Leveraging Earth Observation Data for Malaria Prediction in Nigeria

W. Sami[1]*, M. Sabir[1]*, F. Abbas[1]*, M. T. Quddoos[2], and U. Nazir[1,3]

[1] School of Computer and IT, Beaconhouse National University, Pakistan,
{f2021-500, f2021-277, f2021-502, usman.nazir}@bnu.edu.pk
[2] Lahore University of Management Sciences
muhammad.quddoos@lums.edu.pk
[3] Planetary Health Informatics Lab, University of Oxford
usman.nazir@ndorms.ox.ac.uk

Abstract. The study focuses on using environmental indicators to predict infectious diseases, specifically malaria in Nigeria, a heavily affected country. The research combines machine learning algorithms with earth observation data, categorizing the data into four main types: malaria incidence rates from 2000 to 2017, environmental variables, nighttime lights data, and distance from water bodies data. Two machine learning models, Long Short-Term Memory (LSTM) and Transformer, were trained and tested. Results showed that the LSTM model was the most efficient, providing more precise and accurate predictions for malaria incidence rates. The Transformer model showed improved performance after increasing the number of epochs, but the LSTM consistently delivered efficient results. The findings suggest that environmental variables play a significant role in the spread of malaria. It also highlights the importance of selecting the appropriate model for optimal results. This methodology can be applied beyond Nigeria to examine various health problems in different geographical regions. The paper showcases how government policies can utilize artificial intelligence to address health hazards and promote a more secure and healthy future for communities.

Keywords: Malaria Prediction · Machine Learning · Climate Change · Environmental Factors · Spatio-Temporal Analysis · Malaria Incidence Rate · Healthcare Policy · Disease Prevention.

1 Impact Statement

1.1 Climate Impact

The climatic factors, in particular, temperature and precipitation, have a significant impact on malaria parasite's biological activities, and geographical distribution. The research on how the climate change could impact the malaria incidences reveal an overall increase of risk as more regions are becoming prone to climatic conditions suitable for malaria transmission. In the regions where this disease is less common, the incidence rates are prone to climate change, that may result in a huge number of loss of life, especially in the regions of South Asia, South America, and parts of Africa [1]. An emerging societal crisis could be caused by the expected rise in infectious disease prevalence and a global spread due to climate change.This problem has been highlighted, especially with respect to malaria and other diseases that are vector-transmitted [2]. Malaria parasite's incubation cycle is highly affected by its surrounding's temperature, making temperature a main determinant of outbreak of malaria.

To achieve effective allocation of disease control resources, and to estimate the effect of climate change on global outbreak of malaria, Epidemiological models are becoming more and more common. Spread of malaria due to temperature is mostly incorporated into these models by using average of monthly temperatures, however there are fluctuations all through the diurnal period [3]. With the help

* These authors contributed equally to this work

of satellite sensor data, it is possible to develop early-warning systems for malaria like diseases, that is causing 1 to 2 million deaths per annum, around the world [4]. Furthermore, out of the few medical problems, malaria has been modelled by several research groups, making it feasible to carry out first inter-comparison for affects on health with respect to climate change in future [5]. However, the changes in the pattern of outbreaks of malaria must be predicted, considering the local environmental factors, to install malaria control programs effectively [1].

1.2 Economic Impact

Malaria and poverty are significantly related [6]. The affect of malaria on the health sector and on the economic sector is similar. Malaria is said to be an illness of poverty as well as a reason of poverty in the modern day Africa. In comparison, the economic growth per year of a country without malaria is higher than a country that is free from the burden of malaria. In some parts of Africa, economists claim that malaria is a cause of 'growth penalty' of approximately 1.3 percent per year. If this impact is calculated over several years, a significant affect over GDP can be observed in countries with malaria, compared to the countries without malaria, highly limiting the economic growth of that region. Some countries are reported to have a total public health expenditure of up to 40 percent dedicated to malaria only, 30 to 50 percent of patients being admitted, and up to 50 percent of follow-up patients [7].

The Enugu State study further emphasizes on the adverse economic effects of malaria, which primarily impacts the low-income population, affecting homes, medical sector, and overall growth. The issue puts a direct as well as indirect cost on households, with limited healthcare budget, resulting in lower worker efficiency due to sickness and absence from work [8]. To reduce the adverse affects of malaria on economy, it is essential for the economic development policy makers to consider methods for prevention and control of this disease [9].

2 Introduction

Malaria has been a major threat to public health in West Africa, with an estimated reports of 68 million cases with 194000 deaths only in Nigeria by 2021. Nigeria contributes to Global burden of Malaria with 27% of the worldwide cases [10]. According to Annals of Tropical Medicine & Parasitology's report, out of 7713 patients diagnosed with malaria, 4239 were found to be having pure Plasmodium falciparum malaria [11]. Malaria is contracted by us by bite of a female Anopheles mosquito, that is the primary cause of this disease. Growth and development of these mosquito's is dependent on many environmental factors [12]. Thus, several environmental factors contribute to the spread of malaria such as degree of rainfall contributing to the increased humidity in the region, temperature and distance of spatial location from water bodies, nightlight and the extent of vegetation index in the region are known to be significant predictors [13]. Hence, applying spatial analysis of risk elements like climate factors that act as an aid in transmission is crucial [14]. Some studies conducted in West African countries using statical model have shown a positive correlation among rainfall, temperature, and malaria incidence rates, while in other areas there is a positive association. Hence, it is crucial to understand the level of impact of climate change over diverse areas for malaria incidence [12]. Current systems are considered insufficient in accuracy to be integrated with clinical decisions. Despite all the work done on the prediction of this disease, Nigeria remains one of the most affected country by this disease [15].

As the technology is making its way through the medical domain in recent times, the potential of using predictive models is becoming significant [16]. Machine Learning has made it possible to identify useful patterns and make data meaningful. Mapping the climate factors using the spatio-temporal datasets with malaria incidence data from past years, predictions in these spatial locations can be made for the future outbreaks of malaria. This integration of environment variables with past data has the ability to significantly improve the prediction system of this disease, contributing proactively to the health sector against a rapidly changing climate. This integration of predictive model with the healthcare domain, particularly for the diseases like malaria, will be beneficial for the early warning and preventive systems in Nigeria. Incorporation of environmental factors with the malaria incidence rate will not only increase the accuracy of results, but it will also help identify the most contributing factor towards the outbreak of malaria, allowing the health care agencies to allocate resources effectively, and reduce the impact of this outbreak.

2.1 Contribution

Our major contributions to mapping health equities in South Asia, as compared to our previous research article [17], we transitioned from traditional AI training models to a newer technology of transformers due to its improved performance with the complex tasks. Furthermore, we extended our feature set to include a new feature of distance from water bodies of a particular spatial location to more comprehensively look at the environmental factors influencing our investigation.

Different datasets were merged, including Nighttime lights and distance of a spatial location from water bodies, that were drawn out directly from satellite imagery using Earth Engine to create a custom dataset. This dataset was then integrated with publicly available datasets of different environmental indicators like rainfall, temperature etc, and the incidence rate of malaria, again aligned with the particular spatial locations and time stamps. This method helped us achieve a total of 43 features against each spatial location and each year starting from 2000 till 2017. Consequently, this final and detailed dataset led to improved model performance, and it will later be helpful to a larger research community.

3 Related Work

Malaria is a major cause of deaths and illness in many parts of Africa [18]. There are no dominant researches on climate-based prediction systems for malaria in sub-Saharan African countries, hence we chose Nigeria as the starting point for training model for predicting outbreak of malaria in the region. However, some previously built models for this regard can be found.

"An auto-regressive integrated moving average (ARIMA) and seasonal ARIMA (SARIMA) are two time-series models that have been previously used in malaria prediction" (Nkiruka, Prasad, and Clement, 2021). They work by first training the data, then anticipating the future time points. The auto-regressive (AR) section depicts regression on delayed parameters of interest. An immediate value is displayed on the moving part for the linear combinations of the error terms. Each feature included is responsible in making the model a better-fit and more accurate [19].

Adeola et al. investigated the malaria incidence and environmental factors in the South African region of Nkomazi.They calculated and evaluated monthly environmental "variables such as the normalised difference vegetation index (NDVI), the enhanced vegetation index (EVI), the normalised difference water index (NDWI), the land surface temperature for night (LSTN) and day (LSTD), and rainfall using seasonal autoregressive integrated moving average (SARIMA) models" (Adeola et al., 2019), with several lag time periods. Their 'SARIMA' model predicted the likely outbreaks of malaria for the upcoming three months by figuring out that the major contributor to this outbreak was the rainfall feature [20].

In another research, ARIMA model was applied to the malaria incidence data of Afghanistan [21]. It was found out that peak levels of malaria occurred in the months from July to September while they were minimum in the month of July in Sub-Saharan countries. It suggested that rainfall and humidity were the dominant contributors towards this incidence of malaria. The ARIMA model proved to be good for dealing with seasonal data features, however, it is not very efficient when it comes to the weather variables .

Olusola et al. introduced another statical model, used a combination of time-series models, and dynamic regression models to determine the co-relation between malaria incidence and the climate factors in Azure. The results suggested that as the temperature increases, so does the probability of malaria outbreak in that region, leading to a rise in both incoming patients and outpatients of malaria [22].

Another statistical model for predicting malaria is the "stochastic lattice-based malaria (SLIM)" approach, which was proposed in Kenya, and uses the expected climate change to predict the fluctuations in amount of Anopheles vector, the life circle of Plasmodium parasites, and malaria spread. The research analysed the effect of environmental variability on the pattern of malaria spread by investigating differences in the plasmodium's sporogonic cycle provoked by the rise in surrounding temperature, along with difference in the "gonotrophic cycle of Anopheles vector" that result from changes made by vegetation to elevated (CO2) [23].

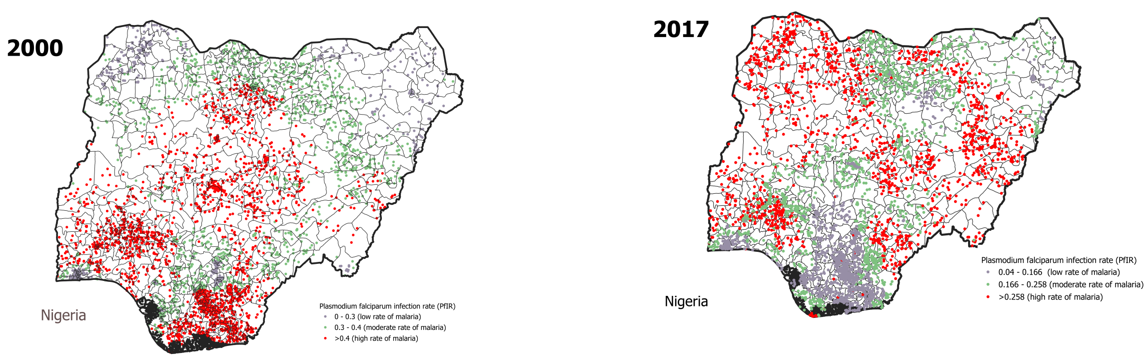
The approach presented by this study that predicts malaria incidences on account of 44 features including climatic variables using the AI model "LSTM". No latest research has used this method to predict the burden of this disease. This study has played its role by identifying unknown environmental factors that are causing a high incidence of malaria in Nigeria. Subsequently, the insights from this model will aid in comprehending the variables involved in outbreak of malaria in the selected area for the upcoming year [19].

4 Methodology

In this section, the methodology used from beginning to end of this investigation has been briefed. As we moved from the initial step of data sourcing till the final step of visualisation, the details of all major steps are highlighted below:

4.1 Data Gathering

1. **Malaria Incidence Rates:** Malaria Incidence rates for Nigeria included the rate of plasmodium falciparum at 4232 spatial locations of Nigeria. The datasets are time series, covering the the time intervals from 2000 to 2017. This data was acquired from Demographic and Health Survey (DHS) given by the US agency for International growth



(a) Plasmodium Falciparum rate for year 2000

(b) Plasmodium Falciparum rate for year 2017

Fig. 1: Malaria Incidence rates for Nigeria

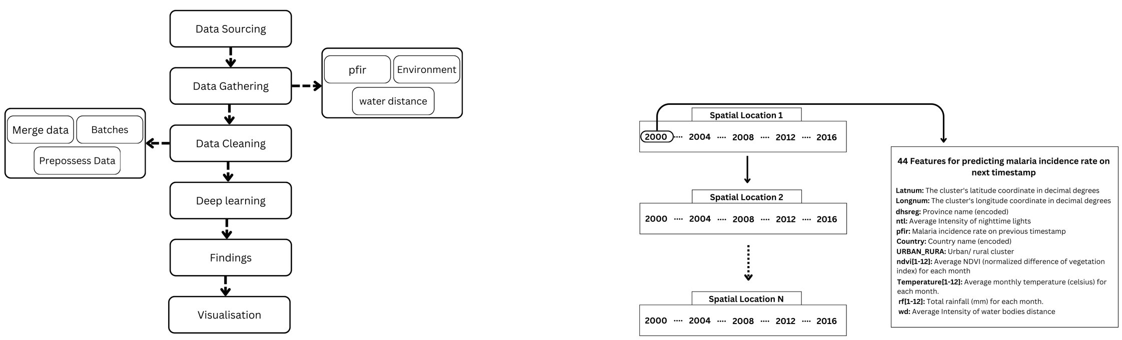
2. **Environmental Indicators:**
From 2000 to 2017, we gathered data on environmental factors contributing to malaria outbreak against each of the 4232 spatial locations. These datasets include temperature (in degrees Celsius), rainfall (in millimeters), and the average Normalized Vegetation Index (NDVI) against each month in these years. These datasets are acquired from International Food Policy Research Institute's (IFPRI) 2020 Advancing Research on Nutrition and Agriculture (AReNA) project, which is mentioned on their official website, https://www.ifpri.org/.

3. **Nighttime Lights Data:**

   Data regarding nighttime lights was obtained from two satellites using the earth engine. The first satellite named DMSP OLS (Defense Meteorological Program Operational Linescan System) was used to gather data from year 2000 t0 2012, and the second satellite The second satellite, VIIRS (Visible Infrared Imaging Ra- diometer Suite) Nighttime Day/Night Band Composites Version 1, was used to get data for the years 2012 to 2017.

4. **Waterbodies Distance Data:**

   Data concerning distance of water bodies from spatial locations was drawn using geo-spatial analysis using data from earth observations. The computation process included the processing of satellite imagery at a 30 metres nominal scale using the JRC/GSW14/YearlyHistory dataset. Water bodies that were larger in size were recognized, and their distance from our 4232 spatial locations was measured.



(a) (b)

Fig. 2: (a): Methodology (b): Data is collected over time at a specific spatial point, with each time step (2000, 2001, ... 2016) including 44 unique features at that particular location.

4.2 Modeling Process

LSTM A four-layered M-LSTM model consisting of 100 LSTM units per layer with a fully linked layer of linear regression at its end was used for predicting the malaria incidence rates. Past data of malaria incidence rates and time-series data of different environmental variables were sourced as part of the modelling strategy. Using the concepts of feature engineering, a final dataset was created containing time-series sequence of environmental data, along with malaria incidence rates. To achieve promising results, three splits were made on data into testing, validation and training sets. Implemented this four layered LSTM (long short-term memory) to identify and capture the temporal patterns and training it in order to reduce the prediction errors, constituting the basic methodology. model parameters were optimised by tuning the hyper parameters, and generalisation was evaluated as part of validation process. Testing part was able to yield an unbiased metrics of performance.

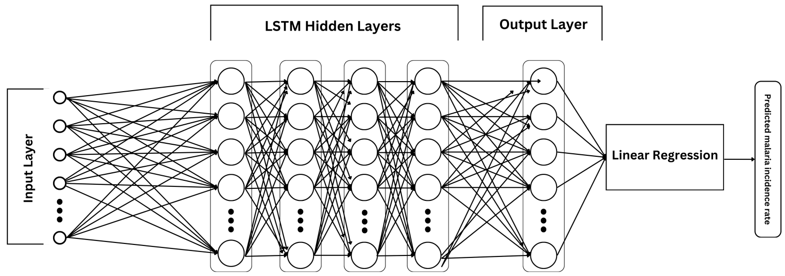


Fig. 3: The M-LSTM model comprises four hidden layers, each containing 100 LSTM units. Subsequently, a fully connected layer is employed, followed by linear regression to estimate the malaria incidence rate for the year 2017 through spatio-temporal splitting and for a random year through spatial splitting.

Transformer A multi-layered Transformer model with 512 embedding dimensions, 8 attention heads, and 6 encoder layers was utilized for predicting malaria incidence rates. A linear regression layer was added to the model to refine predictions. Data preparation included smoothly combining time series of environmental data with malaria rates, ensuring thorough training, validation, and testing sets. A structured process of adjusting hyper-parameters was carried out to improve model parameters, increasing flexibility and reliability in forecasts. Validation thoroughly evaluated the model's ability to perform well in various environmental conditions, while impartial testing verified its consistency in providing precise predictions.

4.3 Data Pre-processing

Data preprocessing can be divided into the following steps:

1. **Data Cleaning and Merging**
   Data cleaning and merging is a crucial and critical initial step in machine learning, especially when datasets are complex and larger in size. Process of data cleaning involved identifying errors, null values and inconsistencies in the datasets and rectifying them. It also involves correcting any anomalies in the data. after the process of data cleaning has been completed, datasets from difference sources are merged to create a final and comprehensive dataset. This step also involved aligning the data, as our datasets were time-series and involved data of same spatial locations. This effective data cleaning and merging was the building blocks for an improved and accurate predictive modeling.
2. **Temporal Batching**
   The step of the temporal batching involved splitting the data into consistent time periods so that the model can identify patterns over the specific intervals of time. This approach helps in recording the delayed effected caused by climatic variables on the outbreak of malaria. For instance, rainfall data can be divided into 12 months for each year to better foresee the rainfall's effect on the prevalence of malaria. This method improves the predictive capability of the model by understanding the past climate variable's impact on the future outbreak of this disease
3. **Training and Testing Split**
   2000-2016 data was used for training and 2017 data was used to test the accuracy of the model. This split is highly important to judge the model's accuracy. Model evaluation is done by using the testing set, issues such as over-fitting are identified where the model does not work accurately on the new data. Properly splitting the data ensures the accuracy of model in predicting the future outbreaks of malaria in Nigeria

5 Qualitative evaluation

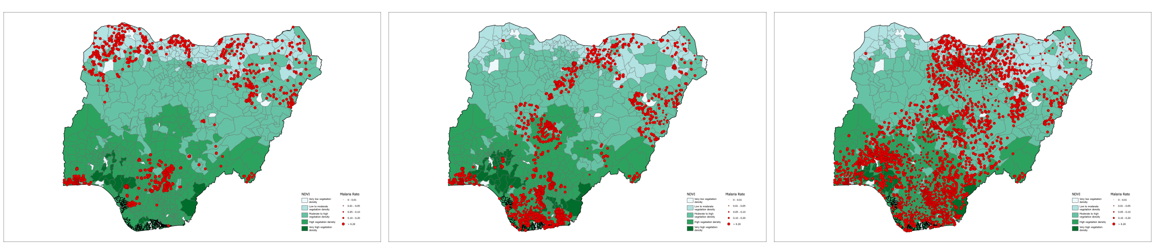
The qualitative analysis for this research helped open various point of views and contextual details related to impact of environmental variables on the outbreak of malaria in the region of Nigeria. We identified patterns that were recurring in the data, especially during the step of visualisation, which provided the insights into the common factors that were present on the incidence point of malaria. The qualitative investigation not only complimented the quantitative evaluation, but it also helped us understand the problem and its contributing factors to a deeper level. The integration of these qualitative findings with quantitative data helped us get a holistic view of this research. Visualisations of malaria incidence rate over the region of Nigeria in 2017 were a major help in evaluating and further understanding the conditions over the affected areas

6 Environmental Factors and Malaria Incidence Rate

Understanding how factors like NDVI, rainfall, temperature, nightlight, and proximity to water bodies affect malaria transmission patterns is crucial. This knowledge helps us develop predictive models and carry out targeted interventions to reduce the spread of malaria.

6.1 Normalized Difference Vegetation Index (NDVI)

The Normalized Difference Vegetation Index (NDVI) is important for understanding the rate of Plasmodium falciparum, the most deadly malaria parasite. High NDVI values indicate dense, healthy vegetation, which provides optimal breeding and resting habitats for malaria vectors. These vectors thrive in vegetated areas with standing water, leading to higher malaria transmission rates. Effective vegetation management, such as draining standing water, clearing dense vegetation, and using larvicides, is essential for controlling vector populations and reducing the incidence of P. falciparum. Different types of vegetation, including dense forests, wetlands, agricultural fields, and urban green spaces, impact malaria transmission by providing breeding and resting sites for vectors.

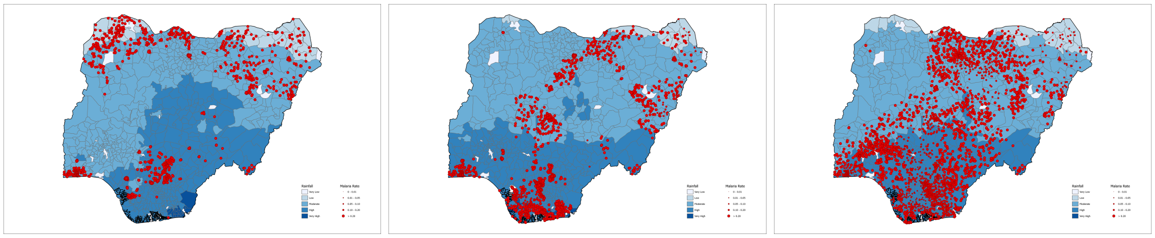


(a) 2001 (b) 2010 (c) 2017

Fig. 4: Comparison of NDVI and DHS malaria incidence rates in Nigeria: Low NDVI suggest reduced probability of encountering malaria at district level, while high NDVI are associated with increased risk of the disease.

6.2 Rainfall

The amount of rain we get has a big impact on how many mosquitoes there are and that affects how likely malaria is to spread. When we have enough rain, it creates standing water where mosquitoes like to breed. Places with lots of rain end up with loads of these breeding spots, leading to a rapid rise in mosquito numbers and increasing the risk of malaria. The timing of rainy periods also affects when malaria is most likely to be spread. It's really important to manage water and drainage in these areas to reduce mosquito breeding and control malaria.

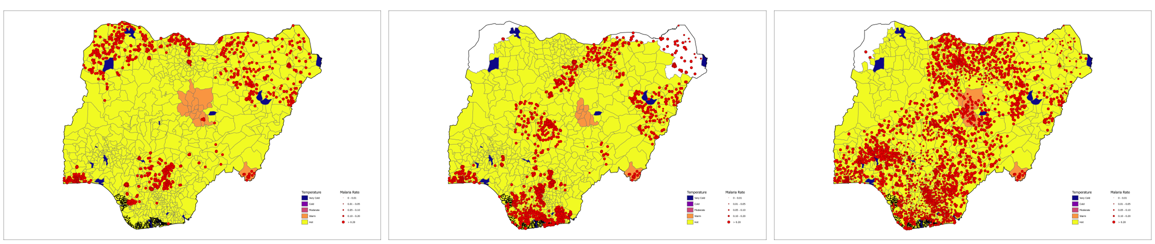


(a) 2001 (b) 2010 (c) 2017

Fig. 5: Comparison of rainfall and DHS malaria incidence rates in Nigeria: Low rainfall suggest reduced probability of encountering malaria at district level, while high rainfall are associated with increased risk of the disease.

6.3 Temperature

The temperature has a big impact on the lives of both the Anopheles mosquito and the Plasmodium parasite, which causes malaria. The best temperature for mosquito growth and parasite maturation is usually between 20-30°C. Warmer temperatures can make mosquitoes grow faster, bite more often, and speed up the development of the malaria parasite inside them. This means that places with higher temperatures often have more cases of malaria. But if it gets too hot, it can actually be harmful to the mosquitoes. It's really important to keep an eye on temperature changes so we can predict when malaria might become a problem and take action in time.

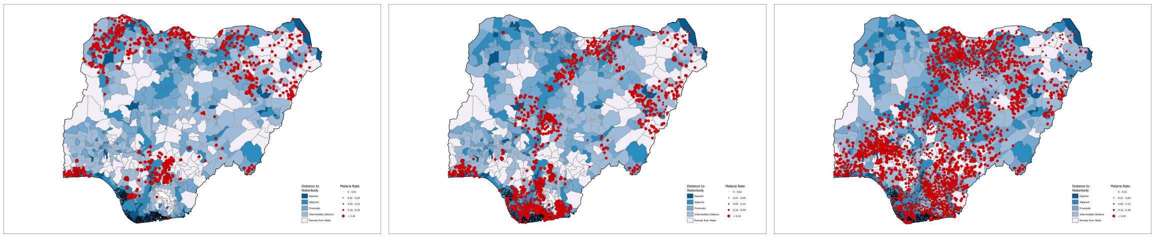


(a) 2001 (b) 2010 (c) 2017

Fig. 6: Comparison of temperature and DHS malaria incidence rates in Nigeria: Low temperatures suggest reduced probability of encountering malaria at district level, while high temperatures are associated with increased risk of the disease.

6.4 Water Bodies Distance

Living near water bodies such as rivers, lakes, and ponds can have a big impact on the spread of malaria. These places are where Plasmodium falciparum parasite, and they can cause higher rates of malaria in nearby areas. The closer people live to water bodies, the more likely they are to encounter mosquitoes, leading to more cases of malaria. On the other hand, areas far from water bodies tend to have fewer mosquitoes and lower rates of malaria. It's important to manage water effectively and make changes to the environment to reduce mosquito breeding near where people live in order to control malaria in these areas.

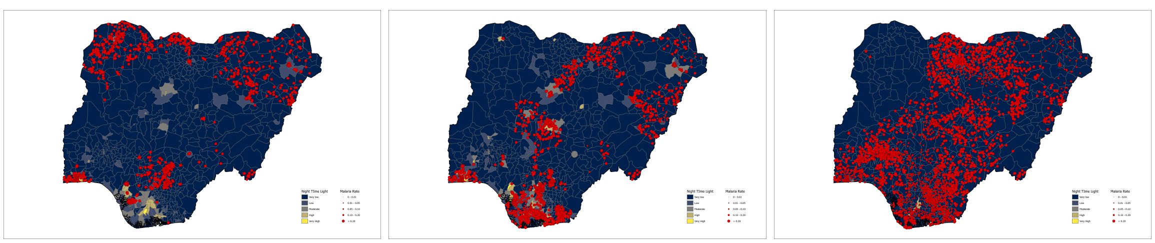


(a) 2001 (b) 2010 (c) 2017

Fig. 7: Comparison of water bodies distance and DHS malaria incidence rates in Nigeria: more water bodies distance suggest reduced probability of encountering malaria at district level, while less water bodies distance are associated with increased risk of the disease

6.5 Nightlight

The level of nighttime light is often used as a measure of human activity and urban development. Studies have found that areas with bright nighttime lights tend to have more cases of malaria. This is because well-lit areas usually indicate bustling and crowded neighborhoods, providing plenty of spots for mosquitoes to breed due to human activities and ample hosts for their blood meals. In urban areas with poor housing and sanitation, mosquitoes find even better conditions for breeding and transmitting malaria. Moreover, the more people move around and gather in brightly lit areas, the more malaria can spread. As a result, regions with intense nighttime lights may see higher rates of malaria. This calls for specific steps to control mosquito populations and public health interventions.



(a) 2001 (b) 2010 (c) 2017

Fig. 8: Comparison of nightlight and DHS malaria incidence rates in Nigeria: Low nightlight suggest reduced probability of encountering malaria at district level, while high nightlight are associated with increased risk of the disease.

7 Quantitative evaluation

Two models were utilized in forecasting malaria incidence rates: Long Short-Term Memory (LSTM) and Transformer. Key metrics such as Root Mean Square Error (RMSE), and R-squared (R^2) were calculated to provide a comprehensive assessment of each model's predictive capabilities.

7.1 Data Splitting Approaches

The training set was exercised to examine malaria incidence rates from 2000 to 2016, contrarily the testing set was used for doping out the anticipated malaria incidence rate in 2017. This procedure capacitated us to figure out how vigorously the models execute in diverse terrestrial and spatial conditions.

7.2 Model Performance Results

The table lends a review on how each model executed while predicting the malaria incidence rate in Nigeria.

Country	Model	Prediction-year	MSE	RMSE
Nigeria	M-LSTM (proposed)	2017	0.0003	0.01
Nigeria	Transformer [24]	2017	0.003	0.05

Table 1: Experimental Findings: Mean squared error (MSE) and root mean squared error (RMSE) were calculated for two models: LSTM and transformer . The performance was evauluated by testing these models on a dataset of spatial locations of Nigeria.

7.3 Performance Metrics

Mean squared error (MSE) and Root mean squared error (RMSE) are broadly use as a performance metrics to analyse the certainty of the predictions in machine learning models. These metrics are mainly accessible in regression tasks, where the intents are to predict continuous values. Analysing, how both MSE and RMSE are appropriate for various models such as LSTM and Transformer as follow:

MSE: The Mean Square Error (MSE) is a measurement that computes the average of the squared variances between the predicted and observed values.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{1}$$

RMSE: The root mean square error (RMSE) measures the differences between predicted values from a model and the actual observed values.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2}$$

7.4 Bivariate Maps

In the qualitative assessment, bivariate maps comparing true and predicted malaria incidence rates are depicted in the Fig. 9. In the case of the proposed model, the true and predicted values exhibit a consistent correlation. For instance, in areas where the true malaria incidence rate is high, the predicted malaria incidence rate also tends to be high, as observed in regions with dark blue coloring in the bivariate map.

(a) M-LSTM (Proposed) (b) Transformer

Fig. 9: Predicted versus true malaria incidence rates for each district of Nigeria. These findings are based on the analysis of the country-specific test dataset.

7.5 Analysis of Results

The performance of two models was scrutinized in forecasting the malaria incidence rate. The LSTM model's veracity enhanced as the number of training span raised. Starting off with 2 epochs, the Mean Squared Error (MSE) was 0.002 with a Root Mean Squared Error (RMSE) of 0.04. After 100 epochs, the MSE decreased to 0.0003, and the RMSE reduced to 0.01, proposing a better fit with more training. Conversely, the Transformer model initially had an MSE of 0.004 and an RMSE of 0.06 at 2 epochs. Having an increase up to 100 epochs, the MSE improved to 0.003 and the RMSE to 0.05. Results from transformer model had higher accuracy even on fewer epochs, however, overall it did not surpass the LSTM model. Lastly, the LSTM model achieved the most satisfactory performance, with an MSE of 0.0003 and an RMSE of 0.01 at 100 epochs, demonstrating excellent predictive accuracy for the malaria incidence rate.

7.6 Results Insights

The thorough assessment using the spatio-temporal analysis highlights the advantages and drawbacks of each model. Although the Transformer model is recognized for its reliable performance and computational efficiency, the LSTM model demonstrated the best predictive accuracy for the assigned task. The LSTM model displayed notable enhancements with additional training epochs.These findings emphasize the significance of model selection and tuning in achieving optimal prediction outcomes for malaria incidence rates.

8 Conclusion

In this paper, we introduced a spatio-temporal deep learning approach to predict malaria outbreaks by integrating earth observation measurements with advanced machine learning techniques. We used Nigeria as a case study where we predicted annual district-level malaria incidence rates using historical data from previous years. The data gathered, for this research, is categorized into four categories: 1) Malaria Incidence Rate, which is a time series dataset, covering the time interval from 2000 to 2017. This is collected from Demographic and Health Survey conducted by the US agency for the International growth. 2) Environmental Indicator includes data gathered on environmental factors responsible for the malaria outbreak. This data is obtained from International Food Policy Research Institute's official website, covering the time span from 2000 to 2017. 3) Nighttime Lights Data is acquired from two satellites using the earth engine. 4) Waterbodies Distance Data is extracted through geospatial analysis using data from earth observation. Using the data sources mentioned above, we executed three models,

namely LSTM, and Transformer, to make pertinent predictions. However, the results obtained from the LSTM exhibited more predictive accuracy and precision.
The scope and the utility of this work are not limited to the disease and geography discussed in this paper. This can be utilized for analysing different health hazards in different geographical contexts, and for making informed decisions to ward off societal crises that pose existential risk to the humanity. Most importantly, this study is a perfect embodiment of the interplay of AI and governance to usher the human beings toward a sustainable and safer future.

References

1. W. Martens, L. W. Niessen, J. Rotmans, T. H. Jetten, and A. J. McMichael, "Potential impact of global climate change on malaria risk.," *Environmental health perspectives*, vol. 103, no. 5, pp. 458–464, 1995.
2. K. D. Lafferty, "The ecology of climate change and infectious diseases," *Ecology*, vol. 90, no. 4, pp. 888–900, 2009.
3. K. P. Paaijmans, A. F. Read, and M. B. Thomas, "Understanding the link between malaria risk and climate," *Proceedings of the National Academy of Sciences*, vol. 106, no. 33, pp. 13844–13849, 2009.
4. D. J. Rogers, S. E. Randolph, R. W. Snow, and S. I. Hay, "Satellite imagery in the study and forecast of malaria," *Nature*, vol. 415, no. 6872, pp. 710–715, 2002.
5. C. Caminade, S. Kovats, J. Rocklov, A. M. Tompkins, A. P. Morse, F. J. Colón-González, H. Stenlund, P. Martens, and S. J. Lloyd, "Impact of climate change on global malaria distribution," *Proceedings of the National Academy of Sciences*, vol. 111, no. 9, pp. 3286–3291, 2014.
6. J. L. Gallup and J. D. Sachs, "The economic burden of malaria,"
7. R. B. Malaria, "Economic costs of malaria," *World Health Organization*, vol. 20, no. 1, 2011.
8. O. Onwujekwe, N. Uguru, E. Etiaba, I. Chikezie, B. Uzochukwu, and A. Adjagba, "The economic burden of malaria on households and the health system in enugu state southeast nigeria," *PloS one*, vol. 8, no. 11, p. e78362, 2013.
9. S. C. Datta and J. J. Reimer, "Malaria and economic development," *Review of Development Economics*, vol. 17, no. 1, pp. 1–15, 2013.
10. World Health Organization, "Report on malaria in nigeria 2022," 2022. Accessed: 2024-05-28.
11. L. A. Salako, F. O. Ajayi, A. Sowunmi, and O. Walker, "Malaria in nigeria: A revisit," *Annals of Tropical Medicine & Parasitology*, vol. 84, no. 5, pp. 435–445, 1990. PMID: 2256767.
12. O. Nkiruka, R. Prasad, and O. Clement, "Prediction of malaria incidence using climate variability and machine learning," *Informatics in Medicine Unlocked*, vol. 22, p. 100508, 2021.
13. H. Hasyim, A. Nursafingi, U. Haque, D. Montag, D. A. Groneberg, M. Dhimal, U. Kuch, and R. Müller, "Spatial modelling of malaria cases associated with environmental factors in south sumatra, indonesia," *Malaria Journal*, vol. 17, no. 1, p. 87, 2018.
14. O. A. Okunlola and O. T. Oyeyemi, "Spatio-temporal analysis of association between incidence of malaria and environmental predictors of malaria transmission in nigeria," *Scientific Reports*, vol. 9, no. 1, p. 17500, 2019.
15. B. J. Brown, P. Manescu, A. A. Przybylski, F. Caccioli, G. Oyinloye, M. Elmi, M. J. Shaw, V. Pawar, R. Claveau, J. Shawe-Taylor, M. A. Srinivasan, N. K. Afolabi, G. Rees, A. E. Orimadegun, W. A. Ajetunmobi, F. Akinkunmi, O. Kowobari, K. Osinusi, F. O. Akinbami, S. Omokhodion, W. A. Shokunbi, I. Lagunju, O. Sodeinde, and D. Fernandez-Reyes, "Data-driven malaria prevalence prediction in large densely populated urban holoendemic sub-saharan west africa," *Scientific Reports*, vol. 10, p. 15918, Sept. 2020.
16. A. Oguntimilehin, A. Adetunmbi, and A. O.B, "A review of predictive models on diagnosis and treatment of malaria fever," *International Journal of Computer Science and Mobile Computing*, vol. 45, pp. 1087–1093, 04 2015.
17. U. Nazir, A. Ejaz, M. Quddoos, M. Uppal, and S. Khalid, "Towards a spatio-temporal deep learning approach to predict malaria outbreaks using earth observation measurements in south asia,"
18. E. Mbunge, R. C. Millham, M. N. Sibiya, and S. Takavarasha, "Application of machine learning models to predict malaria using malaria cases and environmental risk factors," in *2022 Conference on Information Communications Technology and Society (ICTAS)*, pp. 1–5, IEEE, 2022.
19. O. Nkiruka, R. Prasad, and O. Clement, "Prediction of malaria incidence using climate variability and machine learning," *Informatics in medicine Unlocked*, vol. 22, p. 100508, 2021.

20. A. M. Adeola, J. O. Botai, J. M. Olwoch, H. C. d. W. Rautenbach, O. M. Adisa, C. De Jager, C. M. Botai, and M. Aaron, "Predicting malaria cases using remotely sensed environmental variables in nkomazi, south africa," *Geospatial Health*, vol. 14, no. 1, 2019.
21. M. Y. Anwar, J. A. Lewnard, S. Parikh, and V. E. Pitzer, "Time series analysis of malaria in afghanistan: using arima models to predict future trends in incidence," *Malaria journal*, vol. 15, pp. 1–10, 2016.
22. O. S. Makinde, G. J. Abiodun, and O. T. Ojo, "Modelling of malaria incidence in akure, nigeria: negative binomial approach," *GeoJournal*, vol. 86, no. 3, pp. 1327–1336, 2021.
23. P. V. Le, P. Kumar, M. O. Ruiz, C. Mbogo, and E. J. Muturi, "Predicting the direct and indirect impacts of climate change on malaria in coastal kenya," *PloS one*, vol. 14, no. 2, p. e0211258, 2019.
24. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need.(nips), 2017," *arXiv preprint arXiv:1706.03762*, vol. 10, p. S0140525X16001837, 2017.