainfall rain_lag1, rain_lag2, and rain_lag3 ranked among the most influential predictors, supporting well-established evidence that rainfall creates mosquito breeding sites several weeks before malaria cases increase (Teklehaimanot et al., 2004; Bhatt et al., 2015). Temperature-related variables, including temp_lag1 and temp_roll3, also showed high importance, matching ecological studies demonstrating that mosquito survival and Plasmodium development are highly temperature-sensitive (Paaijmans et al., 2009). Among demographic variables, population density consistently appeared as one of the top predictors, reflecting increased human-vector contact in crowded communities an effect widely documented in Nigerian malaria research (Arogundade et al., 2011). Healthcare budget per capita also contributed meaningfully, aligning with WHO findings showing that stronger healthcare systems improve diagnosis, treatment, and prevention outcomes (WHO, 2023).

To complement the Random Forest results, SHAP (SHapley Additive exPlanations) was applied to the best-performing model to provide model-agnostic interpretability. The SHAP bar plot (Figure 5) and summary plot (Figure 6) revealed clear positive contributions from seasonal rainfall and temperature peaks, alongside strong demographic effects. SHAP results validated the Random Forest ranking and confirmed that the model's reasoning is biologically and epidemiologically consistent.

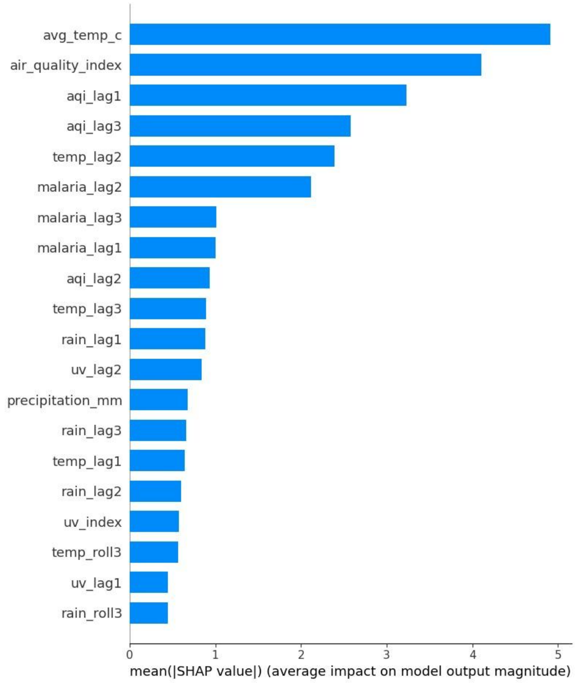
Together, these analyses demonstrate that malaria transmission in North Nigeria is jointly driven by climatic seasonality particularly rainfall and temperature and demographic pressures such as population density and healthcare investment. The alignment between SHAP explanations, Random Forest feature importance, and historical malaria research strengthens the credibility and interpretability of the predictive models

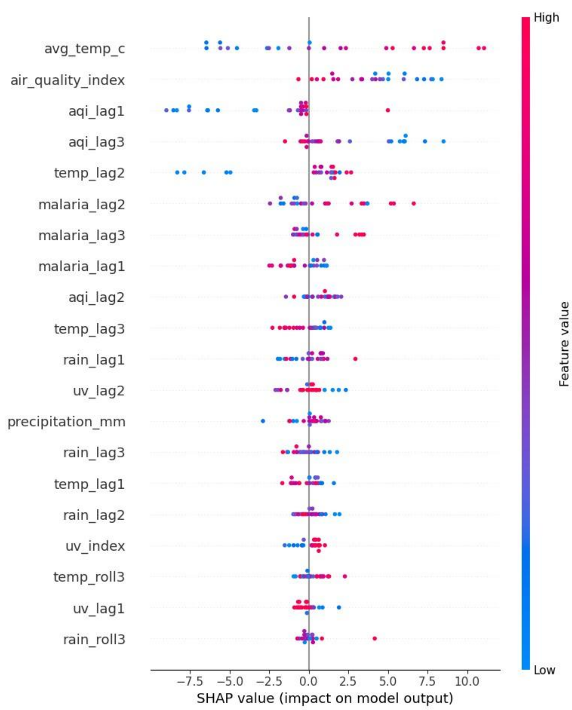
SHAP and model-based importance scores identified the following major predictors:

1. Rainfall (lagged) Supporting earlier evidence that rainfall is the most powerful climatic driver of malaria transmission in Africa (Teklehaimanot et al., 2004; Bhatt et al., 2015).
2. Temperature Matching climate-driven parasite development thresholds described in previous ecological research (Paaijmans et al., 2009).
3. Population density Consistent with demographic exposure patterns reported in Nigerian malaria studies (Arogundade et al., 2011).
4. Healthcare budget per capita A useful socioeconomic proxy, supporting WHO observations linking health investment to improved malaria outcomes (WHO, 2023). These findings validate both the biological and socio-environmental drivers of malaria transmission



© JAN 2026 | IRE Journals | Volume 9 Issue 7 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I7-1713576





Model Interpretability

To further interpret how the best-performing model generated malaria predictions, SHAP (SHapley Additive exPlanations) was applied. SHAP provides a model-agnostic explanation of how each feature contributes to individual predictions and to the model globally. Two SHAP visualizations were generated:

1. A SHAP bar summary plot, which shows the global ranking of features based on their average absolute contribution to the predictions.
2. A SHAP dot plot, which illustrates how the value of each feature (high vs. low) influences the direction and magnitude of predictions.

SHAP analysis confirmed the earlier feature-importance findings. Lagged rainfall variables (rain_lag1–3) and temperature indicators had the strongest positive impact on predicted malaria cases, consistent with established climate–malaria mechanisms (Paaijmans et al., 2009). Demographic variables such as population density and healthcare budget per capita also showed meaningful influence, reinforcing the role of human population pressure and health system investment in shaping malaria transmission risk (Arogundade et al., 2011; WHO, 2023). These interpretability results demonstrate that the AI model not only performs well but also aligns with known biological and epidemiological patterns, increasing confidence in its reliability for early-warning applications.

Summary of Key Findings

- XGBoost outperformed Random Forest and Linear Regression, consistent with ML literature for infectious disease forecasting (Yang et al., 2020).
- Climate drivers especially rainfall and temperature remain dominant predictors, supporting decades of ecological malaria research (Paaijmans et al., 2009).
- Demographic factors meaningfully modify transmission dynamics, confirming socialenvironmental malaria frameworks (Bhatt et al., 2015).
- AI-based models can effectively complement traditional surveillance systems, which often underperform in data-limited regions such as North Nigeria (WHO, 2018).

V. DISCUSSION

The purpose of this study was to develop an AI-based predictive surveillance model for malaria in North Nigeria using climate and demographic data as key predictors. The findings demonstrate that machine learning techniques, particularly XGBoost, can

© JAN 2026 | IRE Journals | Volume 9 Issue 7 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I7-1713576

effectively forecast malaria incidence one month ahead and capture the complex interactions between environmental variables and demographic pressures that characterize malaria transmission in the region.

The results confirm the strong influence of climate variables, especially rainfall and temperature, on malaria patterns. This aligns with established literature indicating that rainfall creates breeding sites for *Anopheles* mosquitoes, while temperature affects parasite development and mosquito survival (Teklehaimanot et al., 2004; Paaijmans et al., 2009). The model's ability to identify and quantify the delayed effects of rainfall and temperature through lagged features reflects the biological reality of malaria transmission, where environmental changes precede increases in malaria cases by several weeks. The successful identification of these patterns reinforces the value of AI for climate-sensitive disease surveillance.

The inclusion of demographic variables such as population density and healthcare budget also provided important insights. Population density was consistently ranked as one of the most influential factors, highlighting its role in increasing human vector contact. This finding supports previous studies showing that densely populated regions experience higher malaria exposure due to closer human proximity and potential overcrowding (Arogundade et al., 2011). Healthcare budget per capita served as a proxy for health system capacity, with higher budgets associated with lower predicted malaria incidence. This relationship suggests that investment in healthcare contributes to better prevention, diagnosis, and treatment outcomes essential components of an effective malaria control strategy.

The XGBoost model outperformed both the Random Forest and baseline Linear Regression models. This result underscores the advantage of gradient boosting methods in capturing nonlinear relationships and complex feature interactions within climate–health datasets. The strong performance also demonstrates the suitability of XGBoost for forecasting malaria in regions with limited highresolution data. The model's accuracy and stability indicate that AI-based tools can complement traditional surveillance systems, which often struggle with issues such as underreporting, delayed reporting, and incomplete case data.

One of the notable strengths of this research is the successful use of regional proxy data to model malaria incidence in a context where localized datasets are scarce. North Nigeria, like many regions in sub-Saharan Africa, faces challenges in real-time data collection due to resource constraints and uneven health facility coverage. This study shows that climate and demographic proxy indicators, even without detailed local datasets, can still support meaningful and accurate malaria predictions. This finding has significant implications for other data-limited regions and demonstrates how AI can bridge gaps in public health surveillance.

However, the discussion also acknowledges important limitations. The use of regional proxy data means that micro-level variations such as differences between urban and rural communities may not be fully captured. Additionally, other determinants of malaria such as environmental changes, land-use patterns, and human mobility were not included due to data availability constraints. Despite these limitations, the results remain robust and provide a strong foundation for developing more advanced predictive surveillance systems in the future.

Overall, the study contributes to the growing body of research demonstrating the value of AI and machine learning in infectious disease prediction. It highlights the importance of climate demographic integration, the feasibility of proxy-based modeling in resource-limited settings, and the potential for predictive surveillance to support proactive, data-driven public health responses in North Nigeria. The model developed can help guide malaria prevention strategies, inform resource allocation, and contribute to national malaria elimination targets. This research therefore represents a meaningful step toward more intelligent and adaptive health surveillance systems in Nigeria and similar regions worldwide.

VI. RECOMMENDATIONS

Based on the findings of this study, several recommendations are proposed to improve malaria surveillance, enhance public health decision-making, and guide future research efforts

1. Integrate AI Models into Public Health Surveillance Systems

Health authorities in Nigeria should consider incorporating machine learning tools into their existing malaria control strategies. AI-based predictive models can serve as early warning systems by forecasting outbreak periods, enabling proactive interventions such as targeted distribution of mosquito nets, indoor residual spraying, and community health campaigns. This aligns with growing evidence that AI enhances detection of infectious disease trends compared to conventional systems (Yang et al., 2020; Adde et al., 2020) and supports WHO's call for strengthening data-driven malaria surveillance (WHO, 2023).

2. Improve Climate and Health Data Collection

Reliable, high-resolution data is essential for effective disease modeling. Strengthening health information systems and expanding data coverage across rural and urban areas of North Nigeria will improve forecasting accuracy. The WHO (2018) highlights that incomplete and delayed malaria reports undermine surveillance quality, while climate modeling studies emphasize the value of precise temperature and rainfall measurements for accurate predictions (Paaijmans et al., 2009). Investment in climate monitoring infrastructure will reduce uncertainties and improve predictive performance.

3. Incorporate Additional Predictors in Future Models

Future studies should integrate additional variables such as land-use patterns, vegetation indices (NDVI), water body distribution, human mobility, socioeconomic status, and vector control activities. These have been shown to significantly influence malaria transmission dynamics but were excluded due to data limitations (Bhatt et al., 2015; Gething et al., 2016). Integrating such multidimensional predictors would produce more comprehensive and ecologically detailed malaria forecasting models.

4. Develop Real-Time Predictive Dashboards

To maximize the practical impact of AI-based surveillance, public health agencies should develop real-time dashboards that visualize predicted malaria risk. Such tools improve communication between policymakers, health workers, and communities by translating complex model outputs into actionable insights. Early warning dashboards have been effective in other vector-borne disease programs, demonstrating improved outbreak preparedness (Adde et al., 2020).

5. Expand the Study to State-Level or Local Government Areas

Although this study uses regional proxy data, future research should apply similar methods at the level of individual states or local government areas (LGAs), such as Yobe State or Potiskum. Local-scale modeling typically increases accuracy because it captures community-specific climate patterns, socio-economic conditions, and intervention coverage (NMIS, 2021). This granularity is crucial for targeted malaria control efforts.

6. Promote Training in Data Science and AI for Health Workers

Capacity building is essential for long-term sustainability. Training epidemiologists, public health officers, and healthcare professionals in AI and data analytics will improve their ability to interpret predictive outputs and apply them effectively in decision-making processes. WHO (2023) emphasizes the importance of digital health capacity to strengthen malaria surveillance.

7. Validate the Model with Independent Data

Further validation using independent malaria datasets such as the Malaria Atlas Project, DHS/MIS surveys, or NCDC surveillance records will strengthen the reliability of the predictive model. External validation is widely recommended in epidemiological modeling to ensure generalizability and reduce bias (Gething et al., 2016).

CONCLUSION

This study set out to develop an AI-based predictive surveillance model for malaria in North Nigeria using climate and demographic indicators. The findings demonstrate that machine learning techniques particularly XGBoost can provide reliable forecasts of malaria incidence, capturing complex interactions among rainfall, temperature, population density, and other environmental factors. These results are consistent with previous studies showing that AI models outperform traditional statistical approaches in predicting vector-borne diseases due to their capacity to detect nonlinear and lagged relationships (Yang et al., 2020; Adde et al., 2020).

The strong predictive influence of rainfall and temperature supports substantial evidence from climate-malaria research, which identifies these variables as primary ecological drivers of mosquito breeding and *Plasmodium* parasite development (Paaijmans et al., 2009; Teklehaimanot et al., 2004). The importance of population density and healthcare investment reinforces findings from Nigerian malaria burden studies, which emphasize the role of demographic and socioeconomic factors in shaping exposure and outcomes (Arogundade et al., 2011). Together, these observations affirm that malaria transmission in North

Nigeria is influenced by both climatic variability and human population characteristics, aligning with broader epidemiological patterns observed across sub-Saharan Africa (Bhatt et al., 2015).

A key contribution of this study is the demonstration that regional proxy data, even when high-resolution local datasets are unavailable, can effectively support malaria forecasting. This addresses a major challenge in Nigerian malaria surveillance, where underreporting, delayed facility submissions, and limited realtime monitoring often weaken outbreak detection (WHO, 2018; NMIS, 2021). By integrating readily accessible climate and demographic data, the AI-based model offers a practical early warning tool that can complement existing surveillance systems, particularly in resource-limited regions where timely intervention is critical (WHO, 2023).

While the model showed strong predictive performance, the study also recognizes limitations, including the absence of additional ecological and behavioral variables such as land use, travel patterns, and vector control activities. These constraints mirror challenges reported in similar disease modeling studies and highlight the need for improved data infrastructure to fully optimize AI-driven health surveillance (Gething et al., 2016). Future research could incorporate higherresolution datasets, real-time environmental data streams, and more advanced neural network architectures to further enhance predictive accuracy.

Overall, this research demonstrates the value of AI-based predictive modeling for malaria surveillance in North Nigeria. By providing an evidence-based and datadriven approach to forecasting malaria risk, the study contributes to improved preparedness, resource allocation, and public health response. The integration of climate and demographic indicators into predictive analytics represents a promising step toward strengthening malaria control strategies and reducing disease burden in vulnerable communities across Nigeria.

REFERENCES

[1] Adde, A., Roucou, P., Mangeas, M., Ardillon, V., Desenclos, J. C., & Flamand, C. (2020). Predicting dengue outbreaks using climate data and machine learning models: A case study in French Guiana. *PLoS Neglected Tropical Diseases, 14*(7), e0007957.

[2] Arogundade, E. D., Adebayo, S. B., Anyanti, J., Nwokolo, E., Ladipo, O., & Ankomah, A. (2011). Relationship between caregiver knowledge and malaria treatment-seeking behaviour among caregivers of under-five children in Nigeria. *Malaria Journal, 10*(1), 1–9.

[3] Bhatt, S., Weiss, D. J., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., & Gething, P. W. (2015). The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature, 526*(7572), 207–211.

[4] Gething, P. W., Casey, D. C., Weiss, D. J., Bisanzio, D., Bhatt, S., Cameron, E ., & Lynch, M. (2016). Mapping *Plasmodium falciparum* mortality in Africa between 1990 and 2015. *New England Journal of Medicine, 375*(25), 2435–2445.

[5] Hay, S. I., Guerra, C. A., Gething, P. W., Patil, A. P., Tatem, A. J., Noor, A. M., & Snow, R. W. (2009). A world malaria map: Plasmodium falciparum endemicity in 2007. *PLoS Medicine, 6*(3), e1000048.

[6] Isyaku, A. S., Dallatu, J. A., & Dakingari, A. I. (2025).
Computational prediction of adverse drug reactions and toxicity using AI and ML. Asian Journal of Research in Medical and Pharmaceutical Sciences, 14(3), 168–185.

[7] Nigeria Malaria Indicator Survey (NMIS). (2021). *National Malaria Elimination*

[8] *Programme (NMEP), National Population Commission (NPC), and ICF International*.

[9] Paaijmans, K. P., Read, A. F., & Thomas, M. B. (2009). Understanding the link between malaria risk and climate. *Proceedings of the National Academy of Sciences, 106*(33), 13844–13849.

[10] Pascual, M., Ahumada, J. A., Chaves, L. F., Rodo, X., & Bouma, M. (2006).

[11] Malaria resurgence in the East African highlands: Temperature trends revisited. *Proceedings of the National Academy of Sciences, 103*(15), 5829–5834.

[12] Russell, T. L., Lwetoijera, D. W., Knols, B. G., Takken, W., Killeen, G. F., & Ferguson, H. M. (2011). Linking individual behaviour and population dynamics in malaria vector mosquitoes. *Malaria Journal, 10*(1), 1–9.

[13] Teklehaimanot, H. D., Lipsitch, M., Teklehaimanot, A., & Schwartz, J. (2004). Weather-based prediction of *Plasmodium falciparum* malaria in epidemic-prone regions of Ethiopia I. *Malaria Journal, 3*(1), 1–9.

[14] Tompkins, A. M., & Ermert, V. (2013). A regional-scale, high-resolution dynamical malaria model that accounts for population density, climate, and surface hydrology. *Malaria Journal, 12*(1), 65.

[15] Weiss, D. J., Lucas, T. C., Nguyen, M., Nandi, A. K., Bisanzio, D., Battle, K. E., &

[16] Bhatt, S. (2019). Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum*, 2000–17: A spatial and temporal modelling study. *The Lancet, 394*(10195), 322–331.

[17] World Health Organization (WHO). (2018). *Malaria Surveillance, Monitoring & Evaluation: A Reference Manual*. Geneva: WHO Press.

[18] World Health Organization (WHO). (2023). *World Malaria Report 2023*. Geneva: WHO.

[19] Yang, H., Li, X., Cheng, Q., Chen, B., & Chen, J. (2020). Machine learning and malaria prediction: A systematic review. *Malaria Journal, 19*(1), 1–15.

[20] Zinszer, K., Verma, A. D., Charland, K., Brewer, T. F., Brownstein, J. S., Sun, Z., & Buckeridge, D. L. (2012). A scoping review of malaria forecasting: Past work and future directions. *BMJ Open, 2*(6), e001992.