Modeling Malaria Incidence Associated with Environmental Risk Factors in Ethiopia using the Geographically Weighted Regression Model, 2015-2016

Berhanu Berga Dadi (berhanuslu@gmail.com)
Universitat Jaume I. Departament de Matemàtiques

Research

Keywords: Ethiopia. Geographically weighted regression. Malaria incidence. Non-stationary. Spatial heterogeneity. Risk factors

Posted Date: September 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-907839/v1

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

**Background:** In Ethiopia, still, malaria is killing and affecting a lot of people of any age group somewhere in the country at any time. However, due to limited research, little is known about the spatial patterns and correlated risk factors on the wards scale.

**Methods:** In this research, we explored spatial patterns and evaluated related potential environmental risk factors in the distribution of malaria incidence in Ethiopia in 2015 and 2016. Hot Spot Analysis (Getis-Ord Gi* statistic) was used to assess the clustering patterns of the disease. The ordinary least square (OLS), geographically weighted regression (GWR), and semiparametric geographically weighted regression (s-GWR) models were compared to describe the spatial association of potential environmental risk factors with malaria incidence.

**Results:** Our results revealed a heterogeneous and highly clustered distribution of malaria incidence in Ethiopia during the study period. The s-GWR model best explained the spatial correlation of potential risk factors with malaria incidence and was used to produce predictive maps. The GWR model revealed that the relationship between malaria incidence and elevation, temperature, precipitation, relative humidity, and normalized difference vegetation index (NDVI) varied significantly among the wards. During the study period, the s-GWR model provided a similar conclusion, except in the case of NDVI in 2015, and elevation and temperature in 2016, which were found to have a global relationship with malaria incidence. Hence, precipitation and relative humidity exhibited a varying relationship with malaria incidence among the wards in both years.

**Conclusions:** This finding could be used in the formulation and execution of evidence-based malaria control and management program to allocate scare resources locally at the wards level. Moreover, these study results provide a scientific basis for malaria researchers in the country.

**Background**

Malaria is still the world, mainly widespread disease parasitic that kills a lot of people. As world research depicts, both types of malaria Plasmodium falciparum and Plasmodium vivax are the cause of many deaths and illness in the world. For example, due to Plasmodium falciparum 2.6 billion and Plasmodium vivax, 2.5 billion populations were at risk (Ge et al., 2017). In terms of malaria death, Africa is leading the world by 90%, and the other 10% was in the rest of the world (Alemu et al., 2011). Africa also ranks the world in malaria incidence in 2017, which is about 92% or 200 million malaria incidence and followed by South East Asia, which is only 5% of malaria incidence (WHO, 2018). The annual report of malaria showed, over 80% of malaria incidence were in sub-Saharan Africa, of which over two million died due to disease (Alemu et al., 2011).

Both parasite and vectors depend on temperature for growth (Shapiro, Whitehead and Thomas, 2017). Most of the time in Africa, the malaria incidence depends on temperature above 28° (Mordecai et al., 2013). Malaria transmission profoundly is affected by climate and environmental risk factors (Ge et al., 2017). Plasmodium parasites are the leading cause of malaria to exist in humans. Infected female Anopheles mosquitoes, also known as “malaria vectors” are the main parasite that is killing humans by transmitting from one place to another. Plasmodium falciparum and Plasmodium vivax are well-recognised parasites among the five types of pests that are killing a lot of people globally (WHO, 2016).

Usually, in Africa, the incidence of malaria depends on temperature; for example, in temperature above 28°, the prevalence of the disease is reduced (Mordecai et al., 2013). Both the climate and environmental risk factors, namely relative humidity, precipitation, and temperature), ecological and socioeconomic variables mostly affect malaria transmission (Ge et al., 2017).

Malaria is highly sensitive to climate-related disease; the study showed that the occurrence of short-term variations in climate factors such as precipitation, temperature, and relative humidity could result in a measurable malaria epidemic (Ankamah, Nokoe and Iddrisu, 2018). Currently, studies have been conducted to examine the effect of climate factor risk on malaria incidence in Port Harcourt. The result of the study showed that the occurrence of malaria significantly dependent on the increase in rainfall and a decrease in temperature (Weli and Efe, 2015). A similar study has been conducted in Ghana. The finding of the survey showed temperature maximum was better to predict the malaria epidemic in the country than minimum temperature (Ankamah, Nokoe and Iddrisu, 2018).

Malaria is a leading cause of social and public health problems globally, including Ethiopia (WHO, 2018). In Ethiopia, around 4–5 million malaria incidences have reported annually. The malaria case prevalent was about 75%, which make over 50 million people at risk (Alemu et al., 2011). Moreover, in the country, the most favorable temperature for malaria mosquito's parasite epidemic ranges between 22°C and above 32°C (Craig, Le Sueur and Snow, 1999).

Recent studies showed that the life of vector-borne diseases had been highly influenced by interannual and interdecadal climate inconsistency (Ankamah, Nokoe and Iddrisu, 2018). Therefore, global climate change has been playing a great role in the malaria epidemic in Ethiopia. Researchers have been approved that there was a link between climate variability and malaria Incidence in Ethiopia (Ankamah, Nokoe and Iddrisu, 2018). It is confirmed by the study that the malaria occurrence elsewhere in Ethiopia and Senegal were strong relationships between climatic variability and rainfall (Alemu et al., 2011).

In Ethiopia, still, malaria is killing and affecting a lot of people of any age group somewhere in the country at any time. Furthermore, there are a large number of inpatient and outpatient due to malaria case in Ethiopia (Alemu et al., 2011); it is a significant loss in terms of life and money for the country. Therefore, researching modeling malaria incidence associated with environmental risk factors in of Ethiopia is very relevant.
Methods

Study area

Ethiopia is located in the eastern part of African content approximately 30°-150N latitude and 33°-48°E, longitude. Its land and water coverage is 1,000,000 and 104,300 square kilometres, respectively (Hagose, 2017). The total population of the country is 83.7 million (Ethiopian et al., 2014). Ethiopia is one of the most densely populated areas in the world. The topography of the land varies from lowland to mountainous landscapes. The elevation in the study area varies between 1290 and 3000 meters above sea level (Fazzini, Bisci and Billi, 2015).

The study area map uses the World Geodetic System (WGS84) map projection as its reference coordinate system for data analysis.

Data Collection

Malaria incidence data were collected from Ethiopian Public Health Institute for all wards. The Ethiopia Public Health Institute summarized weekly malaria incidence in each ward in the study area between 2015 and 2016 (96 weeks in all). The total number of malaria incidence were 39592.14 over 2015–2016 in the study area. Malaria incidence was computed as malaria incidence divided by population and multiplied by 1000.

The climate and environmental predictors considered in this study, as well as their descriptions, are listed in Table 1. The selected environmental variables are monthly Normalized difference vegetation index (NDVI), elevation, Relative humidity, Temperature and Precipitation. The dataset of elevation, Relative humidity, Temperature and Precipitation is provided by Ethiopian Metrology Agency (EMA). This dataset had station data collected from 132 stations from the country. Whereas, Normalized difference vegetation index (NDVI) along with the spatial reference of the study area was downloaded from the Moderate Resolution Imaging Spectroradiometer (MODIS) instruments on-board the Terra and Aqua satellites (https://ladsweb.nascom.nasa.gov/). For this research work, the MODIS Terra NDVI product (MOD13A3 Version 6), a monthly level-3 composite with a 1 km spatial resolution, was applied to describe the vegetation coverage of each county in each month.

| Variable name                        | Code  | Source                                      |
|--------------------------------------|-------|---------------------------------------------|
| Elevation                            | Elevation | Ethiopian Metrology Agency                  |
| Relative humidity                    | RH    | Ethiopian Metrology Agency                  |
| Temperature                          | TM    | Ethiopian Metrology Agency                  |
| Precipitation                        | PR    | Ethiopian Metrology Agency                  |
| Population density                   | Pop   | Ethiopian Central Statistical Agency        |
| Average Normalized difference vegetation index | NDVI  | MOD13A3 product with Resolution/scale of Raster 1 km |

Different methods were used in other studies to interpolate environmental data by using deterministic techniques automatically or to estimate the values statistically at the grid x y co-ordinates (Berke, 2004). I applied kriging interpolation to get the value of elevation, Relative humidity, temperature and precipitation for the entire study area. Population data (2012) were collected from the Ethiopian Central Statistical Agency and were used to compute the malaria incidence.

Spatial analysis of malaria

Initially, before applying regression analysis, we generate a malaria risk map, or hot-spots map, using local Gi* statistics (Yeshiwondim et al., 2009):

$$G_{i}^{*}(d) = \frac{W_{ij}(d)z_{i}}{\sum_{j}x}$$  \hspace{1cm} (1)

Where, \(w_{ij}\) is a geospatial weight matrix at a given distance lag in kilometers \(d\), \(w_{ij}(d)\) is 1 for location distance from \(j\) to \(i\) is within \(d\); otherwise \(w_{ij}(d)\) is 0. The existence of hotspot of malaria indicators will be determined based on the value of Z-score. A high positive value of \(Z > 1.96\), showed that the position distinct \(i\) is surrounded by relatively high malaria incidence region. In contrast, a high negative Z-score value indicates that the location separates \(i\) is surrounded by relatively low (cold spot) malaria incidence in distinct areas. Otherwise, random distribution of malaria incidence for high and negative value of \(Z \geq 1.96\) and \(\leq 1.96\) (Yeshiwondim et al., 2009).

Anselin's Local Indicators of Spatial Association (LISA) method, particularly the Local Moran's I Statistic (Anselin, 1995), was used to identify and map the local clusters of high malaria incidence rates. LISA calculates a measure of spatial association for each ward locations. A local Moran's I autocorrelation statistic at the location i (Acharya et al., 2018)can be expressed as
\[ I_i = z_i \sum_{j} w_{ij} z_j \quad (2) \]

where \( z_i \) and \( z_j \) are the standardized scores of attribute values for unit \( i \) and \( j \), and \( j \) is among the identified neighbors of \( i \) according to the weight's matrix \( w_{ij} \).

**Regression analysis**

Geographically weighted regression (Ge et al., 2017) is a suitable method for spatially varying relationship data analysis. In regression data analysis, for example, ordinary least squares (OLS), generally assume fixed relationships between dependent and independent variables in the study area. However, Geographically weighted regression (GWR) lets the regression parameters to differ locally by disseminating location-wise parameter estimates for all independent variables (Fotheringham, Charlton and Brunsdon, 2002). GWR determines the spatial variability of coefficients within the study area, and the explanatory power of explanatory variables is spatially measured for local analysis (Ge et al., 2017). It has been widely used as a tool to explore non-stationarity, and its execution has been improved with new contributions, such as new sets of kernel functions (Chasco, García and Vicéns, 2007), best bandwidth selection (Páez, Uchida and Miyamoto, 2002; Ge et al., 2017), and optimum distance metric selection (Lu et al., 2015; Ge et al., 2017).

In past years, GWR has been used in several fields, for example, environmental science and meteorological science discipline (e.g., Pasculli et al., 2014; Videras, 2014; Ge et al., 2017; Yao and Zhang, 2013), Geographic information science and Remote Sensing (e.g., Gao and Li, 2011; Ge et al., 2017; Su, Xiao and Zhang, 2012; Wang, Kockelman and Wang, 2011), and mainly public health and disease (e.g. Ge et al., 2017; McKinley et al., 2013). In the studies with human health, the GWR methods have been applied to discover the spatial dissimilarities of heavy metals in the soil (McKinley et al., 2013), climate variables (Ge et al., 2017), air pollution (Jephcote and Chen, 2012), and socioeconomic variables (Chi et al., 2013).

The model GWR is appropriate for non-stationary variables (Fotheringham, Charlton and Brunsdon, 2002). In the first step, the average annual malaria incidence in all two years (2015–2016) was tested for spatial heterogeneity (non-stationarity) with Global Moran's I statistic.

In many studies, to deal with data with zero malaria incidence, the malaria case data were adjusted by a Bayesian model (Ge et al., 2017). The malaria data we used had not consisted of a large number of locations with zero malaria incidence; therefore, we didn't apply the Bayesian model for this study.

In regression, multicollinearity could occur if one explanatory variable was a linear function of another explanatory variable and formerly observed in GWR modeling (Hasyim et al., 2018). The independent variables "altitude," "relative humidity," "precipitation," "NDVI," and "temperature" were tested for multicollinearity. To investigate the collinearity problem among the independent variables, we used indices that are based on the predicted variance of modeling Variance Inflation Factor (VIF) (Halimi et al., 2014). We considered the most often applied criteria that establish that variables with VIF greater than 4 warrant further investigations, and those with VIF greater than 10 indicate serious multicollinearity.

Ordinary Least Squares was initially used before the GWR model to examine global statistical relationships between dependent and explanatory variables, including the multicollinearity assumption. At this level, the presence of local variation in relationships was not taken into account in regression. The OLS regression model was used to assess the global relationship between malaria incidence and the selected environmental risk factors. The method of least square expressed in the following equation (Acharya et al., 2018):

\[ y_i = a + \sum_{j=1}^{k} a_j x_{ij} + \epsilon_i \quad (3) \]

Where \( y_i \) is the \( i \)th examination of the response variable, \( a_j x_{ij} \) is the \( j \)th examination of the \( j \)th explanatory variable, and \( \epsilon_i \) is the error terms. The global model assumes that the rate of neighborhood ward \( i \) is independent of neighboring \( j \) and that residuals usually distributed in terms with zero mean.

Since the study area was characterized by spatial heterogeneity, we used the GWR model as an alternative to examine the local relationship between the dependent and independent variables (Hasyim et al., 2018). With the discussed dependent and independent variables, the GWR model can be formalized as

\[ y(u) = \beta_0 (\text{vec} u) + \sum_{i=1}^{n} \beta_i (\text{vec} u) x_i + \epsilon \quad (4) \]

Where \( y \) is the value of malaria incidence at the location \( u \), \( x_i \) is the value of explanatory variable \( t \) at the location \( u \), \( \beta_i (u) \) is the regression coefficient at the location \( u \), and \( \epsilon \) is the random error with mean 0 and variance \( \sigma^2 \).
In the GWR model, each explanatory variable has different regression parameters due to spatially varying parameters in weighted analysis regression (Mar’ah, Djuraidah and Wigena, 2017).

The GWR model that has both local and global parameters is known as Mixed Geographically Weighted Regression or Semi-parametric Geographically Weighted Regression (Nakaya et al., 2005). According to (Mar’ah, Djuraidah and Wigena, 2017), a stepwise procedure that allows all possible mixture of global and local parameters was tested, and the optimum mixed/semi-parametric model was selected based on the smallest corrected Akaike Information Criterion (AICc) value. The spatial variability test (F-Test) was used by the model to determine local parameters in the model (Mei, Wang and Zhang, 2006). The specified local and global parameters depend on the confidence interval of GWR coefficients (Mar’ah, Djuraidah and Wigena, 2017).

In the GWR models, a weight matrix is calculated to calibrate the model and distinguish the spatial association among nearby wards. A fixed Gaussian kernel function was applied for the weighting scheme (Hasyim et al., 2018). The optimal distance threshold was determined by minimizing the AICc of the model.

A Gaussian kernel is appropriate for fixed kernels as it can prevent the risk of there being no data within a kernel (Nakaya, 2016). The golden search method was applied to determine the optimal bandwidth size for geographically weighting efficiently. The optimal bandwidth and the related weighting function were attained by selecting the lowest AICc score. The fixed Gaussian kernel for geographical weighting used in this study (Nakaya, 2016) is as follows:

\[ w_{ij} = \exp\left[-\left(d_{ij}/b\right)^2\right] \]

Where \( w_{ij} \) is the weight value of the observation at the location \( j \) to estimate the coefficient at location \( i \), \( d_{ij} \) is the Euclidean distance between \( i \) and \( j \), and \( b \) is the size of the fixed bandwidth given by the distance metric. The positive [negative] association between response and explanatory variables can be indicated by a positive [negative] regression coefficient \( \beta_k(u) \) of the explanatory variable \( t \) at the location \( u \). If one explanatory variable \( X_t \) (i.e. environmental risk factor) has a positive [negative] coefficient at the location \( u \), it means that when \( X_t \) increases at the location \( u \), it is expected that the malaria incidence (Y) increases [decreases] at the location \( u \) assuming all other factors remain constant.

Finally, we applied Semiparametric Geographically Weighted Regression (s-GWR) models treating some predictors as global while others as local. The s-GWR model is expressed in the following equation (Fotheringham, Charlton and Brunsdon, 2002):

\[ y_i = \sum_{j=1}^{k_g} \alpha_j x_{ij} + \sum_{j=k_g+1}^{k} \beta_j (u, v_i) x_{ij} + e_i \]

(5) where for observation \( i \), \( (u, v_i) \) is the geographical location, \( \{\alpha_1 \ldots \alpha_{k_g}\} \) are the \( k_g \) global coefficients associated with the set of global explanatory variables \( \{x_1 \ldots x_{k_g}\} \), \( \{\beta_1 (u, v) \ldots \beta_{k} (u, v)\} \) are the \( k_l \) local coefficient functions associated with the set of local explanatory variables \( \{x_{k_g+1} \ldots x_{k}\} \).

The selected environmental variables (Elevation, precipitation, temperature, NDVI, and relative humidity) correspond to the explanatory variables, and the pre-processed annual average malaria incidence is the response variable in the GWR and s-GWR models of 2015 and 2016. OLS regressions were also fitted for comparison purposes. Diagnostic information provided includes the overall \( R^2 \), AICc, and the analysis of spatial autocorrelation of the residuals.

Some results of the GWR and s-GWR models (e.g., local \( R^2 \), local coefficients, and estimated incidence and residuals) were mapped using the ArcGIS10.6. Mapping local parameters make a straightforward interpretation based on recognized characteristics and spatial context of the study area (Goodchild and Janelle, 2004). On the other hand, only mapping the predictor's local coefficients does not provide a way of knowing whether they are significant anywhere on the study region (Matthews and Yang, 2012). Accordingly, statistically significant wards where pseudor values exceed \( \pm 1.96 \) were considered as relevant (Ehikses et al., 2014; Acharya et al., 2018; Wabiri et al., 2016; Matthews and Yang, 2012).

**Results**

**The distribution of malaria incidence in 2015 and 2016**

A total of 19687.31 incidence in 2015 and 456807.8 incidence in 2016 in a total population of approximately 99,870,000 and 102,400,000 in Ethiopia were recorded in 2015 and 2016, respectively. This translates to overall annual malaria incidence of 0.197 per 1000 and 4.461 per 1000 inhabitants in 2015 and 2016, respectively.

Figure 3 Depicts the annual malaria incidence distribution in 2015 and 2016 at wards level in the country using Hot Spot Analysis (Getis-Ord Gi* statistic). Accordingly, in 2015 the yearly malaria incidence hot-spot distribution was along the north and northwestern region of the country (Fig. 3.a). In contrast, in 2016, the annual malaria incidence hot-spot distribution was along the northern part of the county (Fig. 3.b). These results also highlight the spatial nonstationary of malaria incidence.

**Ordinary Least Squares model**
The coefficients of the Ordinary Least Squares models have the same value for all points within the study area (Table 2) and (Table 3). Thus, the global regression models could not capture the process for spatial heterogeneity and varying relationships in the data. In the 2015 model (Table 4), none of the regression coefficients is significantly different from zero at the 5% significance level (p-value > 0.05), though the coefficient of temperature (TM2015) is significant at the 10% significance level. In the 2016 model (Table 5) all coefficients are significant at the 5% level (p-value ≤ 0.05), except NDVI2016.

### Table 2  
**Summary of OLS Results - Model Variables for 2015**

| Variable  | Coefficient | StdError | t-Statistic | Probability | Robust_SE | Robust_t | Robust_pr (p-value) | VIF |
|-----------|-------------|----------|-------------|-------------|-----------|----------|---------------------|-----|
| Intercept | -37.078     | 65.507   | -0.566      | 0.571       | 71.985    | -0.515   | 0.606               | …… |
| Elevation | -0.009      | 0.008    | -1.120      | 0.262       | 0.009     | -0.962   | 0.335               | 1.687 |
| TM2015    | 3.083       | 1.598    | 1.930       | 0.053       | 1.777     | 1.735    | 0.083               | 3.416 |
| RH2015    | 0.193       | 0.521    | 0.371       | 0.710       | 0.581     | 0.332    | 0.739               | 3.856 |
| PR2015    | 0.246       | 0.166    | 1.481       | 0.138       | 0.198     | 1.242    | 0.214               | 3.252 |
| NDVI2015  | -0.002      | 0.002    | -1.441      | 0.149       | 0.002     | -1.589   | 0.112               | 2.208 |

### Table 3  
**Summary of OLS Results - Model Variables for 2016**

| Variable  | Coefficient | StdError | t-Statistic | Probability | Robust_SE | Robust_t | Robust_pr (p-value) | VIF |
|-----------|-------------|----------|-------------|-------------|-----------|----------|---------------------|-----|
| Intercept | -9252.870   | 2844.844 | -3.252      | 0.001*      | 3677.108  | -2.516   | 0.012*              | …… |
| Elevation | 2.895       | 0.342    | 8.451       | 0.000*      | 0.551     | 5.251    | 0.000*              | 1.731 |
| TM2016    | 411.115     | 69.412   | 5.922       | 0.000*      | 100.407   | 4.094    | 0.001*              | 3.807 |
| PR2016    | 6.261       | 4.973    | 1.258       | 0.208       | 2.211     | 2.831    | 0.004*              | 2.096 |
| RH2016    | -51.043     | 21.510   | -2.372      | 0.018*      | 25.473    | -2.003   | 0.045*              | 3.951 |
| NDVI2016  | -0.027      | 0.082    | -0.331      | 0.741       | 0.053     | -0.508   | 0.611               | 1.655 |

### Table 4  
**Comparison of goodness-of-fit results and residual analysis of the GWR and OLS models**

| Year | AICc | Adjusted R² | RSS | Global Moran’s I of residual (significance-score) |
|------|------|-------------|-----|-----------------------------------------------|
|      | GWR  | OLS         | GWR | OLS                                          |
| 2015 | 7311.88 | 7683.95 | 0.515 | 0.012 | 715.43 | 3199364.47 | 0.007 | 0.395 | (p = 0.143, Z = 1.463) | (p = 0.000, Z = 10.662) |
| 2016 | 12349.73 | 12772.77 | 0.608 | 0.182 | 2250710287 | 5753288780 | -0.104 | 0.316 | (p = 0.020, Z = -2.319) | (p = 0.000, Z = 23.683) |

### Table 5  
**Comparison of GWR and s-GWR models performances based on AICc**

| Year | Model                          | AICc  | Improvement |
|------|--------------------------------|-------|-------------|
| 2015 | GWR model (before L -> G selection) | 7311.88 | 38.194     |
|      | S-GWR model (after L -> G selection) | 7273.689 |            |
| 2016 | GWR model (before L -> G selection) | 12349.729 | 45.010     |
|      | S-GWR model (after L -> G selection) | 12304.718 |            |

In the two models (2015, 2016), all independent variables have VIF < 4, so there is no evidence of multicollinearity among them.

**Geographically Weighted Regression model**

The GWR models were used to explore the local effects of variables on malaria incidence in all wards in 2015 and 2016. The independent variables were temperature, elevation, relative humidity, precipitation, and predictor variable derived from remote sensing data (NDVI).
The pseudo-t statistics in the GWR model indicate the statistical significance of locally varying coefficients for the explanatory variables. Figure 4 depicts the spatial distribution of pseudo-t values for all independent variables for both years in the study area. Pseudo-t values were computed by dividing independent coefficient estimates by their standard errors, with statistical significance defined as a pseudo-t-value greater than or equal to 1.96 (positive relationship) or pseudo-t value smaller than or equal to -1.96 (negative relationship) (Nakaya et al., 2005; Kuo et al., 2017). The non-significant coefficients are represented in yellow in Fig. 4, with a statistically significant positive association in red/orange and negative statistically significant relationship in green/light green. Figure 4(a-j) shows local coefficients for independent variables for both years in the GWR models. It effectively reveals how the direction and strength of the relationship between each predictor and response variable vary over space. Table 4, Table 6, and Table 7 summarize the values of the maps of GWR local coefficients in Figure 4.(a-j), and also show global adjustment measures (R², Adjusted R² and AICc). Despite the higher Adjusted R² in 2016 model, the 2015 model has a better global fit considering its lowest value of the AICc. All these results are further discussed below.

| Year | Fitness measures | OLS (global model) | GWR (local model) | s-GWR (mixed model) |
|------|-----------------|--------------------|-------------------|---------------------|
| 2015 | AlCc            | 7683.95            | 7311.88           | 7273.69             |
|      | R²              | 0.021              | 0.630             | 0.642               |
|      | Adjusted R²     | 0.012              | 0.515             | 0.538               |
| 2016 | AlCc            | 12772.77           | 12349.73          | 12304.72            |
|      | R²              | 0.189              | 0.683             | 0.685               |
|      | Adjusted R²     | 0.182              | 0.609             | 0.624               |

| Global coefficients |
|---------------------|
| Variable            | Estimate | Standard Error | t(Estimate/SE) | DIFF of Criterion |
| NDVI2015            | -3.560280| 3.673756       | -0.969112      | 33.827555         |

| Local Coefficients |
|--------------------|
| Variable            | Minimum | Lower quartile | Median | Upper quartile | maximum |
| Intercept           | -191.420| 7.910          | 27.407 | 69.213         | 516.043 |
| Elevation           | -254.355| -41.107        | -14.286| -4.255         | 209.723 |
| TM2015              | -171.932| -12.620        | 1.321  | 16.662         | 505.169 |
| RH2015              | -436.018| -13.593        | 6.545  | 18.316         | 745.869 |
| PR2015              | -286.025| -15.893        | -1.410 | 8.455          | 464.719 |

R² = 0.642, Adjusted R² = 0.578, AICc = 7273.690

In 2015, temperature coefficients showed a positive and negative association with malaria case per wards and were significant in some wards located to the northwestern, southwestern part of the study areas Fig. 4g. Elevation 2015 estimated coefficients showed a positive and negative relationship with malaria incidence per ward and was significant in some wards located to the northern, southwestern, and southern part of the study areas (Fig. 4i). In 2015, the estimated NDVI coefficients showed a positive and negative association with malaria incidence per wards and were significant in some wards located to the northwestern, northeastern and southern part of the country Fig. 4a. In 2015 precipitation estimated coefficients showed a positive and negative association with malaria incidence per wards and were significant in some wards located to the northwestern, north-central, and south-central parts of the study area (Fig. 4c). In 2015 relative humidity estimated coefficients showed a positive and negative association with malaria incidence per wards and were significant in some wards located to the northern and western part of wards in the study area (Fig. 4e). In 2016 NDVI estimated coefficients showed only positive association with malaria incidence per wards and were significant in some wards located to the northeastern wards of the country (Fig. 4b).

In 2016 Precipitation estimated coefficients showed an only negative relationship with malaria incidence per wards and were significant in some wards located to the northern part of the country (Fig. 4d). In 2016 relative humidity, temperature, and elevation estimated coefficients showed a positive and negative association with malaria incidence per wards and were significant in some wards located to the northern part of the country (Fig. 4h, f, and j).
Table 4: Depicts the comparison of the GWR and OLS models based on several indicators. For both years, the sum of the residuals of squares (RSS) was summarized to evaluate the model error, and Global Moran’s I of residuals were tested along with the associated significance levels. The AICc values showed that the GWR model of each year fitted better than the corresponding OLS models. The spatial autocorrelation of residuals was not entirely removed in the 2015 GWR model, but the Global Moran’s I statistic was closer to zero in GWR than in the OLS models.

Global Moran’s I results showed (Table 4) there is significant autocorrelation in the residuals of the GWR model in 2016, and confirms the variables considered in this study were unable to appropriately predict the spatial distribution of malaria incidence in the entire study area. That was due to the scarce population in some wards or missing explanatory variables. In contrast, the Global Moran’s I result of spatial autocorrelation of residuals of the 2015 model was not statistically significant so that the model was well specified.

Semiparametric Geographically Weighted Regression

Semiparametric Geographically Weighted Regression (s-GWR) models were investigated. The GWR model with all local variables (before L -> G selection) was compared with s-GWR models (after L -> G selection), where local variables were step by step selected to become global variables. The best s-GWR models had an AICc of 7273.689 in 2015, and 12304.718 in 2016 (Table 5), thus they performed better than the GWR models (Table 8). The s-GWR models were further used to explore the local and global relationships of the explanatory variables in connection to malaria case.

Table 8: Summary of s-GWR model coefficients in 2016

| Variable | Estimate | Standard Error | t(Estimate/SE) | DIFF of Criterion |
|----------|----------|----------------|---------------|-------------------|
| Intercept| -4.715983| 259.525511     | -0.018172     | 3.588635          |
| Elevation| -69.872590| 153.268985     | -0.455882     | 17.736322         |
| TM2016   | 263.031167| 237.138194     | 1.109189      | 8.655711          |

Local Coefficients

| Variable | Minimum | Lower quartile | Median  | Upper quartile | maximum |
|----------|---------|----------------|---------|----------------|---------|
| PR2016   | -29917.885 | -76.106       | 3.564   | 57.030         | 1709.727 |
| RH2016   | -11305.206 | 86.091        | 226.052 | 292.885        | 7926.942 |
| NDVI2016 | -1485.792  | -119.889      | -32.930 | 21.994         | 13194.320 |

\[ R^2 = 0.685, \text{ Adjusted } R^2 = 0.624, \text{ AICc} = 12304.719 \]

Table 6 summarizes the result obtained from GWR 4.0 (Nakaya et al., 2005; Nakaya, 2016), where comparison of OLS, GWR, and s-GWR in terms of AICc, \( R^2 \), and adjusted \( R^2 \). The global regression model OLS explained only 1.2% in 2015 and 18.2% in 2016 of the variability of malaria incidence, whereas the variability explained by the GWR models increased to 51.5% in 2015 and 60.9% in 2016, and a little more with the s-GWR models (53.8% in 2015, and 62.4% in 2016). The model fit was significantly improved with the s-GWR model, reducing the AICc values from 7684 to 7274 in 2015, and from 12773 to 12305 in 2016. In summary, both s-GWR models performed better than the other competing models, thus they are considered the final models for malaria incidence in this study.

The results of geographic variability test and local to global variable selection approach were based on DIFF of Criterion (Table 7 and Table 8) suggesting no spatial variability in the negative coefficient of NDVI in 2015, and negative coefficient of elevation and positive coefficient of temperature in 2016 (Nakaya, 2016). Therefore, NDVI is a global explanatory variable, while the other four variables have a local varying explanation power in the 2015 model. In 2016, both elevation and temperature variables remained as global, while the other three independent variables are local.

The s-GWR model with NDVI as global predictor and elevation, temperature, precipitation, and relative humidity as local predictors corresponds to the final model found in 2015. In 2016, s-GWR model with elevation and temperature as global predictors and NDVI, precipitation and relative humidity as local predictors is the final model.

The variation of the estimated local coefficients and associated t statistics are shown in (Fig. 5 and Fig. 6 below).

Discussion

In this study, the effects environmental variables on malaria incidence were measured by OLS, GWR and s-GWR models for each year, 2015 and 2016, across 679 wards in Ethiopia. In the study area, the high-risk region for malaria, and spatial clustering appeared in the distribution of malaria incidence for both years. All three models considered the same set of explanatory variables, which were temperature, elevation, relative humidity, precipitation, and a predictor variable derived from remote sensing data (NDVI).
The results of this study showed that malaria incidence in Ethiopia during the study period heterogeneously distributed and spatially clustered at the ward level in the country. The results are consistent with research findings from past studies conducted in various malaria-endemic regions of the world (Delmelle et al., 2016; Wijayanti et al., 2016; Acharya et al., 2018; Lin and Wen, 2011).

This research is the first ward-level malaria study using the s-GWR model in entire Ethiopia, which explained the modeling malaria incidence associated with environmental risk factors in the country. The finding could be helpful for ward-level planning, policy formulation, and implementation of malaria control.

Our study showed the importance of the semiparametric geographical modeling approach of local-level risk factors analysis by contrasting global (OLS), local (GWR), and mixed (s-GWR) model. Our analysis exhibited the drawbacks of the OLS method to explain spatial variation of malaria incidence in terms of predictive performance and model accuracy and complexities evaluated to the GWR model. We showed that model goodness-of-fit could be improved through the implementation of

the s-GWR model. These findings are concurrent with malaria study in Ghana (Ehlkes et al., 2014), and dengue fever in Jhapa district, Nepal (Acharya et al., 2018). However, when independent variables do not show spatial non-stationarity, the ordinary least squares regression model is generally recommended to avoid the model complexity instead of GWR or s-GWR (Ramezankhani et al., 2017).

As a rule of thumb, a “serious” difference between GWR and OLS models generally regarded as one where the dissimilarity in AICc values between the models is at least 3 (Fotheringham, Charlton and Brunsdon, 1998). The s-GWR models had the smallest AICc values for 2015 and 2016, so it was the best model.

The Global Moran’s I of the residuals of the final s-GWR model in 2015 was −0.059589 (z score = -2.653625 and p-value = 0.012), which indicates that in 2015 there was significant spatial autocorrelation in the residuals of the model, thus it was not correctly specified (i.e. key explanatory variables are likely to be missing). In 2016, the Moran’s I of the residuals of the final s-GWR model was −0.079349 (z score = -3.622420 and p-value = 0.053), so there is no significant spatial autocorrelation in the residual and the model was properly specified.

A significant benefit of the s-GWR model is the ability to visually represent the varying strength of association between the response and explanatory variables (Buck, 2016). Moreover, it facilitates the interpretation based on spatial context and known characteristics of the study area (Goodchild and Janelle, 2004). The variation in local $R^2$ over the wards revealed significant regional differences in the malaria incidence transmission process in the study area (Fig. 7a-b). The local $R^2$ depicted that the local model had higher performance in hotspot areas when it compared to the other parts of the study area matching with previous parallel studies from Nepal (Acharya et al., 2018), Colombia (Delmelle et al., 2016) and South Africa (Manyangadze et al., 2016).

Figure 7. Local R$^2$ (a, b), residual distribution (c, d) in the s-GWR-based prediction model of the malaria case in 2015 and 2016

Our final mixed s-GWR models show that the distribution of annual malaria incidence is heterogeneous (Fig. 5 and Fig. 6) as observed in other studies (Pinchoff et al., 2015; Rulisa et al., 2013; Parker et al., 2015).

According to researchers in Ethiopia, Brazil and Cambodia (Alemu et al., 2011; De Castro et al., 2006; Dysoley et al., 2008; Hasyim et al., 2018) the environmental risk factors were significantly correlation with malaria incidence that vary strongly at the village level. Identifying the malaria incidence clusters (areas with a high number of malaria incidences) is critical in developing or improving malaria planning and control strategies at the ward scale in the country. Dissimilarities of malaria pattern exist between different regions (Guthmann et al., 2002). Thus in our study also indicated that the pattern of malaria incidence distribution is not the same in the study area; it changes from year to year in the country. The malaria incidence clusters may point to the wards that require prompt notice in terms of planning and implementation of the disease control strategies.

The heterogeneity of malaria incidence determined by a variety of ecological, biological, and sociological factors (Pinchoff et al., 2015). As noted by researcher (Ehlkes et al., 2014), nearly all studies assume homogeneous influence of independent variables (Mushinzimana et al., 2006; Dambach et al., 2012; Stefani et al., 2011) but this may not always be most appropriate (Nakaya et al., 2005). In our study, the analysis result showed that assuming that some variables vary at the local level, while others have a global effect, substantially improves the model performance. Permitting spatial heterogeneity within the regression model allows clear interpretation regarding the true nature of the potential relationship (Ehlkes et al., 2014). That could be due to the Long-lasting insecticidal nets (LLINs) distributed to some of the wards that have malaria incidence in the country. Long-lasting insecticidal nets (LLINs) are a tool to control malaria vector in malaria epidemic areas effectively (Masaninga et al., 2018). When assessing the relationship between environmental variables and malaria incidence, one should think about the pathways in which these variables under study lie (Ehlkes et al., 2014). For instance, the environmental variables: temperature, NDVI, elevation, relative humidity, and precipitation, which influence the malaria inciencde considered in this study as they determine the plenty of mosquito or their breeding habitats.

In Ethiopia, malaria control strategies include indoor residual spraying (IRS) and LLINs are applied based on the local setting (Loha et al., 2019). Those factors tend to reduce the incidence of malaria. The interaction between these factors and malaria incidence may bring out unexpected results, defying the norms regarding the association between environmental risk factors and malaria disease.
In this research, there was an association between elevation and malaria incidence. Internationally, Anopheline species diversity and density decrease from the lowlands to highlands (Hasyim et al., 2018). Therefore, poor inhabitants living in forested lowland areas in Papua, Indonesia, were found to be at a higher risk of malaria disease than those in the highlands (Hananidita and Tampubolon, 2016).

In contrast, a positive association between elevation and plenty of Anopheles mosquitoes has noticed in the highlands of Ethiopia, Colombia, and Ecuador, mainly in warmer years (Siraj et al., 2014; Pinault and Hunter, 2011; Alimi et al., 2015). It has been accepted that malaria transmission possible decreases as the altitude increases (Chikodzi, 2013; Meyrowitsch et al., 2011). In our study, also we noted elevation was significant in 2015 and showed its expected negative relationship with malaria incidence in some of the wards in the northern and southern wards, but also showed a positive correlation in some wards to the western part of the country (Fig. 5a).

In Ethiopia, precipitation was significantly correlated with malaria incidence in tropical areas (Midekisa et al., 2015). Moreover, in Botswana, precipitation showed association with the incidence of clinical malaria incidence (Chirebvu et al., 2016). Variations in monthly rainfall in rural Tanzania primarily correlated with malaria incidence (Thomson et al., 2017). In South Africa, the number of malaria incidence was significantly positively associated with higher winter precipitation (Kleinschmidt et al., 2001). In this study, coefficients of precipitation in 2015 showed the expected positive and negative relationship with malaria incidence in some wards in the country. Precipitation was significant in some rural wards located in the northwestern, western, central, and southwestern part of the country as depicted in Fig. 5b. In Ethiopia, minimum temperatures significantly correlated with malaria incidence in cold areas (Midekisa et al., 2015). In this study also local coefficients of temperature in 2015 showed positive and negative relationship with malaria incidence and were significant in some wards located in the northern and western part of the country, as depicted in Fig. 5c.

Precipitation creates oviposition sites for female mosquitoes, whereas relative humidity is a crucial parameter for adult mosquito daily survival (Day, 2016). Anopheline mosquitoes need stagnant water to wind up their larval and pupal development. Thus, precipitation and relative humidity affect the transmission of malaria by providing water to create aquatic habitats. In this study also local coefficients of relative humidity in 2015 showed the expected positive and negative relationship with malaria incidence, and they were significant in some wards located in the northwestern and western part of the country, as depicted in Fig. 5d. Anopheles (Cellia) leucosphyrus is the type of malaria that can be transmitted in forested areas of Sumatra (Elyazar et al., 2013). In 2016, NDVI local coefficients showed an only positive relationship with malaria incidence in some wards in the country. NDVI was significant in some wards located in the northern part of the country, as it showed in Fig. 6a. In 2016 precipitation local coefficients showed an only negative relationship with malaria incidence and were significant in some wards located in the northern part of the country, as it showed in Fig. 6b. In 2016 relative humidity local coefficients showed a positive and negative relationship with malaria incidence in some wards in the country. Relative humidity was significant in some wards located in the northern part of the country (Fig. 6c). This indicated that s-GWR successfully captured the spatial stationary and non-stationary to model the factors that influence the spread of malaria incidence.

The weak positive and weak negative relationships between environmental risk factors and malaria occurrences in some of the wards could be due to the protective effect of malaria control factors, for example, vector control methods including LLINs and IRS. These malaria interferences contribute significantly to the decline in malaria incidence, mainly in areas progressing towards malaria elimination (Meyrowitsch et al., 2011). Researchers (Gwitira et al., 2015) also noted that in malaria incidence where there is effective malaria control, there would be weak correlations among habitat suitability and malaria incidence. This was observed in this study in 2015, NDVI was a weak association with malaria incidence in the country. In 2016 elevation and temperature were also weak correlations with malaria incidence in the country.

Temperature, precipitation, and relative humidity are frequently used to predict for the spatial, seasonal, and interannual variation for malaria transmission, such as the dynamic malaria model forecasting malaria occurrence with seasonal climate (Hoshen and Morse, 2004). Land use, relative humidity, elevation, and precipitation have been identified by GWR to determine the regional vulnerability to malaria incidence in Purworejo, Indonesia (Hasyim et al., 2018). The GWR model revealed here in our study that elevation, temperature, precipitation, relative humidity, and NDVI significantly influence malaria incidence in some wards in Ethiopia. Similarly, in 2015 elevation, temperature, precipitation, and relative humidity have been identified by s-GWR and were significantly influence malaria incidence in some of the wards in Ethiopia. Similarly, in 2016 precipitation, NDVI, and relative humidity have been identified by the s-GWR model and were significantly influence malaria incidence in some wards in Ethiopia. However, s-GWR model allowing for spatial heterogeneity explains better the relationship of malaria incidence with environmental risk factors in Ethiopia. Similarly, in Venezuela, the GWR model analysis showed that ecological relations that act on different scales play a role in malaria transference and that modeling increases the understanding of important spatiotemporal inconsistency (Hasyim et al., 2018).

**Conclusion**

This research explored and analyzed the modeling spatial distribution of malaria incidence and its association with environmental risk factors in Ethiopia. The finding of this research showed that malaria incidence distribution in Ethiopia was heterogeneous and highly clustered at the ward level. All environmental variables considered (elevation, temperature, relative humidity, participation, and NDVI) were the most important risk factors responsible for the spatial variation of the disease incidence.

The key task for malaria elimination should be built systems and tools to reduce disease burden where malaria transmission is high locally. By comparing GWR and s-GWR against the global regression model, in both 2015 and 2016, it becomes apparent that GWR and s-GWR models yielded
new information about malaria incidence that varies over space. In our study, the variability of malaria incidence over space was due to environmental and geographical local differences (Loha and Lindtjørn, 2010). The s-GWR models provided better fits when compared with the results of the local GWR and global OLS models.

The findings of this research have a direct implication for health policy planning and decision making. Malaria being a highly focal disease, health authorities should always consider selecting district-geographical areas: in this case, the wards rather than zones for control and intervention program. The method adopted could be an essential tool to locate such high-risk areas. Moreover, this research shows the relevance of a mixed geographical regression modeling approach in geostatistics analysis of disease and other phenomena influenced by complex environmental factors at the ward or local scale.

This research inherits some limitations which need to address in the future study. We could not include some essential social-economic variables such as Gross Domestic Product (GDP) and migration patterns in our analysis due to data unavailability. Regardless of these limitations, this is the first spatially explicit malaria incidence study in Ethiopia to map and explore environmental risk factors in the entire country at the ward-level. The methodological framework implemented in this study is convertible in other regions and at different spatial scales depending upon the data availability, as well as to other climate-related diseases. Moreover, this study demonstrates the importance of a mixed s-GWR modeling approach in the spatial analysis of malaria incidence affected by complex environmental risk factors at the ward-level.

This research also revealed the importance of local geographical modeling (e.g., s-GWR) approach to improve the knowledge about malaria incidence and its determinants, so that this study can be used for the control and management of malaria disease at the ward level. Future studies should consider including more risk factors that may further improve the performance of the s-GWR models in determining the local variation of malaria incidence.

**Abbreviations**

| AIC – Akaike Information Criterion |
| AICc – Akaike Information Criterion |
| EMA - Ethiopian Metrology Agency |
| GDP – Gross Domestic Product |
| GIS – Geographic Information System |
| GWR – Geographically Weighted Regression |
| IRS – Indoor Residual Spraying |
| LLIN – Long-Lasting Insecticidal Nets |
| MODIS – Moderate Resolution Imaging Spectroradiometer |
| NDVI – Normalized Difference Vegetation Index |
| S-GWR – Semiparametric Geographically Weighted Regression |
| VIF - Variance Inflation Factor |

**Declarations**

**Availability of data and materials**

The data sets used and/or analyzed during the current study are available at Ethiopian Public Health Institute and Ethiopian Metrology Agency.

**Acknowledgments**

I want to thank also all consortium of Erasmus Mundus Master's program in Geospatial Technologies. Grateful acknowledgment expressed to Supervisors Prof.Dr.Ana Cristina Costa, Universidade NOVA de Lisboa,

**Funding**

Not applicable.
Author Contributions
The author contributed to conceptual and study design, data analysis, reviewed the paper critically then gave final approval to be published and agreed to be responsible for all aspects of the work. The author read and approved the final manuscript.

Ethics Declarations

Ethics approval and consent to participate
The study was reviewed and approved by the ethical and review committee of Kotobe Metropolitan University. A letter of permission to conduct the study was obtained from the Kotobe Metropolitan University. A Confidentiality assured for all the information gathered from the logbook by giving a unique ID number and all information was used on for this study only.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Additional information

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations

References

1. Acharya BK, et al. Modeling the spatially varying risk factors of dengue fever in Jhapa district, Nepal, using the semi-parametric geographically weighted regression model. Int J Biometeorol. 2018. doi:10.1007/s00484-018-1601-8.

2. Alemu A, et al. Climatic variables and malaria transmission dynamics in Jimma town, South West Ethiopia. Parasites Vectors. 2011;4(1):1–11. doi:10.1186/1756-3305-4-30.

3. Alimi TO, et al. Predicting potential ranges of primary malaria vectors and malaria in northern South America based on projected changes in climate, land cover and human population. Parasites Vectors. 2015. doi:10.1186/s13071-015-1033-9.

4. Ankamah S, Nokoe KS, Iddrisu WA. Modelling Trends of Climatic Variability and Malaria in Ghana Using Vector Autoregression. Malaria Research Treatment. 2018. doi:10.1155/2018/6124321.

5. Anselin L. Local Indicators of Spatial Association—LISA. Geographical Analysis. 1995. doi:10.1111/j.1538-4632.1995.tb00338.x.

6. Berke O. Exploratory disease mapping: Kriging the spatial risk function from regional count data. International Journal of Health Geographies. 2004. doi:10.1186/1476-072X-3-18.

7. Buck KD. Modelling of geographic cancer risk factor disparities in US counties. Appl Geogr. 2016. doi:10.1016/j.apgeog.2016.08.001.

8. De Castro MC, et al. Malaria risk on the Amazon frontier. Proc Natl Acad Sci USA. 2006. doi:10.1073/pnas.0510576103.

9. Chasco C, García I, Vicéns J. "Modeling Spatial Variations in Household Disposable Income with Geographically Weighted Regression." Munich Personal RePEc Archive, 2007.

10. Chi SH, et al. Can Geographically Weighted Regression improve our contextual understanding of obesity in the US? Findings from the USDA Food Atlas. Appl Geogr. 2013. doi:10.1016/j.apgeog.2013.07.017.

11. Chikodzi D. Spatial Modelling of Malaria Risk Zones Using Environmental, Anthropogenic Variables and Geographical Information Systems Techniques. Journal of Geosciences Geomatics. 2013. doi:10.12691/jgg-1-1-2.

12. Chirebvu E, et al. Clinical Malaria Transmission Trends and Its Association with Climatic Variables in Tubu Village, Botswana: A Retrospective Analysis. PloS one. 2016. doi:10.1371/journal.pone.0139843.

13. Craig M, Le Sueur D, Snow B. A climate-based distribution model of malaria transmission in sub-Saharan Africa. Parasitol Today. 1999. doi:10.1016/S0169-4758(99)01396-4.

14. Dambach P, et al. Utilization of combined remote sensing techniques to detect environmental variables influencing malaria vector densities in rural West Africa. International Journal of Health Geographics. 2012. doi:10.1186/1476-072X-11-8.
15. Day JF. Mosquito oviposition behavior and vector control. *Insects* doi. 2016. 10.3390/insects7040065.
16. Delmelle E, et al. A spatial model of socioeconomic and environmental determinants of dengue fever in Cali, Colombia. *Acta Trop.* 2016. doi:10.1016/j.actatropica.2016.08.028.
17. Dysoley L, et al. Changing patterns of forest malaria among the mobile adult male population in Chumkiri District, Cambodia. *Acta Trop.* 2008. doi:10.1016/j.actatropica.2007.01.007.
18. Ehlkes L, et al. Geographically weighted regression of land cover determinants of Plasmodium falciparum transmission in the Ashanti Region of Ghana. *International Journal of Health Geographics.* 2014. doi:10.1186/1476-072X-13-35.
19. Elyazar IRF, et al. (2013) “The distribution and bionomics of Anopheles malaria vector mosquitoes in Indonesia,” in *Advances in Parasitology.* doi: 10.1016/B978-0-12-407705-8.00003-3.
20. Ethiopian R et al. Projecting ethiopian demographics from 2012–2050 using the spectrum suite of models. (July). 2014. doi:10.13140/RG.2.1.2996.1767.
21. Fazzini M, Bisci C, Billi P. The Climate of Ethiopia Massimiliano. *Choice Reviews Online.* 2015. doi:10.5860/choice.192361.
22. Fotheringham AS, Charlton ME, Brunsdon C. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment Planning A.* 1998. doi:10.1068/a301905.
23. Fotheringham S, Charlton M, Brunsdon C. (2002) "Chp 2: Geographically Weighted Regression: The Basics," in *Geographically Weighted Regression: The analysis of spatially varying relationships.*
24. Gao J, Li S. Detecting spatially non-stationary and scale-dependent relationships between urban landscape fragmentation and related factors using Geographically Weighted Regression. *Appl Geogr.* 2011. doi:10.1016/j.apgeog.2010.06.003.
25. Ge Y, et al. Geographically weighted regression-based determinants of malaria incidences in northern China. *Trans GIS.* 2017. doi:10.1111/tgis.12259.
26. Goodchild MF, Janelle DG. Spatially Integrated Social Science Thinking Spatially in the Social Sciences. *Oxford University Press;* 2004.
27. Guthmann JP, et al. Environmental factors as determinants of malaria risk. A descriptive study on the northern coast of Peru. *Tropical Medicine International Health.* 2002. doi:10.1046/j.1365-3156.2002.00883.x.
28. Gwirita I, et al. Modelled habitat suitability of a malaria causing vector (Anopheles arabiensis) relates well with human malaria incidences in Zimbabwe. *Appl Geogr.* 2015. doi:10.1016/j.apgeog.2015.03.010.
29. Hago L. Strategic Analysis of Sesame (Sesamum indicum L.) Market Chain in Ethiopia a Case of Humera District. *International Journal of Plant Soil Science.* 2017. doi:10.9734/ijpps/2017/31928.
30. Halimi M, et al. Modelling spatial relationship between climatic conditions and annual parasite incidence of malaria in southern part of Sistan&Balouchstan Province of Iran using spatial statistic models. *Asian Pacific Journal of Tropical Disease.* 2014. doi:10.1016/S2222-1808(14)60434-5.
31. Hanandita W, Tampubolon G. Geography and social distribution of malaria in Indonesian Papua: A cross-sectional study. *International Journal of Health Geographics.* 2016. doi:10.1186/s12942-016-0043-y.
32. Hasym H, et al. Spatial modelling of malaria cases associated with environmental factors in South Sumatra, Indonesia. *Malaria Journal BioMed Central.* 2018;17(1):1–15. doi:10.1186/s12936-018-2230-8.
33. Hoshen MB, Morse AP. A weather-driven model of malaria transmission. *Malaria Journal.* 2004. doi:10.1186/1475-2875-3-32.
34. Jephcote C, Chen H. Environmental injustices of children’s exposure to air pollution from road-transport within the model British multicultural city of Leicester: 2000-09. *Sci Total Environ.* 2012. doi:10.1016/j.scitotenv.2011.11.040.
35. Kleinschmidt I, et al. Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in KwaZulu Natal, South Africa. *Am J Epidemiol.* 2001. doi:10.1093/aje/153.12.1213.
36. Kuo CC, et al. Significance of major international seaports in the distribution of murine typhus in Taiwan. *PLoS Neglected Tropical Diseases.* 2017. doi:10.1371/journal.pntd.0005430.
37. Lin CH, Wen TH. Using geographically weighted regression (GWR) to explore spatially varying relationships of immature mosquitoes and human densities with the incidence of dengue. *International Journal of Environmental Research Public Health.* 2011. doi:10.3390/ijerph8072798.
38. Loha E, Lindtjørn B. Model variations in predicting incidence of Plasmodium falciparum malaria using 1998–2007 morbidity and meteorological data from south Ethiopia. *Malaria Journal.* 2010. doi:10.1186/1475-2875-9-166.
39. Manyangadze T, et al. Risk factors and micro-geographical heterogeneity of Schistosoma haematobium in Ndumo area, uMkhanyakude district, KwaZulu-Natal, South Africa. *Acta Trop.* 2016. doi:10.1016/j.actatropica.2016.03.028.
42. Mar’ah Z, Djuraidah A, Wigena AH. (2017) "Semi-parametric Geographically Weighted Regression Modelling using Linear Model of Coregionalization," ResearchGate.Net, (July). Available at: https://www.researchgate.net/profile/Zakiyah_Marah/publication/318285119_Semi-parametric_Geographically_Weighted_Regression_Modelling_using_Linear_Model_of_Coregionalization/links/5960da5ea6fdccc9b1f96ccf/Semi-parametric-Geographically-Weighted-Regression.

43. Masaninga F, et al. Insecticide-treated nets mass distribution campaign: Benefits and lessons in Zambia. Malaria Journal. 2018. doi:10.1186/s12936-018-2314-5.

44. Matthews SA, Yang T-C. Mapping the results of local statistics. Demographic Research. 2012. doi:10.4054/demres.2012.26.6.

45. McKinley JM, et al. Investigating local relationships between trace elements in soils and cancer data. Spatial Statistics. 2013. doi:10.1016/j.spasta.2013.05.001.

46. Mei CL, Wang N, Zhang WX. Testing the importance of the explanatory variables in a mixed geographically weighted regression model. Environment Planning A. 2006. doi:10.1068/a3768.

47. Meyrowitsch DW, et al. Is the current decline in malaria burden in sub-Saharan Africa due to a decrease in vector population? Malaria Journal. 2011. doi:10.1186/1475-2875-10-188.

48. Midekisa A, et al. Seasonal associations of climatic drivers and malaria in the highlands of Ethiopia. Parasites Vectors. 2015. doi:10.1186/s13071-015-0954-7.

49. Mordecai EA, et al. Optimal temperature for malaria transmission is dramatically lower than previously predicted. Ecol Lett. 2013. doi:10.1111/ele.12015.

50. Mushinzimana E, et al. Landscape determinants and remote sensing of anopheline mosquito larval habitats in the western Kenya highlands. Malaria Journal. 2006. doi:10.1186/1475-2875-5-1.

51. Nakaya T, et al. Geographically weighted Poisson regression for disease association mapping. Stat Med. 2005. doi:10.1002/sim.2129.

52. Nakaya T. (2016) “GWR4.09 User Manual,” GWR 4 Development Team, p. 40.

53. Páez A, Uchida T, Miyamoto K. A general framework for estimation and inference of geographically weighted regression models: 1. Location-specific kernel bandwidths and a test for locational heterogeneity. Environment Planning A. 2002. doi:10.1068/a34110.

54. Parker DM, et al. Microgeography and molecular epidemiology of malaria at the Thailand-Myanmar border in the malaria pre-elimination phase. Malaria Journal. 2015. doi:10.1186/s12936-015-0712-5.

55. Pasculli A, et al. A modelling methodology for the analysis of radon potential based on environmental geology and geographically weighted regression. Environ Model Softw. 2014. doi:10.1016/j.envsoft.2014.01.006.

56. Pinault LL, Hunter FF. New highland distribution records of multiple Anopheles species in the Ecuadorian Andes. Malaria Journal. 2011. doi:10.1186/1475-2875-10-236.

57. Pinchoff J, et al. Spatial patterns of incident malaria cases and their household contacts in a single clinic catchment area of Chongwe District, Zambia. Malaria Journal. 2015. doi:10.1186/s12936-015-0793-1.

58. Ramezankhani R, et al. Environmental risk factors for the incidence of cutaneous leishmaniasis in an endemic area of Iran: A GIS-based approach. Spatial Spatio-temporal Epidemiology. 2017. doi:10.1016/j.sste.2017.03.003.

59. Rulisa S, et al. Malaria Prevalence, Spatial Clustering and Risk Factors in a Low Endemic Area of Eastern Rwanda: A Cross Sectional Study. PLoS ONE. 2013. doi:10.1371/journal.pone.0069443.

60. Shapiro LLM, Whitehead SA, Thomas MB. Quantifying the effects of temperature on mosquito and parasite traits that determine the transmission potential of human malaria. PLoS Biol. 2017. doi:10.1371/journal.pbio.2003489.

61. Siraj AS, et al. Altitudinal changes in malaria incidence in highlands of Ethiopia and Colombia. Science. 2014. doi:10.1126/science.1244325.

62. Stefani A, et al. Studying relationships between environment and malaria incidence in Camopi (French Guiana) through the objective selection of buffer-based landscape characterisations. International Journal of Health Geographics. 2011. doi:10.1186/1476-072X-10-65.

63. Su S, Xiao R, Zhang Y. Multi-scale analysis of spatially varying relationships between agricultural landscape patterns and urbanization using geographically weighted regression. Appl Geogr. 2012. doi:10.1016/j.apgeog.2011.06.005.

64. Thomson MC, et al. Using rainfall and temperature data in the evaluation of national malaria control programs in Africa. Am J Trop Med Hyg. 2017. doi:10.4269/ajtmh.16-0696.

65. Videras J. Exploring spatial patterns of carbon emissions in the USA: a geographically weighted regression approach. Popul Environ. 2014. doi:10.1007/s11111-014-0211-6.

66. Wabiri N, et al. Assessing the spatial nonstationarity in relationship between local patterns of HIV infections and the covariates in South Africa: A geographically weighted regression analysis. Spatial Spatio-temporal Epidemiology. 2016. doi:10.1016/j.sste.2015.12.003.

67. Wang Y, Kockelman KM, Wang X. Anticipation of land use change through use of geographically weighted regression models for discrete response. Transp Res Rec. 2011. doi:10.3141/2245-14.

68. Weli VE, Efe SI. Climate and Epidemiology of Malaria in Port Harcourt Region, Nigeria. American Journal of Climate Change. 2015. doi:10.4236/ajcc.2015.41004.
Figures

Figure 1

Map of the study area covering 679 wards (counties) of the entire country, Ethiopia.

Figure 2

Flow chart of the research approach
Figure 3

Hot-spot (Getis-Ord Gi* statistic) results of malaria incidence in 2015 (a) and 2016 (b)
Figure 4

GWR Pseudo t-values for independent variables in 2015 (left) and 2016 (right)
Figure 5
S-GWR Pseudo t-values for independent variables in 2015 with significance levels
Figure 6
s-GWR Pseudo t-values for independent variables in 2016 with significance levels
Figure 7

Local R² (a, b), residual distribution (c, d) in the s-GWR-based prediction model of the malaria case in 2015 and 2016