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.

Artificial intelligence, statistical, mechanistic and hybrid methods for malaria forecasting and early warning: a critical comparative review of studies published in 2023–2026

Aristote Dimitri Fankam De Yaingain

aristote.deyaingain@eth-foundation.org

EquiTech Health Foundation

Systematic Review

Keywords: malaria, early warning system, forecasting, deep learning, physics-informed neural networks, Bayesian models, foundation models, sub-Saharan Africa

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11692680/v1

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

Additional Declarations: No competing interests reported.

Abstract

Malaria early warning has moved from climate-lagged regression to deep learning, time-series foundation models and physics-informed hybrids within a few years, yet recent methods have rarely been compared on common operational criteria. This narrative critical review, based on PubMed (299 records), Europe PMC preprint (164 records) and targeted searches, examines studies published from 2023 to November 2026, with a few earlier works retained where later papers build on them. Four method families (mechanistic, statistical and Bayesian, machine and deep learning, hybrid and physics-informed) are compared on seven criteria: data needs, forecast horizon, interpretability, generalization, uncertainty quantification, deployability, and equity and governance. Mechanistic models are the most interpretable but show modest skill. Bayesian spatio-temporal and structural time-series models handle uncertainty and local use best. Transformer-based models report the highest accuracy, but validation is often a single hold-out year or a random split. Hybrid physics-informed models are promising but rest on few studies, and recent evidence shows that naive neural-mechanistic hybrids can fail when only case counts are observed. The quality of routine surveillance data, notably in DHIS2, limits all families. The field needs shared benchmarks, rolling-origin and spatial validation, probabilistic scoring, data-quality auditing and explicit governance. Hybrid approaches are a priority for low-resource settings, provided they are benchmarked against established methods on identical data.

1. Background

The World Health Organization estimated about 282 million malaria cases and 610,000 deaths worldwide in 2024, of which the WHO African Region accounted for 265 million cases (94%) and 579,000 deaths (95%) [1]. Forecasts that arrive weeks to months ahead could steer vector control, chemoprevention timing and stock management, yet an operational tool of this kind is still missing in most of sub-Saharan Africa [2].

Two recent syntheses frame the field. Mategula et al. screened 6,233 records and retained 30 early-warning studies from 16 countries; time-series models dominated and the absence of standardized evaluation criteria limited any judgment of impact [3]. To our knowledge, no recent review contrasts mechanistic, statistical, machine-learning and hybrid approaches against one common grid of operational criteria, while new classes of models (Transformers, time-series foundation models, simulation-grounded networks) have appeared since 2023 [4–7].

This review therefore (i) describes the four families and their most recent representative studies, (ii) compares them on seven criteria that matter for deployment in low-resource settings, (iii) identifies weaknesses in validation and reporting, including the quality of the routine data feeding the models, and (iv) outlines where physics-informed hybrids could add value. Ethical and governance issues are treated explicitly, because predictive tools only help if the underlying data are governed fairly [8–10].

2. Review approach

This is a narrative critical review, not a PRISMA systematic review. It covers studies published from January 2023 to the search date of November 9, 2026, with a few earlier works retained as foundations [11, 12]. PubMed was searched on November 9, 2026 with a registered string combining “malaria” with forecasting, prediction and early-warning terms, restricted to 2023 onward; it returned 299 records. Titles were screened by the author with LLM assistance and 109 were flagged as potentially relevant. Fifteen flagged records not already in the reference list were prioritized, and their abstracts were read: seven were retained (section 3), three were excluded as out of scope (a news-and-perspectives article, a study predicting laboratory parameters in patients, and a sample-size tutorial), and five could not be verified from an accessible source and were left out. Europe PMC was searched on November 10, 2026 for preprints: a broad string returned 1,405 records, and a string requiring “malaria” in the title and forecasting terms in the title or abstract (2023–2026) returned 164, none already indexed as journal articles. After title screening, abstracts of 15 were read; seven were retained (section 3) and eight were set aside because the design was already represented or the task was not forecasting. OpenAlex was searched on November 10, 2026 as an open substitute for Scopus and Web of Science: a broad string returned 1,460 records and a narrowed string (malaria in the title; forecasting, prediction, early-warning or time-series terms in the title or abstract; diagnostic and imaging terms excluded) returned 514, of which 411 remained after removing records already found in PubMed or Europe PMC. A keyword filter on titles flagged 136 journal articles or preprints; abstracts of 18 were read, of which eight were retained (section 3), two were already cited, and eight were set aside because they were descriptive trend analyses, intervention evaluations, or classification of individual patients or outbreak years. The other flagged titles were not individually assessed, and the earlier targeted searches (each method family, and the reference lists of the scoping review by Mategula et al. [3]) supplemented the PubMed results. Studies were included if they had an explicit predictive component for malaria incidence, risk or transmission; image-based diagnosis is out of scope (one such paper is cited only for the general problem of domain shift [13]). Preprints are included but flagged as not peer reviewed.

Each family is judged on seven criteria:

1. Data requirements and availability in low-resource settings
2. Forecast horizon (lead time)
3. Interpretability
4. Generalization across regions and climate zones
5. Uncertainty quantification
6. Operational deployability
7. Equity, bias and data governance

Use of a large language model. An LLM (Claude, Anthropic) was used to assist with literature searching, reading of source pages, drafting and language editing of this manuscript. The author checked the cited sources and takes full responsibility for the content; the LLM is not an author.

Limitation. Scopus and Web of Science were not searched, because institutional access was unavailable; OpenAlex was used instead, and not all flagged records were screened at abstract level. The selection should be read as indicative rather than exhaustive.

3. The four method families

3.1 Mechanistic models

Mechanistic models simulate vector and parasite dynamics from climate and demography. The two weather-driven models most used in Africa are the Liverpool Malaria Model (LMM) and VECTRI, both of which have been coupled to ECMWF monthly and seasonal forecasts in prototype pan-African systems [2, 11]. Thomson et al. showed early on that seasonal multi-model climate forecasts can add several months of lead time to malaria warnings in Botswana [11]. More recently, Olaniyan et al. evaluated VECTRI driven by ECMWF System 5 forecasts at lead times of 1 to 7 months over West Africa: the model showed statistically significant probabilistic skill for identifying hyper-endemic zones, but also notable biases and low correlations, and they concluded that it could serve as an early-warning tool at seasonal scale [2].

Individual-based simulators such as OpenMalaria are mainly used for intervention planning rather than forecasting; in Benin, for example, calibrated simulations guided the geographic versus demographic extension of seasonal malaria chemoprevention, giving qualitative support to decisions rather than forecasts.

Process-based stochastic models are now being fitted to local data. Angelakis et al. built a stochastic metapopulation model of western Kenya that includes immunity, human and vector mobility, interventions and local climate; over 2008–2019 it captured small-scale heterogeneity between subpopulations and can project incidence under climate and intervention scenarios [14].

Mechanistic and statistical models are now being compared head-to-head in preprints. In Somalia, eight statistical time-series models and an SIRS model were fitted to annual incidence for 2000–2022: ARFIMA was the most accurate on a hold-out set (MAPE 7.44%), whereas the SIRS model was less accurate (MAPE 24.15%) but estimated a basic reproduction number of 1.19 and identified transmission and recovery rates as intervention targets [15]. In the North Bank Region of The Gambia, a Liverpool Malaria Model simulation, a distributed lag non-linear model and three machine-learning algorithms were combined; random forest performed best in a 24-month out-of-sample window (mean absolute error of about 85 cases), and rainfall lagged by three months was the most important predictor [16].

Critical note. Results depend strongly on structural choices. Talib et al. showed that using explicit convection changed simulated transmission differently in LMM and VECTRI, with LMM-predicted risk falling by roughly half in parts of the East African highlands while VECTRI changed little [17]. The gain in interpretability thus comes with modest skill that is difficult to calibrate.

3.2 Statistical and Bayesian models

Bayesian spatio-temporal models fitted with integrated nested Laplace approximation (INLA) dominate recent African studies. Njotto et al. linked Tanzanian mainland cases to lagged climate and bed-net coverage over 2016–2023 [18]; Armando et al. combined INLA with distributed lag non-linear models in Mozambique [19]; and a KwaZulu-Natal study used zero-inflated Bayesian models for low-transmission settings [20]. In Burkina Faso, forecast quality differed between climatic zones [21]. In Madagascar, a spatially resolved early-warning workflow with monthly automated updates was developed to serve local health actors [22]. In Mozambique, Yaari et al. applied a set of off-the-shelf statistical forecasting models to weekly district-level malaria and diarrhoeal-disease reports from 2016 to 2022; all models beat a forecast based on historical expectation up to eight weeks ahead, with comparable or better performance at 16 weeks, and a weighted ensemble was the most accurate [23]. A preprint from the Democratic Republic of Congo fitted five forecasting models (SARIMA, ETS, TBATS, a neural autoregression and ARFIMA) to DHIS2 data on severe malaria in pregnancy and found that the best model differed between provinces [24]. In Senegal, a SARIMAX model with maximum relative humidity as an exogenous variable predicted 2024 well ($R^2 = 0.69$) but failed in 2025 ($R^2 = 0.03$), a divergence the authors attributed to dual-active-ingredient nets decoupling malaria morbidity from climate [25].

Several recent national and sub-national studies give a more cautious picture. In Togo, generalized additive models using lagged satellite climate and vegetation data were fitted for 40 health districts and three target groups (training 2013–2016, test 2017); case numbers were forecast inaccurately, and the authors concluded that the models could not be used as such in malaria control [26]. In western Kenya, empirical dynamic modeling with convergent cross-mapping identified causal links between lagged climate variables, bed-net use and incidence and gave good short-term forecasts [27]. In Mozambique, a Bayesian spatio-temporal model with lagged climate covariates (2001–2018) classified high- and low-incidence seasons with an AUC of 0.897 and a lead time of up to four months [28]. In India, polynomial and ARIMA models were compared across eight states, the log-transformed polynomial regression performed best, and an interactive R Shiny tool was released for program staff [29]. In Bahir Dar, Ethiopia, a SARIMAX model with weather regressors reached a MAPE of 28.3%, yet none of the weather variables was statistically significant [30].

Classical time-series methods remain widely used. Mulenga et al. fitted a seasonal ARIMA model to six years of routine data from the Zambezi district of Zambia [31], and Akindeh et al. proposed a wavelet–STL pre-processing step combined with ARIMA, SARIMA or Bayesian structural time series for all ten regions of Cameroon; the Bayesian variant gave calibrated probabilistic forecasts, while R^2 values for the ARIMA-type models were typically between 0.49 and 0.77 [32].

Critical note. Negative results are informative: the Togo and Bahir Dar studies [26, 30] show that climate-lagged models can fail at fine spatial scales or in urban settings. These models quantify uncertainty natively and are the best validated for local use, but they are mostly explanatory, depend on stationarity assumptions, and their horizons are short (typically one to three months).

3.3 Machine learning and deep learning

Classical machine learning remains common: in northern Benin, a support vector machine was the best of the tested algorithms for climate-based outbreak warning, and projections were extended to 2050 under climate scenarios [33]. Deep learning has pushed reported accuracy further. On 23 years of daily data from Limpopo, South Africa, a Transformer reached an AUROC of about 80%, 20–40% above statistical and gradient-boosting baselines [4], and a Temporal Fusion Transformer later reported R^2 up to 0.95 on weekly malaria data [5]. Earlier work from the same group predicted outbreaks up to nine months ahead from sea-surface-temperature variability [12]. In the Peruvian Amazon, Pan et al. used a self-organizing map to derive a dynamic sea-surface-temperature index that outperformed the standard ENSO index in a generalized linear model, with lead times beyond three months [34].

Several recent country studies compare algorithms on local data. In The Gambia, eight machine-learning algorithms were compared for district-level outbreak prediction from meteorological data; extreme gradient boosting and decision trees reached the highest test accuracy (93.3%) followed by random forests (91.5%), and adding non-climatic features was important [35]. In Sierra Leone, an ANN and a SARIMA model were compared on 2018–2023 case data, and the ANN captured non-linear dynamics even without explicit climate inputs [36]. In Brazil's Legal Amazon, LSTM, GRU, SVR, random forest, XGBoost and ARIMA were compared for weekly cases; random forest most often gave the lowest errors, and K-means spatial clustering further improved accuracy [37]. In Amazonas state, an earlier preprint by the same group found ARIMA marginally better than deep-learning and machine-learning models, with clustering of municipalities as an aid to monitoring [38]. In Guinea-Bissau, XGBoost, random forest and SVM were benchmarked against SARIMA, ETS and a spline GLMM on 924 monthly observations from 11 regions; XGBoost was best (R^2 = 0.78 on the 2021 test set and 0.91 on 2022) and outperformed the conventional models [39].

Earth-observation approaches have been tested at national scale: Nazir et al. trained an LSTM-based model on 2000–2016 satellite-derived covariates and tested it on 2017 [40], and a later study applied LSTM and Transformer models to Nigerian data, where the LSTM outperformed the Transformer [41]. In Ibadan, Nigeria, a specialist hybrid combining a linear component with a recurrent network outperformed stacked hybrids on 22 years of prevalence data [42]. In south-eastern Senegal, XGBoost outperformed two other gradient-boosting variants across 13 health districts (R^2 = 0.90), with Shapley values used to rank the drivers [43]. In the Adamaoua region of Cameroon, an LSTM outperformed random forests, support vector machines and regularized regressions (accuracy 76%, RMSE = 0.08) [44]. Chen et al. addressed uncertainty with heteroscedastic quantification to "catch the spikes" in outbreak data [45]. In Ethiopia, a hybrid framework combining multi-output regression and time-series forecasting was proposed for the Amhara region [46].

Time-series foundation models are the newest entrants. In an HFMD forecasting study, TimesFM and Moirai were compared with ARIMA and LSTM, and showed good potential at the cost of greater computing resources [47]; a medRxiv preprint reported that pre-trained foundation models can be

competitive with SARIMA, Prophet and mechanistic models in data-limited epidemic settings [7]. These results concern other diseases and have not been replicated for malaria in Africa.

Critical note. Headline scores are hard to compare: single-year hold-outs, random splits, and metrics ranging from AUROC to R^2 all appear. A very high accuracy on imbalanced outbreak data says little without a confusion matrix, and studies rarely test the model outside its training region. The evidence also conflicts: machine learning beat conventional models in Guinea-Bissau [39], whereas ARIMA-type models were best in Amazonas [38] and Somalia [15], so no family can be declared superior from single-site comparisons.

3.4 Hybrid and physics-informed models

Hybrids embed epidemiological structure in a learning model. Rajnarayanan et al. introduced temperature- and altitude-dependent transmission in an SIR-SI model and estimated its parameters with ANN, RNN and PINN approaches, reporting better performance than purely data-driven estimation [48]. PINNs have been adapted to mosquito-population ODE systems, with customized normalization and loss balancing needed to handle stiff, multi-scale dynamics [49]. A neural-ODE study applied to HIV and malaria compartmental models reported that the framework learns transmission patterns while preserving the interpretability of differential equations [50]. In a medRxiv comparison for sub-Saharan Africa, a spatiotemporal Transformer gave the highest accuracy, while a neural ODE had slightly lower point accuracy but better temporal continuity and uncertainty behavior with Monte Carlo dropout [51]. Statistical–machine-learning hybrids add uncertainty: a Ghana preprint pairs Gaussian process regression with Holt-Winters smoothing under rolling-origin validation [52]. In Goa, India, hybrids that add an ARMA component to random forest, SVM or XGBoost lowered the RMSE to 0.5–15, against 13–37 for the machine-learning models alone and 5–41 for the time-series models; the machine-learning models were less accurate but more precise, whereas the time-series models were more accurate but less precise [53].

A BiLSTM trained on CMIP6 rainfall and temperature anomalies reproduced VECTRI-simulated malaria incidence for Senegal and Nigeria (R^2 between 0.93 and 0.98) and was benchmarked against six machine-learning models and checked against national surveillance data [54]; because its training target is the output of a mechanistic model, it acts as a fast emulator of that model rather than an independent forecaster. Simulation-grounded foundation models represent another hybrid route: Mantis is trained entirely on mechanistic simulations and forecasts out of the box across diseases and regions, but has been evaluated mostly on influenza and COVID-19 data [6].

Critical note. Hybridization is not automatically beneficial. Su et al. catalog failure modes of neural–mechanistic combinations when only case counts are observed and the disease dynamics are non-stationary, and show that robust performance requires making non-stationarity explicit [55]. Evidence for malaria comes largely from simulations, national aggregates or preprints, and we found no multi-country benchmark against LMM, VECTRI or Transformers on identical data.

4. Comparative assessment

No family wins on all seven criteria. The ratings in Table 1 are our qualitative synthesis of the studies in section 3, not a pooled estimate.

Table 1
Qualitative comparison of the four families.

Criterion	Mechanistic (LMM, VECTRI)	Statistical / Bayesian (INLA, GAM, BSTS, SARIMA)	ML / deep learning (SVM, XGBoost, LSTM, Transformer)	Hybrid / physics-informed (PINN, neural ODE, GP+smoothing)
Data needs	Climate, hydrology, population; case data mainly to calibrate	Case series plus climate covariates	Largest: long, high-frequency series (23 years daily in one study)	Case series plus model structure; structure may offset short series
Horizon	1–7 months with seasonal forecasts [2]	Typically 1–3 months; 8–16 weeks in Mozambique [23]	Weeks to months; 3–9 months with SST predictors [12, 34]	Up to 12 months reported in a Ghana preprint [52]
Interpretability	High	High (coefficients, lags)	Low to moderate	Moderate to high if parameters keep biological meaning
Generalization	Transferable in principle but sensitive to model structure [17]	Zone-dependent [21]	Rarely tested outside training region	Not tested at scale
Uncertainty	Ensembles, probabilistic skill scores	Native credible intervals	Rarely reported; emerging quantile/heteroscedastic methods [45]	Available in GP and Bayesian variants [52]
Deployability	Open-source codes; pilot pan-African systems	Automated monthly workflows exist [22]	Needs compute and skills; few operational cases	Research stage
Equity / governance	Independent of case-data quality but coarse	Uses local data; local scale feasible	Bias risk if trained on non-African data	Same as the data it learns from

5. Critical gaps

The main weakness of the field is less model choice than how models are evaluated, fed with data, and handed over to decision makers.

- **Metrics cannot be compared.** Studies report AIC, RMSE and R^2 [3], AUROC [4], rank probability skill scores [2], relative error reductions [40], classification accuracy [35] and RMSE or MAE on different scales [37, 53]. Proper scoring rules for probabilistic forecasts (CRPS, weighted interval score, log score) and calibration checks are almost never used. A shared protocol is missing.
- **Validation designs differ in strictness.** An Ethiopian preprint used a random 70/30 split with cross-validation [46]; the M-LSTM was tested on a single year [40]. By contrast the Ghana preprint used rolling-origin validation with Diebold–Mariano tests [52]. Random splits leak temporal information and say little about true forecasting skill; spatial hold-outs are rare.
- **Uncertainty is often missing.** Probabilistic outputs are standard in Bayesian work and emerging in deep learning [45], yet many headline results remain point forecasts.
- **Models drift when interventions change.** A model trained on historical climate–case relationships can fail when control measures alter transmission: in Senegal, a SARIMAX model's R^2 fell from 0.69 to 0.03 after the introduction of dual-active-ingredient nets [25]. Forecasts should be re-validated regularly, and intervention coverage should be a model input rather than an afterthought.
- **Routine data quality is an unexamined input.** Nearly all models learn from DHIS2-type data. Evaluations show timeliness and completeness problems: in KwaZulu-Natal only 61% of cases were notified within 24 hours [56]; in eastern Uganda concordance between DHIS2 and facility source documents fell below the WHO standard of 80% [57]; and a four-country evaluation documented discrepancies between registers, monthly forms and DHIS2 [58]. Forecast skill estimated on such data may reflect reporting behavior as much as transmission.
- **Forecasts are not tied to actions.** Mategula et al. found that the lack of standard criteria limited evaluation of early-warning impact [3]; predictive analytics must be linked to clear action pathways and national systems [59].
- **Governance and equity are under-reported.** Data privacy, informed consent and transparent governance are repeatedly flagged [10], and the WHO guidance on AI for health places ethics, human rights and equity at the center of design and deployment [8, 9]. Monitoring with disaggregated data is needed to detect drift and emerging inequities, and foundation or satellite-based inputs should be tested for bias on African data.

6. Perspectives and recommendations

Hybrid physics–AI models are a plausible route to forecasts that are both accurate and defensible, provided they are benchmarked properly. We propose six priorities.

1. **A shared benchmark.** Open case and climate datasets from several African countries, scored with rolling-origin validation, spatial hold-outs and at least one proper probabilistic score, so that LMM,

VECTRI, INLA models, Transformers, foundation models and hybrids can be compared on identical data.

2. **A mechanistic skeleton with a learned correction.** Climate-dependent transmission functions constrain the model [48], while a neural or Gaussian-process component absorbs what the equations miss and supplies uncertainty [52]. Following Su et al., non-stationarity and partial observability should be handled explicitly [55].
3. **Cautious use of foundation-model inputs.** Zero-shot and simulation-grounded models [6, 7] should be tested for bias and transfer on African malaria data before adoption.
4. **Data-quality-aware forecasting.** Model reporting delay and under-reporting, and audit routine data before training [56–58].
5. **Country-specific design.** Models should be co-developed with national malaria programs and fitted to their data and decision calendars [59].
6. **Governance from the start.** Interoperable pipelines, disaggregated monitoring, consent and alignment with continental frameworks [8–10].

7. Conclusions

Recent malaria forecasting has produced credible models in each of four families, but none is proven operationally superior because studies use different data, metrics and validation schemes, and because the routine data they rely on are imperfect. Bayesian models offer the most reliable uncertainty for local use, deep learning the highest reported accuracy, and mechanistic models the clearest biology. Hybrid physics-informed approaches could combine these strengths, but only after common benchmarks, probabilistic scoring, data-quality safeguards and governance are in place. These are the conditions under which AI can support malaria elimination equitably.

Abbreviations

AIC
Akaike information criterion
ANN
artificial neural network
ARIMA
autoregressive integrated moving average
AUROC
area under the receiver operating characteristic curve
BSTS
Bayesian structural time series
CRPS
continuous ranked probability score
DHIS2
District Health Information Software 2

ECMWF
European Center for Medium-Range Weather Forecasts
GAM
generalized additive model
GP
Gaussian process
HFMD
hand, foot and mouth disease
INLA
integrated nested Laplace approximation
LLM
large language model
LMM
Liverpool Malaria Model
LSTM
long short-term memory
ODE
ordinary differential equation
PINN
physics-informed neural network
RMSE
root mean squared error
RNN
recurrent neural network
SARIMA
seasonal ARIMA
SST
sea-surface temperature
STL
seasonal-trend decomposition using LOESS
TFT
Temporal Fusion Transformer
VECTRI
VECtor-borne disease community model of ICTP, TRIeste
WHO
World Health Organization.

Declarations

Ethics approval and consent to participate.

Not applicable.

Consent for publication.

Not applicable.

Competing interests.

ADFY declares no competing interests.

Funding.

This work received no specific funding.

Author Contribution

A.D.F.Y. conceived the review, selected and analyzed the literature, and wrote the manuscript.

Acknowledgments.

Not applicable.

Authors' information. ADFY is affiliated with the Faculty of Science, University of Dschang, and the EquiTech Health Foundation, Cameroon.

Availability of data and materials.

Not applicable. No new datasets were generated; all sources are cited in the reference list.

References

1. World Health Organization. World malaria report 2025: addressing the threat of antimalarial drug resistance. Geneva: World Health Organization. 2025. ISBN 978-92-4-011782-2. Licence: CC BY-NC-SA 3.0 IGO. https://iris.who.int/(accessed November 2, 2026).
2. Olaniyan EA, Tompkins AM, Caminade C. Predicting malaria hyper endemic zones in West Africa using a regional scale dynamical malaria model. Front Trop Dis. 2024;5:1322502. 10.3389/fitd.2024.1322502.
3. Mategula D, Gichuki J, Barnes KI, Giorgi E, Terlouw DJ. Advancing early warning systems for malaria: progress, challenges, and future directions, a scoping review. PLOS Glob Public Health.

2025;5(5):e0003751. 10.1371/journal.pgph.0003751.

4. Pillay MT, Minakawa N, Kim Y, Kgalane N, Ratnam JV, Behera SK, et al. Utilizing a novel high-resolution malaria dataset for climate-informed predictions with a deep learning transformer model. Sci Rep. 2023. 10.1038/s41598-023-50176-3.
5. Pillay MT, Le MTQ, Takamatsu Y, Phong TV, Kgalane N, Minakawa N. Application of a Temporal Fusion Transformer and long-term climate and disease data to assess the predictive power and understand the drivers for malaria and dengue. Int J Environ Res Public Health. 2026;23(1):75. 10.3390/ijerph23010075.
6. Dudley C, Magdaleno R, Harding C, Sharma A, Martin E, Eisenberg M. Mantis: a simulation-grounded foundation model for disease forecasting. arXiv:2508.12260. 2025 (preprint).
7. Kalahasti S, Faucher B, Wang B, Ascione C, Carbajal R, Enault M, Cassis CV, Launay T, Guerrisi C, Boëlle PY, Baldo F, Valdano E. Foundation time series models for forecasting and policy evaluation in infectious disease epidemics. medRxiv. 2025. 10.1101/2025.02.24.25322795(preprint).
8. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: WHO. 2021. ISBN 9789240029200. https://who.int/publications/i/item/9789240029200(accessed November 9, 2026).
9. Goodman KW, Litewka S, Malpani R, Pujari S, Reis AA. Global health and big data: the WHO's artificial intelligence guidance. S Afr J Sci. 2023;119(5/6). 10.17159/sajs.2023/14725.
10. Obeagu EI, Ogenyi FC. Artificial intelligence in African malaria control programs: opportunities and risks. 2025. 10.1097/MS9.0000000000004259
11. Thomson MC, Doblas-Reyes FJ, Mason SJ, et al. Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. Nature. 2006;439:576–9. 10.1038/nature04503 . (historical reference).
12. Martineau P, Behera SK, Nonaka M, Jayanthi R, Ikeda T, Minakawa N, et al. Predicting malaria outbreaks from sea surface temperature variability up to 9 months ahead in Limpopo, South Africa, using machine learning. Front Public Health. 2022. 10.3389/fpubh.2022.962377.
13. Kabore KK, Guel D. Addressing challenges in data quality and model generalization for malaria detection. arXiv:2501.00464. 2024 (preprint; image-based diagnosis, cited only for domain shift).
14. Angelakis A, Beloconi A, Nyawanda BO, Khagayi S, Kariuki S, Munga S, et al. Assessing the impact of climate and control interventions on spatio-temporal malaria dynamics using a stochastic metapopulation model. PLoS Comput Biol. 2026;22(3):e1014004. 10.1371/journal.pcbi.1014004.
15. Yusuf IM. Modeling and projecting malaria incidence in Somalia: a dual-approach framework combining statistical forecast accuracy with mechanistic epidemiological insight. Research Square. 2026. 10.21203/rs.3.rs-8670933/v1(preprint).
16. Okafor UA. Development of a process-based and machine learning early warning system for malaria transmission dynamics under multi-scenario climate change: a subnational assessment of the North Bank Region of The Gambia. Research Square. 2026. 10.21203/rs.3.rs-10987061/v1(preprint).

17. Talib J, Abatan AA, Hoek Spaans R, et al. The effect of explicit convection on simulated malaria transmission across Africa. PLoS ONE. 2024;19(4):e0297744.
18. Njotto LL, Senyoni W, Cronie O, et al. Bayesian spatio-temporal modeling and prediction of malaria cases in Tanzania mainland (2016–2023). Int J Health Geogr. 2025;24. 10.1186/s12942-025-00408-8.
19. Armando CJ, Rocklöv J, Sidat M, et al. Climate variability, socio-economic conditions and vulnerability to malaria infections in Mozambique 2016–2018. Front Public Health. 2023. 10.3389/fpubh.2023.1162535.
20. Ebhuoma O, Gebreslasie M, Villena OC, Arab A. Environmental and geographical factors influence malaria transmission in KwaZulu-Natal province, South Africa. Geospat Health. 2025;20(1):1370. 10.4081/gh.2025.1370.
21. Traoré N, Millogo O, Sié A, Vounatsou P. Impact of climate variability and interventions on malaria incidence and forecasting in Burkina Faso. Int J Environ Res Public Health. 2024;21(11):1487. 10.3390/ijerph21111487.
22. Evans MV, Ihantamalala FA, Randriamihaja M et al. Increasing the resolution of malaria early warning systems for use by local health actors. medRxiv. 2024. 10.1101/2024.04.17.24305948(preprint).
23. Yaari R, Galanti M, Zepeda-Tello R, Chicumbe S, Jani I, Cassy A, et al. Infectious disease forecasting to support public health: use of readily available methods to predict malaria and diarrhoeal diseases in Mozambique. J Glob Health. 2025;15:04114. 10.7189/jogh.15.04114.
24. Akilimali A, Onesime J, Muse AH, Hassan MA, Nadarajah S, Aizeque A et al. Spatiotemporal dynamics and forecasting of severe malaria incidence among pregnant women in the Democratic Republic of Congo (2020–2024): a retrospective observational study. Research Square. 2026. 10.21203/rs.3.rs-8707200/v1(preprint).
25. Aw Y, Kane M, Morlighem C, Tine EG, Ly AY, Nourdine MA, et al. Integrating climate variability and Anopheles gambiae s.l. bionomics into a SARIMAX model for malaria forecasting in Keur Socé, Senegal. Open Res Afr. 2026. 10.12688/openresafrica.16883.1(. version 1, not yet peer reviewed).
26. Thomas A, Bakaï TA, Atcha-Oubou T, Tchadjobo T, Rabilloud M, Voirin N. Exploring malaria prediction models in Togo: a time series forecasting by health district and target group. BMJ Open. 2024. 10.1136/bmjopen-2022-066547.
27. Nyawanda BO, Kariuki S, Khagayi S, Bigogo G, Danquah I, Munga S, et al. Forecasting malaria dynamics based on causal relations between control interventions, climatic factors, and disease incidence in western Kenya. J Glob Health. 2024. 10.7189/jogh.14.04208.
28. Armando CJ, Rocklöv J, Sidat M, Tozan Y, Mavume AF, Sewe MO. Spatio-temporal modelling and prediction of malaria incidence in Mozambique using climatic indicators from 2001 to 2018. Sci Rep. 2025. 10.1038/s41598-025-97072-6.
29. Ghosh SK, Ananthakumar U, Chougale PD, Somaraj AB. Forecasting malaria in Indian states: a time series approach with R Shiny integration. Malar J. 2025. 10.1186/s12936-025-05526-z.

30. Gelaw TT, Abera MA. Forecasting malaria incidence in a resource-limited urban setting with climate variables as exogenous regressors: time series analysis using a SARIMAX model in Bahir Dar, Ethiopia. Malar J. 2026. 10.1186/s12936-026-05823-1.
31. Mulenga EM, Spangenberg ED, Kashila J. Predicting malaria trends in Zambezi district of Zambia using ARIMA modeling techniques: a 2019–2024 time series analysis. J Appl Math. 2026;2026(1):1040598.
32. Akindeh MN, Adegoke T, Sheetal SP, Innocent A, Andrillene LDW, Madeleine PF, et al. Towards reliable malaria forecasting: a noise-resilient wavelet-STL hybrid framework for Cameroon. Epidemiol Biostat Public Health. 2026;21(1). 10.54103/2282-0930/30118.
33. Gbaguidi GJ, Topanou N, Leal Filho W, Ketoh GK. Towards an intelligent malaria outbreak warning model based intelligent malaria outbreak warning in the northern part of Benin, West Africa. BMC Public Health. 2024;24:450. 10.1186/s12889-024-17847-w.
34. Pan M, Hu S, Janko MM, Zaitchik BF, Takahashi K, Lescano AG, et al. A machine learning-based dynamic SST index for long-lead malaria prediction in the Peruvian Amazon. GeoHealth. 2026;10(1). e2025GH001529.
35. Khan O, Ajadi JO, Hossain MP. Predicting malaria outbreak in The Gambia using machine learning techniques. PLoS ONE. 2024;19(5):e0299386. 10.1371/journal.pone.0299386.
36. Jalloh SW, Malenje B, Imboga H, Hodges MH. Forecasting malaria cases using climate variability in Sierra Leone. Malar J. 2025. 10.1186/s12936-025-05389-4.
37. Monteiro KHC, Rocha ES, Morais LA, Santos EG, Neto SRS, Sampaio V, et al. Integrating machine learning and spatial clustering for malaria case prediction in Brazil's Legal Amazon. BMC Infect Dis. 2025;25. 10.1186/s12879-025-11193-x.
38. Barboza MFX, Monteiro KHC, Dourado RA, Sampaio VS, Endo PT. Enhancing malaria prediction in Amazonas, Brazil: a comparative analysis of deep learning models based on clustering. Research Square. 2023. 10.21203/rs.3.rs-3338998/v1(preprint).
39. Nanque MC, Allaya MM, Loum MA, Dum VN, Nante JE, Gning LD et al. Predicting malaria incidence using meteorologicnvironmental, and spatial factors in Guinea-Bissau: a machine learning approach. Research Square. 2026. 10.21203/rs.3.rs-10972587/v1(preprint).
40. Nazir U, Quddoos MT, Uppal M, Khalid S. Predicting malaria outbreaks using earth observation measurements and spatiotemporal deep learning modelling: a South Asian case study from 2000 to 2017. Lancet Planet Health. 2024. 10.1016/S2542-5196(24)00082-2.
41. Sami W, Abbas F, Sabir M et al. Data-driven prediction of malaria outbreaks in Nigeria. In: IGARSS 2025 – IEEE International Geoscience and Remote Sensing Symposium. 2025:1636–1640. 10.1109/IGARSS55030.2025.11244048
42. Fisher T, Rojas-Galeano S, Fernandez-Reyes D. Specialist hybrid models with asymmetric training for malaria prevalence prediction. Front Public Health. 2023;11:1207624. 10.3389/fpubh.2023.1207624.

43. Ly AY, Allaya MM, Loum MA, Gning LD, Sy O, Bakhoum LN, et al. Forecasting malaria incidence in the Southeast districts of Senegal using a machine learning approach. BMC Artif Intell. 2025. 10.1186/s44398-025-00007-4.
44. Naroum E, Maka Maka E, Abboubakar H, Dayang P, Bamana AB, Garga B, et al. Comparative analysis of deep learning and machine learning techniques for forecasting new malaria cases in Cameroon's Adamaoua region. Intell Based Med. 2025;12:100220. 10.1016/j.ibmed.2025.100220.
45. Chen F, Qi Q, Qiu J, Zhang K, Li X. Catching the spikes: heteroscedastic uncertainty quantification for enhanced malaria prediction. NeurIPS 2024 Workshops (AIM-FM).
46. Azezew K, Tesema A, Mekuria B, Kassie A, Embiale A, Salau AO, Asresa T. Hybrid predictive modeling of malaria incidence in the Amhara Region, Ethiopia: integrating multi-output regression and time-series forecasting. arXiv:2510.01302. 2025 (preprint).
47. Wang Y, Huang G, Chen C, Li Q, Xu X. A comparative study of time series foundation models for hand, foot, and mouth disease forecasting: TimesFM, Moirai, and traditional approaches. Front Public Health. 2025;13:1634138. 10.3389/fpubh.2025.1634138.
48. Rajnarayanan A, Kumar M, Tridane A. Analysis of a mathematical model for malaria using data-driven approach. Sci Rep. 2025;15:27272. 10.1038/s41598-025-12078-4.
49. Viet Cuong D, Lalić B, Petrić M, Thanh Binh N, Roantree M. Adapting physics-informed neural networks to improve ODE optimization in mosquito population dynamics. PLoS ONE. 2024;19(12):e0315762.
50. Ndubuogaranya FC, Ehigie JO, Okunuga SA. Neural ordinary differential equations for parameter estimation, simulation and forecasting of infectious disease dynamics: applications to HIV/AIDS and malaria models. Ife J Sci. 2026;28(2):261–79. 10.4314/ijs.v28i2.2.
51. Srivastava V, Kumar M, Vaidya S, Tridane A. Advanced spatio-temporal neural networks for malaria prevalence forecasting in Sub-Saharan Africa. medRxiv. 2025. 10.1101/2025.11.12.25340054(preprint).
52. Ansah-Narh T, Asare Afrane Y, Bremang Tandoh J. Hybrid probabilistic forecasting of under-five malaria admissions in Ghana: a Gaussian process regression with Holt-Winters smoothing. arXiv:2606.00834. 2026 (preprint; reported as accepted in Artificial Intelligence in Medicine).
53. Hussain SSA, Bedi S, Yadav CP, Mohanty AK, Mahatme K, Tyagi S, et al. Hybrid models combining trend and seasonality components with machine learning algorithms provide accurate forecasting of malaria incidence. PLOS Glob Public Health. 2025;5(10):e0004500. 10.1371/journal.pgph.0004500.
54. Diouf I, Dieng A, Barry MB, Talla C, Oyebiyi O, Babalola AS, et al. Impact of climate change on malaria incidence in Senegal and Nigeria: a machine learning approach. Front Clim. 2026. 10.3389/fclim.2026.1877760.
55. Su Y, Lee R, Cui J, Ramakrishnan N. How (not) to hybridize neural and mechanistic models for epidemiological forecasting. arXiv:2602.06323. 2026 (preprint).

56. Mabona M, Zwane T, Raman J, Kuonza L, Mhlongo B, Phafane P. Evaluation of the malaria case surveillance system in KwaZulu-Natal Province, South Africa, 2022: a focus on DHIS2. Malar J. 2024;23:47. 10.1186/s12936-024-04873-7.
57. Nuwasiima S, Mpimbaza A, Muteebwa L, Nagawa E, Arinaitwe E, Kiberu F, et al. Concordance of data on key malaria indicators between DHIS2 and source documents, and influencing factors at public primary health facilities in eastern Uganda: a mixed methods study. Malar J. 2025;24(1):270. 10.1186/s12936-025-05519-y.
58. Konaté-Touré A, Ngufor C, Mpimbaza A, Atobatele S, Akpiroroh E, Ssewante N, et al. Concordance of rapid diagnostic test results between health facility registers and health management information systems: a multi-country evaluation. Malar J. 2025;25(1):70. 10.1186/s12936-025-05753-4.
59. de Hoop T, Shu E, de Milliano M, Seidenfeld D. Designing national government systems to integrate malaria early warning systems with precision. American Institutes for Research;, March. 2026. https://www.air.org/sites/default/files/2026-03/Malaria-Early-Warning-Systems-Brief-508-March-2026.pdf (accessed November 2, 2026).