Spatio-Temporal Analysis of Malaria Cases in Ghana:

A Data-Driven Approach to Intervention Evaluation and Resource Allocation

Report Submitted By:
Prince Oppong Boakye
MSc. Geoinformatics and Spatial Data Science
IFGI, University of Münster
Matriculation Number: 554352

Institut für Geoinformatik
Universität Münster

1.0 Introduction

Assessing the spatio-temporal patterns of malaria incidences in Ghana will allow assessing the effectiveness of intervention strategies such as insecticide-treated nets and indoor spray. This project will assist practitioners and public health professionals to make data-driven interventions toward malaria.

1.1 Goals

- **Identify high-risk areas:** Map malaria hotspots over time to guide resource prioritization.
- **Measure intervention change:** Compare changes in malaria prevalence before and after interventions.
- **Offer real-time insights:** Allow users to map trends, correlations, and spatial relationships on an interactive dashboard.

1.2 Personas – Target User Groups

- **Dr. Grace Mensah (Public Health Specialist)**
  - Needs current and accurate data for appropriate resource allocation.
  - Needs tools to show temporal variation trends and intervention impacts on malaria.
- **Abena Osei (Community Health Worker):**
  - Requires mobile-friendly tools to identify high-risk areas and notify communities.
  - Requires alerts about real-time population data and places where there is an outbreak of malaria.

1.3 User Stories

1.3.1 Identifying High-Risk Areas

As Dr. Grace Mensah,
I would like to map malaria hotspots on a map for the past five years so that I can prioritize the allocation of resources to the most affected areas.

1.3.2 Monitoring Intervention Effectiveness

As Dr. Grace Mensah,
I need to know the malaria incidence rates before and after distribution of bed nets so that I may be able to assess the performance of the intervention.

1.3.3 Real-Time Risk Alerts

As Abena Osei,
I would also like real-time alerts sent to me when locations become more at risk for malaria, so I can do community outreach there first.

1.3.4 Understanding Seasonal Pattern

Being Abena Osei,
I would like to understand how rainfall and temperature correlate with malaria outbreaks to educate the community about preventive measures before the rainy season.

2.0 Design Rationale

The project design of the app was obtained through the requirements of user groups and what tasks they are trying to perform. Important design principles include:

2.1 Visualization Design

- **Temporal Trends**: Line graphs were selected to depict the time-trend changes in malaria incidence and treatment rates to allow users to see the effect of their interventions.
- **Spatial Hotspots**: Interactive choropleth maps were implemented to visualize malaria incidence by region and allow users to know the areas at the highest risk.
- **Comparison Dashboard**: Bar charts and faceted maps were used to illustrate malaria incidence before and after interventions so users could see the differentiating effect of interventions.

2.2 Interaction Design

- **Real-Time Exploration**: The Shiny dashboard allows users to explore the data interactively by filtering for individual years, measures, and regions.
- **Dynamic Alerts**: Dynamic alerts are another aspect built into the design-but only for future versions of the app-which would enable the user to be informed of emerging hotspots in real-time.
- **User-Centric Design**: The dashboard is designed with easy-to-learn controls and easily accessible visuals that should enable non-computer literate users to navigate the dashboard with good usability.

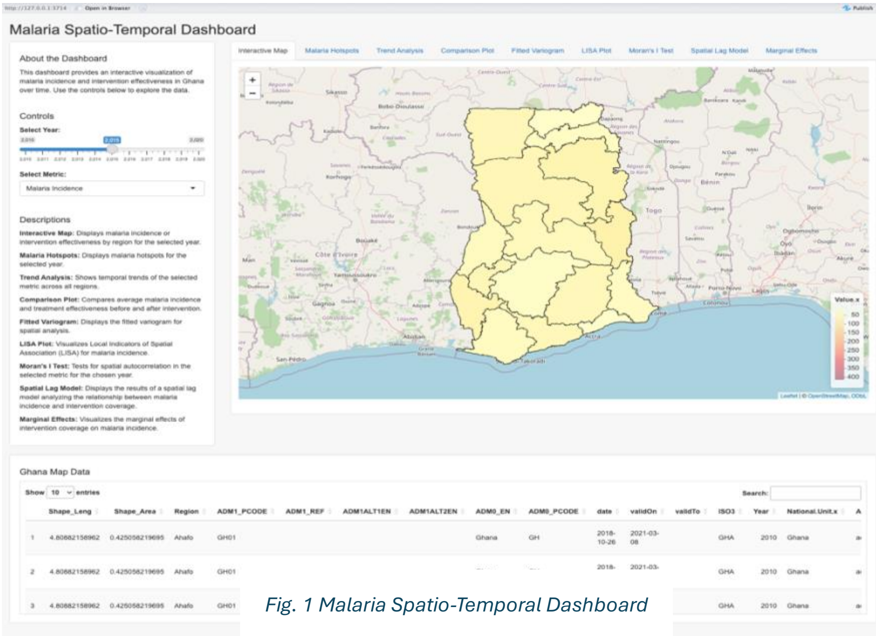


Fig. 1 Malaria Spatio-Temporal Dashboard

2.3 Alignment with Persona Needs

- **Dr. Grace Mensah**: Dashboard containing insights on trends, evaluation of interventions, and effective resource allocation.
- **Abena Osei**: Interactive maps and real-time insights into community outreach and educational campaigns.

3.0 Project Description

The project offers an interactive dashboard to analyze spatio-temporal trends with respect to malaria incidence and the effectiveness of interventions in Ghana. Data is acquired through different sources covering incidence, intervention coverage, and treatment effectiveness.

3.1 Functionality

- **Interactive Map**: The map shows the malaria incidence and intervention effectiveness for a selected region-wise year.
- **Temporal Trends**: Generate time series plots of malaria incidence and treatment rates over varying time intervals.
- **Comparison Dashboard**: The bar chart with faceted maps allows for comparison of malaria incidence before and after intervention.
- **Spatial Analysis**: Tools for hotspot identification, spatial autocorrelation (Moran's I), and variogram analysis.
- **Data Table**: Provides easy and sortable access to region and time-year malaria data.

3.2 Planned vs. Realized Features

Feature	Planned	Realized
Interactive Map	Yes	Yes
Temporal Trends	Yes	Yes
Comparison Dashboard	Yes	Yes
Spatial Autocorrelation	Yes	Yes
Real-Time Alerts	Yes	No
Mobile-Friendly Interface	Yes	No

Table 1 Planned vs Realized Features

3.3 Git Repository

The code and data for this project can be found in this Github repository:
https://github.com/iprincegh/Spatio-temporal-Analysis-of-Data-in-R/tree/main

3.4 Docker Image

The application is available as a Docker image (Click Here).Run the following commands to access the dashboard locally:

```
docker pull iprince/malariaanalysisgh-dashboard
docker run -d --rm -p 3838:3838 iprince/malariaanalysisgh-dashboard
```

4.0 Analysis Processes

The proposed step-by-step analytical procedure started at data gathering and preparation and ended up in spatial modeling. Here is a more elaborate discussion of individual steps involved:

4.1. Data Preparation and Cleaning

- **Data Sources:**
  - Subnational unit-level malaria incidence data from the Malaria Atlas Project.
  - Intervention data on bed net coverage and use and indoor residual spraying.
  - Treatment data: Metrics on effective treatment rates.
  - Ghana shapefile: Used GIS data of Ghana administrative regions.
- **Data Cleaning:**
  - Removing extra spaces and renaming columns for uniformity.
  - Dealing with missing values mostly by filtering incomplete records out.
  - Aggregated malaria incidence by region and year for hotspot analysis.
- **Data Merging:**
  - Stacking malaria incidence, intervention, and treatment data for a single dataset
  - Merging the combined dataset with ghana shapefile for spatio-analysis.

4.2 Exploratory Analysis

- **Descriptive Statistics:**
  - Described malaria incidence and treatment data, which provided an understanding of the overall trends.
- **Temporal Trends:**
  - Line graphs represents changes in time to malaria incidence and treatment rates.
  - Interactive visualization using ggplot2 and plotly.

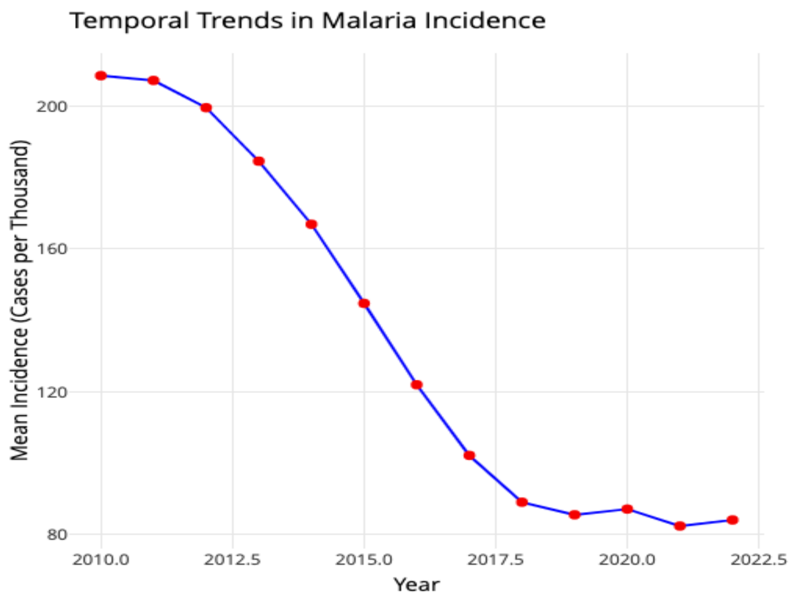


Fig.2 Temporal Trends

- **Spatial Trends:**
  - To visualize malaria incidence by region, choropleth maps were created.
  - Interactive maps created with tmap and leaflet.

4.3 Hotspot Identification

- **Aggregation:**
  - Average malaria incidence by region and year was calculated.
- **Visualization:**
  - Interactive maps showing on malaria hotspots on select years (2015 used in the project).

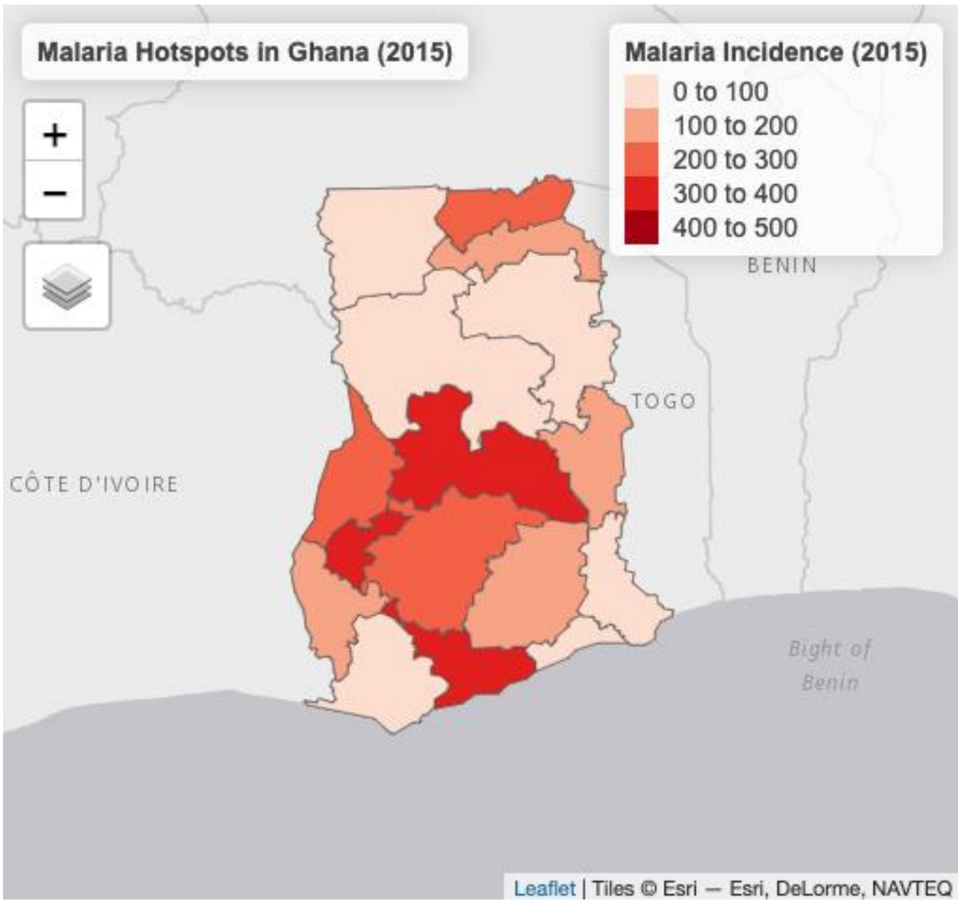


Fig.3 Malaria Hotspots in Ghana, Year = 2015

4.4 Impact of Interventions

- **Comparison of Periods:**
  - Calculated average malaria incidence before and after the intervention year (2012).
  - Observed changes on treatment coverage and malaria incidence.

- **Visualization:**
  - Bar charts with a faceted map comparison of intervention periods.

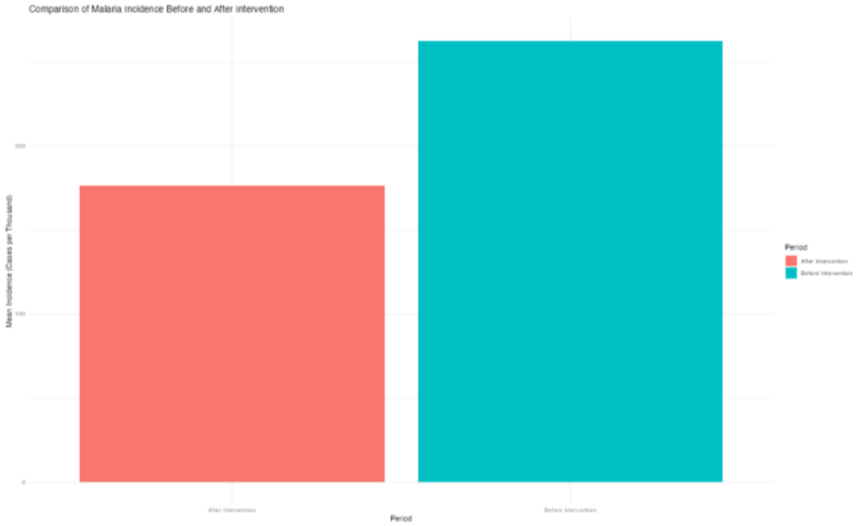


Fig.4 Comparison of Malaria Incidence Bar Chart

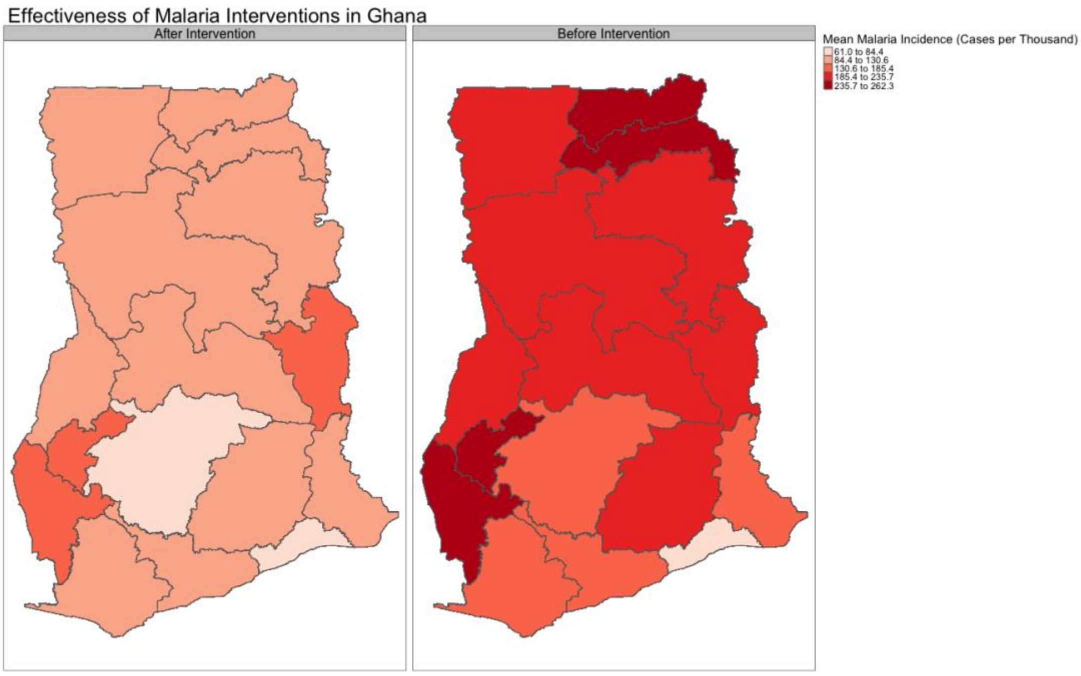


Fig.5 Comparison of Malaria Incidence faceted Map

4.5 Spatial Correlation and Modeling

- **Spatial Autocorrelation:**
  - Moran's I test for global spatial dependencies.
  - Local Moran's I analysis in looking for spatial clusters with high/low incidences.

```
        Moran I test under randomisation

data:  values
weights: lw

Moran I statistic standard deviate = 0.36271, p-value = 0.3584
alternative hypothesis: greater
sample estimates:
Moran I statistic       Expectation          Variance
    -7.812767e-04     -3.484321e-03      5.553676e-05
```

Fig.6 Moran I Test Under Randomization

- **Variogram Analysis:**
  - Fitted a variogram model for spatial dependence analysis.

```
> # Fit variogram model
> model_initial <- vgm(psill = 8000, model = "Exp", range = 200, nugget = 1500)
> vm_fit <- fit.variogram(v.m, model = model_initial, fit.method = 6)
> print(vm_fit)
  model     psill    range
1   Nug 14973.929   0.0000
2   Exp  2932.188 176.7804
```

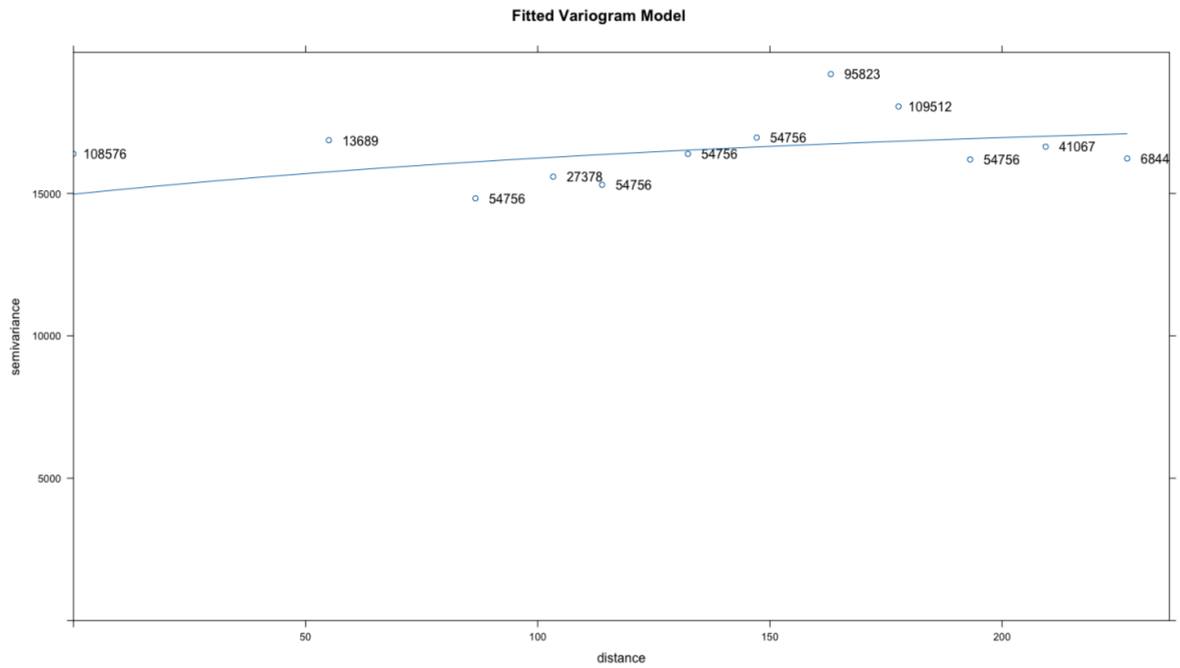
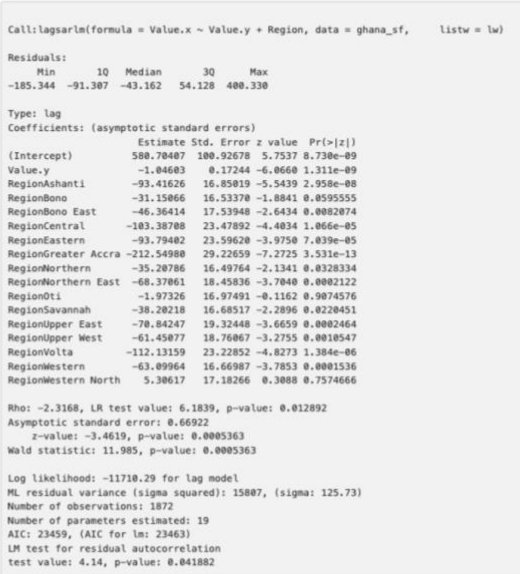


Fig.7 Fitted Variogram Model

- **Spatial Lag Model:**
    - How other regions influence the malaria incidence.
    - Analyzed residuals and marginal effects for interventions.



```
Call:lagsarlm(formula = Value.x ~ Value.y + Region, data = ghana_sf,     listw = lw)

Residuals:
     Min       1Q   Median       3Q      Max
-185.344  -91.307  -43.162   54.128  400.330

Type: lag
Coefficients: (asymptotic standard errors)
                       Estimate Std. Error z value  Pr(>|z|)
(Intercept)           580.70407  100.92678  5.7537 8.730e-09
Value.y                -1.04603    0.17244 -6.0660 1.311e-09
RegionAshanti         -93.41626   16.85019 -5.5439 2.958e-08
RegionBono            -31.15066   16.53370 -1.8841 0.0595555
RegionBono East       -46.36414   17.53948 -2.6434 0.0082074
RegionCentral        -103.38708   23.47892 -4.4034 1.066e-05
RegionEastern         -93.79402   23.59620 -3.9750 7.039e-05
RegionGreater Accra -212.54980   29.22659 -7.2725 3.531e-13
RegionNorthern        -35.20786   16.49764 -2.1341 0.0328334
RegionNorthern East   -68.37061   18.45836 -3.7040 0.0002122
RegionOti              -1.97326   16.97491 -0.1162 0.9074576
RegionSavannah        -38.20218   16.68517 -2.2896 0.0220451
RegionUpper East      -70.84247   19.32448 -3.6659 0.0002464
RegionUpper West      -61.45077   18.76067 -3.2755 0.0010547
RegionVolta          -112.13159   23.22852 -4.8273 1.384e-06
RegionWestern         -63.09964   16.66987 -3.7853 0.0001536
RegionWestern North     5.30617   17.18266  0.3088 0.7574666

Rho: -2.3168, LR test value: 6.1839, p-value: 0.012892
Asymptotic standard error: 0.66922
    z-value: -3.4619, p-value: 0.0005363
Wald statistic: 11.985, p-value: 0.0005363

Log likelihood: -11710.29 for lag model
ML residual variance (sigma squared): 15807, (sigma: 125.73)
Number of observations: 1872
Number of parameters estimated: 19
AIC: 23459, (AIC for lm: 23463)
LM test for residual autocorrelation
test value: 4.14, p-value: 0.041882
```

Fig.8 Residual Effects

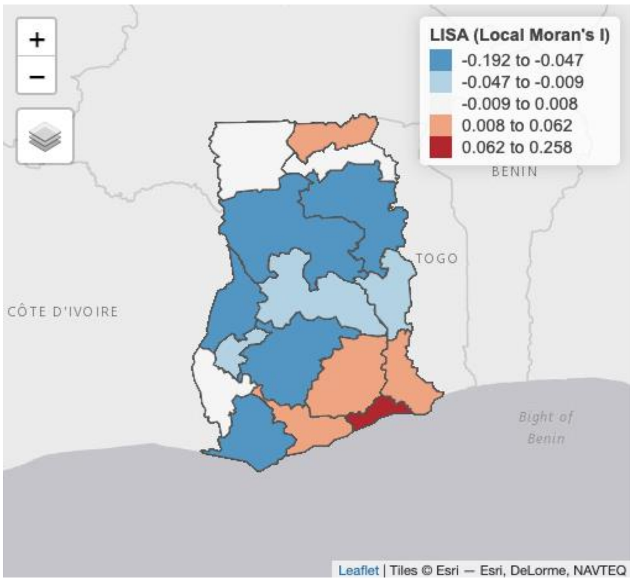


Fig.9 LISA (Local Moran's I)

4.6 Dashboard Development

- **Integration:**
    - Everything visualization and analysis in one package.
- **Interactivity:**
    - Year, metric, and region controls were included.
    - Maps and charts will update in real-time.

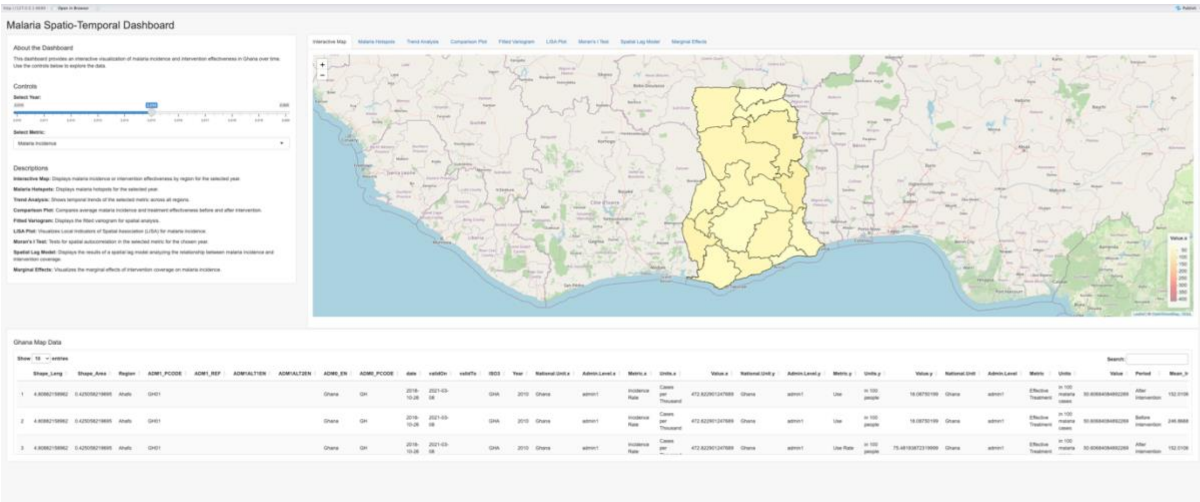


Fig.10 Application Dashboard

5.0 Reflection and Discussion

5.1 Evaluation of the Final Product

How the final product measures against the key objectives of the project:

- **Strengths:**
  - The interactive dashboard shows evidence of malaria trends and assesses the effectiveness of interventions.
  - Visualizations are lucid and intuitive for non-technical users.
  - Deepening knowledge of spatial modeling tools: Moran's I and variogram.
- **Weaknesses:**
  - Real-time notifications and mobile-responsive aspects should be enhanced.
  - More data sources can be brought in: environmental data (such as rainfall and temperature).

5.2 What Could Have Been Done Better

- **Real-Time Alerts**: The dashboard would be enhanced by implementing alerts for hotspots that will serve community health workers.
- **Mobile Optimization**: Enhancements of the dashboard would make it readily accessible to the field workforce.
- **User Testing**: Doing user testing involving the target audience would certainly yield some important feedback that would influence the improvement of the interface and usability.

5.3 What Worked Well

- Successfully managed to pull a range of data into a single dashboard.
- It embraced interactive visualizations for better engagement of users, superimposing on leaflet maps and plotly graphs.
- Provided soft insights into spatial concerns by employing tools for analyzing malaria distributions and impacts of interventions.

5.4 Future Improvements

- **Integration of environmental data sources:** rainfall and temperature, essential, as such could highlight seasonal patterns or cyclic behavior.
- More user testing workshops with public health specialists and community health workers will be very important--gathering information that would make the dashboard content stronger.
- The output - Dashboard will be hosted on the cloud for a wider reach and live updates.

6.0 Conclusion

This project will be a tool for the analysis and visualization of spatiotemporal trends in incidence of malaria in Ghana. Answering the needs of public health specialists and community health workers, it supports data-driven decision-making and resource allocation. The project has met its core objective with its current version, and more functions and usability enhancements will surely result from further versions with future user testing.

7.0 References

Anselin, L. (1995). Local indicators of spatial association—LISA. Geographical Analysis, 27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x*

Bhatt, S., Weiss, D. J., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., ... & Gething, P. W. (2015). The effect of malaria control on Plasmodium falciparum *in Africa between 2000 and 2015.* Nature, 526*(7572), 207–211. https://doi.org/10.1038/nature15535*

Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). Applied spatial data analysis with R *(2nd ed.). Springer. https://doi.org/10.1007/978-1-4614-7618-4*

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., ... & Hyndman, R. (2021). Shiny: Web application framework for R. *RStudio. Retrieved from https://shiny.rstudio.com/*

Cressie, N. (1993). Statistics for spatial data. *Wiley.*

Gething, P. W., Patil, A. P., Smith, D. L., Guerra, C. A., Elyazar, I. R. F., Johnston, G. L., ... & Hay, S. I. (2011). Climate change and the global malaria recession. Nature, 478*(7370), 376–379. https://doi.org/10.1038/nature10415*

Ghana Health Service (GHS). (2021). National Malaria Control Programme Report. *Retrieved from https://www.ghanahealthservice.org/*

Lindsay, S. W., & Birley, M. H. (1996). Climate change and malaria transmission. Annals of Tropical Medicine & Parasitology, 90*(6), 573–588. https://doi.org/10.1080/00034983.1996.11813087*

Malaria Atlas Project (MAP). (2023). Subnational malaria data. *Retrieved from https://malariaatlas.org/*

Murray, C. J. L., Rosenfeld, L. C., Lim, S. S., Andrews, K. G., Foreman, K. J., Haring, D., ... & Lopez, A. D. (2012). Global malaria mortality between 1980 and 2010: A systematic analysis. The Lancet, 379*(9814), 413–431. https://doi.org/10.1016/S0140-6736(12)60034-8*

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. The R Journal, 10*(1), 439–446. https://doi.org/10.32614/RJ-2018-009*

R Core Team. (2023). R: A language and environment for statistical computing. *R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/*

Sievert, C. (2020). Interactive web-based data visualization with R, plotly, and shiny. *CRC Press.*

Tanser, F. C., Sharp, B., & le Sueur, D. (2003). Potential effect of climate change on malaria transmission in Africa. The Lancet, 362*(9398), 1792–1798. https://doi.org/10.1016/S0140-6736(03)14841-2*

Tennekes, M. (2018). tmap: Thematic maps in R. Journal of Statistical Software, 84*(6), 1–39. https://doi.org/10.18637/jss.v084.i06*

Venables, W. N., & Ripley, B. D. (2002). Modern applied statistics with S *(4th ed.). Springer. https://doi.org/10.1007/978-0-387-21706-2*

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis *(2nd ed.). Springer. https://doi.org/10.1007/978-3-319-24277-4*

Wickham, H., & Grolemund, G. (2017). R for data science. *O'Reilly Media.*

World Health Organization (WHO). (2020). World malaria report. *Retrieved from https://www.who.int/publications/i/item/9789240015791*