Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Machine Learning and Explainable AI for Multi-State Classification of Malaria Transmission Dynamics in Kenya

Maurice Wanyonyi

mauricewanyonyi27@gmail.com

African Institute for Capacity Development (AICAD)

Research Article

Keywords: Malaria transmission, machine learning, explainable artificial intelligence, multi-state classification

Posted Date: June 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9735076/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Machine Learning and Explainable AI for Multi-State Classification of Malaria Transmission Dynamics in Kenya

Maurice Wanyonyi[1,2*]

[1*]Department of Mathematics and Statistics, University of Embu, Embu, 60100, 6, Kenya.
[2]Research and Development, African Institute for Capacity Development (AICAD), 46179, Nairobi, 00100, Kenya.

Corresponding author(s). E-mail(s): mauricewanyonyi27@gmail.com;

Abstract

Malaria remains a major public health challenge in sub–Saharan Africa, with pronounced spatial and temporal variation in transmission intensity that complicates effective control strategies. Accurate classification of transmission states is essential for guiding targeted interventions and strengthening early warning systems. This study develops a machine learning framework for the classification of malaria transmission states in Kenya using monthly panel data from 47 counties spanning the period 2015 to 2025. Transmission was categorised into four operationally relevant states based on incidence thresholds. Four supervised learning models, namely multinomial logistic regression, random forest, extreme gradient boosting, and support vector machine, were trained using temporally lagged features and evaluated under a forward chaining validation scheme to preserve temporal structure. Model performance was assessed using accuracy, macro averaged F1 score, Matthews correlation coefficient, and Brier score, complemented by calibration analysis. Extreme gradient boosting achieved the best overall performance, with accuracy of 0.9918, macro averaged F1 score of 0.9647, and Matthews correlation coefficient of 0.9831, alongside the lowest Brier score of 0.0031, indicating highly reliable probability estimates. Feature importance analysis revealed that lagged incidence, vegetation index, precipitation, and insecticide treated net coverage were the most influential predictors. Partial dependence analysis demonstrated nonlinear relationships and clear seasonal patterns in transmission dynamics. The findings show that machine learning approaches can accurately classify malaria transmission states while providing interpretable and well calibrated outputs for decision making. This framework

offers a practical tool for supporting malaria surveillance and resource allocation. Further validation in different epidemiological settings is recommended to assess generalisability.

Keywords: Malaria transmission, machine learning, explainable artificial intelligence, multi-state classification

1 Introduction

Malaria remains a major public health burden in sub Saharan Africa, with an estimated 249 million cases and 608000 deaths worldwide in 2022, of which 94% of cases and 95% of deaths occurred in the African Region [1]. In Kenya, malaria transmission exhibits substantial spatial and temporal heterogeneity, ranging from highly endemic regions around Lake Victoria and coastal zones to epidemic prone highland areas and arid regions with low transmission [2, 3]. Approximately 70% of the population is at risk, and malaria accounts for over 20% of outpatient visits in high burden settings [4]. Despite sustained scale up of insecticide treated nets, indoor residual spraying, and artemisinin based combination therapies, progress in malaria control has slowed, with recent evidence of resurgence linked to insecticide resistance, climate variability, and declining intervention effectiveness [5, 6]. These challenges highlight the need for improved analytical approaches that can support timely and targeted intervention strategies.

Accurate characterisation and prediction of malaria transmission dynamics are essential for guiding resource allocation, informing policy decisions, and strengthening early warning systems. Traditional modelling approaches, including compartmental models and statistical time series methods, have provided important insights into malaria dynamics [7, 8]. However, these approaches often rely on restrictive assumptions such as linear relationships, stationarity, and predefined functional forms, which limit their ability to represent the complex and nonlinear interactions among climatic, environmental, and intervention related factors. In addition, their capacity to incorporate high dimensional data and capture context specific variability across space and time remains limited.

Machine learning methods offer a flexible and data driven alternative by learning patterns directly from observed data without requiring explicit specification of underlying relationships [9, 10]. These approaches are well suited to modelling complex epidemiological systems, as they can capture nonlinear dependencies, represent interactions among multiple predictors, and improve predictive accuracy in settings characterised by heterogeneity and uncertainty. Consequently, machine learning has gained increasing attention in malaria research.

Several studies have demonstrated the potential of machine learning in this domain. Ayoka and Nnadi [11] applied machine learning techniques to predict malaria prevalence in Nigeria, while Baba Adamu et al. [12] examined the influence of climate variability on malaria transmission. Morang?a et al. [13] used machine learning to classify clinical malaria outcomes based on haematological indicators. Taconet et al. [14]

developed interpretable models to analyse environmental drivers of mosquito biting rates in Burkina Faso, and Zheng et al. [15] investigated climate driven transmission patterns in Tanzania. Additional work has explored integration with spatial clustering and mathematical modelling frameworks [16, 17]. While these studies provide valuable insights, most focus on continuous outcomes such as incidence or prevalence and often do not directly align with operational decision making frameworks used in malaria control.

Despite recent progress, several important gaps remain. First, limited attention has been given to modelling malaria transmission as discrete operational states that correspond to actionable risk categories used in public health practice. Second, many existing studies emphasise predictive accuracy without adequately assessing probabilistic calibration, even though reliable probability estimates are critical for risk based decision making. Third, the application of explainable artificial intelligence methods remains limited, reducing transparency and hindering practical adoption of machine learning models in public health settings. Fourth, computational considerations such as training time, scalability, and inference efficiency are rarely evaluated, despite their importance for implementation in resource constrained environments.

To address these gaps, this study develops a comprehensive and interpretable machine learning framework for the classification of malaria transmission states in Kenya using spatio temporal panel data. The study models transmission as discrete epidemiologically meaningful categories, evaluates multiple machine learning algorithms using both predictive and calibration metrics, and integrates explainable artificial intelligence methods to enhance interpretability. In addition, computational performance is assessed to ensure that the proposed framework is suitable for practical deployment in real world settings.

The contributions of this study are fourfold. First, it introduces a multi state classification framework that directly predicts operationally relevant transmission categories. Second, it provides a rigorous comparative evaluation of machine learning models using both accuracy based and probabilistic calibration measures. Third, it integrates explainable artificial intelligence techniques to identify key drivers of malaria transmission and improve model transparency. Fourth, it evaluates computational efficiency to support implementation in routine surveillance systems.

Overall, this study presents a robust, interpretable, and scalable framework for modelling malaria transmission dynamics in Kenya, with direct relevance for evidence based malaria control and early warning systems.

2 Materials and methods

2.1 Study Design and Data Sources

This study adopts a quantitative longitudinal design using a panel dataset comprising monthly observations for all 47 counties in Kenya from January 2015 to December 2025. The unit of analysis is the county month, resulting in a balanced panel with complete temporal coverage across all spatial units.

Malaria incidence data were obtained from the Kenya Ministry of Health District Health Information System version two and supplemented with Kenya Malaria

Indicator Surveys conducted in 2015, 2020, and 2025. Environmental variables, including temperature, precipitation, and Normalised Difference Vegetation Index, were obtained from the Climate Hazards Group InfraRed Precipitation with Station data. Intervention related variables, including insecticide treated net coverage, were derived from Malaria Indicator Surveys and routine distribution records. Static geographical variables, including elevation and population density, were obtained from the Shuttle Radar Topography Mission and the Kenya National Bureau of Statistics.

All datasets were harmonised to a common monthly temporal resolution and aligned spatially at the county level using consistent administrative identifiers.

The outcome variable is a categorical malaria transmission state denoted by $S_{i,t} \in \{0, 1, 2, 3\}$, where $i = 1, \ldots, 47$ indexes counties and t indexes time in months. The transmission state is derived from malaria incidence per 1000 population, denoted by $I_{i,t}$, as follows:

$$S_{i,t} = \begin{cases} 0 & \text{if } I_{i,t} < 5 \\ 1 & \text{if } 5 \leq I_{i,t} < 20 \\ 2 & \text{if } 20 \leq I_{i,t} < 100 \\ 3 & \text{if } I_{i,t} \geq 100 \end{cases} \tag{1}$$

These thresholds are consistent with epidemiological stratification approaches used in malaria surveillance and reflect operational categories relevant for intervention planning [2, 3].

2.2 Data Preprocessing

All preprocessing steps were implemented within a training data pipeline to avoid information leakage. Continuous variables were standardised using statistics computed from the training data only and subsequently applied to validation and test sets.

Missing values in environmental variables were imputed using linear interpolation for short gaps and seasonal mean imputation for longer gaps. Intervention and demographic variables were imputed using last observation carried forward when appropriate to preserve temporal consistency.

Outliers were assessed using interquartile range criteria. Observations were retained unless clear data entry errors were identified, in which case corrections were applied based on source verification. All preprocessing procedures were implemented using Python libraries including pandas and numpy.

2.3 Feature Engineering and Temporal Structure

Let $X_{i,t} \in \mathrm{R}^p$ denote the vector of observed covariates for county i at time t. To capture temporal dependence and delayed effects, lagged features were constructed.

For lag order $l \in \{1, 2\}$, lagged covariates are defined as:

$$X_{i,t}^{(l)} = X_{i,t-l} \tag{2}$$

Lagged outcome states were included to represent persistence in transmission dynamics:

$$S_{i,t}^{(l)} = S_{i,t-l} \tag{3}$$

The final feature vector is defined as:

$$\tilde{X}_{i,t} = [X_{i,t-1}, X_{i,t-2}, S_{i,t-1}, S_{i,t-2}] \tag{4}$$

All lagged variables were constructed using past information only to ensure temporal validity.

2.4 Problem Formulation

The modelling objective is to estimate the conditional probability distribution of the current transmission state given lagged predictors. This is formulated as a multi class classification problem:

$$\hat{S}_{i,t} = \arg \max_{k \in \{0,1,2,3\}} P(S_{i,t} = k \mid \tilde{X}_{i,t}; \theta) \tag{5}$$

where θ represents model parameters.

2.5 Machine Learning Models

Four supervised learning algorithms were implemented to capture different modelling assumptions. All models were developed in Python using scikit learn version 1.2.2 and xgboost version 1.7.3. A fixed random seed of 42 was used to ensure reproducibility.

2.5.1 Multinomial Logistic Regression

Class probabilities were modelled using:

$$P(S = k \mid X) = \frac{\exp(\beta_k^\top X)}{\sum_{j=0}^{3} \exp(\beta_j^\top X)} \tag{6}$$

Parameters were estimated using maximum likelihood with ℓ_2 regularisation.

2.5.2 Random Forest

Random Forest constructs an ensemble of decision trees trained on bootstrap samples:

$$\hat{S}(X) = \text{mode}\,\{T_1(X), \ldots, T_B(X)\} \tag{7}$$

Splits were selected by maximising reduction in Gini impurity.

Extreme Gradient Boosting

The model updates predictions sequentially:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(X_i) \tag{8}$$

with objective function:

$$\mathcal{L} = \sum_i l(y_i, \hat{y}_i) + \sum_k \left(\gamma T_k + \frac{1}{2} \lambda \|w_k\|^2 \right) \tag{9}$$

2.5.3 Support Vector Machine

The optimisation problem is defined as:

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_i \xi_i \tag{10}$$

A radial basis function kernel was used to capture nonlinear relationships.

2.6 Hyperparameter Tuning and Temporal Validation

To preserve temporal structure and prevent information leakage, a forward chaining validation strategy was implemented. The dataset was divided into a training period from 2015 to 2020 and a test period from 2021 to 2025.

Within the training data, time ordered cross validation with five folds was used. Hyperparameters were selected using grid search to maximise macro averaged F1 score.

All model selection steps were conducted strictly within the training data to ensure unbiased evaluation.

2.7 Performance Metrics

Model performance was evaluated using multiple complementary metrics.

$$\text{Accuracy} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}(\hat{S}_i = S_i) \tag{11}$$

Macro averaged precision, recall, and F1 score were computed to account for class balance. The Matthews Correlation Coefficient was used as a robust summary metric:

$$MCC = \frac{\sum_{k,l,m} C_{kk}C_{lm} - C_{kl}C_{mk}}{\sqrt{(\sum_k T_k)\,(\sum_k P_k)}} \tag{12}$$

2.8 Calibration Assessment

Calibration was assessed to evaluate the agreement between predicted probabilities and observed outcomes. Reliability diagrams were constructed using ten probability bins.

The Brier score was computed as:

$$BS = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K}(y_{ik} - \hat{p}_{ik})^2 \tag{13}$$

Lower values indicate better calibration.

2.9 Explainable Artificial Intelligence

Interpretability was assessed using complementary approaches.

SHapley Additive exPlanations quantify feature contributions:

$$\phi_j = \sum_{S \subseteq \{1,\ldots,p\} \setminus \{j\}} \frac{|S|!(p - |S| - 1)!}{p!} \left[f(S \cup \{j\}) - f(S) \right] \tag{14}$$

Partial dependence functions estimate marginal effects:

$$\hat{f}_k(x_j) = \frac{1}{N} \sum_{i=1}^{N} P(S = k \mid X_{i,\setminus j}, x_j) \tag{15}$$

Local Interpretable Model agnostic Explanations provide local approximations:

$$\xi(X) = \arg\min_{g \in G} \mathcal{L}(f, g, \pi_X) + \Omega(g) \tag{16}$$

2.10 Reproducibility

All analyses were conducted using Python version 3.13. The computational environment included the following key libraries: `numpy`, `pandas`, `scikit-learn`, `matplotlib`, and `seaborn`. Additional packages were used for model interpretation and evaluation, including `shap`, `lime`, and `xgboost`.

Code, processed data, and documentation will be made publicly available in an open-access repository to ensure full reproducibility. Random seeds were fixed across all experiments, and all modelling workflows were implemented using deterministic pipelines wherever possible to ensure consistency of results.

2.11 Ethical Considerations

This study used secondary aggregated data obtained from national surveillance systems. No individual level or identifiable data were included. Ethical approval and informed consent were therefore not required.

3 Results

3.1 Dataset Characteristics

The dataset comprised 47 counties observed monthly from January 2015 to December 2025, yielding a total of 6,204 county month observations. The outcome variable, malaria transmission state, was derived from incidence per 1000 population and categorised into four classes: low (0), moderate (1), high (2), and very high (3). The class distribution was approximately balanced, with moderate transmission occurring most frequently, followed by low, high, and very high states. This distribution reflects the spatial and temporal variability of malaria transmission across Kenya.

Summary statistics for all variables are presented in Table 1. Environmental variables exhibited expected variability, with a mean temperature of 26.97°C and mean precipitation of 101.20 mm. The Normalised Difference Vegetation Index averaged 0.53, indicating moderate vegetation density. Insecticide treated net coverage showed a mean of 0.63, reflecting progressive scale up over the study period. Malaria incidence

per 1000 population had a mean of 11.03 cases, consistent with moderate transmission levels. Lagged variables displayed similar distributions to their contemporaneous counterparts, indicating temporal persistence in the underlying processes.

Table 1 Summary statistics of the features used in the analysis.

Feature	Mean	Std
Temperature (°C)	26.97	2.19
Precipitation (mm)	101.20	57.98
NDVI	0.53	0.10
ITN coverage	0.63	0.18
Elevation (m)	1343.27	875.84
Population density (persons/km^2)	2523.26	1499.12
Incidence per 1000	11.03	5.62

3.2 Model Performance Comparison

Four machine learning models were evaluated: multinomial logistic regression, random forest, extreme gradient boosting, and support vector machine. Performance was assessed using accuracy, macro averaged precision, recall, F1 score, Matthews correlation coefficient, specificity, area under the receiver operating characteristic curve, and Brier score.

Table 2 summarises the results. Extreme gradient boosting achieved the highest overall performance, with an accuracy of 0.9918, macro F1 score of 0.9647, and Matthews correlation coefficient of 0.9831. It also produced the lowest Brier score of 0.0031, indicating highly reliable probability estimates.

Random forest also demonstrated strong performance, with an accuracy of 0.9869 and a Matthews correlation coefficient of 0.9730, although its macro F1 score was lower at 0.7383. Multinomial logistic regression achieved moderate performance across all metrics, with an accuracy of 0.9476 and macro F1 score of 0.7331. The support vector machine exhibited the lowest performance, with an accuracy of 0.6792 and macro F1 score of 0.2022.

Table 2 Performance comparison of the four classifiers. The best value in each column is highlighted in bold.

Model	Acc.	Prec.	Rec.	F1	MCC	Spec.	AUC	Brier	Train (s)	Inf. (s)
LR	0.9476	0.7511	0.7188	0.7331	0.8905	0.9696	0.9704	0.02307	2.2775	0.00360
RF	0.9869	0.7391	0.7380	0.7383	0.9730	0.9932	0.9999	0.01065	1.6900	0.03357
XGBM	**0.9918**	**0.9667**	**0.9628**	**0.9647**	**0.9831**	**0.9956**	**0.9998**	**0.00313**	**0.6363**	**0.02414**
SVM	0.6792	0.1698	0.2500	0.2022	0.0000	0.7500	0.9026	0.08683	9.9940	0.74623

Note: LR = Multinomial Logistic Regression, RF = Random Forest, XGBM = Extreme Gradient Boosting Machine, SVM = Support Vector Machine, Acc. = Accuracy, Prec. = Macro Precision, Rec. = Macro Recall, Spec. = Macro Specificity, Train = Training time in seconds, Inf. = Inference time in seconds.

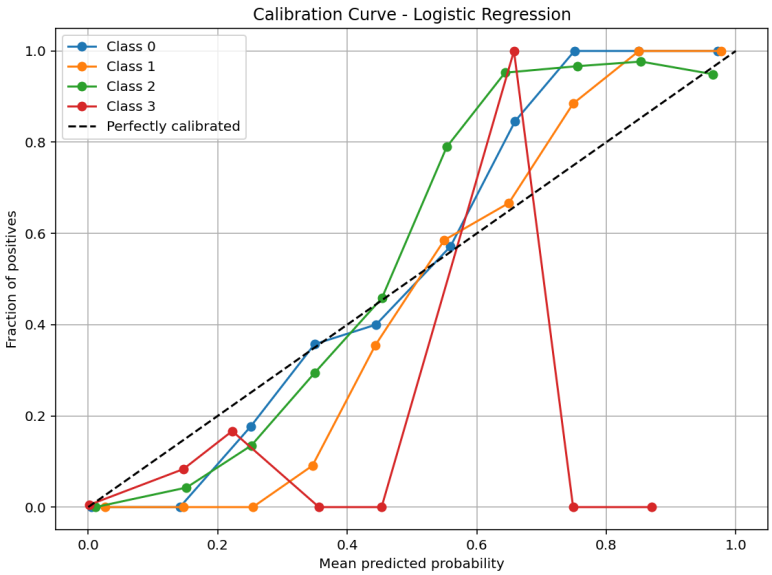


Fig. 1 Calibration curve (reliability diagram) for Logistic Regression. The diagonal represents perfect calibration.

3.3 Calibration Analysis

Calibration curves were used to evaluate the agreement between predicted probabilities and observed outcomes (Figures 1 to 4). Extreme gradient boosting demonstrated strong calibration across all classes, with curves closely aligned to the diagonal. Random forest also exhibited good calibration, although minor deviations were observed at higher probability levels.

Multinomial logistic regression showed systematic deviations from the diagonal, particularly for moderate and high transmission classes, indicating less reliable probability estimates. The support vector machine produced poorly calibrated probabilities, with predictions concentrated near extreme values, consistent with its higher Brier score.

3.4 Learning Curves and Model Generalisation

Learning curves were analysed to assess model generalisation and potential overfitting (Figures 5 to 8). Extreme gradient boosting and random forest exhibited high training accuracy with validation scores stabilising at similarly high levels, indicating strong generalisation with minimal overfitting.

Multinomial logistic regression showed convergence of training and validation performance at a lower level, suggesting limited model capacity. In contrast, the support vector machine displayed substantial variability and a noticeable gap between training and validation performance, indicating instability in learning.

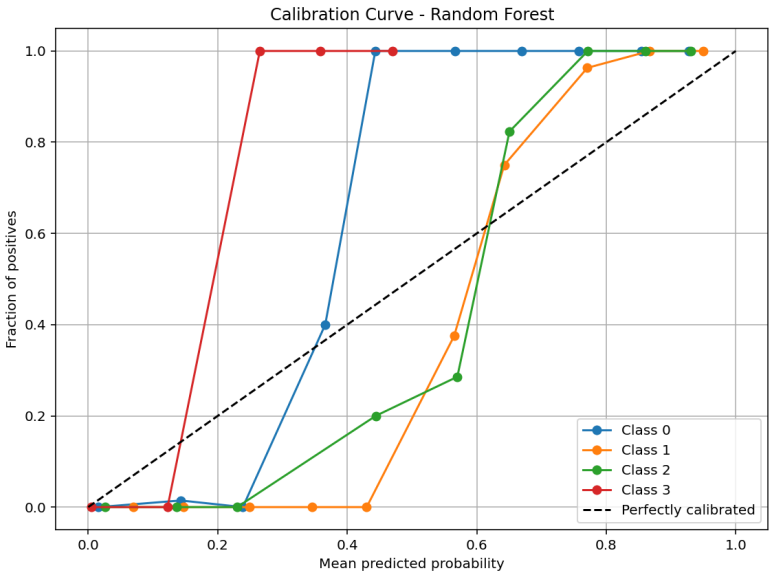


Fig. 2 Calibration curve for Random Forest.

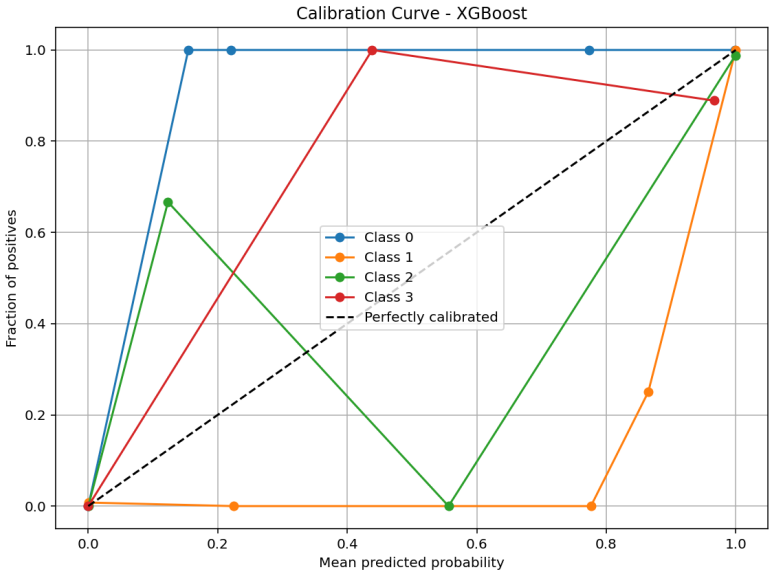


Fig. 3 Calibration curve for XGBoost.

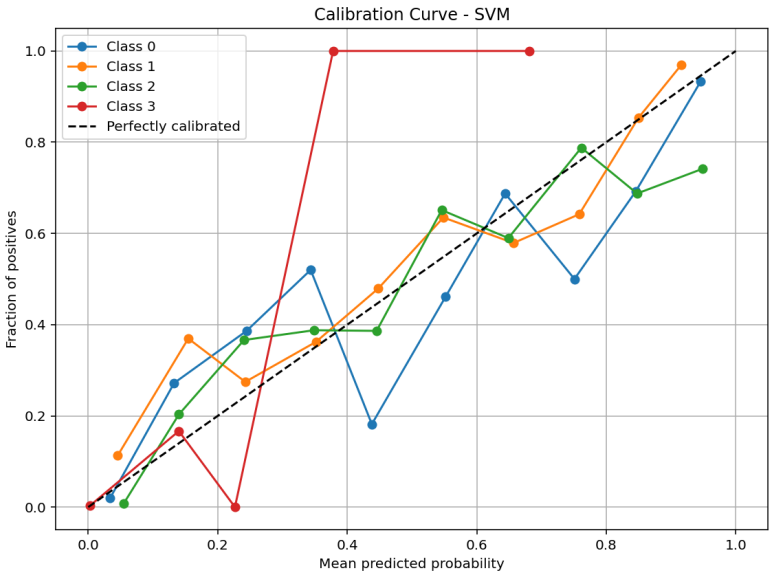


Fig. 4 Calibration curve for SVM.

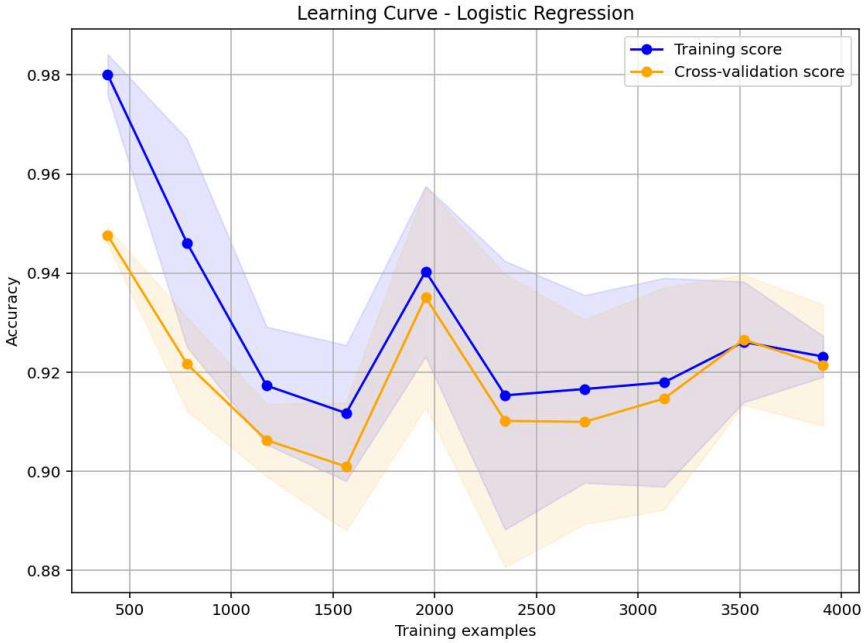


Fig. 5 Learning curve for Logistic Regression showing training and cross-validation accuracy as a function of training set size.

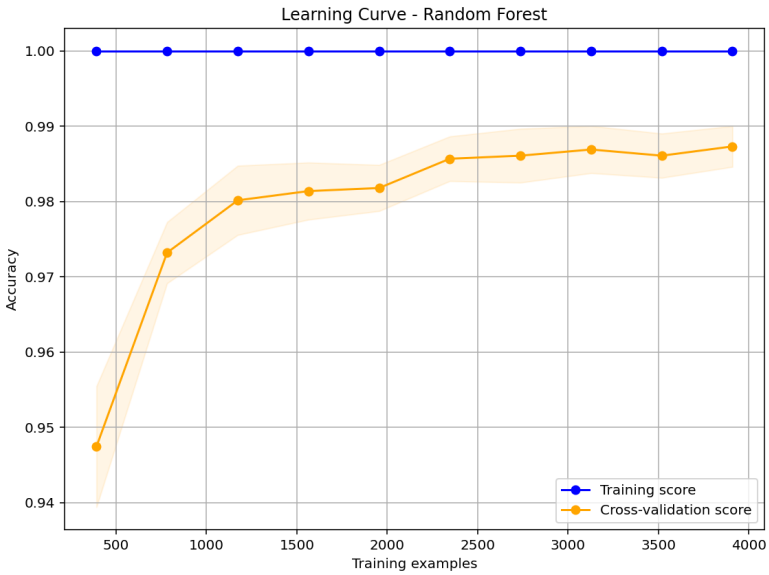


Fig. 6 Learning curve for Random Forest.

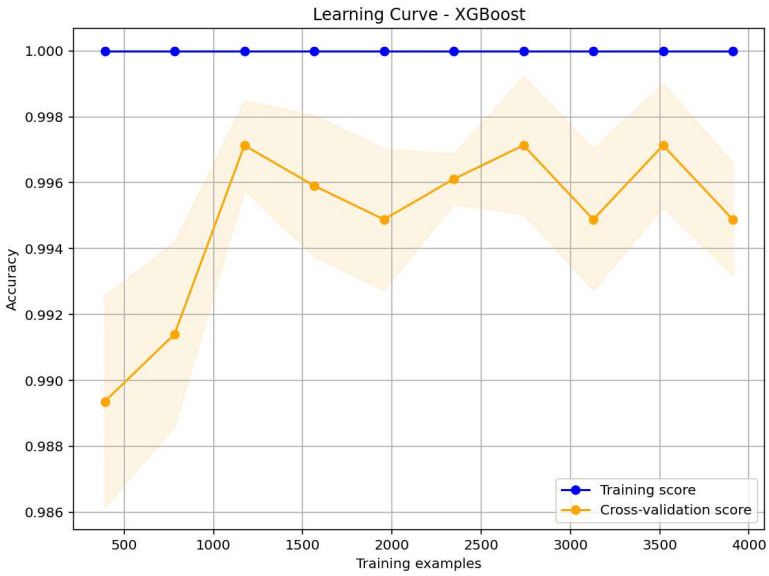


Fig. 7 Learning curve for XGBoost.

3.5 Confusion Matrix Analysis

The confusion matrix for the best performing model, extreme gradient boosting, is shown in Figure 9. The model achieved high classification accuracy across all transmission states. Misclassifications primarily occurred between adjacent classes, such as

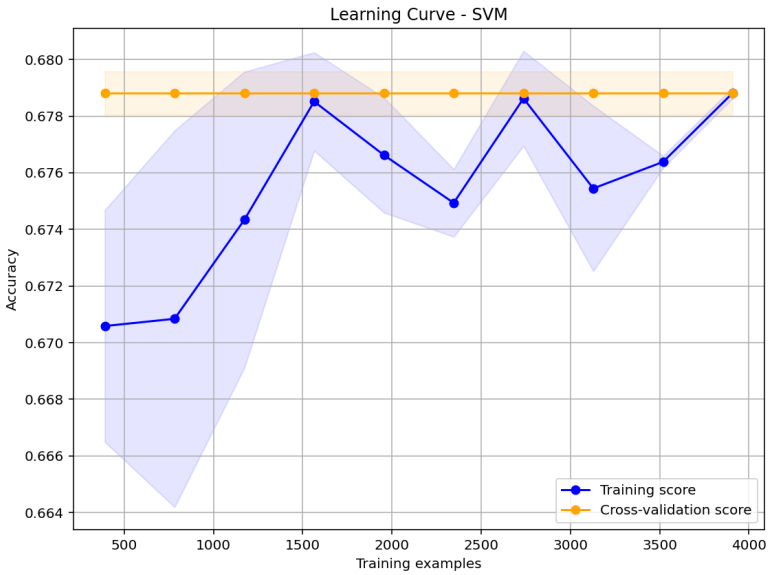


Fig. 8 Learning curve for SVM.

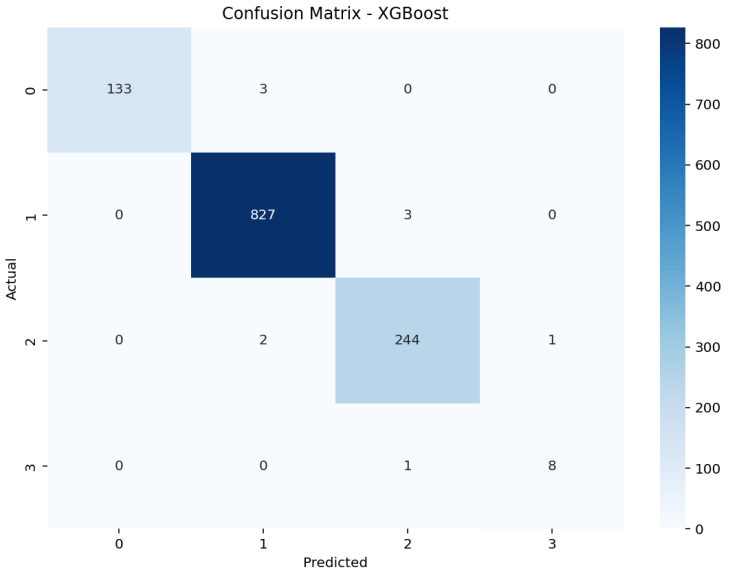


Fig. 9 Confusion matrix for the best performing model (XGBoost).

low and moderate or moderate and high transmission states. Misclassification between extreme categories was minimal, indicating that the model preserved the ordinal structure of transmission intensity.

3.6 Discrimination Performance

Receiver operating characteristic curves (Table 2) and precision recall curves (Figures 10?13) were used to evaluate model discrimination. Extreme gradient boosting achieved an area under the receiver operating characteristic curve of 0.9998, indicating excellent class separation. Random forest demonstrated comparable receiver operating characteristic performance, although precision recall performance was slightly reduced for less frequent classes.

Multinomial logistic regression showed moderate discrimination ability, while the support vector machine exhibited lower performance across both receiver operating characteristic and precision recall analyses.

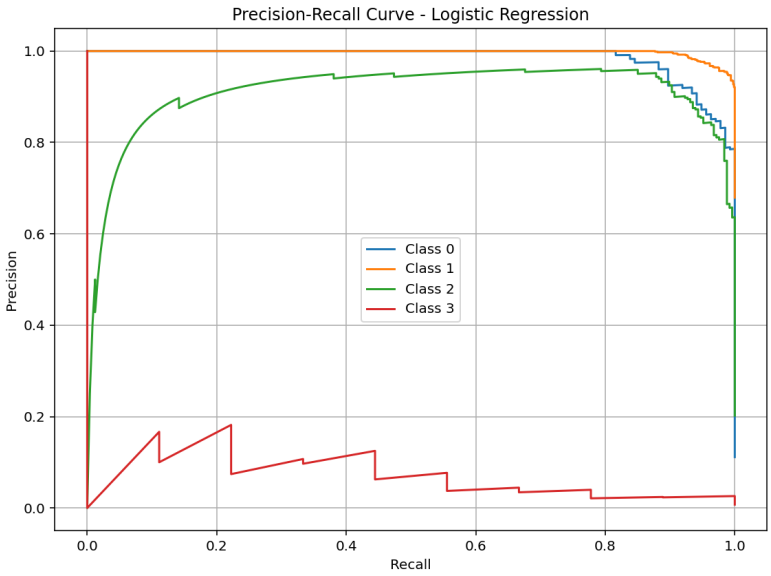


Fig. 10 Precision-recall curves for Logistic Regression.

3.7 Computational Efficiency

Training time and inference latency were evaluated to assess computational efficiency (Table 2). Extreme gradient boosting required the shortest training time at 0.6363 seconds and maintained low inference latency. Random forest required slightly longer training time, while logistic regression exhibited fast inference but moderate training time. The support vector machine was the most computationally intensive model for both training and inference.

3.8 Model Interpretability

Explainable artificial intelligence methods were applied to the best performing model to enhance interpretability.

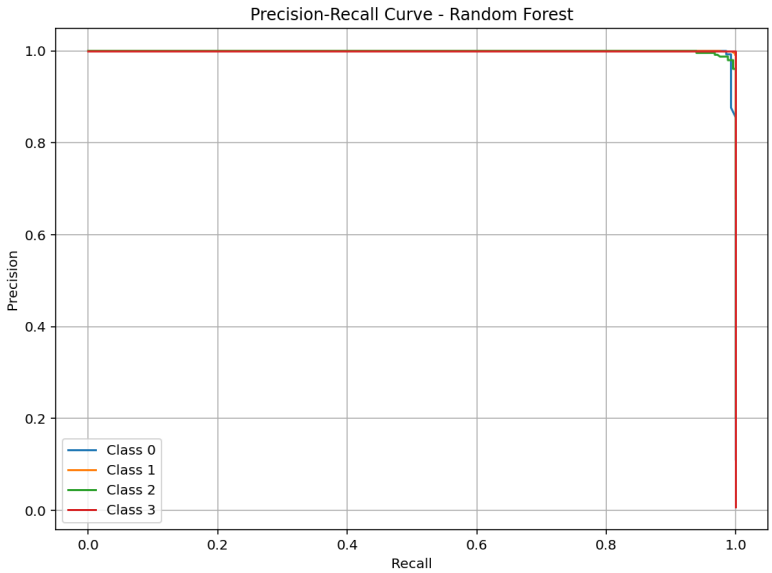


Fig. 11 Precision-recall curves for Random Forest.

3.8.1 Global Feature Importance

Feature contributions were quantified using SHapley Additive exPlanations. Mean absolute SHAP values for each predictor, aggregated across all classes, are presented in Figure 14. Lagged malaria incidence (one month) emerged as the most influential variable, with substantially higher contribution compared to other predictors. Environmental variables, including Normalised Difference Vegetation Index and precipitation, also demonstrated strong influence on model predictions, while insecticide treated net coverage showed moderate contribution. Lagged versions of these variables retained notable importance, indicating the persistence of temporal effects. In contrast, elevation and population density contributed less to the model, reflecting their relatively static nature.

The distribution of SHAP values for the most important features is illustrated in Figure 15. Higher values of lagged incidence were associated with positive SHAP values, corresponding to increased predicted probability of higher transmission states. Similar positive associations were observed for vegetation index and precipitation. In contrast, insecticide treated net coverage exhibited predominantly negative SHAP values, indicating lower predicted transmission probability at higher coverage levels. The dispersion of SHAP values across observations suggests variability in feature effects, particularly for lagged predictors, highlighting heterogeneity in the underlying relationships captured by the model.

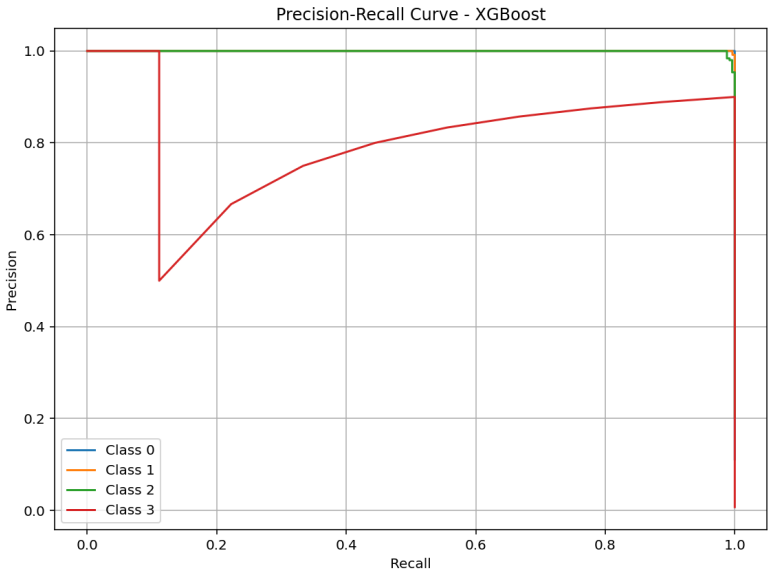


Fig. 12 Precision-recall curves for XGBoost.

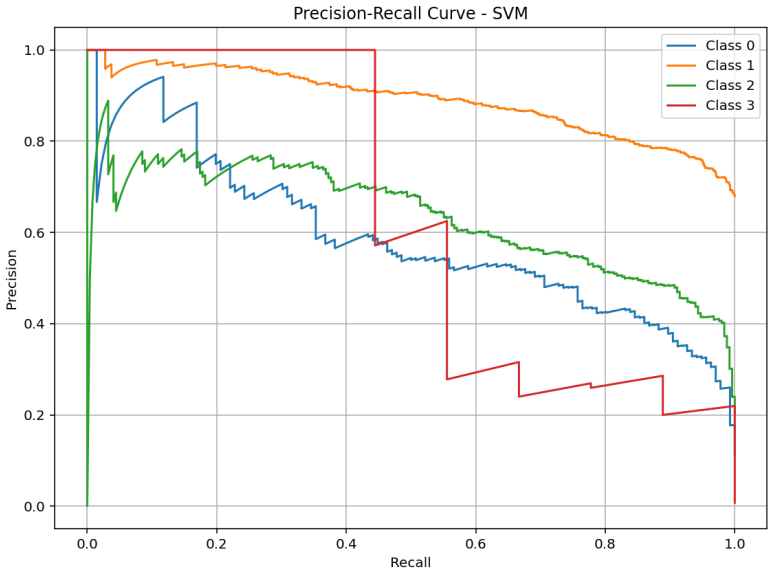


Fig. 13 Precision-recall curves for SVM.

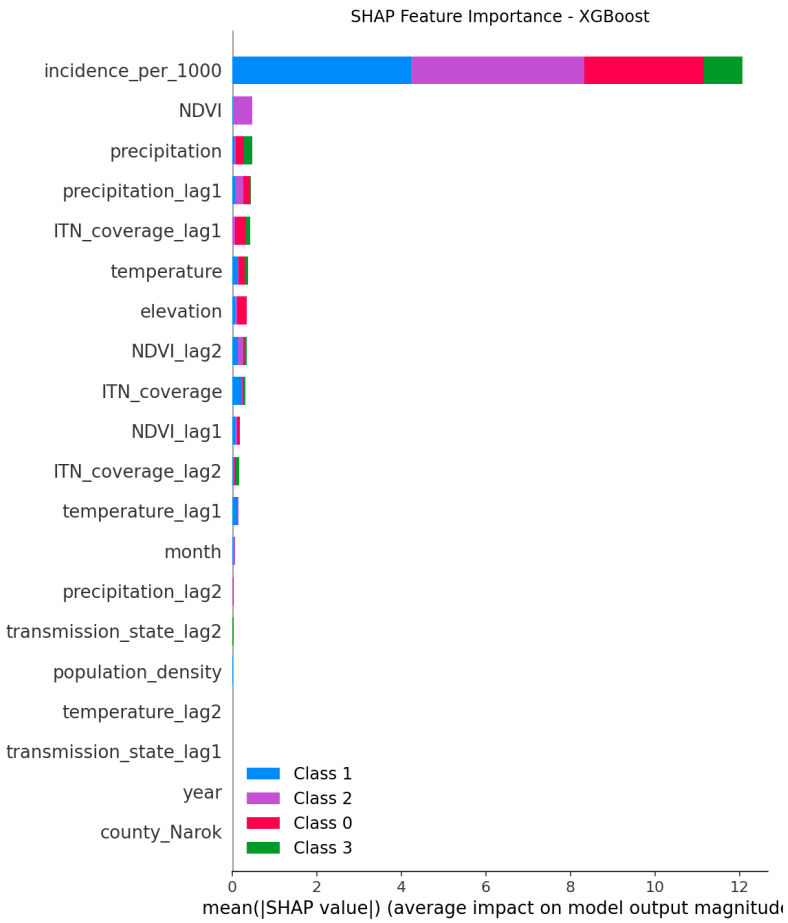


Fig. 14 Mean absolute SHAP values for the top features, indicating global feature importance across all transmission states.

3.8.2 Partial Dependence Plots

Partial dependence plots were used to examine the marginal effect of selected predictors on the predicted probability of each transmission state, while averaging over the distribution of all other variables (Figure 16 and Figures 17–20).

The relationship between temperature and predicted transmission probabilities is shown in Figure 16. The probability of moderate and high transmission states increased with temperature up to approximately 28°C, after which the relationship stabilised or showed a slight decline. In contrast, the probability of very high transmission increased steadily across the observed temperature range, while the probability of low transmission decreased as temperature increased.

Additional partial dependence plots (Figures 17–20) illustrate the effects of year, month, and temperature across transmission states. The probability of moderate transmission remained relatively stable over time but exhibited clear seasonal variation,

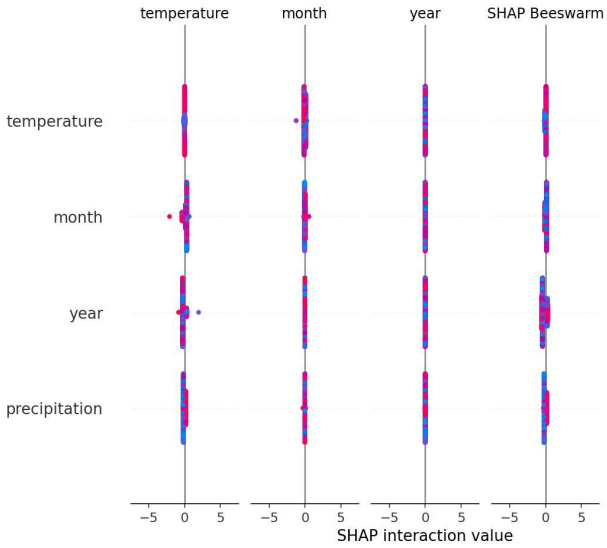


Fig. 15 Beeswarm plot of SHAP values for the most important features. Each point represents a single observation, with colour indicating feature value (red = high, blue = low).

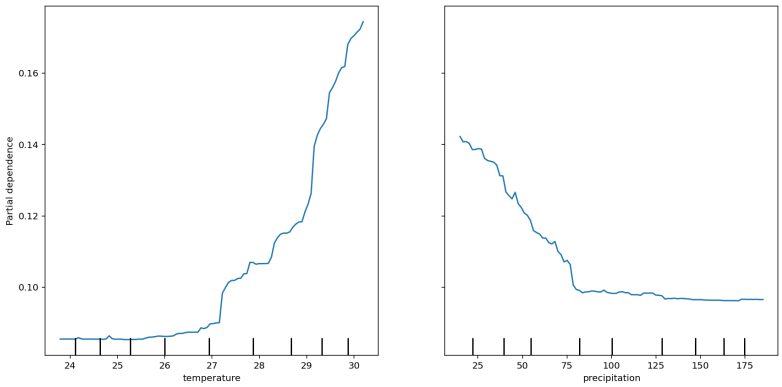


Fig. 16 Partial dependence plot showing the effect of temperature on predicted probabilities across transmission states.

with higher probabilities during periods corresponding to the long and short rainy seasons. High transmission showed a gradual decline over the study period, whereas very high transmission remained low for most years. Low transmission displayed an inverse pattern, with higher probabilities observed in periods associated with lower rainfall.

These results indicate that the model captures nonlinear relationships and seasonal variation in malaria transmission, consistent with the temporal structure of the data.

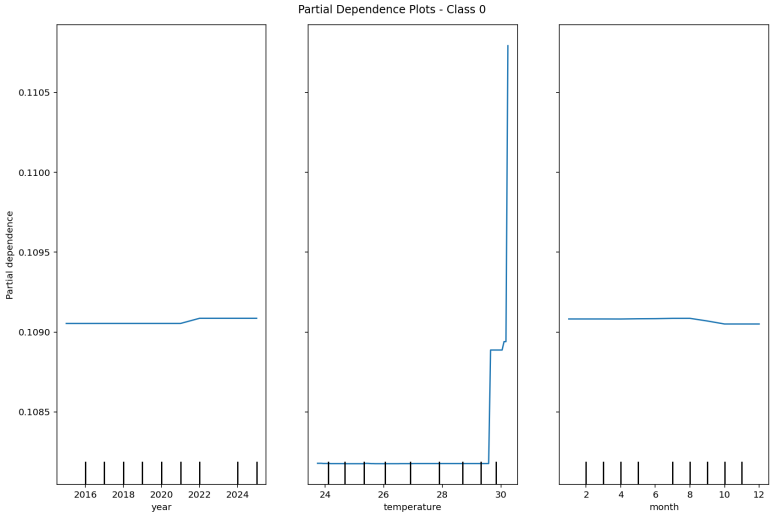


Fig. 17 Partial dependence plots for class 0 (low transmission) with respect to year, month, and temperature.

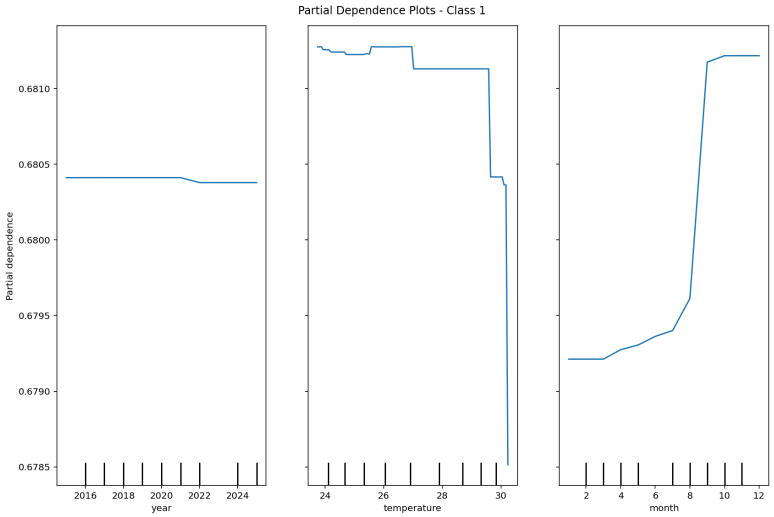


Fig. 18 Partial dependence plots for class 1 (moderate transmission).

3.8.3 Local Explanations with LIME

While SHAP and PDP provide global interpretability, LIME was used to explain individual predictions. Figures 21?23 illustrate three representative examples. In each case, the explanation highlights which features contributed most to the predicted class, along with their contribution weights. For instance, a prediction of moderate transmission was driven primarily by incidence per 1,000 falling within a certain range

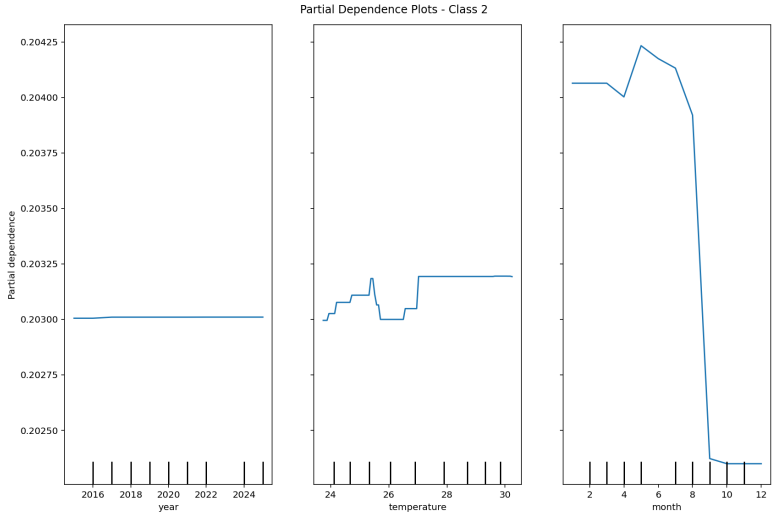


Fig. 19 Partial dependence plots for class 2 (high transmission).

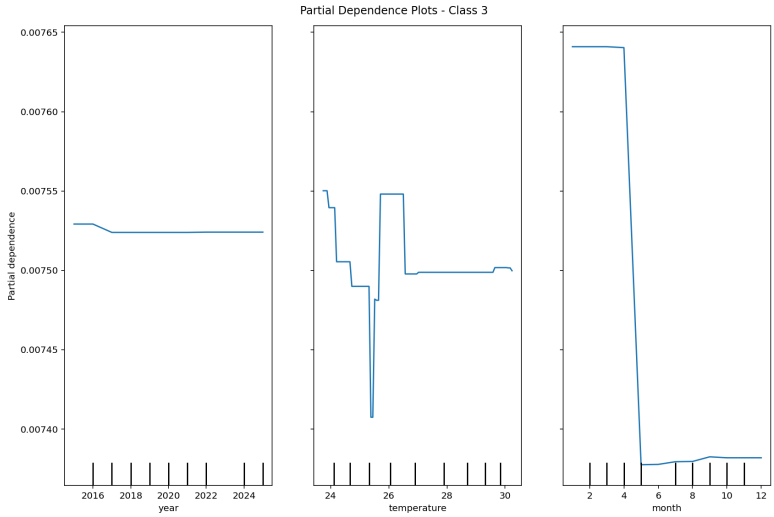


Fig. 20 Partial dependence plots for class 3 (very high transmission).

and the county being in a high?risk region (e.g., counties around Lake Victoria). These local explanations allow public health officials to understand why a specific county?month was classified into a particular transmission state, thereby facilitating targeted interventions.

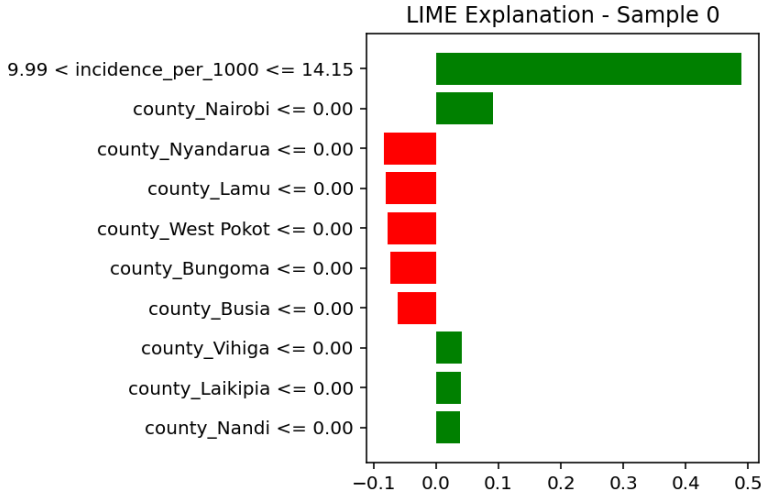


Fig. 21 LIME explanation for an instance predicted as class 0 (low transmission). The features and their contributions are shown.

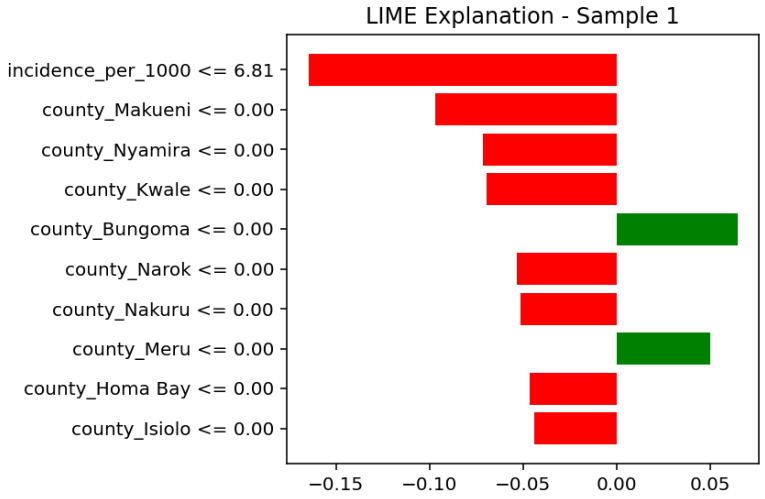


Fig. 22 LIME explanation for an instance predicted as class 1 (moderate transmission).

4 Discussion

This study developed and evaluated a machine learning framework for classifying malaria transmission states in Kenya using spatio temporal panel data. The results show that ensemble tree based methods, particularly extreme gradient boosting, achieved superior predictive performance, strong probabilistic calibration, and consistent interpretability across multiple evaluation criteria. These findings demonstrate

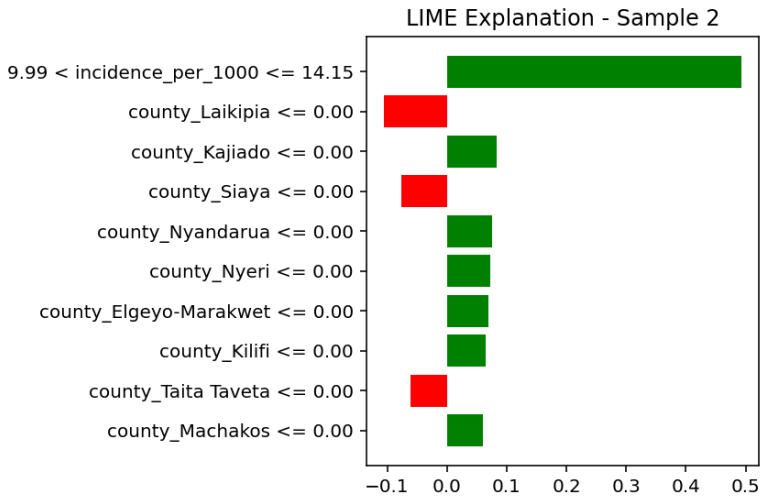


Fig. 23 LIME explanation for an instance predicted as class 2 (high transmission).

that data driven approaches can provide accurate and reliable tools for malaria surveillance and decision making in endemic settings.

4.1 Model Performance and Predictive Accuracy

Extreme gradient boosting consistently achieved the highest performance across all evaluation metrics. This finding is consistent with previous studies that report strong performance of boosting methods in epidemiological prediction tasks [9, 18]. The superior performance can be attributed to the ability of boosting algorithms to capture nonlinear relationships and complex interactions among climatic, environmental, and intervention related variables.

However, the very high performance observed in this study should be interpreted with caution. Malaria transmission exhibits strong temporal persistence, and the inclusion of lagged predictors, particularly past incidence and transmission states, provides substantial predictive information. As a result, part of the predictive performance reflects the continuity of transmission dynamics rather than purely novel pattern discovery. This highlights the importance of temporal validation and careful interpretation of model outputs in epidemiological applications.

Random forest also demonstrated strong performance, although with slightly lower macro averaged F1 score. This difference may reflect the sequential learning strategy of boosting, which iteratively reduces prediction errors. In contrast, multinomial logistic regression showed lower performance, suggesting limited capacity to represent nonlinear dependencies. The support vector machine exhibited the lowest performance, which may be due to sensitivity to parameter selection and challenges in handling complex temporal structures.

4.2 Importance of Probabilistic Calibration

In addition to predictive accuracy, this study emphasised the importance of probabilistic calibration. Extreme gradient boosting produced well calibrated probability estimates, as indicated by low Brier scores and calibration curves closely aligned with observed outcomes. In contrast, logistic regression and support vector machine models produced less reliable probability estimates.

Calibration is critical in public health decision making because it determines whether predicted risks correspond to actual observed outcomes. Models that are poorly calibrated may produce misleading confidence in predictions, which can result in inefficient allocation of resources or delayed interventions. The results therefore support the inclusion of calibration assessment as a standard component of model evaluation in malaria prediction studies.

4.3 Interpretation of Key Drivers of Transmission

The explainability analysis provides insight into the drivers of malaria transmission. Lagged incidence emerged as the most influential predictor, confirming the strong temporal dependence of malaria dynamics. Environmental variables such as vegetation index and precipitation were also important, reflecting their role in creating suitable conditions for mosquito breeding and survival [4, 5]. Insecticide treated net coverage was associated with reduced transmission probability, consistent with the protective effect of vector control interventions.

Partial dependence analysis revealed nonlinear relationships between temperature and transmission probability, with increasing risk up to a threshold level followed by stabilisation. This pattern aligns with established biological mechanisms governing mosquito development and parasite replication [7]. Seasonal variation in predicted transmission further reflects rainfall driven dynamics that are characteristic of malaria transmission in East Africa [2, 15].

While these findings are consistent with existing epidemiological knowledge, it is important to note that the interpretations are based on statistical associations rather than causal relationships. Explainability methods provide useful approximations of model behaviour but should not be interpreted as definitive evidence of causality.

4.4 Implications for Malaria Surveillance and Control

The classification of malaria transmission into discrete states offers clear practical advantages for public health planning. Unlike continuous prediction of case counts, state based classification aligns directly with operational categories used in malaria control programs. This facilitates identification of areas requiring intensified intervention, enhanced surveillance, or resource reallocation.

The combination of high predictive performance, reliable probability estimates, and interpretability suggests that the proposed framework can support early warning systems and decision support tools. In particular, calibrated probabilities can assist decision makers in prioritising high risk areas under resource constraints. The relatively low computational cost of the best performing model further supports its feasibility for integration into routine surveillance systems.

4.5 Comparison with Existing Studies

Previous studies have applied machine learning techniques to malaria prediction, focusing primarily on continuous outcomes such as incidence or prevalence [11–13]. In contrast, this study models malaria transmission as discrete operational states, which are more directly aligned with decision making frameworks used in practice.

In addition, this study incorporates calibration assessment and explainability methods within a unified framework. These aspects are often overlooked in existing work, despite their importance for practical implementation. The use of forward chaining validation further distinguishes this study by providing a more realistic evaluation of model performance under temporal dependence, in contrast to random validation approaches that may lead to optimistic estimates [18].

4.6 Limitations

Several limitations should be considered. First, the analysis relies on routinely collected surveillance data, which may be affected by reporting delays, underreporting, and variation in diagnostic practices. These factors can introduce measurement error and influence model performance.

Second, the use of lagged outcome variables, while improving predictive accuracy, may limit the ability of the model to generalise to settings with different transmission dynamics. Future work should evaluate model performance under alternative feature specifications to better understand this dependency.

Third, the classification thresholds used to define transmission states may not fully capture regional heterogeneity. Sensitivity analysis using alternative thresholds could provide further insight into the robustness of the results.

Fourth, not all relevant predictors were available, including entomological indicators, indoor residual spraying coverage, and socioeconomic factors. Inclusion of such variables may improve model performance and enhance understanding of transmission dynamics.

Finally, the study focuses on Kenya, and the extent to which the findings generalise to other regions remains uncertain.

4.7 Future Research

Future research should prioritise external validation using independent datasets from other regions to assess generalisability. Incorporating spatial dependence through spatial modelling approaches may improve representation of geographic transmission patterns.

Integration with mechanistic models could enable scenario analysis under varying environmental and intervention conditions. In addition, the development of operational decision support tools that combine predictions, uncertainty estimates, and interpretability outputs could facilitate adoption by public health practitioners.

Further investigation of advanced temporal modelling approaches, including deep learning methods, may provide additional improvements in capturing complex transmission dynamics.

5 Conclusion

This study developed and evaluated a machine learning framework for the classification of malaria transmission states in Kenya using longitudinal surveillance data. The findings show that ensemble tree based methods, particularly extreme gradient boosting, achieve high predictive performance, produce reliable probability estimates, and provide consistent interpretability across multiple evaluation criteria.

The results highlight the importance of temporal information in modelling malaria dynamics. The inclusion of lagged incidence and environmental variables improves predictive capability and reflects the persistence and seasonality of transmission. Key predictors identified by the model, including vegetation index, precipitation, and insecticide treated net coverage, are consistent with established epidemiological evidence, supporting the credibility of the model outputs. The incorporation of probabilistic calibration further strengthens the framework by ensuring that predicted risks are meaningful and suitable for decision making.

Modelling transmission as discrete states provides a practical advantage by aligning predictions with operational categories used in malaria control programs. This approach facilitates more direct translation of model outputs into public health action, including early warning, prioritisation of high risk areas, and efficient allocation of limited resources. In this context, the proposed framework offers a useful tool for strengthening malaria surveillance and supporting evidence based intervention strategies.

However, the strong predictive performance observed in this study is partly influenced by temporal dependence in the data, and caution is required when generalising the results to settings with different transmission patterns. Further validation using independent datasets and alternative feature specifications is necessary to assess robustness and generalisability. In addition, incorporating spatial information, entomological indicators, and health system variables may provide a more comprehensive representation of transmission dynamics.

Future research should focus on external validation, integration with spatial and mechanistic modelling approaches, and development of operational decision support systems that combine prediction, uncertainty, and interpretability. Embedding such frameworks within routine surveillance platforms has the potential to enhance real time monitoring and improve the effectiveness of malaria control efforts.

In conclusion, this study demonstrates that machine learning methods, when combined with calibration assessment and interpretability techniques, can provide accurate, transparent, and operationally relevant tools for modelling malaria transmission in endemic settings.

Acknowledgements

We acknowledge the Kenya Ministry of Health for providing access to malaria surveillance data through the District Health Information System (DHIS2).

Ethics Approval and Consent to Participate

This study used secondary aggregated data obtained from publicly available sources and national surveillance systems. No individual-level, identifiable, or confidential data were used. Therefore, ethical approval and informed consent to participate were not required.

Consent for Publication

Not applicable.

Ethics Declaration

The study was conducted using anonymized aggregated secondary data and did not involve human participants directly. Ethical approval was therefore not required.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Competing Interests

The author declare no competing interests.

Data Availability Statement

The data underlying the results presented in this study are publicly available in Zenodo and can be accessed through the following link https://doi.org/10.5281/zenodo.20051237.

References

[1] World Health Organization: World Malaria Report 2023. WHO, Geneva (2023)

[2] Omondi, C.J., Okiro, E.A., Snow, R.W.: Spatial and temporal dynamics of malaria in kenya: 2015–2020. Malaria Journal **21**(1), 245 (2022)

[3] Macharia, P.M., Okiro, E.A., Snow, R.W.: Sub-national mapping of malaria risk in kenya. BMJ Global Health **6**(4), 004956 (2021)

[4] Mogeni, P., Williams, T.N., Bejon, P.: Effect of temperature on malaria transmission in western kenya. American Journal of Tropical Medicine and Hygiene **103**(2), 789–796 (2020)

[5] Dieng, M.D.B., Munga, S., Tompkins, A.M., Zornoza, M.G., Caminade, C., Fersch, B., Kunstmann, H.: High resolution physically based modelling reveals malaria incidence reduction by vector control measures. Scientific Reports (2026)

[6] Camponovo, F., Jeandron, A., Skrip, L.A., Golumbeanu, M., Champagne, C., Symons, T.L., Pothin, E.: Malaria treatment for prevention: a modelling study of the impact of routine case management on malaria prevalence and burden. BMC Infectious Diseases **24**(1), 1267 (2024)

[7] Okuneye, K., Gumel, A.B.: Analysis of a temperature- and rainfall-dependent model for malaria transmission dynamics. Mathematical Biosciences **287**, 72–92 (2017)

[8] Lusiyana, N., Ahdika, A.: Evaluating recurrent episodes of malaria incidence in timika, indonesia, through a markovian multiple-state model. Infectious Disease Modelling **7**(3), 261–276 (2022)

[9] Mbunge, E., Milham, R.C., Sibiya, M.N., Takavarasha Jr, T.: Machine learning techniques for predicting malaria: Unpacking emerging challenges and opportunities for tackling malaria in sub-saharan africa. In: Computer Science On-line Conference, pp. 327–344. Springer, Cham (2023)

[10] Ashaolu, J.O., Akanji, T.S., Ayansola, V.I., Olawale-Succes, O.O., Sunday, A.J., Some, S.Y.: From risk factors to predictive modelling: applying machine learning to childhood malaria surveillance in resource-limited settings. BMC Infectious Diseases **25**(1), 1693 (2025)

[11] Ayoka, T.O., Nnadi, C.O.: Application of machine learning models in predicting malaria prevalence in nigeria: an analysis of the 2015–2020 demographic and health surveys. Journal of Parasitic Diseases, 1–13 (2025)

[12] Baba-Adamu, M., Ngamdu, M.B., Adamu, U., Jajere, A.A.: Machine learning analyses of climate variability trends and malaria transmission dynamics in yobe state, nigeria. UMYU Scientifica **4**(2), 258–269 (2025)

[13] Morang'a, C.M., Amenga–Etego, L., Bah, S.Y., Appiah, V., Amuzu, D.S., Amoako, N., *et al.*: Machine learning approaches classify clinical malaria outcomes based on haematological parameters. BMC Medicine **18**(1), 375 (2020)

[14] Taconet, P., Porciani, A., Soma, D.D., Mouline, K., Simard, F., Koffi, A.A., *et al.*: Data-driven and interpretable machine-learning modeling to explore the fine-scale environmental determinants of malaria vectors biting rates in rural burkina faso. Parasites & Vectors **14**(1), 345 (2021)

[15] Zheng, J.X., Lu, S.N., Li, Q., Li, Y.J., Xue, J.B., Gavana, T., *et al.*: Deciphering the climate-malaria nexus: A machine learning approach in rural southeastern tanzania. Public Health **238**, 124–130 (2025)

[16] Monteiro, K.H.d.C., Silva Rocha, É., Morais, L.A., Santos, E.G., Neto, S.R.d.S., Sampaio, V., Endo, P.T.: Integrating machine learning and spatial clustering for malaria case prediction in brazil's legal amazon. BMC Infectious Diseases **25**(1), 802 (2025)

[17] Golumbeanu, M., Yang, G.J., Camponovo, F., Stuckey, E.M., Hamon, N., Mondy, M., *et al.*: Leveraging mathematical models of disease dynamics and machine learning to improve development of novel malaria interventions. Infectious Diseases of Poverty **11**(03), 37–53 (2022)

[18] Khan, O., Ajadi, J.O., Hossain, M.P.: Predicting malaria outbreak in the gambia using machine learning techniques. PLoS One **19**(5), 0299386 (2024)