ISSN 2282-0930 • Epidemiology Biostatistics and Public Health - 2026, Volume 21, Issue 1

Towards Reliable Malaria Forecasting: A Noise-Resilient Wavelet-STL Hybrid Framework for Cameroon

Akindeh Mbuh Nji[(1,2,3,8)*], *Taiwo M. Adegoke*[(1, 7)*], *Sheetal P. Silal*[(8)], *Innocent M. Ali*[(3,9)], *Andrillene L. D. Wondeu*[(1)], *Pascal F. Madeleine*[(1,2)], *Masso T. Sylvanie*[(1,2)], *Calvino Fomboh Tah*[(1,2)], *Christophe Antonio-Nkondjio*[(10)], *Jude D. Bigoga*[(2,3)], *Rose G. F. Leke*[(3)], *Gillian Stresman*[(4,6)], *Liwang Cui*[(4,5)], *Wilfred F. Mbacham*[(1,2,3)]

(1) The Laboratory for Public Health Research Biotechnologies, MARCAD Plus Program, The Biotechnology Centre, University of Yaoundé 1, BP 8094, Yaoundé, Centre Region, Cameroon (ROR: 022zbs961)
(2) Department of Biochemistry, Faculty of Science, University of Yaounde, Cameroon (ROR: 022zbs961)
(3) ICEMR for Central Africa, FINISTECH, Cameroon
(4) ICEMR for Central Africa, University of South Florida (ROR: 032db5x82)
(5) Center for Global Health and Infectious Diseases Research, College of Public Health, University of South Florida,Tampa, Florida, USA (ROR: 032db5x82)
(6) Department of Epidemiology, College of Public Health, University of South Florida, 13201 Bruce B Downs Blvd, Tampa, FL, USA (ROR: 032db5x82)
(7) Department of Mathematics and Statistics, Abiola Ajimobi Technical University, Ibadan (ROR: 05tb13r23)
(8) Modelling and Simulation Hub, Africa, Department of Statistical Sciences, University of Cape Town, Cape Town, South Africa (ROR: 03p74gp79)
(9) Department of Biochemistry, Faculty of Science, BP 67, University of Dschang, Dschang, West Region, Cameroon (ROR: 0566t4z20)
(10) Malaria Research Laboratory, OCEAC, Yaoundé, Centre, PO Box 288, Cameroon (ROR: 02fywtq82)

** These auhors contributed equally to this work*

CORRESPONDING AUTHOR: Akindeh Mbuh Nji and Taiwo M. Adegoke, The Laboratory for Public Health Research Biotechnologies, MARCAD Plus Program, The Biotechnology Centre, University of Yaoundé 1, BP 8094, Yaoundé, Centre Region, Cameroon. Email: akindeh@gmail.com, adegoketaiwom@gmail.com

SUMMARY

Background: In Cameroon, malaria is a significant public health issue, with heterogenous distribution and seasonal change making control and containment harder to plan. Medium-term projections are needed to provide reliable early warning and optimal prevention resource allocation. Traditional methods, including ARIMA and SARIMA, are dogged by noisy surveillance data as well as structural non-stationarity, and the hybrids currently being applied rarely measure the uncertainty in the forecasts.
Methods: We introduce a multi-step hybrid forecasting pipeline to combine wavelet-based denoising, robust Seasonal-Trend decomposition with Loess (STL), and state-of-the-art remainder modeling. The remaining component was decomposed by using ARIMA, SARIMA, or Bayesian Structural Time Series (BSTS) and forecasts were recreated out of all components. The analysis was conducted on monthly malaria incidence in the ten administrative regions of Cameroon and 24-month future projections were developed. RMSE, MAE, R2, and information criteria were used to evaluate model performance, and uncertainty was measured using analytical intervals (ARIMA/SARIMA) and posterior predictive distributions (BSTS).
Results: The Wavelet STL preprocessing significantly enhanced model stability, and model accuracy in all regions. The predictive performance of ARIMA and SARIMA models was similar, and the R2 values were between 0.49 and 0.77 following seasonal adjustment. The similarity implies that the use of STL decomposition actually eliminated seasonal variations in the time series, thus, limiting the benefit of the explicit representation of seasonality in SARIMA. Thus, the other non-seasonal dynamics were likewise captured in both models. In many areas BSTS was significantly higher or as large as ARIMA/SARIMA, and obtained higher R2 values. Notably, BSTS offered probabilistic predictions that were calibrated, which allowed to effectively measure the forecast uncertainty. The results presented in this paper indicate that the hybrid pipeline suggested is both noisy and uncertain, and can provide forecasts of malaria cases in the villages

DOI: 10.54103/2282-0930/30118
Accepted: 13th November 2025

of Cameroon over a period of 24 months. The uncertainty in the forecast was explicitly measured by 95% prediction interval and the model validation revealed that most of the cases observed were within the interval which implies a good predictive ability.
Conclusion: The Wavelet-STL hybrid model is the next step in malaria prediction by combining the denoising, structural decomposition, and probabilistic models. Its deployment in the regions of Cameroon shows innovative methodological value and direct applicability to early warning systems. The method can be easily extended to other infectious diseases with seasonal spread and noisy surveillance data.

Keywords: Malaria forecasting, ARIMA, SARIMA, Bayesian Structural Time Series, Wavelet transformation

INTRODUCTION

In Sub-Saharan Africa, malaria is one of the most prevalent diseases causing morbidity and mortality, and in the case of Cameroon, one of the most victimized countries [1, 2]. The dynamics of transmission differ significantly between ecological zones, leading to intense regional and strong seasonal heterogeneity. Medium-term predictions (12-24 months) are then essential to early warning, intervention planning and effective preventive and therapeutic resource allocation [3, 4]. However, malaria surveillance series are both susceptible to noises, non-stationary, and often interrupted by outbreaks or intervention programs, which makes them extremely difficult to forecast.

Autoregressive-based models such as Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) have longstanding applications in the prediction of infectious diseases [5, 6]. Although useful in modelling short-run dependencies, the models are very susceptible to noise, structural break and model specifications [5]. Hybrids of wavelet and ARIMA have been recently proposed to improve robustness by filtering with wavelets prior to ARIMA estimation [7, 8], although they generally do not explicitly separate trend and seasonality. Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) models have also been considered as methods of machine learning [9, 10], but they typically need large training sets, are not interpretable, and are challenging to implement in national health information systems.

To address the limitations, we build a multi-stage hybrid forecasting pipeline that combines wavelet-based denoising, a powerful seasonal-trend decomposition (STL) and higher order modeling of the stochastic residue. In particular, the pipeline consists of: (i) Maximal Overlap Discrete Wavelet Transform (MODWT) denoising, to remove the high-frequency noise, (ii) STL, in order to extract deterministic trend and seasonality, (iii) the rest of the forecasts are modeled by means of ARIMA, SARIMA, or Bayesian Structural Time Series (BSTS), and (iv) reconstructing predictions based on component predictions. Figure 1 gives a schematic view of the framework. This design, unlike the previous ones, clearly separates deterministic and stochastic dynamics, is more interpretable, and has calibrated probabilistic forecasts with the Bayesian version. The novelty is a systematic application and comparison of Wavelet-STL-ARIMA, Wavelet-STL-SARIMA, and Wavelet-STL-BSTS to all ten Cameroon regions, showing innovation in methodology and to the operational aspect in controlling malaria. In addition to malaria, the framework can be generalized to other infectious diseases where surveillance data is noisy and seasonal, allowing it to contribute to the broader epidemic forecasting and public health preparedness field.

METHODS

Study Area

Cameroon is located in Central Africa and is known for its rich geographic and cultural diversity. The country stretches from 1° to 13° north and 8° to 16° east, covering a range of landscapes, including coastal plains, mountains, and tropical rainforests. It has a population of over 25 million people, spread across its ten regions: Adamaoua, Centre, East, Far North, Littoral, North, Northwest, Southwest, West, and South. Health outcomes and patterns of disease in Cameroon are shaped by differences in climate, economic conditions, and access to healthcare across regions. This diversity affects how diseases spread, how healthcare resources are distributed, and how well public health systems function.

Data Sources

The study used reported malaria case data obtained from the national District Health Information System 2 (DHIS2) database which the Ministry of Public Health in Cameroon maintains. The dataset contains monthly malaria-related case reports which cover all ten administrative regions throughout the period from January 2015 to December 2024. The DHIS2 platform serves as a primary electronic health information system in Cameroon for both healthcare data reporting and surveillance routines.

Study Design and Settings

Our pipeline is a multi-stage hybrid forecast using

Cameroon regional malaria incidence. The national DHIS2 surveillance system included monthly case counts that were aggregated to the ten administrative regions and analyzed as region-specific univariate time series. The contribution is methodological: a single pipeline that (i) denoises, (ii) structurally decomposes, (iii) models the stochastic remainder using multiple base learners, (iv) recomposes prediction and (v) may or may not ensemble the base learners.

Data Framework

We examine monthly series of malaria incidence in each of the ten administrative regions in Cameroon. The modeling structure assumes modeling regions in isolation, allowing the heterogeneous dynamics of time but comparing regions with each other.

Figure 1. Conceptual flow of the Wavelet–STL–Hybrid pipeline

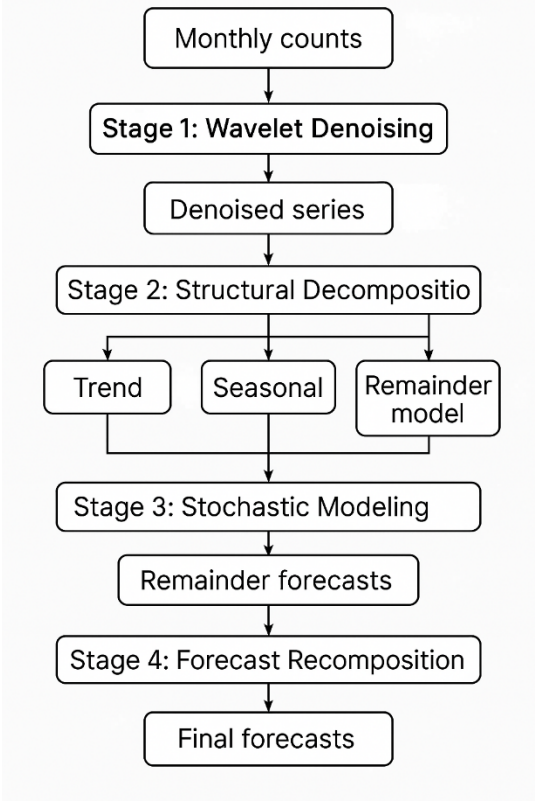


Figure 1 shows the conceptual framwork for the study. Regional malaria series are denoised via MODWT, decomposed into trend (T_t), seasonal (S_s), and remainder (R_t) using STL. The remainder is modeled by ARIMA, SARIMA, or BSTS. Forecasts are recomposed as $\hat{Y}_{t+h} = \hat{T}_{t+h} + \hat{S}_{t+h} + \hat{R}_{t+h}$.

Stage 1: Wavelet Denoising

Let represents the monthly malaria counts at time. To reduce noise in measurements and temporary variations [8], we use the Maximal Overlap Discrete Wavelet Transform (MODWT) using a Daubechies la8 filter at four levels of decomposition:

$$Y_t = \sum_{j=1}^{J} W_{j,t} + V_{J,t} \quad (1)$$

Detail coefficients $\{W_{j,t}\}$ are soft-thresholded using a universal threshold $\lambda = \hat{\sigma}\sqrt{2\log n}$, with $\hat{\sigma}$ estimated from the median absolute deviation of the first-level coefficients. The denoised series $\tilde{Y}_t$ is reconstructed via the inverse MODWT:

$$\tilde{Y}_t = \mathrm{IMODWT}(\{\widetilde{W}_{j,t}\}, V_{J,t}) \quad (2)$$

Stage 2: Structural Decomposition

The denoised series $\tilde{Y}_t$ is decomposed into trend (T_t), seasonal (S_t), and remainder (R^t) components using robust STL [11]

$$\tilde{Y}_t = T_t + S_t + R_t, \qquad \sum_{m=1}^{12} S_{t+m} = 0 \quad (3)$$

In this case, seasonality will be limited to a periodical cycle with period of 1 month (s=12). A non-seasonal ARIMA model is used to predict future values of the trend and the deterministic propagation of seasonality is achieved through calendar repetition:

$$\hat{S}_{t+h} = S_{(t+h)\bmod 12} \quad (4)$$

Stage 3: Modeling the Stochastic Remainder

The stochastic component R_t has an irregular and autoregressive dynamics which are not captured by the less irregular trend or deterministic seasonality [5]. We are comparing 3 competing modeling strategies:

Wavelet–STL–ARIMA

We fit an ARIMA (*p,d,q*) model to R_t:

$$\phi(B)(1-B)^d R_t = \theta(B)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,\sigma^2) \quad (5)$$

where *B* is the backshift operator, $\phi(B)$ and $\theta(B)$ are autoregressive and moving-average polynomials, and is the differencing order.

Wavelet–STL–SARIMA

When residual seasonal dependence remains, we use a Seasonal ARIMA $(p,d,q)\times(P,D,Q)_{12}$ specification:

$$\Phi(B^{12})\phi(B)(1-B)^d(1-B^{12})^D R_t = \Theta(B^{12})\theta(B)\varepsilon_t \quad (6)$$

This allows seasonal autoregressive (*P*), differencing (*D*), and moving-average (*Q*) components to capture remaining cyclic autocorrelation.

Wavelet–STL–BSTS

For a fully Bayesian treatment, we specify a Bayesian Structural Time Series (BSTS) model with a local level and slope for R_t:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (7)$$

$$\beta_t = \beta_{t-1} + \zeta_t \quad (8)$$

$$R_t = \mu_t + \varepsilon_t \quad (9)$$

with $\eta_t \sim \mathcal{N}(0,\sigma_\eta^2), \zeta_t \sim \mathcal{N}(0,\sigma_\zeta^2)$ and $\varepsilon_t \sim \mathcal{N}(0,\sigma_\varepsilon^2)$.

Posterior inference is conducted via MCMC, providing predictive distributions and credible intervals.

Stage 4: Forecast Recomposition

For each modeling approach $m \in \{\text{ARIMA}, \text{SARIMA}, \text{BSTS}\}$, forecasts are recombined as:

$$\hat{Y}_{t+h}^{(m)} = \hat{T}_{t+h} + \hat{S}_{t+h} + \hat{R}_{t+h}^{(m)} \qquad (10)$$

This ensures interpretability: long-term structure is driven by T_t and S^t, while short-term stochasticity is modeled by R_t.

i. ARIMA/SARIMA: The analysis form of the h-step prediction error variance in the Gaussian ARIMA model provides forecast uncertainty. Specifically, if $\hat{Y}_{t+h}$ denotes the point forecast, the forecast error variance is

$$\text{Var}\left(Y_{t+h} - \hat{Y}_{t+h}\right) = \sigma^2 \sum_{i=0}^{h-1} \psi_i^2 \qquad (11)$$

where σ^2 is the innovation variance and $\{\psi_i\}$ are the impulse response weights implied by the fitted ARIMA or SARIMA model. Under the assumption of normally distributed innovations, prediction intervals are computed as

$$\hat{Y}_{t+h} \pm z_{\alpha/2}\sqrt{\text{Var}(Y_{t+h} - \hat{Y}_{t+h})} \qquad (12)$$

for nominal levels 1 – α = 0.80 and 0.95. These intervals are then shifted by the deterministic forecasts of trend and seasonality obtained from STL decomposition to yield full-series uncertainty bands.

ii. BSTS: Conversely, Bayesian Structural Time Series models have no closed-form prediction intervals. Rather the posterior predictive distribution defines uncertainty. We generate *M* draws from the predictive density of the remainder R_{t+h} via MCMC, recompose with deterministic components, and compute empirical quantiles. The 80% and 95% credible intervals are thus defined as the 10th–90th percentiles and 2.5th–97.5th percentiles of the simulated paths, respectively

Training and Evaluation Protocol

Modelling is done separately in each region and the forecast horizon is held constant at 24 months. Performance on the recomposed fitted series is assessed based on standard accuracy measures, including the Root Mean Squared Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). We also provide the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) of the remainder models as an evaluation of relative parsimony in the case of the ARIMA and SARIMA pipelines. Further evaluations on forecast uncertainty are done using empirical interval coverage, average interval width to give a measure of probabilistic calibration.

RESULTS

Figure 2 gives a spatial distribution and annual pattern of the reported number of malaria cases in the 10 administrative regions of Cameroon over a period of 2015 to 2024. The maps are of individual years, and the regions are color-shaded depending on the total number of malaria cases that were registered with the scale light blue representing lower cases and dark blue representing higher cases. Throughout the decade, the Extreme-North region was recorded to have the highest burden of malaria since it was darker in shade in all years. In contrast, regions such as Central, South, and Littoral appear lighter throughout most of the period, indicating a comparatively lower burden. However, some fluctuations can be observed. A slight increase in malaria case intensity in the North and Adamaoua regions from 2019 onward. Spatial pattern displays an evident drop to the south, with the Northern regions producing a cumulative number of cases than southern counterparts per year. This trend conforms with the established malaria ecology in Cameroon with arid northern areas experiencing shorter and intense transmission periods, which southern areas in Cameroon are relatively more stable and well-covered health systems. The time series of all maps shows that the geography of distribution of the cases is relatively stable, but there exist yearly changes in size, which probably points to a mix of climatic factors, coverage of interventions and health-attaining behaviour.

3.1 Model Performance

Table 1 shows the accuracy in-sample measures over all regions and all methods. The performance differed depending on the model selection and regional epidemiology profile.

3.1.1 ARIMA

Wavelet-STL-ARIMA pipeline had average to poor results in all regions, with minimum RMSE of 0.95 in South and a maximum of 25.1 in Far North. The coefficients of determination (R^2) were always positive ranging between 0.49 and 0.77. The model showed good model fit (R2>0.65) in regions with more stable incidence, including the South and the West, but larger residual errors in high-burden and volatile regions including the Far North, despite similar levels of explanatory power (R^2=0.77). The rest of the models, reflected in information criteria (AIC/BIC) suggested that short-term autocorrelation following the STL decomposition could be well explained using ARIMA in a relatively parsimonious way.

3.1.2 SARIMA

The STL adjusted Wavelet-SARIMA model obtained accuracy scores very similar to ARIMA, with RMSE and R2 scores matching within two decimal points in most areas. This indicates that there were minimal residual seasonality in the remaining component after

ISSN 2282-0930 • Epidemiology Biostatistics and Public Health - 2026, Volume 21, Issue 1

STL decomposition. Examples include using SARIMA in the Centre and Littoral regions, where seasonal parameters were chosen, but they did not correspond to quantifiable increases in accuracy. These findings highlight the utility of STL to isolate and eliminate seasonal structure before remainder modeling to maximize the marginal utility of SARIMA.

3.1.3 BST

Wavelet-STL-BSTS pipeline was always better than or equally good as ARIMA and SARIMA in regions. RMSEs were consistently lower/similar and R-squares increased in some instances: Adamawa (0.55 vs. 0.49 ARIMA), Centre (0.72 vs. 0.67) and East (0.64 vs. 0.60). Interestingly, BSTS demonstrated the best fit (R^2=0.79) in the Far North (as the most variable) and, still, allowed meaningful interpretation of state variables. In addition to point precision, the Bayesian model offered calibrated predictive distributions, with a probabilistic view that ARIMA and SARIMA do not have.

Figure 2. Malaria cases by region

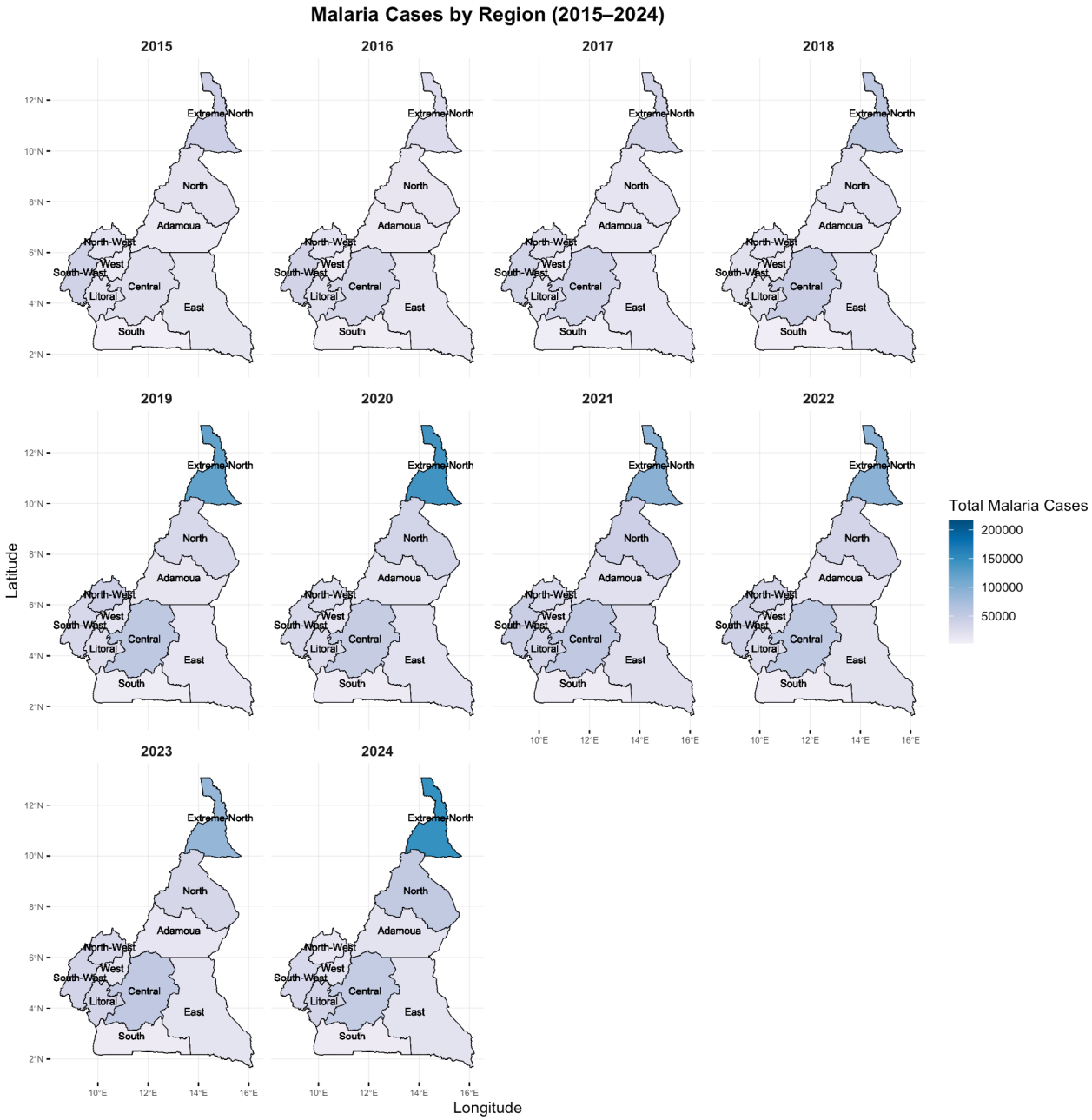


Table 1. Comparative in-sample performance of Wavelet–STL–ARIMA, Wavelet–STL–SARIMA, and Wavelet–STL–BSTS across the ten administrative regions of Cameroon. ARIMA and SARIMA metrics include RMSE, MAE, R^2, AIC, and BIC for the remainder models, while BSTS metrics include RMSE, MAE, and R^2

Region	ARIMA					SARIMA					BSTS		
	RMSE	MAE		AIC	BIC	RMSE	MAE		AIC	BIC	RMSE	MAE	
Adamawa	3.27	2.63	0.494	-7.34	-1.80	3.27	2.63	0.494	-7.34	-1.80	3.09	2.53	0.547
Centre	7.45	5.93	0.671	388.00	399.00	7.47	5.95	0.669	392.00	403.00	6.90	5.64	0.717
East	2.42	1.95	0.603	51.10	62.20	2.37	1.92	0.619	29.90	49.30	2.30	1.88	0.641
Far North	25.10	21.10	0.770	667.00	678.00	24.90	21.10	0.774	635.00	649.00	24.00	20.70	0.789
Littoral	3.80	2.92	0.646	32.90	41.20	3.80	2.92	0.646	32.90	41.20	3.80	2.92	0.646
North	7.56	6.27	0.706	304.00	318.00	7.56	6.27	0.706	304.00	318.00	7.55	6.28	0.706
Northwest	6.54	5.40	0.419	338.00	349.00	6.52	5.38	0.423	336.00	350.00	6.14	5.18	0.489
South	0.95	0.76	0.672	-380.00	-366.00	0.95	0.76	0.673	-381.00	-373.00	0.95	0.76	0.674
Southwest	4.85	3.94	0.621	-83.30	-66.60	4.85	3.94	0.621	-83.30	-66.60	4.81	3.90	0.628
West	2.77	2.25	0.662	68.70	82.60	2.77	2.24	0.661	67.50	81.30	2.70	2.21	0.679

ISSN 2282-0930 • Epidemiology Biostatistics and Public Health - 2026, Volume 21, Issue 1

Figure 3. Regional malaria forecasts using ARIMA models with Wavelet–STL preprocessing

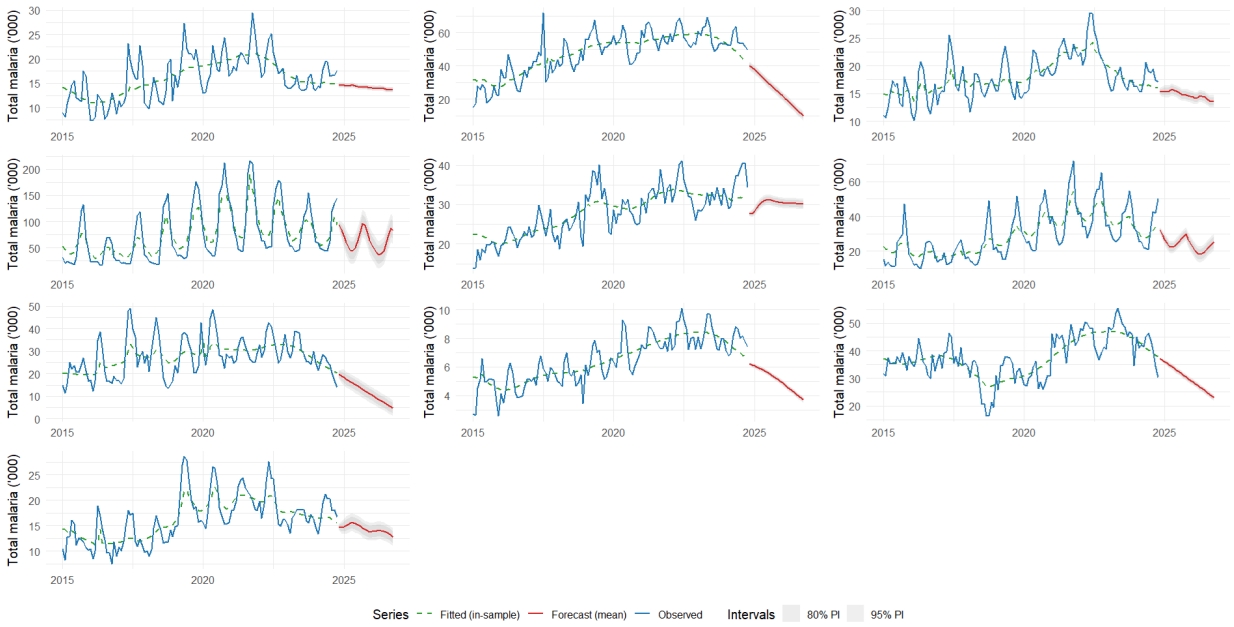


Figure 4. Regional malaria forecasts using SARIMA models with Wavelet–STL preprocessing

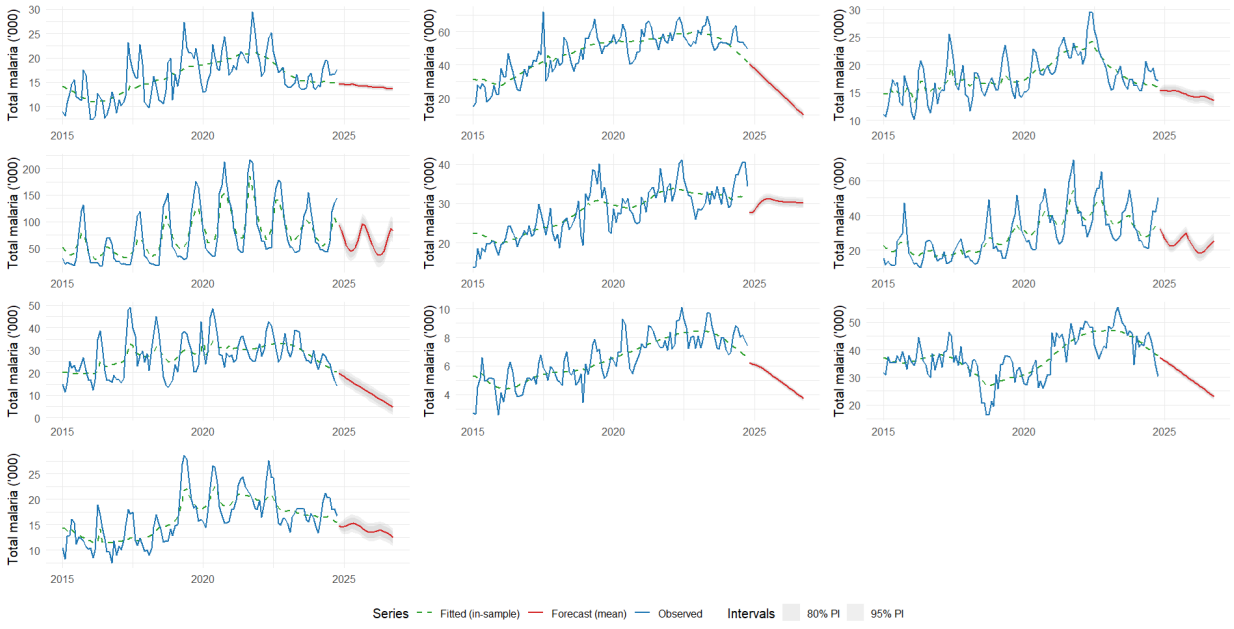
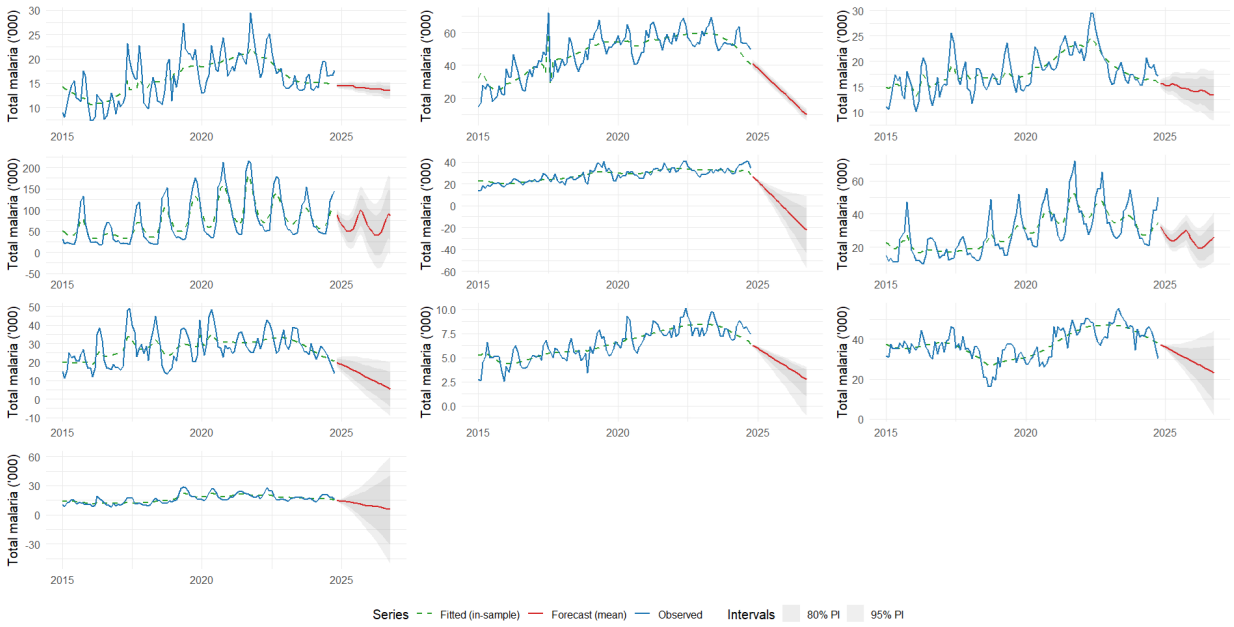


Figure 5. Regional malaria forecasts using the Wavelet–STL–BSTS model



RESULTS

The findings suggest that combining the use of wavelet denoising and STL decomposition in the forecasting system improves stability and accuracy of forecasts, which are better than the traditionally used ARIMA and SARIMA models. Isolating deterministic trend and seasonal signals makes the stochastic rest less complex, enabling more parsimonious models to give robust performance. The similar results between ARIMA and SARIMA indicate that STL is useful in removing seasonal dependencies and thus it requires no explicit seasonal differentiation.

The findings further highlight the advantages of BSTS models, which showed superior performance in settings characterized by high malaria burden and volatile transmission dynamics. This underscores the strength of Bayesian state-space methods in capturing residual uncertainty and generating credible intervals, a feature that is particularly valuable in public health forecasting.

Figures 2-4 illustrate the regional malaria predictions derived via the three Wavelet-STL pipelines. The incidence is observed (black line), 80% and 95% forecasts (red line) fitted in-sample (dashed line) and 24-month forecasts (red line) with 80% and 95% prediction intervals represented (shaded bands). In general, the plots show that Wavelet–STL preprocessing can be used to generate smooth and interpretable series within each of the ten administrative regions. The ARIMA prediction (Figure 2) is consistent with movements of relatively stable areas like the South and West, though forecast errors and broader bands are apparent in the more volatile areas including the Far North. The introduction of seasonal terms using SARIMA (Figure 3) yield almost identical trajectories, as does ARIMA; this is also in line with quantitative measures of residual seasonality of the time series after STL decomposition; there are only slight variations in a few areas, like the Centre and Far North.

Contrastingly, BSTS model was able to outperform the ARIMA and SARIMA model in most regions and had superior explanatory power and reduced forecast errors (Figure 4). In other parts, BSTS recorded higher or equal R^2 values as ARIMA and SARIMA in almost every location, with its maximum being 0.79 in the Far North and Adamawa. The best improvements were in the high-burden and volatile areas where BSTS dropped RMSE of 25.10 (ARIMA) and 24.90 (SARIMA) to 24.00 in the Far North and had a rise in R^2 of 0.77-0.77 to 0.79. On the same note, in the Northwest, BSTS raised R^2 to 0.42-0.42 under ARIMA/SARIMA to 0.49, and reduced RMSE and MAE. Performance improvement, on the contrary, was modest in more stable areas like North and Littoral, and all the models demonstrated similar fits. These findings suggest that BSTS is better able to accommodate regional heterogeneity and temporal volatility and offer a better predictive performance especially in high-burden regions.

Combined, the figures indicate that while the ARIMA and SARIMA pipelines show similar performance following a wavelet and STL preprocessing, the BSTS version has the added benefit of uncertainty-conscious forecasting, which generates plausible and adaptive forecast intervals that enhance the usefulness of the pipeline in malaria early warning and resource planning.

These results are consistent with earlier work

demonstrating the benefits of wavelet-based hybrid approaches in epidemiological time-series analysis [7, 8]. However, the present study extends this evidence base by embedding STL decomposition and systematically contrasting frequentist and Bayesian approaches. Importantly, the probabilistic forecasts produced by BSTS offer practical advantages for malaria early warning systems, where decision-makers rely not only on point forecasts but also on reliable quantification of uncertainty to inform timely interventions [1].

CONCLUSION

This work presents and confirms a multi-stage hybrid pipeline of malaria incidence forecasting in Cameroon by showing that Wavelet-STL preprocessing gave better predictive power by enhancing R^2 between 0.42 -0.77 in baseline models to 0.55 -0.79 after preprocessing, along with uniformly lowering RMSE through regions. Of the rest of the models, ARIMA and SARIMA were found to give stronger baselines, although BSTS invariably gave higher accuracy or equal accuracy and also provided calibrated probabilistic intervals. These findings highlight the significance of noise reduction, structural decomposition, and probabilistic modeling in epidemiological forecasting.

Recommendation

Methodologically, the Wavelet-STL-BSTS pipeline is our recommendation of the most suitable malaria early warning system, especially in areas characterized by unpredictable transmission patterns. In environments where computational resources are limited, the Wavelet-STL-ARIMA variant can also be considered as a viable alternative, since it is also equally accurate in less volatile areas. At the policy level, such a framework could be incorporated into national surveillance systems like DHIS2 so that health authorities can produce rolling 12–24-month projections of resource utilization.

FUNDING

This research was funded in whole or in part by Science for Africa Foundation to DEL-22-001 with support from Wellcome Trust and the UK Foreign, Commonwealth & Development Office and is part of the EDCTP 2 programme supported by the European Union. For purposes of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. Additional support was provided through the National Institutes of Health (NIH), National Institute of Allergy and Infectious Diseases (NIAID), through the International Centers of Excellence for Malaria Research (ICEMR) program under grant number U19AI181593.

REFERENCES

1. World Health Organization. *World malaria report 2023*. Geneva: WHO; 2023.
2. Snow RW, Sartorius B, Kyalo D, Maina J, Amratia P. The prevalence of *Plasmodium falciparum* in sub-Saharan Africa since 1900. *Nature* 2017;550(7677):515–8.
3. Bhatt S, Weiss DJ, Cameron E, et al. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature* 2015;526(7572):207–11.
4. World Health Organization. *Malaria early warning systems: a framework for planning and implementation*. Geneva: WHO; 2022.
5. Box GEP, Jenkins GM, Reinsel GC, Ljung GM. *Time series analysis: forecasting and control*. 5th ed. Hoboken (NJ): Wiley; 2015.
6. Findley DF, Monsell BC, Bell WR, Otto MC, Chen BC. New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. *J Bus Econ Stat* 1998;16(2):127–52.
7. Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 2003;50:159–75.
8. Percival DB, Walden AT. *Wavelet methods for time series analysis*. Cambridge: Cambridge University Press; 2000.
9. Adhikari R, Agrawal RK. An introductory study on time series modeling and forecasting. *arXiv preprint* arXiv:1302.6613; 2018.
10. Chae S, Kwon S, Lee D. Predicting infectious disease using deep learning and big data. *Int J Environ Res Public Health* 2018;15(8):1596.
11. Cleveland RB, Cleveland WS, McRae JE, Terpenning I. STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*. 1990;6(1):3–73.